Part 2

Statistics Scope

Contant table

Rényi Entropy for Mixture Model of Multivariate Skew Laplace distribution

Salah H. Abid[1] **Uday J. Quaez**[2]

1Department of Mathematics, College of Education, Al-Mustansiriya
2Department of Mathematics, College of Education, Al-Mustansiriya
abidsalah@gmail.com , uoday1977@gmail.com

Abstract. Rényi entropy is an important concept developed by Rényi in information theory. In this paper, we study in detail this measure of information in cases multivariate skew Laplace distributions then we extend this study to the class of mixture model of multivariate skew Laplace distributions. The upper and lower bounds of Rényi entropy of mixture model are determined. In addition, an asymptotic expression for Rényi entropy is given by the approximation. Finally, we give a real data example to illustrate the behavior of entropy of the mixture model under consideration.

1. Introduction

The distributions mixture model is an essential thing for a lot of implementations like, image processing, data mining and density estimation [1-4]. Azzalini et. al. [5] proposed a multivariate skew normal distribution. Lin and Pyne [6-7] studied some developments of mixture of skew t distributions and skew normal distributions in multivariate case. the mixture of multivariate skew Laplace distributions was introduced by Doğru et al. [8]. In the entropy theory, Shannon in [9] proposed an information measure of a random vector z belong to R^d as follows

$$H(z;\vartheta)= -E\left(\log(p(x;\vartheta))\right) \qquad (1)$$

where, $p(z;\vartheta)$ is a probability density function of z. Shannon measure was generalized by Rényi [10] as follows:

$$R_\omega(z;\vartheta) = \begin{cases} \frac{1}{1-\omega}\log\left(E(p(z;\vartheta))^{\omega-1}\right), & \omega \neq 1 \\ -E\left(\log(p(z;\vartheta))\right), & \omega = 1 \end{cases} \qquad (2)$$

The relation between both measures is shown by $H(z;\vartheta) = \lim_{\omega\to 1} R_\omega(z;\vartheta)$. Also for any $0 < \omega_1 < \omega_2$ then $R_{\omega_2}(z;\vartheta) < R_{\omega_1}(z;\vartheta)$ [1]. Contreras-Reyes et. al. in [11] and [3] determined the bounds of entropy of mixture of skew normal distribution. Some results of entropy for asymmetric exclusion process was introduced by Wood [12].

In this work, the mixture of skew Laplace distributions in the multivariate case is considered. The expression of (Shannon, Rényi) entropy of this model is derived. Also, the bounds of entropy are determined. Finally, a real data example is given.

2. Preliminaries

Multivariate skew Laplace distribution was proposed by Azzalini [13]. A d-random vector x has a multivariate skew Laplace distribution denoted by $(x\sim MSL_d(\xi,\Box,\lambda))$ if its probability density function is given as follows

$$g_d(x;\xi,\Box,\lambda) = \frac{\det(\Box)^{-\frac{1}{2}}}{2^d\beta(\pi)^{\frac{1}{2}(d-1)}\Gamma\left(\frac{d+1}{2}\right)}\exp\begin{Bmatrix} -\beta\sqrt{(x-\xi)'\Box^{-1}(x-\xi)} \\ +(x-\xi)'\Box^{-1}\lambda \end{Bmatrix} \qquad (3)$$

where, $\beta = \sqrt{1+\lambda'\Psi^{-1}\lambda}$, ξ is a location vector in R^d, Ψ is a scale matrix and λ is a skewness vector. The stochastic representation of $x \sim MSL_d(\xi,\Psi,\lambda)$ can be obtained as a mixture of variance and mean of multivariate normal distribution and inverse gamma distribution [14]

$$x = \xi + z^{-1}\lambda + \sqrt{z^{-1}}\Psi^{\frac{1}{2}}y \tag{4}$$

where, $x \sim MSL_d(\xi,\Psi,\lambda)$, $y \sim MN_d(0,I_d)$ and $z \sim IG(\frac{d+1}{2},\frac{1}{2})$, the notations IG and MN_d represent inverse gamma distribution and multivariate normal distributions respectively. Arslan [14] showed that the conditional distribution of x given y will be normal distribution with mean $\xi + z^{-1}\lambda$ and variance $z^{-1}\Psi$. The joint density function of x and z is

$$g(x,z) = \frac{\det(\Psi)^{-\frac{1}{2}}\exp\{(x-\xi)'\Psi^{-1}(x-\xi)\}}{2^d\beta(\pi)^{\frac{1}{2}(d-1)}\Gamma\left(\frac{d+1}{2}\right)} z^{-\frac{5}{2}} \exp\left\{-\frac{1}{2}\begin{pmatrix}(x-\xi)'\Psi^{-1}. \\ (x-\xi)z+\beta^2z^{-1}\end{pmatrix}\right\} \tag{5}$$

and the conditional density function of z given x is

$$g(z|x) = \frac{\beta}{\sqrt{2\pi}}\exp\left\{\begin{matrix}-\beta z^{-\frac{5}{2}}\sqrt{(x-\xi)'\Psi^{-1}(x-\xi)} \\ -\frac{1}{2}((x-\xi)'\Psi^{-1}(x-\xi)z+\beta^2z^{-1})\end{matrix}\right\} \tag{6}$$

we note clearly that when $\lambda = 0$ then the distribution of skew Laplace can be reduced to the symmetric Laplace distribution. The expected and variance of x were derived as [14]

$$E(x) = \xi + (d+1)\lambda \tag{7}$$

$$Var(x) = (d+1)(\Psi + 2\lambda\lambda') \tag{8}$$

The probability density function of mixture model of these distributions denoted by $(x \sim MMSL_d(\xi,\Psi,\lambda,\Upsilon))$ is

$$g(x;\xi,\Psi,\lambda,\Upsilon) = \sum_{i=1}^{m}\Upsilon_i\, g(x;\xi_i,\Psi_i,\lambda_i) \tag{9}$$

where, $\Upsilon_i \geq 0$,$\sum_{i=1}^{m}\Upsilon_i = 1$, $\xi = \{\xi_1,\xi_2,\dots,\xi_m\}$, $\Psi = \{\Psi_1,\Psi_2,\dots,\Psi_m\}$,$\lambda = \{\lambda_1,\lambda_2,\dots,\lambda_m\}$. Let $\kappa = (\kappa_1,\kappa_2,\dots,\kappa_m)$ be a set of m latent allocations of densities of observations x. then $g(x;\xi,\Psi,\lambda,\Upsilon) = \prod_{j=1}^{m} g(x;\xi,\Psi,\lambda,\kappa_j)$, where $Pr(\kappa_j = i|\Upsilon) = \Upsilon_i$ then for each density of j-th component in (9) is shown as follows

$x_j|(\kappa_j = i) \stackrel{d}{=} \xi_i + z_j^{-1}\lambda_i + \sqrt{z_j^{-1}}\Psi_i^{\frac{1}{2}}y_j$, $j = 1,2,\dots,m$, where $y_j \sim MN_d(0,I_d)$ and $z_j \sim IG(\frac{d+1}{2},\frac{1}{2})$.
Directly, from equations (7-8), we get

$$E(x) = \sum_{i=1}^{m}\Upsilon_i(\xi_i + (d+1)\lambda_i) \tag{10}$$

$$Var(x) = \sum_{i=1}^{m}\Upsilon_i\begin{pmatrix}(d+1)(\Psi_i + 2\lambda_i\lambda_i') \\ +(\xi_i + (d+1)\lambda_i)(\xi_i + (d+1)\lambda_i)'\end{pmatrix} - E(x)\,E(x)' \tag{11}$$

Lemma 1. [15] The upper bound for any random vector z with zero mean and variance $\mathfrak{R}$ (not necessary normal) is

$$H(z;\vartheta) \le \frac{1}{2}\log\left(\det(2\pi e\Re)\right) \tag{12}$$

with equality if $z \sim MN_d(0,\Re)$.

Lemma 2. Let $x \sim MSL_d(\xi,\square,\lambda)$. Then

(i) $$E[(x-\xi)'\square^{-1}\lambda] = (d+1)\lambda'\square^{-1}\lambda \tag{13}$$

(ii) $$E\left[\beta\sqrt{(x-\xi)'\square^{-1}(x-\xi)}\right] = d + (d+1)\lambda'\square^{-1}\lambda \tag{14}$$

Proof: Directly,

$$E[(x-\xi)'\square^{-1}\lambda] = E\left(Tr((x-\xi)'\square^{-1}\lambda)\right) = E\left(Tr\left(\lambda'\square^{-1}(x-\xi)\right)\right) = Tr\left(\lambda'\square^{-1}E(x-\xi)\right)$$

$$= Tr(\lambda'\square^{-1}(d+1)\lambda) = (d+1)\lambda'\square^{-1}\lambda$$

Now, to prove part (ii), if we use the conditional density function in equation (6) then the conditional expectation of z^{-1} given x can be written as

$$E(z^{-1}|x) = \frac{1}{\beta^2}\left(1 + \beta\sqrt{(x-\xi)'\square^{-1}(x-\xi)}\right) \tag{15}$$

where, $z^{-1} \sim G(\frac{d+1}{2},\frac{1}{2})$ and $\beta = \sqrt{1+\lambda'\square^{-1}\lambda}$. By taking the expectation of above equation, we obtain

$$E\left(E(z^{-1}|x)\right) = \frac{1}{\beta^2}\left(1 + E\left[\beta\sqrt{(x-\xi)'\square^{-1}(x-\xi)}\right]\right)$$

Therefore,

$$E(z^{-1}) = \frac{1}{\beta^2}\left(1 + E\left[\beta\sqrt{(x-\xi)'\square^{-1}(x-\xi)}\right]\right)$$

But $Z^{-1} \sim G(\frac{d+1}{2},\frac{1}{2})$ then

$$d+1 = \frac{1}{\beta^2}\left(1 + E\left[\beta\sqrt{(x-\xi)'\square^{-1}(x-\xi)}\right]\right)$$

Lemma 3. [16] (Multinomial Theorem**)** Let $r_1, r_2, \dots, r_\eta \ge 0$ and $\gamma_1, \gamma_2, \dots, \gamma_\eta \ge 0$ be real numbers. Then the equation

$$\left(\sum_{i=1}^{\eta} \gamma_i r_i\right)^{\omega} = \sum_{\square_i \in B} \frac{\omega!}{\prod_{i=1}^{\eta} \square_i!} \prod_{i=1}^{\eta} (\gamma_i r_i)^{\square_i} \tag{16}$$

is accomplished where, $B = \left\{\square_i \in N, \sum_{i=1}^{\eta} \square_i = \omega, i = 1,2,\dots,\eta\right\}$ and $\sum_{\square_i \in B} \frac{\omega!}{\prod_{j=1}^{\eta} \square_j!} = \eta^{\omega}$.

Lemma 4. [16] Let $y_1, y_2, \dots, y_\eta$ and $a_1, a_2, \dots, a_\eta$ be real numbers. Then for any positive integer ω, we have

$$\left(\sum_{i=1}^{\eta} a_i y_i\right)^{\omega} = \left(\sum_{i=1}^{\eta} a_i\right)^{\omega} y_\eta^{\omega} + \sum_{i=1}^{\eta-1}\left(\sum_{k=1}^{i} a_k\right)^{\omega}\left(y_i^{\omega} - y_{i+1}^{\omega}\right)$$
$$+ \sum_{\square_t \in A} \frac{\omega!}{\prod_{j=1}^{\eta} \square_j!}\left(\prod_{t=1}^{i}(\gamma_t)^{\square_t}\right)\left[\left(\prod_{t=1}^{i-1} y_t^{\square_t}\right) - y_i^{\omega-\square_i}\right] \tag{17}$$

is satisfied. where, $A = \{\square_t \in N; 0 < \square_t < \omega, \sum_{i=1}^{\eta} \square_i = \omega, \square_{t+1} = \square_{t+2} = \cdots = \square_{\eta} = 0 \}$

Lemma 5. [16] Let $\gamma_1, \gamma_2, \ldots, \gamma_{\eta} \geq 0$ and $r_1, r_2, \ldots, r_{\eta} \geq 0$ be real numbers and let $p \geq 0$ and $0 \leq \omega \leq p$. Then,

$$\left(\sum_{i=1}^{\eta} \gamma_i r_i\right)^{\omega} \|r\|_p^{p-\omega} \geq \sum_{i=1}^{\eta-1} (i)^{1-\frac{\omega}{p}} \left(\sum_{\square=1}^{i} \gamma_{\square}\right)^{\omega} \left(r_i^p - r_{i+1}^p\right) + \eta^{1-\frac{\omega}{p}} \left(\sum_{\square=1}^{\eta} \gamma_{\square}\right)^{\omega} r_{\eta}^p \quad (18)$$

where, $\|r\|_p = \left(\sum_{\square=1}^{\eta} r_{\square}{}^p\right)^{\frac{1}{p}}$

Proposition 6. Let $x_0 \sim ML_d(\xi, \square)$. Then, for any $\omega > 0$, the Shannon and Rènyi entropies of x are

$$H(x_0; \xi, \square) = \log\left(\frac{\det(4e^2\pi\square)^{\frac{1}{2}}\Gamma\left(\frac{d+1}{2}\right)}{\sqrt{\pi}}\right) \quad (19)$$

$$R_{\omega}(x_0; \xi, \square) = \log\left[\frac{\omega^{-\frac{d}{1-\omega}} \det(4\pi\square)^{\frac{1}{2}}\Gamma\left(\frac{d+1}{2}\right)}{\sqrt{\pi}}\right], 0 < \omega, \omega \neq 1 \quad (20)$$

Proof: from the equations (1-2), Then the proof is immediate.

3. Results and Discussion

3.1. Rényi Entropy of Multivariate Skew Laplace Distributions

Proposition 7. Let $x_0 \sim ML_d(\xi, \square)$ and $x \sim MSL_d(\xi, \square, \lambda)$. Then

$$H(x; \xi, \square, \lambda) = H(x_0; \xi, \square, v) + \log(\beta) \quad (21)$$

where, $\beta = \sqrt{1 + \|\lambda\|_{\square^{-1}}^2}$ and $\|\lambda\|_{\square^{-1}}^2 = \lambda\square^{-1}\lambda'$

Proof: from equation (3), we get

$$H(X; \xi, \square, \lambda) = -\log\left(\frac{\det(\square)^{-\frac{1}{2}}}{2^d \beta (\pi)^{\frac{1}{2}(d-1)} \Gamma\left(\frac{d+1}{2}\right)}\right) + E\left(\beta\sqrt{(x-\xi)'\square^{-1}(x-\xi)}\right) - E\left((x-\xi)'\square^{-1}\lambda\right)$$

Using lemma 2., we obtain

$$H(X; \xi, \square, \lambda) = -\log\left(\frac{\det(\square)^{-\frac{1}{2}}}{2^d \beta (\pi)^{\frac{1}{2}(d-1)} \Gamma\left(\frac{d+1}{2}\right)}\right) + d$$

Lemma 8. Let $x \sim MSL_d(\xi, \square, \lambda)$. Then

$$\int_{R^d} \left(g(x; \xi, \square, \lambda)\right)^{\omega} dx = \left(\frac{\omega^{-\frac{d}{1-\omega}} \det(4\pi\square)^{\frac{1}{2}} \Gamma\left(\frac{d+1}{2}\right)\beta}{\sqrt{\pi}}\right)^{(1-\omega)} \quad (22)$$

Proof: Replacing $\text{▯}^{-\frac{1}{2}}\lambda$ by $\tilde{\lambda}$, $\omega^{-2}\text{▯}$ by ▯_{ω} and using change of variables $y = \text{▯}_{\omega}^{-\frac{1}{2}}(x-\xi)$ in equation (3), we have

$$\int_{R^d}\left(g(x;\xi,\text{▯},\lambda)\right)^{\omega}dx = \frac{\omega^{-\omega d}\det(\text{▯}_{\omega})^{-\frac{1}{2}\omega+\frac{1}{2}}}{2^{\omega d}\beta^{\omega}(\pi)^{\frac{1}{2}\omega(d-1)}\left(\Gamma\left(\frac{d+1}{2}\right)\right)^{\omega}}\int_{R^d}\exp\{-\beta\sqrt{y'y}+y'\tilde{\lambda}\}dy$$

where, $y \sim ML_d(0, I_d, \tilde{\lambda})$.

Corollary 9. If $x_0 \sim ML_d(\xi,\text{▯}))$ and $x \sim MSL_d(\xi,\text{▯},\lambda)$, then Rényi entropy can be written as

$$R_{\omega}(x;.\xi,\text{▯},\lambda) = R_{\omega}(X_0;\xi,\text{▯}) + \log(\beta) \qquad (23)$$

Proof : Taking logarithm for both sides of equation (22) and multiplying by $\frac{1}{1-\omega}$, we get

$$R_{\omega}(x;.\xi,\text{▯},\lambda,v) = \log\left[\frac{\omega^{-\frac{d}{1-\omega}}\left(\det(4\pi\text{▯})\right)^{\frac{1}{2}}\Gamma\left(\frac{d+1}{2}\right)}{\sqrt{\pi}}\right] + \log(\beta) \qquad (24)$$

Proposition 6. gives us the result of this corollary.

3.2. Approximate Rényi Entropy of Mixture Model

Lemma 10. Let $x \sim MMSL_d(\xi,\text{▯},\lambda,\curlyvee)$. Then

$$\mathcal{T}_{lower} \leq H(x;\xi,\text{▯},\lambda,\curlyvee) \leq \mathcal{T}_{upper} \qquad (25)$$

where,

$$\mathcal{T}_{upper} = \frac{1}{2}\log\left(\det(2\pi e R)\right) \qquad (26)$$

$$\mathcal{T}_{lower} = \sum_{i=1}^{m}\curlyvee_i\log\left(\frac{\left(\det(4e^2\pi\text{▯}_i)\right)^{\frac{1}{2}}\Gamma\left(\frac{d+1}{2}\right)}{\sqrt{\pi}}\right) + \sum_{i=1}^{n}\curlyvee_i\log\left(\sqrt{1+\|\lambda_i\|^2_{\text{▯}_i^{-1}}}\right) \qquad (27)$$

$$R = \sum_{i=1}^{n}\curlyvee_i\left((d+1)(\text{▯}_i+2\lambda_i\lambda_i{}') + \mathcal{K}_i\mathcal{K}_i'\right) \qquad (28)$$

$$\mathcal{K}_i = \xi_i + (d+1)\lambda_i - E(x) \qquad (29)$$

Proof: By applying lemma 1., then we get on the upper bound. From the probability density function of mixture of distributions, we have

$$H(x;\xi,\text{▯},\lambda,\curlyvee) = -E\left(\log\left(\sum_{j=1}^{m}\curlyvee_j\, g(x;\xi_j,\text{▯}_j,\lambda_j)\right)\right)$$

But the function $-\log(x)$ is a concave then from Jensen's inequality, we get

$$H(x;\xi,\text{▯},\lambda,\curlyvee) \geq \sum_{j=1}^{m}\curlyvee_j H(X;\xi_j,\text{▯}_j,\lambda_j)$$

By using proposition 7. then the proof is finished.

Lemma 11. Let $x \sim MMSL_d(\xi, \square, \lambda, \curlyvee)$. Then

$$R_\omega(x; \xi, \square, \lambda, \curlyvee) \leq \mathcal{B}_{Upper} \tag{30}$$

where,

$$\mathcal{B}_{Upper} = \frac{1}{1-\omega} \log\Big\{\exp\big((1-\omega) R_\omega(x; \xi_m, \square_m, \lambda_m)\big) + \sum_{i=1}^{m-1} \Big(\sum_{\square=1}^{i} \curlyvee_\square\Big)^\omega \begin{pmatrix} \exp\big((1-\omega) R_\omega(x; \xi_i, \square_i, \lambda_i)\big) - \\ \exp\big((1-\omega) R_\omega(x; \xi_{i+1}, \square_{i+1}, \lambda_{i+1})\big) \end{pmatrix}\Big\} \tag{31}$$

Proof:

$$\big(g(x; \xi, \square, \lambda, \curlyvee)\big)^\omega = \Big(\sum_{i=1}^{m} \curlyvee_i\, g(x; \xi_i, \square_i, \lambda_i)\Big)^\omega$$

By using lemma 5., we obtain

$$\Big(\sum_{i=1}^{m} \curlyvee_i\, g(x; \xi_i, \square_i, \lambda_i)\Big)^\omega \geq g(x; \xi_m, \square_m, \lambda_m)^\omega + \sum_{i=1}^{m-1} \Big(\sum_{\square=1}^{i} \curlyvee_\square\Big)^\omega \begin{pmatrix} \big(g(x; \xi_i, \square_i, \lambda_i)\big)^\omega \\ -\big(g(x; \xi_{i+1}, \square_{i+1}, \lambda_{i+1})\big)^\omega \end{pmatrix}$$

If we take the integration of this inequality, then

$$\int_{R^d} \big(g(x; \xi, \square, \lambda, \curlyvee)\big)^\omega dx \geq \int_{R^d} g(x; \xi_m, \square_m, \lambda_m)^\omega dx + \sum_{i=1}^{m-1} \Big(\sum_{\square=1}^{i} \curlyvee_\square\Big)^\omega \int_{R^d} \begin{bmatrix} \big(g(x; \xi_i, \square_i, \lambda_i)\big)^\omega \\ -\big(g(x; \xi_{i+1}, \square_{i+1}, \lambda_{i+1})\big)^\omega \end{bmatrix} dx$$

Again, by taking logarithm for both sides and multiplying by $\frac{1}{1-\omega}$ of above inequality , we get

$$R_\omega(X; \xi, \square, \lambda, \curlyvee) \leq \frac{1}{1-\omega} \log\Big\{\int_{R^d} g(x; \xi_m, \square_m, \lambda_m)^\omega dx + \sum_{i=1}^{m-1} \Big(\sum_{\square=1}^{i} \curlyvee_\square\Big)^\omega \int_{R^d} \begin{bmatrix} \big(g(x; \xi_i, \square_i, \lambda_i)\big)^\omega \\ -\big(g(x; \xi_{i+1}, \square_{i+1}, \lambda_{i+1})\big)^\omega \end{bmatrix} dx\Big\}$$

Lemma 12. Let $x \sim MMSL_d(\xi, \square, \lambda, \curlyvee)$. Then for each positive integers $\square_1, \square_2, \dots, \square_m \ni \sum_{i=1}^{m} \square_i = \omega$ the approximation

$$\frac{1}{\omega} \log\Big\{\Big(\frac{\omega!}{\square_1!\, \square_2! \dots \square_m!}\Big) \prod_{i=1}^{m} \big(\curlyvee_i g(x; \xi_i, \square_i, \lambda_i)\big)^{\square_i}\Big\} \cong -\sum_{i=1}^{m} \gamma_i \log\Big(\frac{\gamma_i}{\curlyvee_i g(x; \xi_i, \square_i, \lambda_i)}\Big) \tag{32}$$

is accomplished as $\omega \to \infty$, where $\gamma_i = \frac{\square_i}{\omega}, i = 1,2, \dots, m$

Proof:

$$\frac{1}{\omega} \log\Big\{\Big(\frac{\omega!}{\square_1!\, \square_2! \dots \square_m!}\Big) \prod_{i=1}^{m} \big(\curlyvee_i g(x; \xi_i, \square_i, \lambda_i)\big)^{\square_i}\Big\} = \frac{1}{\omega} \log(\omega!) - \frac{1}{\omega} \sum_{i=1}^{m} \log(\square_i!) + \frac{1}{\omega} \sum_{i=1}^{m} \square_i \log\big(\curlyvee_i g(x; \xi_i, \square_i, \lambda_i)\big)$$

By using approximation value of factorial in the last equation, we have

$$\frac{1}{\omega}\log\left\{\left(\frac{\omega!}{\square_1!\,\square_2!\ldots\square_m!}\right)\prod_{i=1}^{m}\left(\gamma_i g(x;\xi_i,\square_i,\lambda_i)\right)^{\square_i}\right\} = \log(\omega)-1+\frac{1}{2\omega}\log(2\pi\omega)-\frac{1}{\omega}\sum_{i=1}^{m}\square_I\log(\square_i)$$

$$+\frac{1}{\omega}\sum_{i=1}^{m}\square_i-\frac{1}{2\omega}\sum_{i=1}^{m}\log(2\pi\square_i)+\sum_{i=1}^{m}\gamma_i\log\left(\gamma_i g(x;\xi_i,\square_i,\lambda_i)\right)$$

But $\gamma_i=\frac{\square_i}{\omega},\; i=1,2,\ldots,m$, this gives $\sum_{i=1}^{m}\gamma_i=1$. Consequently,

$$\frac{1}{\omega}\log\left\{\left(\frac{\omega!}{\square_1!\,\square_2!\ldots\square_m!}\right)\prod_{i=1}^{m}\left(\gamma_i g(x;\xi_i,\square_i,\lambda_i)\right)^{\square_i}\right\}=\frac{1}{2\omega}\left[\log\left(\frac{(2\pi\omega)^{1-n}}{\prod_{i=1}^{n}\gamma_i}\right)\right]-\sum_{i=1}^{m}\gamma_i\log\left(\frac{\gamma_i}{\gamma_i g(x;\xi_i,\square_i,\lambda_i)}\right)$$

But $\lim_{\omega\to\infty}\sum_{i=1}^{m}\gamma_i\log\left(\frac{\gamma_i}{\gamma_i g(x;\xi_i,\square_i,\lambda_i)}\right)=0$. Then the result of this lemma is accomplished.

Lemma 13. The approximation

$$R_\omega(X;\xi,\square,\lambda,\gamma)\cong\frac{1}{1-\omega}\log\left(\sum_{\square_i\in B}\left(\prod_{i=1}^{m}(\gamma_i)^{-\square_i}\right)\left(\prod_{i=1}^{m}\gamma_i^{\square_i}\exp\left((1-\omega)R_{\square_i}(X;\xi_i,\square_i,\lambda_i)\right)\right)\right) \tag{33}$$

is satisfied as $\omega\to\infty$.where, $\sum_{\square_i\in B}\frac{\omega!}{\prod_{i=1}^{m}\square_i!}=m^{\omega}$, $B=\{\square_i\in N,\square_i>0,\sum_{i=1}^{m}\square_i=\omega, i=1,2,\ldots,m\}$

Proof:

$$\int_{R^d}\left(g(x;\xi,\square,\lambda,\gamma)\right)^{\omega}dx=\int_{R^d}\left(\sum_{i=1}^{m}\gamma_i\, g(x;\xi_i,\square_i,\lambda_i)\right)^{\omega}$$

By applying multinomial theorem on above equation, we obtain

$$\int_{R^d}\left(g(x;\xi,\square,\lambda,\gamma)\right)^{\omega}dx=\int_{R^d}\sum_{\square_i\in B}\frac{\omega!}{\prod_{i=1}^{m}\square_i!}\prod_{i=1}^{m}\left(\gamma_i g(x;\xi_i,\square_i,\lambda_i)\right)^{\square_i}dx \tag{34}$$

where, $\sum_{\square_i\in B}\frac{\omega!}{\prod_{i=1}^{m}\square_i!}=m^{\omega}$, $B=\{\square_i\in N,\square_i>0,\sum_{i=1}^{m}\square_i=\omega, i=1,2,\ldots,m\}$. By replacing right side of equation (34) in equation (32), we get

$$\int_{R^d}\left(g(x;\xi,\square,\lambda,\gamma)\right)^{\omega}=\sum_{\square_i\in B}\prod_{i=1}^{m}(\gamma_i)^{-\square_i}\int_{R^d}\prod_{i=1}^{m}\left(\gamma_i g(x;\xi_i,\square_i,\lambda_i)\right)^{\square_i}dx$$

Therefore,

$$\int_{R^d}\left(g(x;\xi,\square,\lambda,\gamma)\right)^{\omega}=\sum_{\square_i\in B}\left[\prod_{i=1}^{m}(\gamma_i)^{-\square_i}\right]\left[\prod_{i=1}^{m}\int_{R^d}\left(\gamma_i g(x;\xi_i,\square_i,\lambda_i)\right)^{\square_i}dx\right]$$

If we take logarithm for both sides and multiplying by $\frac{1}{1-\omega}$ of the last approximation, then the result of proof is hold.

Lemma 14. Consider $x\sim MMSL_d(\xi,\square,\lambda,\gamma)$, then the lower bound of Rényi entropy is

$$R_\omega(X;\xi,\square,\lambda,v,\gamma)\geq\mathcal{B}_{Lower} \tag{35}$$

where, $\mathcal{B}_{Lower}=\frac{1}{1-\omega}\log\left(\sum_{\square_i\in B}\frac{\omega!}{\prod_{i=1}^{m}\square_i!}\prod_{i=1}^{m}(\gamma_i)^{\square_i}\exp\left\{\frac{(1-\omega)}{\omega}\sum_{i=1}^{m}\square_i R_\omega(X;\xi_i,\square_i,\lambda_i)\right\}\right)$

Proof:

$$R_\omega(X;\xi,\square,\lambda,\upsilon)=\frac{1}{1-\omega}\log\left(\int_{R^d}\left(\sum_{i=1}^{m}\upsilon_i\, g(x;\xi_i,\square_i,\lambda_i)\right)^{\omega}dx\right)$$

By applying multinomial theorem on above equation, we obtain

$$\int_{R^d}\big(g(x;\xi,\square,\lambda,\upsilon)\big)^{\omega}dx=\sum_{\square_i\in B}\frac{\omega!}{\prod_{i=1}^{m}\square_i!}\prod_{i=1}^{m}(\upsilon_i)^{\square_i}\int_{R^d}\prod_{i=1}^{m}\big(g(x;\xi_i,\square_i,\lambda_i)\big)^{\square_i}dx$$

Applying generalized HÖlder's Inequality, we obtain

$$\int_{R^d}\big(g(x;\xi,\square,\lambda,\upsilon)\big)^{\omega}dx\le\sum_{\square_i\in B}\frac{\omega!}{\prod_{i=1}^{m}\square_i!}\prod_{i=1}^{m}(\upsilon_i)^{\square_i}\prod_{i=1}^{m}\left(\int_{R^d}\big(g(x;\xi_i,\square_i,\lambda_i)\big)^{p_i\square_i}dx\right)^{\frac{1}{p_i}}$$

where, $p_1,p_2,\dots,p_m>0$ and $\sum_{i=1}^{m}\frac{1}{p_i}=1$. Then

$$\int_{R^d}\big(g(x;\xi,\square,\lambda,\upsilon)\big)^{\omega}dx\le\sum_{\square_i\in B}\frac{\omega!}{\prod_{i=1}^{m}\square_i!}\prod_{i=1}^{m}(\upsilon_i)^{\square_i}\exp\left\{\sum_{i=1}^{m}\left(\frac{(1-p_i\square_i)}{p_i}R_{p_i\square_i}(X;\xi_i,\square_i,\lambda_i)\right)\right\}$$

If we choose $p_i=\frac{\omega}{\square_i}$, $i=1,2,\dots,m$ then $\sum_{i=1}^{m}\frac{1}{p_i}=\sum_{i=1}^{m}\frac{\square_i}{\omega}=1$ and $1\le\frac{\omega}{\square_i}\le\omega$, therefore,

$$\int_{R^d}\big(g(x;\xi,\square,\lambda,\upsilon)\big)^{\omega}dx\le\sum_{\square_i\in B}\frac{\omega!}{\prod_{i=1}^{m}\square_i!}\prod_{i=1}^{m}(\upsilon_i)^{\square_i}\exp\left\{\frac{(1-\omega)}{\omega}\sum_{i=1}^{m}\square_i R_\omega(X;\xi_i,\square_i,\lambda_i)\right\}$$

If we take the logarithm for both sides and multiplying by $\frac{1}{1-\omega}$ then the proof is ccompleted.

Theorem 15. Let $x\sim MMSL_d(\xi,\square,\lambda,\upsilon)$. Then the approximate form of Renyi entropy of x is

$$\begin{aligned}R_\omega(X;\xi,\square,\lambda,\upsilon)=\frac{1}{2(1-\omega)}\Bigg\{&\log\left(\sum_{\square_i\in B}\frac{\omega!}{\prod_{i=1}^{m}\square_i!}\prod_{i=1}^{m}(\upsilon_i)^{\square_i}\exp\left\{\frac{(1-\omega)}{\omega}\sum_{i=1}^{m}\square_i R_\omega(X;\xi_i,\square_i,\lambda_i)\right\}\right)\\&+(1-\omega)R_\omega(X;\xi_m,\square_m,\lambda_m)\\&+\log\left(\sum_{i=1}^{m-1}\left(\sum_{\square=1}^{i}\upsilon_\square\right)^{\omega}\left(\begin{array}{l}\exp\big((1-\omega)R_\omega(X;\xi_i,\square_i,\lambda_i)\big)\\-\exp\big((1-\omega)R_\omega(X;\xi_{i+1},\square_{i+1},\lambda_{i+1})\big)\end{array}\right)\right)\Bigg\}\end{aligned}\qquad(36)$$

Proof: The proof is directed from lemmas 11. and 14., by taking the mean of Renyi entropy bounds.

Example 1. Consider $x\sim MMSL_d(0,\square,\lambda,\upsilon)$ with the following cases:

Case (1): d=1

m=2, $\upsilon=(0.2,0.8)$, $\square=(1.5,5)$, $\lambda=(0.3\ ,4)$

m=3 , $\upsilon=(0.2,0.3,0.5)$, $\square=(1.5,5,3)$, $\lambda=(0.3\ ,4,2.2)$

m=4, $\upsilon=(0.1,0.2,0.2,0.5)$, $\square=(1.5,5,3,2)$, $\lambda=(0.3\ ,4,2.2,1)$

m=5 , $\upsilon=(0.2,0.2,0.2,0.2,0.2)$, $\square=(1.5,5,3,2,5)$, $\lambda=(0.3\ ,4,2.2,1,2.1)$

Case (2): d=2

m=2, $\upsilon=(0.2,0.8)$, $\square=\left(\begin{pmatrix}0.7 & 0.3\\0.3 & 3\end{pmatrix},\begin{pmatrix}0.12 & 0.13\\0.13 & 3\end{pmatrix}\right)$, $\lambda=\left(\begin{pmatrix}0.16\\0.59\end{pmatrix},\begin{pmatrix}2.3\\3.1\end{pmatrix}\right)$

$m=3$, $\Upsilon = (0.2,0.3,0.5)$, $\square = \left(\begin{pmatrix}0.7 & 0.3\\0.3 & 3\end{pmatrix},\begin{pmatrix}0.12 & 0.13\\0.13 & 3\end{pmatrix},\begin{pmatrix}0.18 & 0.6\\0.6 & 4\end{pmatrix}\right)$,

$\lambda = \left(\begin{pmatrix}0.16\\0.59\end{pmatrix},\begin{pmatrix}2.3\\3.1\end{pmatrix},\begin{pmatrix}2.6\\1\end{pmatrix}\right)$

$m=4$, $\Upsilon = (0.1,0.2,0.2,0.5)$,

$\square = \left(\begin{pmatrix}0.7 & 0.3\\0.3 & 3\end{pmatrix},\begin{pmatrix}0.12 & 0.13\\0.13 & 3\end{pmatrix},\begin{pmatrix}0.18 & 0.6\\0.6 & 4\end{pmatrix},I_2\right)$, $\lambda = \left(\begin{pmatrix}0.16\\0.59\end{pmatrix},\begin{pmatrix}2.3\\3.1\end{pmatrix},\begin{pmatrix}2.6\\1\end{pmatrix},\begin{pmatrix}0.6\\1\end{pmatrix}\right)$.

$m=5$, $\Upsilon = (0.2,0.2,0.2,0.2,0.2)$, $\square = \left(\begin{pmatrix}0.7 & 0.3\\0.3 & 3\end{pmatrix},\begin{pmatrix}0.12 & 0.13\\0.13 & 3\end{pmatrix},\begin{pmatrix}0.18 & 0.6\\0.6 & 4\end{pmatrix},I_2,I_2\right)$,

$\lambda = \left(\begin{pmatrix}0.16\\0.59\end{pmatrix},\begin{pmatrix}2.3\\3.1\end{pmatrix},\begin{pmatrix}2.6\\1\end{pmatrix},\begin{pmatrix}0.6\\1\end{pmatrix},\begin{pmatrix}1\\1\end{pmatrix}\right)$.

Case (3): d=3

$m=2$, $\Upsilon = (0.2,0.8)$, $\square = \left(\begin{pmatrix}0.7 & 0.3 & 0.5\\0.3 & 3 & 0.3\\0.5 & 0.3 & 1\end{pmatrix},\begin{pmatrix}5 & 0.3 & 2\\0.3 & 5 & 1\\2 & 1 & 3\end{pmatrix}\right)$, $\lambda = \left(\begin{pmatrix}0.16\\0.59\\0.1\end{pmatrix},\begin{pmatrix}2.3\\3.1\\1.5\end{pmatrix}\right)$

$m=3$, $\Upsilon = (0.2,0.3,0.5)$, $\square = \left(\begin{pmatrix}0.7 & 0.3 & 0.5\\0.3 & 3 & 0.3\\0.5 & 0.3 & 1\end{pmatrix},\begin{pmatrix}5 & 0.3 & 2\\0.3 & 5 & 1\\2 & 1 & 3\end{pmatrix},I_3\right)$,

$\lambda = \left(\begin{pmatrix}0.16\\0.59\\0.1\end{pmatrix},\begin{pmatrix}2.3\\3.1\\1.5\end{pmatrix},\begin{pmatrix}2\\1\\2\end{pmatrix}\right)$

$m=4$, $\Upsilon = (0.1,0.2,0.2,0.5)$, $\square = \left(\begin{pmatrix}0.7 & 0.3 & 0.5\\0.3 & 3 & 0.3\\0.5 & 0.3 & 1\end{pmatrix},\begin{pmatrix}5 & 0.3 & 2\\0.3 & 5 & 1\\2 & 1 & 3\end{pmatrix},I_3,I_3\right)$,

$\lambda = \left(\begin{pmatrix}0.16\\0.59\\0.1\end{pmatrix},\begin{pmatrix}2.3\\3.1\\1.5\end{pmatrix},\begin{pmatrix}2\\1\\2\end{pmatrix},\begin{pmatrix}1\\2\\1\end{pmatrix}\right)$

$m=5$, $\Upsilon = (0.2,0.2,0.2,0.2,0.2)$, $\square = \left(\begin{pmatrix}0.7 & 0.3 & 0.5\\0.3 & 3 & 0.3\\0.5 & 0.3 & 1\end{pmatrix},\begin{pmatrix}5 & 0.3 & 2\\0.3 & 5 & 1\\2 & 1 & 3\end{pmatrix},I_3,I_3,I_3\right)$,

$\lambda = \left(\begin{pmatrix}0.16\\0.59\\0.1\end{pmatrix},\begin{pmatrix}2.3\\3.1\\1.5\end{pmatrix},\begin{pmatrix}2\\1\\2\end{pmatrix},\begin{pmatrix}1\\2\\1\end{pmatrix},\begin{pmatrix}2\\1\\2\end{pmatrix}\right)$

Table 1. Shannon entropy of $x \sim MMSL_d(\xi,\square,\lambda,\Upsilon)$ is computed of the cases m = 2,3,4 and 5 in one , two and three dimensions.

d	m	Approximate Shannon entropy			
		C_{lower}	C_{upper}	$H(X;\mu,S,\delta,\varepsilon)$	*Error*
1	2	3.3841	3.7180	3.5511	0.1669
	3	3.1475	3.5316	3.3396	0.1921
	4	2.9356	3.3246	3.1301	0.1945
	5	3.0050	3.4062	3.2056	0.2006
2	2	5.1119	6.0311	5.5715	0.4596
	3	5.1747	7.1005	6.1376	0.9629
	4	4.9633	6.8127	5.8880	0.9247
	5	4.9650	6.7900	5.8775	0.9125
3	2	8.8184	9.4704	9.1444	0.3260
	3	8.0156	9.7645	8.8900	0.8744
	4	7.8976	9.4687	8.6831	0.7855
	5	7.8305	9.6221	8.7263	0.8958

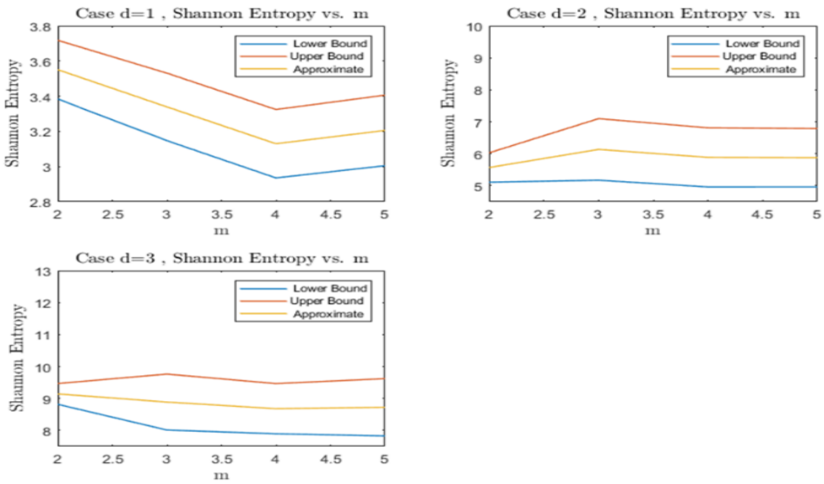


Figure 1. The horizontal and vertical lines represent the values of parameter m and Shannon entropy of $x \sim MMSL_d(\xi, \square, \lambda, v, \Upsilon)$ in example 1.

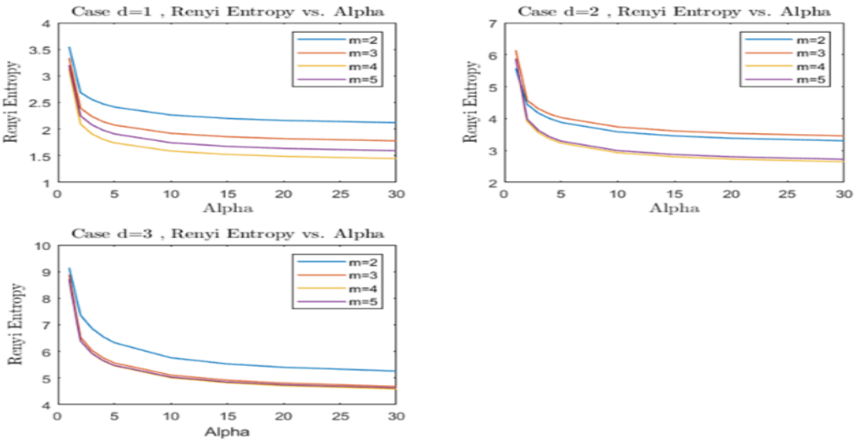


Figure 2. illustrates the relationship between Rènyi entropy of x~ $\mathbf{MMSL_d}$ (ξ,Ψ,λ,ν,ɤ) and parameter ω in example 1.

4. Conclusions

Our statistical tools were extended to determine the expressions of bounds of the class of mixture model of these distributions. Using such a pair of bounds, the approximate formula of entropy can be determined. In fact, there is not an analytical method to find the exact value of the Rényi entropy of mixture model of distributions therefore, our approximation is effective and more accurate. We have seen through the example given in this article that the error in the values of Rényi entropy by approximation was almost acceptable.

References

[1] P. José, "Information Theoretic Learning Renyi's Entropy and Kernal Perspectives," Springer Science +Business Media, LLC, 2010.

[2] S. Lee, G. McLachlan," Finite Mixtures of Multivariate Skew t-distributions," Statistics and Computing 24(2), 2014.

[3] J. Contreras-Reyes, D. Cortés," Bounds on Rényi and Shannon Entropies for Finite Mixtures of Multivariate Skew-Normal Distributions: Applications to Swordfish (Xiphias gladius Linnaeus)," Entropy 11, 382, 2016.

[4] R. Arellano-Valle, R. Contreras-Reyes, J. Genton," Shannon Entropy and Mutual information for Multivariate Skew Elliptical Distributions," Scandinavian Journal of Statistics Theory and Applications, 2012.

[5] A. Azzalini, A. Dalla Valle," The Multivariate Skew-normal distribution," Biometrika, 83(4), pp. 715-726, 1996.

[6] T. Lin, J. Lee, H. Wan," Robust Mixture Modeling Using the Skew t-Distribution ," Statistics and Computing,2007.

[7] S. Pyne, X. Hu, K. Wang, E. Rossin, T. Lin, L. Maier, C. Baecher-Allan, G. McLachlan, P. Tamayo, D. Hafler, J. Mesirov," Automated High-dimensional Flow Cytometry Data Analysis ," Proceedings of the National Academy of Sciences 106(21), 2009.

[8] F. Doğru, Y. Bulut, O. Arslan,." Finite Mixture of multivariate skew Laplace Distribution," arXiv:1702.00628, 2016.

[9] C. E. Shannon," A mathematical Theory of Communication," Bell systems technology, 27, pp. 379-423, 1948.

[10] A. Rényi," On Measures of Information and entropy," Proceeding of the Fourth Berkeley Symposium on Mathematics, Statistics and Probability pp. 547-561, 1961.

[11] E. Javier, J. Contreras-Reyes," Rényi Entropy and Complexity Measure for Skew-Gaussian Distributions and Related Families," arXiv:1406.0111v2, 2016.

[12] R. Wood, R. Blythe, M. Evans," Rényi Entropy of the Totally Asymmetric Exclusion Process," arXiv:1708.00303, 2017.

[13] A. Azzalini, A. Capitanio," Distribution Generated by Perturbation of Symmetry with Emphasis on a Multivariate Skew t Distribution," Journal of the Royal Statistical Society: series B, 65 (2), pp.367-389, 2003.

[14] O. Arslan," An alternative multivariate skew Laplace Distribution: properties and estimation," Statistical Papers, 51(4), pp. 865-887, 2010.

[15] T. Cover, J. Thomas," Elements of Information Theory," Wily and Son, 2nd New York. NY, USA, 2006.

[16] G. Bennett," Lower Bounds for Matrices," Elsevier Science Publishing Co, 1986.

Generalized Gamma – Generalized Gompertz Distribution

M A A Boshi[1], S H Abid[2] and N H Al-Noor[1]

[1] **Mathematics Department, College of Science, Mustansiriyah University, Baghdad, Iraq**
[2] **Mathematics Department, Education College, Mustansiriyah University, Baghdad, Iraq**
E-mail: drnadialnoor@gmail.com, nadialnoor@uomustansiriyah.edu.iq

Abstract: In this paper, we present a generated class of continuous distributions named H-G distributions. A new family of this class named generalized gamma - G along with one of its special cases, generalized gamma - generalized Gompertz distribution, are discussed. The cumulative distribution, probability density, reliability and hazard rate functions are introduced. Furthermore, the most vital statistical properties for instance, the r-th moment, characteristic function, quantile and simulated data, Shannon entropy, relative entropy and stress strength are obtained.
Keywords. Generalized Gamma distribution, Generalized Gompertz distribution, Shannon entropy, Relative entropy, Stress strength.

1. Introduction

Many procedures have been suggested and considered to generalized families of probability distributions based on the extension of the common continuous distributions via using differential equations, compounding weighting, adding parameter(s) to the baseline distribution, and so on. These generalized families aim to create more and more flexible distributions for modeling different types of data. The generalized families were initiated to Mudholkar et al. [1][2][3], Marshall and Olkin [4], Gupta et al. [5]. Recently many new generated families can be seen in, among others, Ahmad et al. [6], Hamedani et al. [7], Alizadeh et al. [8], Hosseini et al. [9], Korkmaz [10], Jamal et al. [11]. In this article, a new generated family of continuous distributions besides one of its special cases are presented and proposed. This family is based on composing two cumulative distribution functions "say H and G" with each other.

Assume that $G(x)$ and $g(x)$ are any baseline cdf and pdf "i.e. cumulative distribution function and probability density function respectively" of a continuous random variable X. Also assume that $H(.)$ and $h(.)$ respectively represent the cdf and pdf of any continuous distribution with the interval $[0, \infty)$. The overall formula of reliability function for this class named H - G is given by

$$R(x)_{H-G} = \int_0^{-\ln G(x)} h(x)\, dx = H(-\ln G(x)) \qquad (1)$$

Corresponding to (1), the general formulas of cdf and its associated pdf will be

$$F(x)_{H-G} = 1 - R(x)_{H-G} = 1 - H\big(-\ln G\,(x)\big) \qquad (2)$$

$$f(x)_{H-G} = \frac{d}{dx}[F(x)_{H-G}] = \frac{g(x)}{G(x)}\, h(-\ln G(x)) \qquad (3)$$

2. The remains of this article are prearranged as follows: In section 2, a brief detail about the generalized Gamma – G distributions is provided. Sections 3 and 4 respectively address the Generalized Gamma – Generalized Gompertz distribution with its essential statistical properties. Finally, the conclusions are presented in section 5.

3. Generalized Gamma – G distributions

Stacy [12] presented the generalized gamma (GG) distribution "sometimes called by his name as Stacy distribution" as an elastic family for modeling data in the set of shape and hazard rate function. The GG distribution includes special sub-models, among others, exponential, gamma, Weibull, and Rayleigh. It has been used in a number of research fields like hydrology, engineering, reliability "or survival" analysis as well as a statistical model of speech signals [13].

Let $H(.)$ and $h(.)$ that mentioned in $(1),(2)$ and (3) be the cdf and pdf of GG distribution [14] with positive parameters $a\ ,d\ ,p$ as

$$H(-\ln G(x)\,;a,d,p) = \frac{\gamma\left[\frac{d}{p},\left(\frac{-1}{a}\ln G(x)\right)^{p}\right]}{\Gamma\left(\frac{d}{p}\right)} \qquad (4)$$

$$h(-\ln G(x)\,;a,d,p) = \frac{1}{\Gamma\left(\frac{d}{p}\right)}\frac{p}{a^{d}}\,[-\ln G(x)]^{d-1}\,e^{-\left(\frac{-1}{a}\ln G(x)\right)^{p}} \qquad (5)$$

where $\Gamma(\cdot)$ and $\gamma(\cdot\,,\,\cdot)$ are respectively the gamma and incomplete gamma functions.

By substituting (4) and (5) in (2) and (3), the cdf and pdf of new family named generalized gamma – G (for short $GG-G$) distributions will be

$$F(x)_{GG-G} = 1-\frac{\gamma\left[\frac{d}{p},\left(\frac{-1}{a}\ln G(x)\right)^{p}\right]}{\Gamma\left(\frac{d}{p}\right)} \qquad (6)$$

$$f(x)_{GG-G} = \frac{1}{\Gamma\left(\frac{d}{p}\right)}\frac{p}{a^{d}}\frac{g(x)}{G(x)}[-\ln G(x)]^{d-1}\,e^{-\left(\frac{-1}{a}\ln G(x)\right)^{p}} \qquad (7)$$

It should be noted that we need to find the general expanded formula to $f(x)_{GG-G}$ which is important for obtaining the basic statistical properties when dealing with some special cases.

By using $e^{-z} = \Sigma_{i=0}^{\infty}\frac{(-1)^{i}}{i!}\,z^{i}$ the pdf in (7) can be rewritten as

$$f(x)_{GG-G} = \frac{\frac{p}{a^{d}}}{\Gamma\left(\frac{d}{p}\right)}\frac{g(x)}{G(x)}[-\ln G(x)]^{d-1}\Sigma_{i=0}^{\infty}\frac{(-1)^{i}}{i!}\left[\frac{-1}{a}\ln G(x)\right]^{pi}$$

$$= \frac{\frac{p}{a^{d}}}{\Gamma\left(\frac{d}{p}\right)}\Sigma_{i=0}^{\infty}\frac{(-1)^{i}}{i!\,a^{pi}}\frac{g(x)}{G(x)}[-\ln G(x)]^{d+pi-1}$$

For $i \geq 1$, using $[-\ln z]^{a} = \sum_{k,l=0}^{\infty}\sum_{j=0}^{k}\frac{(-1)^{j+k+l}\,a}{a-j}\,C_{k}^{k-a}C_{j}^{k}C_{l}^{a+k}\,P_{j,k}\,z^{l}$ we have

$$[-\ln G(x)]^{d+pi-1} = \sum_{k,l=0}^{\infty}\sum_{j=0}^{k}\frac{(-1)^{j+k+l}(d+pi-1)}{d+pi-1-j}\, C_k^{k-d-pi+1} C_j^k C_l^{d+pi-1+k} P_{j,k}\, [G(x)]^l$$

Now $f(x)_{GG-G}$ become

$$f(x)_{GG-G} = \frac{\frac{p}{a^d}}{\Gamma\left(\frac{d}{p}\right)}\sum_{i=0}^{\infty}\frac{(-1)^i}{i!\; a^{pi}}\frac{g(x)}{G(x)}\sum_{k,l=0}^{\infty}\sum_{j=0}^{k}\frac{(-1)^{j+k+l}(d+pi-1)}{d+pi-1-j} C_k^{k-d-pi+1} C_j^k C_l^{d+pi-1+k} P_{j,k}\, [G(x)]^l$$

Then the pdf in (7) can be rewritten with an expanded formula as

$$f(x)_{GG-G} = \frac{\frac{p}{a^d}}{\Gamma\left(\frac{d}{p}\right)}\sum_{i,k,l=0}^{\infty}\sum_{j=0}^{k}\frac{(-1)^{i+j+k+l}\;(d+pi-1)}{i!\, a^{pi}(d+pi-1-j)} C_k^{k-d-pi+1} C_j^k C_l^{d+pi-1+k} P_{j,k}\; g(x)[G(x)]^{l-1} \qquad (8)$$

where $P_{j,0} = 1$ for $j \geq 0$ and $P_{j,k} = k^{-1}\sum_{m=1}^{k}\frac{(-1)^m[m(j+1)-k]}{m+1} P_{j,k-m}$ for $k = 1,2,\dots$.

4. Generalized Gamma – Generalized Gompertz distribution

El-Gohary [15] proposed a new generalization of Gompertz, exponential and generalized exponential distributions named as the generalized Gompertz (GGo). The main feature of GGo is a flexibility that may be skewed to both sides "right and left", and the most common distributions are special sub-models of it [16].

Let $G(x)$ and $g(x)$ in (6) and (7) be the cdf and pdf of generalized Gompertz [15] with three positive parameters α, β and λ given respectively by

$$G(x;\alpha,\beta,\lambda) = \left(1 - e^{-\frac{\alpha}{\beta}(e^{\beta x}-1)}\right)^{\lambda} \qquad (9)$$

$$g(x;\alpha,\beta,\lambda) = \alpha\lambda e^{\beta x} e^{-\frac{\alpha}{\beta}(e^{\beta x}-1)}\left(1 - e^{-\frac{\alpha}{\beta}(e^{\beta x}-1)}\right)^{\lambda-1} \qquad (10)$$

According to (6) the cdf of new distribution named generalized gamma – generalized Gompertz (for short $GG-GGo$) distribution will be

$$F(x)_{GG-GGo} = 1 - \frac{\gamma\left[\frac{d}{p},\left(\frac{1}{a}\ln\left[1 - e^{-\frac{\alpha}{\beta}(e^{\beta x}-1)}\right]^{-\lambda}\right)^{p}\right]}{\Gamma\left(\frac{d}{p}\right)} \qquad (11)$$

The pdf of $GG-GGo$ distribution can be obtained according to (7) as

$$f(x)_{GG-GGo} = \frac{\frac{p}{a^d}\alpha\lambda}{\Gamma\left(\frac{d}{p}\right)} e^{\beta x} e^{-\frac{\alpha}{\beta}\left(e^{\beta x}-1\right)} \left(1-e^{-\frac{\alpha}{\beta}\left(e^{\beta x}-1\right)}\right)^{-1}$$

$$\left(\ln\left[1-e^{-\frac{\alpha}{\beta}\left(e^{\beta x}-1\right)}\right]^{-\lambda}\right)^{d-1} e^{-\left(\frac{1}{a}\ln\left[1-e^{-\frac{\alpha}{\beta}\left(e^{\beta x}-1\right)}\right]^{-\lambda}\right)^{p}} \qquad (12)$$

Now since

$$g(x)[G(x)]^{l-1} = \alpha\lambda e^{\beta x} e^{-\frac{\alpha}{\beta}\left(e^{\beta x}-1\right)} \left(1-e^{-\frac{\alpha}{\beta}\left(e^{\beta x}-1\right)}\right)^{\lambda-1} \left(\left[1-e^{-\frac{\alpha}{\beta}\left(e^{\beta x}-1\right)}\right]^{\lambda}\right)^{l-1}$$

$$= \alpha\lambda e^{\beta x} e^{-\frac{\alpha}{\beta}\left(e^{\beta x}-1\right)} \left(1-e^{-\frac{\alpha}{\beta}\left(e^{\beta x}-1\right)}\right)^{\lambda l-1} \qquad (13)$$

Then according to (8) and (13), the expansion formula for the pdf of $GG-GGo$ distribution can be obtained by

$$f(x)_{GG-GGo} = \frac{\frac{p}{a^d}\alpha\lambda}{\Gamma\left(\frac{d}{p}\right)} \sum_{i,k,l=0}^{\infty}\sum_{j=0}^{k} \frac{(-1)^{i+j+k+l}\ (d+pi-1)}{i!\,a^{pi}(d+pi-1-j)}\ C_k^{k-d-pi+1} C_j^k$$

$$C_l^{d+pi-1+k}\ P_{j,k}\ e^{\beta x}\ e^{-\frac{\alpha}{\beta}\left(e^{\beta x}-1\right)} \left(1-e^{-\frac{\alpha}{\beta}\left(e^{\beta x}-1\right)}\right)^{\lambda l-1} \qquad (14)$$

The $GG-GGo$ reliability and hazard rate functions can be obtained by

$$R(x)_{GG-GGo} = \frac{\gamma\left[\frac{d}{p},\left(\frac{1}{a}\ln\left[1-e^{-\frac{\alpha}{\beta}\left(e^{\beta x}-1\right)}\right]^{-\lambda}\right)^{p}\right]}{\Gamma\left(\frac{d}{p}\right)} \qquad (15)$$

$$D(x)_{GG-GGo} = \frac{\frac{p}{a^d}\alpha\lambda e^{\beta x} e^{-\frac{\alpha}{\beta}\left(e^{\beta x}-1\right)} \left(\ln\left[1-e^{-\frac{\alpha}{\beta}\left(e^{\beta x}-1\right)}\right]^{-\lambda}\right)^{d-1} e^{-\left(\frac{1}{a}\ln\left[1-e^{-\frac{\alpha}{\beta}\left(e^{\beta x}-1\right)}\right]^{-\lambda}\right)^{p}}}{\left(1-e^{-\frac{\alpha}{\beta}\left(e^{\beta x}-1\right)}\right)\gamma\left[\frac{d}{p},\left(\frac{1}{a}\ln\left[1-e^{-\frac{\alpha}{\beta}\left(e^{\beta x}-1\right)}\right]^{-\lambda}\right)^{p}\right]} \qquad (16)$$

Figures 1 - 4 show some shapes of $GG-GGo$ functions with specific choices of its parameters.

5. Statistical properties of the $GG-GGo$ distribution

Here, the most necessary properties of $GG-GGo$ distribution are defined.

r-th Moment: The $GG-GGo$ non-central r-th moment can be gained as follows

$$E(X^r)_{GG-GGo} = \int_0^\infty x^r f(x)_{GG-GGo}\, dx = \frac{\frac{p}{a^d}}{\Gamma\left(\frac{d}{p}\right)} \sum_{i,k,l=0}^{\infty} \sum_{j=0}^{k} \frac{(-1)^{i+j+k+l}\ (d+pi-1)}{i!\, a^{pi}(d+pi-1-j)}\ C_k^{k-d-pi+1} C_j^k$$

$$C_l^{d+pi-1+k} P_{j,k} \int_0^\infty x^r\, \alpha\lambda\, e^{\beta x}\, e^{-\frac{\alpha}{\beta}(e^{\beta x}-1)} \left(1-e^{-\frac{\alpha}{\beta}(e^{\beta x}-1)}\right)^{\lambda l-1} dx$$

$$= \frac{\frac{p}{a^d}}{\Gamma\left(\frac{d}{p}\right)} \sum_{i,k,l=0}^{\infty} \sum_{j=0}^{k} \frac{(-1)^{i+j+k+l}\ (d+pi-1)}{i!\, a^{pi}(d+pi-1-j)}\ C_k^{k-d-pi+1} C_j^k$$

$$C_l^{d+pi-1+k} P_{j,k} \frac{1}{l} \int_0^\infty x^r\, \alpha\lambda l\, e^{\beta x}\, e^{-\frac{\alpha}{\beta}(e^{\beta x}-1)} \left(1-e^{-\frac{\alpha}{\beta}(e^{\beta x}-1)}\right)^{\lambda l-1} dx$$

Let $I^{**} = \int_0^\infty x^r\, \alpha\lambda l\, e^{\beta x}\, e^{-\frac{\alpha}{\beta}(e^{\beta x}-1)} \left(1-e^{-\frac{\alpha}{\beta}(e^{\beta x}-1)}\right)^{\lambda l-1} dx$ which it represents the r-th moment of GGo distribution [15] with parameters α, β and λl as

$I^{**} = \alpha\lambda l\, \Gamma(r+1)\ \sum_{s=0}^{\infty}\sum_{m=0}^{\infty} C_s^{\lambda l-1} \frac{(-1)^{s+m}}{\Gamma(m+1)} \frac{e^{\frac{\alpha}{\beta}(s+1)}}{\left[\frac{\alpha}{\beta}(s+1)\right]^{-m}} \left[\frac{-1}{\beta(m+1)}\right]^{r+1}$, then $E(X^r)_{GG-GGo}$ will be

$$E(X^r)_{GG-GGo} = \frac{\frac{p}{a^d}}{\Gamma\left(\frac{d}{p}\right)} \sum_{i,k,l=0}^{\infty} \sum_{j=0}^{k} \frac{(-1)^{i+j+k+l}\ (d+pi-1)}{i!\, a^{pi}(d+pi-1-j)}$$

$$C_k^{k-d-pi+1} C_j^k C_l^{d+pi-1+k} P_{j,k} \frac{1}{l} \alpha\lambda l\, \Gamma(r+1)$$

$$\sum_{s=0}^{\infty}\sum_{m=0}^{\infty} C_s^{\lambda l-1} \frac{(-1)^{s+m}}{\Gamma(m+1)} \frac{e^{\frac{\alpha}{\beta}(s+1)}}{\left[\frac{\alpha}{\beta}(s+1)\right]^{-m}} \left[\frac{-1}{\beta(m+1)}\right]^{r+1}$$

Thus the $GG-GGo$ non-central r-th moment is given by

$$E(X^r)_{GG-GGo} = \frac{\frac{p}{a^d}\alpha\lambda}{\Gamma\left(\frac{d}{p}\right)} \sum_{i,k,l,s,m=0}^{\infty} \sum_{j=0}^{k} \frac{(-1)^{i+j+k+l+s+m}\ (d+pi-1)}{i!\, a^{pi}(d+pi-1-j)}$$

$$C_k^{k-d-pi+1} C_j^k C_l^{d+pi-1+k} C_s^{\lambda l-1} P_{j,k} \frac{\Gamma(r+1)}{\Gamma(m+1)}$$

$$\frac{e^{\frac{\alpha}{\beta}(s+1)}[\alpha(s+1)]^m}{\beta^{m+r+1}} \left[\frac{-1}{m+1}\right]^{r+1} \qquad (17)$$

With $(r = 1,2,3,4)$ extra properties like the mean, variance, coefficient of skewness and coefficient of kurtosis can be attained.

Characteristic Function: The $GG-GGo$ characteristic function can be found by

$$Q_X(t)_{GG-GGo} = \sum_{r=0}^{\infty} \frac{(it)^r}{r!} E(X^r)_{GG-GGo} = \frac{\frac{p}{a^d}\alpha\lambda}{\Gamma\left(\frac{d}{p}\right)} \sum_{i,k,l,s,m,r=0}^{\infty} \sum_{j=0}^{k} \frac{(-1)^{i+j+k+l+s+m}}{i!\, a^{pi}}$$
$$\frac{(d+pi-1)}{(d+pi-1-j)} C_k^{k-d-pi+1} C_j^k C_l^{d+pi-1+k} C_s^{\lambda l-1} P_{j,k} \frac{(it)^r}{r!} \frac{\Gamma(r+1)}{\Gamma(m+1)}$$
$$\frac{e^{\frac{\alpha}{\beta}(s+1)}[\alpha(s+1)]^m}{\beta^{m+r+1}} \left[\frac{-1}{m+1}\right]^{r+1} \qquad (18)$$

Quantile Function: The $100\, q^{th}$ quantile of $GG-GGo$ random variable is defined as a solution of $P(x \le x_q) = F(x_q)_{GG-GGo}$ with respect to x_q where $x_q > 0$ and $0 < q < 1$. Therefore it may be obtained by solving numerically the following non-linear equation

$$q_{GG-GGo} = 1 - \frac{\gamma\left[\frac{d}{p}, \left(\frac{1}{a} \ln\left[1-e^{-\frac{\alpha}{\beta}\left(e^{\beta x_q}-1\right)}\right]^{-\lambda}\right)^p\right]}{\Gamma\left(\frac{d}{p}\right)} \qquad (19)$$

A random variable X has $GG-GGo$ distribution can be simulated by solving numerically the following nonlinear equation

$$(1-U)\Gamma\left(\frac{d}{p}\right) - \gamma\left[\frac{d}{p}, \left(\frac{1}{a} \ln\left[1-e^{-\frac{\alpha}{\beta}\left(e^{\beta x}-1\right)}\right]^{-\lambda}\right)^p\right] = 0 \qquad (20)$$

where U has the standard uniform distribution**.**

Shannon Entropy: The $GG-GGo$ Shannon entropy can be obtained by $-\int_0^{\infty} \ln(f(x)_{GG-GGo}) f(x)_{GG-GGo}\, dx$. Adopting the natural logarithm to the pdf in (12) and using $e^z = \Sigma_{i=0}^{\infty} \frac{z^i}{i!}$ we have

$$\ln f(x)_{GG-GGo} = \ln\left(\frac{\frac{p}{a^d}\alpha\lambda}{\Gamma\left(\frac{d}{p}\right)}\right) + \beta x - \frac{\alpha}{\beta}\left(e^{\beta x}-1\right) - \ln\left[1-e^{-\frac{\alpha}{\beta}\left(e^{\beta x}-1\right)}\right]$$
$$+(d-1)\ln\left(\ln\left[1-e^{-\frac{\alpha}{\beta}\left(e^{\beta x}-1\right)}\right]^{-\lambda}\right) - \left(\frac{1}{a}\ln\left[1-e^{-\frac{\alpha}{\beta}\left(e^{\beta x}-1\right)}\right]^{-\lambda}\right)^p$$
$$= \ln\left(\frac{\frac{p}{a^d}\alpha\lambda}{\Gamma\left(\frac{d}{p}\right)}\right) + \beta x - \frac{\alpha}{\beta}e^{\beta x} + \frac{\alpha}{\beta} - \ln\left[1-e^{-\frac{\alpha}{\beta}\left(e^{\beta x}-1\right)}\right]$$
$$+(d-1)\ln\left(\ln\left[1-e^{-\frac{\alpha}{\beta}\left(e^{\beta x}-1\right)}\right]^{-\lambda}\right) - \left(\frac{1}{a}\ln\left[1-e^{-\frac{\alpha}{\beta}\left(e^{\beta x}-1\right)}\right]^{-\lambda}\right)^p$$

$$= \ln\left(\frac{\frac{p}{a^d}\alpha\lambda}{\Gamma\left(\frac{d}{p}\right)}\right) + \frac{\alpha}{\beta} + \beta x - \frac{\alpha}{\beta}\sum_{k=0}^{\infty}\frac{\beta^k}{k!}x^k - \ln\left[1 - e^{-\frac{\alpha}{\beta}\left(e^{\beta x}-1\right)}\right]$$
$$+(d-1)\ln\left(\lambda \ln\left[1 - e^{-\frac{\alpha}{\beta}\left(e^{\beta x}-1\right)}\right]^{-1}\right)$$
$$-\frac{(-\lambda)^p}{a^p}\left(\ln\left[1 - e^{-\frac{\alpha}{\beta}\left(e^{\beta x}-1\right)}\right]\right)^p$$

$$= \ln\left(\frac{\frac{p}{a^d}\alpha\lambda}{\Gamma\left(\frac{d}{p}\right)}\right) + \frac{\alpha}{\beta} + (d-1)\ln(\lambda) + \beta x - \frac{\alpha}{\beta}\sum_{k=0}^{\infty}\frac{\beta^k}{k!}x^k$$
$$-\ln\left[1 - e^{-\frac{\alpha}{\beta}\left(e^{\beta x}-1\right)}\right] + (d-1)\ln\left(-\ln\left[1 - e^{-\frac{\alpha}{\beta}\left(e^{\beta x}-1\right)}\right]\right)$$
$$-\frac{(-\lambda)^p}{a^p}\left(\ln\left[1 - e^{-\frac{\alpha}{\beta}\left(e^{\beta x}-1\right)}\right]\right)^p$$

Now $GG - GGo$ Shannon entropy is given by

$$SH_{GG-GGo} = -\left\{\begin{array}{l} \ln\left(\frac{\frac{p}{a^d}\alpha\lambda}{\Gamma\left(\frac{d}{p}\right)}\right) + \frac{\alpha}{\beta} + (d-1)\ln(\lambda) + \beta E(X) \\ -\frac{\alpha}{\beta}\sum_{k=0}^{\infty}\frac{\beta^k}{k!}E(X^k) - E\left(\ln\left[1 - e^{-\frac{\alpha}{\beta}\left(e^{\beta X}-1\right)}\right]\right) \\ +(d-1)E\left[\ln\left(-\ln\left[1 - e^{-\frac{\alpha}{\beta}\left(e^{\beta X}-1\right)}\right]\right)\right] \\ -\frac{(-\lambda)^p}{a^p}E\left[\left(\ln\left[1 - e^{-\frac{\alpha}{\beta}\left(e^{\beta X}-1\right)}\right]\right)^p\right] \end{array}\right\} \quad (21)$$

where $E(X)$ and $E(X^k)$ as in (17) with $r = 1$ and $r = k$ respectively. Also we need to obtain the following expectations

$E\left(\ln\left[1 - e^{-\frac{\alpha}{\beta}\left(e^{\beta X}-1\right)}\right]\right)$, $E\left[\ln\left(-\ln\left[1 - e^{-\frac{\alpha}{\beta}\left(e^{\beta X}-1\right)}\right]\right)\right]$ and $E\left[\left(\ln\left[1 - e^{-\frac{\alpha}{\beta}\left(e^{\beta X}-1\right)}\right]\right)^p\right]$.

Now $E\left(\ln\left[1 - e^{-\frac{\alpha}{\beta}\left(e^{\beta X}-1\right)}\right]\right) = \int_0^{\infty}\ln\left[1 - e^{-\frac{\alpha}{\beta}\left(e^{\beta x}-1\right)}\right]f(x)\,dx$. By using

$\ln(1-z) = -\Sigma_{i=1}^{\infty}\frac{z^i}{i}; |z| < 1, e^{-z} = \Sigma_{i=0}^{\infty}\frac{(-1)^i}{i!}z^i,$

$(a+b)^n = \Sigma_{k=0}^{\infty} C_k^n\, a^{n-k}\, b^k = \Sigma_{k=0}^{\infty} C_k^n\, a^k\, b^{n-k}\,; n \geq 0, C_k^n = \frac{n!}{k!\,(n-k)!}$ is Binomial coefficients and $e^z = \Sigma_{i=0}^{\infty}\frac{z^i}{i!}$ we have

$$\ln\left[1-e^{-\frac{\alpha}{\beta}(e^{\beta x}-1)}\right] = -\sum_{n=1}^{\infty}\frac{e^{-\frac{\alpha n}{\beta}(e^{\beta x}-1)}}{n}$$

$$= -\sum_{n=1}^{\infty}\sum_{i=0}^{\infty}\frac{(-1)^i}{n\,i!}\left(\frac{\alpha n}{\beta}\right)^i\left(e^{\beta x}-1\right)^i$$

$$= -\sum_{n=1}^{\infty}\sum_{i=0}^{\infty}\frac{(-1)^i}{n\,i!}\left(\frac{\alpha n}{\beta}\right)^i\sum_{i_1=0}^{i}C_{i_1}^{i}\,e^{i_1\beta x}(-1)^{i-i_1}$$

$$= \sum_{n=1}^{\infty}\sum_{i=0}^{\infty}\sum_{i_1=0}^{i}\frac{(-1)^{2i-i_1+1}}{n\,i!}\left(\frac{\alpha n}{\beta}\right)^i C_{i_1}^{i}\sum_{i_2=0}^{\infty}\frac{(i_1\beta)^{i_2}}{i_2!}x^{i_2}$$

$$= \sum_{n=1}^{\infty}\sum_{i=0}^{\infty}\sum_{i_1=0}^{i}\sum_{i_2=0}^{\infty}\frac{(-1)^{2i-i_1+1}}{n\,i!}\left(\frac{\alpha n}{\beta}\right)^i C_{i_1}^{i}\frac{(i_1\beta)^{i_2}}{i_2!}x^{i_2}$$

Therefore

$$E\left(\ln\left[1-e^{-\frac{\alpha}{\beta}(e^{\beta X}-1)}\right]\right) = \sum_{n=1}^{\infty}\sum_{i=0}^{\infty}\sum_{i_1=0}^{i}\sum_{i_2=0}^{\infty}\frac{(-1)^{2i-i_1+1}}{n\,i!}\left(\frac{\alpha n}{\beta}\right)^i C_{i_1}^{i}\frac{(i_1\beta)^{i_2}}{i_2!}E(X^{i_2}) \qquad (22)$$

For $\ln\left(-\ln\left[1-e^{-\frac{\alpha}{\beta}(e^{\beta x}-1)}\right]\right)$ using the same agreement $\ln(1-z), (a+b)^n$, e^{-z}, e^{z} and $\ln z = \Sigma_{k=0}^{\infty}\frac{(-1)^k(z-1)^{k+1}}{k+1}$; $0 < z \le 2$ we have

$$\ln\left(-\ln\left[1-e^{-\frac{\alpha}{\beta}(e^{\beta x}-1)}\right]\right) = \ln\left(\Sigma_{n=1}^{\infty}\frac{e^{-\frac{\alpha n}{\beta}(e^{\beta x}-1)}}{n}\right)$$

$$= \Sigma_{k=0}^{\infty}\frac{(-1)^k}{k+1}\left(\sum_{n=1}^{\infty}\frac{e^{-\frac{\alpha n}{\beta}(e^{\beta x}-1)}}{n}-1\right)^{k+1}$$

$$= \Sigma_{k=0}^{\infty}\frac{(-1)^k}{k+1}\Sigma_{j=0}^{k+1}C_j^{k+1}\left(\Sigma_{n=1}^{\infty}\frac{e^{-\frac{\alpha n}{\beta}(e^{\beta x}-1)}}{n}\right)^j(-1)^{k+1-j}$$

$$= \Sigma_{k=0}^{\infty}\Sigma_{j=0}^{k+1}\frac{(-1)^{2k+1-j}}{k+1}C_j^{k+1}\left(\Sigma_{n=1}^{\infty}\frac{e^{-\frac{\alpha n}{\beta}(e^{\beta x}-1)}}{n}\right)^j$$

$$= \Sigma_{k=0}^{\infty}\Sigma_{j=0}^{k+1}\frac{(-1)^{2k+1-j}}{k+1}C_j^{k+1}$$

$$\Sigma_{n_1=1}^{\infty}\Sigma_{n_2=1}^{\infty}\cdots\Sigma_{n_j=1}^{\infty}\frac{e^{-\frac{\alpha}{\beta}(n_1+n_2+\cdots+n_j)(e^{\beta x}-1)}}{n_1 n_2\ldots n_j}$$

$$= \Sigma_{k=0}^{\infty}\Sigma_{j=0}^{k+1} \frac{(-1)^{2k+1-j}}{k+1} C_j^{k+1}$$
$$\Sigma_{n_1=1}^{\infty}\Sigma_{n_2=1}^{\infty} \dots \Sigma_{n_j=1}^{\infty}\Sigma_{i=0}^{\infty} \frac{(-1)^i}{n_1 n_2 \dots n_j\, i!}$$
$$\left[\frac{\alpha}{\beta}(n_1 + n_2 + \dots + n_j)\right]^i \left(e^{\beta x} - 1\right)^i$$

$$= \Sigma_{k=0}^{\infty}\Sigma_{j=0}^{k+1} \frac{(-1)^{2k+1-j}}{k+1} C_j^{k+1}$$
$$\Sigma_{n_1=1}^{\infty}\Sigma_{n_2=1}^{\infty} \dots \Sigma_{n_j=1}^{\infty}\Sigma_{i=0}^{\infty} \frac{(-1)^i}{n_1 n_2 \dots n_j\, i!}$$
$$\left[\frac{\alpha}{\beta}(n_1 + n_2 + \dots + n_j)\right]^i \Sigma_{t=0}^{i} C_t^i e^{t\beta x}(-1)^{i-t}$$

$$= \Sigma_{k=0}^{\infty}\Sigma_{j=0}^{k+1} \frac{(-1)^{2k-j+2i-t+1}}{k+1} C_j^{k+1}$$
$$\Sigma_{n_1=1}^{\infty}\Sigma_{n_2=1}^{\infty} \dots \Sigma_{n_j=1}^{\infty}\Sigma_{i=0}^{\infty} \frac{1}{n_1 n_2 \dots n_j\, i!}$$
$$\left[\frac{\alpha}{\beta}(n_1 + n_2 + \dots + n_j)\right]^i \Sigma_{t=0}^{i} C_t^i\, \Sigma_{r=0}^{\infty} \frac{(t\beta)^r}{r!}\, x^r$$

Therefore $E\left[\ln\left(-\ln\left[1 - e^{-\frac{\alpha}{\beta}(e^{\beta X}-1)}\right]\right)\right]$ will be

$$E\left[\ln\left(-\ln\left[1 - e^{-\frac{\alpha}{\beta}(e^{\beta X}-1)}\right]\right)\right] = \Sigma_{k=0}^{\infty}\Sigma_{j=0}^{k+1}\, \Sigma_{n_1,n_2,\dots,n_j=1}^{\infty}\Sigma_{i=0}^{\infty}\Sigma_{t=0}^{i}\Sigma_{r=0}^{\infty} C_j^{k+1}\, C_t^i$$
$$\frac{(-1)^{2k-j+2i-t+1}}{(k+1)n_1 n_2 \dots n_j\, i!} \left[\frac{\alpha}{\beta}(n_1 + n_2 + \dots + n_j)\right]^i \frac{(t\beta)^r}{r!} E(X^r) \qquad (23)$$

Also for $\left(\ln\left[1 - e^{-\frac{\alpha}{\beta}(e^{\beta x}-1)}\right]\right)^p$ by using $\ln(1-z)$, e^{-z}, $(a+b)^n$ and e^z we get

$$\left(\ln\left[1 - e^{-\frac{\alpha}{\beta}(e^{\beta x}-1)}\right]\right)^p = \left(-\sum_{n=1}^{\infty} \frac{\left(e^{-\frac{\alpha}{\beta}(e^{\beta x}-1)}\right)^n}{n}\right)^p$$

$$= (-1)^p \sum_{n_1=1}^{\infty}\sum_{n_2=1}^{\infty} \dots \sum_{n_p=1}^{\infty} \frac{e^{-\frac{\alpha}{\beta}(n_1+n_2+\dots+n_p)(e^{\beta x}-1)}}{n_1 n_2 \dots n_p}$$

$$= \sum_{n_1=1}^{\infty}\sum_{n_2=1}^{\infty} \dots \sum_{n_p=1}^{\infty}\sum_{i=0}^{\infty} \frac{(-1)^{i+p}}{n_1 n_2 \dots n_p\, i!}$$
$$\left(\frac{\alpha}{\beta}(n_1 + n_2 + \dots + n_p)\right)^i \left(e^{\beta x} - 1\right)^i$$

$$= \sum_{n_1=1}^{\infty}\sum_{n_2=1}^{\infty} \dots \sum_{n_p=1}^{\infty}\sum_{i=0}^{\infty} \frac{(-1)^{i+p}}{n_1 n_2 \dots n_p\, i!}$$
$$\left(\frac{\alpha}{\beta}(n_1 + n_2 + \dots + n_p)\right)^i \sum_{i_1=0}^{i} C_{i_1}^i e^{\beta i_1 x} (-1)^{i-i_1}$$

$$= \sum_{n_1=1}^{\infty}\sum_{n_2=1}^{\infty}\cdots\sum_{n_p=1}^{\infty}\sum_{i=0}^{\infty}\sum_{i_1=0}^{i}\frac{(-1)^{2i-i_1+p}}{n_1\dots n_p\ i!}$$
$$\left(\frac{\alpha}{\beta}(n_1+n_2+\cdots+n_p)\right)^i C_{i_1}^i\sum_{r=0}^{\infty}\frac{(\beta i_1)^r}{r!}x^r$$

Therefore $E\left[\left(\ln\left[1-e^{-\frac{\alpha}{\beta}(e^{\beta X}-1)}\right]\right)^p\right]$ will be

$$E\left[\left(\ln\left[1-e^{-\frac{\alpha}{\beta}(e^{\beta X}-1)}\right]\right)^p\right] = \sum_{n_1,n_2,\dots,n_p=1}^{\infty}\sum_{i=0}^{\infty}\sum_{i_1=0}^{i}\sum_{r=0}^{\infty}\frac{(-1)^{2i-i_1+p}}{n_1 n_2\dots n_p\ i!}$$
$$\left(\frac{\alpha}{\beta}(n_1+n_2+\cdots+n_p)\right)^i C_{i_1}^i\frac{(\beta i_1)^r}{r!}E(X^r) \qquad (24)$$

Now from (21) the $GG-GGo$ Shannon entropy is given by

$$SH_{GG-GGo} = \ln\left(\frac{\Gamma\left(\frac{d}{p}\right)}{\frac{p}{a^d}\alpha\lambda}\right)-\frac{\alpha}{\beta}-(d-1)\ln(\lambda)-\beta E(X)$$
$$+\alpha\sum_{k=0}^{\infty}\frac{\beta^{k-1}}{k!}E(X^k)+E\left(\ln\left[1-e^{-\frac{\alpha}{\beta}(e^{\beta X}-1)}\right]\right)$$
$$-(d-1)E\left[\ln\left(-\ln\left[1-e^{-\frac{\alpha}{\beta}(e^{\beta X}-1)}\right]\right)\right]$$
$$+\frac{(-\lambda)^p}{a^p}E\left[\left(\ln\left[1-e^{-\frac{\alpha}{\beta}(e^{\beta X}-1)}\right]\right)^p\right] \qquad (25)$$

where

$E\left(\ln\left[1-e^{-\frac{\alpha}{\beta}(e^{\beta X}-1)}\right]\right)$, $E\left[\ln\left(-\ln\left[1-e^{-\frac{\alpha}{\beta}(e^{\beta X}-1)}\right]\right)\right]$ and $E\left[\left(\ln\left[1-e^{-\frac{\alpha}{\beta}(e^{\beta X}-1)}\right]\right)^p\right]$ as in (22), (23) and (24) respectively.

$E(X)$ and $E(X^k)$ as in (17) with $r=1$ and $r=k$ respectively.

Relative Entropy: The $GG-GGo$ relative entropy can be obtained from $\int_0^{\infty}\ln\left(\frac{f(x)_{GG-GGo}}{f_1(x)_{GG-GGo}}\right)f(x)_{GG-GGo}\,dx$. Taking the natural logarithm of $f(x)_{GG-GGo}$ in (12) w.r.t. $f_1(x)_{GG-GGo}$ with parameters $(a_1, d_1, p_1, \alpha_1, \beta_1, \lambda_1)$ we get

$$\ln\left(\frac{f(x)_{GG-GGo}}{f_1(x)_{GG-GGo}}\right) = \ln\left(\frac{\frac{p}{a^d}\alpha\lambda\Gamma\left(\frac{d_1}{p_1}\right)}{\frac{p_1}{a_1^{d_1}}\alpha_1\lambda_1\Gamma\left(\frac{d}{p}\right)}\right)+\frac{\alpha}{\beta}-\frac{\alpha_1}{\beta_1}+(d-1)\ln(\lambda)-(d_1-1)\ln(\lambda_1)$$
$$+\beta x-\beta_1 x-\frac{\alpha}{\beta}\sum_{k=0}^{\infty}\frac{\beta^k}{k!}x^k+\frac{\alpha_1}{\beta_1}\sum_{k=0}^{\infty}\frac{\beta_1{}^k}{k!}x^k-\ln\left[1-e^{-\frac{\alpha}{\beta}(e^{\beta x}-1)}\right]$$
$$+\ln\left[1-e^{-\frac{\alpha_1}{\beta_1}(e^{\beta_1 x}-1)}\right]+(d-1)\ln\left(-\ln\left[1-e^{-\frac{\alpha}{\beta}(e^{\beta x}-1)}\right]\right)$$
$$-(d_1-1)\ln\left(-\ln\left[1-e^{-\frac{\alpha_1}{\beta_1}(e^{\beta_1 x}-1)}\right]\right)$$
$$-\frac{(-\lambda)^p}{a^p}\left(\ln\left[1-e^{-\frac{\alpha}{\beta}(e^{\beta x}-1)}\right]\right)^p+\frac{(-\lambda_1)^{p_1}}{a_1^{p_1}}\left(\ln\left[1-e^{-\frac{\alpha_1}{\beta_1}(e^{\beta_1 x}-1)}\right]\right)^{p_1}$$

Then $GG-GGo$ relative entropy will be

$$RE_{GG-GGo} = \ln\left(\frac{\frac{p}{a^{d}}\alpha\lambda\Gamma\left(\frac{d_1}{p_1}\right)}{\frac{p_1}{a_1^{d_1}}\alpha_1\lambda_1\Gamma\left(\frac{d}{p}\right)}\right) + \frac{\alpha}{\beta} - \frac{\alpha_1}{\beta_1} + (d-1)\ln(\lambda) - (d_1-1)\ln(\lambda_1)$$
$$+(\beta-\beta_1)E(X) - \alpha\sum_{k=0}^{\infty}\frac{\beta^{k-1}}{k!}E(X^k)$$
$$+\alpha_1\sum_{k=0}^{\infty}\frac{\beta_1{}^{k-1}}{k!}E(X^k) - E\left(\ln\left[1-e^{-\frac{\alpha}{\beta}(e^{\beta x}-1)}\right]\right)$$
$$+E\left(\ln\left[1-e^{-\frac{\alpha_1}{\beta_1}(e^{\beta_1 x}-1)}\right]\right) + (d-1)E\left[\ln\left(-\ln\left[1-e^{-\frac{\alpha}{\beta}(e^{\beta x}-1)}\right]\right)\right]$$
$$-(d_1-1)E\left[\ln\left(-\ln\left[1-e^{-\frac{\alpha_1}{\beta_1}(e^{\beta_1 x}-1)}\right]\right)\right]$$
$$-\frac{(-\lambda)^p}{a^p}E\left[\left(\ln\left[1-e^{-\frac{\alpha}{\beta}(e^{\beta x}-1)}\right]\right)^p\right] + \frac{(-\lambda_1)^{p_1}}{a_1^{p_1}}E\left[\left(\ln\left[1-e^{-\frac{\alpha_1}{\beta_1}(e^{\beta_1 x}-1)}\right]\right)^{p_1}\right] \qquad (26)$$

where
$E(X)$ and $E(X^k)$ as in (17) with $r=1$ and $r=k$ respectively and the other expectations recall from (22), (23) and (24) with specified parameters.

Stress Strength: Let Y and X be present the stress strength random variables that independent of each other follows respectively $GG-GGo$ with different parameters, then the stress strength can be obtained by

$$SS_{GG-GGo} = P(Y<X) = \int_0^{\infty} f_X(x)_{GG-GGo}\, F_Y(x)\, dx$$

where

$$F_Y(x) = 1 - \frac{\gamma\left[\frac{d_1}{p_1}, \left(\frac{1}{a_1}\ln\left[1-e^{-\frac{\alpha_1}{\beta_1}(e^{\beta_1 x}-1)}\right]^{-\lambda_1}\right)^{p_1}\right]}{\Gamma\left(\frac{d_1}{p_1}\right)}$$

Now in order to find the stress strength we firstly expansion the incomplete gamma function in the $F_Y(x)$ according to $\gamma(u,z) = \Sigma_{i=0}^{\infty}\frac{(-1)^i z^{u+i}}{i!(u+i)}$ as

$$F_Y(x) = 1 - \frac{1}{\Gamma\left(\frac{d_1}{p_1}\right)}\Sigma_{m=0}^{\infty}\frac{(-1)^m}{m!\left(\frac{d_1}{p_1}+m\right)}\left[\left(\frac{1}{a_1}\ln\left[1-e^{-\frac{\alpha_1}{\beta_1}(e^{\beta_1 x}-1)}\right]^{-\lambda_1}\right)^{p_1}\right]^{\frac{d_1}{p_1}+m}$$
$$= 1 - \frac{1}{\Gamma\left(\frac{d_1}{p_1}\right)}\Sigma_{m=0}^{\infty}\frac{(-1)^m\left(\frac{-\lambda_1}{a_1}\right)^{d_1+mp_1}}{m!\left(\frac{d_1}{p_1}+m\right)}\left(\ln\left[1-e^{-\frac{\alpha_1}{\beta_1}(e^{\beta_1 x}-1)}\right]\right)^{d_1+mp_1}$$

Recall the same steps of getting $\left(\ln\left[1-e^{-\frac{\alpha}{\beta}(e^{\beta x}-1)}\right]\right)^p$ we can get

$$\left(\ln\left[1-e^{-\frac{\alpha_1}{\beta_1}\left(e^{\beta_1 x}-1\right)}\right]\right)^{d_1+mp_1} = \sum_{n_1=1}^{\infty}\sum_{n_2=1}^{\infty}\dots\sum_{n_{d_1+mp_1}=1}^{\infty}\sum_{i=0}^{\infty}\sum_{i_1=0}^{i}\frac{(-1)^{2i-i_1+d_1+mp_1}}{n_1\dots n_{d_1+mp_1}\; i!} C_{i_1}^{i}\left(\frac{\alpha_1}{\beta_1}\left(n_1+n_2+\dots+n_{d_1+mp_1}\right)\right)^{i}\sum_{r=0}^{\infty}\frac{(\beta_1 i_1)^r}{r!}x^r$$

Then $F_Y(x)$ will be

$$F_Y(x) = 1-\frac{1}{\Gamma\left(\frac{d_1}{p_1}\right)}\sum_{n_1=1}^{\infty}\sum_{n_2=1}^{\infty}\dots\sum_{n_{d_1+mp_1}=1}^{\infty}\sum_{i=0}^{\infty}\sum_{i_1=0}^{i}\sum_{m=0}^{\infty}\frac{(-1)^{2i-i_1+d_1+mp_1+m}}{n_1\dots n_{d_1+mp_1}\; i!}\frac{\left(\frac{-\lambda_1}{a_1}\right)^{d_1+mp_1}}{m!\left(\frac{d_1}{p_1}+m\right)}C_{i_1}^{i}\left(\frac{\alpha_1}{\beta_1}\left(n_1+n_2+\dots+n_{d_1+mp_1}\right)\right)^{i}\sum_{r=0}^{\infty}\frac{(\beta_1 i_1)^r}{r!}x^r \qquad (27)$$

Based on (27) the $GG-GGo$ stress strength can be obtained as

$$SS_{GG-GGo} = 1-\frac{1}{\Gamma\left(\frac{d_1}{p_1}\right)}\sum_{n_1=1}^{\infty}\sum_{n_2=1}^{\infty}\dots\sum_{n_{d_1+mp_1}=1}^{\infty}\sum_{i=0}^{\infty}\sum_{i_1=0}^{i}\sum_{m=0}^{\infty}\sum_{r=0}^{\infty}\frac{(-1)^{2i-i_1+d_1+mp_1+m}}{n_1\dots n_{d_1+mp_1}\; i!}\frac{\left(\frac{-\lambda_1}{a_1}\right)^{d_1+mp_1}}{m!\left(\frac{d_1}{p_1}+m\right)}C_{i_1}^{i}\left(\frac{\alpha_1}{\beta_1}\left(n_1+n_2+\dots+n_{d_1+mp_1}\right)\right)^{i}\frac{(\beta_1 i_1)^r}{r!}E(X^r) \qquad (28)$$

where $E(X^r)$ as in (17) .

Conclusions

Here we present H-G generated class of continuous distributions on the basis of composing two cumulative distribution functions. A new family of this class named generalized gamma - G distributions along with generalized gamma - generalized Gompertz as a special case are discussed. The essential functions "cdf, pdf, reliability and hazard rate" are presented. In addition, the vital statistical properties such as r-th moment, characteristic function, quantile function and simulated data, Shannon and relative entropies besides stress strength are attained.

References

[1] Mudholkar G S and Kollia G D 1994 Generalized Weibull family: A structural analysis *Commun. Statist. Theor. Meth.* **23**(4) pp 1149-71.

[2] Mudholkar G S and Srivastava D K 1993 Exponentiated Weibull family for analyzing bathtub failure rate data *IEEE Transactions on Reliability* **42**(2) pp 299-302.

[3] Mudholkar G S, Srivastava D K and Freimer M 1995 The exponentiated Weibull family: A reanalysis of the bus-motor-failure data *Technometrics* **37**(4) pp 436-45.

[4] Marshall A W and Olkin I 1997 A new method for adding a parameter to a family of distributions with application to the exponential and Weibull families *Biometrika* **84**(3) pp 641-52.

[5] Gupta R C, Gupta P I and Gupta R D 1998 Modeling failure time data by Lehmann alternatives *Communications in statistics-Theory and Methods* **27**(4) pp 887-904

[6] Ahmad Z, Elgarhy M and Hamedani G G 2018 A new Weibull-X family of distributions: properties, characterizations and applications *Journal of Statistical Distributions and Applications* **5**(5) pp 1-18.

[7] Hamedani G G, Altun E, Korkmaz M Ç, Yousof H M and Butt N S 2018 A new extended G family of

continuous distributions with mathematical properties, characterizations and regression modelling *Pak.j.stat.oper.res.* **14**(3) pp 737-58.
[8] Alizadeh M, Rasekhi M, Yousof H M and Hamedani G G 2018 The transmuted Weibull-G family of distributions *Hacettepe Journal of Mathematics and Statistics* **47**(6) pp 1671-89.
[9] Hosseini B, Afshari M and Alizadeh M 2018 The generalized odd Gamma-G family of distributions: properties and applications *Austrian Journal of Statistics* **47**(2) pp 69-89.
[10] Korkmaz M A 2019 A new family of the continuous distributions: The extended Weibull-G family *Commun. Fac. Sci. Univ. Ank. Ser.* **68**(1) pp 248-70.
[11] Jamal F and Nasir M 2019 Generalized Burr X family of distributions *International Journal of Mathematics and Statistics* **19**(1) pp 55-73.
[12] Stacy E W 1962 A generalization of the gamma distribution *Annals of Mathematical Statistics* **33**(3) pp 1187-92.
[13] Shin J W, Chang J H and Kim N S 2005 Statistical modeling of speech signals based on generalized gamma distribution *IEEE Signal Processing Letters* **12**(3) pp 258-61.
[14] Nadarajah S and Rocha R 2016 Newdistns: An R package for new families of distributions *Journal of Statistical Software* **69**(10) pp 1-32.
[15] El-Gohary A, Alshamrani A and Al-Otaibi A N 2013 The generalized Gompertz distribution *Applied Mathematical Modelling* **37**(1-2) pp 13-24.
[16] Tahmasebi S and Jafari A A 2016 Generalized Gompertz-power series distributions *Hacettepe Journal of Mathematics and Statistics* **45**(5) pp 1579-1604.

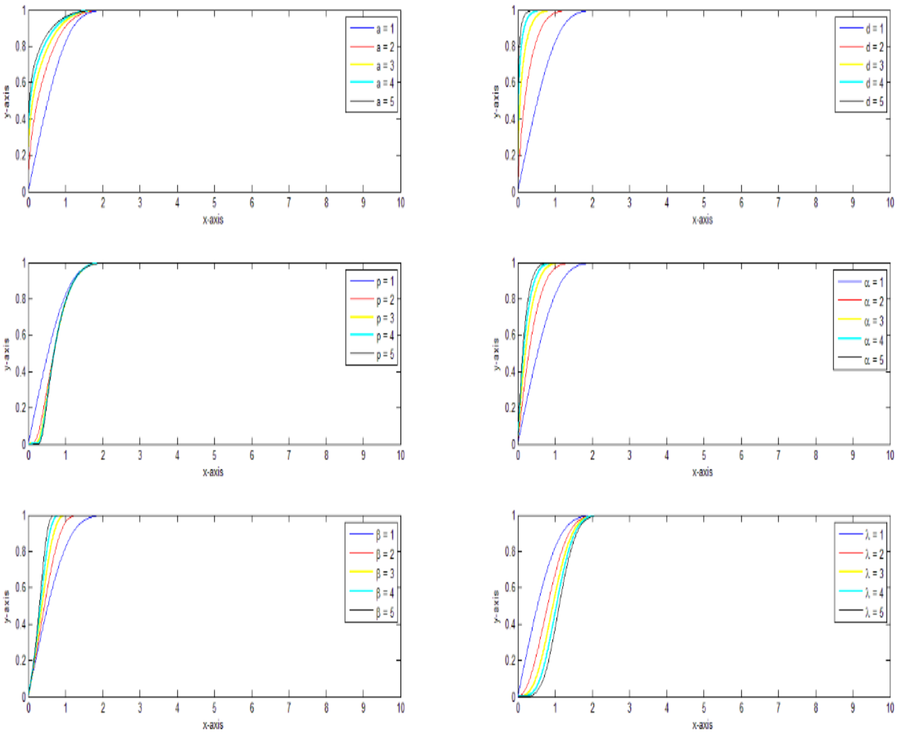


Figure 1. Plots of the $GG - GGo$ cumulative distribution for some parameter values and the others equal to 1.

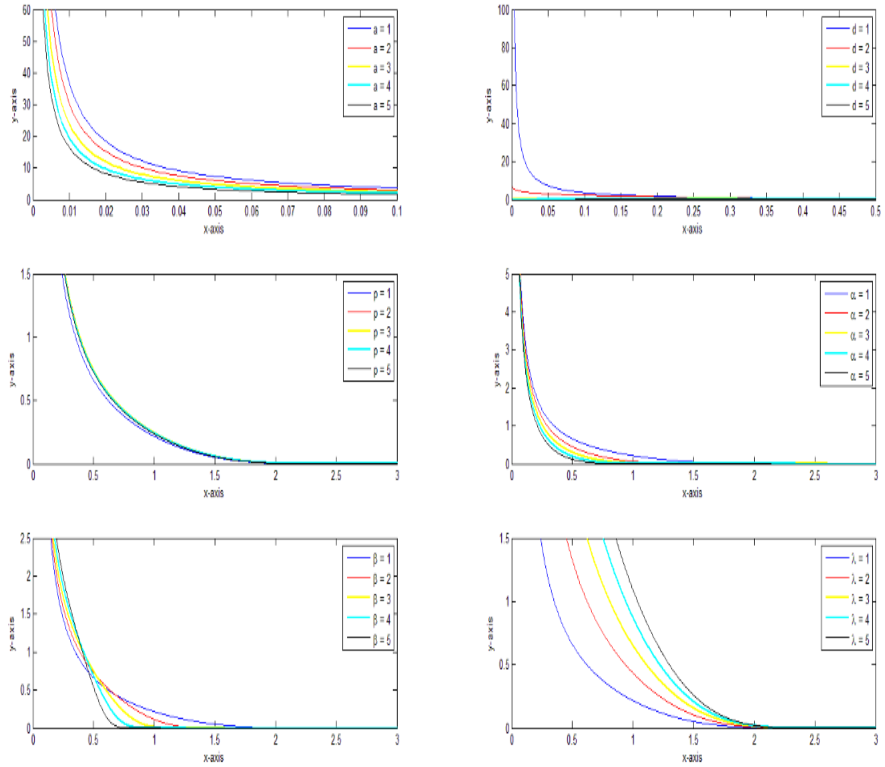


Figure 2. Plots of the $GG-GGo$ density for some parameter values and the others equal to 1.

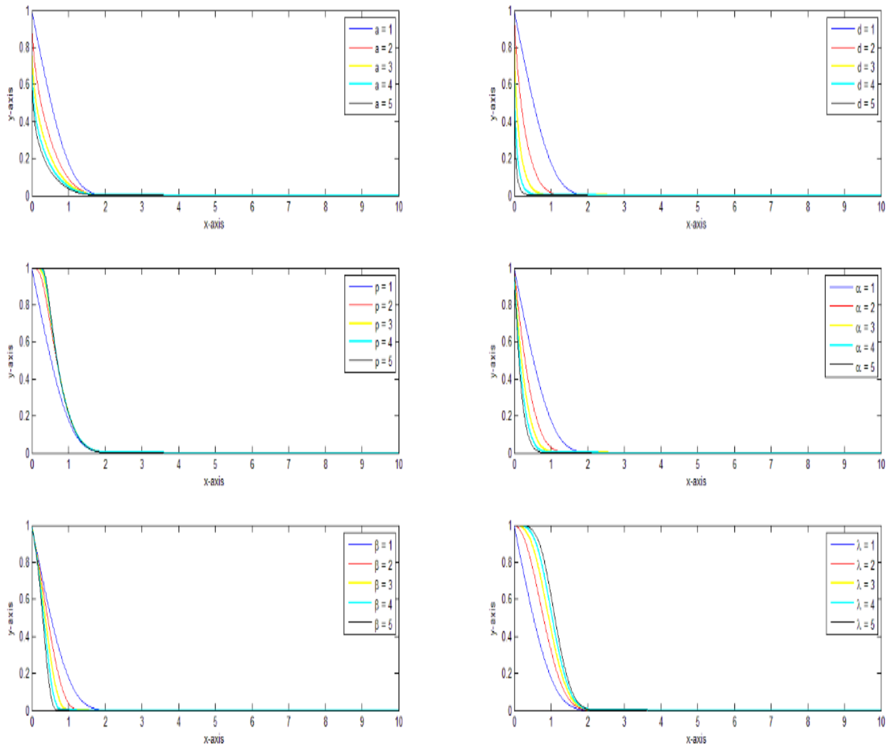


Figure 3. Plots of the $GG-GGo$ reliability function for some parameter values and the others equal to 1.

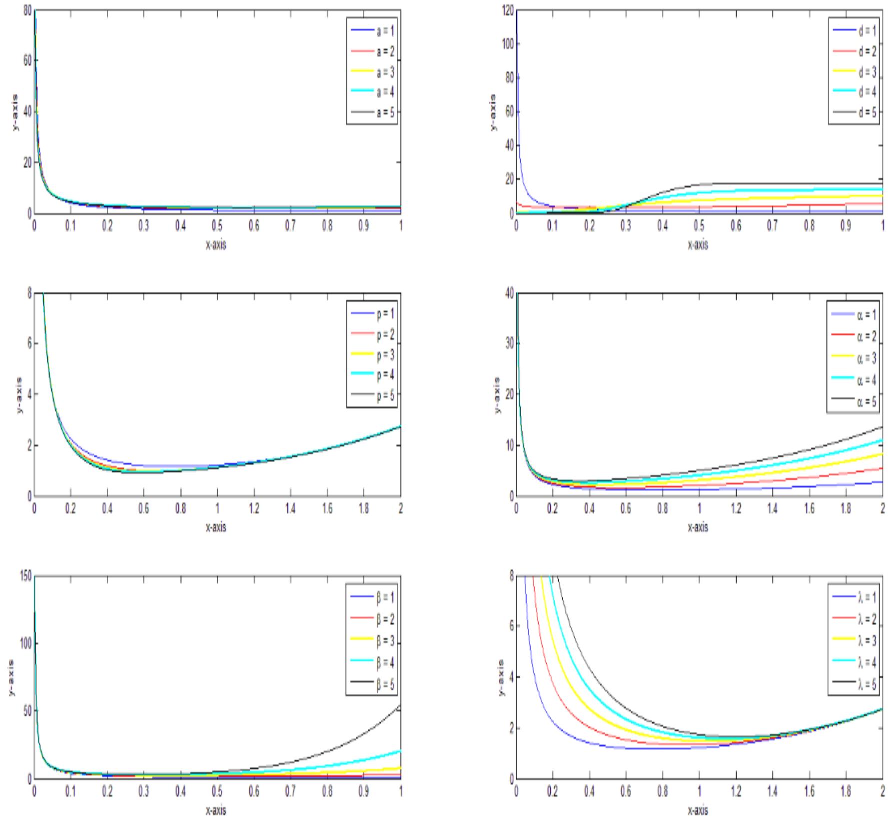


Figure 4. Plots of the $GG - GGo$ hazard rate function for some parameter values and the others equal to 1.

Properties of two Doubly-Truncated Generalized Distributions

Salah Hamza Abid **Hind Husham Jani**

abidsalah@gmail.com **hind.mathematic13@gmail.com**
Mathematics department / University of Mustansiriyah / Bagdad- Iraq

Abstract : In this paper, some properties of doubly truncated generalized gamma distribution and doubly truncated Generalized Invers Weibull distribution are derived. These properties are the reliability and hazard functions, rth raw moments, stress-strength reliability, Shannon entropy and relative entropy.
Keywords: Doubly truncated Generalized Gamma distribution, Doubly truncated Generalized Inverse Weibull distribution, hazard function, characteristic function, Shannon entropy, relative entropy, Stress-strength reliability.

I. **Introduction**

In recent years, many authors have concentrated their consideration on the suggestion of new and more flexible probability distributions, established using different techniques to represent a set of data. Properties of a distribution are very useful to show the ability of that distribution.

The doubly truncated distributions are more realistic to represent phenomena, without losing the generality, since the truncated parameters can take any values.

In this paper, properties of doubly truncated generalized gamma distribution (DTGG) and Doubly truncated Generalized Invers Weibull distribution (DTGIW) are derived. These properties are the rth raw moments, stress-strength reliability, Shannon entropy and relative entropy.

II. **The generalized gamma distribution**

The generalized Gamma distribution GGD is a continuous probability distribution with three parameters, Presented by Stacy in 1962. It is contains some of important densities as special cases as Exponential, Gamma, Weibull , half-Normal and lognormal distributions. A lot of literature has been written about GGD, some of which will be mentioned below. Khodabin and Ahmadabadi in 2010 derived some other properties of GGD with Kullback-Leibler discrimination , Akaike and Bayesian information criterion. Cordeiro et al. 2011 derived another generalization of Stacy's GGD using exponentiated method, and applied it to life time and survival analysis. Cox and Matheson in 2014 compared exponentiated Weibull (EW) and matching GG distributions graphically and using the Kullback-Leibler distance. They found that the survival functions for the EW and matching GG are graphically indistinguishable, and only the hazard functions can sometimes be seen to be slightly different. In 2017, Abid and Abdulrazak presented [0,1] truncated Frechet GGD. They derived the distribution properties such as reliability function, hazard function, the rth raw moment function, Stress-Strength reliability, Shannon and Relative entropies. Barriga in 2018 defined a new extension of the GGD based on the generator pioneered by Marshall and Olkin in 1997. It is shown by Kiche et al in 2019 that GGD has three sub-families and its application to the analysis of a survival data has also been explored.

The probability density function of the GG variable is,

$$f(x) = \frac{p}{\epsilon^d}\, x^{d-1}\, e^{-\left(\frac{x}{\epsilon}\right)^p} / \Gamma(d/p) \qquad (1)$$

And the cumulative distribution function is,

$$F(x) = \gamma((x/\epsilon)^p, d/p\ )/\Gamma(d/p) \qquad (2)$$

Where $\Gamma(.)$ denotes the gamma function which defined as $\Gamma(s) = \int_0^\infty t^{s-1}e^{-t}\,dt$, and $\gamma\left(\left(\frac{x}{\epsilon}\right)^p, \frac{d}{p}\right)$ denotes the lower incomplete gamma function, which is generally can written as

$$\gamma(z,s) = \int_0^s t^{z-1}e^{-t}\,dt = \sum_{s_1=0}^{\infty} \frac{(-1)^{s_1}}{s_1!}\frac{z^{s+s_1}}{s+s_1} \qquad (3)$$

The upper incomplete gamma function also is, $\Gamma(s,x) = \int_x^\infty t^{s-1}e^{-t}\,dt,$ where, $\gamma(z,s) = \Gamma(s) - \Gamma(s,x)$.

1 II.1 Essential properties of *DTGG* (ϵ, d, p, a, b)

We consider here the doubly truncated Generalized Gamma Distribution *DTGG* (ϵ, d, p, a, b) of random variable X, where the lower and upper limits are a and b respectively, then the probability density function (pdf) and the commutative density function (cdf) are respectively,

$$g(x) = \frac{\frac{p}{\epsilon^d} x^{d-1} e^{-\left(\frac{x}{\epsilon}\right)^p}}{\gamma\left(\left(\frac{b}{\epsilon}\right)^p, \frac{d}{p}\right) - \gamma\left(\left(\frac{a}{\epsilon}\right)^p, \frac{d}{p}\right)} \quad , a < x < b \qquad (4)$$

$$G(x) = \frac{\gamma\left(\left(\frac{x}{\epsilon}\right)^p, \frac{d}{p}\right) - \gamma\left(\left(\frac{a}{\epsilon}\right)^p, \frac{d}{p}\right)}{\gamma\left(\left(\frac{b}{\epsilon}\right)^p, \frac{d}{p}\right) - \gamma\left(\left(\frac{a}{\epsilon}\right)^p, \frac{d}{p}\right)} \quad , a < x < b \qquad (5)$$

Then, the reliability and the hazard functions of X are respectively,

$$R(x) = 1 - G(x) = \frac{\gamma\left(\left(\frac{b}{\epsilon}\right)^p, \frac{d}{p}\right) - \gamma\left(\left(\frac{x}{\epsilon}\right)^p, \frac{d}{p}\right)}{\gamma\left(\left(\frac{b}{\epsilon}\right)^p, \frac{d}{p}\right) - \gamma\left(\left(\frac{a}{\epsilon}\right)^p, \frac{d}{p}\right)} \qquad (6)$$

$$h(x) = \frac{g(x)}{R(x)} = \frac{\frac{p}{\epsilon^d} x^{d-1} e^{-\left(\frac{x}{\epsilon}\right)^p}}{\gamma\left(\left(\frac{b}{\epsilon}\right)^p, \frac{d}{p}\right) - \gamma\left(\left(\frac{x}{\epsilon}\right)^p, \frac{d}{p}\right)} \qquad (7)$$

Since, $\int_a^b x^{d-1} e^{-\left(\frac{x}{\epsilon}\right)^p} dx = \frac{\epsilon^d}{p} \left\{\gamma\left(\left(\frac{b}{\epsilon}\right)^p, \frac{d}{p}\right) - \gamma\left(\left(\frac{a}{\epsilon}\right)^p, \frac{d}{p}\right)\right\}$ (8)

Then, the rth raw moment of *DTGG* (ϵ, d, p, a, b) distribution is,

$$E(X^r) = \epsilon^r \frac{\left\{\gamma\left(\left(\frac{b}{\epsilon}\right)^p, \frac{d+r}{p}\right) - \gamma\left(\left(\frac{a}{\epsilon}\right)^p, \frac{d+r}{p}\right)\right\}}{\left\{\gamma\left(\left(\frac{b}{\epsilon}\right)^p, \frac{d}{p}\right) - \gamma\left(\left(\frac{a}{\epsilon}\right)^p, \frac{d}{p}\right)\right\}} \qquad (9)$$

Then, the characteristic function can easily get by using the relation,

$$Q_x(t) = E\left(e^{ixt}\right) = \sum_{r=0}^{\infty} \frac{(it)^r}{r!} E(x^r) \ , \ \text{since} \ e^{ixt} = \sum_{r=0}^{\infty} \frac{(it)^r}{r!} x^r$$

2 II.2 Stress-Strength Reliability

Inferences about $R = P[Y < X]$, where X and Y are two independent random variables, is very common in the reliability literature. For example, if X is the strength of a component which is subject to a stress Y, then R is a measure of system performance and arises in the context of mechanical reliability of a system. The system fails if and only if at any time the applied stress is greater than its strength. Let Y and X be the stress and the strength random variables, independent of each other, follow respectively $DTGG(\epsilon_2, d_2, p_2, a, b)$ and $DTGG(\epsilon_1, d_1, p_1, a, b)$, then,

$$R = P(Y < X) = \int_a^b f_X(x) F_Y(x)\ dx$$

$$R = \int_a^b \frac{\frac{p_1}{\epsilon_1^{d_1}}}{\gamma\left(\left(\frac{b}{\epsilon_1}\right)^{p_1}, \frac{d_1}{p_1}\right) - \gamma\left(\left(\frac{a}{\epsilon_1}\right)^{p_1}, \frac{d_1}{p_1}\right)} x^{d_1-1} e^{-\left(\frac{x}{\epsilon_1}\right)^{p_1}} \left[\frac{\gamma\left(\left(\frac{x}{\epsilon_2}\right)^{p_2}, \frac{d_2}{p_2}\right) - \gamma\left(\left(\frac{a}{\epsilon_2}\right)^{p_2}, \frac{d_2}{p_2}\right)}{\gamma\left(\left(\frac{b}{\epsilon_2}\right)^{p_2}, \frac{d_2}{p_2}\right) - \gamma\left(\left(\frac{a}{\epsilon_2}\right)^{p_2}, \frac{d_2}{p_2}\right)}\right] dx$$

Let $K = \frac{1}{\left[\gamma\left(\left(\frac{b}{\epsilon_1}\right)^{p_1}, \frac{d_1}{p_1}\right) - \gamma\left(\left(\frac{a}{\epsilon_1}\right)^{p_1}, \frac{d_1}{p_1}\right)\right]\left[\gamma\left(\left(\frac{b}{\epsilon_2}\right)^{p_2}, \frac{d_2}{p_2}\right) - \gamma\left(\left(\frac{a}{\epsilon_2}\right)^{p_2}, \frac{d_2}{p_2}\right)\right]}$

$$R = K \frac{p_1}{\epsilon_1^{d_1}} \left[\int_a^b x^{d_1-1} e^{-\left(\frac{x}{\epsilon_1}\right)^{p_1}} \gamma\left(\left(\frac{x}{\epsilon_2}\right)^{p_2}, \frac{d_2}{p_2}\right) dx - \int_a^b x^{d_1-1} e^{-\left(\frac{x}{\epsilon_1}\right)^{p_1}} \gamma\left(\left(\frac{a}{\epsilon_2}\right)^{p_2}, \frac{d_2}{p_2}\right) dx\right]$$

By using (3), we get

$$R = K \frac{p_1}{\epsilon_1^{d_1}} \left[\begin{array}{c} \sum_{s_1=0}^{\infty} \frac{(-1)^{s_1}}{s_1!} \frac{(\epsilon_2)^{-(d_2+s_1p_2)}}{\left(\frac{d_2}{p_2}+s_1\right)} \int_a^b x^{d_1+d_2+s_1p_2-1} e^{-\left(\frac{x}{\epsilon_1}\right)^{p_1}} dx \\ -\gamma\left(\left(\frac{a}{\epsilon_2}\right)^{p_2}, \frac{d_2}{p_2}\right) \int_a^b x^{d_1-1} e^{-\left(\frac{x}{\epsilon_1}\right)^{p_1}} dx \end{array}\right]$$, Then by (8)

$$R = K \frac{p_1}{\epsilon_1^{d_1}} \left[\begin{array}{c} \sum_{s_1=0}^{\infty} \frac{(-1)^{s_1}}{s_1!} \frac{(\epsilon_2)^{-(d_2+s_1p_2)}}{\left(\frac{d_2}{p_2}+s_1\right)} \frac{\epsilon_1^{d_1+d_2+s_1p_2}}{p_1} \\ \left\{\gamma\left(\left(\frac{b}{\epsilon_1}\right)^{p_1}, \frac{d_1+d_2+s_1p_2}{p_1}\right) - \gamma\left(\left(\frac{a}{\epsilon_1}\right)^{p_1}, \frac{d_1+d_2+s_1p_2}{p_1}\right)\right\} \\ -\gamma\left(\left(\frac{a}{\epsilon_2}\right)^{p_2}, \frac{d_2}{p_2}\right) \frac{\epsilon_1^{d_1}}{p_1} \left\{\gamma\left(\left(\frac{b}{\epsilon_1}\right)^{p_1}, \frac{d_1}{p_1}\right) - \gamma\left(\left(\frac{a}{\epsilon_1}\right)^{p_1}, \frac{d_1}{p_1}\right)\right\} \end{array}\right]$$

So the stress- strength reliability of *DTGG* is,

$$R = K \left[\begin{array}{c} \sum_{s_1=0}^{\infty} \frac{(-1)^{s_1}}{s_1!} \frac{(\epsilon_2)^{-(d_2+s_1p_2)}}{\left(\frac{d_2}{p_2}+s_1\right)} \epsilon_1^{d_2+s_1p_2} \\ \left\{\gamma\left(\left(\frac{b}{\epsilon_1}\right)^{p_1}, \frac{d_1+d_2+s_1p_2}{p_1}\right) - \gamma\left(\left(\frac{a}{\epsilon_1}\right)^{p_1}, \frac{d_1+d_2+s_1p_2}{p_1}\right)\right\} \\ -\gamma\left(\left(\frac{a}{\epsilon_2}\right)^{p_2}, \frac{d_2}{p_2}\right) \left\{\gamma\left(\left(\frac{b}{\epsilon_1}\right)^{p_1}, \frac{d_1}{p_1}\right) - \gamma\left(\left(\frac{a}{\epsilon_1}\right)^{p_1}, \frac{d_1}{p_1}\right)\right\} \end{array}\right] \quad (10)$$

II.3 Shannon entropy

The Shannon entropy of a random variable X is a measure of variation of the uncertainty. It is defined for a random variable X with values in a finite set X as $H = E(-\ln(g(x)))$. So, the Shannon entropy of *DTGG* random variable is,

$$H = E\left(-\ln\left(\frac{\frac{p}{\epsilon^d}}{\gamma\left(\left(\frac{b}{\epsilon}\right)^p,\frac{d}{p}\right)-\gamma\left(\left(\frac{a}{\epsilon}\right)^p,\frac{d}{p}\right)} X^{d-1}\, e^{-\left(\frac{X}{\epsilon}\right)^p}\right)\right)$$

$$= E\left(-\ln\left(\frac{\frac{p}{\epsilon^d}}{\gamma\left(\left(\frac{b}{\epsilon}\right)^p,\frac{d}{p}\right)-\gamma\left(\left(\frac{a}{\epsilon}\right)^p,\frac{d}{p}\right)}\right) + -(d-1)\ln(X) + \left(\frac{X}{\epsilon}\right)^p\right)$$

$$= -\ln\left(\frac{\frac{p}{\epsilon^d}}{\gamma\left(\left(\frac{b}{\epsilon}\right)^p,\frac{d}{p}\right)-\gamma\left(\left(\frac{a}{\epsilon}\right)^p,\frac{d}{p}\right)}\right) - (d-1)E(\ln(X)) + E\left(\left(\frac{X}{\epsilon}\right)^p\right)$$

Since, $E(\ln(X)) = \int_a^b \ln(x) \frac{\frac{p}{\epsilon^d}}{\gamma\left(\left(\frac{b}{\epsilon}\right)^p,\frac{d}{p}\right)-\gamma\left(\left(\frac{a}{\epsilon}\right)^p,\frac{d}{p}\right)} x^{d-1}\, e^{-\left(\frac{x}{\epsilon}\right)^p}\, dx$

Let $A_1 = \frac{\frac{p}{\epsilon^d}}{\gamma\left(\left(\frac{b}{\epsilon}\right)^p,\frac{d}{p}\right)-\gamma\left(\left(\frac{a}{\epsilon}\right)^p,\frac{d}{p}\right)}$, then,

$$E(\ln(X)) = A_1 \int_a^b \ln(x)\, x^{d-1}\, e^{-\left(\frac{x}{\epsilon}\right)^p}\, dx = A_1 \int_a^b \ln(x)\, x^{d-1} \sum_{s_1=0}^{\infty} \frac{\left(-\left(\frac{x}{\epsilon}\right)^p\right)^{s_1}}{s_1!}\, dx$$

$$= A_1 \sum_{s_1=0}^{\infty} \frac{\left(-\left(\frac{1}{\epsilon}\right)^p\right)^{s_1}}{s_1!} \int_a^b \ln(x)\, x^{d+ps_1-1}\, dx$$

Since, $\int_a^b x^m \ln(x)\, dx = x^{m+1}\left\{\frac{\ln(x)}{m+1} - \frac{1}{(m+1)^2}\right\}$, then,

$$E(\ln(X)) = A_1 \sum_{s_1=0}^{\infty} \frac{(-1)^{s_1}}{s_1!(\epsilon)^{ps_1}}\left[b^{d+ps_1}\left\{\frac{\ln(b)}{d+ps_1} - \frac{1}{(d+ps_1)^2}\right\} - a^{d+ps_1}\left\{\frac{\ln(a)}{d+ps_1} - \frac{1}{(d+ps_1)^2}\right\}\right] \quad (11)$$

And by using (9), we get,

$$E\left(\left(\frac{X}{\epsilon}\right)^p\right) = (\epsilon)^{-p} E((X)^p) = \frac{\left\{\gamma\left(\left(\frac{b}{\epsilon}\right)^p,\frac{d+p}{p}\right)-\gamma\left(\left(\frac{a}{\epsilon}\right)^p,\frac{d+p}{p}\right)\right\}}{\left\{\gamma\left(\left(\frac{b}{\epsilon}\right)^p,\frac{d}{p}\right)-\gamma\left(\left(\frac{a}{\epsilon}\right)^p,\frac{d}{p}\right)\right\}} \quad (12)$$

Then the entropy of DTGG random variable is:-

$$H = -\ln\left(\frac{\frac{p}{\epsilon^d}}{\gamma\left(\left(\frac{b}{\epsilon}\right)^p,\frac{d}{p}\right)-\gamma\left(\left(\frac{a}{\epsilon}\right)^p,\frac{d}{p}\right)}\right)$$

$$-(d-1)A_1 \sum_{s_1=0}^{\infty} \frac{(-1)^{s_1}}{s_1!(\epsilon)^{ps_1}}\left[b^{d+ps_1}\left\{\frac{\ln(b)}{d+ps_1} - \frac{1}{(d+ps_1)^2}\right\} - a^{d+ps_1}\left\{\frac{\ln(a)}{d+ps_1} - \frac{1}{(d+ps_1)^2}\right\}\right] +$$

$$\frac{\left\{\gamma\left(\left(\frac{b}{\epsilon}\right)^p,\frac{d+p}{p}\right)-\gamma\left(\left(\frac{a}{\epsilon}\right)^p,\frac{d+p}{p}\right)\right\}}{\left\{\gamma\left(\left(\frac{b}{\epsilon}\right)^p,\frac{d}{p}\right)-\gamma\left(\left(\frac{a}{\epsilon}\right)^p,\frac{d}{p}\right)\right\}} \quad (13)$$

II.4 The relative entropy

The relative entropy (or the Kullback–Leibler divergence) is a measure of the difference between two probability distributions G and G^*. In applications G typically represents the "true" distribution of data, observations, or a precisely calculated theoretical distribution, while G^* typically represents a theory, model, description, or approximation of G. Specifically, the Kullback–Leibler divergence of G^* from G, denoted $D_{KL}(G||G^*)$, is a measure of the information gained when one revises ones beliefs from the prior probability distribution G^* to the posterior probability distribution G. More exactly, it is the amount of information that is lost when G^* is used to approximate G, defined operationally as the expected extra number of bits required to code samples from G using a code optimized for G^* rather than the code optimized for G.

The relative entropy $D_{KL} = (G||G^*)$ for $DTGG$ random variable is,

$$D_{KL} = (G||G^*) = \int_a^b g(x)\ln\left(\frac{g(x)}{g^*(x)}\right)\,dx$$

$$= \int_a^b g(x)\ln\big(g(x)\big)\,dx - \int_a^b g(x)\ln\big(g^*(x)\big)\,dx$$

$$= -H - \int_a^b g(x)\ln\big(g^*(x)\big)\,dx$$

So, $$\int_a^b g(x)\ln\big(g^*(x)\big)\,dx = \int_a^b \frac{\frac{p_1}{\epsilon_1{}^{d_1}}}{\gamma\left(\left(\frac{b}{\epsilon_1}\right)^{p_1},\frac{d_1}{p_1}\right)-\gamma\left(\left(\frac{a}{\epsilon_1}\right)^{p_1},\frac{d_1}{p_1}\right)} x^{d_1-1}\, e^{-\left(\frac{x}{\epsilon_1}\right)^{p_1}}$$

$$\ln\left(\frac{\frac{p_2}{\epsilon_2{}^{d_2}}}{\gamma\left(\left(\frac{b}{\epsilon_2}\right)^{p_2},\frac{d_2}{p_2}\right)-\gamma\left(\left(\frac{a}{\epsilon_2}\right)^{p_2},\frac{d_2}{p_2}\right)} x^{(d_2-1)}\, e^{-\left(\frac{x}{\epsilon_2}\right)^{p_2}}\right)dx$$

$$= A_2 + (d_2-1)E(\ln(X)) - E\left(\left(\frac{X}{\epsilon_2}\right)^{p_2}\right)$$

Where, $A_2 = \ln\left(\frac{\frac{p_2}{\epsilon_2{}^{d_2}}}{\gamma\left(\left(\frac{b}{\epsilon_2}\right)^{p_2},\frac{d_2}{p_2}\right)-\gamma\left(\left(\frac{a}{\epsilon_2}\right)^{p_2},\frac{d_2}{p_2}\right)}\right)$. By using (11) and (12), we get,

$$\int_a^b g(x)\ln\big(g^*(x)\big)\,dx = A_2 + (d_2-1)\left[A_1\sum_{s_1=0}^{\infty}\frac{(-1)^{s_1}}{s_1!(\epsilon_1)^{p_1 s_1}}\left[\begin{array}{l} b^{d_1+p_1s_1}\left\{\frac{\ln(b)}{d_1+p_1s_1}-\frac{1}{(d_1+p_1s_1)^2}\right\} \\ -a^{d_1+p_1s_1}\left\{\frac{\ln(a)}{d_1+p_1s_1}-\frac{1}{(d_1+p_1s_1)^2}\right\}\end{array}\right]\right]$$

$$-\left[\epsilon_2{}^{-p_1}\epsilon_1{}^{p_2}\frac{\left\{\gamma\left(\left(\frac{b}{\epsilon_1}\right)^{p_1},\frac{d_1+p_2}{p_1}\right)-\gamma\left(\left(\frac{a}{\epsilon_1}\right)^{p_1},\frac{d_1+p_2}{p_1}\right)\right\}}{\left\{\gamma\left(\left(\frac{b}{\epsilon_1}\right)^{p_1},\frac{d_1}{p_1}\right)-\gamma\left(\left(\frac{a}{\epsilon_1}\right)^{p_1},\frac{d_1}{p_1}\right)\right\}}\right]$$

Then, the relative entropy is,

$$D_{KL} = (G||G^*) = \ln(A_1) + (d_1-1)\left[A_1\sum_{s_1=0}^{\infty}\frac{(-1)^{s_1}}{s_1!(\epsilon_1)^{p_1 s_1}}\left[\begin{array}{l} b^{d_1+p_1s_1}\left\{\frac{\ln(b)}{d_1+p_1s_1}-\frac{1}{(d_1+p_1s_1)^2}\right\} \\ -a^{d_1+p_1s_1}\left\{\frac{\ln(a)}{d_1+p_1s_1}-\frac{1}{(d_1+p_1s_1)^2}\right\}\end{array}\right]\right]$$

$$-\left[\frac{\left\{\gamma\left(\left(\frac{b}{\epsilon_1}\right)^{p_1},\frac{d_1+p_1}{p_1}\right)-\gamma\left(\left(\frac{a}{\epsilon_1}\right)^{p_1},\frac{d_1+p_1}{p_1}\right)\right\}}{\left\{\gamma\left(\left(\frac{b}{\epsilon_1}\right)^{p_1},\frac{d_1}{p_1}\right)-\gamma\left(\left(\frac{a}{\epsilon_1}\right)^{p_1},\frac{d_1}{p_1}\right)\right\}}\right]-A_2-$$

$$(d_2-1)\left[A_1\sum_{s_1=0}^{\infty}\frac{(-1)^{s_1}}{s_1!(\epsilon_1)^{p_1s_1}}\left[\begin{array}{l} b^{d_1+p_1s_1}\left\{\frac{\ln(b)}{d_1+p_1s_1}-\frac{1}{(d_1+p_1s_1)^2}\right\} \\ -a^{d_1+p_1s_1}\left\{\frac{\ln(a)}{d_1+p_1s_1}-\frac{1}{(d_1+p_1s_1)^2}\right\}\end{array}\right]\right]$$

$$+\left[{\epsilon_2}^{-p_1}{\epsilon_1}^{p_2}\frac{\left\{\gamma\left(\left(\frac{b}{\epsilon_1}\right)^{p_1},\frac{d_1+p_2}{p_1}\right)-\gamma\left(\left(\frac{a}{\epsilon_1}\right)^{p_1},\frac{d_1+p_2}{p_1}\right)\right\}}{\left\{\gamma\left(\left(\frac{b}{\epsilon_1}\right)^{p_1},\frac{d_1}{p_1}\right)-\gamma\left(\left(\frac{a}{\epsilon_1}\right)^{p_1},\frac{d_1}{p_1}\right)\right\}}\right] \quad (14)$$

III. The generalized inverse weibull distribution

A three parameter generalized inverse Weibull distribution (GIWD) with decreasing and unimodal failure rate is introduced and studied by de Gusmao et al in 2011. They provided a comprehensive treatment of the mathematical properties of GIWD. The mixture model of two generalized inverse Weibull distributions is investigated. They also proposed a location-scale regression model based on the log- GIWD for modeling lifetime data. In addition, some diagnostic tools for sensitivity analysis is developed. Khan and King in 2014, introduced five parameter transmuted GIWD. They derived moments, moment generating function, entropy, mean deviation, Bonferroni and Lorenz curves. Maximum likelihood for estimating the model parameters is used and based on the observed information matrix is obtained. Elbatal and Muhammed in 2014, presented the Exponentiated GIWD. They derived the moment generating function and the rth moment. Expressions for the density, moment generating function and rth moment of the order statistics also are obtained. They discussed the parameters estimation by maximum likelihood and provide the information matrix. GIW- GIW distribution is proposed by Abid et al in 2019 as new distribution. The probability density function, cumulative distribution function, reliability and hazard rate functions are introduced. Furthermore, they derived most important statistical properties of GIW- GIW distribution such as Shannon entropy, relative entropy, stress-strength model. Salem in 2019 studied the Marshall–Olkin GIWD. The new distribution is flexible and contains sub-models such as inverse exponential, inverse Rayleigh, Weibull, inverse Weibull, Marshall–Olkin inverse Weibull and Fréchet distributions. Some properties are obtained. Maximum likelihood, least square estimators, interval estimators, estimators, fisher information matrix and asymptotic confidence intervals are described.

The probability density and cumulative functions of GIWD random variable with three parameters $\alpha>0,\ \beta>0$ and $\sigma>0$ are respectively given by,

$$f(x)=\sigma\beta\alpha^{\beta}\,x^{-(\beta+1)}\,e^{-\sigma\left(\frac{\alpha}{x}\right)^{\beta}},\qquad x>0 \quad (15)$$

$$F(x)=e^{-\sigma\left(\frac{\alpha}{x}\right)^{\beta}},\qquad x>0 \quad (16)$$

III.1 Essential properties of *DTGIW*$(\boldsymbol{\alpha},\boldsymbol{\sigma},\boldsymbol{\beta},\boldsymbol{a},\boldsymbol{b})$

The pdf and cdf of $DTGIW(\alpha,\sigma,\beta,a,b)$ random variable are respectively,

$$g(x)=\frac{\sigma\beta\alpha^{\beta}\,x^{-(\beta+1)}\,e^{-\sigma\left(\frac{\alpha}{x}\right)^{\beta}}}{e^{-\sigma\left(\frac{\alpha}{b}\right)^{\beta}}-e^{-\sigma\left(\frac{\alpha}{a}\right)^{\beta}}},\quad a<x<b \quad (17)$$

$$G(x) = \frac{e^{-\sigma\left(\frac{\alpha}{x}\right)^{\beta}} - e^{-\sigma\left(\frac{\alpha}{a}\right)^{\beta}}}{e^{-\sigma\left(\frac{\alpha}{b}\right)^{\beta}} - e^{-\sigma\left(\frac{\alpha}{a}\right)^{\beta}}} \qquad , a < x < b \tag{18}$$

Also the reliability and hazard functions of $DTGIW(\alpha, \sigma, \beta, a, b)$ distribution are respectively,

$$R(x) = \frac{e^{-\sigma\left(\frac{\alpha}{b}\right)^{\beta}} - e^{-\sigma\left(\frac{\alpha}{x}\right)^{\beta}}}{e^{-\sigma\left(\frac{\alpha}{b}\right)^{\beta}} - e^{-\sigma\left(\frac{\alpha}{a}\right)^{\beta}}} \tag{19}$$

$$H(x) = \frac{\sigma\beta\,\alpha^{\beta}\, x^{-(\beta+1)}\, e^{-\sigma\left(\frac{\alpha}{x}\right)^{\beta}}}{e^{-\sigma\left(\frac{\alpha}{b}\right)^{\beta}} - e^{-\sigma\left(\frac{\alpha}{x}\right)^{\beta}}} \tag{20}$$

So, the rth raw moment is,

$$E(X^r) = \int_a^b x^r \, \frac{\sigma\beta\alpha^{\beta}\, x^{-(\beta+1)}\, e^{-\sigma\left(\frac{\alpha}{x}\right)^{\beta}}}{e^{-\sigma\left(\frac{\alpha}{b}\right)^{\beta}} - e^{-\sigma\left(\frac{\alpha}{a}\right)^{\beta}}}\, dx$$

$$= \frac{\sigma\beta\,\alpha^{\beta}}{e^{-\sigma\left(\frac{\alpha}{b}\right)^{\beta}} - e^{-\sigma\left(\frac{\alpha}{a}\right)^{\beta}}} \int_a^b x^{r-\beta-1}\, e^{-\sigma\left(\frac{\alpha}{x}\right)^{\beta}}\, dx$$

$$= \frac{\sigma\beta\,\alpha^{\beta}}{e^{-\sigma\left(\frac{\alpha}{b}\right)^{\beta}} - e^{-\sigma\left(\frac{\alpha}{a}\right)^{\beta}}} \int_a^b x^{r-\beta-1} \sum_{s_1=0}^{\infty} \frac{\left(-\sigma\left(\frac{\alpha}{x}\right)^{\beta}\right)^{s_1}}{s_1!}\, dx$$

$$= \frac{\sigma\beta\,\alpha^{\beta}}{e^{-\sigma\left(\frac{\alpha}{b}\right)^{\beta}} - e^{-\sigma\left(\frac{\alpha}{a}\right)^{\beta}}} \sum_{s_1=0}^{\infty} \frac{\left(-\sigma(\alpha)^{\beta}\right)^{s_1}}{s_1!} \int_a^b x^{r-\beta-\beta s_1-1}\, dx$$

$$= \frac{\sigma\beta\,\alpha^{\beta}}{e^{-\sigma\left(\frac{\alpha}{b}\right)^{\beta}} - e^{-\sigma\left(\frac{\alpha}{a}\right)^{\beta}}} \sum_{s_1=0}^{\infty} \frac{\left(-\sigma(\alpha)^{\beta}\right)^{s_1}}{s_1!} \left[\frac{b^{r-\beta(s_1+1)} - a^{r-\beta(s_1+1)}}{r-\beta(s_1+1)}\right]$$

Let $= \frac{\sigma\beta\,\alpha^{\beta}}{e^{-\sigma\left(\frac{\alpha}{b}\right)^{\beta}} - e^{-\sigma\left(\frac{\alpha}{a}\right)^{\beta}}}$, then,

$$E(x^r) = v \sum_{s_1=0}^{\infty} \frac{\left(-\sigma(\alpha)^{\beta}\right)^{s_1}}{s_1!} \left[\frac{b^{r-\beta-\beta s_1} - a^{r-\beta-\beta s_1}}{r-\beta-\beta s_1}\right] \tag{21}$$

Then, the characteristic function can easily get by using the relation,

$$Q_x(t) = E\left(e^{ixt}\right) = \sum_{r=0}^{\infty} \frac{(it)^r}{r!} E(x^r) \quad , \quad \text{since} \quad e^{ixt} = \sum_{r=0}^{\infty} \frac{(it)^r}{r!}\, x^r$$

III.2 Stress-Strength Reliability

Suppose $X \sim DTGIW(\alpha_1, \sigma_1, \beta_1, a, b)$ and $Y \sim DTGIW(\alpha_2, \sigma_2, \beta_2, a, b)$ with unknown parameters $\alpha_1, \sigma_1, \beta_1, \alpha_2, \sigma_2, \beta_2, a, b$, where X and Y are independently distributed, then the stress-stress reliability function is,

$$R=\int_a^b \frac{\sigma_1\beta_1\alpha_1{}^{\beta_1}x^{-(\beta_1+1)}e^{-\sigma_1\left(\frac{\alpha_1}{x}\right)^{\beta_1}}e^{-\sigma_2\left(\frac{\alpha_2}{x}\right)^{\beta_2}}}{\left(e^{-\sigma_1\left(\frac{\alpha_1}{b}\right)^{\beta_1}}-e^{-\sigma_1\left(\frac{\alpha_1}{a}\right)^{\beta_1}}\right)\left(e^{-\sigma_2\left(\frac{\alpha_2}{b}\right)^{\beta_2}}-e^{-\sigma_2\left(\frac{\alpha_2}{a}\right)^{\beta_2}}\right)}\,dx$$

$$-\int_a^b \frac{\sigma_1\beta_1\alpha_1{}^{\beta_1}x^{-(\beta_1+1)}e^{-\sigma_1\left(\frac{\alpha_1}{x}\right)^{\beta_1}}e^{-\sigma_2\left(\frac{\alpha_2}{a}\right)^{\beta_2}}}{\left(e^{-\sigma_1\left(\frac{\alpha_1}{b}\right)^{\beta_1}}-e^{-\sigma_1\left(\frac{\alpha_1}{a}\right)^{\beta_1}}\right)\left(e^{-\sigma_2\left(\frac{\alpha_2}{b}\right)^{\beta_2}}-e^{-\sigma_2\left(\frac{\alpha_2}{a}\right)^{\beta_2}}\right)}\,dx$$

$$=\int_a^b \frac{\sigma_1\beta_1\alpha^{\beta_1}x^{-(\beta_1+1)}e^{-\sigma_1\left(\frac{\alpha}{x}\right)^{\beta_1}}e^{-\sigma_2\left(\frac{\alpha_2}{x}\right)^{\beta_2}}}{\left(e^{-\sigma_1\left(\frac{\alpha_1}{b}\right)^{\beta_1}}-e^{-\sigma_1\left(\frac{\alpha_1}{a}\right)^{\beta_1}}\right)\left(e^{-\sigma_2\left(\frac{\alpha_2}{b}\right)^{\beta_2}}-e^{-\sigma_2\left(\frac{\alpha_2}{a}\right)^{\beta_2}}\right)}\,dx$$

$$-\frac{e^{-\sigma_2\left(\frac{\alpha_2}{a}\right)^{\beta_2}}}{\left(e^{-\sigma_2\left(\frac{\alpha_2}{b}\right)^{\beta_2}}-e^{-\sigma_2\left(\frac{\alpha_2}{a}\right)^{\beta_2}}\right)}\int_a^b \frac{\sigma_1\beta_1\alpha_1{}^{\beta_1}x^{-(\beta_1+1)}e^{-\sigma_1\left(\frac{\alpha_1}{x}\right)^{\beta_1}}}{\left(e^{-\sigma_1\left(\frac{\alpha_1}{b}\right)^{\beta_1}}-e^{-\sigma_1\left(\frac{\alpha_1}{a}\right)^{\beta_1}}\right)}dx$$

Let $K_1=\frac{\sigma_1\beta_1\alpha^{\beta_1}}{\left(e^{-\sigma_1\left(\frac{\alpha_1}{b}\right)^{\beta_1}}-e^{-\sigma_1\left(\frac{\alpha_1}{a}\right)^{\beta_1}}\right)\left(e^{-\sigma_2\left(\frac{\alpha_2}{b}\right)^{\beta_2}}-e^{-\sigma_2\left(\frac{\alpha_2}{a}\right)^{\beta_2}}\right)}$ and $K_2=\frac{e^{-\sigma_2\left(\frac{\alpha_2}{a}\right)^{\beta_2}}}{\left(e^{-\sigma_2\left(\frac{\alpha_2}{b}\right)^{\beta_2}}-e^{-\sigma_2\left(\frac{\alpha_2}{a}\right)^{\beta_2}}\right)}$, then

$$R=K_1\int_a^b x^{-(\beta_1+1)}\sum_{s_1=0}^{\infty}\frac{\left(-\sigma_1\left(\frac{\alpha_1}{x}\right)^{\beta_1}\right)^{s_1}}{s_1!}\sum_{s_2=0}^{\infty}\frac{\left(-\sigma_2\left(\frac{\alpha_2}{x}\right)^{\beta_2}\right)^{s_2}}{s_2!}\,dx-K_2$$

$$=K_1\sum_{s_1=0}^{\infty}\frac{(-\sigma_1(\alpha_1)^{\beta_1})^{s_1}}{s_1!}\sum_{s_2=0}^{\infty}\frac{(-\sigma_2(\alpha_2)^{\beta_2})^{s_2}}{s_2!}\int_a^b x^{-\beta_1-\beta_1 s_1-\beta_2 s_2-1}\,dx-K_2$$

Then the stress-strength reliability function will be,

$$R=K_1\sum_{s_1=0}^{\infty}\frac{(-\sigma_1(\alpha_1)^{\beta_1})^{s_1}}{s_1!}\sum_{s_2=0}^{\infty}\frac{(-\sigma_2(\alpha_2)^{\beta_2})^{s_2}}{s_2!}\left[\frac{b^{-(\beta_1(1+s_1)+\beta_2 s_2)}-a^{-(\beta_1(1+s_1)+\beta_2 s_2)}}{-(\beta_1(1+s_1)+\beta_2 s_2)}\right]-K_2 \qquad (22)$$

III.3 Shannon entropy

The Shannon entropy of $DTGIW(\alpha,\sigma,\beta,a,b)$ can be found as,

$$H=E\left(-\ln\left(\frac{\sigma\beta\alpha^{\beta}}{e^{-\sigma\left(\frac{\alpha}{b}\right)^{\beta}}-e^{-\sigma\left(\frac{\alpha}{a}\right)^{\beta}}}\right)+(\beta+1)\ln(X)+\sigma\left(\frac{\alpha}{X}\right)^{\beta}\right)$$

$$H=-\ln\left(\frac{\sigma\beta\alpha^{\beta}}{e^{-\sigma\left(\frac{\alpha}{b}\right)^{\beta}}-e^{-\sigma\left(\frac{\alpha}{a}\right)^{\beta}}}\right)+(\beta+1)E(\ln(X))+\sigma E\left(\left(\frac{\alpha}{X}\right)^{\beta}\right)$$

Let $A_1=\frac{\sigma\beta\alpha^{\beta}}{\left[e^{-\sigma\left(\frac{\alpha}{b}\right)^{\beta}}-e^{-\sigma\left(\frac{\alpha}{a}\right)^{\beta}}\right]}$, then

$$E(\ln(X))=\int_a^b \ln(x)\frac{\sigma\beta\alpha^{\beta}x^{-(\beta+1)}e^{-\sigma\left(\frac{\alpha}{x}\right)^{\beta}}}{e^{-\sigma\left(\frac{\alpha}{b}\right)^{\beta}}-e^{-\sigma\left(\frac{\alpha}{a}\right)^{\beta}}}\,dx$$

$$= A_1 \int_a^b \ln(x)\, x^{-(\beta+1)} \sum_{s_1=0}^{\infty} \frac{\left(-\sigma\left(\frac{\alpha}{x}\right)^{\beta}\right)^{s_1}}{s_1!} dx$$

$$= A_1 \sum_{s_1=0}^{\infty} \frac{(-\sigma(\alpha)^{\beta})^{s_1}}{s_1!} \int_a^b \ln(x)\, x^{-(\beta(1+s_1)+1)}\, dx$$

Since $\int_a^b x^{-m} \ln(x)\, dx = -\frac{\ln(x)}{(m-1)x^{m-1}} - \frac{1}{(m-1)^2 x^{m-1}} = \left\{-\frac{\ln(x)}{m-1} - \frac{1}{(m-1)^2}\right\} \frac{1}{x^{m-1}}$, then,

$$E(\ln(X)) = A_1 \sum_{s_1=0}^{\infty} \frac{(-\sigma(\alpha)^{\beta})^{s_1}}{s_1!\beta} \left\{\frac{-1}{b^{(\beta(1+s_1))}}\left\{\frac{\ln(b)}{(1+s_1)} + \frac{1}{\beta(1+s_1)^2}\right\} + \frac{1}{a^{(\beta(1+s_1))}}\left\{\frac{\ln(a)}{(1+s_1)} + \frac{1}{\beta(1+s_1)^2}\right\}\right\} \quad (23)$$

With $m = (\beta(1+s_1)+1)$.

Let $A_2 = \sigma\beta\alpha^{2\beta} / \left[exp - \sigma\left(\frac{\alpha}{b}\right)^{\beta} - exp\left(-\sigma\left(\frac{\alpha}{a}\right)^{\beta}\right)\right]$, then,

$$E\left(\alpha^{\beta} X^{-\beta}\right) = A_2 \int_a^b (x)^{-2\beta-1} e^{-\sigma\left(\frac{\alpha}{x}\right)^{\beta}} dx = A_2 \int_a^b (x)^{-2\beta-1} \sum_{s_1=0}^{\infty} \frac{\left(-\sigma\left(\frac{\alpha}{x}\right)^{\beta}\right)^{s_1}}{s_1!} dx$$

$$= A_2 \sum_{s_1=0}^{\infty} \frac{(-\sigma(\alpha)^{\beta})^{s_1}}{s_1!} \int_a^b (x)^{-(\beta(2+s_1)+1)}\, dx$$

$$= A_2 \sum_{s_1=0}^{\infty} \frac{(-\sigma(\alpha)^{\beta})^{s_1}}{s_1!} \left[\frac{(b)^{-(\beta(2+s_1))} - (a)^{-(\beta(2+s_1))}}{-(\beta(2+s_1))}\right] \quad (24)$$

Finally, the Shannon entropy will be,

$$H = (\beta+1)\left[A_1 \sum_{s_1=0}^{\infty} \frac{(-\sigma(\alpha)^{\beta})^{s_1}}{s_1!\beta} \left\{\frac{-1}{b^{(\beta(1+s_1))}}\left\{\frac{\ln(b)}{(1+s_1)} + \frac{1}{\beta(1+s_1)^2}\right\} + \frac{1}{a^{(\beta(1+s_1))}}\left\{\frac{\ln(a)}{(1+s_1)} + \frac{1}{\beta(1+s_1)^2}\right\}\right\}\right]$$

$$+\sigma A_2 \sum_{s_1=0}^{\infty} \frac{(-\sigma(\alpha)^{\beta})^{s_1}}{s_1!} \left[\frac{(b)^{-(\beta(2+s_1))} - (a)^{-(\beta(2+s_1))}}{-(\beta(2+s_1))}\right] - \ln(A_1) \quad (25)$$

III.4 The relative entropy

The relative entropy $D_{KL} = (G||G^*)$ for *DTGIW* random variable is,

$D_{KL} = (G||G^*) = -H - \int_a^b g(x) \ln(g^*(x))\, dx$, then,

$$\int_a^b g(x) \ln(g^*(x))\, dx = \int_a^b \left(\left[\frac{\sigma_1\beta_1\alpha_1{}^{\beta_1} x^{-(\beta_1+1)} e^{-\sigma_1\left(\frac{\alpha_1}{x}\right)^{\beta_1}}}{\left(e^{-\sigma_1\left(\frac{\alpha_1}{b}\right)^{\beta_1}} - e^{-\sigma_1\left(\frac{\alpha_1}{a}\right)^{\beta_1}}\right)}\right] \ln\left(\frac{\sigma_2\beta_2\alpha_2{}^{\beta_2} x^{-(\beta_2+1)} e^{-\sigma_2\left(\frac{\alpha_2}{x}\right)^{\beta_2}}}{\left(e^{-\sigma_2\left(\frac{\alpha_2}{b}\right)^{\beta_2}} - e^{-\sigma_2\left(\frac{\alpha_2}{a}\right)^{\beta_2}}\right)}\right)\right) dx$$

$$= \ln\left(\frac{\sigma_2\beta_2\alpha_2{}^{\beta_2}}{\left(e^{-\sigma_2\left(\frac{\alpha_2}{b}\right)^{\beta_2}} - e^{-\sigma_2\left(\frac{\alpha_2}{a}\right)^{\beta_2}}\right)}\right) - \left[\frac{\sigma_1\beta_1\alpha_1{}^{\beta_1}(\beta_2+1)}{\left(e^{-\sigma_1\left(\frac{\alpha_1}{b}\right)^{\beta_1}} - e^{-\sigma_1\left(\frac{\alpha_1}{a}\right)^{\beta_1}}\right)}\right] \int_a^b x^{-(\beta_1+1)} e^{-\sigma_1\left(\frac{\alpha_1}{x}\right)^{\beta_1}} (\ln(x))\, dx -$$

$$\left[\frac{\sigma_1\beta_1\alpha_1{}^{\beta_1}\sigma_2\alpha_2{}^{\beta_2}}{\left(e^{-\sigma_1\left(\frac{\alpha_1}{b}\right)^{\beta_1}} - e^{-\sigma_1\left(\frac{\alpha_1}{a}\right)^{\beta_1}}\right)}\right] \int_a^b x^{-(\beta_1+\beta_2+1)} e^{-\sigma_1\left(\frac{\alpha_1}{x}\right)^{\beta_1}} dx$$

By using the same steps for equations (23) and (24) , we get,

$$D_{KL} = -(\beta_1 + 1)\left[A_{11}\sum_{s_1=0}^{\infty}\frac{\left(-\sigma_1(\alpha_1)^{\beta_1}\right)^{s_1}}{s_1!\beta_1}\left\{\begin{matrix}\frac{-1}{b^{(\beta_1(1+s_1))}}\left\{\frac{\ln(b)}{(1+s_1)}+\frac{1}{\beta_1(1+s_1)^2}\right\}\\ +\frac{1}{a^{(\beta_1(1+s_1))}}\left\{\frac{\ln(a)}{(1+s_1)}+\frac{1}{\beta_1(1+s_1)^2}\right\}\end{matrix}\right\}\right]$$

$$-\sigma_1 A_{21}\sum_{s_1=0}^{\infty}\frac{\left(-\sigma_1(\alpha_1)^{\beta_1}\right)^{s_1}}{s_1!}\left[\frac{(b)^{-(\beta_1(2+s_1))}-(a)^{-(\beta_1(2+s_1))}}{-(\beta_1(2+s_1))}\right]+\ln(A_{11})$$

$$-\ln\left(\frac{\sigma_2\beta_2{\alpha_2}^{\beta_2}}{\left(e^{-\sigma_2\left(\frac{\alpha_2}{b}\right)^{\beta_2}}-e^{-\sigma_2\left(\frac{\alpha_2}{a}\right)^{\beta_2}}\right)}\right)+A_3\sum_{s_1=0}^{\infty}\frac{\left(-\sigma_1(\alpha_1)^{\beta}\right)^{s_1}}{s_1!\beta_1}\left\{\begin{matrix}\frac{-1}{b^{(\beta_1(1+s_1))}}\left\{\frac{\ln(b)}{(1+s_1)}+\frac{1}{\beta_1(1+s_1)^2}\right\}\\ +\frac{1}{a^{(\beta_1(1+s_1))}}\left\{\frac{\ln(a)}{(1+s_1)}+\frac{1}{\beta_1(1+s_1)^2}\right\}\end{matrix}\right\}$$

$$+A_4\sum_{s_1=0}^{\infty}\frac{\left(-\sigma_1(\alpha_1)^{\beta_1}\right)^{s_1}}{s_1!}\left[\frac{(b)^{-(\beta_1(1+s_1)+\beta_2)}-(a)^{-(\beta_1(1+s_1)+\beta_2)}}{-(\beta_1(1+s_1)+\beta_2)}\right] \qquad (26)$$

Where $A_{11} = \frac{\sigma_1\beta_1{\alpha_1}^{\beta_1}}{e^{-\sigma_1\left(\frac{\alpha_1}{b}\right)^{\beta_1}}-e^{-\sigma_1\left(\frac{\alpha_1}{a}\right)^{\beta_1}}}$, $A_{21} = \frac{\sigma_1\beta_1{\alpha_1}^{2\beta_1}}{e^{-\sigma_1\left(\frac{\alpha_1}{b}\right)^{\beta_1}}-e^{-\sigma_1\left(\frac{\alpha_1}{a}\right)^{\beta_1}}}$, $A_3 = \frac{\sigma_1\beta_1{\alpha_1}^{\beta_1}(\beta_2+1)}{e^{-\sigma_1\left(\frac{\alpha_1}{b}\right)^{\beta_1}}-e^{-\sigma_1\left(\frac{\alpha_1}{a}\right)^{\beta_1}}}$

and $A_4 = \frac{\sigma_1\beta_1{\alpha_1}^{\beta_1}\sigma_2{\alpha_2}^{\beta_2}}{e^{-\sigma_1\left(\frac{\alpha_1}{b}\right)^{\beta_1}}-e^{-\sigma_1\left(\frac{\alpha_1}{a}\right)^{\beta_1}}}$.

IV- Summary and Conclusion

Distributions are used to represent set(s) of data in statistical analysis. The composing of some distributions with each other's in some way to generate new distributions more flexible than the others to model real data . In this paper, we derived Properties of DTGG and DTGIW distributions, since doubly truncated distributions are more realistic to represent phenomena. We provided forms of rth raw moment, reliability function, hazard rate function, Shannon entropy function and Relative entropy function. This paper deals also with the determination of stress-strength R=p[y<x] when x (strength) and y (stress) are two independent DTGG (DTGIW) distribution with different parameters.

Reference

[1] Abid, S. and Abdulrazak, R. (2017) “ [0,1] truncated Frechet-G generator of distributions”, Applied Mathematics, 7(3), 51-66.

[2] Abid,S., Al-Noor,N. and Boshi, M. (2019) “On the generalized inverse Weibull Distribution”, AIP Conference Proceedings 2086, 030002 (2019)

[3] Barriga,G., Cordeiro, G., Dey,D. Cancho, V., Louzada,F. and Suzuki, A. (2018) “The Marshall-Olkin generalized gamma distribution”, Communications for Statistical Applications and Methods, Vol. 25, No. 3, 245–261.

[4] Cordeiro, G., Ortega, E. and Silva, G. (2011) “ The exponentiated generalized gamma distribution with application to lifetime data”, J. of stat. comp. and simulation, vol. 81, no. 7, 827-842.

[5] Cox, C. and Matheson , M (2014) “A Comparison of the Generalized Gamma and Exponentiated Weibull Distributions”, Stat Med., 20; 33(21): 3772–3780.

[6] de Gusmão, F., Ortega, E. and Cordeiro, G. (2011) “The generalized inverseWeibull distribution”, Stat Papers, DOI 10.1007/s00362-009-0271-3

[7] Elbatal, I. and Muhammed, H. (2014) “Exponentiated Generalized Inverse Weibull Distribution”, Applied Mathematical Sciences, Vol. 8, 2014, no. 81, 3997 – 4012

[8] Khan, M. and King, R. (2014) “Transmuted generalized Inverse Weibull distribution”, Journal of Applied Statistical Science, Vol. 20, No. 3, pp. 213-230
[9] Khodabin, M. and Ahmadabadi, A. (2010) “Some properties of generalized gamma distribution”, Mathematical Sciences Vol. 4, No. 1 , 9-28.
[10] Kiche, J., Ngesa, O. and Orwa, G. (2019) “On Generalized Gamma Distribution and Its Application to Survival Data”, International Journal of Statistics and Probability; Vol. 8, No. 5, 85-102.
[11] Salem, H. (2019) “The Marshall–Olkin Generalized Inverse Weibull Distribution:
[12] Properties and Application”, Modern Applied Science; Vol. 13, No. 2, 54-65.
[13] Stacy, E. (1962) “A generalization of gamma distribution”, Annals of Mathematical Statistics, 33, 1187-92.

On the Doubly-Truncated Generalized Gompertz and Marshal-Olkin extended Uniform Distributions

Salah Hamza Abid **Hind Husham Jani**

abidsalah@gmail.com **hind.mathematic13@gmail.com**

Mathematics department / University of Mustansiriyah / Bagdad- Iraq

Abstract :It is well known that the truncated distributions are more matches with reality. Therefore, in this paper, some properties of doubly truncated generalized Gompertz distribution and doubly truncated Marshal-Olkin extended Uniform distribution are derived. These properties are the reliability and hazard functions, rth raw moments, stress-strength reliability, Shannon entropy and relative entropy.

Keywords: Doubly truncated Generalized Gompertz distribution, Doubly truncated Marshal-Olkin extended Uniform distribution, hazard function, characteristic function, Shannon entropy, relative entropy, Stress-strength reliability.

IV. **Introduction**

It is well known that the Statistical distributions are very salutary in describing and predicting real life phenomena. There are many probability distributions of which some can be fitted more strictly to the frequency of the data than others, depending on the characteristics of the phenomenon and of the distribution.

The doubly truncated distributions are more realistic to represent phenomena, without losing the generality, since the truncated parameters can take any values.

In this paper, properties of doubly truncated generalized Gompertz distribution (DTGGO) and Doubly truncated Marshall-Olkin extended Uniform distribution (DTMOEU) are derived. These properties are the rth raw moments, The characteristic function, stress-strength reliability, Shannon entropy and relative entropy.

V. **The generalized Gompertz distribution**

El-Gohary et al in 2013 introduced the generalized Gompertz distribution (GGO). The main advantage of this new distribution is that it has increasing or constant or decreasing or bathtub curve failure rate depending upon the shape parameter. This property makes GGD is very useful in survival analysis. Some statistical properties are derived and some issues related with parameters estimation are discussed. Khan et al in 2017 introduced the transmuted generalized Gompertz distribution. They studied its statistical properties. Explicit expressions are derived for the quantile, moments, moment generating function and entropies. Maximum likelihood estimation is used to estimate the model parameters. De Andrade et al in 2019 presented the exponentiated generalized extended Gompertz distribution. The hazard function of this distribution includes inverted bathtub and bathtub shapes. For the new model, several properties are derived and the maximum likelihood estimation is discussed. Mazucheli et al in 2019 derived a new distribution from the Gompertz distribution by a transformation of the type X = exp(-y) , where Y has the Gompertz distribution. The proposed distribution encompasses the behavior of and provides better fits than some well-known lifetime distributions. Karamikabir et al in 2019 proposed an extended generalized Gompertz (EGGo) family of densities. Certain statistical properties of EGGo family including distribution shapes, hazard function, skewness, limit behavior, moments and order statistics are discussed. The flexibility of this family is assessed by its application to real data sets and comparison with other competing distributions. The performances of the some estimators are discussed. Boshi et al in 2019 introduced generalized Gompertz - generalized Gompertz distribution. The probability density, cumulative distribution, reliability and hazard rate functions are derived. The most essential statistical properties of this new distribution such as the rth raw moments function, characteristic function, quantiel, median, Shannon and relative entropies along with stress strength model are obtained.

The probability density function of the GGO variable is,

$$f(x) = \theta\lambda e^{\mu x}\, e^{-\frac{\lambda}{\mu}(e^{\mu x}-1)}\left(1 - e^{-\frac{\lambda}{\mu}(e^{\mu x}-1)}\right)^{\theta-1} \quad (1)$$

And the cumulative distribution function is,

$$F(x) = \left(1 - e^{-\frac{\lambda}{\mu}(e^{\mu x}-1)}\right)^{\theta} \quad (2)$$

II.1 Essential properties of $DTGGO\ (\theta, \lambda, \mu, a, b)$

We consider here the doubly truncated Generalized Gompertz Distribution $DTGGO\ (\theta, \lambda, \mu, a, b)$ of random variable X, where the lower and upper limits are a and b respectively, then the probability density function (pdf) and the commutative density function (cdf) are respectively,

$$g(x) = \frac{f(x)}{F(b)-F(a)}$$

$$= \frac{\theta\lambda e^{\mu x}\, e^{-\frac{\lambda}{\mu}(e^{\mu x}-1)}\left(1-e^{-\frac{\lambda}{\mu}(e^{\mu x}-1)}\right)^{\theta-1}}{\left(1-e^{-\frac{\lambda}{\mu}(e^{\mu b}-1)}\right)^{\theta}-\left(1-e^{-\frac{\lambda}{\mu}(e^{\mu a}-1)}\right)^{\theta}} \quad (3)$$

$$G(x) = \frac{F(x)-F(a)}{F(b)-F(a)}$$

$$= \frac{\left(1-e^{-\frac{\lambda}{\mu}(e^{\mu x}-1)}\right)^{\theta}-\left(1-e^{-\frac{\lambda}{\mu}(e^{\mu a}-1)}\right)^{\theta}}{\left(1-e^{-\frac{\lambda}{\mu}(e^{\mu b}-1)}\right)^{\theta}-\left(1-e^{-\frac{\lambda}{\mu}(e^{\mu a}-1)}\right)^{\theta}} \quad (4)$$

Then, the reliability and the hazard functions of X are respectively,

$$R(x) = 1 - G(x)$$

$$R(x) = \frac{\left(1-e^{-\frac{\lambda}{\mu}(e^{\mu b}-1)}\right)^{\theta}-\left(1-e^{-\frac{\lambda}{\mu}(e^{\mu x}-1)}\right)^{\theta}}{\left(1-e^{-\frac{\lambda}{\mu}(e^{\mu b}-1)}\right)^{\theta}-\left(1-e^{-\frac{\lambda}{\mu}(e^{\mu a}-1)}\right)^{\theta}} \quad (5)$$

$$H(x) = \frac{g(x)}{R(x)}$$

$$= \frac{\theta\lambda e^{\mu x}\, e^{-\frac{\lambda}{\mu}(e^{\mu x}-1)}\left(1-e^{-\frac{\lambda}{\mu}(e^{\mu x}-1)}\right)^{\theta-1}}{\left(1-e^{-\frac{\lambda}{\mu}(e^{\mu b}-1)}\right)^{\theta}-\left(1-e^{-\frac{\lambda}{\mu}(e^{\mu x}-1)}\right)^{\theta}} \quad (6)$$

The rth moment of $DTGGO\ (\theta, \lambda, \mu, a, b)$ distribution is,

$$E(X^r) = \int_a^b x^r\ g(x)\ dx = \int_a^b x^r \frac{\theta\lambda e^{\mu x}\ e^{-\frac{\lambda}{\mu}(e^{\mu x}-1)}\left(1-e^{-\frac{\lambda}{\mu}(e^{\mu x}-1)}\right)^{\theta-1}}{\left(1-e^{-\frac{\lambda}{\mu}(e^{\mu b}-1)}\right)^{\theta}-\left(1-e^{-\frac{\lambda}{\mu}(e^{\mu a}-1)}\right)^{\theta}}\ dx$$

Let $A = \frac{\theta\lambda}{\left(1-e^{-\frac{\lambda}{\mu}(e^{\mu b}-1)}\right)^{\theta}-\left(1-e^{-\frac{\lambda}{\mu}(e^{\mu a}-1)}\right)^{\theta}}$, then,

$$E(X^r) = A\int_a^b x^r e^{\mu x}\ e^{-\frac{\lambda}{\mu}(e^{\mu x}-1)}\ \left(1-e^{-\frac{\lambda}{\mu}(e^{\mu x}-1)}\right)^{\theta-1}\ dx$$

Since $(x+y)^n = \sum_{k=0}^{n}\binom{n}{k}(x)^{n-k}(y)^k$, then,

$$\left(1-e^{-\frac{\lambda}{\mu}(e^{\mu x}-1)}\right)^{\theta-1} = \sum_{j_1=0}^{\theta-1}\binom{\theta-1}{j_1}(1)^{\theta-1-j_1}(-1)^{j_1}e^{-\frac{\lambda}{\mu}(e^{\mu x}-1)j_1}$$, and then,

$$E(X^r) = A\sum_{j_1=0}^{\theta-1}\binom{\theta-1}{j_1}(-1)^{j_1}\ \int_a^b x^r e^{\mu x}\ e^{-\frac{\lambda}{\mu}(e^{\mu x}-1)(j_1+1)}\ dx$$

$$= A\sum_{j_1=0}^{\theta-1}\binom{\theta-1}{j_1}(-1)^{j_1}\ \int_a^b x^r e^{\mu x}\ \sum_{s_1=0}^{\infty}\frac{\left(-\frac{\lambda}{\mu}(e^{\mu x}-1)(j_1+1)\right)^{s_1}}{s_1!}\ dx$$

$$= A\sum_{j_1=0}^{\theta-1}\binom{\theta-1}{j_1}(-1)^{j_1}\ \sum_{s_1=0}^{\infty}\frac{\left(-\frac{\lambda}{\mu}(j_1+1)\right)^{s_1}}{s_1!}\ \int_a^b x^r e^{\mu x}\left((e^{\mu x}-1)\right)^{s_1} dx$$

Since, $\left((e^{\mu x}-1)\right)^{s_1} = \sum_{j_2=0}^{s_1}\binom{s_1}{j_2}(-1)^{j_2}e^{-\mu x(j_2-s_1)}$, then,

$$E(X^r) = A\sum_{j_1=0}^{\theta-1}\binom{\theta-1}{j_1}(-1)^{j_1}\ \sum_{s_1=0}^{\infty}\frac{\left(-\frac{\lambda}{\mu}(j_1+1)\right)^{s_1}}{s_1!}\sum_{j_2=0}^{s_1}\binom{s_1}{j_2}(-1)^{j_2}\ \int_a^b x^r e^{-\mu x(j_2-s_1-1)}\ dx$$

$$= A\sum_{j_1=0}^{\theta-1}\binom{\theta-1}{j_1}(-1)^{j_1}\ \sum_{s_1=0}^{\infty}\frac{\left(-\frac{\lambda}{\mu}(j_1+1)\right)^{s_1}}{s_1!}\sum_{j_2=0}^{s_1}\binom{s_1}{j_2}(-1)^{j_2}$$

$$\left[\int_0^b x^r e^{-\mu x(j_2-s_1-1)}\ dx - \int_0^a x^r e^{-\mu x(j_2-s_1-1)}\ dx\right]$$

$$= A\sum_{j_1=0}^{\theta-1}\binom{\theta-1}{j_1}(-1)^{j_1}\ \sum_{s_1=0}^{\infty}\frac{\left(-\frac{\lambda}{\mu}(j_1+1)\right)^{s_1}}{s_1!}\sum_{j_2=0}^{s_1}\binom{s_1}{j_2}(-1)^{j_2}$$

$$\left[(\mu(j_2-s_1-1))^{-(r+1)}\begin{Bmatrix}\gamma((r+1),\mu b(j_2-s_1-1))\\ -\gamma((r+1),\mu a(j_2-s_1-1))\end{Bmatrix}\right] \quad (7)$$

Then, the characteristic function can easily get by using the relation,

$$Q_x(t) = E\left(e^{ixt}\right) = \sum_{r=0}^{\infty}\frac{(it)^r}{r!}E(x^r)\ \ ,\ \ \text{since}\ \ e^{ixt} = \sum_{r=0}^{\infty}\frac{(it)^r}{r!}\ x^r$$

II.2 Stress-Strength Reliability

Inferences about $R = P[Y < X]$, where X and Y are two independent random variables, is very common in the reliability literature. For example, if X is the strength of a component which is subject to a stress Y, then R is a measure of system performance and arises in the context of mechanical reliability of a system. The system fails if and only if at any time the applied stress is greater than its strength. Let Y and X be the stress and the strength random variables, independent of each other, follow respectively $DTGGO(\theta_2, \lambda_2, \mu_2, a, b)$ and $DTGGO(\theta_1, \lambda_1, \mu_1, a, b)$, then,

$$R = P(Y < X) = \int_a^b f_X(x) F_Y(x)\ dx$$

$$= \int_a^b \frac{\theta_1 \lambda_1 e^{\mu_1 x} e^{-\frac{\lambda_1}{\mu_1}(e^{\mu_1 x}-1)} \left(1-e^{-\frac{\lambda_1}{\mu_1}(e^{\mu_1 x}-1)}\right)^{\theta_1 -1}}{\left(1-e^{-\frac{\lambda_1}{\mu_1}(e^{\mu_1 b}-1)}\right)^{\theta_1} - \left(1-e^{-\frac{\lambda_1}{\mu_1}(e^{\mu_1 a}-1)}\right)^{\theta_1}} \frac{\left(1-e^{-\frac{\lambda_2}{\mu_2}(e^{\mu_2 x}-1)}\right)^{\theta_2} - \left(1-e^{-\frac{\lambda_2}{\mu_2}(e^{\mu_2 a}-1)}\right)^{\theta_2}}{\left(1-e^{-\frac{\lambda_2}{\mu_2}(e^{\mu_2 b}-1)}\right)^{\theta_2} - \left(1-e^{-\frac{\lambda_2}{\mu_2}(e^{\mu_2 a}-1)}\right)^{\theta_2}} dx$$

$$= \int_a^b \frac{\theta_1 \lambda_1 e^{\mu_1 x} e^{-\frac{\lambda_1}{\mu_1}(e^{\mu_1 x}-1)} \left(1-e^{-\frac{\lambda_1}{\mu_1}(e^{\mu_1 x}-1)}\right)^{\theta_1 -1} \left(1-e^{-\frac{\lambda_2}{\mu_2}(e^{\mu_2 x}-1)}\right)^{\theta_2}}{\left[\left(1-e^{-\frac{\lambda_1}{\mu_1}(e^{\mu_1 b}-1)}\right)^{\theta_1} - \left(1-e^{-\frac{\lambda_1}{\mu_1}(e^{\mu_1 a}-1)}\right)^{\theta_1}\right]\left[\left(1-e^{-\frac{\lambda_2}{\mu_2}(e^{\mu_2 b}-1)}\right)^{\theta_2} - \left(1-e^{-\frac{\lambda_2}{\mu_2}(e^{\mu_2 a}-1)}\right)^{\theta_2}\right]} dx$$

$$- \int_a^b \frac{\theta_1 \lambda_1 e^{\mu_1 x} e^{-\frac{\lambda_1}{\mu_1}(e^{\mu_1 x}-1)} \left(1-e^{-\frac{\lambda_1}{\mu_1}(e^{\mu_1 x}-1)}\right)^{\theta_1 -1} \left(1-e^{-\frac{\lambda_2}{\mu_2}(e^{\mu_2 a}-1)}\right)^{\theta_2}}{\left[\left(1-e^{-\frac{\lambda_1}{\mu_1}(e^{\mu_1 b}-1)}\right)^{\theta_1} - \left(1-e^{-\frac{\lambda_1}{\mu_1}(e^{\mu_1 a}-1)}\right)^{\theta_1}\right]\left[\left(1-e^{-\frac{\lambda_2}{\mu_2}(e^{\mu_2 b}-1)}\right)^{\theta_2} - \left(1-e^{-\frac{\lambda_2}{\mu_2}(e^{\mu_2 a}-1)}\right)^{\theta_2}\right]} dx$$

Let $K_1 = \dfrac{\theta_1 \lambda_1}{\left[\left(1-e^{-\frac{\lambda_1}{\mu_1}(e^{\mu_1 b}-1)}\right)^{\theta_1} - \left(1-e^{-\frac{\lambda_1}{\mu_1}(e^{\mu_1 a}-1)}\right)^{\theta_1}\right]\left[\left(1-e^{-\frac{\lambda_2}{\mu_2}(e^{\mu_2 b}-1)}\right)^{\theta_2} - \left(1-e^{-\frac{\lambda_2}{\mu_2}(e^{\mu_2 a}-1)}\right)^{\theta_2}\right]}$

And $K_2 = \dfrac{\left(1-e^{-\frac{\lambda_2}{\mu_2}(e^{\mu_2 a}-1)}\right)^{\theta_2}}{\left[\left(1-e^{-\frac{\lambda_2}{\mu_2}(e^{\mu_2 b}-1)}\right)^{\theta_2} - \left(1-e^{-\frac{\lambda_2}{\mu_2}(e^{\mu_2 a}-1)}\right)^{\theta_2}\right]}$, then,

$$R = K_1 \left[\int_a^b e^{\mu_1 x}\, e^{-\frac{\lambda_1}{\mu_1}(e^{\mu_1 x}-1)} \left(1 - e^{-\frac{\lambda_1}{\mu_1}(e^{\mu_1 x}-1)}\right)^{\theta_1 -1} \left(1 - e^{-\frac{\lambda_2}{\mu_2}(e^{\mu_2 x}-1)}\right)^{\theta_2} dx \right] - K_2$$

Since, $\left(1 - e^{-\frac{\lambda_1}{\mu_1}(e^{\mu x}-1)}\right)^{\theta_1 -1} = \sum_{j_1=0}^{\theta_1 -1} \binom{\theta_1 -1}{j_1} (1)^{\theta_1 -1-j_1} (-1)^{j_1} \left(e^{-\frac{\lambda_1}{\mu_1}(e^{\mu_1 x}-1) j_1}\right)$ and $\left(1 - e^{-\frac{\lambda_2}{\mu_2}(e^{\mu_2 x}-1)}\right)^{\theta_2} = \sum_{j_2=0}^{\theta_2} \binom{\theta_2}{j_2} (1)^{\theta_2 - j_2} (-1)^{j_2} \left(e^{-\frac{\lambda_2}{\mu_2}(e^{\mu_2 x}-1) j_2}\right)$, then,

$$R = K_1 \left[\begin{array}{c} \sum_{j_1=0}^{\theta_1-1} \binom{\theta_1 - 1}{j_1} (-1)^{j_1} \sum_{j_2=0}^{\theta_2} \binom{\theta_2}{j_2} (-1)^{j_2} \\ \int_a^b e^{\mu_1 x}\, e^{-\frac{\lambda_1}{\mu_1}(e^{\mu_1 x}-1)(j_1+1)} \left(e^{-\frac{\lambda_2}{\mu_2}(e^{\mu_2 x}-1) j_2} \right) dx \end{array} \right] - K_2$$

$$= K_1 \left[\begin{array}{c} \sum_{j_1=0}^{\theta_1-1} \binom{\theta_1-1}{j_1} (-1)^{j_1} \sum_{j_2=0}^{\theta_2} \binom{\theta_2}{j_2} (-1)^{j_2} \\ \int_a^b e^{\mu_1 x} \sum_{s_1=0}^{\infty} \frac{\left(-\frac{\lambda_1}{\mu_1}(e^{\mu_1 x}-1)(j_1+1)\right)^{s_1}}{s_1!} \, dx \\ \sum_{s_2=0}^{\infty} \frac{\left(-\frac{\lambda_2}{\mu_2}(e^{\mu_2 x}-1) j_2\right)^{s_2}}{s_2!} \end{array} \right] - K_2$$

$$= K_1 \left[\begin{array}{c} \sum_{j_1=0}^{\theta_1-1} \binom{\theta_1-1}{j_1} (-1)^{j_1} \sum_{j_2=0}^{\theta_2} \binom{\theta_2}{j_2} (-1)^{j_2} \sum_{s_1=0}^{\infty} \frac{\left(-\frac{\lambda_1}{\mu_1}(j_1+1)\right)^{s_1}}{s_1!} \\ \sum_{s_2=0}^{\infty} \frac{\left(-\frac{\lambda_2}{\mu_2} j_2\right)^{s_2}}{s_2!} \int_a^b e^{\mu_1 x} \, (e^{\mu_1 x} - 1)^{s_1} (e^{\mu_2 x} - 1)^{s_2} \, dx \end{array} \right] - K_2$$

Since $(e^{\mu_1 x} - 1)^{s_1} = \sum_{j_3=0}^{s_1} \binom{s_1}{j_3} (e^{\mu_1 x})^{s_1 - j_3} (-1)^{j_3}$ and

$(e^{\mu_2 x} - 1)^{s_2} = \sum_{j_4=0}^{s_2} \binom{s_2}{j_4} (e^{\mu_2 x})^{s_2 - j_4} (-1)^{j_4}$ then,

$$R = K_1 \left[\begin{array}{c} \sum_{j_1=0}^{\theta_1-1} \binom{\theta_1-1}{j_1} (-1)^{j_1} \sum_{j_2=0}^{\theta_2} \binom{\theta_2}{j_2} (-1)^{j_2} \sum_{s_1=0}^{\infty} \frac{\left(-\frac{\lambda_1}{\mu_1}(j_1+1)\right)^{s_1}}{s_1!} \\ \sum_{s_2=0}^{\infty} \frac{\left(-\frac{\lambda_2}{\mu_2} j_2\right)^{s_2}}{s_2!} \sum_{j_3=0}^{s_1} \binom{s_1}{j_3} (-1)^{j_3} \sum_{j_4=0}^{s_2} \binom{s_2}{j_4} (-1)^{j_4} \\ \int_a^b e^{-\mu_1 x (j_3 - s_1 - 1)} e^{-\mu_2 x (j_4 - s_2)} \, dx \end{array} \right] - K_2$$

$$= K_1 \left[\begin{array}{c} \sum_{j_1=0}^{\theta_1-1} \binom{\theta_1-1}{j_1} (-1)^{j_1} \sum_{j_2=0}^{\theta_2} \binom{\theta_2}{j_2} (-1)^{j_2} \sum_{s_1=0}^{\infty} \frac{\left(-\frac{\lambda_1}{\mu_1}(j_1+1)\right)^{s_1}}{s_1!} \\ \sum_{s_2=0}^{\infty} \frac{\left(-\frac{\lambda_2}{\mu_2} j_2\right)^{s_2}}{s_2!} \sum_{j_3=0}^{s_1} \binom{s_1}{j_3} (-1)^{j_3} \sum_{j_4=0}^{s_2} \binom{s_2}{j_4} (-1)^{j_4} \\ \int_a^b e^{-(\mu_1 j_3 - \mu_1 s_1 - \mu_1 - \mu_2 j_4 - \mu_2 s_2)x} \, dx \end{array} \right] - K_2$$

$$= K_1 \left[\begin{array}{c} \sum_{j_1=0}^{\theta_1-1} \binom{\theta_1-1}{j_1} (-1)^{j_1} \sum_{j_2=0}^{\theta_2} \binom{\theta_2}{j_2} (-1)^{j_2} \sum_{s_1=0}^{\infty} \frac{\left(-\frac{\lambda_1}{\mu_1}(j_1+1)\right)^{s_1}}{s_1!} \\ \sum_{s_2=0}^{\infty} \frac{\left(-\frac{\lambda_2}{\mu_2} j_2\right)^{s_2}}{s_2!} \sum_{j_3=0}^{s_1} \binom{s_1}{j_3} (-1)^{j_3} \sum_{j_4=0}^{s_2} \binom{s_2}{j_4} (-1)^{j_4} \\ \frac{e^{-(\mu_1 j_3 - \mu_1 s_1 - \mu_1 - \mu_2 j_4 - \mu_2 s_2)b} - e^{-(\mu_1 j_3 - \mu_1 s_1 - \mu_1 - \mu_2 j_4 - \mu_2 s_2)a}}{(\mu_1 j_3 - \mu_1 s_1 - \mu_1 - \mu_2 j_4 - \mu_2 s_2)} \end{array} \right] + K_2 \qquad (8)$$

II.3 Shannon entropy

The Shannon entropy of a random variable X is a measure of variation of the uncertainty. It is defined for a random variable X with values in a finite set X as $H = E\left(-\ln\left(g(x)\right)\right)$. So, the Shannon entropy of *DTGGO* random variable is,

$$H = E\left(-\ln \mathrm{A} - \mu X + \frac{\lambda}{\mu}\left(e^{\mu X}-1\right) - (\theta-1)\ln\left(1-e^{-\frac{\lambda}{\mu}\left(e^{\mu X}-1\right)}\right)\right)$$

$$= -\ln \mathrm{A} \ - \mu E(X) + \frac{\lambda}{\mu}E\left(e^{\mu X}\right) - \frac{\lambda}{\mu} - (\theta-1)E\left(\ln\left(1-e^{-\frac{\lambda}{\mu}\left(e^{\mu X}-1\right)}\right)\right)$$

Since, $E(X) = A\sum_{j_1=0}^{\theta-1}\binom{\theta-1}{j_1}(-1)^{j_1}\sum_{s_1=0}^{\infty}\frac{\left(-\frac{\lambda}{\mu}(j_1+1)\right)^{s_1}}{s_1!}\sum_{j_2=0}^{s_1}\binom{s_1}{j_2}(-1)^{j_2}$

$$\left[\left(\mu(j_2-s_1-1)\right)^{-2}\left\{\begin{matrix}\gamma\left((2),\mu b(j_2-s_1-1)\right)\\ -\gamma\left((2),\mu a(j_2-s_1-1)\right)\end{matrix}\right\}\right]$$

And $e^{\mu x} = \sum_{\varepsilon=0}^{\infty}\frac{(\mu x)^{\varepsilon}}{\varepsilon!}$ Then,

$$E\left(e^{\mu X}\right) = \sum_{\varepsilon=0}^{\infty}\frac{\mu^{\varepsilon}}{\varepsilon!}E\left(x^{\varepsilon}\right) = A\sum_{j_1=0}^{\theta-1}\binom{\theta-1}{j_1}(-1)^{j_1}\sum_{s_1=0}^{\infty}\frac{\left(-\frac{\lambda}{\mu}(j_1+1)\right)^{s_1}}{s_1!}\sum_{j_2=0}^{s_1}\binom{s_1}{j_2}(-1)^{j_2}$$

$$\sum_{\varepsilon=0}^{\infty}\frac{\mu^{\varepsilon}}{\varepsilon!}\left[\left(\mu(j_2-s_1-1)\right)^{-(\varepsilon+1)}\left\{\begin{matrix}\gamma\left((\varepsilon+1),\mu b(j_2-s_1-1)\right)\\ -\gamma\left((\varepsilon+1),\mu a(j_2-s_1-1)\right)\end{matrix}\right\}\right]$$

And,

$$E\left(\ln\left(1-e^{-\frac{\lambda}{\mu}\left(e^{\mu X}-1\right)}\right)\right) = A\left[\int_a^b\left(\begin{matrix}\ln\left(1-e^{-\frac{\lambda}{\mu}\left(e^{\mu x}-1\right)}\right)e^{\mu x}\\ e^{-\frac{\lambda}{\mu}\left(e^{\mu x}-1\right)}\left(1-e^{-\frac{\lambda}{\mu}\left(e^{\mu x}-1\right)}\right)^{\theta-1}\end{matrix}\right)dx\right]$$

$$= A\sum_{j_1=0}^{\theta-1}\binom{\theta-1}{j_1}(-1)^{j_1}\sum_{s_3=0}^{\infty}\frac{\left(-\frac{\lambda}{\mu}(j_1+1)\right)^{s_3}}{s_3!}\sum_{j_3=0}^{s_3}\binom{s_3}{j_3}(-1)^{j_3}\int_a^b\ln\left(1-e^{-\frac{\lambda}{\mu}\left(e^{\mu x}-1\right)}\right)e^{\mu x(j_3-s_3-1)}\ dx$$

Since, $ln(x) = -\left[x+\frac{x^2}{2}+\frac{x^3}{3}+\cdots\right] = -\sum_{u=1}^{\infty}\frac{x^u}{u},\quad x^2<1 \ and \ x=-1$

$$= A\sum_{j_1=0}^{\theta-1}\binom{\theta-1}{j_1}(-1)^{j_1}\sum_{s_3=0}^{\infty}\frac{\left(-\frac{\lambda}{\mu}(j_1+1)\right)^{s_3}}{s_3!}\sum_{j_3=0}^{s_3}\binom{s_3}{j_3}(-1)^{j_3}\left[\int_a^b-\sum_{l=1}^{\infty}\frac{e^{-\frac{\lambda l}{\mu}\left(e^{\mu x}-1\right)}}{l}e^{\mu x(j_3-s_3-1)}\ dx\right]$$

$$= -A\sum_{j_1=0}^{\theta-1}\binom{\theta-1}{j_1}(-1)^{j_1}\sum_{s_3=0}^{\infty}\frac{\left(-\frac{\lambda}{\mu}(j_1+1)\right)^{s_3}}{s_3!}\sum_{j_3=0}^{s_3}\binom{s_3}{j_3}(-1)^{j_3}\sum_{l=1}^{\infty}\frac{1}{l}\left[\int_a^b e^{-\frac{\lambda l}{\mu}\left(e^{\mu x}-1\right)}e^{\mu x(j_3-s_3-1)}\ dx\right]$$

$$= -A\sum_{j_1=0}^{\theta-1}\binom{\theta-1}{j_1}(-1)^{j_1}\sum_{s_3=0}^{\infty}\frac{\left(-\frac{\lambda}{\mu}(j_1+1)\right)^{s_3}}{s_3!}\sum_{j_3=0}^{s_3}\binom{s_3}{j_3}(-1)^{j_3}\sum_{l=1}^{\infty}\frac{e^{-\frac{\lambda l}{\mu}}}{l}\left[\int_a^b\sum_{m=1}^{\infty}\frac{\left(-\frac{\lambda l}{\mu}\right)^m e^{m\mu x}}{m!}e^{\mu x(j_3-s_3-}\right.$$

$$= -A\sum_{j_1=0}^{\theta-1}\binom{\theta-1}{j_1}(-1)^{j_1}\sum_{s_3=0}^{\infty}\frac{\left(-\frac{\lambda}{\mu}(j_1+1)\right)^{s_3}}{s_3!}\sum_{j_3=0}^{s_3}\binom{s_3}{j_3}(-1)^{j_3}\sum_{l=1}^{\infty}\sum_{m=0}^{\infty}\frac{e^{-\frac{\lambda l}{\mu}}}{lm!}\left(-\frac{\lambda l}{\mu}\right)^m\left[\int_a^b e^{\mu x(m+j_3-s_3\cdot}\right.$$

$$= -A\sum_{j_1=0}^{\theta-1}\binom{\theta-1}{j_1}(-1)^{j_1}\sum_{s_3=0}^{\infty}\frac{\left(-\frac{\lambda}{\mu}(j_1+1)\right)^{s_3}}{s_3!}\sum_{j_3=0}^{s_3}\binom{s_3}{j_3}(-1)^{j_3}\sum_{l=1}^{\infty}\sum_{m=0}^{\infty}\frac{e^{-\frac{\lambda l}{\mu}}}{lm!}\left(-\frac{\lambda l}{\mu}\right)^m\frac{e^{\mu(m+j_3-s_3-1)b}-e^{\mu(m+j_3-s_3}}{\mu(m+j_3-s_3-1)}$$

Then the entropy of DTGGO random variable is,

$$H = -\ln A - \mu A\sum_{j_1=0}^{\theta-1}\binom{\theta-1}{j_1}(-1)^{j_1}\sum_{s_1=0}^{\infty}\frac{\left(-\frac{\lambda}{\mu}(j_1+1)\right)^{s_1}}{s_1!}\sum_{j_2=0}^{s_1}\binom{s_1}{j_2}(-1)^{j_2}$$
$$\left[(\mu(j_2-s_1-1))^{-1}\gamma(2,b)-(\mu(j_2-s_1-1))^{-1}\gamma(2,a)\right]$$
$$+\frac{\lambda}{\mu}A\sum_{j_1=0}^{\theta-1}\binom{\theta-1}{j_1}(-1)^{j_1}\sum_{s_1=0}^{\infty}\frac{\left(-\frac{\lambda}{\mu}(j_1+1)\right)^{s_1}}{s_1!}\sum_{j_2=0}^{s_1}\binom{s_1}{j_2}(-1)^{j_2}$$
$$\sum_{\varepsilon=0}^{\infty}\frac{(\mu x)^{\varepsilon}}{\varepsilon!}\left[(\mu(j_2-s_1-1))^{-(\varepsilon+1)}\left[\gamma((\varepsilon+1),\mu b(j_2-s_1-1))-\gamma((\varepsilon+1),\mu a(j_2-s_1-1))\right]\right]$$
$$+(\theta-1)\left[A\sum_{j_1=0}^{\theta-1}\binom{\theta-1}{j_1}(-1)^{j_1}\sum_{s_3=0}^{\infty}\frac{\left(-\frac{\lambda}{\mu}(j_1+1)\right)^{s_3}}{s_3!}\sum_{j_3=0}^{s_3}\binom{s_3}{j_3}(-1)^{j_3}\sum_{l=1}^{\infty}\sum_{m=1}^{\infty}\frac{e^{-\frac{\lambda l}{\mu}}}{lm!}\left(-\frac{\lambda l}{\mu}\right)^m\frac{e^{\mu(m+j_3-s_3-1)b}-}{\mu(m+j_3\cdot}\right.$$

(9)

II.4 The relative entropy

The relative entropy (or the Kullback–Leibler divergence) is a measure of the difference between two probability distributions G and G^*. In applications G typically represents the "true" distribution of data, observations, or a precisely calculated theoretical distribution, while G^* typically represents a theory, model, description, or approximation of G. Specifically, the Kullback–Leibler divergence of G^* from G, denoted $D_{KL}(G||G^*)$, is a measure of the information gained when one revises ones beliefs from the prior probability distribution G^* to the posterior probability distribution G. More exactly, it is the amount of information that is lost when G^* is used to approximate G, defined operationally as the expected extra number of bits required to code samples from G using a code optimized for G^* rather than the code optimized for G.

The relative entropy $D_{KL} = (G||G^*)$ for *DTGG*O random variable is,

$$D_{KL} = (G||G^*) = \int_a^b g(x)\ln\left(\frac{g(x)}{g^*(x)}\right)dx$$
$$= \int_a^b g(x)\ln(g(x))\,dx - \int_a^b g(x)\ln(g^*(x))\,dx$$
$$= -H - \int_a^b g(x)\ln(g^*(x))\,dx$$

So, $\int_a^b g(x)\ln\left(g^*(x)\right)\,dx = \int_a^b \frac{\theta_1\lambda_1 e^{\mu_1 x} e^{-\frac{\lambda_1}{\mu_1}(e^{\mu_1 x}-1)}\left(1-e^{-\frac{\lambda_1}{\mu_1}(e^{\mu_1 x}-1)}\right)^{\theta_1-1}}{\left(1-e^{-\frac{\lambda_1}{\mu_1}(e^{\mu_1 b}-1)}\right)^{\theta_1}-\left(1-e^{-\frac{\lambda_1}{\mu_1}(e^{\mu_1 a}-1)}\right)^{\theta_1}}$

$$\ln\left(\frac{\theta_2\lambda_2 e^{\mu_2 x} e^{-\frac{\lambda_2}{\mu_2}(e^{\mu_2 x}-1)}\left(1-e^{-\frac{\lambda_2}{\mu_2}(e^{\mu_2 x}-1)}\right)^{\theta_2-1}}{\left(1-e^{-\frac{\lambda_2}{\mu_2}(e^{\mu_2 b}-1)}\right)^{\theta_2}-\left(1-e^{-\frac{\lambda_2}{\mu_2}(e^{\mu_2 a}-1)}\right)^{\theta_2}}\right) dx$$

Let $A_2 = \frac{\theta_2\lambda_2}{\left(1-e^{-\frac{\lambda_2}{\mu_2}(e^{\mu_2 b}-1)}\right)^{\theta_2}-\left(1-e^{-\frac{\lambda_2}{\mu_2}(e^{\mu_2 a}-1)}\right)^{\theta_2}}$, $A_1 = \frac{\theta_1\lambda_1}{\left(1-e^{-\frac{\lambda_1}{\mu_1}(e^{\mu_1 b}-1)}\right)^{\theta_1}-\left(1-e^{-\frac{\lambda_1}{\mu_1}(e^{\mu_1 a}-1)}\right)^{\theta_1}}$, then,

$$\int_a^b g(x)\ln\left(g^*(x)\right)\,dx = Ln(A_2) + \mu_2 E(x) - \frac{\lambda_2}{\mu_2}E\left(e^{\mu_2 x}-1\right) + (\theta_2-1)Ln\left(1-e^{-\frac{\lambda_2}{\mu_2}(e^{\mu_2 x}-1)}\right)$$

Since, $E(X) = A_1 \sum_{j_1=0}^{\theta_1-1}\binom{\theta_1-1}{j_1}(-1)^{j_1}\sum_{s_1=0}^{\infty}\frac{\left(-\frac{\lambda_1}{\mu_1}(j_1+1)\right)^{s_1}}{s_1!}\sum_{j_2=0}^{s_1}\binom{s_1}{j_2}(-1)^{j_2}$

$$\left[\left(\mu_1(j_2-s_1-1)\right)^{-2}\left\{\begin{array}{c}\gamma\left((2),\mu_1 b(j_2-s_1-1)\right)\\ -\gamma\left((2),\mu_1 a(j_2-s_1-1)\right)\end{array}\right\}\right]$$

And, $-\frac{\lambda_2}{\mu_2}E\left(e^{\mu_2 x}-1\right) = \frac{\lambda_2}{\mu_2}\left(1-Ee^{\mu_2 x}\right)$, where

$$E\left(e^{\mu_2 X}\right) = A_1\sum_{j_1=0}^{\theta_1-1}\binom{\theta_1-1}{j_1}(-1)^{j_1}\sum_{s_1=0}^{\infty}\frac{\left(-\frac{\lambda_1}{\mu_1}(j_1+1)\right)^{s_1}}{s_1!}\sum_{j_2=0}^{s_1}\binom{s_1}{j_2}(-1)^{j_2}\sum_{\varepsilon=0}^{\infty}\frac{(\mu_2 x)^{\varepsilon}}{\varepsilon!}\left[\left(\mu_1(j_2-s_1-1)\right)^{-(\varepsilon+1)}\left\{\begin{array}{c}\gamma\left((\varepsilon+1),\mu b(j_2-s_1-1)\right)\\ -\gamma\left((\varepsilon+1),\mu a(j_2-s_1-1)\right)\end{array}\right\}\right]$$

And,

$$(\theta_2-1)Ln\left(1-e^{-\frac{\lambda_2}{\mu_2}(e^{\mu_2 x}-1)}\right) = (1-\theta_2)E\sum_{u=1}^{\infty}\frac{e^{-\frac{u\lambda_2}{\mu_2}e^{(\mu_2 x-1)}}}{u}$$

$$= (1-\theta_2)\sum_{u=1}^{\infty}\frac{e^{\frac{u\lambda_2}{\mu_2}}}{u}Ee^{-\frac{u\lambda_2}{\mu_2}e^{\mu_2 x}} = (1-\theta_2)\sum_{u=1}^{\infty}\frac{e^{\frac{u\lambda_2}{\mu_2}}}{u}E\sum_{l_1=0}^{\infty}\frac{-\frac{u\lambda_2}{\mu_2}^{l_1}}{l_1!}\, e^{l_1\mu_2 x}$$

$$= (1-\theta_2)\sum_{u=1}^{\infty}\frac{e^{\frac{u\lambda_2}{\mu_2}}}{u}\sum_{l_1=0}^{\infty}\frac{-\frac{u\lambda_2}{\mu_2}^{l_1}}{l_1!}\, E\, e^{l_1\mu_2 x}$$

$$= (1-\theta_2)\sum_{u=1}^{\infty}\sum_{l_1=0}^{\infty}\frac{-\frac{u\lambda_2}{\mu_2}^{l_1}}{l_1!}\frac{e^{\frac{u\lambda_2}{\mu_2}}}{u}\ \sum_{l_2=0}^{\infty}\frac{(l_1\mu_2 x)^{l_2}}{l_2!}$$

$$= (1-\theta_2)\sum_{u=1}^{\infty}\sum_{l_1=0}^{\infty}\sum_{l_2=0}^{\infty}\frac{(l_1\mu_2)^{l_2}}{l_2!}\frac{-\frac{u\lambda_2}{\mu_2}^{l_1}}{l_1!}\frac{e^{\frac{u\lambda_2}{\mu_2}}}{u}\ EX^{l_2}$$

$$= (1-\theta_2)\sum_{u=1}^{\infty}\sum_{l_1=0}^{\infty}\sum_{l_2=0}^{\infty}\frac{(l_1\mu_2)^{l_2}}{l_2!}\frac{-\frac{u\lambda_2}{\mu_2}^{l_1}}{l_1!}\frac{e^{\frac{u\lambda_2}{\mu_2}}}{u}$$

$$A_1\sum_{j_1=0}^{\theta_1-1}\binom{\theta_1-1}{j_1}(-1)^{j_1}\ \sum_{s_1=0}^{\infty}\frac{\left(-\frac{\lambda_1}{\mu_1}(j_1+1)\right)^{s_1}}{s_1!}\sum_{j_2=0}^{s_1}\binom{s_1}{j_2}(-1)^{j_2}$$

$$\left[\left(\mu(j_2-s_1-1)\right)^{-(r+1)}\left\{\begin{matrix}\gamma\left((l_2+1),\mu b(j_2-s_1-1)\right)\\ -\gamma\left((l_2+1),\mu a(j_2-s_1-1)\right)\end{matrix}\right\}\right]$$

Then, the relative entropy is,

$$D_{KL} = \ln A_1 + \mu_1 A_1\sum_{j_1=0}^{\theta_1-1}\binom{\theta_1-1}{j_1}(-1)^{j_1}\ \sum_{s_1=0}^{\infty}\frac{\left(-\frac{\lambda_1}{\mu_1}(j_1+1)\right)^{s_1}}{s_1!}\sum_{j_2=0}^{s_1}\binom{s_1}{j_2}(-1)^{j_2}$$

$$\left[\left(\mu_1(j_2-s_1-1)\right)^{-1}\gamma(2,b)-\left(\mu_1(j_2-s_1-1)\right)^{-1}\gamma(2,a)\right]$$

$$-\frac{\lambda_1}{\mu_1}A_1\sum_{j_1=0}^{\theta_1-1}\binom{\theta_1-1}{j_1}(-1)^{j_1}\ \sum_{s_1=0}^{\infty}\frac{\left(-\frac{\lambda_1}{\mu_1}(j_1+1)\right)^{s_1}}{s_1!}\sum_{j_2=0}^{s_1}\binom{s_1}{j_2}(-1)^{j_2}$$

$$\sum_{\varepsilon=0}^{\infty}\frac{\mu_1^{\varepsilon}}{\varepsilon!}\left[\left(\mu_1(j_2-s_1-1)\right)^{-(\varepsilon+1)}\left[\gamma\left((\varepsilon+1),\mu_1 b(j_2-s_1-1)\right)-\gamma\left((\varepsilon+1),\mu_1 a(j_2-s_1-1)\right)\right]\right]$$

$$-(\theta_1-1)\left[\begin{matrix}A_1\sum_{j_1=0}^{\theta_1-1}\binom{\theta_1-1}{j_1}(-1)^{j_1}\ \sum_{s_3=0}^{\infty}\frac{\left(-\frac{\lambda_1}{\mu_1}(j_1+1)\right)^{s_3}}{s_3!}\sum_{j_3=0}^{s_3}\binom{s_3}{j_3}(-1)^{j_3}\\ \sum_{l=1}^{\infty}\sum_{m=1}^{\infty}\frac{e^{-\frac{\lambda_1 l}{\mu_1}}}{lm!}\left(-\frac{\lambda_1 l}{\mu_1}\right)^{m}\frac{e^{\mu_1(m+j_3-s_3-1)b}-e^{\mu_1(m+j_3-s_3-1)a}}{\mu_1(m+j_3-s_3-1)}\end{matrix}\right]-$$

$$\mu_2 A_1\sum_{j_1=0}^{\theta_1-1}\binom{\theta_1-1}{j_1}(-1)^{j_1}\ \sum_{s_1=0}^{\infty}\frac{\left(-\frac{\lambda_1}{\mu_1}(j_1+1)\right)^{s_1}}{s_1!}\sum_{j_2=0}^{s_1}\binom{s_1}{j_2}(-1)^{j_2}$$

$$\left[\left(\mu_1(j_2-s_1-1)\right)^{-2}\left\{\begin{matrix}\gamma\left((2),\mu_1 b(j_2-s_1-1)\right)\\ -\gamma\left((2),\mu_1 a(j_2-s_1-1)\right)\end{matrix}\right\}\right]-\frac{\lambda_2}{\mu_2}$$

$$\left(\begin{matrix}1-A_1\sum_{j_1=0}^{\theta_1-1}\binom{\theta_1-1}{j_1}(-1)^{j_1}\ \sum_{s_1=0}^{\infty}\frac{\left(-\frac{\lambda_1}{\mu_1}(j_1+1)\right)^{s_1}}{s_1!}\sum_{j_2=0}^{s_1}\binom{s_1}{j_2}(-1)^{j_2}\sum_{\varepsilon=0}^{\infty}\frac{(\mu_2 x)^{\varepsilon}}{\varepsilon!}\\ \left[\left(\mu_1(j_2-s_1-\ 1)\right)^{-(\varepsilon+1)}\left\{\begin{matrix}\gamma\left((\varepsilon+1),\mu b(j_2-s_1-1)\right)\\ -\gamma\left((\varepsilon+1),\mu a(j_2-s_1-1)\right)\end{matrix}\right\}\right]\end{matrix}\right)+(\theta_2-1)$$

$$\sum_{u=1}^{\infty}\sum_{l_1=0}^{\infty}\sum_{l_2=0}^{\infty}\frac{(l_1\mu_2)^{l_2}}{l_2!}\frac{-\frac{u\lambda_2}{\mu_2}^{l_1}}{l_1!}\frac{e^{\frac{u\lambda_2}{\mu_2}}}{u}A_1\sum_{j_1=0}^{\theta_1-1}\binom{\theta_1-1}{j_1}(-1)^{j_1}\ \sum_{s_1=0}^{\infty}\frac{\left(-\frac{\lambda_1}{\mu_1}(j_1+1)\right)^{s_1}}{s_1!}\sum_{j_2=0}^{s_1}\binom{s_1}{j_2}(-1)^{j_2}$$

$$\left[\left(\mu(j_2 - s_1 - 1)\right)^{-(r+1)} \begin{Bmatrix} \gamma\left((l_2+1), \mu b(j_2 - s_1 - 1)\right) \\ -\gamma\left((l_2+1), \mu a(j_2 - s_1 - 1)\right) \end{Bmatrix}\right] \quad (10)$$

VI. **Marshall-Olkin extended Uniform distribution**

Jose and Krishna in 2011 introduced the Marshall-Olkin Extended Uniform (MOEU) distribution. MOEU distribution is expressed as a mixture distribution with exponential distribution as mixing density. They derived also the limiting distributions of sample maxima and sample minima. Record value properties of the new distribution are discussed. Simulation studies are conducted to Estimate unknown parameters. The MOEU stress-strength model R is obtained in 2015 by Abid and Hassan, where the stress and the strength are independent MOEU distributions with different scale parameters and different shape parameters. Different methods to estimate R and MOEU distribution parameters are studied, maximum likelihood estimator, method of moments estimator, percentiles estimator, least squares estimator, weighted least squares estimator, L-moment estimator and regression estimator. An empirical study was conducted to support the theoretical aspect. In 2015 Abid and Hassan introduced the Beta Marshall-Olkin extended uniform (BMOEU) distribution. The *r*th moment, , the cumulative distribution function, the reliability function and the hazard rate function is obtained for the new distribution. The BMOEU strength-stress model with different eight parameters will be derived here. In reliability theory, a combination of two distributions failure rate model for reliability studies is paid much attention. In 2015 also Abid and Hassan derived the failure rate model of MOEU(α,θ) and every one of MOEU(a,b), MOEU(a,θ), Uniform(θ), truncated exponential (λ,θ), truncated Weibull(λ,k,θ), truncated Frechet (a,b,θ), truncated Rayleigh (σ^2,θ), doubly truncated Cauchy(a,b,θ)and doubly truncated Gumbel (a,b,θ) distributions.

The probability density and cumulative functions of MOEU random variable are respectively given by,

$$f(x,\varphi,\tau) = \frac{\varphi\tau}{(\varphi\tau+(1-\varphi)x)^2} \qquad ,0 < x < \tau, \varphi > 0. \quad (11)$$

$$F(x,\varphi,\tau) = \frac{x}{(\varphi\tau+(1-\varphi)x)}\,, \qquad 0 < x < \tau, \varphi > 0. \quad (12)$$

<u>III.1 Essential properties of *DTMOEU*$(\boldsymbol{\alpha},\boldsymbol{\sigma},\boldsymbol{\beta},\boldsymbol{a},\boldsymbol{b})$</u>

The pdf and cdf of *DTGIW* random variable are respectively,

$$g(x) = \frac{f(x)}{F(b)-F(a)} \qquad ,0 < a < x < b \ \ ,b < \tau \ and\ \varphi,\tau > 0$$

$$= \frac{\lambda(\lambda+(1-\varphi)a)(\lambda+(1-\varphi)b)}{(\lambda+(1-\varphi)x)^2[b(\lambda+(1-\varphi)a)-a(\lambda+(1-\varphi)b)]}\ , \text{Where, } \varphi\tau = \lambda > 0. \quad (13)$$

$$G(x) = \frac{F(x)-F(a)}{F(b)-F(a)}$$

$$= \frac{(\lambda+(1-\varphi)b)[x(\lambda+(1-\varphi)a)-a(\lambda+(1-\varphi)x)]}{(\lambda+(1-\varphi)x)[b(\lambda+(1-\varphi)a)-a(\lambda+(1-\varphi)b)]}, \qquad a < x < b \quad (14)$$

Also the reliability and hazard functions are respectively,

$$R(x) = 1 - G(x)$$

$$= \frac{b(\lambda+(1-\varphi)x)(\lambda+(1-\varphi)a)-x(\lambda+(1-\varphi)b)(\lambda+(1-\varphi)a)}{(\lambda+(1-\varphi)x)[b(\lambda+(1-\varphi)a)-a(\lambda+(1-\varphi)b)]} \quad (15)$$

$$H(x) = \frac{g(x)}{R(x)}$$

$$= \frac{\lambda(\lambda+(1-\varphi)a)(\lambda+(1-\varphi)b)}{(\lambda+(1-\varphi)x)[b(\lambda+(1-\varphi)x)(\lambda+(1-\varphi)a)-x(\lambda+(1-\varphi)b)(\lambda+(1-\varphi)a)]} \quad (16)$$

So, the rth raw moment is,

$$E(X^r) = \int_a^b x^r \, g(x) \, dx$$

$$= \int_a^b x^r \frac{\lambda(\lambda+(1-\varphi)a)(\lambda+(1-\varphi)b)}{(\lambda+(1-\varphi)x)^2[b(\lambda+(1-\varphi)a)-a(\lambda+(1-\varphi)b)]} \, dx$$

Suppose $A = \frac{\lambda(\lambda+(1-\varphi)a)(\lambda+(1-\varphi)b)}{[b(\lambda+(1-\varphi)a)-a(\lambda+(1-\varphi)b)]}$, then,

$$E(X^r) = A \int_a^b \frac{x^r}{(\lambda+(1-\varphi)x)^2} \, dx$$

$$= \frac{A}{(1-\varphi)^{r+1}} [h(b) - h(a)]$$

Since, $\int_a^b \frac{x^m}{(c+dx)^n} \, dx = \frac{1}{d^{(m+1)}} \left(\sum_{\substack{s_1=0 \\ s_1 \neq m-n+1}}^{m} \frac{m!(-c)^{s_1}(c+dx)^{m-n-s_1+1}}{(m-s_1)!s_1!(m-n-s_1+1)} + \frac{m!(-c)^{m-n+1}}{(m-n+1)!(n-1)!} log|(c+dx)| \right)$

Then, $E(X^r) = \frac{A}{(1-\varphi)^{r+1}} [h(b) - h(a)]$, where,

$$h(\delta) = \sum_{\substack{s_1=0 \\ s_1 \neq r-1}}^{m} \frac{r!(-\lambda)^{s_1}(\lambda+(1-\varphi)\delta)^{r-s_1-1}}{(r-s_1)!s_1!(r-s_1-1)} + \frac{r!(-\lambda)^{r-1}}{(r-1)!} log|(\lambda + (1-\varphi)\delta)| \quad (17)$$

Then, the characteristic function can easily get by using the relation,

$$Q_x(t) = E\left(e^{ixt}\right) = \sum_{r=0}^{\infty} \frac{(it)^r}{r!} E(x^r) \quad , \quad \text{since} \quad e^{ixt} = \sum_{r=0}^{\infty} \frac{(it)^r}{r!} \, x^r$$

III.2 Stress-Strength Reliability

Suppose $X \sim DTMOEU(\varphi_1, \lambda_1, a, b)$ and $Y \sim DTMOEU(\varphi_2, \lambda_2, a, b)$ with unknown parameters $\varphi_1, \lambda_1, \varphi_2, \lambda_2, a, b$, where X and Y are independently distributed, then the stress-stress reliability function is,

$$R = P(Y < X) = \int_a^b g_X(x) G_Y(x) \, dx$$

$$g_X(x) = \frac{\lambda_1(\lambda_1+(1-\varphi_1)a)(\lambda_1+(1-\varphi_1)b)}{(\lambda_1+(1-\varphi_1)x)^2[b(\lambda_1+(1-\varphi_1)a)-a(\lambda_1+(1-\varphi_1)b)]} \quad , \ a < x < b$$

$$G_Y(x) = \frac{(\lambda_2+(1-\varphi_2)b)[x(\lambda_2+(1-\varphi_2)a)-a(\lambda_2+(1-\varphi_2)x)]}{(\lambda_2+(1-\varphi_2)x)[b(\lambda_2+(1-\varphi_2)a)-a(\lambda_2+(1-\varphi_2)b)]} \quad , \ a < x < b$$

$$R = \int_a^b \left[\frac{\lambda_1(\lambda_1+(1-\varphi_1)a)(\lambda_1+(1-\varphi_1)b)}{(\lambda_1+(1-\varphi_1)x)^2[b(\lambda_1+(1-\varphi_1)a)-a(\lambda_1+(1-\varphi_1)b)]}\right] \frac{(\lambda_2+(1-\varphi_2)b)\begin{bmatrix} x(\lambda_2+(1-\varphi_2)a) \\ -a(\lambda_2+(1-\varphi_2)x) \end{bmatrix}}{(\lambda_2+(1-\varphi_2)x)\begin{bmatrix} b(\lambda_2+(1-\varphi_2)a) \\ -a(\lambda_2+(1-\varphi_2)b) \end{bmatrix}} \, dx$$

Let $B = \frac{(\lambda_2+(1-\varphi_2)b)[\lambda_1(\lambda_1+(1-\varphi_1)a)(\lambda_1+(1-\varphi_1)b)]}{\begin{bmatrix} b(\lambda_1+(1-\varphi_1)a) \\ -a(\lambda_1+(1-\varphi_1)b) \end{bmatrix}\begin{bmatrix} b(\lambda_2+(1-\varphi_2)a) \\ -a(\lambda_2+(1-\varphi_2)b) \end{bmatrix}}$

$$R = B \int_a^b \left[\frac{[x(\lambda_2+(1-\varphi_2)a)-a(\lambda_2+(1-\varphi_2)x)]}{(\lambda_1+(1-\varphi_1)x)^2(\lambda_2+(1-\varphi_2)x)}\right] dx$$

$$= B\int_a^b \left[\frac{x(\lambda_2+(1-\varphi_2)a)}{(\lambda_1+(1-\varphi_1)x)^2(\lambda_2+(1-\varphi_2)x)}\right] dx - B\int_a^b \left[\frac{a}{(\lambda_1+(1-\varphi_1)x)^2}\right] dx$$

Let, $C = B(\lambda_2 + (1-\varphi_2)a)$

Since $\int_a^b \frac{dx}{(c+dx)^n} = \frac{-1}{(n-1)d(c+dx)^{n-1}}$ where *c* and *d* are constant, then,

$$R = C\int_a^b \left[\frac{x}{(\lambda_1+(1-\varphi_1)x)^2(\lambda_2+(1-\varphi_2)x)}\right] dx + \frac{aB}{(1-\varphi_1)}\left[\frac{1}{(\lambda_1+(1-\varphi_1)b)} - \frac{1}{(\lambda_1+(1-\varphi_1)a)}\right]$$

Suppose $y = \lambda_2 + (1-\varphi_2)x \rightarrow x = \frac{y-\lambda_2}{(1-\varphi_2)} \rightarrow dx = \frac{dy}{(1-\varphi_2)}$

Then $x = a \rightarrow y = \lambda_2 + (1-\varphi_2)a = a_1$ and $x = b \rightarrow y = \lambda_2 + (1-\varphi_2)b = b_1$, then,

$$R = C\int_{a_1}^{b_1} \left[\frac{\frac{y-\lambda_2}{(1-\varphi_2)}}{\left(\lambda_1+(1-\varphi_1)\frac{y-\lambda_2}{(1-\varphi_2)}\right)^2 y}\right] \frac{dy}{(1-\varphi_2)} + B\left[\frac{a}{(1-\varphi_1)(\lambda_1+(1-\varphi_1)b)} - \frac{a}{(1-\varphi_1)(\lambda_1+(1-\varphi_1)a)}\right]$$

$$= C\int_{a_1}^{b_1} \left[\frac{y-\lambda_2}{\left(\lambda_1(1-\varphi_2)+(1-\varphi_1)(y-\lambda_2)\right)^2 y}\right] dy + B\left[\frac{a}{(1-\varphi_1)(\lambda_1+(1-\varphi_1)b)} - \frac{a}{(1-\varphi_1)(\lambda_1+(1-\varphi_1)a)}\right]$$

$$\int_{a_1}^{b_1} \left[\frac{(y-\lambda_2)\,dy}{\left(\lambda_1(1-\varphi_2)+(1-\varphi_1)(y-\lambda_2)\right)^2 y}\right] = \int_{a_1}^{b_1} \left[\frac{dy}{\left(\lambda_1(1-\varphi_2)+(1-\varphi_1)(y-\lambda_2)\right)^2}\right]$$

$$-\lambda_2\int_{a_1}^{b_1} \left[\frac{dy}{\left(\lambda_1(1-\varphi_2)+(1-\varphi_1)(y-\lambda_2)\right)^2 y}\right]$$

Since $\int \frac{dx}{(c+lx)^2} = \frac{-1}{d(c+lx)}$ and $\int \frac{dx}{(c+lx)^2 x} = \frac{-1}{c^2}\left[Ln\left|\frac{c+lx}{x}\right| + \frac{lx}{c+lx}\right]$, then,

$$\int_{a_1}^{b_1} \left[\frac{y-\lambda_2}{\left(\lambda_1(1-\varphi_2)+(1-\varphi_1)(y-\lambda_2)\right)^2 y}\right] =$$
$$\left\{\frac{1}{(1-\varphi_1)[\lambda_1(1-\varphi_2)-\lambda_2(1-\varphi_1)+(1-\varphi_1)a_1]} - \frac{1}{(1-\varphi_1)[\lambda_1(1-\varphi_2)-\lambda_2(1-\varphi_1)+(1-\varphi_1)a_1]}\right\} +$$
$$\lambda_2\left\{\frac{1}{[\lambda_1(1-\varphi_2)-\lambda_2(1-\varphi_1)]^2}\right\}\left[\left\{Ln\left|\frac{\lambda_1(1-\varphi_2)-\lambda_2(1-\varphi_1)+(1-\varphi_1)b_1}{b_1}\right| + \frac{(1-\varphi_1)b_1}{\lambda_1(1-\varphi_2)-\lambda_2(1-\varphi_1)+(1-\varphi_1)b_1}\right\} -\right.$$
$$\left.\left\{Ln\left|\frac{\lambda_1(1-\varphi_2)-\lambda_2(1-\varphi_1)+(1-\varphi_1)a_1}{a_1}\right| + \frac{(1-\varphi_1)a_1}{\lambda_1(1-\varphi_2)-\lambda_2(1-\varphi_1)+(1-\varphi_1)a_1}\right\}\right]$$

, then

$$R =$$
$$C\left[\left\{\frac{1}{(1-\varphi_1)[\lambda_1(1-\varphi_2)-\lambda_2(1-\varphi_1)+(1-\varphi_1)a_1]} - \frac{1}{(1-\varphi_1)[\lambda_1(1-\varphi_2)-\lambda_2(1-\varphi_1)+(1-\varphi_1)a_1]}\right\} +\right.$$
$$\lambda_2\left\{\frac{1}{[\lambda_1(1-\varphi_2)-\lambda_2(1-\varphi_1)]^2}\right\}\left[\left\{Ln\left|\frac{\lambda_1(1-\varphi_2)-\lambda_2(1-\varphi_1)+(1-\varphi_1)b_1}{b_1}\right| + \frac{(1-\varphi_1)b_1}{\lambda_1(1-\varphi_2)-\lambda_2(1-\varphi_1)+(1-\varphi_1)b_1}\right\} -\right.$$
$$\left.\left.\left\{Ln\left|\frac{\lambda_1(1-\varphi_2)-\lambda_2(1-\varphi_1)+(1-\varphi_1)a_1}{a_1}\right| + \frac{(1-\varphi_1)a_1}{\lambda_1(1-\varphi_2)-\lambda_2(1-\varphi_1)+(1-\varphi_1)a_1}\right\}\right]\right]$$
$$+B\left[\frac{a}{(1-\varphi_1)(\lambda_1+(1-\varphi_1)b)} - \frac{a}{(1-\varphi_1)(\lambda_1+(1-\varphi_1)a)}\right] \quad (18)$$

III.3 Shannon entropy

The Shannon entropy of *DTMOEU* random variable can be found as,

$$\mathrm{H}=\mathrm{E}\left(-\ln\left(\frac{\lambda(\lambda+(1-\varphi)a)(\lambda+(1-\varphi)b)}{(\lambda+(1-\varphi)x)^2\begin{bmatrix}b(\lambda+(1-\varphi)a)\\-a(\lambda+(1-\varphi)b)\end{bmatrix}}\right)\right)=\mathrm{E}\left(\begin{matrix}-\ln\left(\frac{\lambda(\lambda+(1-\varphi)a)(\lambda+(1-\varphi)b)}{[b(\lambda+(1-\varphi)a)-a(\lambda+(1-\varphi)b)]}\right)\\+2\ln(\lambda+(1-\varphi)x)\end{matrix}\right)$$

Suppose $D=\frac{\lambda(\lambda+(1-\varphi)a)(\lambda+(1-\varphi)b)}{[b(\lambda+(1-\varphi)a)-a(\lambda+(1-\varphi)b)]}$ Then $\mathrm{H}=-\ln(D)+2\mathrm{E}(\ln(\lambda+(1-\varphi)x))$

$$\mathrm{E}(\ln(\lambda+(1-\varphi)x))=\int_a^b\frac{\ln(\lambda+(1-\varphi)x)\begin{bmatrix}\lambda(\lambda+(1-\varphi)a)\\(\lambda+(1-\varphi)b)\end{bmatrix}}{(\lambda+(1-\varphi)x)^2\begin{bmatrix}b(\lambda+(1-\varphi)a)\\-a(\lambda+(1-\varphi)b)\end{bmatrix}}dx=D\int_a^b\frac{\ln(\lambda+(1-\varphi)x)}{(\lambda+(1-\varphi)x)^2}dx$$

Let $y=\lambda+(1-\varphi)x\rightarrow x=\frac{y-\lambda}{(1-\varphi)}\rightarrow dx=\frac{dy}{(1-\varphi)}$

Then $x=a\rightarrow y=\lambda+(1-\varphi)a=a_1$ And $x=b\rightarrow y=\lambda+(1-\varphi)b=b_1$

$$\mathrm{E}(\ln(\lambda+(1-\varphi)x))=\frac{D}{(1-\varphi)}\int_{a_1}^{b_1}\frac{\ln(y)}{y^2}dy\ =\frac{-D}{(1-\varphi)}\left\{\frac{|1+Ln(b_1)|}{b_1}-\frac{|1+Ln(a_1)|}{a_1}\right\},$$

, Since $\int\frac{\ln(x)}{x^n}dx=-\frac{\ln(x)}{x}-\frac{1}{x}$, then the Shannon entropy is,

$$\mathrm{H}=-\ln(D)-\frac{2D}{(1-\varphi)}\left[\left(\frac{\ln(b_1)+1}{b_1}\right)-\left(\frac{\ln(a_1)+1}{a_1}\right)\right]\qquad(19)$$

III.4 The relative entropy

The relative entropy $D_{KL}=(G||G^*)$ for *DTGIW* random variable is,

$D_{KL}=(G||G^*)=-H-\int_a^b g(x)\ln\left(g^*(x)\right)\,dx$, then,

$$\int_a^b g(x)\,ln\left(g^*(x)\right)\,dx=\int_a^b\left[\begin{matrix}\frac{\lambda_1(\lambda_1+(1-\varphi_1)a)(\lambda_1+(1-\varphi_1)b)}{(\lambda_1+(1-\varphi_1)x)^2\begin{bmatrix}b(\lambda_1+(1-\varphi_1)a)\\-a(\lambda_1+(1-\varphi_1)b)\end{bmatrix}}\\ ln\left(\frac{\lambda_2(\lambda_2+(1-\varphi_2)a)(\lambda_2+(1-\varphi_2)b)}{(\lambda_2+(1-\varphi_2)x)^2\begin{bmatrix}b(\lambda_2+(1-\varphi_2)a)\\-a(\lambda_2+(1-\varphi_2)b)\end{bmatrix}}\right)\end{matrix}\right]dx$$

$$=\int_a^b\left[\begin{matrix}\frac{\lambda_1(\lambda_1+(1-\varphi_1)a)(\lambda_1+(1-\varphi_1)b)}{(\lambda_1+(1-\varphi_1)x)^2\begin{bmatrix}b(\lambda_1+(1-\varphi_1)a)\\-a(\lambda_1+(1-\varphi_1)b)\end{bmatrix}}\\ ln\left(\frac{\lambda_2(\lambda_2+(1-\varphi_2)a)(\lambda_2+(1-\varphi_2)b)}{\begin{bmatrix}b(\lambda_2+(1-\varphi_2)a)\\-a(\lambda_2+(1-\varphi_2)b)\end{bmatrix}}\right)\end{matrix}\right]dx-\int_a^b\left[\begin{matrix}\frac{\lambda_1(\lambda_1+(1-\varphi_1)a)(\lambda_1+(1-\varphi_1)b)}{(\lambda_1+(1-\varphi_1)x)^2\begin{bmatrix}b(\lambda_1+(1-\varphi_1)a)\\-a(\lambda_1+(1-\varphi_1)b)\end{bmatrix}}\\ 2\,ln((\lambda_2+(1-\varphi_2)x))\end{matrix}\right]dx$$

Suppose $A_1=\frac{\lambda_1(\lambda_1+(1-\varphi_1)a)(\lambda_1+(1-\varphi_1)b)}{[b(\lambda_1+(1-\varphi_1)a)-a(\lambda_1+(1-\varphi_1)b)]}$ And $A_2=ln\left(\frac{\lambda_2(\lambda_2+(1-\varphi_2)a)(\lambda_2+(1-\varphi_2)b)}{\begin{bmatrix}b(\lambda_2+(1-\varphi_2)a)\\-a(\lambda_2+(1-\varphi_2)b)\end{bmatrix}}\right)$

Then $\int_a^b g(x)\,ln\left(g^*(x)\right)\,dx=A_2-2A_1\int_a^b\left[\frac{ln((\lambda_2+(1-\varphi_2)x))}{(\lambda_1+(1-\varphi_1)x)^2}\right]dx$

Suppose $u=\lambda_1+(1-\varphi_1)x\ \rightarrow x=\frac{u-\varphi_1\tau_1}{(1-\varphi_1)}\rightarrow dx=\frac{du}{(1-\varphi_1)}$

Then $x = a \rightarrow u = \lambda_1 + (1-\varphi_1)a = a_2$

And $x = b \rightarrow u = \lambda_1 + (1-\varphi_1)b = b_2$

Then $\int_a^b g(x)\, ln\big(g^*(x)\big)\ dx = A_2 - \frac{2A_1}{(1-\varphi_1)} \int_{a_2}^{b_2} \left[\frac{ln\left(\frac{(1-\varphi_1)\lambda_2-(1-\varphi_2)\lambda_1}{(1-\varphi_1)} + \frac{(1-\varphi_2)}{(1-\varphi_1)}u\right)}{u^2}\right] du$

Since $\int_{a_2}^{b_2} \left[\frac{ln(g+fx)}{x^2}\right] dx = \frac{f}{g}\ln x - \left(\frac{1}{x}+\frac{f}{g}\right) ln(g+fx)$

$$\int_a^b g(x)\, ln\big(g^*(x)\big)\ dx = A_2 - \frac{2A_1}{(1-\varphi_1)} \left[\left(\begin{array}{c} \frac{(1-\varphi_2)}{[(1-\varphi_1)\lambda_2-(1-\varphi_2)\lambda_1]} \ln b_2 \\ - \left[\begin{array}{c} \left(\frac{1}{b_2} + \frac{(1-\varphi_2)}{[(1-\varphi_1)\lambda_2-(1-\varphi_2)\lambda_1]}\right) \\ ln\left(\frac{(1-\varphi_1)\lambda_2-(1-\varphi_2)\lambda_1}{(1-\varphi_1)} + \frac{(1-\varphi_2)}{(1-\varphi_1)} b_2\right) \end{array} \right] \end{array} \right) - \left(\begin{array}{c} \frac{(1-\varphi_2)}{[(1-\varphi_1)\lambda_2-(1-\varphi_2)\lambda_1]} \ln a_2 \\ - \left[\begin{array}{c} \left(\frac{1}{a_2} + \frac{(1-\varphi_2)}{[(1-\varphi_1)\lambda_2-(1-\varphi_2)\lambda_1]}\right) \\ ln\left(\frac{(1-\varphi_1)\lambda_2-(1-\varphi_2)\lambda_1}{(1-\varphi_1)} + \frac{(1-\varphi_2)}{(1-\varphi_1)} a_2\right) \end{array} \right] \end{array} \right) \right]$$

Then,

$$D_{KL} = \ln(D) + \frac{2D}{(1-\varphi)} \left[\left(\frac{\ln(b_1)+1}{b_1}\right) - \left(\frac{\ln(a_1)+1}{a_1}\right) \right] -$$

$$A_2 + \frac{2A_1}{(1-\varphi_1)} \left[\left(\begin{array}{c} \frac{(1-\varphi_2)}{[(1-\varphi_1)\lambda_2-(1-\varphi_2)\lambda_1]} \ln b_2 \\ - \left[\begin{array}{c} \left(\frac{1}{b_2} + \frac{(1-\varphi_2)}{[(1-\varphi_1)\lambda_2-(1-\varphi_2)\lambda_1]}\right) \\ ln\left(\frac{(1-\varphi_1)\lambda_2-(1-\varphi_2)\lambda_1}{(1-\varphi_1)} + \frac{(1-\varphi_2)}{(1-\varphi_1)} b_2\right) \end{array} \right] \end{array} \right) - \left(\begin{array}{c} \frac{(1-\varphi_2)}{[(1-\varphi_1)\lambda_2-(1-\varphi_2)\lambda_1]} \ln a_2 \\ - \left[\begin{array}{c} \left(\frac{1}{a_2} + \frac{(1-\varphi_2)}{[(1-\varphi_1)\lambda_2-(1-\varphi_2)\lambda_1]}\right) \\ ln\left(\frac{(1-\varphi_1)\lambda_2-(1-\varphi_2)\lambda_1}{(1-\varphi_1)} + \frac{(1-\varphi_2)}{(1-\varphi_1)} a_2\right) \end{array} \right] \end{array} \right) \right] \quad (20)$$

IV- Summary and Conclusion

Distributions are used to represent set(s) of data in statistical analysis. The composing of some distributions with each other's in some way to generate new distributions more flexible than the others to model real data . In this paper, we derived Properties of DTGGO and DTMOEU distributions, since doubly truncated distributions are more realistic to represent phenomena. We provided forms of rth raw moment, reliability function, hazard rate function, Shannon entropy function and Relative entropy function. This paper deals also with the determination of stress-strength R=p[y<x] when x (strength) and y (stress) are two independent DTGGO (DTMOEU) distribution with different parameters.

References

[1] Abid, S. and Hassan , H. (2015) " The Marshall-Olkin Extended Uniform Stress-Strength Model", American Journal of Mathematics and Statistics, 5(1): 1-10.

[2] Abid, S. and Hassan , H. (2015) " Some Additive Failure Rate Models Related with MOEU Distribution ", American Journal of Systems Science, 4(1): 1-10.

[3] Abid, S. and Hassan , H. (2015) " The Beta Marshall-Olkin Extended Uniform Distribution ", Journal of Safety Engineering, 4(1): 1-7.

[4] Boshi, M. & Abid, S. and Al-Noor, N. (2019) "generalized Gompertz - generalized Gompertz distribution", IOP Conf. Series: Journal of Physics: Conf. Series 1234 (2019) 012112, 1-12.

[5] De Andrade, T. & Chakraborty, S. & Handique, L. and Silva, F. (2019) " The exponentiated generalized extended Gompertz distribution" , Journal of Data Science,17(2). P. 299 – 330.

[6] El-Gohary, A. & Alshamrani, A. and Adel Al-Otaibi, N. (2013) "The generalized Gompertz distribution", Applied Mathematical Modelling 37 , 13–24.

[7] Jose,K. and Krishna, E. (2011) "Marshall-Olkin Extended Uniform Distribution", ProbStat Forum, Volume 04, October, Pages 78-88.

[8] Karamikabir, H. & Afshari, M. & Alizadeh, M. and Hamedani, G. (2019) "A new extended generalized Gompertz distribution with statistical properties and simulations", COMM. IN STAT.—THEORY AND METHODS, https://doi.org/ 10.1080/036 10926. 2019.1634209.

[9] Khan, M. & King, R. and Hudson, I. (2017) " Transmuted Generalized Gompertz distribution with application", Journal of Statistical Theory and Applications, Vol. 16, No. 1 (March) 65–80.

[10] Mazucheli, J. & Menezes, A. and Dey, S. (2019) "Unit- Gompertz distribution with applications" , STATISTICA, anno LXXIX, n. 1, 25-43.

Rayleigh Rayleigh Distribution: Properties and Applications

N H Al-Noor and N K Assi

Mathematics Department, College of Science, Mustansiriyah University, Baghdad, Iraq
E-mail: drnadialnoor@gmail.com, nadialnoor@uomustansiriyah.edu.iq

Abstract: In this paper, a new compound distribution named Rayleigh Rayleigh (RaRa) is presented. Several structural statistical properties of new distribution containing explicit expressions for the r-th moments, characteristic function, quantile function, order statistics, Shannon and relative entropies, besides stress strength reliability were considered and studied. The unknown parameters of RaRa distribution have been estimated under the maximum likelihood estimation method. Moreover, the RaRa distribution is applied upon a simulation study and real data set in order to evaluate its utility and flexibility.
Keywords. Rayleigh distribution, Compound distribution, Statistical properties, Shannon and relative entropies, Stress strength.

6. Introduction

Practically, it is observed that most common distributions are not sufficiently flexible to accommodate various phenomena of nature. For this purpose, researchers have focused on the expansion of these distributions in order to create a more and more realistic and flexible model for modeling data. The generalized families have been established by Mudholkar et al. [1], Marshall and Olkin [2], and Gupta et al. [3]. Newly, a number of new-generation families can be found in, for example, Ristic and Balakrishnan [4], Alzaatreh et al. [5], Nadarajah et al. [6], Tahir et al. [7], Ahmad et al. [8] and Al-Noor et al. [9].

Among probability distributions, the Rayleigh model is one of the most commonly used distributions. The Rayleigh distribution introduced by Rayleigh and it has appeared as a special case of the Weibull distribution. It plays a key role in modeling and analyzing life-time data such as project effort loading modeling, survival and reliability analysis, theory of communication, physical sciences, technology, diagnostic imaging, applied statistics and clinical research. With regard to this importance and the desire to give greater flexibility to this distribution, several researchers have developed extensive extensions to Rayleigh distribution for example, among others, Kundu and Raqab [10], Voda [11], Merovci [12][13], Merovci and Elbatal [14], Ateeq et al. [15]. In this paper, a new generation family of continuous distributions based on Rayleigh distribution is presented and suggested. This family is built on inspired by the composing two cumulative distribution functions together, say H and G, as follows.

Assume that $G(x)$ and $g(x)$ be any continuous baseline cumulative distribution and probability density functions (cdf and pdf) of a random variable X. Also assume respectively that $H(.)$ and $h(.)$ be the cdf and pdf of any $[0,\infty)$ continuous distribution. The general formula of reliability function for this class named H - G is given by

$$R(x)_{H-G} = \int_0^{-\ln G(x)} h(x)\,dx = H(-\ln G(x)) \qquad (1)$$

The general formulas of cdf and its associated pdf are

$$F(x)_{H-G} = 1 - R(x)_{H-G} = 1 - H\big(-\ln G\,(x)\big) \qquad (2)$$

$$f(x)_{H-G} = \frac{d}{dx}[F(x)_{H-G}] = \frac{g(x)}{G(x)}\, h(-\ln G(x)) \qquad (3)$$

Based on the above general formulas, a new family named Rayleigh–G distributions along with one of its special cases "sub-model" named Rayleigh– Rayleigh distribution are proposed and offered. The remains of this article are established as follows: In section 2, a brief detail about the Rayleigh – G distributions is provided. Sections 3, 4 and 5 respectively address the new Rayleigh – Rayleigh distribution with its structural statistical properties and the maximum likelihood estimation of its parameters. Sections 6 address the numerical illustration via a simulation study and a real data set application. Finally, the conclusions are presented in section 7.

7. Rayleigh – G distributions

Let $H(.)$ and $h(.)$ that mentioned in $(1),(2)$ and (3) be the cdf and pdf of Rayleigh distribution [14] with positive scale parameter θ as

$$H(-\ln G(x);\theta) = 1 - e^{-\frac{\theta}{2}(-\ln G(x))^2} \qquad (4)$$

$$h(-\ln G(x);\theta) = \theta(-\ln G(x))e^{-\frac{\theta}{2}(-\ln G(x))^2} \qquad (5)$$

By substituting (4) and (5) in (2) and (3), the cdf and pdf of new family named Rayleigh – G (for short $Ra-G$) distributions will be

$$F(x)_{Ra-G} = e^{-\frac{\theta}{2}(-\ln G(x))^2} \qquad (6)$$

$$f(x)_{Ra-G} = \frac{g(x)}{G(x)}\theta(-\ln G(x))\, e^{-\frac{\theta}{2}(-\ln G(x))^2} \qquad (7)$$

The pdf in (7) can be rewritten as an expanded formula that is relevant for obtaining the moments as one of the basic statistical properties when dealing with some particular cases. Thus, via using $e^{-z} = \Sigma_{i=0}^{\infty}\frac{(-1)^i}{i!}\, z^i$ we get

$$f(x)_{Ra-G}^{E} = \frac{g(x)}{G(x)}\theta(-\ln G(x))\Sigma_{i=0}^{\infty}\frac{(-1)^i}{i!}\left[\frac{\theta}{2}(-\ln G(x))^2\right]^i$$

and then

$$f(x)_{Ra-G}^{E} = \Sigma_{i=0}^{\infty}\frac{(-1)^i\,\theta^{i+1}}{i!\,2^i}\,\frac{g(x)}{G(x)}[-\ln G(x)]^{2i+1}$$

For $i \geq 1$, using $[-\ln z]^a = \sum_{k,l=0}^{\infty}\sum_{j=0}^{k}\frac{(-1)^{j+k+l}\,a}{a-j}\, C_k^{k-a}C_j^kC_l^{a+k}\,P_{j,k}\,z^l$ we have

$$[-\ln G(x)]^{2i+1} = \sum_{k,l=0}^{\infty}\sum_{j=0}^{k}\frac{(-1)^{j+k+l}(2i+1)}{2i+1-j}\, C_k^{k-2i-1}C_j^kC_l^{2i+k+1}P_{j,k}\,[G(x)]^l$$

Now the above formula of $f(x)_{Ra-G}^{E}$ will be

$$f(x)_{Ra-G}^{E} = \sum_{i=0}^{\infty} \frac{(-1)^i\, \theta^{i+1}}{i!\; 2^i} \frac{g(x)}{G(x)} \sum_{k,l=0}^{\infty} \sum_{j=0}^{k} \frac{(-1)^{j+k+l}(2i+1)}{2i+1-j} C_k^{k-2i-1} C_j^{k} C_l^{2i+k+1} P_{j,k}\, [G(x)]^l$$

Then the expansion formula $f(x)_{Ra-G}^{E}$ for the pdf $f(x)_{Ra-G}$ in (7) will be

$$f(x)_{Ra-G}^{E} = \sum_{i,k,l=0}^{\infty} \sum_{j=0}^{k} \frac{(-1)^{i+j+k+l}\,(2i+1)}{i!\,2^i\,(2i+1-j)}\, C_k^{k-2i-1} C_j^{k} C_l^{2i+k+1}\, P_{j,k}\, \theta^{i+1} g(x)[G(x)]^{l-1} \qquad (8)$$

where $P_{j,0} = 1$ for $j \geq 0$ and $P_{j,k} = k^{-1} \sum_{m=1}^{k} \frac{(-1)^m[m(j+1)-k]}{m+1} P_{j,k-m}$ for $k = 1,2, \dots$.

8. Rayleigh – Rayleigh distribution

Suppose that $G(x)$ and $g(x)$ in $(6),(7)$ and (8) be the cdf and pdf of Ra with positive parameter λ given respectively [14] as

$$G(x;\lambda) = 1 - e^{-\frac{\lambda}{2}x^2} \qquad (9)$$

$$g(x;\lambda) = \lambda x\, e^{-\frac{\lambda}{2}x^2} \qquad (10)$$

The cdf and pdf of new distribution named Rayleigh – Rayleigh (for short $Ra - Ra$) can be found according to (6) and (7) as

$$F(x)_{Ra-Ra} = e^{-\frac{\theta}{2}\left(-\ln\left(1-e^{-\frac{\lambda}{2}x^2}\right)\right)^2} \qquad (11)$$

$$f(x)_{Ra-Ra} = \frac{\theta\lambda x\, e^{-\frac{\lambda}{2}x^2}}{1-e^{-\frac{\lambda}{2}x^2}} \left(-\ln\left(1-e^{-\frac{\lambda}{2}x^2}\right)\right) e^{-\frac{\theta}{2}\left(-\ln\left(1-e^{-\frac{\lambda}{2}x^2}\right)\right)^2} \qquad (12)$$

The expansion formula of $Ra - Ra$ density function can be found according to (8) as

$$f(x)_{Ra-Ra}^{E} = \sum_{i,k,l=0}^{\infty} \sum_{j=0}^{k} \frac{(-1)^{i+j+k+l}\,(2i+1)}{i!\,2^i(2i+1-j)} C_k^{k-2i-1} C_j^{k} C_l^{2i+k+1} P_{j,k}\; \theta^{i+1} \lambda\, x\, e^{-\frac{\lambda}{2}x^2} \left(1-e^{-\frac{\lambda}{2}x^2}\right)^{l-1} \qquad (13)$$

The $Ra - Ra$ reliability and hazard rate functions can be found as

$$R(x)_{Ra-Ra} = 1 - F(x)_{Ra-Ra} = 1 - e^{-\frac{\theta}{2}\left(-\ln\left(1-e^{-\frac{\lambda}{2}x^2}\right)\right)^2} \qquad (14)$$

$$D(x)_{Ra-Ra} = \frac{f(x)_{Ra-Ra}}{R(x)_{Ra-Ra}} = \frac{\theta\lambda x\, e^{-\frac{\lambda}{2}x^2} \left(-\ln\left(1-e^{-\frac{\lambda}{2}x^2}\right)\right) e^{-\frac{\theta}{2}\left(-\ln\left(1-e^{-\frac{\lambda}{2}x^2}\right)\right)^2}}{\left(1-e^{-\frac{\lambda}{2}x^2}\right)\left(1-e^{-\frac{\theta}{2}\left(-\ln\left(1-e^{-\frac{\lambda}{2}x^2}\right)\right)^2}\right)} \qquad (15)$$

Figures 1 - 2 show the shapes of the cdf and pdf of $Ra - Ra$ distribution under a variety of selected values of parameters.

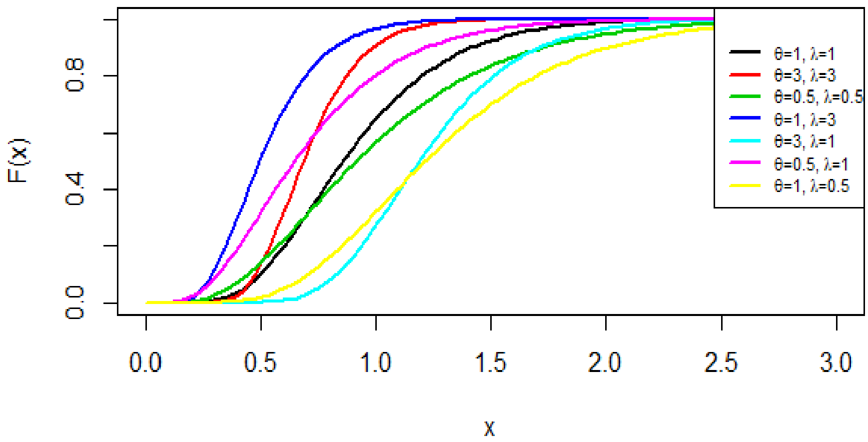


Figure 1. plot of $Ra - Ra$ cdf under a variety values of parameters

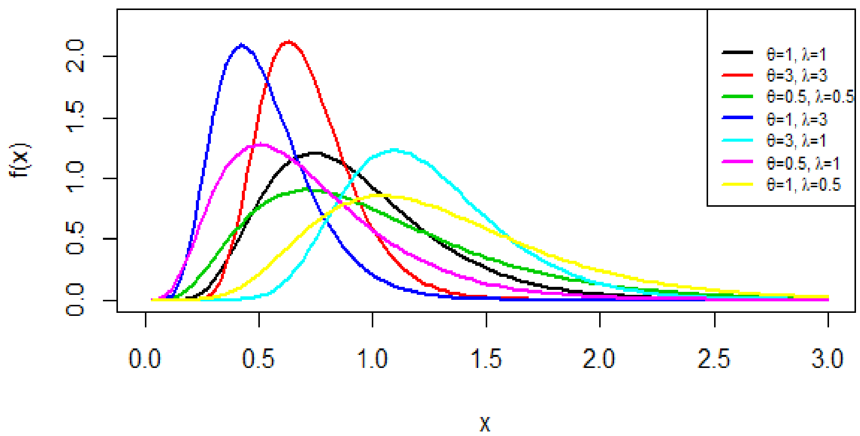


Figure 2. plot of $Ra - Ra$ pdf under a variety values of parameters

9. Statistical properties of the $\boldsymbol{Ra - Ra}$ distribution

In this section, the most necessary statistical properties of the $Ra - Ra$ distribution are given respectively as

4.1 The r-th moment: The r-th moment of $Ra - Ra$ distribution can be obtained as follows

Firstly recall $f(x)^{E}_{Ra-Ra}$ in (13) where $\left(1 - e^{-\frac{\lambda}{2}x^2}\right)^{l-1}$ have two cases $l - 1 > 0$ and $l - 1 < 0$ via using

$$(1 - z)^{b} = \Sigma_{i=0}^{\infty} \frac{(-1)^{i}}{i!} \frac{\Gamma(b+1)}{\Gamma(b-i+1)} z^{i} \quad ; |z| < 1, b > 0$$

$$(1 - z)^{-b} = \Sigma_{i=0}^{\infty} \frac{\Gamma(b+i)}{i!\ \Gamma(b)} z^{i} \quad ; |z| < 1, b > 0$$

where

$$\left(1-e^{-\frac{\lambda}{2}x^2}\right)^{l-1} = \begin{cases} \Sigma_{m=0}^{\infty} \dfrac{(-1)^m}{m!} \dfrac{\Gamma(l)}{\Gamma(l-m)} e^{-\frac{\lambda m}{2}x^2} & for \;\; l-1>0 \\ \Sigma_{m=0}^{\infty} \dfrac{\Gamma(l-1+m)}{m!\;\Gamma(l-1)} e^{-\frac{\lambda m}{2}x^2} & for \;\; l-1<0 \end{cases}$$

Now $f(x)^E_{Ra-Ra}$ for $l-1>0$ will be

$$f(x)^E_{Ra-Ra} = \sum_{i,k,l=0}^{\infty}\sum_{j=0}^{k} \frac{(-1)^{i+j+k+l}\,(2i+1)}{i!\,2^i(2i+1-j)} C_k^{k-2i-1} C_j^k C_l^{2i+k+1} P_{j,k}$$
$$\theta^{i+1}\lambda\, x\, e^{-\frac{\lambda}{2}x^2} \Sigma_{m=0}^{\infty} \frac{(-1)^m}{m!} \frac{\Gamma(l)}{\Gamma(l-m)} e^{-\frac{\lambda m}{2}x^2}$$
$$= \sum_{i,k,l,m=0}^{\infty}\sum_{j=0}^{k} \frac{(-1)^{i+j+k+l+m}\,(2i+1)}{i!\;m!\;2^i(2i+1-j)} \frac{\Gamma(l)}{\Gamma(l-m)} C_k^{k-2i-1} C_j^k C_l^{2i+k+1} P_{j,k}$$
$$\theta^{i+1}\lambda\, x\, e^{-\frac{\lambda(m+1)}{2}x^2}$$

and $f(x)^E_{Ra-Ra}$ for $l-1<0$ will be

$$f(x)^E_{Ra-Ra} = \sum_{i,k,l=0}^{\infty}\sum_{j=0}^{k} \frac{(-1)^{i+j+k+l}\,(2i+1)}{i!\,2^i(2i+1-j)} C_k^{k-2i-1} C_j^k C_l^{2i+k+1} P_{j,k}$$
$$\theta^{i+1}\lambda\, x\, e^{-\frac{\lambda}{2}x^2} \Sigma_{m=0}^{\infty} \frac{\Gamma(l-1+m)}{m!\;\Gamma(l-1)} e^{-\frac{\lambda m}{2}x^2}$$
$$= \sum_{i,k,l,m=0}^{\infty}\sum_{j=0}^{k} \frac{(-1)^{i+j+k+l}\,(2i+1)}{i!\;m!\,2^i(2i+1-j)} \frac{\Gamma(l-1+m)}{\Gamma(l-1)} C_k^{k-2i-1} C_j^k C_l^{2i+k+1} P_{j,k}$$
$$\theta^{i+1}\lambda\, x\, e^{-\frac{\lambda(m+1)}{2}x^2}$$

The $f(x)^E_{Ra-Ra}$ can be rewritten as

$$f(x)^E_{Ra-Ra} = W\lambda\, x\, e^{-\frac{\lambda(m+1)}{2}x^2} \tag{16}$$

where

$$W = \begin{cases} \displaystyle\sum_{i,k,l,m=0}^{\infty}\sum_{j=0}^{k} \frac{(-1)^{i+j+k+l+m}\,(2i+1)}{i!\;m!\;2^i(2i+1-j)} \frac{\Gamma(l)}{\Gamma(l-m)} C_k^{k-2i-1} C_j^k C_l^{2i+k+1} P_{j,k}\, \theta^{i+1} & ;l-1>0 \\ \displaystyle\sum_{i,k,l,m=0}^{\infty}\sum_{j=0}^{k} \frac{(-1)^{i+j+k+l}\,(2i+1)}{i!\;m!\,2^i(2i+1-j)} \frac{\Gamma(l-1+m)}{\Gamma(l-1)} C_k^{k-2i-1} C_j^k C_l^{2i+k+1} P_{j,k}\, \theta^{i+1} & ;l-1<0 \end{cases} \tag{17}$$

Now the r-th moment of $Ra-Ra$ distribution can be obtained as

$$E(X^r)_{Ra-Ra} = \int_0^\infty x^r f(x)^E_{Ra-Ra}\, dx = \int_0^\infty x^r W\lambda\, x\, e^{-\frac{\lambda(m+1)}{2}x^2} dx$$

$$= W \frac{1}{m+1} \int_0^\infty x^r \lambda(m+1)\, x\, e^{-\frac{\lambda(m+1)}{2}x^2} dx$$

where $\int_0^\infty x^r \lambda(m+1)\, x\, e^{-\frac{\lambda(m+1)}{2}x^2} dx$ represents the r-th moment of Ra distribution with parameter $\lambda(m+1)$ i.e. $\int_0^\infty x^r \lambda(m+1) x\, e^{-\frac{\lambda(m+1)}{2}x^2} dx = \left(\frac{2}{\lambda(m+1)}\right)^{\frac{r}{2}} \Gamma\left(1+\frac{r}{2}\right)$.

Thus the $E(X^r)_{Ra-Ra}$ will be

$$E(X^r)_{Ra-Ra} = W \frac{1}{(m+1)^{\frac{r}{2}+1}} \left(\frac{2}{\lambda}\right)^{\frac{r}{2}} \Gamma\left(1+\frac{r}{2}\right) \qquad (18)$$

where W as in (17) and $\Gamma(.)$ is the gamma function.

Additional properties of $Ra-Ra$ distribution for instance the mean, variance, coefficients of kurtosis and skewness can be found with specific value of $r = 1,2,3,4$.

4.2 The characteristic function: The characteristic function of $Ra-Ra$ distribution can be found by

$$Q_X(t)_{Ra-Ra} = \sum_{r=0}^{\infty} \frac{(it)^r}{r!} E(X^r)_{Ra-Ra} = W \sum_{r=0}^{\infty} \frac{(it)^r}{r!\,(m+1)^{\frac{r}{2}+1}} \left(\frac{2}{\lambda}\right)^{\frac{r}{2}} \Gamma\left(1+\frac{r}{2}\right) \qquad (19)$$

where W as in (17) and $\Gamma(.)$ is the gamma function.

4.3 The quantile function: The quantile function of $Ra-Ra$ random variable is defined as a solution of $P(x \le x_{(q)}) = F(x_{(q)})_{Ra-Ra}$ w.r.t. $x_{(q)}$; $x_{(q)} > 0$ and $0 < q < 1$. Therefore it can be found via using (11) as

$$x_{(q)Ra-Ra} = \left[-\frac{2}{\lambda}\ln\left(1 - e^{-\left(-\frac{2}{\theta}\ln(q)\right)^{\frac{1}{2}}}\right)\right]^{\frac{1}{2}} \qquad (20)$$

The median of $Ra-Ra$ random variable can be found via setting $q = 0.5$. A random variable X has the $Ra-Ra$ distribution can be simulated by

$$x = \left[-\frac{2}{\lambda}\ln\left(1 - e^{-\left(-\frac{2}{\theta}\ln(u)\right)^{\frac{1}{2}}}\right)\right]^{\frac{1}{2}} \qquad (21)$$

where U has the standard uniform distribution.

4.4 Order statistics: Consider $x_1, x_2, \dots, x_n$ as a random sample of size n taken independently from $Ra-Ra$ distribution. Let $x_{1:n}, x_{2:n}, \dots, x_{n:n}$ be the corresponding order statistics. Then the pdf of $x_{k:n}$, $k \le n$ can be found via the following standard formula statistics

$$f_{k:n}(x) = \frac{n!}{(k-1)!\,(n-k)!} \left(F(x)\right)^{k-1} \left(1-F(x)\right)^{n-k} f(x) \quad ; 0 \le x_{k:n} < \infty$$

and then based on (11) and (12) we get

$$f_{k:n}(x) = \frac{n!}{(k-1)!\,(n-k)!} \frac{\theta\lambda x\, e^{-\frac{\lambda}{2}x^2}}{1-e^{-\frac{\lambda}{2}x^2}} e^{-\frac{\theta k}{2}\left(-\ln\left(1-e^{-\frac{\lambda}{2}x^2}\right)\right)^2}$$
$$\left(1-e^{-\frac{\theta}{2}\left(-\ln\left(1-e^{-\frac{\lambda}{2}x^2}\right)\right)^2}\right)^{n-k} \left(-\ln\left(1-e^{-\frac{\lambda}{2}x^2}\right)\right) \qquad (22)$$

The joint pdf of order statistics can be found via the following standard formula statistics

$$f_{j,k:n}(x,y) = \frac{n!}{(j-1)!\,(k-j-1)!\,(n-k)!} \left(F(x)\right)^{j-1}\left(F(y)-F(x)\right)^{k-j-1}$$
$$\left(1-F(y)\right)^{n-k} f(x)\, f(y) \quad ;\, 1 \le j \le k\,, 0 \le x \le y < \infty$$

and then

$$f_{j,k:n}(x,y) = \frac{n!}{(j-1)!\,(k-j-1)!\,(n-k)!} \frac{\theta^2\lambda^2 xy\, e^{-\frac{\lambda}{2}(x^2+y^2)}}{\left(1-e^{-\frac{\lambda}{2}x^2}\right)\left(1-e^{-\frac{\lambda}{2}y^2}\right)}$$
$$\left(e^{-\frac{\theta}{2}\left(-\ln\left(1-e^{-\frac{\lambda}{2}y^2}\right)\right)^2} - e^{-\frac{\theta}{2}\left(-\ln\left(1-e^{-\frac{\lambda}{2}x^2}\right)\right)^2}\right)^{k-j-1}$$
$$e^{-\frac{\theta j}{2}\left(-\ln\left(1-e^{-\frac{\lambda}{2}x^2}\right)\right)^2} \left(1-e^{-\frac{\theta}{2}\left(-\ln\left(1-e^{-\frac{\lambda}{2}y^2}\right)\right)^2}\right)^{n-k}$$
$$\left(-\ln\left(1-e^{-\frac{\lambda}{2}x^2}\right)\right)\left(-\ln\left(1-e^{-\frac{\lambda}{2}y^2}\right)\right) e^{-\frac{\theta}{2}\left(-\ln\left(1-e^{-\frac{\lambda}{2}y^2}\right)\right)^2} \qquad (23)$$

4.5 Shannon Entropy: The $Ra-Ra$ Shannon entropy can be found via $SH_{Ra-Ra} = -\int_0^\infty \ln(f(x)_{Ra-Ra})\, f(x)_{Ra-Ra}\, dx$. Taking the natural logarithm of (12) we get

$$\ln(f(x)_{Ra-Ra}) = \ln(\theta\lambda) + \ln(x) - \frac{\lambda}{2}x^2 - \ln\left(1-e^{-\frac{\lambda}{2}x^2}\right) + \ln\left(-\ln\left(1-e^{-\frac{\lambda}{2}x^2}\right)\right)$$
$$-\frac{\theta}{2}\left(-\ln\left(1-e^{-\frac{\lambda}{2}x^2}\right)\right)^2$$

Now SH_{Ra-Ra} is given by

$$SH_{Ra-Ra} = -\left\{\begin{array}{l} \ln(\theta\lambda) + E(\ln(X)) - \frac{\lambda}{2}E(X^2) - E\left(\ln\left(1-e^{-\frac{\lambda}{2}X^2}\right)\right) \\ +E\left(\ln\left(-\ln\left(1-e^{-\frac{\lambda}{2}X^2}\right)\right)\right) - \frac{\theta}{2}E\left(\left(-\ln\left(1-e^{-\frac{\lambda}{2}X^2}\right)\right)^2\right) \end{array}\right\} \qquad (24)$$

Now for $E(\ln(X))$ recall (16) we get

$$E(\ln(X)) = \int_0^\infty \ln(x)\, f(x)^E_{Ra-Ra}\, dx = \int_0^\infty \ln(x)\, W\lambda\, x\, e^{-\frac{\lambda(m+1)}{2}x^2}\, dx$$

Let $y=\frac{\lambda(m+1)}{2}\,x^2 \to x^2=\frac{2y}{\lambda(m+1)} \to x=\left(\frac{2y}{\lambda(m+1)}\right)^{\frac{1}{2}}$ and $dx=\left(\frac{1}{2\lambda(m+1)y}\right)^{\frac{1}{2}}dy$ then

$$E(\ln(X))=W\int_0^\infty \lambda\ln\left(\left(\frac{2y}{\lambda(m+1)}\right)^{\frac{1}{2}}\right)\left(\frac{2y}{\lambda(m+1)}\right)^{\frac{1}{2}} e^{-y}\left(\frac{1}{2\lambda(m+1)y}\right)^{\frac{1}{2}}dy$$

$$=W\int_0^\infty \left(\frac{1}{2}\ln\left(\frac{2y}{\lambda(m+1)}\right)\right)\left(\frac{1}{m+1}\right) e^{-y}\,dy$$

$$=W\frac{1}{2(m+1)}\int_0^\infty\left(\ln\left(\frac{2}{\lambda(m+1)}\right)+\ln(y)\right)e^{-y}\,dy$$

$$=W\frac{1}{2(m+1)}\left[\ln\left(\frac{2}{\lambda(m+1)}\right)\int_0^\infty e^{-y}\,dy+\int_0^\infty \ln(y)\,e^{-y}\,dy\right]$$

Since $\int_0^\infty e^{-y}\,dy=1$ and via using $\int_0^\infty z^{s-1}\ln(z)\,e^{-mz}\,dz=m^{-s}\,\Gamma(s)(\psi(s)-\ln(m))$ we get $\int_0^\infty \ln(y)\,e^{-y}\,dy=\psi(1)$ so

$$E(\ln(X))=W\frac{1}{2(m+1)}\left[\ln\left(\frac{2}{\lambda(m+1)}\right)+\psi(1)\right] \qquad (25)$$

For $E\left(\ln\left(1-e^{-\frac{\lambda}{2}x^2}\right)\right)$ using $\ln(1-z)=-\Sigma_{i=1}^{\infty}\frac{z^i}{i};\ |z|<1$ and e^{-z} we get

$$\ln\left(1-e^{-\frac{\lambda}{2}x^2}\right)=-\sum_{i=1}^{\infty}\frac{1}{i}\,e^{-\frac{\lambda i}{2}x^2}=-\sum_{i=1}^{\infty}\sum_{j=0}^{\infty}\frac{(-1)^j}{j!\;i}\left(\frac{\lambda i}{2}x^2\right)^j=\sum_{i=1}^{\infty}\sum_{j=0}^{\infty}\frac{(-1)^{j+1}}{j!\;i}\left(\frac{\lambda i}{2}\right)^j x^{2j}$$

Then

$$E\left(\ln\left(1-e^{-\frac{\lambda}{2}x^2}\right)\right)=\sum_{i=1}^{\infty}\sum_{j=0}^{\infty}\frac{(-1)^{j+1}}{j!\;i}\left(\frac{\lambda i}{2}\right)^j E(X^{2j}) \qquad (26)$$

where $E(X^{2j})$ as in (18) with $r=2j$.

For $E\left(\ln\left(-\ln\left(1-e^{-\frac{\lambda}{2}x^2}\right)\right)\right)$ using the above formula of $\ln(1-z)$, as well as using

$\ln z=\Sigma_{k=0}^{\infty}\frac{(-1)^k(z-1)^{k+1}}{k+1};\ 0<z\le 2$,

$(a+b)^n=\Sigma_{k=0}^{\infty}C_k^n\,a^{n-k}\,b^k=\Sigma_{k=0}^{\infty}C_k^n\,a^k\,b^{n-k}$; $n\ge 0$, $C_k^n=\frac{n!}{k!\,(n-k)!}$ is Binomial coefficients, and e^{-z} we get

$$\ln\left(-\ln\left(1-e^{-\frac{\lambda}{2}x^2}\right)\right)=\ln\left(\Sigma_{i=1}^{\infty}\frac{1}{i}e^{-\frac{\lambda i}{2}x^2}\right)$$

$$=\Sigma_{k=0}^{\infty}\frac{(-1)^k}{k+1}\left(\sum_{i=1}^{\infty}\frac{e^{-\frac{\lambda i}{2}x^2}}{i}-1\right)^{k+1}$$

$$= \Sigma_{k=0}^{\infty} \frac{(-1)^k}{k+1} \Sigma_{m=0}^{k+1} C_m^{k+1} \left(\Sigma_{i=1}^{\infty} \frac{e^{-\frac{\lambda i}{2} x^2}}{i} \right)^m (-1)^{k+1-m}$$

$$= \Sigma_{k=0}^{\infty} \Sigma_{m=0}^{k+1} \frac{(-1)^{2k+1-m}}{k+1} C_m^{k+1} \left(\Sigma_{i=1}^{\infty} \frac{e^{-\frac{\lambda i}{2} x^2}}{i} \right)^m$$

$$= \Sigma_{k=0}^{\infty} \Sigma_{m=0}^{k+1} \frac{(-1)^{2k+1-m}}{k+1} C_m^{k+1} \Sigma_{i_1=1}^{\infty} \Sigma_{i_2=1}^{\infty} \dots \Sigma_{i_m=1}^{\infty} \frac{e^{-\frac{\lambda}{2}(i_1+i_2+\dots+i_m) x^2}}{i_1 i_2 \dots i_m}$$

$$= \Sigma_{k=0}^{\infty} \Sigma_{m=0}^{k+1} \frac{(-1)^{2k+1-m}}{k+1} C_m^{k+1}$$
$$\Sigma_{i_1=1}^{\infty} \Sigma_{i_2=1}^{\infty} \dots \Sigma_{i_m=1}^{\infty} \Sigma_{j=0}^{\infty} \frac{(-1)^j}{i_1 i_2 \dots i_m \; j!} \left(\frac{\lambda}{2}(i_1 + i_2 + \dots + i_m) \right)^j x^{2j}$$

Then

$$E\left(\ln\left(-\ln\left(1 - e^{-\frac{\lambda}{2} x^2} \right) \right) \right) = \Sigma_{j,k=0}^{\infty} \Sigma_{m=0}^{k+1} \Sigma_{i_1=1}^{\infty} \Sigma_{i_2=1}^{\infty} \dots \Sigma_{i_m=1}^{\infty} \frac{(-1)^{j+2k+1-m}}{j! \, i_1 i_2 \dots i_m \, (k+1)} C_m^{k+1} \left(\frac{\lambda}{2}(i_1 + i_2 + \dots + i_m) \right)^j E(X^{2j}) \quad (27)$$

where $E(X^{2j})$ as in (18) with $r = 2j$.

For $E\left(\left(-\ln\left(1 - e^{-\frac{\lambda}{2} x^2} \right) \right)^2 \right)$ using $(-\ln(1-z))^a = a \, \Sigma_{k=0}^{\infty} C_k^{k-a} \Sigma_{j=0}^{k} \frac{(-1)^{j+k} C_j^k P_{j,k}}{a-j} z^{a+k}$ and e^{-z} we get

$$\left(-\ln\left(1 - e^{-\frac{\lambda}{2} x^2} \right) \right)^2 = 2 \, \Sigma_{k=0}^{\infty} C_k^{k-2} \Sigma_{j=0}^{k} \frac{(-1)^{j+k}}{2-j} C_j^k P_{j,k} \left(e^{-\frac{\lambda}{2} x^2} \right)^{2+k}$$

$$= 2 \, \Sigma_{k=0}^{\infty} \Sigma_{j=0}^{k} \frac{(-1)^{j+k}}{2-j} C_k^{k-2} C_j^k P_{j,k} \, e^{-\frac{\lambda(2+k)}{2} x^2}$$

$$= 2 \, \Sigma_{k=0}^{\infty} \Sigma_{j=0}^{k} \frac{(-1)^{j+k}}{2-j} C_k^{k-2} C_j^k P_{j,k} \, \Sigma_{i=0}^{\infty} \frac{(-1)^i}{i!} \left(\frac{\lambda(2+k)}{2} x^2 \right)^i$$

$$= 2 \, \Sigma_{i,k=0}^{\infty} \Sigma_{j=0}^{k} \frac{(-1)^{i+j+k}}{i! \, (2-j)} C_k^{k-2} C_j^k P_{j,k} \left(\frac{\lambda(2+k)}{2} \right)^i x^{2i}$$

Then

$$E\left(\left(-\ln\left(1 - e^{-\frac{\lambda}{2} x^2} \right) \right)^2 \right) = \Sigma_{i,k=0}^{\infty} \Sigma_{j=0}^{k} \frac{(-1)^{i+j+k}}{i! \, (2-j) 2^{i-1}} C_k^{k-2} C_j^k P_{j,k} \left(\lambda(2+k) \right)^i E(X^{2i}) \quad (28)$$

where $E(X^{2i})$ as in (18) with $r = 2i$ and $P_{j,k} = k^{-1} \Sigma_{m=1}^{k} [k - m(j+1)] \, C_m \, P_{j,k-m}$ for $k = 1,2, \dots$ with $P_{j,0} = 1$ and $C_k = (-1)^{k+1} (k+1)^{-1}$.

Now from (24) the $Ra - Ra$ Shannon entropy can be found by

$$SH_{Ra-Ra} = -\ln(\theta\lambda) - E(\ln(X)) + \frac{\lambda}{2}E(X^2) + E\left(\ln\left(1-e^{-\frac{\lambda}{2}X^2}\right)\right)$$
$$-E\left(\ln\left(-\ln\left(1-e^{-\frac{\lambda}{2}X^2}\right)\right)\right) + \frac{\theta}{2}E\left(\left(-\ln\left(1-e^{-\frac{\lambda}{2}X^2}\right)\right)^2\right) \qquad (29)$$

where

$E(\ln(X))$, $E\left(\ln\left(1-e^{-\frac{\lambda}{2}X^2}\right)\right)$, $E\left(\ln\left(-\ln\left(1-e^{-\frac{\lambda}{2}X^2}\right)\right)\right)$, $E\left(\left(-\ln\left(1-e^{-\frac{\lambda}{2}X^2}\right)\right)^2\right)$ as in (25), (26), (27), (28) respectively and $E(X^2)$ as in (18) with $r = 2$.

4.6 The Relative Entropy: The $Ra - Ra$ relative entropy can be found via $RE_{Ra-Ra} = \int_0^\infty \ln\left(\frac{f(x)_{Ra-Ra}}{f_1(x)_{Ra-Ra}}\right) f(x)_{Ra-Ra}\, dx$. Taking the natural logarithm of the $f(x)_{Ra-Ra}$ in (12) w.r.t. $f_1(x)_{Ra-Ra}$ with parameters (θ_1, λ_1) we get

$$\ln\left(\frac{f(x)_{Ra-Ra}}{f_1(x)_{Ra-Ra}}\right) = \ln(\theta\lambda) - \ln(\theta_1\lambda_1) + \ln(x) - \ln(x) - \frac{\lambda}{2}x^2 + \frac{\lambda_1}{2}x^2 - \ln\left(1-e^{-\frac{\lambda}{2}x^2}\right)$$
$$+\ln\left(1-e^{-\frac{\lambda_1}{2}x^2}\right) + \ln\left(\left(-\ln\left(1-e^{-\frac{\lambda}{2}x^2}\right)\right)\right) - \ln\left(\left(-\ln\left(1-e^{-\frac{\lambda_1}{2}x^2}\right)\right)\right)$$
$$-\frac{\theta}{2}\left(-\ln\left(1-e^{-\frac{\lambda}{2}x^2}\right)\right)^2 + \frac{\theta_1}{2}\left(-\ln\left(1-e^{-\frac{\lambda_1}{2}x^2}\right)\right)^2$$

Then $Ra - Ra$ relative entropy will be

$$RE_{Ra-Ra} = \ln\left(\frac{\theta\lambda}{\theta_1\lambda_1}\right) - \left(\frac{\lambda-\lambda_1}{2}\right)E(X^2) - E\left(\ln\left(1-e^{-\frac{\lambda}{2}X^2}\right)\right)$$
$$+E\left(\ln\left(1-e^{-\frac{\lambda_1}{2}X^2}\right)\right) + E\left(\ln\left(\left(-\ln\left(1-e^{-\frac{\lambda}{2}X^2}\right)\right)\right)\right) - E\left(\ln\left(\left(-\ln\left(1-e^{-\frac{\lambda_1}{2}X^2}\right)\right)\right)\right)$$
$$-\frac{\theta}{2}E\left(\left(-\ln\left(1-e^{-\frac{\lambda}{2}X^2}\right)\right)^2\right) + \frac{\theta_1}{2}E\left(\left(-\ln\left(1-e^{-\frac{\lambda_1}{2}X^2}\right)\right)^2\right) \qquad (30)$$

where
$E(X^2)$ as in (18) with $r = 2$ and the extra expectations respectively as in (25), (26), (27), (28) with indicated parameters.

4.7 The stress strength: Let Y be the stress and X be the strength of $Ra - Ra$ independent random variables with different parameters, then the stress strength $SS_{Ra-Ra} = P(Y < X)$ can be found by

$SS_{Ra-Ra} = \int_0^\infty f_X(x)_{Ra-Ra} F_Y(x)_{Ra-Ra}\, dx$ where

$F_Y(x)_{Ra-Ra} = e^{-\frac{\theta_1}{2}\left(-\ln\left(1-e^{-\frac{\lambda_1}{2}x^2}\right)\right)^2}$. Using e^{-z} , $F_Y(x)_{Ra-Ra}$ can rewritten as

$$F_Y(x)_{Ra-Ra} = \sum_{i=0}^{\infty} \frac{(-1)^i}{i!}\left(\frac{\theta_1}{2}\left(-\ln\left(1-e^{-\frac{\lambda_1}{2}x^2}\right)\right)^2\right)^i$$
$$= \sum_{i=0}^{\infty} \frac{(-1)^i}{i!}\left(\frac{\theta_1}{2}\right)^i\left(-\ln\left(1-e^{-\frac{\lambda_1}{2}x^2}\right)\right)^{2i}$$

For $\left(-\ln\left(1-e^{-\frac{\lambda_1}{2}x^2}\right)\right)^{2i}$ using $(-\ln(1-z))^a$, e^{-z} and follows the previous similar steps of getting $\left(-\ln\left(1-e^{-\frac{\lambda}{2}x^2}\right)\right)^2$ we get

$$\left(-\ln\left(1-e^{-\frac{\lambda_1}{2}x^2}\right)\right)^{2i} = 2i\ \Sigma_{k=0}^{\infty} C_k^{k-2i} \Sigma_{j=0}^{k} \frac{(-1)^{j+k}}{2i-j}\ C_j^k\ P_{j,k}\left(e^{-\frac{\lambda_1}{2}x^2}\right)^{2i+k}$$
$$= 2i\ \Sigma_{k=0}^{\infty}\Sigma_{j=0}^{k} \frac{(-1)^{j+k}}{2i-j}\ C_k^{k-2i} C_j^k\ P_{j,k}\ e^{-\frac{\lambda_1(2i+k)}{2}x^2}$$
$$= 2i\ \Sigma_{k=0}^{\infty}\Sigma_{j=0}^{k} \frac{(-1)^{j+k}}{2i-j} C_k^{k-2i} C_j^k\ P_{j,k}\ \Sigma_{m=0}^{\infty} \frac{(-1)^m}{m!}\left(\frac{\lambda_1(2i+k)}{2}\ x^2\right)^m$$
$$= 2i\ \Sigma_{k,m=0}^{\infty}\Sigma_{j=0}^{k} \frac{(-1)^{j+k+m}}{m!\,(2i-j)} C_k^{k-2i} C_j^k\ P_{j,k}\left(\frac{\lambda_1(2i+k)}{2}\right)^m\ x^{2m}$$

Then

$$F_Y(x)_{Ra-Ra} = \Sigma_{i,k,m=0}^{\infty}\Sigma_{j=0}^{k} \frac{(-1)^{i+j+k+m}\ 2i}{i!\ m!\,(2i-j)} C_k^{k-2i} C_j^k\ P_{j,k}\left(\frac{\theta_1}{2}\right)^i\left(\frac{\lambda_1(2i+k)}{2}\right)^m\ x^{2m} \qquad (31)$$

The $Ra-Ra$ stress strength can be found based on (31)as

$$SS_{Ra-Ra} = \Sigma_{i,k,m=0}^{\infty}\Sigma_{j=0}^{k} \frac{(-1)^{i+j+k+m}\ 2i}{i!\ m!\,(2i-j)} C_k^{k-2i} C_j^k\ P_{j,k}\left(\frac{\theta_1}{2}\right)^i\left(\frac{\lambda_1(2i+k)}{2}\right)^m E(\,X^{2m}\,) \qquad (32)$$

where $E(\,X^{2m})$ as in (18) with $r=2m$ and $P_{j,k} = k^{-1}\ \Sigma_{m=1}^{k}\,[k-m(j+1)]\ C_m\ P_{j,k-m}$ for $k=1,2,\ldots$ with $P_{j,0}=1$ and $C_k = (-1)^{k+1}\,(k+1)^{-1}$.

5 Estimation of Ra-Ra parameters

The method of maximum likelihood estimation is considered here to estimate the parameters of $Ra-Ra$ distribution with complete sample. Consider a complete $Ra-Ra$ random sample, say $x_1, x_2, \ldots, x_n$, with parameter vector $\zeta = (\theta,\lambda)^T$. The natural logarithm likelihood function $\ell(\zeta\,|\underline{x})$ in relation to (12) is

$$\ell(\zeta|\underline{x}) = n\ln(\theta) + n\ln(\lambda) + \Sigma_{i=1}^{n}\ln(x_i) - \frac{\lambda}{2}\Sigma_{i=1}^{n}{x_i}^2 - \Sigma_{i=1}^{n}\ln\left(1-e^{-\frac{\lambda}{2}{x_i}^2}\right)$$
$$+\Sigma_{i=1}^{n}\ln\left(-\ln\left(1-e^{-\frac{\lambda}{2}{x_i}^2}\right)\right) - \frac{\theta}{2}\Sigma_{i=1}^{n}\left(-\ln\left(1-e^{-\frac{\lambda}{2}{x_i}^2}\right)\right)^2 \qquad (33)$$

The maximum likelihood estimates (MLEs) of two parameters (θ, λ) can be found via solving the nonlinear system of natural logarithm likelihood equations $\frac{\partial \ell(\zeta|\underline{x})}{\partial \zeta} = \left(\frac{\partial \ell(\zeta|\underline{x})}{\partial \theta}, \frac{\partial \ell(\zeta|\underline{x})}{\partial \lambda}\right)^T = 0$ through iterative numerical techniques.

6 Numerical illustrations

Numerical illustrations is presented here to exhibit the abilities of $Ra - Ra$ distribution via simulation study and application with real data set.

6.1 Simulation study

In this subsection a simulation study is carried out to exhibit the performances of the MLEs of $Ra - Ra$ distribution. The steps of process are as follows

1. Generate an i.i.d. random samples follow $Ra - Ra$ distribution. The number of replicated samples was made 1000 times each with sizes $n = 25{,}50{,}100$ and 200 .
2. Select the initial "true" values of parameters to be as in Tables (1-3) with $\theta > \lambda$, $\theta < \lambda$ and $\theta = \lambda$.
3. Calculate Bias and RMSE "root mean squared error" where
   $Bias(\hat{\theta}) = \frac{1}{1000}\sum_{i=1}^{1000}(\hat{\theta} - \theta)$ and $RMSE(\hat{\theta}) = \sqrt{\frac{1}{1000}\sum_{i=1}^{1000}(\hat{\theta} - \theta)^2}$
4. Repeat step 3 for other parameter λ .

The empirical results are shown in Tables (1-3). It clearly appears that RMSE values increase as the values of the parameters increasing and the RMSE values decrease as the sample size increases.

Table 1. The Bias and RMSE of the Ra-Ra parameters estimation using MLE with $\theta > \lambda$.

n	Par.	Initial	Bias	RMSE	Initial	Bias	RMSE
25	λ	1	0.1164	0.3277	0.5	0.0758	0.2054
	θ	3	1.1336	2.9564	1	0.2784	0.7177
50	λ	1	0.0586	0.2066	0.5	0.0379	0.1282
	θ	3	0.4818	1.4252	1	0.1218	0.3766
100	λ	1	0.0288	0.1341	0.5	0.0182	0.0817
	θ	3	0.2336	0.8214	1	0.0604	0.2261
200	λ	1	0.0133	0.0929	0.5	0.0084	0.0562
	θ	3	0.1040	0.5224	1	0.0266	0.1460

Table 2. The Bias and RMSE of the Ra-Ra parameters estimation using MLE with $\theta < \lambda$.

n	Par.	Initial	Bias	RMSE	Initial	Bias	RMSE
25	λ	3	0.4593	1.2391	1	0.1960	0.5064
	θ	1	0.2812	0.7206	0.5	0.1200	0.3059
50	λ	3	0.2300	0.7751	1	0.0958	0.3084
	θ	1	0.1243	0.3821	0.5	0.0530	0.1679
100	λ	3	0.1094	0.4899	1	0.0449	0.1924
	θ	1	0.0607	0.2265	0.5	0.0263	0.1016
200	λ	3	0.0504	0.3388	1	0.0209	0.1317
	θ	1	0.0268	0.1469	0.5	0.0116	0.0663

Table 3. The Bias and RMSE of the Ra-Ra parameters estimation using MLE with $\theta = \lambda$.

n	Par.	Initial	Bias	RMSE	Initial	Bias	RMSE	Initial	Bias	RMSE
25	λ	0.5	0.0975	0.2529	1	0.1559	0.4173	3	0.3479	0.9802
	θ	0.5	0.1193	0.3041	1	0.2875	0.7309	3	1.1461	2.9769
50	λ	0.5	0.0480	0.1546	1	0.0771	0.2594	3	0.1769	0.6250
	θ	0.5	0.0532	0.1684	1	0.1251	0.3843	3	0.4899	1.4447
100	λ	0.5	0.0224	0.0963	1	0.0362	0.1639	3	0.0847	0.3939
	θ	0.5	0.0263	0.1019	1	0.0603	0.2268	3	0.2268	0.8005
200	λ	0.5	0.0104	0.0662	1	0.0167	0.1128	3	0.0387	0.2779
	θ	0.5	0.0116	0.0666	1	0.0267	0.1467	3	0.1018	0.5208

6.2 Real data application

In this subsection, the application of real data set is analyzed to verify the flexibility of the proposed family. The $\mathrm{Ra}-\mathrm{Ra}$ distribution has been compared with four distributions "Gamma Rayleigh (GaRa), Marshal Olkin Rayleigh (MORa), Truncated-Exponential Skew Symmetric Rayleigh (TESRa), and Rayleigh (Ra) distributions. For more details about the compared distributions see [2][4][6]. The R software has been used to compute the analytical measures "negative log-likelihood (NLL), Akaike Information Criteria (AIC), Consistent Akaike Information Criteria (CAIC), Bayesian Information Criteria (BIC), Hanan and Quinn Information Criteria (HQIC)", and values of parameters estimation via MLE method.

2.1.1 The data set taken from Mathers et al. [16] representing the crude mortality rate (CMR) among people who inject drugs. The data set consist of the following observations
2.01, 6.32, 3.52, 2.15, 5.42, 2.04, 2.77, 2.26, 1.95, 1.00, 2.45, 0.74, 0.98, 1.27, 2.77, 3.68, 1.18, 1.09, 1.60, 0.57, 3.33, 0.91, 7.14, 2.08, 3.85, 1.99, 7.76, 2.52, 1.57, 4.67, 4.22, 1.92, 1.59, 4.08, 2.02, 0.84, 6.85, 2.18, 2.04, 1.05, 2.91, 1.37, 2.43, 2.28, 3.74, 1.30, 1.59, 1.83, 3.85, 6.30, 4.83, 0.50, 3.40, 2.33, 4.25, 3.49, 2.12, 0.83, 0.54, 3.23, 4.50, 0.71, 0.48, 2.30, 7.73.
The fitting results for each of the fitted distributions are shown in Table 4.

Table 4. Results of fitting real data set.

Distribution	NLL	AIC	CAIC	BIC	HQIC	MLE
RaRa	118.4921	240.9842	241.1777	245.333	242.7000	$\hat{\theta}$: 0.2526222 $\hat{\lambda}$: 0.0383351
GaRa	120.3703	244.7406	244.9342	249.0894	246.4565	$\hat{\theta}$: 2.2483838 $\hat{\lambda}$: 0.0532035
MORa	119.0476	242.0952	242.2888	246.444	243.8111	$\hat{\theta}$: 0.2742291 $\hat{\lambda}$: 0.0920414
TESRa	119.6947	243.3893	243.5829	247.7381	245.1052	$\hat{\theta}$: 2.5100785 $\hat{\lambda}$: 0.0985120
Ra	123.6520	249.3039	249.3674	251.4783	250.1618	$\hat{\lambda}$: 0.1864005

From Table 4 the newly proposed $Ra - Ra$ distribution displays a precise good representation as it has the lowest values for the analytical measures NLL, AIC, CAIC, BIC, and HQIC. Furthermore, Figures 3 and 4 present respectively the histogram plot of the data set with the other compared distributions and the corresponding empirical cdf plot. The fitted of $\mathrm{Ra}-\mathrm{Ra}$ density is closer to the empirical histogram than the fits of other compared distributions.

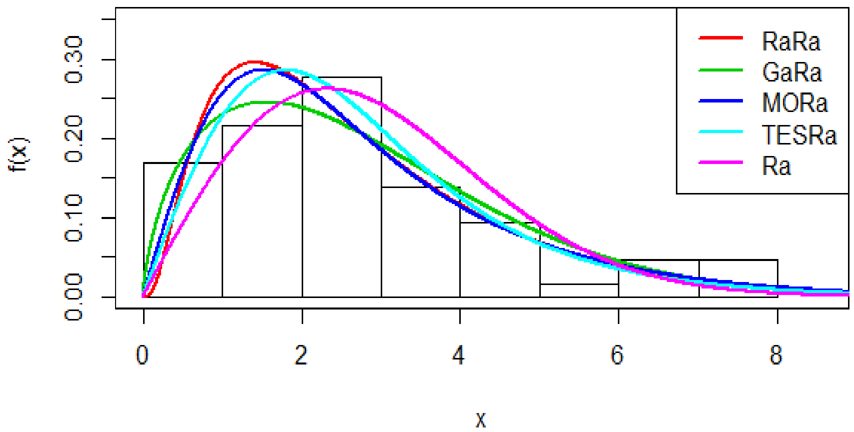


Figure 3. Histogram plot of the dataset with other compared probability distributions

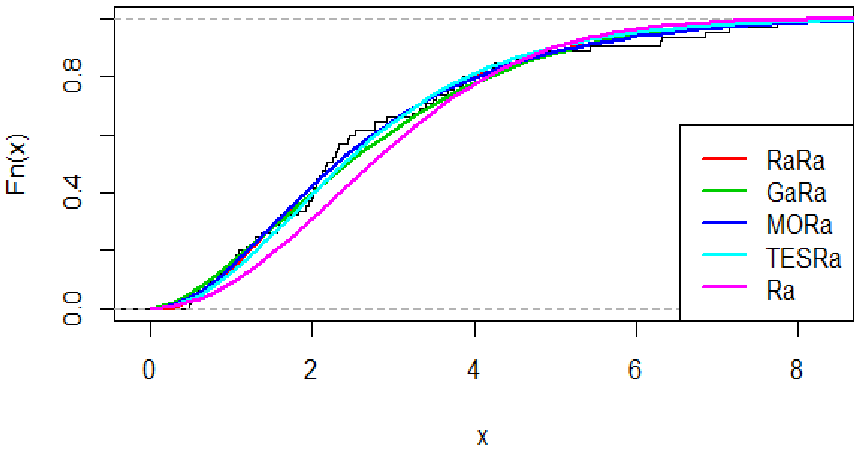


Figure 4. Empirical cdf of the dataset with other compared probability distributions

7. Conclusions

In this article, a new extended family to the Rayleigh distribution built on composing two cumulative distribution functions (cdfs) is adopted. This adoption leads to a new family of continuous distributions, named the Rayleigh - G distributions. General formulas for the basic functions of the new family are investigated. A special case "sub-model" of the new family, named Rayleigh - Rayleigh distribution is considered. The most essential statistical properties of the new distribution are investigated. There is a certain advantage of using the proposed distribution like its cdf has a closed-form. The parameters estimation via the maximum likelihood method is discussed. Numerical illustrations via simulation study and application with real data set are conducted. Through the simulation study, the proficiency and consistency of the maximum likelihood estimators (MLEs) of the proposed distribution are illustrated. In the practical application, the real data set taken from Mathers et al. [16] representing the crude mortality rate (CMR) among people who inject drugs. The proposed distribution "Rayleigh – Rayleigh" reveals better fits "more flexibility" to this real data set than the other compared distributions. This flexibility enables using Rayleigh – Rayleigh distribution in various application areas.

References

[1] Mudholkar G S and Srivastava D K 1993 Exponentiated Weibull family for analyzing bathtub failure rate data *IEEE Transactions on Reliability* **42**(2) pp 299-302.
[2] Marshall A W and Olkin I 1997 A new method for adding a parameter to a family of distributions with application to the exponential and Weibull families *Biometrika* **84**(3) pp 641-52.
[3] Gupta R C, Gupta P I and Gupta R D 1998 Modeling failure time data by Lehmann alternatives

Communications in statistics-Theory and Methods **27**(4) pp 887-904.
[4] Ristic M M and Balakrishnan N 2012 The gamma exponentiated exponential distribution *Journal of Statistical Computation and Simulation* **82**(8) pp 1191-206.
[5] Alzaatreh A, Lee C and Famoye F 2013 A new method for generating families of continuous distributions *Metron* 71(1) pp 63-79.
[6] Nadarajah S, Nassiri V and Mohammadpour A 2014 Truncated-exponential skew-symmetric distributions *Statistics* **48**(4) pp 872-95.
[7] Tahir M, Zubair M, Mansoor M, Cordeiro G M, Alizadeh M and Hamedani G 2016 A new Weibull-G family of distributions *Hacettepe Journal of Mathematics and Statistics* **45**(2) pp 629-47.
[8] Ahmad Z, Elgarhy M and Hamedani G G 2018 A new Weibull-X family of distributions: properties, characterizations and applications *Journal of Statistical Distributions and Applications* **5**(5) pp 1-18.
[9] Al-Noor N H, Abid S H and Boshi M A A 2019 On the exponentiated Weibull distribution *AIP Conference Proceedings* **2183**, 110003
[10] Kundu D and Raqab M Z 2005 Generalized Rayleigh distribution: Different methods of estimations *Computational Statistics and Data Analysis* **49**(1) pp 187-200.
[11] Voda V G 2007 A new generalization of Rayleigh distribution *Reliability: Theory and Applications* **2**(2) pp 47-56.
[12] Merovci F 2013 Transmuted Rayleigh distribution *Austrian Journal of Statistics* **42**(1) pp 21-31.
[13] Merovci F 2014 Transmuted generalized Rayleigh distribution *Journal of Statistics Applications and Probability* **3**(1) pp 9-20.
[14] Merovci F and Elbatal I 2015 Weibull Rayleigh distribution: Theory and applications *Applied Mathematics and Information Sciences* **9**(4) pp 2127-37.
[15] Ateeq K, Qasim T B and Alvi A R 2019 An extension of Rayleigh distribution and applications *Cogent Mathematics and Statistics* **6**(1) pp 1-16.
[16] Mathers B M, Degenhardt L, Bucello C, Lemon J, Wiessing L and Hickman M 2013 Mortality among people who inject drugs: a systematic review and meta-analysis *Bull World Health Organ*. **91**(2) pp 102-23.

Assessment of Teachers’ Knowledge about Tuberculosis at Primary School in Balad City

Wafaa S. Hasanain

Department of Mathematics, College of Science, Mustansiriyah University, Iraq
E-mail: w.s.hasanain@uomustansiriyah.edu.iq
or
E-mail: w_a_fairness@yahoo.com

Abstract: Tuberculosis (TB) is one of the most common infectious diseases worldwide and continues to be a major public health problem for low and middle-income countries. Undoubtedly, Lack of knowledge about tuberculosis among health care and education workers, as well as if knowledge and practices of tuberculosis among students were generally insufficient causes an increased risk of contracting the disease.

Tuberculosis (TB) is a chronic communicable bacterial disease caused by Mycobacterium tuberculosis. The Latest World Health Organization (WHO) Report shows that there were 9.0 Million new TB cases and 1.5 Million tuberculosis deaths. The Transmission of the TB disease by Mycobacterium tuberculosis (a bacterium of a group that includes the causative agents of tuberculosis). takes place by air in the form of sneeze, talk, cough, spit, etc.[1,9,11,12,13]

This applied study attempt to identify, assess and analyze teachers' knowledge about tuberculosis in primary schools. A descriptive design, cross-sectional study was carried out in order to achieve the earlier stated objectives of this study by find out the relationship between teachers' knowledge and social demographic data (sex, age, academic achievement, ….).

The present study lasted for four months by prepared a questionnaires to assess the level of teachers' knowledge, and these questionnaire contains many themes, each theme contained a number of questions to evaluate and analyze teachers ’knowledge of tuberculosis by answering a set of questions (as a variables); (mode of transmission, symptoms and signs, diagnostic features of TB, duration of treatment, prevention methods, risk of developing tuberculosis).The research hypothesis also states that (mycobacterium tuberculosis factor) has a direct impact on TB infection, and to achieve this hypothesis, a questionnaire was distributed to a sample with a size of (58) teachers and the method of Multiple Logistic Regression was used for statistical treatment. Finally, the research concluded a set of results and conclusions included in tables by comparing Likelihood-ratio chi-square statistics and classification table of the observed versus predicted responses.

Keywords: Tuberculosis (TB); Assessment teachers' knowledge; Statistical analysis, Describing Demographic Characteristics, Logistic Regression

1. Introduction

In recent years, some models of modern statistical applications have increased in the analysis of categorical data, especially in the medical, social, cognitive and other fields, and logistic regression analysis is one of the most important statistical models used in describing and analyzing data for these phenomena.

The importance of using multiple logistic regression has increased because it is concerned with analyzing data when the dependent variable is a nominal scale and includes more than two categories[2,3,5,6,8,10,14], and it depends on choosing one of the categories of the dependent variable as a reference class and then a binary regression analysis is done by comparing this reference class with other categories.

Multiple logistic regression can be represented as:

$$Ln\left[\frac{p}{1-p}\right] = \beta_0 + \beta_1 x_{1j} + \beta_2 x_{2j} + \cdots + \beta_k x_{kj} \qquad (1)$$

Using the logit transformation, we now have

$$p = \left[\frac{exp(\beta_0 + \beta_1 x_{1j} + \beta_2 x_{2j} + \cdots + \beta_k x_{kj})}{1 + exp(\beta_0 + \beta_1 x_{1j} + \beta_2 x_{2j} + \cdots + \beta_k x_{kj})}\right] \qquad (2)$$

The logistic function curve as in (figure1).

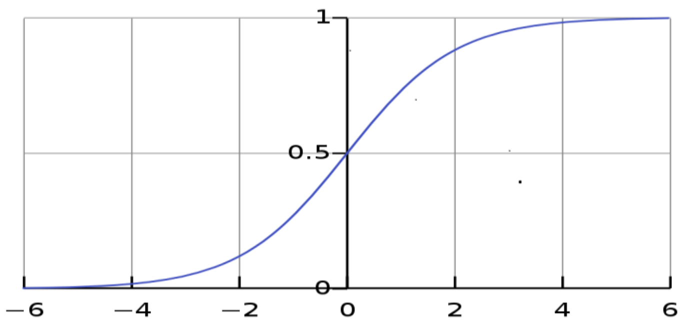


Figure 1. Graphing the logistic function curve[4,7]

2. Design of the Study

A descriptive design, cross-sectional study was carried out in order to achieve the earlier stated objectives of this study, the present study is started for the period on (4 December 2018 to 3 April 2019). The study was carried out to determine the assessment of teacher's knowledge about tuberculosis in balad city as a main objective in this study.

3. Administrative Arrangement

In order to initiate this study and collecting the data, the researcher obtained the necessary official approvals. addition to these approvals, the researcher obtained the agreement from participants before distribution of the questionnaire.

4. Setting of the Study

The current study was conducted on the teachers who are working at the schools in Balad City.

5. Sample of the Study

The population in this study is the entire cohort of teachers who are working at schools in Balad. The target population of this study was the teachers who are working in Balad city,(58) teachers.

5.1. Inclusion Criteria

1- All participating teachers working at schools in Balad city.
2- Both sexes (male and female).
3- Those who are voluntary participated.
4- Those who are working at schools and are engaged in teachers tasks and responsibilities.

5.2 Exclusion criteria

1. Those are not willing to participate in the Study.
2. Respondent did not fill the questionnaire completely.

6. Instrument of the study

A questionnaire was designed by the researcher through adoption and modification of the scales that contribute in achieving the objectives of this study. The researcher was developing the instrument of the study depending on the followings:
1. Extensive review of available literature.
2. The validity of the questionnaire was determined through a panel of experts (6 experts). The questionnaire was appropriately designed and considered valid after taking into consideration their suggestion and after getting all the comments and recommendation in attention.

The questionnaire of the study is composed of two parts which are distributed as follows:

6.1. Part 1: socio-demographic information

1. Sex

1.1. Male
1.2. Female

2. Age
2.1. (30 - 39)
2.2. (40 - 49)
2.3. (50 and above)

3. Marital status:
3.1. Unmarried
3.2. Married
3.3. Divorced

4. Academic achievement:
4.1. Graduated teacher's house
4.2. Graduate Institute:
4.3. College graduate / Bachelor

6.2. Part 2: Assessment of teachers' knowledge about tuberculosis at primary school

5. Tuberculosis is …:
5. 1 Communicable bacterial disease
5.2. Disease due to smoking and alcohol
5.3. Hereditary disease

6. Cause of TB:
6.1. Cold wind

6.2. Bacteria
6.3. Smoking
6.4. Poor hygiene:

7. Mode of transmission:
7.1. Through coughing droplet
7.2. Though sharing dish
7.3. Through shaking hands

8. Signs and symptom:
8.1. Cough for 2 weeks or above
8.2. Weight loss
8.3. On-going fatigue
8.4. Persistent fever

9. diagnostic features of TB (Detection of Tuberculosis):
9.1. Sputum test
9.2. **Blood test**
9.3. Chest X-ray
9.4. **Based on sign and symptom**
9.5. Mantoux test (a test for immunity to tuberculosis using intradermal injection of tuberculin)

10. Duration of treatment:
10.1. (4-6) months
10.2. (6-8) months:
10.3. (8-10) months
10.4. (10-12) months

11. Prevention methods:
11.1. Cover mouth when coughing/sneezing
11.2. Washing hands
11.3. Avoiding handshakes
11.4. Isolating TB patients
11.5. Vaccination
11.6. Sufficient ventilation

12. Risk of Developing Tuberculosis:
12.1. Elderly
12.2. Family and person in close contact to patient
12.3. School Children
12.4. Smokers

7. Pilot study
After review of the questionnaire by experts and its approval, a pilot study was carried out before starting the actual data collection on a purposive sample of teachers (N= 6) from both genders at schools in Balad to achieve the following aims.
1. To determine the reliability of the questionnaires.
2. To estimate the average time required for the data collection of each respondent.
3. The clarity of questions items

The results of pilot study display that

1. The items of the questionnaire were clear and understood.
2. The time required for answering the questionnaire range from (10-15 minutes).

8. Reliability of Questionnaire

In order to determining the reliability of the present study's instrument, Statistical Package for Social Science Program (IBM SPSS) version 24.0 was used for the purpose of reliability' determination.

The Cronbach's Alpha (Alpha Correlation Coefficient) was computed in order to determine the internal consistency of the instrument (table 1).

Table 1. Reliability Analysis of the Scale (n= 6)

Scale	Cronbach's Alpha
Teachers knowledge	0.73

9. Data Collection

The data were collected for the present study through the utilization of the self-administrative questionnaire, for all subjects who were included in the study sample. The researcher distributed the questionnaire for teachers after taking their willing to participate in this study.

11. Descriptive Statistical Data Analysis

After the collection of data, they have been coded and analyzed by the application of statistical procedures and by using (IBM SPSS) program to analyse and assess the results of the study.

The method of analysis used was descriptive statistical. This analysis was performed through computation (Frequencies (F)), and (Percentages (%)), as well as the use of multiple logistic regression analysis for variables.

12. Practical study

12.1 study samples by demographic characteristics

In this section, a practical study has been done to assess the result. After collecting data and relying on the results of the questionnaire, which was distributed to more than one teacher, and after excluding the non-typical answers and conforming to the fundamental rules in answering the questionnaire, the results of 58 questionnaires that represent the sample size in questions were fixed.

Table 2. Distribution of the study samples by demographic characteristics

Demographic data groups	Study	F	%
	Cold wind	12	20.7
Cause of TB	Bacteria	29	50.0
	Smoking and alcohol	2	3.4
	Poor hygiene	15	25.9
Sex	Male	16	27.6
	Female	42	72.4
	30 – 39	11	19.0
Age	40 – 49	42	72.4
	50 and above	5	8.6
	Graduated teachers house	8	13.8
Academic achievement	Graduated Institute	32	55.2
	College graduate / Bachelor	18	31.0
	Unmarried	5	8.6
Marital Status	Married	53	91.4
	Divorce	0	0.0
	Communicable bacterial disease	48	82.8
Tuberculosis is	Disease due to smoking and alcohol	9	15.5
	Hereditary disease	1	1.7
	Through coughing droplet	42	72.4
Mode of transmission	Though sharing dish	14	24.1
	through shaking hands	2	3.4
	cough for 2 weeks or above	17	29.3
Signs and Symptom	weight loss	9	15.5
	ongoing fatigue	8	13.8
	persistent fever	24	41.4
	sputum test	17	29.3
Diagnostic features of TB	blood test	13	22.4
(Detection of Tuberculosis)	chest X- ray	27	46.6
	Based on sign and symptom	1	1.7
	Mantoux test	0	0.0
	4 - 6months	33	56.9
Duration of treatment	6 - 8months	11	19.0
	8 - 10 months	2	3.4
	10-12 months	12	20.7
	cover mouth when coughing / sneezing	6	10.3
	washing hands	2	3.4
Prevention methods	avoiding handshakes	3	5.2
	isolating TB patients	24	41.4
	Vaccination	17	29.3
	sufficient ventilation	6	10.3
	cover mouth when coughing / sneezing	6	10.3
	Elderly	7	12.1
Risk of Developing tuberculosis	family and person in close contact to patient	9	15.5
	school children	23	39.7
	Smokers	19	32.8

(Table 2) shows that most of the study sample were their (Cause of TB; Cold wind, Bacteria, Smoking and alcohol, Poor hygiene) which accounted (20.7, 50.5, 3.4, 25.9)% respectively, whereas most of them (50%) answer bacteria is causes the TB and it is the correct answer. The majority of studied sample was female (72.4%), and male (27.6%). At age group for participants in the study more than a halve age ((40 – 49) years; 72.4%), followed by those who age ((30 – 39) years; 19.0%), those who (50 and above years; 8.6%). More than halve of study group's participants were graduated from Institute, then from college and finally from teachers house as percentages (55.2,31.0, 13.8)% respectively. The most of teachers were married, unmarried and they are accounted (91.4 , 8.6)% respectively. For the Tuberculosis, most teachers diagnosed it as a (Communicable bacterial disease at 82.8 percentage) and this is correct diagnose. For the Mode of transmission the first transmission is due to coughing droplet (72.4%), followed by those who sharing dish (24.1%), and (3.4) for shaking hands. The answer for Signs and Symptom recorded (cough for 2 weeks or above, weight loss, on going fatigue, persistent fever which accounted (29.3, 15.5, 13.8, 41.4)% respectively. The teachers diagnose features of TB as (sputum test, blood test, chest X- ray, based on sign and symptom) with (29.3, 22.4 46.6, 1.7)% respectively. For the study group about duration of treatment they came at these order (4–6 months ; 56.9, 6–8 months ; 19.0, and 8–10 months ; 3.4, 10–12 months ; 20.7) respectively. Followed by Prevention methods ,they recorded (cover mouth when coughing / sneezing, washing hands, avoiding handshakes, isolating TB patients, vaccination, sufficient ventilation, cover mouth when coughing / sneezing) with (10.3, 3.4, 5.2, 41.4, 29.3, 10.3, 10.3)% respectively. Finally risk of developing tuberculosis listed (elderly, family and person in close contact to patient, school children, smokers) with (12.1, 15.5, 39.7, 32.8)%.

12.2 Building the logistic regression model for the dependent variable

After that to test the significance of relationship between the dependent and independent variables included in the study using multiple logistic regression. The factor of (causes of tuberculosis) are chosen as a dependent variable divided into four categories, and the reference category that was established indicates that tuberculosis is the result of infection with TB bacteria, and compared with the independent variables to assess and analyze the knowledge of teachers 'responses to Tuberculosis. The (table 3) indicates the categories and values of dependent variable.

Table 3. categories and values of dependent variable

Categ. Dep. Var.	values
Cold wind	1
Bacteria	2
Smoking and alcohol	3

12.3 Model estimation of sex variable

A multinomial logit model is fit for the full factorial model by applying the (equation 2) and it is performed through an iterative maximum likelihood algorithm , Likelihood is often assumed to be the same as a probability or even as a P-value ***[Wikipedia demonstrates],[Kinear & Gray 2006 p.483]***. (Table 4) indicates the number of Iteration cycles of the derivative of the maximum likelihood function within the categories of the sex variable.

Table 4. Iteration Maximum likelihood of sex variable

Iteration History*

		Cause of TB					
		Cold wind		Smoking and alcohol		Poor hygiene	
Iteration	-2 Log Likelihood	Intercept	[Sex=1]	Intercept	[Sex=1]	Intercept	[Sex=1]
0	23.633	-.882389	.000000	-2.674149	.000000	-.659246	.000000
1	17.219	-1.382389	1.812500	-3.174149	1.812500	-.929087	978175
2	17.082	-1.426332	1.833048	-3.218091	1.833048	-.916149	1.147943
3	17.081	-1.427116	1.832578	-3.218876	1.832578	-.916291	1.139444
4	17.081	-1.427116	1.832581	-3.218876	1.832581	-.916291	1.139434
5	17.081	-1.427116	1.832581	-3.218876	1.832581	-.916291	1.139434

* Source: the results are Preparing by the researcher based on of the data analysis.

From the (table 4) and to obtain the lowest value of the negative logarithm function to get an optimal estimate of the parameters, we stop in the fifth iteration of the negative derivative (-2 Log Likelihood), and we got the lowest value which is (1.139434) to the differences between the parameters, where these differences reached to less than (0.001).

Table 5. Estimates of the logistic regression of cause of TB with sex model

Parameter Estimates

Cause of T B[a]		B	Std. Error	Wald	df	Sig.	Exp(B)	95% Confidence Interval for Exp(B) Lower Bound	Upper Bound
Cold wind	Intercept	-1.427	.455	9.855	1	.002			
	[Sex=1]	1.833	.790	5.388	1	**.020**	1.330	29.371	
	[Sex=2]	0[b]	.	.	0	.	.	.	
Smoking and alcohol	Intercept	-3.219	1.020	9.963	1	.002			
	[Sex=1]	1.833	1.513	1.467	1	.226	.322	121.334	
	[Sex=2]	0[b]	.	.	0	.	.	.	
Poor hygiene	Intercept	-.916	.374	5.997	1	.014			
	[Sex=1]	1.139	.768	2.201	1	.138	.693	14.082	
	[Sex=2]	0[b]	.	.	0	.	.	.	

a. The reference category is: Bacteria.
b. This parameter is set to zero because female is reference category.

(Table 5) depict the regression coefficients, standard error, Wald statistic with degrees of freedom and significance for the indp. var. (sex) and confidence interval for Exp(B) and the constant term would have been the log odds ratio for bacteria. For the dependent variable (the Cause of TB) had been a multinomial categories, and had specified bacteria as the reference group. The constant term (B) is the estimated coefficient and equal to log-odds ratios, and the equation of estimated log. regression

log(cold wind) = -1.427 + 1.833*(sex=1) (3)

log(Smoking and alcohol) = -3.219 + 1.833*(sex=1) (4)

log(Poor hygiene) = -.916 + 1.139*(sex=1) (5)

These estimates tell about the relationship between (sex) variable and the (cause of TB), where the dependent variable is on the logit scale, and from the table it turned out significance the effect of sex when the teacher is male and his choice was cold wind. The fact that the teacher is male means that it increases the probability of choosing (exposure to cold air) versus infection by bacteria (1.833) compared to the female, which represents as a reference category, that means if the sex of teacher is men, then the logarithm of the possibility [Exp(B); this is the odd ratio for the predictors, this is the exponentiation of the coefficients] of choosing cold wind versus bacteria increases by an amount (1.33), whereas the teacher's choice of smoking and alcohol and Poor hygiene reasons did not show any significant.

The Wald columns provide the Wald chi-square value and 2-tailed p-value used in testing the null hypothesis and it is equal to (5.388).

Table 6. Classification table of TB with sex variables

classification

	Predicted				
Observed	Cold wind	Bacteria	Smoking and alcohol	Poor hygiene	Percent Correct
Cold wind	6	6	0	0	50.0%
Bacteria	4	25	0	0	**86.2%**
Smoking and alcohol	1	1	0	0	0.0%
Poor hygiene	5	10	0	0	0.0%
Overall Percentage	27.6%	72.4%	0.0%	0.0%	**53.4%**

we notice that from the classification table (table 6) the overall accuracy of this model to predict cause of TB by using the independent variable (sex) is (53.4%) which is somewhat considered acceptable. We notes that the number of classified and selected the correct answer of teachers (25 teachers) to bacteria at percentage (86.2%) and that they were wrongly chosen by (4 teachers).

12.4 Model estimation of Tuberculosis variable

Now we test the significant between the dependent (cause of Tb) and the independent (Tuberculosis).

Table 7. Iteration maximum likelihood of Tuberculosis variable

Iteration History*

		Cause of TB								
		Cold wind			Smoking and alcohol			Poor hygiene		
Iteration	-2 Log Likelihood	Intercept	[Tuberculosis is=1]	[Tuberculosis is=2]	Intercept	[Tuberculosis is=1]	[Tuberculosis is=2]	Intercept	[Tuberculosis is=1]	[Tuberculosis is=2]
0	27.336	-.882	.000	.000	-2.674	.000	.000	-.659	.000	27.336
1	16.838	-2.882	1.580	4.463	-4.674	1.479	5.000	-2.659	1.922	16.838
2	15.867	-4.018	2.669	5.625	-5.809	2.518	5.881	-3.795	3.064	15.867
3	15.722	-5.061	3.711	6.671	-6.853	3.557	6.851	-4.838	4.107	15.722
4	15.702	-6.077	4.727	7.686	-7.868	4.572	7.868	-5.853	5.122	15.702
5	15.695	-7.082	5.732	8.692	-8.874	5.578	8.874	-6.859	6.128	15.695
6	15.693	-8.084	6.734	9.694	-9.876	6.580	9.876	-7.861	7.130	15.693
7	15.692	-9.085	7.735	10.694	-10.877	7.581	10.877	-8.862	8.131	15.692
8	15.691	-10.085	8.735	11.695	-11.877	8.581	11.877	-9.862	9.131	15.691
9	15.691	-11.085	9.735	12.695	-12.877	9.581	12.877	-10.862	10.131	15.691
10	15.691	-12.085	10.735	13.695	-13.877	10.581	13.877	-11.862	11.131	15.691
11	15.691	-13.085	11.735	14.695	-14.877	11.581	14.877	-12.862	12.131	15.691
12	15.691	-14.085	12.735	15.695	-15.877	12.581	15.877	-13.862	13.131	15.691
13	15.691	-15.085	13.735	16.695	-16.877	13.581	16.877	-14.862	14.131	15.691
14	15.691	-16.085	14.735	17.695	-17.877	14.581	17.877	-15.862	15.131	15.691
15	27.336	-.882	.000	.000	-2.674	.000	.000	-.659	.000	27.336

*Source: the results are Preparing by the researcher based on of the data analysis.

For the (table 7) and to obtain the lowest value of the negative logarithm function to get an optimal estimate of the parameters, we stop in the fifteen iteration of the negative derivative (-2 Log Likelihood), and we got last absolute change in -2 Log Likelihood is .000, and last maximum absolute change in parameters is 1.000000.

Table 8. Estimates of the logistic regression cause of TB with tuberculosis model

Parameter Estimates

Cause of TB	B	Std. Error	Wald	df	Sig.	Exp(B)

	Intercept	-17.085	5128.877	.000	1	.997	
	[Tuberculosis is=1]	15.735	5128.877	.000	1	.998	6819911.866
Cold wind	[Tuberculosis is=2]	18.695	5128.877	.000	1	.997	131526871.70
	[Tuberculosis is=3]	0	.	.	0	.	.
	Intercept	-18.877	1.414	178.171	1	.000	
	[Tuberculosis is=1]	15.581	1.743	79.938	1	**.000**	5845638.598
Smoking and alcohol	[Tuberculosis is=2]	18.877	.000	.	1	.	157832242.10
	[Tuberculosis is=3]	0	.	.	0	.	.
	Intercept	-16.862	4587.407	.000	1	.997	
	[Tuberculosis is=1]	16.131	4587.407	.000	1	.997	10132440.360
Poor hygiene	[Tuberculosis is=2]	17.555	4587.407	.000	1	.997	42088598.410
	[Tuberculosis is=3]	0	.	.	0	.	.

(Table 8) presents the regression coefficients, standard error, Wald statistic with degrees of freedom and significance for the indp. Var., significance .The teacher's choices is the effect of infectious bacterial diseases are transmitted, which corresponds to the fact that the cause of tuberculosis is the bacterial agent. The fact that when the teacher choices communicable bacterial disease means that it increases the probability of choosing (smoking and alcohol) versus infection by bacteria for the cause of tuberculosis by register a number equal to (b=15.581) with sig = 0.00, compared to the rest of tuberculosis (hereditary disease), whilst there is no significant between cause of TB represented by (Cold wind and Poor hygiene) and the tuberculosis, and the equation of estimated log. regression are

$$\log(\text{cold wind}) = -17.09 + 15.74*(\text{ Tuberculosis is}=1) + 15.74*(\text{ Tuberculosis is}=2) \quad (6)$$

$$\log(\text{Smoking and alcohol})= -18.87 +15.58*(\text{ Tuberculosis is}=1) + 18.88*(\text{ Tuberculosis is}=2) \quad (7)$$

$$\log(\text{Poor hygiene}) = -16.86+16.13*(\text{ Tuberculosis is}=1) + 17.56*(\text{ Tuberculosis is}=2) \quad (8)$$

Table 9. Classification table of TB with tuberculosis variables

Classification - TUBERCULOSIS

	Predicted				
observed	cold wind	bacteria	smoking and alcohol	poor hygiene	Percent correct
Cold wind	5	7	0	0	41.7%
Bacteria	1	28	0	0	**96.6%**
Smoking and alcohol	1	1	0	0	0.0%
Poor hygiene	2	13	0	0	0.0%
Overall Percentage	15.5%	84.5%	0.0%	0.0%	**56.9%**

From the (table 9) we find that the ratio of the correct classification percentage to the bacteria is (96.6%), or for those who are chosen bacteria are (28 teachers) and the rest chose cold wind, smoking and alcohol, Poor hygiene with percentage equal to zero. The overall classification rate has reached (56.9%).

Table 10. Iteration Maximum likelihood of Mode of transmission variable

Iteration History											
		Cause of TB									
		Cold wind			Smoking and alcohol			Poor hygiene			
Iteration	-2 Log Likelihood	Intercept	[Mode of transmission=1]	[Mode of transmission=2]	Intercept	[Mode of transmission=1]	[Mode of transmission=2]	Intercept	[Mode of transmission=1]	[Mode of transmission=2]	
0	22.299	-.882389	.000	.000	-2.674	.000	.000	-.659	.000	.000	
1	18.922	.534277	-1.706	-.750	-3.674	.595	2.357	-1.659	.917	1.390	
2	18.475	-.070414	-1.119	-.153	-5.135	2.001	3.561	-3.120	2.383	2.899	
3	18.406	.001883	-1.191	-.225	-6.126	2.991	4.517	-4.111	3.374	3.888	
4	18.383	-1.747419E-5	-1.190	-.223	-7.136	4.001	5.527	-5.121	4.384	4.898	
5	18.374	5.909074E-8	-1.190	-.223	-8.140	5.004	6.530	-6.125	5.387	5.902	
6	18.371	-7.326481E-11	-1.190	-.223	-9.141	6.005	7.532	-7.126	6.388	6.903	
7	18.370	3.275862E-14	-1.190	-.223	-10.141	7.006	8.532	-8.126	7.389	7.903	
8	18.369	-1.740438E-15	-1.190	-.223	-11.142	8.006	9.532	-9.127	8.389	8.904	
9	18.369	-1.087408E-15	-1.190	-.223	-12.142	9.006	10.532	-10.127	9.389	9.904	
10	18.369	1.260736E-15	-1.190	-.223	-13.142	10.006	11.532	-11.127	10.389	10.904	
11	18.369	-1.006894E-15	-1.190	-.223	-14.142	11.006	12.532	-12.127	11.389	11.904	
12	18.369	-1.240451E-15	-1.190	-.223	-15.142	12.006	13.532	-13.127	12.389	12.904	
13	18.369	-1.941337E-15	-1.190	-.223	-16.142	13.006	14.532	-14.127	13.389	13.904	
14	18.369	-1.383785E-16	-1.190	-.223	-17.142	14.006	15.532	-15.127	14.389	14.904	
15	18.369	-1.085473E-15	-1.190	-.223	-18.142	15.006	16.532	-16.127	15.389	15.904	
16	18.369	1.433960E-15	.000	.000	-2.674	.000	.000	-.659	.000	.000	

12.5 Model estimation of Mode of transmission variable

Now we test the significant between the dependent (cause of Tb) and the independent (Mode of transmission).

For the (table 10) and to obtain the lowest value of the negative logarithm function to get an optimal estimate of the parameters, we stop in the sixteen iteration of the negative derivative, Last absolute change in -2 log likelihood is (0.000), and last maximum absolute change in parameters is (1.000000).

Table 11. Estimates of the logistic regression cause of TB with mode of transmission model

Parameter Estimates

Cause of TB[a]		B	Std. Error	Wald	df	Sig.	Exp(B)	95 % confidence interval for exp (B.) Lower Bound	Upper Bound
Cold wind	Intercept	.000	1.414	.000	1	1.000			
	[Mode of transmission=1]	-1.190	1.479	.647	1	.421	.304	.017	
	[Mode of transmission=2]	-.223	1.565	.020	1	.887	.800	.037	5.521
	[Mode of transmission=3]	0[b]	.	.	0	.	.	.	17.196
Smoking and alcohol	Intercept	-19.142	1.095	305.336	1	.000			.
	[Mode of transmission=1]	16.006	1.498	114.197	1	.000	8941132.727	474724.898	
	[Mode of transmission=2]	17.532	.000	.	1	.	41129210.540	41129210.540	168400382.500
	[Mode of transmission=3]	0[b]	.	.	0	.	.	.	41129210.540
Poor hygiene	Intercept	-17.127	5236.361	.000	1	.997			.
	[Mode of transmission=1]	16.389	5236.361	.000	1	.998	13113660.540	.000	
	[Mode of transmission=2]	16.904	5236.361	.000	1	.997	21935577.630	.000	
	[Mode of transmission=3]	0[b]	.	.	0	.	.	.	

a. The reference category is: Bacteria.
b. This parameter is set to zero because it is redundant

(Table 11) presents the regression coeff., standard error, Wald statistic and significance for the indp. var.,(mode of transmission), the value of exp(B), it turned out significance the effect of mode of transmission. when the teacher choices (through coughing droplet) instead of bacterial agent to mode of transmission means that it increases the probability of choosing (smoking and alcohol) versus infection by bacteria for the cause of tuberculosis by register a number equal to (16.006) compared to the rest of tuberculosis (through shaking hands), whilst there is no significant between cause of TB represented by (Cold wind and Poor hygiene) and the tuberculosis, and the estimated equations

log(cold wind) = 0.000-1.190*(Mode of transmission =1) - 0.223*(Mode of transmission) =2 (9)

log(Smoking and alcohol)=-19.142+16.006*(Mode of transmission=1)+17.532*(Mode of transmission=2) (10)

log(Poor hygiene)=-17.127+16.389*(Mode of transmission=1)+16.904*(Mode of transmission=2) (11)

Table 12. Classification table of TB with mode of transmission variables

Classif. - MODE

Obser.	Predicted: Cold wind	bacteria	smoking and alcohol	poor hygiene	Perc. correct
Cold wind	1	11	0	0	**8.3%**
Bacteria	1	28	0	0	**96.6%**
Smoking and alcohol	0	2	0	0	0.0%
Poor hygiene	0	15	0	0	0.0%
overall Percentage	3.4%	96.6%	0.0%	0.0%	**50.0%**

Table (12) exhibit the ratio of the correct classification percentage to the bacteria is (96.6%), or for those who are chosen bacteria are (28 teachers) and the rest chose cold wind (8.3%), smoking and alcohol, Poor hygiene with percentage equal to zero. The overall classification rate has reached a halve percentage (50.0%).

12.6 Model estimation of Signs and Symptom variable

Table 13. Iteration Maximum likelihood of Signs and Symptom variable

Iteration History

		Cause of TB											
		Cold wind				Smoking and alcohol				Poor hygiene			
Iteration	-2 Log Likelihood	Intercept	[Sig. and Symp. =1]	[Sig and Symp.=2]	[Sig. and Symp. =3]	Intercept	[Sig. and Symp.=1]	[Sig. andSymp. m=2]	[Sig. and Sympt.=3]	Intercept	[Sig. and Symp.=1]	[Sig. and Symp.=2]	[Sig. and Symp.=3]
0	39.276	-.882	.000	.000	.000	-2.674	.000	.000	.000	-.659	.000	.000	.000
1	26.937	-1.028	-.011	1.183	-.250	-2.216	-1.752	-1.569	2.167	.363	-1.861	-2.133	-1.056
2	23.135	-1.102	.092	.844	-.316	-2.195	-2.806	-2.901	1.139	.200	-1.891	-3.281	-.893
3	22.600	-1.099	.087	.877	-.286	-2.197	-3.810	-3.913	.847	.201	-1.905	-4.296	-.894
4	22.445	-1.099	.087	.875	-.288	-2.197	-4.812	-4.924	.811	.201	-1.905	-5.307	-.894

5	22.390	-1.099	.087	.875	-.288	-2.197	-5.812	-5.928	.811	.201	-1.905	-6.311	-.894
6	22.369	-1.099	.087	.875	-.288	-2.197	-6.812	-6.929	.811	.201	-1.905	-7.312	-.894
7	22.361	-1.099	.087	.875	-.288	-2.197	-7.813	-7.930	.811	.201	-1.905	-8.313	-.894
8	22.359	-1.099	.087	.875	-.288	-2.197	-8.813	-8.930	.811	.201	-1.905	-9.313	-.894
9	22.358	-1.099	.087	.875	-.288	-2.197	-9.813	-9.930	.811	.201	-1.905	-10.313	-.894
10	22.357	-1.099	.087	.875	-.288	-2.197	-10.813	-10.930	.811	.201	-1.905	-11.313	-.894
11	22.357	-1.099	.087	.875	-.288	-2.197	-11.813	-11.930	.811	.201	-1.905	-12.313	-.894
12	22.357	-1.099	.087	.875	-.288	-2.197	-12.813	-12.930	.811	.201	-1.905	-13.313	-.894
13	22.357	-1.099	.087	.875	-.288	-2.197	-13.813	-13.930	.811	.201	-1.905	-14.313	-.894
14	22.357	-1.099	.087	.875	-.288	-2.197	-14.813	-14.930	.811	.201	-1.905	-15.313	-.894
15	22.357	-1.099	.087	.875	-.288	-2.197	-15.813	-15.930	.811	.201	-1.905	-16.313	-.894
16	22.357	-.882	.000	.000	.000	-2.674	.000	.000	.000	-.659	.000	.000	.000
17	22.357	-1.028	-.011	1.183	-.250	-2.216	-1.752	-1.569	2.167	.363	-1.861	-2.133	-1.056

For the (table 13) we stop in the seventeen iteration of the negative derivative and the last absolute change in -2 log likelihood equal to zero, and last maximum absolute change in parameters equal to one.

Table 14. Estimates of the logistic regression cause of TB with signs and symptom model

Parameter Estimates

cause of T B*		b	std. error	wald	df	sig.	Ep(B)
	Intercept	-1.099	.667	2.716	1	.099	
	[Signs and Symptom=1]	.087	.886	.010	1	.922	
	[Signs and Symptom=2]	.875	.946	.857	1	.355	
Cold wind	[Signs and Symptom=3]	-.288	1.302	.049	1	.825	1.091
	[Signs and Symptom=4]	0	.	.	0	.	2.400
	Intercept	-2.197	1.054	4.345	1	.037	.750
	[Signs and Symptom=1]	-17.813	6674.009	.000	1	.998	.
	[Signs and Symptom=2]	-17.930	.000	.	1	.	
	[Signs and Symptom=3]	.811	1.537	.279	1	.598	1.837E-8
Smoking and alcohol	[Signs and Symptom=4]	0	.	.	0	.	1.633E-8

Poor hygiene	Intercept	.201	.449	.199	1	.655	2.250
	[Signs and Symptom=1]	-1.905	.890	4.579	1	**.032**	.
	[Signs and Symptom=2]	-18.313	3833.090	.000	1	.996	
	[Signs and Symptom=3]	-.894	.976	.839	1	.360	.149
	[Signs and Symptom=4]	0	.	.	0	.	1.114E-8

***** The reference category is: Bacteria.

(Table 14) presents the regression coeff., standard error, Wald statistic with and significance for the indp. var.,(Signs and symptom), the value of exp(B), it turned out significance the effect of Signs and symptom, when the teacher choices (cough for 2 weeks or above) instead of bacterial agent to signs and symptom means that it decreases the probability of choosing (Poor hygiene) versus infection by bacteria for the signs and symptom by register a number equal to (-1.905) compared to the rest of Signs and symptom (persistent fever), whilst there is no significant between cause of TB represented by (Cold wind and Smoking and alcohol) and the signs and symptom, and the estimated equations are:

log(cold wind)= -1.099+0.087*(Signs and Symptom =1)+0.875*(Signs and Symptom =2) -0.288*(Signs and Symptom =3) (12)

log(Smoking and alcohol)= -2.197-17.813*(Signs and Symptom =1) -17.930*(Signs and Symptom =2)+0.811*(Signs and Symptom =3) (13)

log(Poor hygiene) = 0.201-1.905*(Signs and Symptom=1)-18.313*(Signs and Symptom=2) -0.894*(Signs and Symptom =3) (14)

Table 15. Classification table of TB with signs and symptom variables

Classification - SIGN

Observed	Predicted: Cold wind	Bact.	Smoking and alcohol	Poor hyg.	Perc. Corr.
Cold wind	0	9	0	3	0.0%
Bacteria	0	20	0	9	**69.0%**
Smoking and alcohol	0	1	0	1	0.0%
Poor hygiene	0	4	0	11	**73.3%**
Overall Percentage	0.0%	58.6%	0.0%	41.4%	**53.4%**

(Table 15) exhibits the ratio of the correct classification percentage to the bacteria is about (69.0%), or for those who are chosen bacteria are (20 teachers) and the rest chose Poor hygiene

(73.3%), cold wind and smoking and alcohol with percentage equal to zero. The overall classification rate has reached (53.4%).

12.7 Model estimation of Duration of treatment variable

Table 16. Iteration Maximum likelihood of Mode of transmission variable

Iteration History

		Cause of TB											
		Cold wind				Smoking and alcohol				Poor hygiene			
Iteration	-2 Log Likelihood	Intercept	[Duration of treatment=1]	[Duration of treatment=2]	[Duration of treatment=3]	Intercept	[Duration of treatment=1]	[Duration of treatment=2]	[Duration of treatment=3]	Intercept	[Duration of treatment=1]	[Duration of treatment=2]	[Duration of treatment=3]
0	38.995	-.882389	.000	.000	.000	-2.674	.000	.000	.000	-.659	.000	.000	.000
1	19.018	-.174056	-.360	-2.527	-.708	-3.174	1.288	-1.318	.500	.774	-1.700	-2.900	2.433
2	17.512	.018476	-.594	-3.808	-.901	-4.187	2.119	-1.394	1.513	.695	-1.677	-2.998	3.572
3	17.068	-5.441963E-6	-.575	-4.813	-.882	-5.196	3.117	-1.409	2.522	.693	-1.674	-2.996	4.595
4	16.909	7.489158E-9	-.575	-5.822	-.882	-6.198	4.118	-1.416	3.524	.693	-1.674	-2.996	5.603
5	16.851	-3.808443E-12	-.575	-6.825	-.882	-7.198	5.119	-1.419	4.524	.693	-1.674	-2.996	6.605
6	16.829	7.321564E-16	-.575	-7.826	-.882	-8.198	6.119	-1.420	5.524	.693	-1.674	-2.996	7.606
7	16.821	3.633731E-16	-.575	-8.827	-.882	-9.199	7.119	-1.420	6.524	.693	-1.674	-2.996	8.607
8	16.818	1.160435E-16	-.575	-9.827	-.882	-10.199	8.119	-1.420	7.524	.693	-1.674	-2.996	9.607
9	16.817	-8.587968E-16	-.575	-10.827	-.882	-11.199	9.119	-1.420	8.524	.693	-1.674	-2.996	10.607
10	16.817	-5.322556E-17	-.575	-11.827	-.882	-12.199	10.119	-1.420	9.524	.693	-1.674	-2.996	11.607
11	16.817	-3.526540E-16	-.575	-12.827	-.882	-13.199	11.119	-1.420	10.524	.693	-1.674	-2.996	12.607
12	16.817	-8.146650E-16	-.575	-13.827	-.882	-14.199	12.119	-1.420	11.524	.693	-1.674	-2.996	13.607
13	16.817	-3.640179E-16	-.575	-14.827	-.882	-15.199	13.119	-1.420	12.524	.693	-1.674	-2.996	14.607
14	16.817	4.314974E-16	-.575	-15.827	-.882	-16.199	14.119	-1.420	13.524	.693	-1.674	-2.996	15.607
15	16.817	-3.174123E-16	.000	.000	.000	-2.674	.000	.000	.000	-.659	.000	.000	.000

For the (table 16) and to obtain the lowest value of the negative logarithm function to get an optimal estimate of the parameters, we stop in the sixteen iteration of the negative derivative, Last absolute change in -2 log likelihood is (0.000), and last maximum absolute change in parameters is (1.000000).

Table 17. Estimates of the logistic regression cause of TB with duration of treatment model

Parameter Estimates

Cause of TB		B	Std. Error	Wald	df	Sig.	Exp.(B)
Cold wind	Intercept	.000	.816	.000	1	1.000	
	[Duration of treatment=1]	-.575	.917	.394	1	.530	.563
	[Duration of treatment=2]	-16.827	1425.373	.000	1	.991	4.922E-8
	[Duration of treatment=3]	-.882	7395.244	.000	1	1.000	.414
	[Duration of treatment=4]	0	.	.	0	.	.
Smoking and alcohol	Intercept	-17.199	3074.547	.000	1	.996	
	[Duration of treatment=1]	15.119	3074.547	.000	1	.996	3682570.164
	[Duration of treatment=2]	-1.420	4652.199	.000	1	1.000	.242
	[Duration of treatment=3]	14.524	.000	.	1	.	2031762.849
	[Duration of treatment=4]	0	.	.	0	.	.
Poor hygiene	Intercept	.693	.707	.961	1	.327	
	[Duration of treatment=1]	-1.674	.854	3.843	1	**.050**	.188
	[Duration of treatment=2]	-2.996	1.265	5.609	1	**.018**	.050
	[Duration of treatment=3]	16.607	3910.318	.000	1	.997	16303051.70
	[Duration of treatment=4]	0	.	.	0	.	.

The results of (table 17) turned out two significance the effect of duration of treatment, firstly when the teacher choices the period of treatment (4-6 months) means that it decreases the probability of choosing (Poor hygiene) versus infection by bacteria by registering a number equal to (-1.674) compared to the period of treatment (10-12 months), secondly when the teacher choices the period of treatment (6-8 months) means that it decreases the probability of choosing (Poor hygiene) versus infection by bacteria for the duration of treatment by register a number equal to (-2.996) compared to the period of treatment (10-12 months), whilst there is no significant between cause of TB represented by (Cold wind and Smoking and alcohol) and the duration of treatment, and the equations are:

log(cold wind)= 0.000-0.575*(Duration of treatment =1) -16.827*(Duration of treatment =2) + 0.882*(Duration of treatment =3) (15)

log(Smoking and alcohol)= -17.199 + 15.119*(Duration of treatment =1) -1.420*(Duration of treatment =2) + 14.524*(Duration of treatment =3) (16)

log(Poor hygiene)= 0.693-1.674*(Duration of treatment =1) -2.996*(Duration of treatment =2) + 16.607*(Duration of treatment =3) (17)

Table 18. Classification table of TB with Duration of treatment variables

Classification - DURATION

	Predicted				
observed	Cold wind	Bacteria	Smoking and alcohol	Poor hygiene	Percent Correct
cold wind	0	9	0	3	0.0%
bacteria	0	26	0	3	**89.7%**
smoking and alcohol	0	2	0	0	0.0%
poor hygiene	0	7	0	8	**53.3%**
overall Percentage	0.0%	75.9%	0.0%	24.1%	**58.6%**

The results of (table 18) exhibits the ratio of the correct classification percentage to the bacteria is (89.7%), or for those who are chosen bacteria are (26 teachers) and the rest chose Poor hygiene (53.3%), and cold wind, smoking and alcohol with percentage equal to zero. The overall classification rate has reached (58.6%).

13. Discussion and Conclusion

1. The sex variable showed its significant effect on teachers' knowledge of the causes of tuberculosis (bacteria), where the correct classification rate was (86.2%).
2. The variable of teacher diagnosis of the disease showed a significant effect on his knowledge of the causes of tuberculosis (bacteria), where the correct classification rate was (96.6%).
3. The variable to know the methods of transmission of infection showed a significant effect on his knowledge of the causes of tuberculosis (bacteria), where the correct classification rate (96.6%).
4. The variable to know the signs that appear on the patient with tuberculosis showed a significant effect on his knowledge of the causes of tuberculosis (bacteria), where the correct classification rate was (69.0%).
5. The variable of knowledge of the period of treatment required to recover from tuberculosis showed a significant effect on his knowledge of the causes of tuberculosis (bacteria), where the correct classification rate was (89.7%).
6. The results showed that there were no significant effect for variables (age, marital status, academic achievement) on knowledge of TB disease. Further that the teachers' knowledge of

Prevention methods, Risk of Developing Tuberculosis lacked sufficient information, which caused a lack of significant effect.

14. Recommendations

1. Through reviewing the medical studies that confirmed the seriousness of this disease, we recommend the need to work hard to spread a culture of knowledge of tuberculosis through the media and to organize training courses and workshops for members of society in general and for teachers in particular because most teachers lack knowledge of the developments of this disease.
2. Recommendation to use the logistic regression model in other future studies, especially medical and educational.

15. Acknowledgement

The researcher gratefully acknowledge Mustansiriyah University-College of Science Department of Mathematics, the teachers who answered the questionnaire, and your esteemed journal for supporting this research.

16. References

[1] Abbas I and Peter J 2014 *Knowledge and perception on tuberculosis transmission in Tanzania: Multinomial logistic regression analysis of secondary data* Tanzania Journal of Health Research .v16i1.5 Volume 16, Number 1.

[2] Agresti, A 2002 *Categorical Data Analysis*. New York: Wiley-Inter science.

[3] Chan Y H 2004 Biostatistics 202: *Logistic regression analysis*, Singapore Med J Vol. 45(4):149.

[4] David W. Hosmer J and Stanley L Rodney X S 2013 *Applied Logistic*.

[5] Draper N R and smith H 1981 *Applied Regression analysis*, New York.

[6] Hosmer D.W Lemeshow S and Klor J 1988 *Goodness of fit testing for the logistic model when the estimated probabilities*, Biometrical Journal.

[7] Hosmer D.W Lemeshow S 2000 *Applied logistic regression*. 2nd ed. New York, NY: John Wiley & Sons Inc.

[8] International Journal of Mathematics and Statistics Invention (IJMSI) E-ISSN: 2321 – 4767 P- ISSN: 2321 - 4759 www.ijmsi.org Volume 2 Issue 5 || May. 2014 || PP-01-08 www.ijmsi.org 1 | P a g e A Multinomial Logistic Regression Analysis to Study The Socio-Economic Status On Breast Cancer Incidences In Southern Karnataka Madhu B, Ashok N C and S Balasubramanian JSS University, Mysore – 570015, India

[9] Madhukar P Marcel A B David W D and Mario R 2016 Tuberculosis Nature Reviews Disease Primers Vol. 2.

[10] Park H A 2013 *An Introduction to Logistic Regression*: *From Basic Concepts to Interpretation with Particular Attention to Nursing Domain,J Korean Acad Nurs* , vol. 43,P155.

[11] Respir A J *Crit Care Med* Vol.195, P7-8, 2017 ATS Patient Education Series © American Thoracic Society.

[12] World health organizing https://www.who.int/health-topics/tuberculosis#tab=tab_1.

[13] WHO. Global TB report 2017. Annex 4: TB burden estimates, notifications and treatment ;outcomes for individual countries and territories, Who regions and the world. Accessed on February 26,2018 /gtbr2017_annex4.pdf?ua=1

[14] Azer k kh و Taleb H R *Using Multi-response Logistic Regression Model To Determine "the Factors Affecting Eye Determine the Factors Affecting Eye Disease"* , مجلة كلية الادارة والاقتصاد للدراسات الاقتصادية والادارية والمالية المجلد 11, العدد 3.

High dimensional data challenges in estimating multiple linear regression

Jassim N Hussain

Department of Statistics, College of Administration & Economics, University of Kerbala, Iraq.
Email: jasim.nasir@uokerbala.edu.iq

Abstract. Nowadays, High dimensional data are quickly increasing in many areas because of the development of new technology which helping to collect data with a large number of variables in order to better understanding for a given phenomenon of interest. Multiple Linear Regression is a famous technique used to investigate the relationship between one dependent variable and one or more of independent variables and analyzing the effects of them. Fitting this model requests assumptions, one of them is large sample size. High dimensional data does not satisfy this assumption because the sample size is small compared to the number of explanatory variables (k). Consequently, the results of traditional methods to estimate the model can be misleading. Regularization or shrinkage techniques (e.g., LASSO) have been proposed to estimate this model in this case. Nonparametric method was proposed to estimate this model. Average mean square error and root mean square error criteria are used to assess the performance of nonparametric; LASSO and OLS methods in the case of simulation study and analyzing the real dataset. The results of simulation study and the analysis of real data set show that nonparametric regression method is outperformance of LASSO and OLS methods to fit this model with high dimensional data.
Key Words:, Average Mean square error; Data reduction techniques; Nonparametric regression, Regularization methods, Variable selection method.

1. Introduction

In the last decades, high dimensional data becoming increasingly common in different areas. The cases when the number of covariates is larger than the number of observations ($p > n$) is named high dimensional. Consequently, traditional methods of analyzing statistical models (e.g., regression models) produce poor results and are unreliable for inference because of overfitting in such cases. On the other hand, Multiple Linear Regression (MLR) is one of the important procedure for modeling and analyzing data in many disciplines. Numerous assumptions are requested in fitting this model, the important one is large size of sample. Accordingly, the method of estimation are differing depending on the validity of the assumptions in the data set. If the requested assumptions are hold, then the traditional fitting methods (e.g., ordinary least square (OLS) or maximum likelihood (ML)) are used to fit the MLR. Meanwhile, when some of these assumptions do not satisfied, specifically if the size of the sample is small and the number of the covariates (p) may equal or greater than the size of the sample (n). This case is named as high dimensional data, with this case the results of traditional method when fitting MLR model produce poor and unreliable results for inference because of overfitting property. Also, the multicollinearity problem tends to occur when *n* becomes small or *p* becomes large. Therefore, the estimators of $\boldsymbol{\beta}'s$ become unstable [1].

Consequently, when analyzing the MLR with high dimensional data several alternatives methods are used including the variable selection methods, data Reduction Techniques [2,3] and The shrinkage techniques These methods give biased estimators with a smaller variance than OLS estimators [4, 5].

Several studies to explore and to explain the effect of covariates on dependent variable and other topics in regression analysis are conducted. [6] investigated the estimation of MLR parameters models when the original assumptions of OLS estimation are weak. Also, they introduced some MLR models with outliers and get conclusions. [7] stated that the common question is how to relate the response variable (Y) and the explanatory variables (X_i) by employing the analysis of regression. [8]

was developed statistical methods that are capable to detect outliers. He suggested a robust regression to analysis data with outliers. He also, defined performance of outliers in LR and compared some of robust methods using simulation studies. [9] used a genetic algorithm to determine a set of parameters that minimizes the prediction error for MLR model. [10] extended the econometric methods to GLMs to analysis the binary, count and duration response involved in social sciences and business. He found that these methods perform well in the applications for prediction and inference with high dimensional data. [11] used simulation studies to compared the performance of Lasso, Elastic net, Ridge Regression, and Bayesian models with high dimensional data in multivariate regression. They found that the Ridge Regression model was effective in estimating parameter accurately and control over the Type I error rate. [12] investigated the performance of regularization methods to analysis the high dimensional data with different sparse and non-sparse conditions. they studied the prediction, parameter estimation and variable selection properties.

Finally, [13] stated that there are two purposes of analyzing high-dimensional data are firstly to define the relationship between the covariates and response variable for scientific objectives. Secondly to develop effective techniques that can be predict the future observations accurately.

The goal of this paper is to propose nonparametric method (kernel regression (KR)) to estimate the MLR model with high dimensional data as an alternative method and compare its performance with OLS and LASSO methods.

The rest of the paper consists of Section 2 devoted to specify the model and to describe two methods of estimation. Section 3devoted to present the results of simulation study and analyzing real dataset. Section 4 devoted to discuss the results of the study. Finally some conclusion are placed in section 5.

2. Methodology

MLR model is defined as "a form of predictive modeling technique which investigates the relationship between a **dependent (response) variable** (Y) and **independent variables (covariates)** (X_i). This technique is used for forecasting, time series modeling and finding the fundamental effect relationship between the variables" [14]. Three fundamental components determined MLR model; the first component is the number of explanatory variables, the second component is the type of response variable and the third component is the shape of regression line. Meanwhile, the main goal from any data analysis is to get the correct estimation from raw data. Consequently, the important question is "if there is a statistical relationship between a response variable (Y) and explanatory variables (X_i)". Therefore, answering this question means that MLR modeling is conducted in order to present this relationship. MLR is also used to make inferences about the response variable (Y) based on values of a set of explanatory variables (X_i) [7].

Here, we use the MLR model where the response variable (Y) is a linear function of p explanatory variables (X_j) and a random error (U), this model with an intercept. MLR model of ith row of these variables is given by the following equation [15] [16]:

$$y_i = \beta_0 + \beta_1 x_{i2} + \beta_2 x_{i3} + \cdots + \beta_p x_{ip} + u_i \quad (1)$$

MLR model can be written as matrices notation by:

$$\boldsymbol{Y} = \boldsymbol{X\beta} + \boldsymbol{U} \quad (2)$$

where $\boldsymbol{Y}$ is a vector of $(n \times 1)$ observed response values, $\boldsymbol{X}$ is the matrix of the explanatory variables with $(n \times p)$ rank, $\boldsymbol{\beta}$ is the vector of unknown parameters with $(p \times 1)$, and $\boldsymbol{U}$ is the vector of random error terms with $(n \times 1)$rank as follow:

$$\boldsymbol{Y} = \begin{Bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{Bmatrix};\ \boldsymbol{X} = \begin{Bmatrix} 1 & x_{12} & \cdots & x_{1k} \\ 1 & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \cdots & \vdots \\ 1 & x_{n2} & \cdots & x_{nk} \end{Bmatrix};\ \boldsymbol{\beta} = \begin{Bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{Bmatrix};\ \boldsymbol{U} = \begin{Bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{Bmatrix}$$

The goal of regression analysis is to estimate the unknown parameters $\beta's$ and make inference about the future values of response variable.

If the assumptions of MLR model are holding the ordinary least squares (OLS) method is used to estimate the regression parameters (β) using the formula [16]:

$$\widehat{\boldsymbol{\beta}}_{ols} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{Y} \tag{3}$$

Consequently, the estimated MLR model is:

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_{i2} + \cdots + \hat{\beta}_k x_{ik}, \quad i = 1,2,\dots,n; \quad j = 0,1,2,\dots,k \tag{4}$$

The residuals calculated from a sample size n may be defined the OLS criterion from Eq. (4) as follow:

$$Min\ e^2 = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{5}$$

where ($y_i; \hat{y}_i$) are the observed and the estimated response variable for subject *i* respectively, as in Eq. (4). The minimization is performed with respect to the $(p+1)$ parameters $\{\beta_0;\ \beta_1; \dots; \beta_p\}$ with a constrain that $n > (p+1)$ this means the sample size is larger than the number of parameters to be estimated from that sample [4].

There are other important assumptions to obtain a valid estimation of MLR model as in Eq.(4). These assumptions are the error term follows a normal distribution and the error process u_i independent of all explanatory variables where [14]:

$$(u_i \sim N(0,\sigma^2)), \text{ and } \quad E\{u_i|X_i\} = 0; \quad E(u_i) = 0, \quad Cov(u_i, u_j) = \sigma^2 I_n \tag{6}$$

The other important assumption is $n > p$ which is our concern here because the size of the sample must be large to offer enough power for the test. Meanwhile, if the number of observations n, gets closer to or less than p, the number of covariates then there is more variability in the OLS fit [17]. Also, the matrix ($X'X$) is not invertible in this case therefore there is no unique solution for OLS regression [18]. Consequently, there may be some irrelevant variables included in the MLR model.

Hence, in the case of high dimensional data, when the sample size (n) becomes closer to or less then the number of explanatory variables (p). One of four groups of estimation methods can be used based on the validity of the assumptions of the MLR as mentioned above. the important question is which estimation method is appropriate to estimate the unknown parameters ($\beta's$) of this model in this case?. Consequently, the goal of this paper is to compare the performance of two estimation methods OLS and LASSO with KR method as a proposed method to estimate the parameters of MLR model in this case. The following subsections are consisting of a brief concepts for these methods.

2.1 *Regularization method (LASSO)*

When the sample size (n) in the model may be less than or equal to the number of explanatory variables (p), Least Absolute Shrinkage and Selection Operator (LASSO) is the appropriate method for estimating the parameters of MLR models [19]. Also, there are two important considerations when fitting high-dimensional data: model sparsity and prediction ability. Therefore, LASSO is defined as "a regularization methods for simultaneous estimation and variable selection" [20]. LASSO deals with many predictors may be $p \geq n$, and with an ill-conditioned model matrix X (i.e. $X'X$ is not invertible or near singular). However, LASSO is differing from other methods, because it is making the interpretation of the statistical model more plausible by sets many coefficient estimates exactly to zero. Regularization techniques work by introducing a penalty to the OLS estimator as in Eq. (5). Then, LASSO method of estimation is introduced by [21] and is defined as follows:

$L(\beta,\lambda) = \|Y - X\beta\|_2^2 + \lambda\|\beta\|_1$

$$= \sum_{i=1}^{n}(y_i - \hat{y}_i)^2 + \lambda\sum_{j=1}^{k}|\beta_j|$$

$$\hat{\beta}_{Lasso} = \underbrace{argmin}_{\beta}\{\|Y - X\beta\|^2 + \lambda\|\beta\|_1\} \tag{7}$$

The notations in Eq. (7) are defined as in Eq. (1) where the estimated OLS model of MLR contains the cofficients of interest [20]. Eq.(7) can be solved by using quadratic programming techniques such as a coordinate gradient descent algorithm [22], where it becomes as:

$$\hat{\beta}_{Lasso} = \underbrace{argmin}_{\beta} \{(y_i - \hat{y}_i)^2\} \quad (8)$$

Subject to $\sum_{j=1}^{k} |\hat{\beta}_j| \leq s$

the parameter $\lambda \geq 0$ represents the degree of the coefficients of the model which have small weighted or removed from the model. Therefore, larger λ values means greater shrinkage, and LASSO is converted to the OLS estimator when $\lambda = 0$. The optimal value of λ is estimated by using Jackknife cross-validation, which is described in [21]. LASSO method gives optimal values of the $\hat{\beta}_j$ depending on the importance of the variable; where the greatest importance explanatory variables receive higher values, and the smallest importance are allocated coefficients at or near 0 [11]. [19] stated that LASSO method has good experimental and theoretical characteristics for estimation and variable selection. In spite of the LASSO has shown success in many situations, it has some limitations [21] [23]. Therefore we propose KR method to overcome this limitations.

2.2 *Kernel Regression (KR) Model*

Kernel regression (KR) model has been considered as one of the efficient nonparametric regression models, it is proposed by Nadaraya and Watson at 1964. KR method is particularly powerful in high-dimensional and nonlinear settings [24]. In practice, KR method regresses the dependent variables Y onto the similarity of independent variables measured through the kernel function. KR model is weighted average estimators that use kernel functions as weights. Suppose we have the sample of observations $(y_i, x_i)_{i=1}^{n}$, then, the Kernel estimation of $f(y)$ based on the sample is [25]:

$$f_n(y) = Pr_n(Y = y|x_i) = \frac{1}{n}\sum_{i=1}^{n} \mathrm{K}\,(Y = y|x_i) \quad (9)$$

where $\mathrm{K}\,(Y = y|x_i)$ is one of Kernel function given x_i observation. Two choices must be made (the kernel function ($K(.)$ and the smoothing parameter (h)) when working with a kernel estimator. On the other hand, given a choice of kernel K(.) , and a smoothing parameter h, Kernel regression model for one dimension is [26]:

$$\hat{g}(y|x) = \sum_{i=1}^{n} w(x, x_i)\, y_i \quad (10)$$

$$w(x, x_i) = \frac{K\left(\frac{x_i - x}{h}\right)}{\sum_{i=1}^{n} K\left(\frac{x_i - x}{h}\right)}$$

where $w(x, x_i)$ is the weight of Kernel. Because these weights are smoothly varying with x, the kernel regression estimator $\hat{g}(y|x)$ itself is also smoothly varying with x. The selection of $K(.)$ can be easily changed to support of the density to be estimated. Also, the selection of appropriate smoothing parameter h is very important, because of the effect of the h on the shape of the corresponding estimator. When h value is small, we will obtain an undersmoothed estimator, with high variability. Whereas, if the value of h is large, the resulting estimator will be oversmoothed. In practice, we tend to select h by one of two methods; guessing it or using cross-validation method.

The multiple nonparametric regression (MNR) model is written as [27]:

$$y_i = f(x_i^T) + \varepsilon_i$$
$$= f(x_{i1}, x_{i2}, \dots, x_{ip}) + \varepsilon_i \quad (11)$$

where the function f is left unspecified. Moreover, the object of nonparametric regression is to estimate the regression function $f\,(\cdot)$ directly, rather than to estimate parameters. it is difficult to fit the MNR model when there are many predictors, several models have been developed one of them is the additive nonparametric regression (ANR) model as:

$$y_i = \alpha + f_1(x_{i1}) + f_2(x_{i2}) + \cdots + f_p(x_{ip}) + \varepsilon_i \qquad (12)$$

where $f_j(\cdot)$ is representing the partial-regression functions which are assumed smooth, and to be estimated from the data. These partial-regression functions $f_j(\cdot)$ are fitted using one of the simple-regression smoother, such as local polynomial regression.

Therefore, the first step to fit this model is to define a multiple neighborhood around the original point $\left(x_0^T = (x_{01}, x_{02}, \ldots, x_{0p})\right)$ and the second step is to calculate the weight of these points by using the scaled Euclidean distances as a default in LOESS function, where,

$$D(x_i, x_0) = \sqrt{\sum_{j=1}^{p} (z_{ij} - z_{0j})^2},$$

where $z_{ij} = (x_{ij} - \bar{x}_j)/S_j$, and $\bar{x}_j \; and \; S_j$ are the mean and the standard deviation of the jth predictor. Then, the scaled distances is:

$$w_i = W\left[\frac{D(x_i, x_0)}{h}\right]$$

where $W\ (\cdot)$ is an appropriate weight function, such as the tricube and Epanechnikov and others. The third step is constructing a weighted polynomial regression of y on the $x's$; such as, a local linear fit takes the following form:

$$y_i = \alpha + \beta_1(x_{i1} - x_{01}) + \beta_2(x_{i2} - x_{02}) + \cdots + \beta_p(x_{ip} - x_{0p}) + \varepsilon_i \qquad (13)$$

Then the estimated value at x_0 is then simply by $\hat{y} = \alpha$. Performance of kernel is measured by many criteria among them MSE and RMSE as presented in the next subsection.

2.3 *Assessing Criteria of the Estimation Methods.*

There are some useful measures (criteria) to assess the adequacy of the model to the data when we have several alternative methods to estimate the coefficients of the model. The average of mean square error (AMSE)is the first criterion of goodness of fit which is used to select the most appropriate model [28]:

$$MSE_m = \frac{1}{n-k-1}\left(\sum_{i=1}^{n}(y_i - \hat{y}_i)^2\right), \qquad m = 1,2,\ldots,M \qquad (14)$$

Since we repeat the dataset for M times in the simulation study then we will use the average mean square error as:

$$AMSE = 1 \,/\, M\left(\sum MSE_m\right) \qquad where\ m = 1,2,\ldots,M \qquad (15)$$

where M is the number of replications. Almost, the number of observations (n) is less than or equal the number of explanatory variables in the case of high dimensional data. Therefore, the value of AMSE is larger than its value in the other case of fitting MLR model. meanwhile, MSE for the KR model is [29]:

$$MSE(x) = var\,\{\hat{g}_n(x)\} + bias^2\,\{\hat{g}_n(x)\}, \qquad x \in N \qquad (16)$$

We will use these criteria to assess the goodness of fit the estimation method as in the following section.

3. Applying Estimation Methods

The proposed method and other estimation methods are applied using simulation study and analyzing the real dataset to compare the performance of them. The results are as in the following subsections.

3.1 *Simulation Study*

The first variable generated is the dependent variable by using the following model:

$$y_i = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_{11} x_{i20} + \varepsilon_i \qquad (17)$$

The regression coefficients $(\beta_j)$are assumed the first five parameters of the first ten equal to one and the rest five equal to zero and the first five of the second ten parameters are equal to one and the rest five equal to zero. Three levels of σ_ε^2 are used (1, 5, and 9). The first ten explanatory variables (X_j) are sampled independently from uniform distribution U(0,1) and the second ten explanatory variables from standard normal distribution N(0,1). The data sets are generated under the sample sizes (15, 20 and 25) and the experiment was repeated 1000 times. The program code for the data simulation was adapted in SAS and XLSTAT software. Consequently, the values of the Root average of mean square error (RAMSE) for different methods, different sample sizes and different values of variance of the error are as in Table (1):

Table (1): RAMSE for different methods of estimation, different sample sizes and different values of variance of error

criteria	Method	n=15			n=20			n=30		
	σ_ε^2	$=1$	$=5$	$=9$	$=1$	$=5$	$=9$	$=1$	$=5$	$=9$
RAMSE	**OLS**	***	***	***	***	***	***	1.885	2.038	2.227
	LASSO	***	***	***	***	***	***	1.240	2.116	2.281
	KR	4.247	5.321	6.119	5.903	6.743	7.033	6.598	6.769	7.297

(***)The models when $n \leq p$ are not fitted because model is not full rank.

Meanwhile, the results of AMSE for different methods, different sample sizes and different values of variance of error are as in Table (2).

Table (2): AMSE for different methods of estimation and different sample sizes and different values of variance of error

criteria	Method	n=15			n=20			n=30		
	σ_ε^2	$=1$	$=5$	$=9$	$=1$	$=5$	$=9$	$=1$	$=5$	$=9$
	OLS	***	***	***	***	***	***	1.4126	4.1538	4.9602
AMSE	**LASSO**	***	***	***	***	***	***	1.5363	4.4778	5.2005
	KR	20.26	28.316	37.441	34.841	45.473	49.461	43.527	45.82	53.244

(***)The models when $n \leq p$ are not fitted because model is not full rank.

3.2 *Analyzing of Real Data*

Almost newborns come with healthy weight. Most of them born after 37 or 40 weeks weigh between 2.5 Kg and 4.0 Kg. Newborns who are lighter or heavier than the average baby are usually fine. Therefore, a simple random samples with sizes (15, 20,30) consists of all mothers who visit the primary health care center in Babylon province in year 2018, are drawn to compare the performance of the estimation methods presented above. The weight (Kg) of newborn children is a response variable Y which is considered as an indicator of healthy generation, while the risk factors are as following:

mother age (year); Age at marriage (year); Educational attainment of mother; Educational attainment of Husband; Weight of mother (Kg), Contraception Using; Mother's smoking (yes; No); Age of Husband (year); Job of Husband (yes, No); Period of Marriage (year); Number of born children (#); Period of exercise per week (hour); Thyroid disease (yes, no); Mother's sleeping per day (hour); Taking medications (yes, No), Breastfeeding Duration (Month), Mother's job (yes, No); number of dead children (#). The weight of newborns (Kg) as the dependent variable. When we apply OLS, LASSO and KR methods to fit the MLR model for this dataset then the results of assessing criteria are as in Table (3):

Table (3) values of assessing criteria for different samples and different methods

Method	**n=15**			**n=20**			**n=30**		
criteria	MSE	RMSE	R^2	MSE	RMSE	R^2	MSE	RMSE	R^2
OLS	***	***	***	***	***	***	0.262	0.512	0.92
LASSO	***	***	***	0.129	0.3595	0.93	0.3025	0.550	0.48
KR	4.554	2.134	0.67	5.252	2.292	0.50	7.028	2.651	0.33

(***)The models when $n \leq p$ are not fitted because model is not full rank..

4. Discussing the Results

computational and statistical challenges have been introduced by high dimensional data especially in estimating MLR models. High dimensionality brings noise accretion, false correlations and related heterogeneity. These problems of high dimensional data make traditional statistical methods invalid. The results in tables 1 and 2 show that the OLS regression and LASSO cannot fit the MLR model when the sample size $n = 15\ or\ 20$ is small than or equal to the number of explanatory variables $p = 20\ (i.e\ n \leq p)$ because the models are not full rank, OLS solutions for the parameters are not unique and some statistics will be misleading. Meanwhile the KR method gives an accepted criteria of RAMSE and AMSE in these cases. But when the sample size become $n = 30$ with the same number of explanatory variables the results show an improvement of the results of traditional method OLS to give a less values of the criteria. These results show that KR is outperformance of OLS and LASSO methods to fit the MLR with high dimensional data. Also, analyzing real dataset confirms that KR gives better results than other methods with high dimensional data when ($n = 15\ or\ 20$) and the number of covariates equal to $(p = 18)$. Finally all the results of simulation study and analyzing real data show that KR is powerful than other method in estimating MLR with high dimensional data.

5. Conclusions

The above results show that KR method is better than traditional method (OLS) and the Least LASSO to estimate the predicted values of response variable of MLR model with high dimensional data because it has smallest values of RAMSE and AMSE and gives a predicted values of response variable; mean while the other method cannot fit the MLR model when the sample size less than or equal to the number of explanatory variables. Also, the results from analyzing the real dataset are confirming the results of simulation study that KR is the preferred method than others. For the future works may be compare the performance of KR method with the Principal Component and Ridge regression methods to estimate the coefficients of MLR with high dimensional data. Also, may be compare the performance of KR method with other nonparametric methods (e.g. spline method).

References

[1] Fujiwara M, Minamidani T, Nagai I and Wakaki, H 2012 Principal Components Regression by using Generalized Principal Components Analysis, retrieved on 10-6-2019 from: www.math.sci.hiroshima-u.ac.jp/stat/TR/TR12/TR12-09.pdf.

[2] Hastie T, Tibshirani R and Friedman, J 2009 *The elements of statistical learning: Data mining, inference, and prediction.* Springer-Verlag, New York.

[3] Finch W H, Hernandez Finch M E and Moss L 2014 Dimension reduction regression techniques for high dimensional data, *General Linear Model Journal* 40(2) *1-15.*

[4] Finch W H and Hernandez Finch M E 2017 Multivariate Regression with Small Samples: A Comparison of Estimation Methods. *General Linear Model Journal* 43(1) *16-30.*

[5] Lee H, Park Y M and Lee S 2015 Principal Component Regression by Principal Component Selection. *Communications for Statistical Applications and Methods* 22(2) 173–180.

[6] Cankaya S, Kaylaap GT, Sangun L, Tahtali Y and Akar M 2006 A Comparative Study of Estimation Methods for Parameters in Multiple Linear Regression Model. *J. Appl. Anm. Res* 29 43-47.

[7] Alxeopoulos E C 2010) Introduction to Multivariate Regression Analysis, *Hippokratia Journal* 14(1) 23-28.
[8] Alma Ö G 2011 Comparison of Robust Regression Methods in Linear Regression, *Int. J. Contemp. Math. Sciences* 6(9) 409 – 421
[9] Stoica F, and Boitor C G 2014 Using the Breeder GA to Optimize a Multiple Regression Analysis Model used in Prediction of the Mesiodistal Width of Unerupted Teeth, *Int. J. Comput. Commun.*, 9(1) 62-70.
[10] Sapra S 2016 Econometric Modeling with High-dimensional Data in Business and Economics, *The MT International Conference on Business Research, the macro Conference Proceedings, Milano*, 62- 69
[11] Finch W H and Hernandez Finch M E 2016 Regularization Methods for Fitting Linear Models with Small Sample Sizes: Fitting the Lasso Estimator using R. Practical Assessment, *Research & Evaluation.* 21(7) 1-13.
[12] Sirimongkolkasem T and Drikvandi R 2019 On Regularization Methods for Analysis of High Dimensional Data, *Annals of Data Science,* 6(4) 737–763.
[13] Fan J, Han F and Liu H 2014 Challenges of Big Data analysis, *National Science Review* , 1: 293–314.
[14] Ray S 2015 7 types of regression should be Know. retrieved on 15- 5-2019 from, *https://www.analyticsvidhya.com/*
[15] Januaviani T MA, Gusriani N, Joebaedi K, Sukono, Subiyanto and Bon AT 2019 The LASSO (Least Absolute Shrinkage and Selection Operator) Method to Predict Indonesian Foreign Exchange Deposit Data, *Proceedings of the International Conference on Industrial Engineering and Operations Management Bangkok, Thailand*, March 5-7.
[16] AlNasser H 2017 On Ridge Regression and Least Absolute Shrinkage and Selection Operator, *M.SC. thesis, University of Victoria*, USA.
[17] James G, Witten D, Hastie T and Tibshirani R 2013 *An Introduction to Statistical Learning: With Applications in R*, Springer Publishing Company, New York.
[18] Hastie T, Tibshirani R, and Wainwright M 2015 *Statistical Learning with Sparsity: The Lasso and Generalizations,* Taylor Francis Group, Florida, USA.
[19] Jankova J, Shah R D, Buhlmann P and Samworth R J 2019 Goodness-of-Fit testing in high-dimensional generalized linear models, retrieved at 5th Feb 2020 from https://arxiv.org/pdf/1908.03606.pdf.
[20] Lambert-Lacroix S and Zwald L 2011 Robust regression through the Huber's criterion and adaptive lasso penalty, *Electronic Journal of Statistics,* 5, 1015–1053.
[21] Tibshirani R 1996 Regression shrinkage and selection via the lasso, *Journal of the Royal Statistical Society, Series B., 58*, 267-288.
[22] Lopes H F and Marques P 2018 Multiple linear regression, retrieved at 11th Jan 2020, http://hedibert.org/wp-content/uploads/2018/05/pdf
[23] Zou H and Hastie T 2005 Regularization and variable selection via the elastic net, *J. R. Statist. Soc.* B , 67(2), 301–320
[24] Du P, Parmeter C F and Racine J S 2013 Nonparametric Kernel Regression with Multiple Predictors and Multiple Shape Constraints, *Statistica Sinica* 23, 1347-1371.
[25] Marsh L C and Mukhopadhyay K 2014 Count Regression Model Using Kernel Regression, *paper of conference, Researchgate, uploaded by Lawrence C. Marsh on 31 May 2014*
[26] Hardle W 1994 *Applied Nonparametric Regression*, Institute for Statistics and Econometrics, Berlin.
[27] Fox J 2000 *Multiple and Generalized Nonparametric Regression*, Thousand Oaks CA: Sage.
[28] Sprent P and Seemton N C 2001 *Applied nonparametric statistical methods*, 3rd ed., Chapman & Hall/CRC, UK.
[29] Kokonendji C, Kiesse TS, Clarice G B and Demétrio, C G B 2009 Appropriate kernel regression on a count explanatory variable and applications, *Advances and Applications in Statistics,* 12(1), 99-125.

Variable selection in Gamma regression model using binary gray Wolf optimization algorithm

Niam Abdulmunim Al-Thanoon[1] , Omar Saber Qasim[2], Zakariya Yahya Algamal[3]

1 Department of Operations Research and Artificial Intelligence, University of Mosul, Mosul, Iraq
E-mail: niam.munim@uomosul.edu.iq
2 Department of Mathematics, University of Mosul, Mosul, Iraq
E-mail: omar.saber@uomosul.edu.iq
3 Department of Statistics and Informatics, University of Mosul, Iraq.
E-mail address: zakariya.algamal@uomosul.edu.iq

Abstract: In the real life applications, large amounts of variables have been accumulated quickly. Selection of variables is a very useful tool for improving the prediction accuracy by identifying the most relative variables that related to the study. Gamma regression model is one of the most models that applied in several science fields. Gray Wolf optimization algorithm (GWO) is one of the proposed nature-inspired algorithms that can efficiently be employed for variable selection. In this paper, chaotic GWO is proposed to perform variable selection for gamma regression model. The simulation studies and a real data application are used to evaluate the performance of our proposed procedure in terms of prediction accuracy and variable selection criteria. The obtained results demonstrated the efficiency of our proposed methods comparing with other popular methods.

Keyword: Variable selection; gamma regression model; gray Wolf optimization algorithm.

1. Introduction

Gamma regression model is widely applied method for studying automobile insurance claims and medical science (De Jong & Heller, 2008; Dunder, Gumustekin, & Cengiz, 2016a; Malehi, Pourmotahari, & Angali, 2015). Specifically, when the response variable under the study is distributed as gamma distribution (Al-Abood & Young, 1986; Hattab, 2016).

In many real applications, recent developments in technologies have made the possibility to measure a large number of variables. In the regression modeling, the existence of huge number has a negative effect by overfitting the regression model. Therefore, identification of a small subset of important variables from a large number of variables set for accurate prediction is an important role for building predictive regression models (Zakariya Yahya Algamal & Muhammad Hisyam Lee, 2015).

When the number of variables increases, the traditional variable selection methods, such as stepwise selection, forward selection, and backward elimination computationally become an exhaustive search and require a long time for computing. Penalization methods, (lasso) (Tibshirani, 1996), (scad) (Fan & Li, 2001), elastic net (Zou & Hastie, 2005), and adaptive lasso (Zou, 2006), are become an attractive methods for simultaneously performing variable selection and model estimation.

Recently, the naturally inspired algorithms, such as genetic algorithm, particle swarm optimization algorithm, firefly algorithm, and Gray Wolf optimization algorithm, have a great attraction and proved their efficiency as variable selection methods (Sayed, Hassanien, & Azar, 2017). This is because that the main target in variable selection is to minimize the number of selected variables while maintaining the maximum accuracy of prediction, and, therefore, they can be considered as optimization problems (Sindhu, Ngadiran, Yacob, Zahri, & Hariharan, 2017).

Several researchers have employed the naturally inspired algorithms for variable selection in regression models. Broadhurst, Goodacre, Jones, Rowland, and Kell (1997) employed the genetic algorithm for variable selection in linear regression models, with application in chemometrics. Drezner, Marcoulides, and Salhi (1999) proposed to use tabu search algorithm in model selection in the linear regression model. On the other hand, a hybrid algorithm of genetic algorithm and simulated annealing was proposed as a subset selection method in linear regression model by Örkcü (2013). Brusco (2014) performed a comparison of simulated annealing algorithms for variable selection in principal component analysis and discriminant analysis. Besides, the differential evolution algorithm was used as a variable selection in linear regression model by Dünder, Gümüştekin, Murat, and Cengiz (2017). In generalized linear models, the natural inspired algorithms for variable selection are also used, such as, logistic regression model (Pacheco, Casado, & Núñez, 2009; Unler & Murat, 2010), Poisson regression model (Koç, Dünder, Gümüştekin, Koç, & Cengiz, 2017; Massaro & Bozdogan, 2015), and gamma regression model (Dunder, Gumustekin, & Cengiz, 2016b).

The purpose of this paper is to propose chaotic GWO, which is a swarm intelligence technique, as an alternative variable selection method for use in gamma regression model. The proposed algorithm will efficiently help in identifying the most relevant variables in the count data regression model with a high prediction. The superiority of the proposed algorithm is proved though different simulation settings and a real data application.

2. Gamma regression model

In epidemiology, social, and economic studies, positively skewed data are often arisen. Gamma distribution is a well-known distribution that fits such type of data. Gamma regression model (GRM) is used to model the relationship between the non-negative skewed response variable and potentially variables (Uusipaikka, 2009).

Assume y_i is the response variable which is following a gamma distribution with shape parameter ν and scale parameter γ, i.e. $y_i \sim Gamma(\nu,\gamma)$, then the probability density function is defined as

$$f\,(y_i) = \frac{\gamma}{\Gamma\nu}(\gamma y_i)^{\nu-1}e^{-\gamma y_i}, \quad y_i \geq 0, \tag{1}$$

with $E(y) = \nu/\gamma = \theta$ and $\mathrm{var}(y) = \nu/\gamma^2 = \theta^2/\nu$. Given that $\gamma = \nu/\theta$, Eq. (3) can re-parameterized as a function of the mean (θ) and the shape (ν) parameters and written depending on the exponential function as

$$f\,(y_i) = EXP\left\{\frac{y_i(-1/\theta) - \log(-1/\theta)}{1/\nu} + c(y_i,\nu)\right\}, \tag{2}$$

where the canonical link function is $-1/\theta$, the dispersion parameter is $\phi = 1/\nu$ and $c(y_i,\nu) = \nu\log(\nu) + \nu\log(y_i) - \log(y_i) - \log(\Gamma(\nu))$.

GRM is usually modeled using the canonical link function (reciprocal), $\theta_i = -1/\mathbf{x}_i^T\boldsymbol{\beta}$ which is expressed as a linear combination of covariates $\mathbf{x}_i = (x_{i1},...,x_{ip})^T$. The log link function, $\theta_i = \exp(\mathbf{x}_i^T\boldsymbol{\beta})$, is alternatively used rather than the reciprocal link function because it ensures that $\theta_i > 0$.

The maximum likelihood method of Eq. (4) is the most common method of estimating the coefficients of GRM. Assuming that the observations are independent and $\theta_i = -1/\mathbf{x}_i^T\boldsymbol{\beta}$, the log-likelihood function is given by

$$\ell(\boldsymbol{\beta}) = \sum_{i=1}^{n} \left\{ \frac{y_i \mathbf{x}_i^T \boldsymbol{\beta} - \log(\mathbf{x}_i^T \boldsymbol{\beta})}{1/\nu} + c(y_i, \nu) \right\}, \quad (3)$$

the ML estimator is then obtained by computing the first derivative of the Eq. (3) and setting it equal to zero, as

$$\frac{\partial \ell(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} = \frac{1}{\nu} \sum_{i=1}^{n} \left[y_i - \frac{1}{\mathbf{x}_i^T \boldsymbol{\beta}} \right] \mathbf{x}_i = 0. \quad (4)$$

Depending on the iteratively weighted least squares (IWLS) algorithm, in each iteration, the parameters are updated by

$$\boldsymbol{\beta}^{(r+1)} = \boldsymbol{\beta}^{(r)} + I^{-1}(\boldsymbol{\beta}^{(r)}) S(\boldsymbol{\beta}^{(r)}), \quad (5)$$

where $S(\boldsymbol{\beta}) = \partial \ell(\boldsymbol{\beta}) / \partial \boldsymbol{\beta}$ and $I^{-1}(\boldsymbol{\beta}) = \left(-E \left(\partial^2 \ell(\boldsymbol{\beta}) / \partial \boldsymbol{\beta} \partial \boldsymbol{\beta}^T \right) \right)^{-1}$. The final step of the estimated coefficients is defined as

$$\hat{\boldsymbol{\beta}}_{GR} = (\mathbf{X}^T \hat{\mathbf{W}} \mathbf{X})^{-1} \mathbf{X}^T \hat{\mathbf{W}} \hat{\mathbf{u}}, \quad (6)$$

where $\hat{\mathbf{W}} = \text{diag}(\hat{\theta}_i^2)$ and $\hat{\mathbf{u}}$ is a vector where ith element equals" to $\hat{u}_i = \hat{\theta}_i + ((y_i - \hat{\theta}_i) / \hat{\theta}_i^2)$.

3. Chaotic grey wolf optimization algorithm

Mirjalili, Mirjalili, and Lewis (2014) presented a new metaheuristics algorithms as a swarm intelligence, "which is known as the grey wolf optimizer (GWO) algorithm. The GWO simulate the behavior of leadership and hunting in organisms of grey wolf. The GWO simulates the driving hierarchy in the environment and this distinguishes it from the rest of the swarm algorithms. The simulation of hunting in the GWO algorithm is done through the hierarchy of leadership, where the crowd is divided into different groups and levels such as alpha, beta, and omega (Mirjalili et al., 2014).

Gray wolves belong to the Canidae family and are classified as top predators because they belong to the top of the food chain. The first level of the leadership hierarchy is the alpha (α) type and they represent the leaders, they may be female or male, and they are responsible for making all the decisions related to hunting, sleep, time to wake and so on . The second level in the hierarchy of leadership is the beta (β), where these wolves are helping wolves in the first level of the alpha in making decisions. Wolves in the second level (β) respect wolves in the first level (α) and reinforce decision-making and act as their consultant. In the third level, there is a type of omega (γ) and plays the role of scapegoat for the flock. All wolves from other levels are submitted to wolves of the omega type. It may seem that wolves in the third level are not an important person, but it is observed that the group without them face fighting and internal problems. This is due to the venting of vehemence and frustration of all wolves by the omega (γ). This helps in fulfilling the whole pack and preserve the dominance structure (Singh & Hachimi, 2018). Wolves, which are not alpha (α), beta (β), or omega (γ), are called the subordinate or delta (δ), and wolves in this species must be subjugated to alpha (α) and beta (β), but they dominate the omega (γ) wolves.

Mathematical models for each level of the leadership pyramid of the GWO are calculated through the following:

$$\overrightarrow{D} = \left| \vec{c} . \vec{x}_{p(t)} - \vec{x}(t) \right|, \quad (7)$$

$$\vec{x}(t+1)=\vec{x}_{p(t)}-\vec{a}.\vec{d}, \quad (8)$$

where t shows the current iteration, $\vec{x}_p$ indicates the position vector of the prey, $\vec{x}$ represent the position vector of a grey wolf. The vectors $\vec{a}$ and $\vec{c}$ are defined mathematically as follows:

$$\begin{aligned}\vec{a}&=2\vec{L}.\vec{r}_1-\vec{L},\\ \vec{c}&=2.\vec{r}_2,\end{aligned} \quad (9)$$

where the components of $\vec{L}$ are linearly reduced from 2 to 0 over the course of iterations and $\vec{r}_1, \vec{r}_2$ are random vectors in $[0,1]$(Li et al., 2017).

3.1.1 Hunting

There are three main steps that are applied during hunting prey. There are: (1) the search for prey, (2) encircling, and, (3) attacking. The mathematical behavior of the gray wolf algorithm is simulated by assuming that alpha (α), beta (β), and delta (δ) have potential knowledge of the prey location. Mathematical equations in this regard are developed by

$$\overrightarrow{D_\alpha}=\left|\vec{c}_1.\overrightarrow{x_\alpha}-\vec{x}\right|, \quad \overrightarrow{D_\beta}=\left|\vec{c}_2.\overrightarrow{x_\beta}-\vec{x}\right|, \quad \overrightarrow{D_\delta}=\left|\vec{c}_3.\overrightarrow{x_\delta}-\vec{x}\right|, \quad (10)$$

$$\vec{x}_1=\overrightarrow{x_\alpha}-\vec{a}_1.(\overrightarrow{D_\alpha}), \quad \vec{x}_2=\overrightarrow{x_\beta}-\vec{a}_2.(\overrightarrow{D_\beta}), \quad \vec{x}_3=\overrightarrow{x_\delta}-\vec{a}_3.(\overrightarrow{D_\delta}), \quad (11)$$

$$\frac{\vec{x}_1+\vec{x}_2+\vec{x}_3}{3}, \quad (\mathbf{12})$$

$$\begin{aligned}\vec{a}_{(.)}&=2\vec{L}.\vec{r}_1-\vec{L},\\ \vec{c}_{(.)}&=2.\vec{r}_2,\end{aligned} \quad (13)$$

where $\vec{a}$ is a random value in the interval $[-\vec{L},\vec{L}]$. The gray wolves are compelled to attack the prey when random value $a<1$. The prey is searched through exploration ability and attack prey the ability to exploit. The arbitrary values of L are utilized to force the search to move away from the prey (E Emary, Zawbaa, & Grosan, 2018). The arbitrary values of L are applied to force the search to move away from the prey.

The positions of gray wolves are continuously changing in space to whatever point. In some problems such as feature selection, solutions are limited to binary 0 or 1 values. In such case, BGWO is proposed by Eid Emary, Zawbaa, and Hassanien (2016). The wolves update equation is a function of three position vectors namely $x_\alpha, x_\beta, x_\delta$ which can attract each wolf of the flock towards the first three best solutions. In any given time, the aggregation of solutions is in binary form and all the solutions are on the corner of a hypercube. To update the positions of the given wolf based on the basic GWO algorithm, while keeping the binary restriction according to the Eq. (14).

The main updating equation in the bGWO algorithm can be formulated, in this approach as follows (Eid Emary et al., 2016):

$$x_i^{t+1} = crossover(x_1, x_2, x_3), \quad (14)$$

where $crossover(x, y, z)$ is a suitable crossover between solutions x, y, z and x_1, x_2, x_3 are binary vectors representing the effect of a wolf in bGWO, which move towards the alpha; beta; delta gray wolves in order. x_1, x_2, x_3 are calculated using Eqs. (15), (18), and (21),respectively, as

$$x_1^d = \begin{cases} 1 & if \quad (x_\alpha^d + bstep_\alpha^d) \geq 1 \\ 0 & O.W \end{cases}, \quad (15)$$

where x_α^d represents the position vector of the alpha (α) wolf in the dimension d, and $bstep_\alpha^d$ is a binary step in the dimension d which is calculated by the following equation:

$$bstep_\alpha^d = \begin{cases} 1 & if \quad cstep_\alpha^d \geq rand \\ 0 & O.W \end{cases}, \quad (16)$$

where $rand$ is a random number derived from the uniform distribution in the closed period $[0,1]$, and $cstep_\alpha^d$ is the continuous-valued step size for dimension d and can be calculated using the sigmoidal function through Eq. (17).

$$cstep_\alpha^d = \frac{1}{1 + e^{-10\,(a_1^d D_\alpha^d - 0.5)}}, \quad (17)$$

where a_1^d and D_α^d are calculated using Eqs. (9) and (10) in the dimension d.

$$x_2^d = \begin{cases} 1 & if \quad (x_\beta^d + bstep_\beta^d) \geq 1 \\ 0 & O.W \end{cases}, \quad (18)$$

where x_β^d represents the position vector of the beta (β) wolf in the dimension d, and $bstep_\beta^d$ is a binary step in the dimension d which is calculated by the following equation:

$$bstep_\beta^d = \begin{cases} 1 & if \quad cstep_\beta^d \geq rand \\ 0 & O.W \end{cases}, \quad (19)$$

where $rand$ is a random number derived from the uniform distribution in the closed period $[0,1]$, and $cstep_\alpha^d$ is the continuous-valued step size for dimension d and can be calculated using the sigmoidal function through Eq. (20)

$$cstep_\beta^d = \frac{1}{1 + e^{-10\,(a_1^d D_\beta^d - 0.5)}}, \quad (20)$$

where a_1^d *and* D_β^d are calculated using equations (9), and (10) in the dimension d .

$$x_3^d = \begin{cases} 1 & if \quad (x_\delta^d + bstep_\delta^d) \geq 1 \\ 0 & OW \end{cases}, \qquad (21)$$

where x_δ^d represents the position vector of the delta (δ) wolf in the dimension d , and $bstep_\delta^d$ is a binary step in the dimension d which is calculated by the following equation:

$$bstep_\delta^d = \begin{cases} 1 & if \quad cstep_\delta^d \geq rand \\ 0 & \text{O.W} \end{cases}, \qquad (22)$$

where $rand$ is a random number derived from the uniform distribution in the closed period $[0,1]$, and $cstep_\delta^d$ is the continuous-valued step size for dimension d and can be calculated using the sigmoidal function through Eq. (23)

$$cstep_\delta^d = \frac{1}{1+e^{-10\,(a_1^d D_\delta^d - 0.5)}}, \qquad (23)$$

where a_1^d and D_δ^d are calculated using Eqs. (9) and (10) in the dimension d . The crossover process is then applied to each of the solutions $a;b;c$ as shown in the following equation:

$$x_d = \begin{cases} a_d & if \quad rand < 1/3 \\ b_d & if \quad 1/3 \leq rand < 2/3, \\ c_d & \quad OW \end{cases} \qquad (24)$$

where a_d, b_d *and* c_d represent the binary values for the first, second and third parameter in dimension d , x_d is the crossover process output at dimension d , and $rand$ is a random number derived from the uniform distribution in the closed period [0, 1]. The fitness function is defined" as

$$\text{fitness} = \min\left[\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2\right]. \qquad (25)$$

4. Computational results

In this section, the performance of our proposed variable selection method, CGWO is tested. Further, the performance of CGWO is compared with the GWO, Bayesian information criteria (BIC), and Akaike information criteria (AIC) that are defined as, respectively,

$$\text{AIC} = -2\ell(\hat{\boldsymbol{\beta}}) + 2\times q, \qquad (26)$$

$$\text{BIC} = 2\ell(\hat{\boldsymbol{\beta}}) + \log(n)\times q, \qquad (27)$$

where $\ell(\hat{\boldsymbol{\beta}})$ is the log-likelihood for PRM and q is the number of selected variables.

5. Simulation results

In this section, the same simulation settings of Zakariya Y Algamal and Muhammad H Lee (2015) and Wang et al. (2014) are used. The sample size is considered with $n \in \{50, 100, 200\}$.

Simulation 1: In this simulation, 20 explanatory variables are generated from multivariate normal distributions with mean vector $\mathbf{0}$ and covariance matrix $\mathbf{\Sigma}$ which elements $\rho(x_i, x_j) = \rho^{|i-j|}$ with $\rho = 0.5$. The true vector of parameters is given by $\boldsymbol{\beta} = (1.5, -1.5, 1.5, -1.6, 1.5, \underbrace{0, ..., 0}_{5}, 1.5, -1.5, 1.5, -1.6, 1.5, \underbrace{0, ..., 0}_{5})^T$ with 10 true explanatory variables and the rest in non-true variables.

Simulation 2: Here, The true vector of parameters is given by $\boldsymbol{\beta} = (1.6, -0.88, 0.95, -1.10, 0.70, \underbrace{0, ..., 0}_{15})^T$ with 5 true explanatory variables and 15 non-true variables. The explanatory variables are generated as same as simulation 1 with $\rho(x_i, x_j) = 0.5$.

Simulation 3: In this simulation, 8 explanatory variables are generated as same as simulation 1 with $\rho(x_i, x_j) = 0.5^{|i-j|}$. The true parameter vector is given by $\boldsymbol{\beta} = (\underbrace{0.25, ..., 0.25}_{8})^T$.

For all the simulation examples 1 – 3, the response variable is generated according to PRM as $y_i \sim Gamma(\exp(\mathbf{x}_i^T \boldsymbol{\beta}), 0.5)$. For performance evaluation of the CGWO, the mean squared error (MSE) is used as a prediction accuracy criteria, which is defined as $\sum_{i=1}^{n} (y_i - \hat{y}_i)^2 / n$. In terms of variable selection performance, the number of the truly nonzero coefficients which are incorrectly set to zero (I), and the number of the true zero coefficients which are correctly set to zero (C). The higher the values of C, and the lower the values of I, the better the variable selection performance is. All computations of this paper were conducted using R. Based on 300 times of repeating simulation, the averaged MSE, I, and C are listed in Tables 1 – 3, respectively.

It shows from these tables that the CGWO method there has a significant improvement where it has a much better average of MSE than those GWO, AIC, and BIC methods. For instance, in Table 1 when $n = 50$, the MSE reduction by CGWO was about 42.35%, 34.90%, and 30.79% comparing with AIC, BIC, and GWO respectively. Further, regardless of the value of n , the CGWO often shows the smallest MSE among the competitor methods.

In terms of variable selection performance, our proposed method obviously selects a very few irrelevant variables comparing with GWO, AIC, and BIC, where the number of the true zero coefficients which are correctly set to zero is high comparing with others. For example, in Table 3 when $n = 200$, CGWO does not select, on average, about 8 irrelevant variables out of 10 irrelevant variables. While PSO, AIC, and BIC select more than 4 irrelevant variables. On the other hand, CGWO performs very well with the smallest I (the number of the truly nonzero coefficients which are incorrectly set to zero) among all the used methods. This indicates that CGWO misses a very few important variables.

From the results of simulation 3 (Table 3), the model is dense, and, therefore, all the methods have zero values for the criterion C. On the other hand, CGWO is the best because the number of nonzero variables that have been identified as irrelevant variables is smaller compared with GWO, AIC, and BIC. It is worth noting that AIC has inferior performance in all simulation examples comparing with GWO, BIC, and CGWO methods.

Table 1: Simulation 1 results, on average.

Methods	MSE	C	I
	$n=50$		
CGWO	5.933	8.462	0.711
AIC	10.293	5.373	4.188
BIC	9.114	5.112	3.919
GWO	8.573	6.915	3.143
	$n=100$		
CGWO	5.746	8.586	0.767
AIC	10.006	5.497	4.244
BIC	8.917	5.236	3.975
GWO	8.389	7.039	3.199
	$n=200$		
CGWO	5.697	8.623	1.342
AIC	10.057	5.534	4.219
BIC	8.878	5.273	4.55
GWO	8.337	7.076	2.959

Table 2: Simulation 2 results, on average.

Methods	MSE	C	I
	$n=50$		
CGWO	7.405	13.237	1.214
AIC	11.765	7.927	3.52
BIC	10.586	7.142	3.131
GWO	10.045	9.513	2.762
	$n=100$		
CGWO	7.218	13.301	1.246
AIC	11.478	7.991	3.552
BIC	10.389	7.206	3.163
GWO	9.861	9.577	2.794
	$n=200$		
CGWO	7.169	13.312	1.254
AIC	11.529	8.002	3.56
BIC	10.35	7.217	3.171
GWO	9.809	9.588	2.802

Table 3: Simulation 3 results, on average.

Methods	MSE	C	I
	$n=50$		
CGWO	7.079	٠	1.254
AIC	11.439	٠	3.56
BIC	10.26	٠	3.171
GWO	9.719	٠	2.802
	$n=100$		
CGWO	6.892	٠	0.356
AIC	11.152	٠	2.579
BIC	10.063	٠	2.371

GWO	9.535	٠	0.99
	$n = 200$		
CGWO	6.843	٠	0.33
AIC	11.203	٠	2.551
BIC	10.024	٠	2.222
GWO	9.483	٠	0.914

6. Real application result

To make the benefit of the our proposed method in the real application, a chemistry dataset with $(n, p) = (65,15)$, of imidazo[4,5-b]pyridine derivatives (Algamal, Lee, Al-Fakih, & Aziz, 2015). The response of interest is the biological activities (IC_{50}) (Algamal & Lee, 2017). A Chi-square test as a goodness of fit is used to check whether the biological activities variables has the gamma distribution. The result of the test equals to 9.3657 with p-value equals to 0.9534. This indicating that the gamma distribution fits very well to this response variable. The estimation of the dispersion parameter is 0.0066".

Table 4 summarizes the MSE and the selected variables for each used method for the real data application.

As seen from the result of Table 4, CGWO can remarkably reduce the MSE comparing with GWO, AIC, and BIC. In terms of selected variables, on the other hand, it clearly seen from Table 4 that CGWO only select 6 variables out of 15 variables when the gamma model is assumed. CGWO selected the explanatory variables x_1, x_2, x_7, x_8, x_{11}, and x_{15}. These selected variables are identified as relevant variables to the study. Comparing with GWO and BIC, CGWO includes few variables with the MSE is less than them.

Table 4: MSE and the selected variables for the real application

Methods	Selected variables	MSE
CGWO	$x_1, x_7, x_8, x_{11}, x_{15}\ x_2$	1592.21
AIC	$x_1, x_2, x_3, x_6, x_8, x_{10}, x_{11}, x_{14}$	1862.86
BIC	$x_1, x_2, x_3, x_6, x_7, x_{10}, x_{11}, x_{14}, x_{15}$	1831.81
GWO	$x_1, x_2, x_5, x_7, x_8, x_{11}, x_{15}$	1625.05

7. Conclusion

In this paper, the problem of selecting variables in gamma regression model is considered. A chaotic grey wolf optimization algorithm was proposed as a variable selection method. The results obtained from simulation examples and real data applications demonstrated the superiority of the CGWO in terms of MSE, I, and C comparing with GWO, AIC, and BIC methods.

8. REFERENCES

[1] Al-Abood, A. M., & Young, D. H. (1986). Improved deviance goodness of fit statistics for a gamma regression model. Communications in Statistics - Theory and Methods, 15(6), 1865-1874. doi:10.1080/03610928608829223

[2] Algamal, Z. Y., & Lee, M. H. (2015). Adjusted adaptive lasso in high-dimensional Poisson regression model. Modern Applied Science, 9(4), 170-176.

[3] Algamal, Z. Y., & Lee, M. H. (2015). Penalized logistic regression with the adaptive LASSO for gene selection in high-dimensional cancer classification. Expert Systems with Applications, 42(23), 9326-9332. doi:10.1016/j.eswa.2015.08.016

[4] Algamal, Z. Y., & Lee, M. H. (2017). A novel molecular descriptor selection method in QSAR classification model based on weighted penalized logistic regression. Journal of Chemometrics, 31(10), e2915. doi:10.1002/cem.2915

[5] Algamal, Z. Y., Lee, M. H., Al-Fakih, A. M., & Aziz, M. (2015). High-dimensional QSAR prediction of anticancer potency of imidazo[4,5-b]pyridine derivatives using adjusted adaptive LASSO. Journal of Chemometrics, 29(10), 547-556. doi:10.1002/cem.2741

[6] Broadhurst, D., Goodacre, R., Jones, A., Rowland, J. J., & Kell, D. B. (1997). Genetic algorithms as a method for variable selection in multiple linear regression and partial least squares regression, with applications to pyrolysis mass spectrometry. Analytica Chimica Acta, 348(1-3), 71-86.

[7] Brusco, M. J. (2014). A comparison of simulated annealing algorithms for variable selection in principal component analysis and discriminant analysis. Computational Statistics & Data Analysis, 77, 38-53. doi:10.1016/j.csda.2014.03.001

[8] De Jong, P., & Heller, G. Z. (2008). Generalized linear models for insurance data (Vol. 10): Cambridge University Press Cambridge.

[9] Drezner, Z., Marcoulides, G. A., & Salhi, S. (1999). Tabu search model selection in multiple regression analysis. Communications in Statistics - Simulation and Computation, 28(2), 349-367. doi:10.1080/03610919908813553

[10] Dunder, E., Gumustekin, S., & Cengiz, M. A. (2016a). Variable selection in gamma regression models via artificial bee colony algorithm. Journal of Applied Statistics, 1-9. doi:10.1080/02664763.2016.1254730

[11] Dunder, E., Gumustekin, S., & Cengiz, M. A. (2016b). Variable selection in gamma regression models via artificial bee colony algorithm. Journal of Applied Statistics, 45(1), 8-16. doi:10.1080/02664763.2016.1254730

[12] Dünder, E., Gümüştekin, S., Murat, N., & Cengiz, M. A. (2017). Variable selection in linear regression analysis with alternative Bayesian information criteria using differential evaluation algorithm. Communications in Statistics - Simulation and Computation, 47(2), 605-614. doi:10.1080/03610918.2017.1288245

[13] Emary, E., Zawbaa, H. M., & Grosan, C. (2018). Experienced gray wolf optimization through reinforcement learning and neural networks. IEEE transactions on neural networks and learning systems, 29(3), 681-694.

[14] Emary, E., Zawbaa, H. M., & Hassanien, A. E. (2016). Binary grey wolf optimization approaches for feature selection. Neurocomputing, 172, 371-381.

[15] Fan, J., & Li, R. (2001). Variable selection via nonconcave penalized likelihood and its oracle properties. Journal of the American Statistical Association, 96(456), 1348-1360.

[16] Hattab, W. M. (2016). A derivation of prediction intervals for gamma regression. Journal of Statistical Computation and Simulation, 86(17), 3512-3526. doi:10.1080/00949655.2016.1169421

[17] Koç, H., Dünder, E., Gümüştekin, S., Koç, T., & Cengiz, M. A. (2017). Particle swarm optimization-based variable selection in Poisson regression analysis via information complexity-type criteria. Communications in Statistics - Theory and Methods, 1-9. doi:10.1080/03610926.2017.1390129

[18] Li, Q., Chen, H., Huang, H., Zhao, X., Cai, Z., Tong, C., . . . Tian, X. (2017). An Enhanced Grey Wolf Optimization Based Feature Selection Wrapped Kernel Extreme Learning Machine for Medical Diagnosis. Comput Math Methods Med, 2017, 9512741. doi:10.1155/2017/9512741

[19] Malehi, A. S., Pourmotahari, F., & Angali, K. A. (2015). Statistical models for the analysis of skewed healthcare cost data: A simulation study. Health Economics Review, 5, 1-11. doi:10.1186/s13561-015-0045-7

[20] Massaro, T. J., & Bozdogan, H. (2015). Variable subset selection via GA and information complexity in mixtures of Poisson and negative binomial regression models. arXiv preprint arXiv:1505.05229.

[21] Mirjalili, S., Mirjalili, S. M., & Lewis, A. (2014). Grey wolf optimizer. Advances in engineering software, 69, 46-61.

[22] Örkcü, H. (2013). Subset selection in multiple linear regression models: A hybrid of genetic and simulated annealing algorithms. Applied Mathematics and Computation, 219(23), 11018-11028. doi:10.1016/j.amc.2013.05.016

[23] Pacheco, J., Casado, S., & Núñez, L. (2009). A variable selection method based on Tabu search for logistic regression models. European Journal of Operational Research, 199(2), 506-511. doi:10.1016/j.ejor.2008.10.007

[24] Sayed, G. I., Hassanien, A. E., & Azar, A. T. (2017). Feature selection via a novel chaotic crow search algorithm. Neural Computing and Applications. doi:10.1007/s00521-017-2988-6

[25] Sindhu, R., Ngadiran, R., Yacob, Y. M., Zahri, N. A. H., & Hariharan, M. (2017). Sine–cosine algorithm for feature selection with elitism strategy and new updating mechanism. Neural Computing and Applications, 28(10), 2947-2958. doi:10.1007/s00521-017-2837-7

[26] Singh, N., & Hachimi, H. (2018). A New Hybrid Whale Optimizer Algorithm with Mean Strategy of Grey Wolf Optimizer for Global Optimization. Mathematical and Computational Applications, 23(1), 14. doi:10.3390/mca23010014

[27] Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 58(1), 267-288. doi:10.2307/2346178

[28] Unler, A., & Murat, A. (2010). A discrete particle swarm optimization method for feature selection in binary classification problems. European Journal of Operational Research, 206(3), 528-539. doi:10.1016/j.ejor.2010.02.032

[29] Uusipaikka, E. (2009). Confidence intervals in generalized regression models. NW: Chapman & Hall/CRC Press.

[30] Wang, Z., Ma, S., Zappitelli, M., Parikh, C., Wang, C. Y., & Devarajan, P. (2014). Penalized count data regression with application to hospital stay after pediatric cardiac surgery. Stat. Meth. Med. Res., In press. doi:10.1177/0962280214530608

[31] Zou, H. (2006). The adaptive lasso and its oracle properties. Journal of the American Statistical Association, 101(476), 1418-1429. doi:10.1198/016214506000000735

[32] Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 67(2), 301-320. doi:10.1111/j.1467-9868.2005.00503.x

Multicomponent stress-strength system reliability estimation for generalized exponential-poison distribution

Assis.prof.Dr.Awatif R. Mezaal Lec.Ali Ibrahim mansour Prof.Dr.Nada S. Karam Mathematical

Dep., AL-Mustanseriya University, College of Education, Baghdad, Iraq
draraldubaicy@yahoo.com aliameraltaay313@gmail.com karamns6760@gmail.com

Abstract: This study deals with (s – out of – k) Multicomponent Stress(Y) and Strength(X) System Reliability Estimation. Both stress and strength assumed to have Generalized Exponential-Poisson Distribution with common and known scale parameters (θ and λ). The aim here is to estimate the unknown shape parameters (α and β) for X and Y respectively using two methods of estimation ML and Bayes analysis by one prior with five loss functions. Then estimate Reliability using the same methods and compared the results by Mean square error criteria from simulation study to find the best performance of the estimators .the results show that the best estimator for $R_{(s,k)}$ is Bayes estimator under Quadratic loss function using Gamma prior function , followed by GD, GW, GP, MLE and GS estimators, respectively.

1. **Introduction** .A distribution obtained by compounding an exponential distribution with geometric distribution, with decreasing failure rate, known as exponential-geometric distribution is introduced by Adamidis and Loukas (1998) [1]. In the same fashion, Kus (2007) [9] introduced a two-parameter distribution known as exponential-Poisson (EP) distribution, which has decreasing failure rate, by compounding an exponential distribution with a Poisson distribution. The generalization of this distribution is come from Barreto-Souza and Cribari-Neto [3], with failure rate can be decreasing or increasing.The two-parameter exponential-Poisson (EP) with cumulative distribution function (cdf) given as:[2]

$$F(x) = \frac{1-\exp[-\lambda(1-\exp(-\theta x)]}{1-\exp(-\lambda)} \qquad x>0;\ \theta,\lambda>0$$

The random variables X & Y have Generalized Exponential-Poisson distribution with parameters (α, λ, θ) and (β, λ, θ) respectively if cdf's define as:

$$F(x) = \left[\frac{1-\exp[-\lambda(1-\exp(-\theta x))]}{1-\exp(-\lambda)}\right]^{\alpha} = \left[\frac{1-A_x}{1-e^{-\lambda}}\right]^{\alpha} \qquad (1)$$

$$F(y) = \left[\frac{1-exp[-\lambda(1-exp(-\theta y))]}{1-exp(-\lambda)}\right]^{\beta} = \left[\frac{1-A_y}{1-e^{-\lambda}}\right]^{\beta} \qquad (2)$$

where $A_x = e^{-\lambda(1-e^{-\theta x})}$, $A_y = e^{-\lambda(1-e^{-\theta y})}$ for α, β > 0 the shape parameters, where X, Y are called Generalized Exponential-Poisson distribution(GEPD) random variable with scale parameters (θ, λ). The corresponding probability density functions (pdf 's) are define as:

$$f(x) = \frac{\alpha\lambda\theta}{(1-\exp(-\lambda))^{\alpha}} e^{-\theta x} e^{-\lambda(1-e^{-\theta x})} \left[1-e^{-\lambda(1-e^{-\theta x})}\right]^{\alpha-1} \qquad (3)$$

$$= \frac{\alpha\lambda\theta}{(1-e^{-\lambda})^{\alpha}} e^{-\theta x} A_x [1-A_x]^{\alpha-1} \qquad (3\text{-a})$$

$$f(y) = \frac{\beta\lambda\theta}{(1-e^{-\lambda})^{\beta}} e^{-\theta y} e^{-\lambda(1-e^{-\theta y})} [1 - e^{-\lambda(1-e^{-\theta y})}]^{\beta-1} \quad (4)$$

$$= \frac{\beta\lambda\theta}{(1-e^{-\lambda})^{\beta}} e^{-\theta y} A_y [1 - A_y]^{\beta-1} \quad (4\text{-a})$$

If X is the strength of a component subjected to a stress Y, then R is a measure of system performance, the system fails if and only if the applied stress is greater than its strength, which referred as the stress- strength parameter. It arises in the context of mechanical reliability of a system. Another example of R discussed by Surles and Padgett [11] and Kotz et al. [8] involves the comparison of carbon strengths at different gauge lengths.

The rest of the paper is organized as follows. Section 2, introduce the obtained mathematical expression for the Multicomponent model reliability. In Section 3, considering two methods for estimating R, [ML and Bayes analysis] estimation methods. In Section 4, comparing the estimators of R by Monte Carlo simulations. Finally, the results conclusions are given in Section 5.

2. Multicomponent S-out of-K Stress-Strength System Reliability:

The estimation of the stress- strength parameter R is very common in the statistical literature. Several authors have considered estimation of R with different assumption of distributions (Huang et al. [6], Baklizi [2], Ghitany et al. [5],Bagheri[10])

The system when consisting *k* component and we need the work *s* of *k* is called multicomponent stress-strength system (*s*-out of-*k*). This system is studied by Bhattacharyya and Johnson **[**4**]**. Let the strength of the components $X_1, X_2, \dots, X_k$ imposed to a common stress Y, then the reliability of multicomponent stress-strength system is:

$$R_{(s,k)} = \text{Prob (at least s of } X_1, X_2, \dots, X_k \text{ exceed } Y)$$

$$= \sum_{i=s}^{k} C_i^k \int_0^{\infty} [1 - Fx(y)]^i \, [Fx(y)]^{k-i} f(y) dy \quad (5)$$

Let $X_i \sim GEP(\alpha_i, \theta, \lambda)$; i = 1,..,k be the k^th components strength variable which exposed to common stress random variable Y~GEP(β,θ ,λ) independently, the reliability of (s-out of-k) multicomponent stress-strength $R_{(s,k)}$of GEP distribution can be obtained by substitution eq.(1),(4-a) in(5) as:

$$R_{(s,k)} = \sum_{i=s}^{k} C_i^k \int_0^{\infty} \left[1 - \left(\frac{1-A_y}{1-e^{-\lambda}}\right)^{\alpha}\right]^i \left[\left(\frac{1-A_y}{1-e^{-\lambda}}\right)^{\alpha}\right]^{k-i} \frac{\beta\lambda\theta}{(1-e^{-\lambda})^{\beta}} e^{-\theta y} A_y [1 - A_y]^{\beta-1} dy$$

using $(1-x)^n = \sum_{i=0}^{n} C_i^n (-1)^i x^i$, that $\left[1 - \left(\frac{1-A_y}{1-e^{-\lambda}}\right)^{\alpha}\right]^i = \sum_{j=0}^{i} C_j^i (-1)^j \frac{(1-A_y)^{\alpha j}}{(1-e^{-\lambda})^{\alpha j}}$, then :

$$R_{(s,k)} = \sum_{i=s}^{k} C_i^k \int_0^{\infty} \sum_{j=0}^{i} C_j^i (-1)^j \frac{(1-A_y)^{\alpha j}}{(1-e^{-\lambda})^{\alpha j}} \cdot \frac{(1-A_y)^{\alpha(k-i)}}{(1-e^{-\lambda})^{\alpha(k-i)}} \frac{\beta\lambda\theta}{(1-e^{-\lambda})^{\beta}} e^{-\theta y} A_y (1 - A_y)^{\beta-1} dy$$

$$R_{(s,k)} = \sum_{i=s}^{k} C_i^k \sum_{j=0}^{i} C_j^i (-1)^j \frac{\beta\lambda\theta}{(1-e^{-\lambda})^{\alpha(k+j-i)+\beta}} \int_0^{\infty} e^{-\theta y} A_y (1 - A_y)^{\alpha(k+j-i)+\beta-1} dy$$

$$R_{(s,k)} = \sum_{i=s}^{k} C_i^k \sum_{j=0}^{i} C_j^i (-1)^j \frac{\beta\lambda\theta}{(1-e^{-\lambda})^{\alpha(k+j-i)+\beta}} \frac{(1-e^{-\lambda})^{\alpha(k+j-i)+\beta}}{[\alpha(k+j-i)+\beta]\lambda\theta}$$

$$\therefore R_{(s,k)} = \sum_{i=s}^{k} C_i^k \sum_{j=0}^{i} C_j^i \, (-1)^j \frac{\beta}{\beta+\alpha(k+j-i)} \qquad (6)$$

Where s, k, i, j are integers.

Multicomponent S-out of- K models; R Since the concept of stress-strength in engineering has been one of the deciding factors of the failure of devices, this study can be applied to engineering situations.

The following figures of reliability function $R_{(s,k)}$ in (6) are presented below for different values of shape parameters (α,β).

Figure (1): Multicomponent Reliability against parameter α.

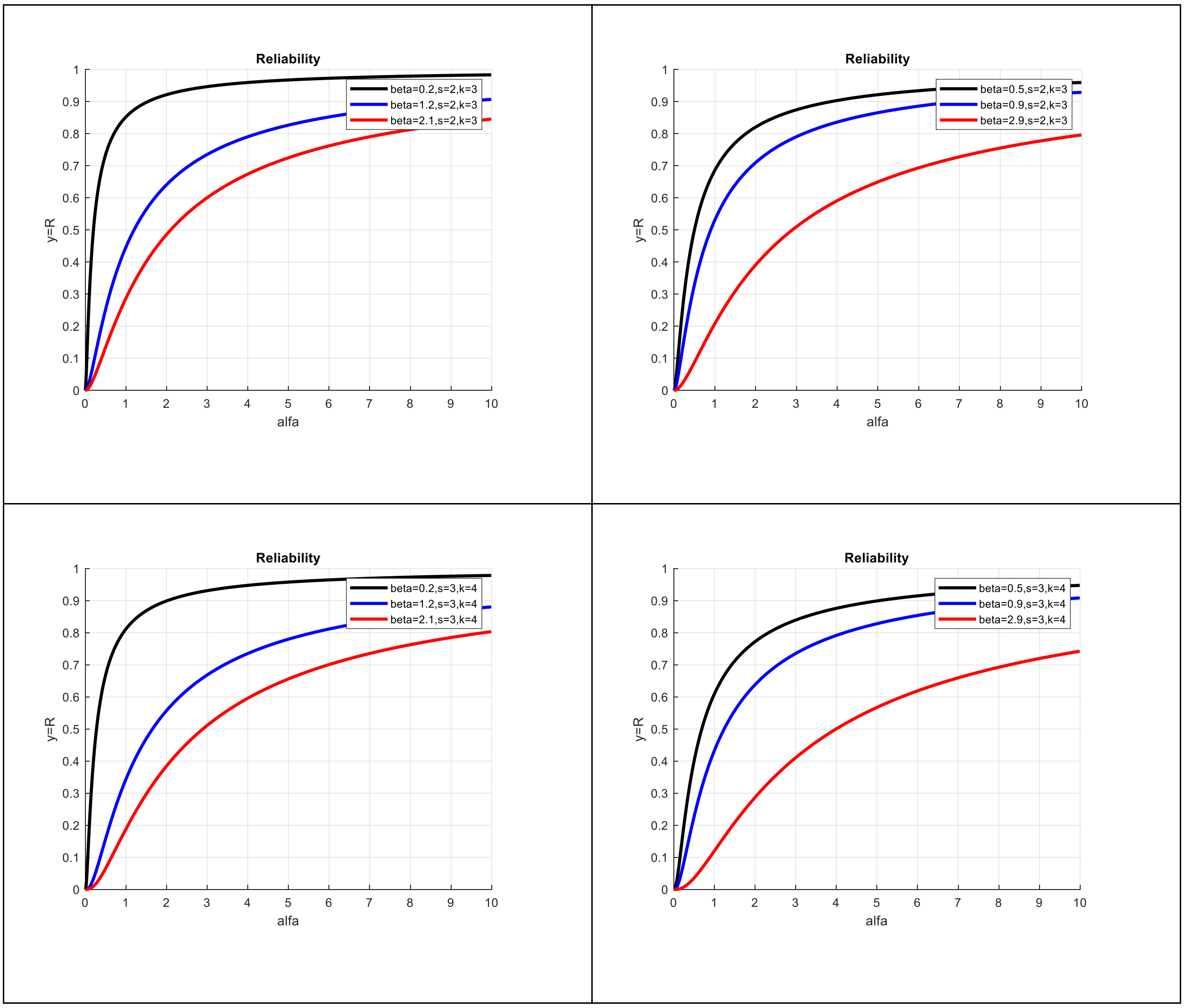
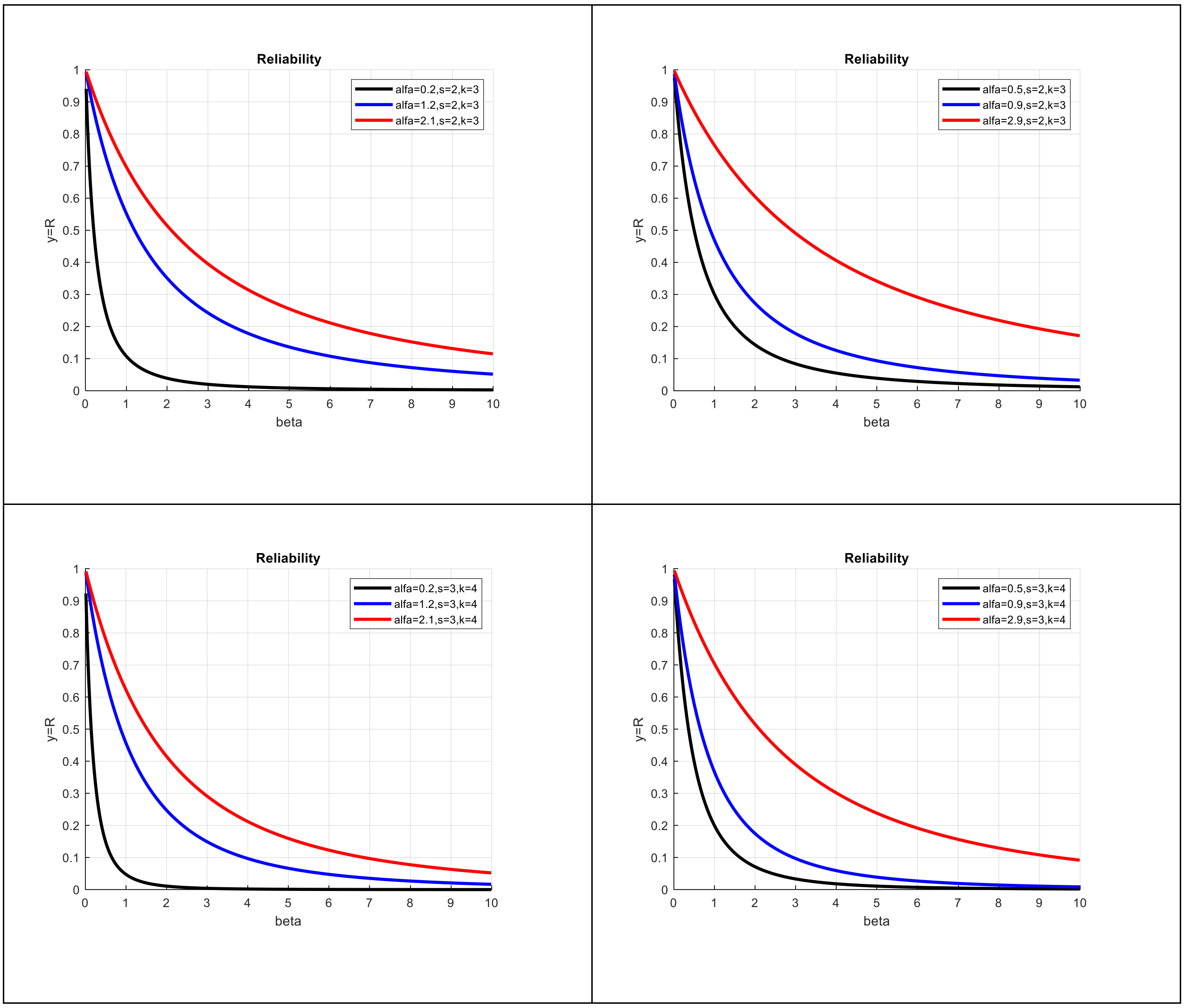


Figure (2): Multicomponent reliability against parameter β.



In figure (1) and for the strength r.v. X, show that reliability value increasing by increasing value of strength shape parameter α, where in figure (2) and for the stress r.v. Y, show that reliability value decreasing by increasing value of stress shape parameter β, knowing these cases with: (s, k) = (2, 3) and (3, 4).

3.Estimation procedures:

The estimation of the unknown parameters α, β and the reliability is done using two estimation methods [MLE, Bayes].

3.1 Maximum Likelihood estimation

Let the two independent $X_1, \dots, X_n$ and $Y_1, \dots, Y_m$ are random sample from GEP, then the likelihood function for eq. (3) is:

$$L = \alpha^n (\lambda\theta)^n e^{-\alpha n \ln\left(1-e^{-\lambda}\right)} e^{-\theta \sum_{i=1}^{n} x_i} \left(\pi_{i=1}^{n} A_{x_i}\right) e^{-\alpha \sum_{i=1}^{n} (1-A_{x_i})^{-1}} e^{\sum_{i=1}^{n} \ln(1-A_{x_i})^{-1}}$$

Let $k = (\lambda\theta)^n e^{-\theta \sum_{i=1}^{n} x_i} \left(\pi_{i=1}^{n} A_{x_i}\right) e^{\sum_{i=1}^{n} \ln(1-A_{x_i})^{-1}}$, then:

$$L = k\alpha^n e^{-\alpha[n ln(1-e^{-\lambda})+\sum_{i=1}^{n} ln(1-A_{x_i})^{-1}]} \qquad (7)$$

If $w_x = nln(1-e^{-\lambda}) + \sum_{i=1}^{n} ln(1-A_{x_i})^{-1}$, then : $L = k\alpha^n e^{-\alpha w_x}$

$$\text{Ln}\, L = lnk + nln\alpha - \alpha w_x$$

$$\frac{\partial lnL}{\partial \alpha} = \frac{n}{\alpha} - w_x \quad \to 0$$

So the ML's estimator for the unknown shape parameters α, β will be as:

$$\hat{\alpha}_{ML} = \frac{n}{w_x} \quad , \quad \hat{\beta}_{ML} = \frac{m}{w_y} \qquad (8)$$

Substation the formulas of eq. 8 in the eq.6 , the ML estimator for $R_{(s,k)}$ say $\hat{R}_{MLe}$ can obtained as:

$$\hat{R}_{MLe} = \sum_{i=s}^{k} C_i^k \sum_{j=0}^{i} C_j^i (-1)^j \frac{\hat{\beta}_{MLe}}{\hat{\beta}_{MLe} + \hat{\alpha}_{MLe}(k+j-i)} \qquad (9)$$

3.2 Bayes analysis

In this section, consider the Bayes estimation of the unknown parameters α, β and reliability with assumed that these parameters as r.v.'s under gamma prior as:

$$g(\alpha) = \frac{b^a}{\Gamma a} \alpha^{a-1} e^{-\alpha b} \qquad \alpha>0;\ b,\ a>0 \qquad (10)$$

The posterior function is:

$$p(\alpha|\underline{x}) = \frac{L(\underline{x}|\alpha)\, g(\alpha)}{\int_0^\infty L(\underline{x}|\alpha)\, g(\alpha)\, d\alpha} \qquad (11)$$

where: L(x|α) is the likelihood for the density function, g(α) is the prior function for the shape parameter. Now using eq.'s (7), (10) in (11), we get:

$$p(\alpha|\underline{x}) = \frac{k\alpha^n e^{-\alpha w_x} \frac{b^a}{\Gamma a}\alpha^{a-1} e^{-\alpha b}}{\int_0^\infty k\alpha^n e^{-\alpha w_x} \frac{b^a}{\Gamma a}\alpha^{a-1} e^{-\alpha b}\, d\alpha} \qquad \therefore p(\alpha|\underline{x}) = \frac{[w_x+b]^{n+a}}{\Gamma n+a} e^{n+a-1} e^{-\alpha[w_x+b]} \qquad (12)$$

3.2.1. Squared error loss function:

$$\hat{\alpha}_s = E(\alpha|\underline{x}) = \int_0^\infty \alpha\, p(\alpha|\underline{x}) d\alpha = \frac{[w_x+b]^{n+a}}{\Gamma n+a} \int_0^\infty \alpha^{n+a} e^{-\alpha[w_x+b]}\, d\alpha$$

Then $$\hat{\alpha}_s = \frac{n+a}{w_x+b} \quad , \quad \hat{\beta}_s = \frac{m+a}{w_y+b} \qquad (13)$$

and the reliability estimation function in eq.6 given by:

$$\hat{R}_{Bs} = \sum_{i=s}^{k} C_i^k \sum_{j=0}^{i} C_j^i (-1)^j \frac{\hat{\beta}_s}{\hat{\beta}_s + \hat{\alpha}_s(k+j-i)} \qquad (14)$$

3.2.2 Precautionary loss function

$$\hat{\alpha}_p = \sqrt{E(\alpha^2|\underline{x})}$$

since $E(\alpha^2|\underline{x}) = \int_0^\infty \alpha^2 p(\alpha|\underline{x})d\alpha = \frac{[w_x+b]^{n+a}}{\Gamma n+a}\int_0^\infty \alpha^{n+a+1} e^{-\alpha[w_x+b]}\, d\alpha$

$$\therefore \hat{\alpha}_p = \sqrt{\frac{(n+a)(n+a+1)}{[w_x+b]^2}} \quad , \quad \hat{\beta}_p = \sqrt{\frac{(m+a)(m+a+1)}{[w_y+b]^2}} \qquad (15)$$

and the reliability estimation function in eq.6 given by:

$$\hat{R}_{Bp} = \sum_{i=s}^{k} C_i^k \sum_{j=0}^{i} C_j^i (-1)^j \frac{\hat{\beta}_p}{\hat{\beta}_p+\hat{\alpha}_p(k+j-i)} \qquad (16)$$

3.2.3. De-Groot loss function

$$\hat{\alpha}_D = \frac{E(\alpha^2|\underline{x})}{E(\alpha|\underline{x})}$$

$$\rightarrow \qquad \hat{\alpha}_D = \frac{n+a+1}{[w_x+b]} \quad , \quad \hat{\beta}_D = \frac{m+a+1}{[w_y+b]} \qquad (17)$$

and the reliability estimation function in eq.8 given by:

$$\hat{R}_{BD} = \sum_{i=s}^{k} C_i^k \sum_{j=0}^{i} C_j^i (-1)^j \frac{\hat{\beta}_D}{\hat{\beta}_D+\hat{\alpha}_D(k+j-i)} \qquad (18)$$

3.2.4. Quadratic loss function

$$\hat{\alpha}_Q = \frac{E(\alpha^{-1}|\underline{x})}{E(\alpha^{-2}|\underline{x})}$$

$$E(\alpha^{-1}|\underline{x}) = \int_0^\infty \alpha^{-1} p(\alpha|\underline{x})d\alpha = \frac{[w_x+b]^{n+a}}{\Gamma n+a}\int_0^\infty \alpha^{-1}\alpha^{n+a-1} e^{-\alpha[w_x+b]}\, d\alpha$$

$$= \frac{[w_x+b]^{n+a}}{\Gamma n+a}\int_0^\infty \alpha^{n+a-2} e^{-\alpha[w_x+b]}\, d\alpha = \frac{w_x+b}{n+a-1}$$

$$E(\alpha^{-2}|\underline{x}) = \frac{[w_x+b]^{n+a}}{\Gamma n+a}\int_0^\infty \alpha^{n+a-3} e^{-\alpha[w_x+b]}\, d\alpha = \frac{[w_x+b]^2}{(n+a-1)(n+a-2)}$$

$$\rightarrow \qquad \hat{\alpha}_Q = \frac{n+a-2}{[w_x+b]} \quad , \quad \hat{\beta}_Q = \frac{m+a-2}{[w_y+b]} \qquad (19)$$

and the reliability estimation function in eq.8 given by:

$$\hat{R}_{BQ} = \sum_{i=s}^{k} C_i^k \sum_{j=0}^{i} C_j^i (-1)^j \frac{\hat{\beta}_Q}{\hat{\beta}_Q+\hat{\alpha}_Q(k+j-i)} \qquad (20)$$

3.2.5. Weighted loss function

$$\hat{\alpha}_W = \frac{1}{E(\alpha^{-1}|\underline{x})}$$

$$\rightarrow \qquad \hat{\alpha}_W = \frac{n+a-1}{[w_x+b]} \quad , \hat{\beta}_w = \frac{m+a-1}{[w_y+b]} \qquad (21)$$

and the reliability estimation function in eq.8 given by:

$$\hat{R}_{Bw} = \sum_{i=s}^{k} C_i^k \sum_{j=0}^{i} C_j^i (-1)^j \frac{\hat{\beta}_w}{\hat{\beta}_w+\hat{\alpha}_w(k+j-i)} \qquad (22)$$

4. Simulation study.

Here, we present results of some numerical three experiments ,results based on Monte Carlo simulation to compare the performance of different estimators proposed in the previous sections and different sample sizes (n,m)=(15,25),(25,15),(25,50),(50,25),(25,70),(70,25),(50,70),(70,50) and the parameters values: (a= 0.4, b=1.2) for Gamma prior, [(α= 0.2,0.9) ,(β=0.5,0.7) ,(θ=0.5,0.7) ,(λ=0.7,1) , (s_1,k_1,s_2,k_2)=(2, 3, 3, 4) for $R_{(s,k)}$. MSE of reliability estimates over the 1000 replications are given in six tables from 1 to 6 as below:

Table -1-: The best estimation method of MSE for $R_{(s,k)}$ when $(\alpha, \beta, \theta, \lambda)$=(0.2, 0.5, 0.5, 0.7) and S=2, K=3, R = 0.2424.

n,m		RML	RGS	RGP	RGD	RGQ	RGW	BEST
15,25	Mean	0.2738	0.2775	0.2806	0.2837	0.2633	0.2708	RGQ
	MSE	0.0073	0.0074	0.0077	0.0080	0.0063	0.0068	
25,15	Mean	0.2619	0.2656	0.2626	0.2596	0.2799	0.2723	RGD
	MSE	0.0064	0.0064	0.0062	0.0060	0.0076	0.0069	
25,50	Mean	0.2656	0.2678	0.2702	0.2725	0.2575	0.2628	RGW
	MSE	0.0055	0.0056	0.0058	0.0060	0.0049	0.0053	
50,25	Mean	0.2604	0.2624	0.2600	0.2577	0.2727	0.2674	RGD
	MSE	0.0049	0.0049	0.0048	0.0046	0.0057	0.0052	
25,70	Mean	0.2663	0.2684	0.2715	0.2746	0.2551	0.2619	RGQ
	MSE	0.0041	0.0042	0.0044	0.0046	0.0035	0.0038	
70,25	Mean	0.2546	0.2564	0.2534	0.2505	0.2695	0.2628	RGD
	MSE	0.0049	0.0049	0.0048	0.0046	0.0058	0.0053	
50,70	Mean	0.2641	0.2654	0.2661	0.2668	0.2625	0.2639	RGQ
	MSE	0.0036	0.0036	0.0037	0.0037	0.0035	0.0036	
70,50	Mean	0.2621	0.2633	0.2626	0.2619	0.2662	0.2647	RGD
	MSE	0.0043	0.0043	0.0043	0.0042	0.0045	0.0044	

Table -2-: The best estimation method of MSE for $R_{(s,k)}$ when $(\alpha, \beta, \theta, \lambda)$ =(0.2, 0.5, 0.5, 0.7) and S=3, K=4, R = 0.1492.

n,m		RML	RGS	RGP	RGD	RGQ	RGW	BEST
15,25	Mean	0.1721	0.1754	0.1780	0.1808	0.1630	0.1695	RGQ
	MSE	0.0047	0.0048	0.0051	0.0053	0.0040	0.0044	

25,15	Mean	0.1732	0.1763	0.1736	0.1710	0.1891	0.1823	RGD
	MSE	0.0053	0.0053	0.0051	0.0049	0.0066	0.0059	
25,50	Mean	0.1743	0.1762	0.1783	0.1804	0.1671	0.1717	RGQ
	MSE	0.0038	0.0039	0.0041	0.0043	0.0033	0.0036	
50,25	Mean	0.1615	0.1633	0.1613	0.1593	0.1722	0.1676	RGD
	MSE	0.0035	0.0035	0.0034	0.0033	0.0041	0.0038	
25,70	Mean	0.1749	0.1767	0.1794	0.1822	.1650	0.1709	RGQ
	MSE	0.0035	0.0035	0.0038	0.0040	0.0028	0.0032	
70,25	Mean	0.1644	0.1660	0.1633	0.1608	0.1776	0.1716	RGD
	MSE	0.0033	0.0033	0.0031	0.0030	0.0041	0.0037	
50,70	Mean	0.1661	0.1672	0.1678	0.1684	0.1647	0.1660	RGQ
	MSE	0.0026	0.0027	0.0027	0.0027	0.0025	0.0026	
70,50	Mean	0.1656	0.1666	0.1660	0.1654	0.1692	0.1679	RGD
	MSE	0.0025	0.0028	0.0027	0.0026	0.0028	0.0027	

Table -3-: The best estimation method of MSE for $R_{(s,k)}$ when$(\alpha, \beta, \theta, \lambda)$ = 0.9, 1.2, 0.7, 1 and S=2, K=3, R = 0.4154.

n,m		RML	RGS	RGP	RGD	RGQ	RGW	BEST
15,25	Mean	0.4192	0.4176	0.4210	0.4245	0.4013	0.4099	RGW
	MSE	0.0087	0.0077	0.0077	0.0078	0.0078	0.0076	
25,15	Mean	0.4121	0.4243	0.4208	0.4173	0.4407	0.4320	RGP
	MSE	0.0076	0.0067	0.0066	0.0067	0.0073	0.0069	
25,50	Mean	0.4201	0.4179	0.4207	0.4234	0.4059	0.4121	RGQ
	MSE	0.0052	0.0050	0.0049	0.0049	0.0048	0.0049	
50,25	Mean	0.4141	0.4222	0.4195	0.4168	0.4342	0.4280	RGD
	MSE	0.0054	0.0052	0.0051	0.0050	0.0053	0.0051	
25,70	Mean	0.4257	0.4216	0.4252	0.4287	0.4063	0.4142	RGW RGS
	MSE	0.0048	0.0043	0.0044	0.0045	0.0044	0.0043	
70,25	Mean	0.4078	0.4173	0.4137	0.4102	0.4327	0.4248	RGD
	MSE	0.0043	0.0039	0.0039	0.0038	0.0042	0.0040	
50,70	Mean	0.4210	0.4213	0.4221	0.4229	0.4179	0.4196	RGQ
	MSE	0.0029	0.0028	0.0028	0.0028	0.0027	0.0028	
70,50	Mean	0.4143	0.4176	0.4168	0.4160	0.4210	0.4218	RGD

	MSE	0.0031	0.0030	0.0029	0.0028	0.0030	0.0030	

Table -4-: The best estimation method of MSE for $R_{(s,k)}$ when$(\alpha, \beta, \theta, \lambda)$ =(0.9, 1.2, 0.7, 1) and S=3, K=4, R = 0.3115.

n,m		RML	RGS	RGP	RGD	RGQ	RGW	BEST
15,25	Mean	0.3253	0.3229	0.3264	0.3299	0.3064	0.3151	RGQ
	MSE	0.0098	0.0085	0.0086	0.0088	0.0081	0.0083	
25,15	Mean	0.3152	0.3268	0.3233	0.3197	0.3436	0.3347	RGD
	MSE	0.0081	0.0074	0.0072	0.0071	0.0083	0.0078	
25,50	Mean	0.3158	0.3137	0.3165	0.3192	0.3016	0.3078	RGQ
	MSE	0.0049	0.0046	0.0046	0.0047	0.0045	0.0047	
50,25	Mean	0.3094	0.3176	0.3148	0.3121	0.3299	0.3235	RGD
	MSE	0.0048	0.0045	0.0045	0.0044	0.0049	0.0047	
25,70	Mean	0.3212	0.3170	0.3206	0.3243	0.3016	0.3095	RGQ
	MSE	0.0052	0.0048	0.0048	0.0050	0.0047	0.0049	
70,25	Mean	0.3082	0.3176	0.3141	0.3105	0.3334	0.3253	RGD
	MSE	0.0045	0.0043	0.0043	0.0042	0.0048	0.0045	
50,70	Mean	0.3127	0.3130	0.3138	0.3146	0.3095	0.3113	RGQ
	MSE	0.0029	0.0028	0.0028	0.0028	0.0027	0.0028	
70,50	Mean	0.3173	0.3204	0.3196	0.3188	0.3239	0.3221	RGD
	MSE	0.0026	0.0025	0.0025	0.0024	0.0026	0.0026	

Table -5-: The best estimation method of MSE for $R_{(s,k)}$ when$(\alpha, \beta, \theta, \lambda)$ = 1.2, 0.8, 1.5, 0.9 and S=2, K=3, R = 0.6136.

n,m		RML	RGS	RGP	RGD	RGQ	RGW	BEST
15,25	Mean	0.6092	0.5987	0.6019	0.6051	0.5835	0.5915	RGD
	MSE	0.0063	0.0059	0.0058	0.0057	0.0068	0.0062	
25,15	Mean	0.6019	0.6017	0.5984	0.5952	0.6166	0.6087	RGQ
	MSE	0.0060	0.0054	0.0055	0.0056	0.0051	0.0052	
25,50	Mean	0.6099	0.6029	0.6054	0.6079	0.5918	0.5976	RGP
	MSE	0.0036	0.0035	0.0034	0.0035	0.0040	0.0037	

50,25	Mean	0.6036	0.6043	0.6017	0.5992	0.6152	0.6096	RGW
	MSE	0.0038	0.0036	0.0036	0.0037	0.0035	0.0034	
25,70	Mean	0.6086	0.6009	0.6041	0.6074	0.5866	0.5939	RGD
	MSE	0.0030	0.0030	0.0029	0.0028	0.0036	0.0032	
70,25	Mean	0.5999	0.6020	0.5988	0.5955	0.6160	0.6089	RGQ
	MSE	0.0036	0.0033	0.0035	0.0036	0.0031	0.0032	
50,70	Mean	0.6013	0.5985	0.5992	0.6000	0.5953	0.5969	RML
	MSE	0.0020	0.0021	0.0021	0.0021	0.0023	0.0022	
70,50	Mean	0.6079	0.6071	0.6064	0.6057	0.6102	0.6087	RGQ
	MSE	0.0020	0.0020	0.0020	0.0020	0.0019	0.0021	

Table -6-: The best estimation method of MSE for $R_{(s,k)}$ when$(\alpha, \beta, \theta, \lambda)$ = (1.2, 0.8, 1.5, 0.9) and S=3, K=4, R = 0.5260.

n,m		RML	RGS	RGP	RGD	RGQ	RGW	BEST
15,25	Mean	0.5134	0.5019	0.5055	0.5091	0.4847	0.4938	RGD
	MSE	0.0082	0.0078	0.0076	0.0074	0.0090	0.0083	
25,15	Mean	0.5151	0.5146	0.5110	0.5074	0.5315	0.5226	RGW
	MSE	0.0080	0.0072	0.0073	0.0075	0.0071	0.0070	
25,50	Mean	0.5212	0.5132	0.5161	0.5189	0.5007	0.5072	RGD
	MSE	0.0046	0.0044	0.0043	0.0042	0.0049	0.0046	
50,25	Mean	0.5135	0.5143	0.5114	0.5086	0.5267	0.5203	RGQ
	MSE	0.0045	0.0042	0.0043	0.0044	0.0040	0.0041	
25,70	Mean	0.5189	0.5102	0.5139	0.5175	0.4940	0.5023	RGD
	MSE	0.0038	0.0038	0.0037	0.0036	0.0046	0.0041	
70,25	Mean	0.5129	0.5153	0.5116	0.5079	0.5312	0.5230	RGQ
	MSE	0.0047	0.0043	0.0044	0.0046	0.0042	0.0043	
50,70	Mean	0.5157	0.5124	0.5132	0.5141	0.5089	0.5107	RGD
	MSE	0.0030	0.0029	0.0029	0.0028	0.0031	0.0030	
70,50	Mean	0.5199	0.5190	0.5182	0.5173	0.5225	0.5207	RGQ
	MSE	0.0028	0.0027	0.0027	0.0027	0.0025	0.0026	

5.Conclusion:

From the tables (1 to 6), we have observed that:

1- MSE value decreasing by increasing sample size (n,m) for MLE and Bayes estimators.
2- The value of $R_{(s,k)}$ decreases as (k) value decreases.
3- In general the best performance was is in Bayes method under Quadratic loss function estimator, followed by GD, GW, GP, MLE and GS estimators, respectively.

Acknowledgments

The authors would like to thank Mustansiriyah University (www.uo Mustansiriyah.edu.iq) Baghdad-Iraq for its support in the present work.

6.Reference

[1] Adamidis, K.; Loukas, S, Alifetime distribution with decreasing failure rate. 1998. *Statistics and Probability Letters*, No.**39**, **p35–42**.

[2] Baklizi, A., Interval estimation of the stress-strength reliability in the two-parameter exponential distribution based on records. *Journal of Statistical Computation and Simulation*, 2014, Vol. **84,** p **2670-2679.**

[3] Barreto-Souza, W. and Cribari-Neto, F., A generalization of the Exponential-Poisson distribution. 2009, *Statistics and Probability Letters*, N0.**79,** p **2493-2500**.

[4] Bhattacharyya G.K. and Johnson R.A., Estimation of reliability in Multicomponent Stress-Strength models ,1974. *Journal of American Statistical Association*, Vol. 69, No.**348**.

[5] Ghitany, M.E., M., D.K. and Aboukhamseen, S.M., Estimation of the reliability of a stress-strength system from power Lindley distributions. 2015, *Com. in Statistics Simulation and Computation,* **No.44,** p **118-136**.

[6] Huang, K., Mi, J. and Wang, Z., Inference about reliability parameter with gamma strength and stress. 2012, *Journal. of Stat. Planning &inference*, Vol.**142**.p **848-854**.

[7] Karam,N.S. and Hind H.. Reliability Estimation in Multi-Component Stress-Strength Based on Burr-III Distribution, 2015, *journal of mathematical theory and modeling*, Vol.**5**, No.**11**.

[8] Kotz, S., Lumelskii, Ya. and Pensky, M., The Stress-Strength Model and Its Generalization: Theory and Applications ,2004. *Journal of the American Statis. Assos*.Vol.**99**, p **903**.

[10] Pak,A.,and Bagheri,N.,On Reliability in a Multi-Component Stress-Strength model with power Lindley distribution.,*Revista colombiana de Estadistica*,July 2018, Volume 41, Issue 2,pp. 251- 267

[11] Surles, J.G. and Padgett, W.J., Inference for reliability and stress-strength for a scaled Burr-type X distribution. 2001, *Lifetime Data Analysis*, No.**7,** p**187-200.**

Sparse minimum average variance estimation via the adaptive elastic net when the predictors correlated

Esraa Rahman[1] and Ali Alkenani[2]

[1]Department of Statistics, College of Administration and Economics, University of Al-Qadisiyah, Al Diwaniyah, Iraq.
[2]Correspondence: Ali Alkenani, E-mail: ali.alkenani@qu.edu.iq
[2]http://orcid.org/0000-0001-5067-2321

Abstract: A new model-free variable selection method was proposed in this article, which is called SMAVE-AdEN. We combined the effective sufficient dimension reduction method MAVE with the variable selection method which is called adaptive Elastic Net (AdEN) to introduce SMAVE-AdEN. The SMAVE-AdEN produces a sparse and accurate estimate when the predictors are highly correlated. The advantage of SMAVE-AdEN is that SMAVE-AdEN extended Adaptive Elastic net (AdEN) to nonlinear and multi-dimensional regression. Also, the SMAVE-AdEN enables MAVE to work with problems were the predictors are highly correlated. In addition, SMAVE-AdEN can exhaustively estimate dimensions, while selecting informative covariates simultaneously. The performance of SMAVE-AdEN is evaluated by both simulation and real data analysis.
Keywords: Dimension reduction, Variable selection, Minimum average variance estimation, *Adaptive Elastic Net.*

1. Introduction

When the amount of predictors p is large, the regression analysis might be highly challenging. A beneficial mechanism to deal with this obstacle is decreasing p-dimensional predictors vector $\mathbf{x}$ with no loss regarding regression information.

The sufficient dimension reduction (SDR) theory has been presented by Cook (1998) to perform the above aim. Assuming y is response variable, also $\mathbf{x} = (x_1, \dots, x_p)^T$ is a $p \times 1$ has been predictor vector. SDR explores $p \times d$ matrix $\mathbf{B}$, in a way that $y \perp\!\!\!\perp \mathbf{x} | \mathbf{x}^T\mathbf{B}$, in which $\perp\!\!\!\perp$ indicates independence. Also, dimension reduction subspace (DRS) is the column space spanned by $\mathbf{B}$. The intersection of all DRS has been referred to as the central subspace ($S_{y|\mathbf{x}}$). The $S_{y|\mathbf{x}}$ has contained all regression information regarding $y|\mathbf{x}$ (Yu and Zhu, 2013). Several methods has been submitted for obtaining $S_{y|\mathbf{x}}$. For instance, SIR (Li, 1991), SAVE (Cook and Weisberg, 1991) as well as PHD (Li, 1992).

When the mean function is of interest, (Cook and Li, 2002) has presented the notion that is related to central mean subspace ($S_{E(y|\mathbf{x})}$). For the purpose of estimate $S_{E(y|\mathbf{x})}$, a number of DR methods were presented, as the iterative Hessian transformation (Cook and Li, 2002) as well as MAVE (Xia et al., 2002).

SDR approaches supplying the researchers with useful tools for gaining adequate DR; nevertheless, these methods suffer from that each DR direction has been linear combination regarding original predictors. This makes the resulting measures not ease.

The process of electing the predictors is a very important in constructing the model of multiple regression. Besides, the selection of the significant predictors for being in the model improves the prediction accuracy related to the model. Furthermore, the small subset of predictors performs the interpretation of the results easy. The regularisation methods were applied for variable selection with regard to regression models from various researchers. See, for instance the Lasso (Tibshirani, 1996), SCAD (Fan and Li, 2001), Elastic Net (EN) (Zou and Hastie, 2005), adaptive Lasso (Zou, 2006) and MCP (Zhang, 2010).

Under SDR perspectives, the views of regularisation methods joined with several SDR methods from many researchers. For instance, Li et al. (2005), Ni et al. (2005), Li and Nachtsheim (2006), Li (2007), Li and Yin (2008). Wang and Yin (2008) combined Lasso with MAVE to produce sparse MAVE (SMAVE) estimate. Wang et al. (2013) proposed penalised MAVE (P-MAVE) through combining bridge penalty with l_1-norms of the rows of a basis matrix. Alkenani and Yu (2013) incorporate MAVE with SCAD, adaptive Lasso and the MCP to produce SCAD-MAVE, ALMAVE and MCP-MAVE, respectively. Wang et al. (2015) combined Lasso with the group-wise MAVE which suggested by Li et al. (2010). Alkenani and Rahman (2020) proposed SMAVE-EN method by combining MAVE with EN penalty to produce sparse and accurate estimates. The Lasso does not have oracle properties and it is not stable. Furthermore, adaptive Lasso achieves the oracle property, yet it inherits the instability related to Lasso with regard to the high-dimensional data. Elastic net handles collinearity, yet lacking oracle property. Zou and Zhan (2015) proposed Adaptive Elastic Net (AdEN) to tackle the limitations of EN. They proposed to employ adaptive Lasso instead of Lasso in EN penalty.

In this article, the SMAVE-AdEN has been presented. The SMAVE-AdEN is a shrinkage estimation method under SDR settings, where there is a set of predictors among which the predictors are highly pairwise correlated. SMAVE-AdEN is a nice combination of MAVE and AdEN. SMAVE-AdEN has advantages over the SMAVE (Wang and Yin, 2008), SPMAVE (Alkenani and Yu, 2013), P-MAVE (Wang et al., 2013) and SMAVE-EN (Alkenani and Rahman, 2020). It benefits from the strength of AdEN. In AdEN, the variable selection and parameters estimation have been implemented in one process and AdEN has the oracle property to select groups of highly correlated variables. The mentioned ability does not hold for Lasso, adaptive lasso, SCAD, MCP, bridge penalties and EN which are employed in the existing methods.

The rest of this article is as follows. In Section 2, a summary of MAVE and SMAVE-EN method is presented. We proposed the SMAVE-AdEN in Section 3. Simulation studies are implemented in Section 4. In Section 5, the methods under consideration are applied to graduate student rate data. The conclusions are given in Section 6.

2. MAVE and SMAVE-EN

In this section, we propose a brief of MAVE and SMAVE-EN. Suppose the following model:

$$y = f(\mathrm{x}_1, \mathrm{x}_2, \dots, \mathrm{x}_p) + \varepsilon, \qquad (1)$$

where y, $\mathbf{x}$ and ε are the response variable, a $p \times 1$ predictor vector and the error term, respectively. In addition, $E(y|\mathbf{x}) = f(\mathrm{x}_1, \mathrm{x}_2, \dots, \mathrm{x}_p)$ and $E(\varepsilon\,|\mathbf{x}) = 0$. For the mean function, the SDR aims to investigate a subspace S such that

$$y \perp\!\!\!\perp E(y|\mathbf{x}) | P_s\mathbf{x}, \qquad (2)$$

where $P_{(.)}$ is a projection operator. The mean DR subspaces achieve (2) (Cook and Li, 2002). If $d = dim(S)$ and $\mathbf{B} = (\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \dots, \boldsymbol{\beta}_d)$ is a basis for S, $\mathbf{x}$ can be replaced with LC $\mathbf{x}^T\boldsymbol{\beta}_1, \mathbf{x}^T\boldsymbol{\beta}_2, \dots, \mathbf{x}^T\boldsymbol{\beta}_d$, $d \le p$ without a loss of information on $E(y|\mathbf{x})$. Cook and Li, (2002) show that the central mean subspace $S_{E(y|\mathbf{x})}$ is the intersection of all subspaces satisfying (2). Many methods were proposed to estimate $S_{E(y|\mathbf{x})}$ and one of the more efficient methods is MAVE.

Xia et al. (2002) proposed MAVE such that the matrix $\mathbf{B}$ is the solution of

$$\min_{\mathbf{B}}\{\mathrm{E}[y - \mathrm{E}(y|\mathbf{x}^T\mathbf{B})]^2\}, \qquad (3)$$

where $\mathbf{B}^T\mathbf{B} = \mathbf{I}_d$. The conditional variance given $\mathbf{x}^T\mathbf{B}$ is

$$\boldsymbol{\sigma}^2_{\mathbf{B}}(\mathbf{x}^T\mathbf{B}) = \mathrm{E}[\{y - \mathrm{E}(y|\mathbf{x}^T\mathbf{B})\}^2\,|\mathbf{x}^T\mathbf{B}]. \qquad (4)$$

Thus,

$$\min_{\mathbf{B}} \mathrm{E}[y - \mathrm{E}(y|\mathbf{x}^T\mathbf{B})]^2 = \min_{\mathbf{B}} \mathrm{E}\{\boldsymbol{\sigma}_{\mathbf{B}}^2(\mathbf{x}^T\mathbf{B})\}. \qquad (5)$$

For any given $\mathbf{x}_0$, $\boldsymbol{\sigma}_{\mathbf{B}}^2(\mathbf{x}^T\mathbf{B})$ can be locally approximated as

$$\boldsymbol{\sigma}_{\mathbf{B}}^2(\mathbf{x}_0^T\mathbf{B}) \approx \sum_{i=1}^{n}\{y_i - \mathrm{E}(y_i|\mathbf{x}_i^T\mathbf{B})\}^2\,\omega_{i0}$$

$$\approx \sum_{i=1}^{n}[y_i - \{a_0 + (\mathbf{x}_i - \mathbf{x}_0)^T\mathbf{B}\,\boldsymbol{b}_0\}]^2\,\omega_{i0},$$

where $\omega_{i0} \geq 0$ are the kernel weights with $\sum_{i=1}^{n}\omega_{i0} = 1$. So, $\mathbf{B}$ can be found by solving

$$\min_{\mathbf{B}:\,\mathbf{B}^T\mathbf{B} = \mathbf{I}_m} \left(\sum_{j=1}^{n}\sum_{i=1}^{n}\left[y_i - \left\{a_j + \left(\mathbf{x}_i - \mathbf{x}_j\right)^T\mathbf{B}\,\boldsymbol{b}_j\right\}\right]^2\omega_{ij}\right). \qquad (6)$$

Alkenani and Rahman (2020) incorporate EN penalty term in (6) to obtain a SMAVE. The SMAVE minimises:

$$\sum_{j=1}^{n}\sum_{i=1}^{n}\left[y_i - \left\{a_j + \left(\mathbf{x}_i - \mathbf{x}_j\right)^T\mathbf{B}\,\boldsymbol{b}_j\right\}\right]^2\omega_{ij} + \lambda_1\|\boldsymbol{\beta}_m\|^2 + \lambda_2\|\boldsymbol{\beta}_m\|_1, \qquad (7)$$

for $m = 1, \ldots, d$.

where, d has been known and it can be estimated by BIC. where, $|\boldsymbol{\beta}_m|^2$ is l_2 norm related with ridge penalty and $|\boldsymbol{\beta}_m|_1$ is $l_1$norm related with Lasso penalty. The minimisation in (7) consists of three parts. The first part is the loss function of MAVE based on the least-squares formulation of MAVE. The second part is the ridge penalty function stabilizes the solution paths, handles the collinearity and, thus, improving the prediction. The third part is the Lasso penalty function which performs parameters estimation and variable selection simultaneously. The EN penalty consists of the second and the third parts. The EN inherit the advantages and disadvantages of ridge and Lasso. Also, $\lambda_1$and λ_2 are the tuning parameters of Elastic Net. Such good properties making SMAVE-EN extremely significant variable selection approach. In spite of its significance, SMAVE-EN have two disadvantages: which are, lacking oracle property and bias for estimating.

In this article, for the reasons mentioned, SMAVE-AdEN is proposed to minimise

$$\sum_{j=1}^{n}\sum_{i=1}^{n}\left[y_i - \left\{a_j + \left(\mathbf{x}_i - \mathbf{x}_j\right)^T\mathbf{B}\,\boldsymbol{b}_j\right\}\right]^2\omega_{ij} + \left(1 + \tfrac{\lambda_2}{n}\right)\{\lambda_1\|\boldsymbol{\beta}_m\|_2^2$$

$$+ \lambda_2\sum_{k=1}^{p} w_k^*\,|\boldsymbol{\beta}_{m,k}|_1\}, \ldots\ldots (8)$$

1. Let $m = 1$, and $\mathbf{B} = \boldsymbol{\beta}_0$, any arbitrary $p \times 1$ vector.

2. For known $\mathbf{B}$, get $(a_j, \boldsymbol{b}_j)$ where $j = 1, \ldots, n$, from

$$\min_{a_j, \boldsymbol{b}_j, j=1,\ldots,n} \left(\sum_{j=1}^{n}\sum_{i=1}^{n}\left[y_i - \left\{a_j + \left(\mathbf{x}_i - \mathbf{x}_j\right)^T\mathbf{B}\,\boldsymbol{b}_j\right\}\right]^2\omega_{ij}\right). \qquad (10)$$

3. For a given $(\hat{a}_j, \hat{\boldsymbol{b}}_j)$, $j = 1, \ldots, n$, solve $\boldsymbol{\beta}_{m\mathrm{SMAVE-AdEN}}$ from

$$\min_{\mathbf{B}:\,\mathbf{B}^T\mathbf{B} = \mathbf{I}_m} \left(\sum_{j=1}^{n}\sum_{i=1}^{n}\left[y_i - \left\{\hat{a}_j + \left(\mathbf{x}_i - \mathbf{x}_j\right)^T\left(\widehat{\boldsymbol{\beta}}_1, \widehat{\boldsymbol{\beta}}_2, \ldots, \widehat{\boldsymbol{\beta}}_{m-1}, \boldsymbol{\beta}_m\right)\hat{\boldsymbol{b}}_j\right\}\right]^2\omega_{ij} + \{\lambda_1\|\boldsymbol{\beta}_m\|_2^2 +\right.$$

$$\left.\lambda_2\sum_{k=1}^{p} w_k^*\,|\boldsymbol{\beta}_{m,k}|_1\right) \qquad (11)$$

4. Replace the mth column of $\mathbf{B}$ by $\widehat{\boldsymbol{\beta}}_{m\mathrm{SMAVE-AdEN}}$ and repeat steps 2 and 3 until convergence.

5. Update $\mathbf{B}$ by $(\widehat{\boldsymbol{\beta}}_{1\mathrm{SMAVE-AdEN}}, \widehat{\boldsymbol{\beta}}_{2\mathrm{SMAVE-AdEN}}, \ldots, \widehat{\boldsymbol{\beta}}_{m\mathrm{SMAVE-AdEN}}, \boldsymbol{\beta}_0)$, and set m to be $m + 1$.

6. If $m < d$, continue steps 2 to 5 until $m = d$.

Where ω_{ij} are the kernel weights and they are computed as follows

$$\omega_{ij} = K_h\left\{\left(\mathbf{x}_i - \mathbf{x}_j\right)^T\widetilde{\mathbf{B}}\right\} / \sum_{i=1}^{n} K_h\left\{\left(\mathbf{x}_i - \mathbf{x}_j\right)^T\widetilde{\mathbf{B}}\right\},$$

K_h is the refined multidimensional Gaussian kernel and $h_{opt} = \mathbb{A}(d)n^{-1/(4+d)}$ is the bandwidth, where $\mathbb{A}(d) = \left\{\frac{4}{(d+2)}\right\}^{1/(4+d)}$, see (Xia et al., 2002).

SMAVE-AdEN combines AdEN into the "OLS formulation" of MAVE. Thus, under the same conditions as those for MAVE and EN, the algorithm well be converging to global minimum. According to the simulations of this study, the algorithm of SMAVE-EN usually converges within seven to twelve iterations. The LARS algorithm of Efron et al. (2004) can be manipulated to get the efficient solution of EN with the order of computational efforts has been majorly comparable to that related to single OLS fit (Zou and T. Hastie, 2005). Compared to MAVE, the penalty term in SMAVE-AdEN appears in the "OLS formulation" of MAVE, so the algorithm is as efficient as MAVE. At an early stage of LARS-EN algorithm, the optimal results have been achieved. After m steps, if the algorithm of LARS-EN is settled, then it demands $O(m^3 + pm^2)$ operations (Zou and T. Hastie, 2005). Due to (Xia et al, 2002), the rate of consistency for the MAVE estimator is $O(h_{opt}^3 \log(n))$. Because of that the rate of consistency of MAVE is a lower than that of AdEN, the rate of consistency of SMAVE-AdEN estimator is controlled by that of MAVE. In summary, under the same conditions of Zou and T. Hastie (2005) and Xia et al. (2002), one may show that the SMAVE-AdEN estimator has the same consistency rate as the MAVE estimator and it is also as efficient as MAVE asymptotically.

3. Simulation study

The goal in this section is to compare the performance of SMAVE-AdEN with the ALMAVE, SCAD-MAVE, MCP-MAVE, SMAVE, P-MAVE and SMAVE-EN methods in terms of prediction accuracy and variable selection. Also, to show the preference of SMAVE-AdEN when the grouped selection is required.

An amount of examples is reported to show the performance of the SMAVE-AdEN. The simulated data consist of a training set, an independent validation set and an independent test set within each example. The training data were employed to fit the models, and the tuning parameters were selected by using the validation data. The test data were employed to compute the mean-squared error (MSE). The notation ./././. refers to the number of observations in the mentioned three sets, respectively.

The ALMAVE, SCAD-MAVE and MCP-MAVE methods were computed using R codes made by Alkenani and Yu (2013). The SMAVE, P-MAVE and SMAVE-EN methods were computed using R codes made by Wang and Yin (2008), Wang et al. (2013) and Alkenani and Rahman (2020), respectively. The R code for SMAVE-AdEN is available from the authors. For each competitor, the tuning parameters was chosen via tenfold cross-validation (C.V).
The examples as follow:

Example1. We generated data from the linear regression model.

$$y = X^T\beta^* + \varepsilon,$$

where β^* is a p −dimensional vector and $\varepsilon \sim N(0, \sigma^2), \sigma = 6$, and X is from multivariate normal distribution with zero mean and covariance $\sum$ whose (j,k) entry is $\sum_{j,k} = \rho^{|j-k|}, 1 \leq k, j \leq p$.

We considered $\rho = 0.5$ and $\rho = 0.75$. Let $p = p_n = [4\, n^{1/2}] - 5$ for $n = 50, 100$. Let $1_m/0_m$ denote a m −vector of $1's/0's$. The vector of true coefficients is $\beta^* = (3.1_q, 3.1_q, 3.1_q, 0_{p-3q})^T$ and $|\mathcal{A}| = 3q$ and $q = [p_n/9]$. In this example $v = \frac{1}{2}$ hence, we used $y = 3$ for computing the adaptive weights in the adaptive elastic-net. Let $\mathcal{A} = \{j : \beta_j^* \neq 0, j = 1, 2, \dots, p\}$, $\gamma > \frac{2\nu}{1-\nu}$, $0 \leq \nu < 1$

$\mathcal{A}$ is the intrinsic dimension of the underlying model.

With regard to each estimator $\hat{\beta}$, the estimation accuracy well be evaluated via mean squared error (MSE) that is specified as $E[(\hat{\beta} - \beta^*)^T \sum(\hat{\beta} - \beta^*)]$. Also, variable selection performance has been gauged via *(C, IC)*, in which *C* has been the number of zero coefficients which have been estimated correctly via zero, also *IC* representing the number of the nonzero coefficients which have been estimated incorrectly via zero.

Example 2. We generated data from the linear regression model

$$y = \frac{x^T \beta_1}{0.5+(1.5+ x^T \beta_2)} + \sigma\varepsilon,$$

where β^* is a p –dimensional vector and $\varepsilon \sim N(0, \sigma^2), \sigma = 6$, and X is from multivariate normal distribution with zero mean and covariance $\sum$ whose (j, k) entry is $\sum_{j,k} = \rho^{|j-k|}, 1 \leq k, j \leq p$. We considered $\rho = 0.5$ and $\rho = 0.75$. Let $p = p_n = [4\, n^{1/2}] - 5$ for $n = 100, 200, 400$. Let $1_m/0_m$ denote a m –vector of $1's/0's1's/0's$. The vector of true coefficients are $\beta_1 = (3.1_q, 3.1_q, 3.1_q, 0_{p-3q})^T$ and $\beta_2 = (0_{p-3q}, 3.1_q, 3.1_q, 3.1_q)^T$ and $|\mathcal{A}| = 3q$ and $q = [p_n/9]$. In this example $v = \frac{1}{2}$ hence, we used $\gamma = 3$ for computing the adaptive weights in the adaptive elastic-net.

$\gamma > \frac{2\nu}{1-\nu}$, $0 \leq \nu < 1$, $\mathcal{A} = \{j : \beta_j^* \neq 0, j = 1, 2, \dots, p\}$

Table 1: Model selection (ME) and fitting results based on 100 replications for $\rho = 0.5$ in example 1

n	***pn***	**|A|**	**Model**	**MSE**	***C***	***IC***
100	35	11	**Truth**		26	0
			SMAVE	8.87 (0.65)	22.74	0.41
			ALMAVE	5.22 (0.31)	24.25	0.13
			SCADMAVE	9.05 (0.53)	22.00	0.52
			MCPMAVE	7.53 (0.62)	22.70	0.34
			SMAVE- EN	4.92 (0.24)	24.75	0.09
			SMAVE- AdEN	4.02 (0.29)	24.90	0.08
200	51	17	**Truth**		36	0
			SMAVE	6.12 (0.57)	33.24	0.12
			ALMAVE	4.01 (0.24)	34.25	0
			SCADMAVE	3.88 (0.33)	33.94	0.08
			MCPMAVE	5.50 (0.50)	33.83	0.09
			SMAVE- EN	3.22 (0.17)	35.02	0
			SMAVE- AdEN	2.94 (0.19)	35.32	0
400	75	25	**Truth**		51	0
			SMAVE	4.11 (0.17)	48.34	0
			ALMAVE	2.61 (0.18)	50.01	0

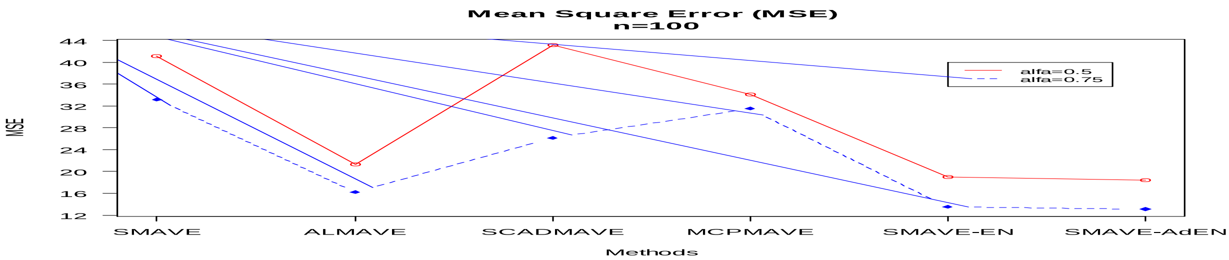
			SCADMAVE	3.02 (0.12)	49.44	0
			MCPMAVE	3.78 (0.22)	48.73	0
			SMAVE- EN	2.13 (0.13)	50.32	0
			SMAVE- AdEN	1.99 (0.10)	50.46	0

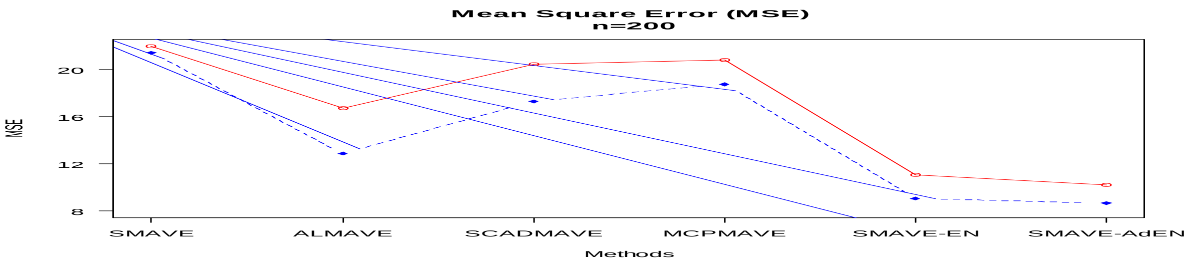
Table 2: ME and fitting results based on 100 replications for $\rho = 0.75$ in example 1

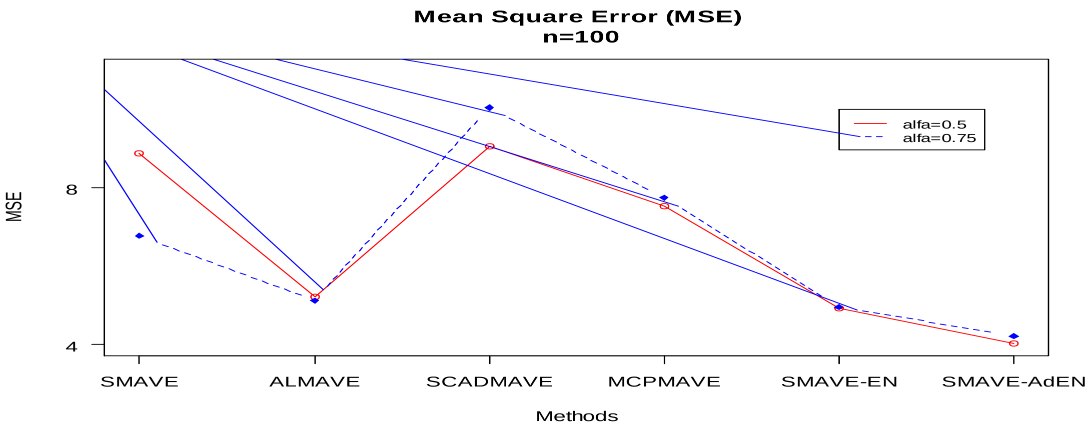
n	***pn***	|A|	**Model**	**MSE**	***C***	***IC***
100	35	11	**Truth**		26	0
			SMAVE	6.77 (0.28)	22.95	0.49
			ALMAVE	5.12 (0.17)	24.75	0.17
			SCADMAVE	10.05 (0.22)	22.13	1.11
			MCPMAVE	7.75 (0.24)	22.82	1.24
			SMAVE- EN	4.95 (0.15)	24.79	0.14
			SMAVE- AdEN	4.21 (0.18)	24.98	0.10
200	51	17	**Truth**		36	0
			SMAVE	6.44 (0.51)	34.11	0.13
			ALMAVE	4.33 (0.23)	35.01	0.03
			SCADMAVE	4.02 (0.31)	34.99	0.55
			MCPMAVE	5.61 (0.42)	34.32	0.89
			SMAVE- EN	3.35 (0.14)	35.41	0
			SMAVE- AdEN	3.12 (0.15)	35.63	0
400	75	25	**Truth**		51	0
			SMAVE	4.33 (0.13)	49.34	0.11
			ALMAVE	2.98 (0.18)	50.33	0
			SCADMAVE	3.45 (0.12)	49.95	0.07
			MCPMAVE	3.90 (0.12)	49.11	0.09
			SMAVE- EN	2.67 (0.09)	50.51	0
			SMAVE- AdEN	2.22 (0.11)	50.67	0

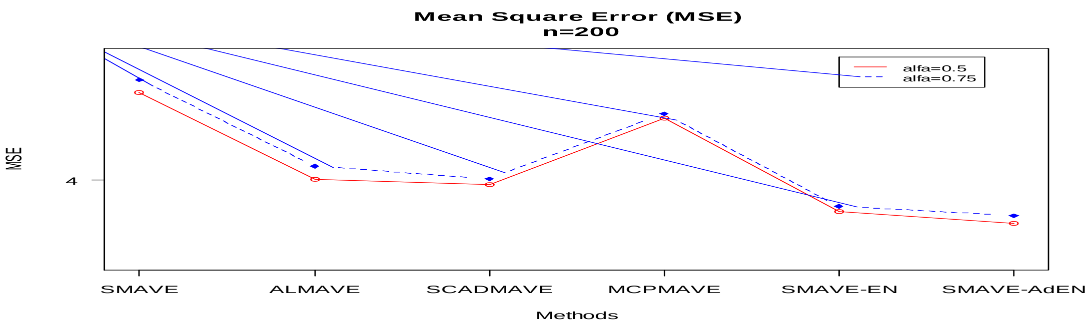
From Table 1 and 2 the prediction results can be summarized as follows. First, it is clear that the SMAVE has the worst performance. SMAVE-AdEN is considerably more accurate than all the considered methods. In general, the ALMAVE and SMAVE-EN was competitor for SMAVE-EN and its performance was better than the rest methods for all the examples. The results of simulation indicate that the SMAVE-AdEN dominates the SMAVE-EN, ALMAVE, SCAD-MAVE, P-MAVE, MCP-MAVE and SMAVE methods under collinearity.

Example 1.


Mean Square Error (MSE)
n=100
MSE
44
40
36
32
28
24
20
16
12
SMAVE
ALMAVE
SCADMAVE
MCPMAVE
SMAVE-EN
SMAVE-AdEN
Methods
alfa=0.5
alfa=0.75


Mean Square Error (MSE)
n=200
MSE
20
16
12
8
SMAVE
ALMAVE
SCADMAVE
MCPMAVE
SMAVE-EN
SMAVE-AdEN
Methods


Mean Square Error (MSE)
n=100
MSE
8
4
SMAVE
ALMAVE
SCADMAVE
MCPMAVE
SMAVE-EN
SMAVE-AdEN
Methods
alfa=0.5
alfa=0.75


Mean Square Error (MSE)
n=200
MSE
4
SMAVE
ALMAVE
SCADMAVE
MCPMAVE
SMAVE-EN
SMAVE-AdEN
Methods
alfa=0.5
alfa=0.75

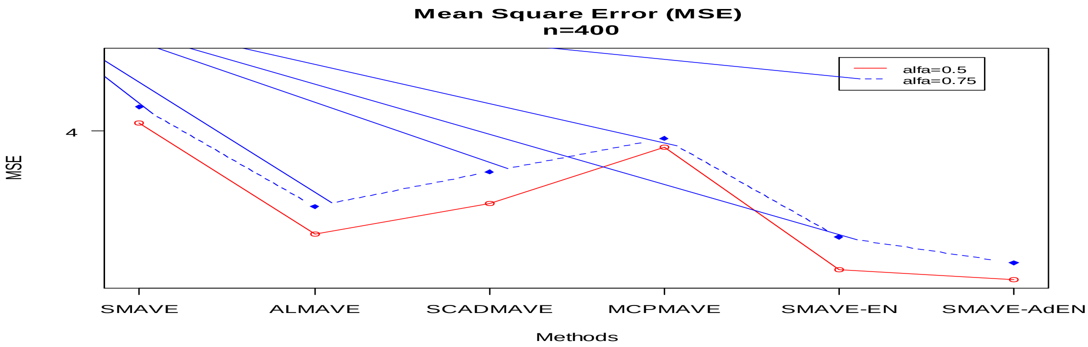


Figure 1: MSE for the considered methods based on example 1.

Table 3: ME and fitting results based on 100 replications for $\rho = 0.5$ in example 2

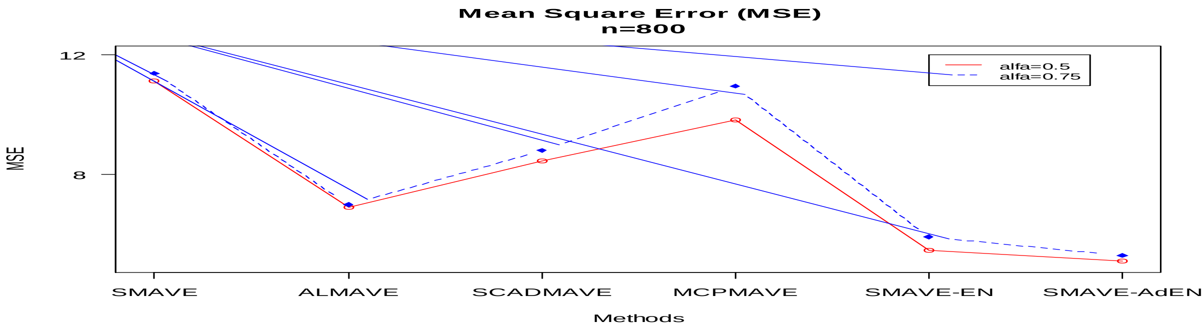


Figure 2: MSE for the considered methods based on example 2.

4. Graduate student rate (GSR) data

In this part, the GSR data has been analysed by SMAVE-AdEN, SMAVE-EN, ALMAVE, SCAD-MAVE, P-MAVE, MCP-MAVE and SMAVE methods.

The performance of the studied methods was compared via computing their prediction MSE on the test data. We apply the most important factors that affect performance of postgraduate students in Iraq data. The data had been collected by the students from the Al-Qadisiyah University in Al-Diwaniya to achieve this aim.

The researcher prepared a questionnaire to obtain the data required to study the most important factors that affect performance of postgraduate students in Iraq (University of Al-Qadisiyah as an example). To approve its validity, the text of the questionnaire was given to a jury to evaluate its appropriateness to investigate the subject under study. The jury consists of tutors of high experience and direct contact with the postgraduate students. The form of the questionnaire was first distributed to a population of postgraduate students in the college of Economy and Administration to explore their opinion about it. After getting their views, the final version of the form became ready to distribute as a procedure to collect the data required. The questionnaire is made up of 57 variables that are viewed to affect performance of the postgraduate students in Iraq. Fifty forms of the questionnaire were distributed to the postgraduate students in University of Al-Qadisiyah. After responding to the questionnaire, the forms were collected. The data were analysed by R code which was written by the author. The suggested method was compared with the methods already available. After analysing the data, we obtained some results which are mentioned in the tables 1and 2.

GSR data contains $n = 50$ observations. The response Y is GSR. The covariates are X_1 (Sex), X_2 (age), X_3 (marital state), X_4 (number of children of the students who are married), X_5 (number of brothers), X_6 (number of sisters), X_7 (students order in the family), X_8 (education level of the older brother whose level is higher than that of the student), X_9 (mother's occupation), X_{10} (father's occupation), X_{11} (education level of the mother), X_{12} (education level of the father), (family income) X_{13}, X_{14} (student income), X_{15} (how far the student is materially and morally responsible in his family), X_{16} (stable and positive family environment), X_{17} (family support to the student), X_{18} (the very ambitions fathers who practice pressure on their children to get higher average in their study) , X_{19} (death of one of the parents), X_{20} (preference of one of the two sexes on the other), X_{21} (computer and internet proficiency of the student), X_{22} (English proficiency of the student), X_{23} (teaching training and experience of the student) , X_{24} (psychology of the student), X_{25} (health of the student (chronic)), X_{26} (treatment of the student inside the classroom), X_{27} (student's participation inside the classroom), X_{28} (availability of appropriate environment in the surroundings of the student), X_{29} (number of hours assigned for reading per a day), X_{30} (worry and confusion during tests), X_{31} (support of the society when the student spends some leisure time with the others), X_{32} (full cooperation and spirit of team work among the students), X_{33} (difficulty of the study material), X_{34} (relation of teachers and students and the way teachers treat students), X_{35} (availability of references and well-equipped labs and classrooms), X_{36} (scientific level of the teachers), X_{37} (presenting lectures in modern and developed methods), X_{38} (student's average in the Baccalaureate), X_{39} (student's score in the competition test), X_{40} (ability to concentrate and paying attention), X_{41} (modernity of the topic of student's thesis), X_{42} (ability of the student to defend his thesis), X_{43} (degree of agreement between the supervisor and the student), X_{44} (the specific specialization of members of the examining committee is like that of the student's thesis), X_{45} (how much objective the members of the examining committee when granting the student's score in his thesis debate), X_{46} (how far the house of the student from the place where he is studding), X_{47} (social class of the student), X_{48} (security in the environment surrounding the student), X_{49} (degree of availability of the electrical pout), X_{50} (the duration between the last degree the student got and his current study), X_{51} (emotional relation (stability) with the other sex), X_{52} (number of years of no-pass during the undergraduate study), X_{53} (how much the student likes his specialization), X_{54} (student's satisfaction of belonging to the university), X_{55} (number of student's in the graduation group), X_{56} (future opportunities available after getting the high certificate), X_{57} (attitude of society towards the postgraduate student).

Table 5: the adjusted R-square values for the model fit based on the real data

		SMAVE-EN	**SMAVE**	**ALMAVE**	**SCADMAVE**	**MCPMAVE**	**SMAVE-AdEN**
Model Fit	*Linear*	0.93	0.77	0.93	0.77	0.77	0.93
	Quadratic	0.94	0.91	0.94	0.91	0.91	0.94
	Cubic	0.94	0.92	0.94	0.92	0.92	0.94
	Quartic	0.94	0.92	0.94	0.92	0.92	0.94

Table 5, reports the adjusted R-squared values for the model fit, based on the GSR data. The studied methods have discovered a nonlinear structure, which can be approximated by a cubic fit. Also, it can be observed that the adjusted R-squared values for the SMAVE-EN and SMAVE-AdEN methods are bigger than the values of adjusted R-squared for the SMAVE, SCADMAVE and MCPMAVE. The adjusted R-squared values for the SMAVE, SCADMAVE and MCPMAVE are similar.

Table6: the prediction error (P.E) of the cubic fit for the studied methods based on the real data

Methods	Prediction error
SMAVE	**0.8212**
ALMAVE	**0.6737**
SCADMAVE	**0.8010**
MCPMAVE	**0.8011**
SMAVE-EN	**0.6481**
SMAVE- AdEN	**0.6443**

From Table 6, it is obvious that the SMAVE-AdEN method has a lower P.E than the SMAVE-EN, SMAVE, ALMAVE, SCADMAVE and MCPMAVE methods. This means that SMAVE-AdEN has better performance than the SMAVE-EN, SMAVE, ALMAVE, SCADMAVE and MCPMAVE methods.

6. Conclusion

In this article, SMAVE-AdEN has been proposed. SMAVE-AdEN combined the AdEN with MAVE method. MAVE can estimate $S_{E(y|\mathbf{x})}$ while AdEN performs a shrinkage estimation and variable selection simultaneously and it encourages groups selection of correlated predictor. The SMAVE-AdEN benefits from the advantages of MAVE and Adaptive Elastic Net. The SMAVE-AdEN enable AdEN to work with nonlinear and multi-dimensional regression. Computationally, the SMAVE-AdEN is proved to be ease implemented with an effective algorithm. From the results of simulation and real data, it is clear that SMAVE-AdEN can have good predictive accuracy, as well as encourages groups variable selection for the strongly correlated predictors under SDR settings.

The proposed approach can be extended to SIR (Li, 1991), SAVE (Cook and Weisberg, 1991) and PHD (Li, 1992). Also, the SMAVE-AdEN can be extended to binary response models. Moreover, robust SMAVE-AdEN is another possible extension of the proposed method.

References

[1] Alkenani, A. and Yu, K. (2013). Sparse MAVE with oracle penalties. Advances and Applications in Statistics 34, 85–105.

[2] Cook, R. (1998). Regression graphics: ideas for studying the regression through graphics. New York, Wily.

[3] Cook, R. D. and Li, B. (2002). Dimension reduction for the conditional mean in regression. The Annals of Statistics 30, 455–474.

[4] Cook, R. D. and Weisberg, S. (1991). Discussion of Li (1991). Journal of the American Statistical Association 86, 328–332.

[5] Efron, B., Hastie, T., Johnstone, I. and Tibshirani, R. (2004). Least angle regression. The Annals of Statistics 32, 407–499.

[6] Fan, J. and Li, R. Z. (2001). Variable selection via non-concave penalized likelihood and its oracle properties. Journal of the American Statistical Association 96,1348–1360.

[7] Li, L. (2007). Sparse sufficient dimension reduction. Biometrika 94, 603–613.

[8] Li, K. (1991). Sliced inverse regression for dimension reduction (with discussion). Journal of the American Statistical Association 86, 316–342.

[9] Li, K. C. (1992). On principal Hessian directions for data visualization and dimension reduction: Another application of Stein’s lemma. Journal of the American Statistical Association 87, 1025–1039.

[10] Li, L., Cook, R. D. and Nachtsheim, C. J. (2005). Model-free variable selection. Journal of the Royal Statistical Society Series B, 67, 285–299.
[11] Li, L., Li, B. and Zhu, L.-X. (2010). Groupwise dimension reduction. Journal of the American Statistical Association. 105, 1188–1201.
[12] Li, L. and Nachtsheim, C. J. (2006). Sparse sliced inverse regression. Technometrics 48, 503–510.
[13] Li, L. and Yin, X. (2008). Sliced Inverse Regression with regularizations. Biometrics 64, 124–131.
[14] Ni, L., Cook, R. D. and Tsai, C. L. (2005). A note on shrinkage sliced inverse regression. Biometrika 92, 242–247.
[15] Stamey, T., Kabalin, J., McNeal, J., Johnstone, I., Freiha, F., Redwine, E. and Yang, N. (1989) Prostate specific antigen in the diagnosis and treatment of adenocarcinoma of the prostate ii: radical prostatectomy treated patients. *J. Urol.*, **16**, 1076–1083.
[16] Tibshirani, R. (1996). Regression shrinkage and selection via the Lasso. Journal of the Royal Statistical Society, Series B 58, 267–288.
[17] Wang, Q. and Yin, X. (2008). A Nonlinear Multi-Dimensional Variable Selection Method for High Dimensional Data: Sparse MAVE. Computational Statistics and Data Analysis 52, 4512–4520.
[18] Wang, T., Xu, P. and Zhu, L. (2013). Penalized minimum average variance estimation. Statist. Sinica 23 543–569.
[19] Wang, T., Xu, P., Zhu, L. (2015). Variable selection and estimation for semiparametric multiple-index models. Bernoulli 21 (1), 242–275.
[20] Xia, Y., Tong, H., Li, W. and Zhu, L. (2002). An adaptive estimation of dimension reduction space. Journal of the Royal Statistical Society Series B 64, 363–410.
[21] Yu, Z. and Zhu, L. (2013). Dimension reduction and predictor selection in semiparametric models. Biometrika, 100, 641-654.
[22] Zhang, C. H. (2010). Nearly unbiased variable selection under minimax concave penalty. Annals of Statistics 38, 894–942.
[23] Zou, H. and Hastie, T. (2005). Regularization and variable selection via the elastic net, Journal of the Royal Statistical Society, Series B 67, 301–320.
[24] Zou, H. (2006). The adaptive Lasso and its oracle properties. Journal of the American Statistical Association 101, 1418–142.

Truncated Rayleigh Pareto Distribution

Reyah Zeadan Khalaf **Kareema Abad Al-Kadim**

College of Education of Pure Sciences/ university of Babylon
Kareema.kadim@yahoo.com **zeaden123@gmail.com**

Abstract. In this paper, we introduced a new distribution which is the truncated Rayleigh Pareto distribution of a variable, and some useful functions, and some mathematical and statistical properties such as density function, the cumulative distribution function, limit function, reliability function, hazard function, stress -strength reliability, engineering arithmetic mean, mode and median, harmonic arithmetic mean, the r^{th} moment about the mean and the origin, coefficients of variation, of kurtosis and Skewedness, order statistic, and some estimation methods.
Keywords: Truncated Rayleigh Pareto distribution, density function, cumulative function, Reliability function, hazard function, reverse hazard function, cumulative hazard function, mode, median and quantile, moment generating function, mea and variance, moments, order statistics, estimation methods.

1.Introduction

In practical and scientific life we need to invent or discover a new distribution or develop a new discovery, so the importance of lifelong distribution is used in many areas of existent life like biostatistics, survival function analysis or reliability, and in this field we will look at the amputated Rayleigh Pareto distribution the new. The aim of the research is to find the new (TRPD) distribution by amputating the period for the distribution (RPD). Using methods such as those used by Wen-liany Hung and ching- yichen (2004), Approximmate MLE of the scale parameter of the truncated Rayleigh distribution under the first – censored data[1], David R-clark, FCAS (2013), A note on the upper- truncated Pareto distribution[2]. Taylor and francis (2014), Parameter Estimation for the Truncated Pareto Distribution[3]. Mathias Reachke (2012), Inference for the truncated exponential distribution[4], and Ahmed Yassin Taqi M.CS. Thesis (2014), Estimation of Parameters and Reliability Function for truncated Logistic Distribution[5]. In the second section, we examined the description of the amputated Rayleigh Pareto distribution and the study of its mathematical and statistical properties, and the methods of estimating and comparing them.

2. Truncated Rayleigh – Pareto distribution.

This is the distribution in which the data is amputated and at the point $x_\circ$, where $x_\circ$ is a constant value. This means that the random drawn values are located between $0 < x < x_\circ$ the pdf for Rayleigh – Pareto distribution is.[6],[7],[8]

$$f_{R.P}(x;q,w,\alpha) = \frac{\alpha}{2q^2 w}\left(\frac{x}{w}\right)^{\alpha-1} e^{-\frac{1}{2q^2}\left(\frac{x}{w}\right)^{\alpha}}, \quad 0 < x \leq x_\circ \;\&\; q,w,\alpha > 0 \qquad (1)$$

2.1 The PDF and CDF of TRPD

The p.d.f of the (TRPD) can be given .

$$g(x;q,w,\alpha) = \frac{f(x;q,w,\alpha)}{F(b)-F(a)} = \frac{\frac{\alpha}{2q^2 w}\left(\frac{x}{w}\right)^{\alpha-1} e^{-\frac{1}{2q^2}\left(\frac{x}{w}\right)^{\alpha}}}{1-e^{-\frac{1}{2q^2}\left(\frac{x}{w}\right)^{\alpha}}}, \quad 0 < x \leq x_\circ \;\&\; q,w,\alpha > 0 \qquad (2)$$

Such as $\int_0^{x_\circ} \frac{\frac{\alpha}{2q^2 w}\left(\frac{x}{w}\right)^{\alpha-1} e^{-\frac{1}{2q^2}\left(\frac{x}{w}\right)^{\alpha}}}{1-e^{-\frac{1}{2q^2}\left(\frac{x}{w}\right)^{\alpha}}}\partial x = 1$

The plot of p.d.f. and p.d.f. for the(TRPD) as follows.

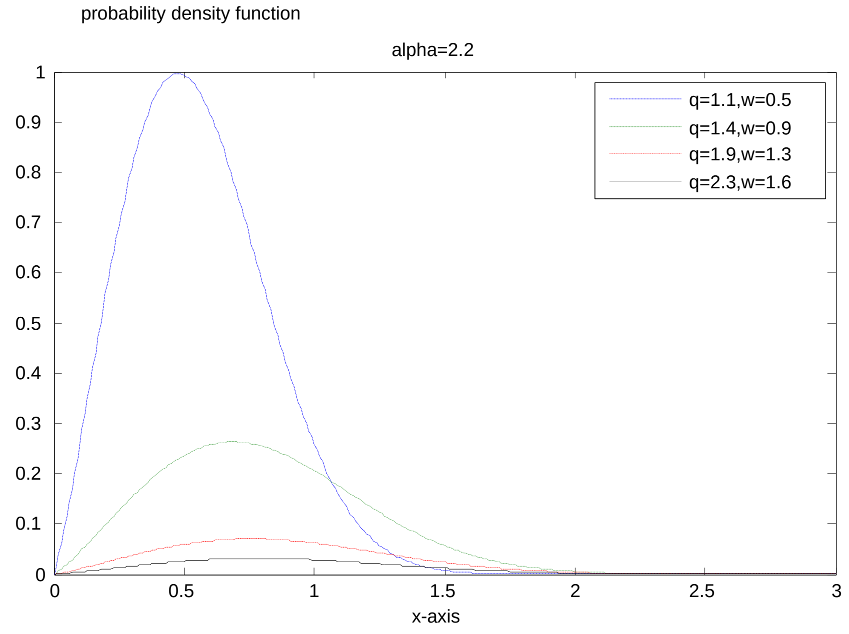


Figure. 1: The plot to p.d.f. for (TRPD), the parameter alpha=2.2 q=1.1,1.4,1.9,2.3 w=0.5,0.9,1.3,1.6

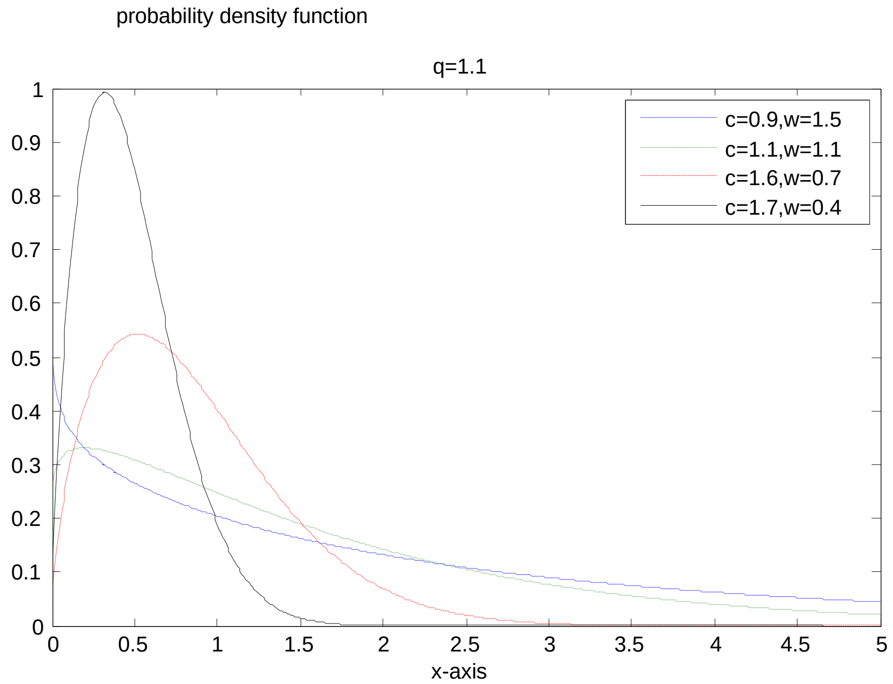


Figure 2 : The plot to p.d.f. for (TRPD), the parameter q=1.1 c=0.9,1.1,1.6,1.7 ; w=1.5,1.1,0.7,0.4

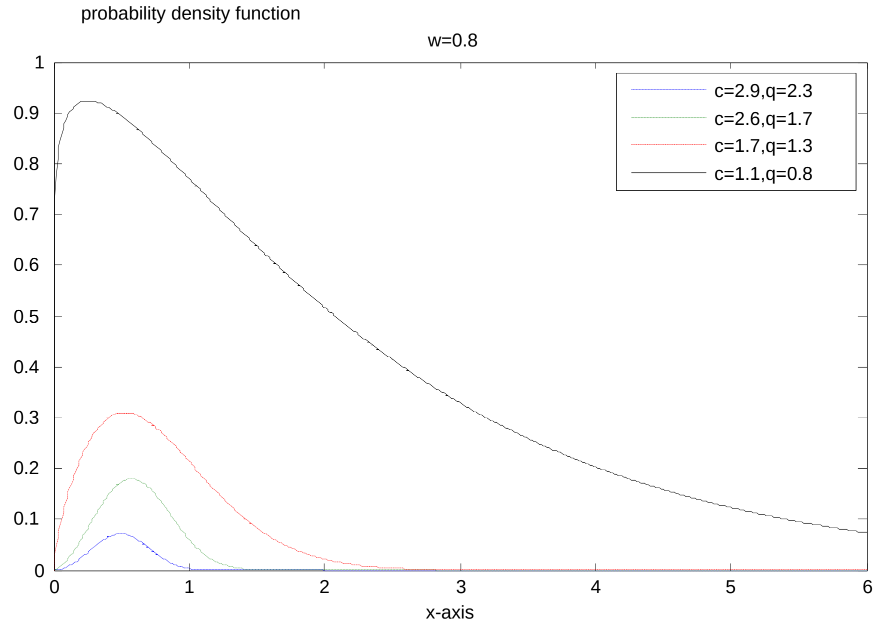
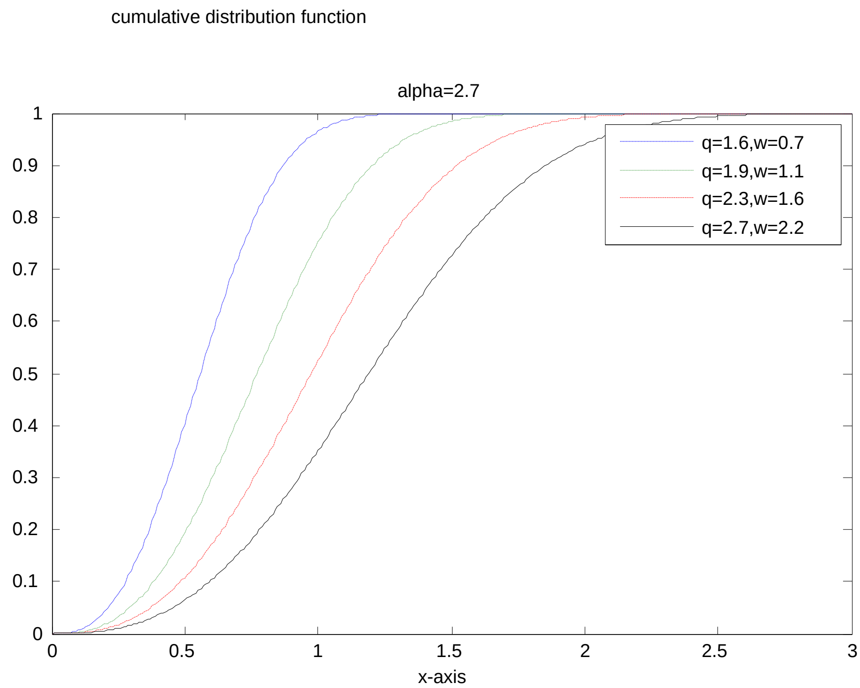


Figure 3: The plot to p.d.f. for (TRPD), the parameter w=0.8 c=2.9,2.6,1.7,1.1 ; q=2.3,1.7,1.3,0.8

And the c.d.f of the (TRPD)is given by

$$G(x;q,w,\alpha)=\int_0^x g(x;q,w,\alpha)\partial x=\int_0^x \frac{\frac{\alpha}{2q^2w}\left(\frac{x}{w}\right)^{\alpha-1}e^{-\frac{1}{2q^2}\left(\frac{x}{w}\right)^{\alpha}}}{1-e^{-\frac{1}{2q^2}\left(\frac{x_\circ}{w}\right)^{\alpha}}}\partial x$$

$$=\frac{1-e^{-\frac{1}{2q^2}\left(\frac{x}{w}\right)^{\alpha}}}{1-e^{-\frac{1}{2q^2}\left(\frac{x_\circ}{w}\right)^{\alpha}}}\,,\qquad 0<x\le x_\circ\ \&\ q,w,\alpha>0 \qquad (3)$$



Figur 4: The plot to c.d.f. for (TRPD), the parameter alpha=2.7; q= 1.6,1.9,2.3,2.7; w= 0.7,1.1,1.6,2.2

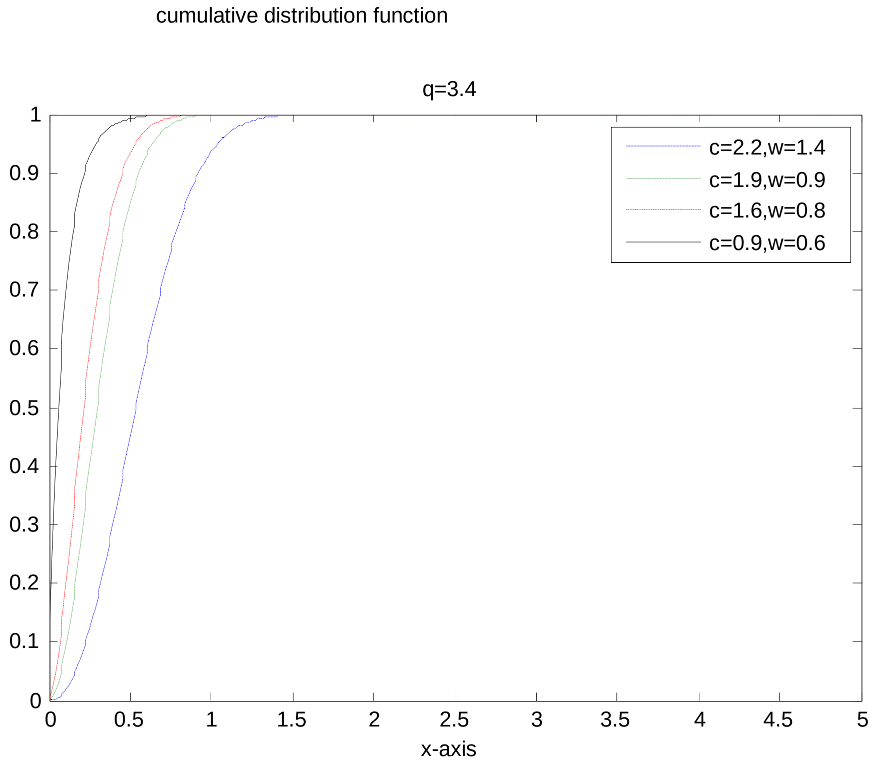


Figure 5: The plot to c.d.f. for (TRPD), the parameter q=3.4; c= 2.2,1.9,1.6,0.9; w= 1.4,0.9,0.8,0.6

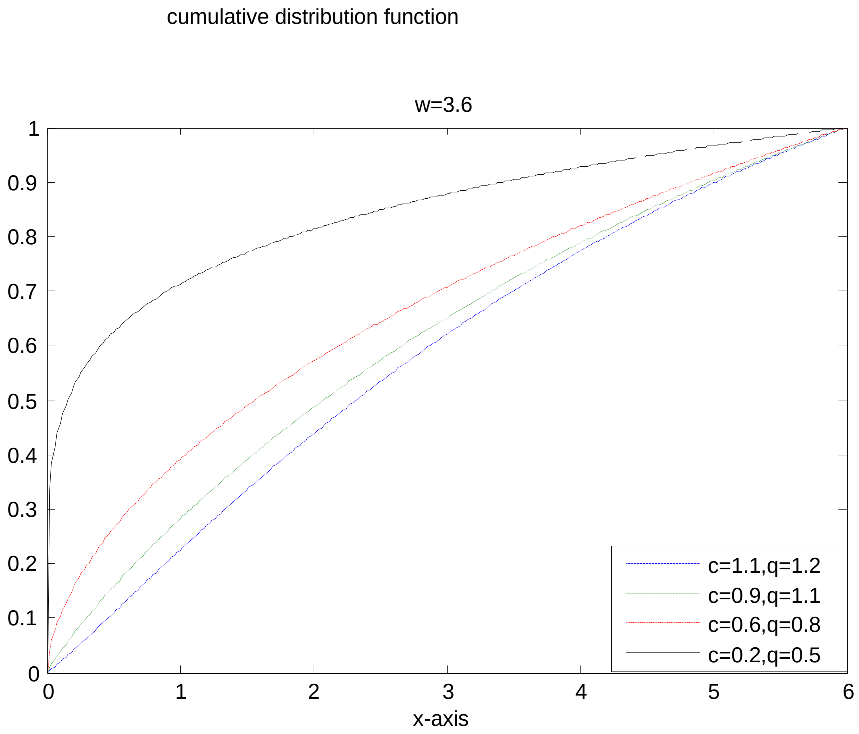


Figure 6: The plot to c.d.f. for (TRPD), the parameter w=3.6; c=1.1,0.9,0.6,0.2; q=1.2,1.1,0.8,0.5

2.2. Limit of p.d.f and c.d.f

The limit of this distribution is given by the form a:

$$\lim_{x\to 0} g(x;q,w,\alpha) = 0 \qquad (4)$$

$$\lim_{x\to 0}\frac{\frac{\alpha}{2q^2w}\left(\frac{x}{w}\right)^{\alpha-1}e^{-\frac{1}{2q^2}\left(\frac{x}{w}\right)^{\alpha}}}{1-e^{-\frac{1}{2q^2}\left(\frac{x_\circ}{w}\right)^{\alpha}}} = \frac{\alpha}{2q^2\left[1-e^{-\frac{1}{2q^2}\left(\frac{x_\circ}{w}\right)^{\alpha}}\right]}\lim_{x\to 0}\left(\frac{x}{w}\right)^{\alpha-1}e^{-\frac{1}{2q^2}\left(\frac{x}{w}\right)^{\alpha}} = 0$$

Also

$$\lim_{x\to x_\circ} g(x;q,w,\alpha) = \lim_{x\to x_\circ}\frac{\frac{\alpha}{2q^2w}\left(\frac{x}{w}\right)^{\alpha-1}e^{-\frac{1}{2q^2}\left(\frac{x}{w}\right)^{\alpha}}}{1-e^{-\frac{1}{2q^2}\left(\frac{x_\circ}{w}\right)^{\alpha}}} = \frac{\frac{\alpha}{2q^2w}\left(\frac{x_\circ}{w}\right)^{\alpha-1}e^{-\frac{1}{2q^2}\left(\frac{x_\circ}{w}\right)^{\alpha}}}{1-e^{-\frac{1}{2q^2}\left(\frac{x_\circ}{w}\right)^{\alpha}}} > 0 \qquad (5)$$

Because $1-e^{-\frac{1}{2q^2}\left(\frac{x_\circ}{w}\right)^{\alpha}} > 0$, and $\lim_{x\to 0}\frac{\alpha}{2q^2w}\left(\frac{x}{w}\right)^{\alpha-1}e^{-\frac{1}{2q^2}\left(\frac{x}{w}\right)^{\alpha}} > 0$

Also the c.d.f. of this distribution is

$$\lim_{x\to 0} G(x;q,w,\alpha) = \lim_{x\to 0}\frac{1-e^{-\frac{1}{2q^2}\left(\frac{x}{w}\right)^{\alpha}}}{1-e^{-\frac{1}{2q^2}\left(\frac{x_\circ}{w}\right)^{\alpha}}} = 0$$

Also,

$$\lim_{x\to x_\circ} G(x;q,w,\alpha) = \lim_{x\to x_\circ}\frac{1-e^{-\frac{1}{2q^2}\left(\frac{x}{w}\right)^{\alpha}}}{1-e^{-\frac{1}{2q^2}\left(\frac{x_\circ}{w}\right)^{\alpha}}} = 1 \qquad (6)$$

i.e $0 \le G(x;q,w,\alpha) \le 1$

3. Some Reliability Functions. [9]

In this section, we introduce some reliability functions for the TRPD.

3.1. Reliability Function

The function of survival or reliability to TRPD are.

$$R(x) = 1 - G(x) = 1 - \frac{1-e^{-\frac{1}{2q^2}\left(\frac{x}{w}\right)^{\alpha}}}{1-e^{-\frac{1}{2q^2}\left(\frac{x_\circ}{w}\right)^{\alpha}}} = \frac{e^{-\frac{1}{2q^2}\left(\frac{x}{w}\right)^{\alpha}} - e^{-\frac{1}{2q^2}\left(\frac{x_\circ}{w}\right)^{\alpha}}}{1-e^{-\frac{1}{2q^2}\left(\frac{x_\circ}{w}\right)^{\alpha}}} \qquad (7)$$

Such as

$$\lim_{x\to 0} R(x) = \lim_{x\to 0} \frac{e^{-\frac{1}{2q^2}\left(\frac{x}{w}\right)^{\alpha}} - e^{-\frac{1}{2q^2}\left(\frac{x_s}{w}\right)^{\alpha}}}{1 - e^{-\frac{1}{2q^2}\left(\frac{x_s}{w}\right)^{\alpha}}} = 1$$

$$\lim_{x\to x_s} R(x) = \lim_{x\to x_s} \frac{e^{-\frac{1}{2q^2}\left(\frac{x}{w}\right)^{\alpha}} - e^{-\frac{1}{2q^2}\left(\frac{x_s}{w}\right)^{\alpha}}}{1 - e^{-\frac{1}{2q^2}\left(\frac{x_s}{w}\right)^{\alpha}}} = 0$$

3.2. Hazard Function

TRPD's hazard function :

$$h(x) = \frac{g(x; q, w, \alpha)}{R(x)} \tag{8}$$

$$h(x) = \frac{\dfrac{\frac{\alpha}{2q^2 w}\left(\frac{x}{w}\right)^{\alpha-1} e^{-\frac{1}{2q^2}\left(\frac{x}{w}\right)^{\alpha}}}{1 - e^{-\frac{1}{2q^2}\left(\frac{x_s}{w}\right)^{\alpha}}}}{\dfrac{e^{-\frac{1}{2q^2}\left(\frac{x}{w}\right)^{\alpha-1}} - e^{-\frac{1}{2q^2}\left(\frac{x_s}{w}\right)^{\alpha}}}{1 - e^{-\frac{1}{2q^2}\left(\frac{x_s}{w}\right)^{\alpha}}}} = \frac{\frac{\alpha}{2q^2 w}\left(\frac{x}{w}\right)^{\alpha-1} e^{-\frac{1}{2q^2}\left(\frac{x}{w}\right)^{\alpha}}}{e^{-\frac{1}{2q^2}\left(\frac{x}{w}\right)^{\alpha}} - e^{-\frac{1}{2q^2}\left(\frac{x_s}{w}\right)^{\alpha}}}, \tag{9}$$

3.3. Reverse Hazard Function

TRPD's Reverse Hazard function :

$$r(x) = \frac{g(x; q, w, \alpha)}{G(x; q, w, \alpha)} = \frac{\dfrac{\frac{\alpha}{2q^2 w}\left(\frac{x}{w}\right)^{\alpha-1} e^{-\frac{1}{2q^2}\left(\frac{x}{w}\right)^{\alpha}}}{1 - e^{-\frac{1}{2q^2}\left(\frac{x_s}{w}\right)^{\alpha}}}}{\dfrac{1 - e^{-\frac{1}{2q^2}\left(\frac{x}{w}\right)^{\alpha}}}{1 - e^{-\frac{1}{2q^2}\left(\frac{x_s}{w}\right)^{\alpha}}}} = \frac{\frac{\alpha}{2q^2 w}\left(\frac{x}{w}\right)^{\alpha-1} e^{-\frac{1}{2q^2}\left(\frac{x}{w}\right)^{\alpha}}}{1 - e^{-\frac{1}{2q^2}\left(\frac{x}{w}\right)^{\alpha}}},$$

Such as $x, \alpha, q, w > o$ **(10)**

3.4. The Cumulative Hazard Function

TRPD's cumulative hazard function

$$H(x) = -Ln\ R(x) = -Ln\ \frac{e^{-\frac{1}{2q^2}\left(\frac{x}{w}\right)^{\alpha}} - e^{-\frac{1}{2q^2}\left(\frac{x_s}{w}\right)^{\alpha}}}{1 - e^{-\frac{1}{2q^2}\left(\frac{x_s}{w}\right)^{\alpha}}}, \tag{11}$$

Such as $x, \alpha, q, w > 0$

3.5. Stress –Strength Reliability.[10]

Let Y represent the pressure applied to a specific device and X represents the strength to maintain the pressure, since the reliability of the stress force is indicated by R = P (Y <X), if X and Y are assumed to be random.

$$R = P(Y < X) = \int_0^{x_\circ} g_x(x) G_y(x)\partial X \qquad \textbf{(12)}$$

$$= \int_0^{x_\circ} \frac{\frac{\alpha}{2q^2 w}\left(\frac{x}{w}\right)^{\alpha-1} e^{-\frac{1}{2q^2}\left(\frac{x}{w}\right)^{\alpha}}}{1 - e^{-\frac{1}{2q^2}\left(\frac{x_\circ}{w}\right)^{\alpha}}} \times \frac{1 - e^{-\frac{1}{2q^2}\left(\frac{x}{w}\right)^{\alpha}}}{1 - e^{-\frac{1}{2q^2}\left(\frac{x_\circ}{w}\right)^{\alpha}}}\partial x$$

$$= \frac{1}{\left(1 - e^{-\frac{1}{2q^2}\left(\frac{x_\circ}{w}\right)^{\alpha}}\right)^2}\int_0^{x_\circ}\left[\frac{\alpha}{2q^2 w}\left(\frac{x}{w}\right)^{\alpha-1} e^{-\frac{1}{2q^2}\left(\frac{x}{w}\right)^{\alpha}} - \frac{\alpha}{2q^2 w}\left(\frac{x}{w}\right)^{\alpha-1}\left(e^{-\frac{1}{2q^2}\left(\frac{x}{w}\right)^{\alpha}}\right)^2\right]\partial x$$

$$= \frac{1}{\left(1 - e^{-\frac{1}{2q^2}\left(\frac{x_\circ}{w}\right)^{\alpha}}\right)^2}\left[\int_0^{x_\circ}\frac{\alpha}{2q^2 w}\left(\frac{x}{w}\right)^{\alpha-1} e^{-\frac{1}{2q^2}\left(\frac{x}{w}\right)^{\alpha}}\partial x - \int_0^{x_\circ}\frac{\alpha}{2q^2 w}\left(\frac{x}{w}\right)^{\alpha-1}\left(e^{-\frac{1}{2q^2}\left(\frac{x}{w}\right)^{\alpha}}\right)^2\partial x\right]$$

$$= \frac{1}{\left(1 - e^{-\frac{1}{2q^2}\left(\frac{x_\circ}{w}\right)^{\alpha}}\right)^2}\left[\left(1 - e^{-\frac{1}{2q^2}\left(\frac{x_\circ}{w}\right)^{\alpha}}\right) - \left(\frac{\alpha}{2q^2 w^{\alpha}}\int_0^{x_\circ} x^{\alpha-1}\left(e^{-\frac{1}{2q^2}\left(\frac{x}{w}\right)^{\alpha}}\right)^2\partial x\right]$$

By using series expansion of $e^{-\frac{1}{2q^2}\left(\frac{x}{w}\right)^{\alpha}}$ then .

$$e^{-\frac{1}{2q^2}\left(\frac{x}{w}\right)^{\alpha}} = \sum_{j=0}^{\infty}\frac{\left(-\frac{1}{2q^2}\right)^j\left(\frac{x}{w}\right)^{\alpha j}}{j!}$$

Then

$$P(Y < X) = \frac{1}{\left(1-e^{-\frac{1}{2q^2}\left(\frac{x_\circ}{w}\right)^{\alpha}}\right)^2}\left[\left(1 - e^{-\frac{1}{2q^2}\left(\frac{x_\circ}{w}\right)^{\alpha}}\right) - \sum_{j=0}^{\infty}\frac{(-1)^{2j}\alpha}{(j!)^2(2q^2)^{2j+1}w^{(2j+1)\alpha}}\int_0^{x_\circ} x^{(2j+1)\alpha-1}\partial x\right.$$

the stress-strength reliability is

$$P(Y < X) = \frac{1}{1-e^{-\frac{1}{2q^2}\left(\frac{x_\circ}{w}\right)^{\alpha}}} - \sum_{j=0}^{\infty}\frac{\alpha(-1)^{2j}(x_\circ)^{(2j+1)\alpha}}{(j!)^2(2q^2)^{2j+1}w^{(2j+1)\alpha}\left(1-e^{-\frac{1}{2q^2}\left(\frac{x_\circ}{w}\right)^{\alpha}}\right)^2(2j+1)\alpha}$$

(13)

4.Some properties of the TRPD :

Proposition(1)

***4.1 mode*: The mode of the TRPD is .**

$$x_{mode} = w\left(\frac{(\alpha-1)(2q^2)}{\alpha}\right)^{\frac{1}{\alpha}} \qquad \textbf{(14)}$$

Proof: $x_{mode} = \frac{\partial Ln\, g(x_{mode},q,w,\alpha)}{\partial x} = 0$ (15)

$$\frac{\partial ln\, g(x_{mode}, q, w, \alpha)}{\partial x} = \frac{\partial}{\partial x}\left[ln \frac{\frac{\alpha}{2q^2 w} (\frac{x}{w})^{\alpha-1} e^{-\frac{1}{2q^2}(\frac{x}{w})^{\alpha}}}{1- e^{-\frac{1}{2q^2}(\frac{x_z}{w})^{\alpha}}} \right] = 0 \Rightarrow$$

$$\frac{\partial}{\partial x}[ln\alpha - ln2q^2 - lnw + (\alpha-1)lnx - (\alpha-1)lnw - \frac{1}{2q^2}(\frac{x}{w})^{\alpha} - ln(1-e^{-\frac{1}{2q^2}(\frac{x_z}{w})^{\alpha}})] = 0$$

$$\Rightarrow \frac{(\alpha-1)}{x_{mode}} - \frac{\alpha}{2q^2 w}(\frac{x_{mode}}{w})^{\alpha-1} = 0 \Rightarrow \frac{(\alpha-1)}{x_{mode}} = \frac{\alpha}{2q^2 w}(\frac{x_{mode}}{w})^{\alpha-1}$$

$$x_{mode} = w(\frac{(\alpha-1)(2q^2)}{\alpha})^{\frac{1}{\alpha}} \quad (16)$$

Proposition(2)

4.2 Median and Quintile:

The quintile x_q and median of the TRPD is

$$x_q = w[-2q^2\, Ln\, (1 - Q\left[1- e^{-\frac{1}{2q^2}(\frac{x_z}{w})^{\alpha}}\right]]^{\frac{1}{\alpha}} \quad (17)$$

$$x_{median} = w[-2q^2\, Ln\, \frac{1}{2}\left[1+ e^{-\frac{1}{2q^2}(\frac{x_z}{w})^{\alpha}}\right]]^{\frac{1}{\alpha}} \quad (18)$$

Proof:

$$G(x_q; q, w, \alpha) = Q(x_Q \leq Q) = Q \quad (19)$$

From equation (2) we obtain.

$$\frac{1- e^{-\frac{1}{2q^2}(\frac{x_Q}{w})^{\alpha}}}{1- e^{-\frac{1}{2q^2}(\frac{x_z}{w})^{\alpha}}} = Q \;\Rightarrow\; 1- e^{-\frac{1}{2q^2}(\frac{x}{w})^{\alpha}} = Q\left[1- e^{-\frac{1}{2q}(\frac{x_z}{w})^{\alpha}}\right] \Rightarrow$$

$e^{-\frac{1}{2q^2}(\frac{x_Q}{w})^{\alpha}} = 1 - Q\left[1- e^{-\frac{1}{2q^2}(\frac{x_z}{w})^{\alpha}}\right]$ Take Ln to the two parties

$$-\frac{1}{2q^2}(\frac{x_Q}{w})^{\alpha} = Ln\left(1- Q\left[1- e^{-\frac{1}{2q^2}(\frac{x_z}{w})^{\alpha}}\right]\right)$$

Then

$$x_Q = w[-2q^2 Ln\,(1 - Q\left[1 - e^{-\frac{1}{2q^2}(\frac{x_o}{w})^\alpha}\right]]^{\frac{1}{\alpha}} \tag{20}$$

To find the median of (TRPD) we set $Q = \frac{1}{2}$ in equation (20).

$$x_{median} = w[-2q^2 Ln\,(1 - \frac{1}{2}\left[1 - e^{-\frac{1}{2q^2}(\frac{x_o}{w})^\alpha}\right]]^{\frac{1}{\alpha}} \Rightarrow$$

$$x_{median} = w[-2q^2 Ln\,\frac{1}{2}\left[1 + e^{-\frac{1}{2q^2}(\frac{x_o}{w})^\alpha}\right]]^{\frac{1}{\alpha}} \tag{21}$$

4.3 Moment Generating Function:

Proposition(3)

The moment generating function of TRPD as follows:

$$M_x(t) = \sum_{j=0}^{\infty}\sum_{s=0}^{\infty} \frac{t^s \alpha(-1)^j\,(x_o)^{s+(j+1)\alpha}}{j!\,s!\,(2q^2)^{j+1} w^{(j+1)}\left[1 - e^{-\frac{1}{2q^2}(\frac{x_o}{w})^\alpha}\right](s+(j+1)\alpha)} \tag{22}$$

Proof:

$$M_x(t) = E(e^{tx}) = \int_0^{x_o} e^{xt} \frac{\frac{\alpha}{2q^2}(\frac{x}{w})^{\alpha-1} e^{-\frac{1}{2q^2}(\frac{x}{w})^\alpha}}{1 - e^{-\frac{1}{2q^2}(\frac{x_o}{w})^\alpha}}$$

By using series expansion of e^{xt} then.

$$e^{tx} = \sum_{s=0}^{\infty} \frac{t^s x^s}{s!}$$

So

$$= \frac{\alpha}{2q^2 w^\alpha\left[1 - e^{-\frac{1}{2q^2}(\frac{x_o}{w})^\alpha}\right]} \sum_{s=0}^{\infty} \frac{t^s}{s!} \int_0^{x_o} x^{s+\alpha-1} e^{-\frac{1}{2q^2}(\frac{x}{w})^\alpha}$$

By using series expansion of $e^{-\frac{1}{2q^2}(\frac{x}{w})^\alpha}$ then.

$$e^{-\frac{1}{2q^2}(\frac{x}{w})^\alpha} = \sum_{j=0}^{\infty} \frac{(-\frac{1}{2q^2})^j (\frac{x}{w})^{j\alpha}}{j!}$$

Then

$$M_x(t)=\sum_{j=0}^{\infty}\sum_{s=0}^{\infty}\frac{t^s\alpha(-1)^j}{j!\,s!\,(2q^2)^{j+1}w^{(j+1)\alpha}\left[1-e^{-\frac{1}{2q^2}(\frac{x_\circ}{w})^\alpha}\right]}\int_0^{x_\circ}x^{s+(j+1)\alpha-1}\partial x$$

$$M_x(t)=\sum_{j=0}^{\infty}\sum_{s=0}^{\infty}\frac{t^s\alpha(-1)^j\ (x_\circ)^{s+(j+1)\alpha}}{j!\,s!\,(2q^2)^{j+1}w^{(j+1)}\left[1-e^{-\frac{1}{2q^2}(\frac{x_\circ}{w})^\alpha}\right](s+(j+1)\alpha)} \tag{23}$$

4.4. Mean and variance:

Proposition (4)

The mean and variance of TRPD

$$E(X)=\sum_{j=0}^{\infty}\frac{\alpha(-1)^j\,(x_\circ)^{(j+1)\alpha+1}}{j!\,(2q^2)^{j+1}(w^{\alpha(j+1)})\left[1-e^{-\frac{1}{2q^2}(\frac{x_\circ}{w})^\alpha}\right](j+1)\alpha+1} \tag{24}$$

$$Var=E(X^2)-[E(X)]^2=\left[\sum_{j=0}^{\infty}\frac{\alpha(-1)^j(x_\circ)^{(j+1)\alpha+2}}{j!\,(2q^2)^{j+1}\ w^{(j+1)\alpha}[1-e^{-\frac{1}{2q^2}(\frac{x_\circ}{w})^\alpha}]}\right]$$

$$-\left[\sum_{j=0}^{\infty}\frac{\alpha(-1)^j\,(x_\circ)^{(j+1)\alpha+1}}{j!\,(2q^2)^{j+1}(w^{\alpha(j+1)})\left[1-e^{-\frac{1}{2q^2}(\frac{x_\circ}{w})^\alpha}\right](j+1)\alpha+1}\right]^2 \tag{25}$$

Proof:

$$E(X)=\int_0^{x_\circ}x\frac{\frac{\alpha}{2q^2w}\left(\frac{x}{w}\right)^{\alpha-1}e^{-\frac{1}{2q^2}(\frac{x}{w})^\alpha}}{1-e^{-\frac{1}{2q^2}(\frac{x_\circ}{w})^\alpha}}\partial x=\frac{\alpha}{2q^2w^\alpha[1-e^{-\frac{1}{2q^2}(\frac{x_\circ}{w})^\alpha}]}\int_0^{x_\circ}x^\alpha e^{-\frac{1}{2q^2}(\frac{x}{w})^\alpha}\partial x$$

$$e^{-\frac{1}{2q^2}(\frac{x}{w})^\alpha}=\sum_{j=0}^{\infty}\frac{(-\frac{1}{2q^2})^j\ (\frac{x}{w})^{j\alpha}}{j!}$$

Then

$$E(X)=\frac{\alpha}{2q^2w^\alpha[1-e^{-\frac{1}{2q^2}(\frac{x_\circ}{w})^\alpha}]}\sum_{j=0}^{\infty}\frac{(-1)^j}{j!\ w^{j\alpha}\ (2q^2)^j}\int_0^{x_\circ}x^{(j+1)\alpha}\,\partial x$$

We can find $\int_0^{x_\circ}x^{(j+1)\alpha}\,\partial x$ as the follows.

$$\int_0^{x_\circ}x^{(j+1)\alpha}\partial x=\left[\frac{x^{(j+1)\alpha+1}}{(j+1)\alpha+1}\right]_0^{x_\circ}=\frac{x_\circ^{(j+1)\alpha+1}}{(j+1)\alpha+1}$$

Then

$$E(X) = \sum_{j=0}^{\infty} \frac{\alpha(-1)^j x_\circ^{(j+1)\alpha+1}}{j!(2q^2)^{j+1} w^{\alpha(j+1)} [1-e^{-\frac{1}{2q^2}(\frac{x_\circ}{w})^\alpha}] (j+1)\alpha+1} \quad (26)$$

The variance of TRPD as follows.

$$E(X^2) = \int_0^{x_\circ} x^2 \frac{\frac{\alpha}{2q^2 w} (\frac{x}{w})^{\alpha-1} e^{-\frac{1}{2q^2}(\frac{x}{w})^\alpha}}{1-e^{-\frac{1}{2q^2}(\frac{x_\circ}{w})^\alpha}} \partial x = \sum_{j=0}^{\infty} \frac{\alpha(-1)^j x_\circ^{(j+1)\alpha+2}}{j!(2q^2)^{j+1} w^{\alpha(j+1)} [1-e^{-\frac{1}{2q^2}(\frac{x_\circ}{w})^\alpha}] (j+1)\alpha+2}$$

Then

$$Var = [\sum_{j=0}^{\infty} \frac{\alpha(-1)^j (x_\circ)^{(j+1)\alpha+2}}{j!\,(2q^2)^{j+1} w^{(j+1)\alpha} [1-e^{-\frac{1}{2q^2}(\frac{x_\circ}{w})^\alpha}]}] - [\sum_{j=0}^{\infty} \frac{\alpha(-1)^j (x_\circ)^{(j+1)\alpha+1}}{j!\,(2q^2)^{j+1} (w^{\alpha(j+1)}) \left[1-e^{-\frac{1}{2q^2}(\frac{x_\circ}{w})^\alpha}\right] (j+1)\alpha+1}]^2 \quad (27)$$

4.5 Moments

Proposition (5)

The r^{th} moment about the origin and mean of the TRPD.

$$E(X^r) = \sum_{j=0}^{\infty} \frac{\alpha(-1)^j (x_\circ)^{(j+1)\alpha+r}}{j!\,(2q^2)^{j+1} w^{(j+1)\alpha} [1-e^{-\frac{1}{2q^2}(\frac{x_\circ}{w})^\alpha}] (j+1)\alpha+r}$$

***r* = 1,2,3** **(28)**

$$E(X-\mu)^r = \sum_{i=0}^{r} C_i^r (-\mu)^i \sum_{j=0}^{\infty} \frac{\alpha(-1)^j (x_\circ)^{(j+1)\alpha+1}}{j!\,(2q^2)^{j+1} (w^{\alpha(j+1)}) \left[1-e^{-\frac{1}{2q^2}(\frac{x_\circ}{w})^\alpha}\right] (j+1)\alpha+1}$$

r = 1,2,3,…. **(29)**

Proof: The r-th about the origin is:

$$E(X^r) = \int_0^{x_\circ} X^r g(x;q,w,\alpha) \partial x = \int_0^{x_\circ} X^r \frac{\frac{\alpha}{2q^2 w} (\frac{x}{w})^{\alpha-1} e^{-\frac{1}{2q^2}(\frac{x}{w})^\alpha}}{1-e^{-\frac{1}{2q^2}(\frac{x_\circ}{w})^\alpha}} \partial x$$

$$= \frac{\alpha}{2q^2 w^\alpha [1-e^{-\frac{1}{2q^2}(\frac{x_\circ}{w})^\alpha}]} \int_0^{x_\circ} X^{\alpha+r-1} e^{-\frac{1}{2q^2}(\frac{x}{w})^\alpha} \partial x$$

$$= \frac{\alpha}{2q^2 w^\alpha [1 - e^{-\frac{1}{2q^2}(\frac{x_\circ}{w})^\alpha}]} \sum_{j=0}^{\infty} \frac{(-\frac{1}{2q^2})^j (\frac{x}{w})^{j\alpha}}{j!} \int_0^{x_\circ} x^{(j+1)\alpha + r - 1} \partial x$$

Then

$$E(X^r) = \sum_{j=0}^{\infty} \frac{\alpha(-1)^j (x_\circ)^{(j+1)\alpha+r}}{j!\,(2q^2)^{j+1} w^{(j+1)\alpha} [1 - e^{-\frac{1}{2q^2}(\frac{x_\circ}{w})^\alpha}][(j+1)\alpha + r]}$$

(30)

r^{th} moment about mean is given by:

$$E(X-\mu)^r = \int_0^{x_\circ} (X-\mu)^r \frac{\frac{\alpha}{2q^2 w} (\frac{x}{w})^{\alpha-1} e^{-\frac{1}{2q^2}(\frac{x}{w})^\alpha}}{1 - e^{-\frac{1}{2q^2}(\frac{x_\circ}{w})^\alpha}} \partial x$$

We have

$$(X-\mu)^r = \sum_{i=0}^{r} C_i^r \,(x)^{r-i} (-\mu)^i$$

Thus r^{th} moment about mean is

$$E(X-\mu)^r = \sum_{i=0}^{r} \sum_{j=0}^{\infty} C_i^r \,(-\mu)^i \frac{\alpha(-1)^j (x_\circ)^{(j+1)\alpha+r-i}}{j!\,(2q^2)^{j+1} (w^{\alpha(j+1)}) \left[1 - e^{-\frac{1}{2q^2}(\frac{x_\circ}{w})^\alpha}\right] (j+1)\alpha + r - i}$$

(31)

5. Order Statistics

Suppose that $x_1, x_2, \ldots, x_n$ denoted a random sample of size n from a TRPD with $G(x; q, w, \alpha)$ and $g(x; q, w, \alpha)$ in the equation (2) and (3) . Let $Y_{k1}, Y_{k2}, \ldots, Y_{kn}$ express the corresponding order statistics ;then.

The p.d.f of $X_{k:n}$ is given by :

$$g_{k,n}(x, q, w, \alpha) = \frac{n!}{(k-1)!(n-k)!} g(x; q, w, \alpha) G(x; q, w, \alpha))^{k-1} [1 - g(x; q, w, \alpha)]^{n-k}$$

From the p.d.f of TRPD we have the p.d.f of the r-th order statistics

$$g_{k,n}(x,q,w,\alpha) = \frac{n!}{(k-1)!(n-k)!}\left[\frac{\frac{\alpha}{2q^2 w}\left(\frac{x}{w}\right)^{\alpha-1} e^{-\frac{1}{2q^2}\left(\frac{x}{w}\right)^{\alpha}}}{1-e^{-\frac{1}{2q^2}\left(\frac{x}{w}\right)^{\alpha}}}\right]\left[\frac{1-e^{-\frac{1}{2q^2}\left(\frac{x}{w}\right)^{\alpha}}}{1-e^{-\frac{1}{2q^2}\left(\frac{x}{w}\right)^{\alpha}}}\right]^{k-1}\left[1-\frac{1-e^{-\frac{1}{2q^2}\left(\frac{x}{w}\right)^{\alpha}}}{1-e^{-\frac{1}{2q^2}\left(\frac{x}{w}\right)^{\alpha}}}\right]^{n-k} =$$

$$\frac{n!}{(k-1)!(n-k)!}\left[\frac{\frac{\alpha}{2q^2 w}\left(\frac{x}{w}\right)^{\alpha-1} e^{-\frac{1}{2q^2}\left(\frac{x}{w}\right)^{\alpha}}}{1-e^{-\frac{1}{2q^2}\left(\frac{x}{w}\right)^{\alpha}}}\right]\left[\frac{1-e^{-\frac{1}{2q^2}\left(\frac{x}{w}\right)^{\alpha}}}{1-e^{-\frac{1}{2q^2}\left(\frac{x}{w}\right)^{\alpha}}}\right]^{k-1}\left[\frac{e^{-\frac{1}{2q^2}\left(\frac{x}{w}\right)^{\alpha}}-e^{-\frac{1}{2q^2}\left(\frac{x}{w}\right)^{\alpha}}}{1-e^{-\frac{1}{2q^2}\left(\frac{x}{w}\right)^{\alpha}}}\right]^{n-k} \qquad \textbf{(32)}$$

Then the p.d.f of the maximum, minimum and the median are explain as follows.

When k=1, we have the p.d.f of minimum

$$g_{1,n}(x,q,w,\alpha) = \frac{n\frac{\alpha}{2q^2 w}\left(\frac{x}{w}\right)^{\alpha-1} e^{-\frac{1}{2q^2}\left(\frac{x}{w}\right)^{\alpha}}}{1-e^{-\frac{1}{2q^2}\left(\frac{x}{w}\right)^{\alpha}}}\left[\frac{e^{-\frac{1}{2q^2}\left(\frac{x}{w}\right)^{\alpha}}-e^{-\frac{1}{2q^2}\left(\frac{x}{w}\right)^{\alpha}}}{1-e^{-\frac{1}{2q^2}\left(\frac{x}{w}\right)^{\alpha}}}\right]^{n-1}$$

(33)

When k = n, we have the p.d.f of maximum

$$g_{n,n}(x,q,w,\alpha) = \frac{n\frac{\alpha}{2q^2 w}\left(\frac{x}{w}\right)^{\alpha-1} e^{-\frac{1}{2q^2}\left(\frac{x}{w}\right)^{\alpha}}}{1-e^{-\frac{1}{2q^2}\left(\frac{x}{w}\right)^{\alpha}}}\left[\frac{1-e^{-\frac{1}{2q^2}\left(\frac{x}{w}\right)^{\alpha}}}{1-e^{-\frac{1}{2q^2}\left(\frac{x}{w}\right)^{\alpha}}}\right]^{n-1} \qquad \textbf{(34)}$$

3 – When k= m+1 , we have the p.d.f of the median

$$g_{m+1,n}(x,q,w,\alpha) = \frac{n!}{m!(n-m-1)!}\left[\frac{\frac{\alpha}{2q^2 w}\left(\frac{x}{w}\right)^{\alpha-1} e^{-\frac{1}{2q^2}\left(\frac{x}{w}\right)^{\alpha}}}{1-e^{-\frac{1}{2q^2}\left(\frac{x}{w}\right)^{\alpha}}}\right]\left[\frac{1-e^{-\frac{1}{2q^2}\left(\frac{x}{w}\right)^{\alpha}}}{1-e^{-\frac{1}{2q^2}\left(\frac{x}{w}\right)^{\alpha}}}\right]^{m}\left[\frac{e^{-\frac{1}{2q^2}\left(\frac{x}{w}\right)^{\alpha}}-e^{-\frac{1}{2q^2}\left(\frac{x}{w}\right)^{\alpha}}}{1-e^{-\frac{1}{2q^2}\left(\frac{x}{w}\right)^{\alpha}}}\right]^{n-m-1}$$

(35)

6. The Coefficients of Variation, Kurtosis and Skewedness.[11][12]

Proposition (6)

The coefficients of variation , skewedness and kurtosis of the TRPD are respectively as .

$$CS = \frac{E(X-\mu)^3}{\sigma^3} \qquad \textbf{(36)}$$

Let $CS = \frac{A}{B}$

By equation (28) we put r=3 then

$$A = E(X-\mu)^3 = \sum_{i=0}^{3}\sum_{j=0}^{\infty} C_i^3\, (-\mu)^i \frac{\alpha(-1)^j\,(x_{\circ})^{(j+1)\alpha+3-i}}{j!\,(2q^2)^{j+1}(w^{\alpha(j+1)})\left[1-e^{-\frac{1}{2q^2}(\frac{x_{\circ}}{w})^{\alpha}}\right](j+1)\alpha+3-i}$$

$$B = \sigma^3 = [E(X-\mu)^2]^{\frac{3}{2}}$$
$$= [\sum_{i=0}^{2}\sum_{j=0}^{\infty} C_i^2\, (-\mu)^i \frac{\alpha(-1)^j\,(x_{\circ})^{(j+1)\alpha+2-i}}{j!\,(2q^2)^{j+1}(w^{\alpha(j+1)})\left[1-e^{-\frac{1}{2q^2}(\frac{x_{\circ}}{w})^{\alpha}}\right](j+1)\alpha+2-i}]^{\frac{3}{2}}$$

$$CK = \frac{E(X-\mu)^4}{\sigma^4} \quad (37)$$

Let $CK = \frac{C}{D}$

Where

$C = E(X-\mu)^4$

By equation (28) we put r=4 then

$$C = \sum_{i=0}^{4}\sum_{j=0}^{\infty} C_i^3\, (-\mu)^i \frac{\alpha(-1)^j\,(x_{\circ})^{(j+1)\alpha+4-i}}{j!\,(2q^2)^{j+1}(w^{\alpha(j+1)})\left[1-e^{-\frac{1}{2q^2}(\frac{x_{\circ}}{w})^{\alpha}}\right](j+1)\alpha+4-i}$$

$$D = \sigma^4 = [E(X-\mu)^2]^2$$

Then

$$D = [\sum_{i=0}^{2}\sum_{j=0}^{\infty} C_i^2\, (-\mu)^i \frac{\alpha(-1)^j\,(x_{\circ})^{(j+1)\alpha+2-i}}{j!\,(2q^2)^{j+1}(w^{\alpha(j+1)})\left[1-e^{-\frac{1}{2q^2}(\frac{x_{\circ}}{w})^{\alpha}}\right](j+1)\alpha+2-i}]^2$$

$$CV = \frac{\sigma}{\mu} = \frac{E}{F} \quad (38)$$

Let $\sigma = E$

$$E = \sqrt{\sum_{i=0}^{2}\sum_{j=0}^{\infty} C_i^2\, (-\mu)^i \frac{(-1)^j\,(x_{\circ})^{(j+1)\alpha+2-i}}{j!\,(2q^2)^{j+1}(w^{\alpha(j+1)})\left[1-e^{-\frac{1}{2q^2}(\frac{x_{\circ}}{w})^{\alpha}}\right](j+1)\alpha+2-i}}$$

$$F = E(X) = \sum_{j=0}^{\infty} \frac{\alpha(-1)^j\, x_{\circ}^{(j+1)\alpha+1}}{j!\,(2q^2)^{j+1} w^{\alpha(j+1)}[1-e^{-\frac{1}{2q^2}(\frac{x_{\circ}}{w})^{\alpha}}]\,(j+1)\alpha+1},$$

Proposition (7)

The harmonic mean is given by:

$$H = E\left(\frac{1}{X}\right) = \sum_{j=0}^{\infty} \frac{\alpha(-1)^j \, x_{\circ}^{(j+1)\alpha-1}}{j!\,(2q^2)^{j+1} w^{\alpha(j+1)} [1- \, e^{-\frac{1}{2q^2}(\frac{x_{\circ}}{w})^{\alpha}}]\,(j+1)\alpha-1},$$

(39)

Proof:

$$E\left(\frac{1}{X}\right) = \int_0^{x_{\circ}} \frac{1}{X}\, g(x;q,w,\alpha)\partial x$$

$$= \int_0^{x_{\circ}} \frac{1}{X} \frac{\frac{\alpha}{2q^2 w}(\frac{x}{w})^{\alpha-1} e^{-\frac{1}{2q^2}(\frac{x}{w})^{\alpha}}}{1-e^{-\frac{1}{2q^2}(\frac{x_{\circ}}{w})^{\alpha}}} = \frac{\alpha}{2q^2 w^{\alpha}[1-e^{-\frac{1}{2q^2}(\frac{x_{\circ}}{w})^{\alpha}}]} \int_0^{x_{\circ}} X^{\alpha-2}\, e^{-\frac{1}{2q^2}(\frac{x}{w})^{\alpha}} \partial x$$

Thus

$$E\left(\frac{1}{X}\right) = \frac{\alpha}{2q^2 w^{\alpha}[1- \, e^{-\frac{1}{2q^2}(\frac{x_{\circ}}{w})^{\alpha}}]} \sum_{j=0}^{\infty} \frac{(-\frac{1}{2q^2})^j \, (\frac{x}{w})^{j\alpha}}{j!} \int_0^{x_{\circ}} x^{(j+1)\alpha-2} \partial x$$

We can find $\int_0^{x_{\circ}} x^{(j+1)\alpha-2} \partial x$ as follows

$$\int_0^{x_{\circ}} x^{(j+1)\alpha-2} \partial x = \frac{x_{\circ}^{\alpha(j+1)-1}}{\alpha(j+1)-1}$$

Thus

$$E\left(\frac{1}{X}\right) = \sum_{j=0}^{\infty} \frac{\alpha(-1)^j (\,x_{\circ})^{(j+1)\alpha-1}}{j!\,(\,2q^2)^{j+1}\, w^{(j+1)\alpha} \left[1-e^{-\frac{1}{2q^2}(\frac{x_{\circ}}{w})^{\alpha}}\right] (j+1)\alpha-1}$$

(40)

Proposition (8)

The geometric mean is by the form :

$$G = E(\sqrt{X}) = \sum_{j=0}^{\infty} \frac{\alpha(-1)^j (\,x_{\circ})^{(j+1)\alpha+\frac{1}{2}}}{j!\,(\,2q^2)^{j+1}\, w^{(j+1)\alpha} \left[1-e^{-\frac{1}{2q^2}(\frac{x_{\circ}}{w})^{\alpha}}\right] (j+1)\alpha+\frac{1}{2}}$$

(41)

Proof:

$$G = E(\,\sqrt{X}) = \int_0^{x_{\circ}} \sqrt{X}\, g(x;q,w,\alpha)\partial x$$

$$= \int_0^{x_\circ} \sqrt{X}\, \frac{\frac{\alpha}{2q^2 w}(\frac{x}{w})^{\alpha-1} e^{-\frac{1}{2q^2}(\frac{x}{w})^{\alpha}}}{1-e^{-\frac{1}{2q^2}(\frac{x_\circ}{w})^{\alpha}}} = \frac{\alpha}{2q^2 w^{\alpha}[1-e^{-\frac{1}{2q^2}(\frac{x_\circ}{w})^{\alpha}}]} \int_0^{x_\circ} X^{\alpha-\frac{1}{2}} e^{-\frac{1}{2q^2}(\frac{x}{w})^{\alpha}} \partial x$$

Thus

$$E(\sqrt{X}) = \frac{\alpha}{2q^2 w^{\alpha}[1-\ e^{-\frac{1}{2q^2}(\frac{x_\circ}{w})^{\alpha}}]} \sum_{j=0}^{\infty} \frac{(-\frac{1}{2q^2})^j\,(\frac{x}{w})^{j\alpha}}{j!} \int_0^{x_\circ} x^{(j+1)\alpha-\frac{1}{2}} \partial x$$

Then

$$G = E(\sqrt{X}) = \sum_{j=0}^{\infty} \frac{\alpha(-1)^j (x_\circ)^{(j+1)\alpha+\frac{1}{2}}}{j!\,(2q^2)^{j+1}\, w^{(j+1)\alpha}\left[1-e^{-\frac{1}{2q^2}(\frac{x_\circ}{w})^{\alpha}}\right](j+1)\alpha+\frac{1}{2}} \quad \textbf{(42)}$$

7. Estimation Methods

We shall discuss some methods to estimate the unknown parameters of TRPD.

7.1. Maximum Likelihood Estimation.

Estimating the parameters for TRPD using the maximum likelihood estimation method for p.d.f are:

$$L = \prod_{i=1}^{n} g(X_i, \theta) = \frac{\frac{\alpha^n}{(2q^2)^n w^n}\,(\frac{\prod_{i=1}^{n} x_i}{w^n})^{\alpha-1} e^{-\frac{1}{2q^2}\sum_{i=1}^{n}\frac{(x_i)^{\alpha}}{w^{\alpha}}}}{1-\ e^{-\frac{1}{2q^2}(\frac{x_\circ}{w})^{\alpha}}}$$

We take the log of both sides to the likelihood function.

$$\ell\ = Ln \frac{\frac{\alpha^n}{(2q^2)^n w^n}\,(\frac{\prod_{i=1}^{n} x_i}{w^n})^{\alpha-1} e^{-\frac{1}{2q^2}\sum_{i=1}^{n}\frac{(x_i)^{\alpha}}{w^{\alpha}}}}{1-\ e^{-\frac{1}{2q^2}(\frac{x_\circ}{w})^{\alpha}}}$$

$$= n\ln\alpha - n\ln(2q^2) - n\ln w + (\alpha-1)\sum_{i=1}^{n} \ln x_i - n(\alpha-1)\ln w - \frac{1}{2q^2}\sum_{i=1}^{n}\frac{x_i^{\alpha}}{w^{\alpha}} -$$

$$\ln\left(1-e^{-\frac{1}{2q^2}(\frac{x_\circ}{w})^{\alpha}}\right) \quad \textbf{(43)}$$

$$\frac{\partial Ln\ell}{\partial \alpha}=\frac{n}{\alpha}+\sum_{i=1}^{n} lnx_i-nlnw-\frac{1}{2q^2}\sum_{i=1}^{n}(\frac{x_i}{w})^{\alpha}ln(\frac{x_i}{w})-\frac{\frac{1}{2q^2}(\frac{x_\circ}{w})^{\alpha}ln(\frac{x_\circ}{w})e^{-\frac{1}{2q^2}(\frac{x_\circ}{w})^{\alpha}}}{1-e^{-\frac{1}{2q^2}(\frac{x_\circ}{w})^{\alpha}}} \quad (44)$$

$$\frac{\partial Ln\ell}{\partial q}=\frac{-2n}{q}+\frac{1}{q^3}\sum_{i=1}^{n}\frac{x_i^{\alpha}}{w^{\alpha}}-\frac{\frac{1}{q^3}(\frac{x_\circ}{w})^{\alpha}e^{-\frac{1}{2q^2}(\frac{x_\circ}{w})^{\alpha}}}{1-e^{-\frac{1}{2q^2}(\frac{x_\circ}{w})^{\alpha}}} \quad (45)$$

$$\frac{\partial Ln\ell}{\partial w}=\frac{-n}{w}-\frac{n(\alpha-1)}{w}+\frac{\alpha}{2q^2}\frac{\sum_{i=1}^{n}x_i^{\alpha}}{w^{\alpha+1}}-\frac{\frac{\alpha}{2q^2}\frac{(x_\circ)^{\alpha}}{w^{\alpha+1}}e^{-\frac{1}{2q^2}(\frac{x_\circ}{w})^{\alpha}}}{1-e^{-\frac{1}{2q^2}(\frac{x_\circ}{w})^{\alpha}}}$$

(46)

Now equaling each these equations to zero.

$$\frac{n}{\hat{\alpha}}+\sum_{i=1}^{n} lnx_i-nLn\hat{w}-\frac{1}{2\hat{q}^2}\sum_{i=1}^{n}(\frac{x_i}{\hat{w}})^{\alpha}ln(\frac{x_i}{\hat{w}})-\frac{\frac{1}{2\hat{q}^2}(\frac{x_\circ}{\hat{w}})^{\hat{\alpha}}ln(\frac{x_\circ}{\hat{w}})e^{-\frac{1}{2\hat{q}^2}(\frac{x_\circ}{\hat{w}})^{\hat{\alpha}}}}{1-e^{-\frac{1}{2\hat{q}^2}(\frac{x_\circ}{\hat{w}})^{\hat{\alpha}}}}=0$$

(47)

$$\frac{-2n}{\hat{q}}+\frac{1}{\hat{q}^3}\sum_{i=1}^{n}\frac{x_i^{\hat{\alpha}}}{\hat{w}^{\hat{\alpha}}}-\frac{\frac{1}{\hat{q}^3}(\frac{x_\circ}{\hat{w}})^{\hat{\alpha}}e^{-\frac{1}{2\hat{q}^2}(\frac{x_\circ}{\hat{w}})^{\hat{\alpha}}}}{1-e^{-\frac{1}{2\hat{q}^2}(\frac{x_\circ}{\hat{w}})^{\hat{\alpha}}}}=0 \quad (48)$$

$$\frac{-n}{\hat{w}}-\frac{n(\hat{\alpha}-1)}{\hat{w}}+\frac{\hat{\alpha}}{2\hat{q}^2}\frac{\sum_{i=1}^{n}x_i^{\hat{\alpha}}}{\hat{w}^{\hat{\alpha}+1}}-\frac{\frac{\hat{\alpha}}{2\hat{q}^2}\frac{(x_\circ)^{\hat{\alpha}}}{\hat{w}^{\hat{\alpha}+1}}e^{-\frac{1}{2\hat{q}^2}(\frac{x_\circ}{\hat{w}})^{\hat{\alpha}}}}{1-e^{-\frac{1}{2\hat{q}^2}(\frac{x_\circ}{\hat{w}})^{\hat{\alpha}}}}=0 \quad (49)$$

The solution of these equation are that the MLEś that can be found of the parameters p,b,α by numerical method.

7.2. The Least Square Method(LS):**[13]**

The method defined is. Let $Y_1,Y_2,\ldots,Y_n$ is a random sample of size n from a distribution function G(.) and $Y_{(1)},Y_{(2)},\ldots,Y_{(n)}$ indicates the order statistics of the observed sample. It is well know that

$$E\left(G\left(Y_{(i)}\right)\right)=\frac{i}{n+1}$$

Obtain the estimators by minimizing

$$\sum_{i=1}^{n}(G(Y_{(i)})-\frac{i}{n+1})^2$$

So $\sum_{i=1}^{n}[\left(\frac{1-e^{-\frac{1}{2q^2}(\frac{x}{w})^\alpha}}{1-e^{-\frac{1}{2q^2}(\frac{x_i}{w})^\alpha}}\right)-\frac{i}{n+1}]^2$ **(50)**

The $\hat{q}_{LSE}$ can be found by derive the equation (50) with respect to q, it is given.

$$\sum_{i=1}^{n}2[(\frac{1-e^{-\frac{1}{2\hat{q}^2}(\frac{x}{w})^\alpha}}{1-e^{-\frac{1}{2\hat{q}^2}(\frac{x_i}{w})^\alpha}})-\frac{i}{n+1}]$$
$$\times[\frac{(1-e^{-\frac{1}{2\hat{q}^2}(\frac{x_i}{w})^\alpha})(-\frac{1}{\hat{q}^3}(\frac{x}{w})^\alpha e^{-\frac{1}{2\hat{q}^2}(\frac{x}{w})^\alpha})-(1-e^{-\frac{1}{2\hat{q}^2}(\frac{x}{w})^\alpha})(-\frac{1}{\hat{q}^3}(\frac{x_i}{w})^\alpha e^{-\frac{1}{2\hat{q}^2}(\frac{x_i}{w})^\alpha})}{(1-e^{-\frac{1}{2q^2}(\frac{x_i}{w})^\alpha})^2}]=0$$

Then

$$\sum_{i=1}^{n}[(\frac{1-e^{-\frac{1}{2\hat{q}^2}(\frac{x}{w})^\alpha}}{1-e^{-\frac{1}{2\hat{q}^2}(\frac{x_i}{w})^\alpha}})-\frac{i}{n+1}]$$
$$\times[\frac{(1-e^{-\frac{1}{2\hat{q}^2}(\frac{x_i}{w})^\alpha})(-\frac{1}{\hat{q}^3}(\frac{x}{w})^\alpha e^{-\frac{1}{2\hat{q}^2}(\frac{x}{w})^\alpha})-(1-e^{-\frac{1}{2\hat{q}^2}(\frac{x}{w})^\alpha})(-\frac{1}{\hat{q}^3}(\frac{x_i}{w})^\alpha e^{-\frac{1}{2\hat{q}^2}(\frac{x_i}{w})^\alpha})}{(1-e^{-\frac{1}{2q^2}(\frac{x_i}{w})^\alpha})^2}]=0 \quad (51)$$

The $\hat{w}_{LES}$ can be found by derive the equation (50) with respect to w, it is given.

$$\sum_{i=1}^{n}2[(\frac{1-e^{-\frac{1}{2q^2}(\frac{x}{\hat{w}})^\alpha}}{1-e^{-\frac{1}{2q^2}(\frac{x_i}{\hat{w}})^\alpha}})\frac{i}{n+1}]\times[\frac{(1-e^{-\frac{1}{2q^2}(\frac{x_i}{\hat{w}})^\alpha})(\frac{\alpha}{2q^2}\frac{x^\alpha}{(\hat{w})^{\alpha+1}}e^{-\frac{1}{2q^2}(\frac{x}{\hat{w}})^\alpha})-(1-e^{-\frac{1}{2q^2}(\frac{x}{\hat{w}})^\alpha})(\frac{1}{2q^2}\frac{x_i^\alpha}{(\hat{w})^{\alpha+1}}e^{-\frac{1}{2q^2}(\frac{x_i}{\hat{w}})^\alpha})}{(1-e^{-\frac{1}{2q^2}(\frac{x_i}{\hat{w}})^\alpha})^2}]=0$$

Then

$$\sum_{i=1}^{n}[(\frac{1-e^{-\frac{1}{2q^2}(\frac{x}{\hat{w}})^\alpha}}{1-e^{-\frac{1}{2q^2}(\frac{x_i}{\hat{w}})^\alpha}})\frac{i}{n+1}]\times[\frac{\left(1-e^{-\frac{1}{2q^2}(\frac{x_i}{\hat{w}})^\alpha}\right)\left(\frac{\alpha}{2q^2}\frac{x^\alpha}{(\hat{w})^{\alpha+1}}e^{-\frac{1}{2q^2}(\frac{x}{\hat{w}})^\alpha}\right)-\left(1-e^{-\frac{1}{2q^2}(\frac{x}{\hat{w}})^\alpha}\right)\left(\frac{1}{2q^2}\frac{x_i^\alpha}{(\hat{w})^{\alpha+1}}e^{-\frac{1}{2q^2}(\frac{x_i}{\hat{w}})^\alpha}\right)}{(1-e^{-\frac{1}{2q^2}(\frac{x_i}{\hat{w}})^\alpha})^2}]=0 \quad (52)$$

The $\hat{\alpha}_{LES}$ can be found by derive the equation (50) with respect to α, it is given.

$$\sum_{i=1}^{n}2[(\frac{1-e^{-\frac{1}{2q^2}(\frac{x}{w})^{\hat{\alpha}}}}{1-e^{-\frac{1}{2q^2}(\frac{x_i}{w})^{\hat{\alpha}}}})\frac{i}{n+1}]\times[\frac{(1-e^{-\frac{1}{2q^2}(\frac{x_i}{w})^{\hat{\alpha}}})(\frac{1}{2q^2}(\frac{x}{w})^{\hat{\alpha}}ln(\frac{x}{w})e^{-\frac{1}{2q^2}(\frac{x}{w})^{\hat{\alpha}}})-(1-e^{-\frac{1}{2q^2}(\frac{x}{w})^{\hat{\alpha}}})(\frac{1}{2q^2}(\frac{x_i}{w})^{\hat{\alpha}}ln(\frac{x_i}{w})e^{-\frac{1}{2q^2}(\frac{x_i}{w})^{\hat{\alpha}}})}{(1-e^{-\frac{1}{2q^2}(\frac{x_i}{w})^{\hat{\alpha}}})^2}]=0$$

Then

$$\sum_{i=1}^{n}\left[\left(\frac{1-e^{-\frac{1}{2q^2}(\frac{x}{w})^{\hat{\alpha}}}}{1-e^{-\frac{1}{2q^2}(\frac{x_s}{w})^{\hat{\alpha}}}}\right)\frac{i}{n+1}\right]\times\left[\frac{(1-e^{-\frac{1}{2q^2}(\frac{x_s}{w})^{\hat{\alpha}}})(\frac{1}{2q^2}(\frac{x}{w})^{\hat{\alpha}}ln(\frac{x}{w})e^{-\frac{1}{2q^2}(\frac{x}{w})^{\hat{\alpha}}})-(1-e^{-\frac{1}{2q^2}(\frac{x}{w})^{\hat{\alpha}}})(\frac{1}{2q^2}(\frac{x_s}{w})^{\hat{\alpha}}ln(\frac{x_s}{w})e^{-\frac{1}{2q^2}(\frac{x_s}{w})^{\hat{\alpha}}})}{(1-e^{-\frac{1}{2q^2}(\frac{x_s}{w})^{\hat{\alpha}}})^2}\right]=0 \quad (53)$$

The solution to these equations from (51) – (53) is by numerical method.

7.3.Percentile Estimation Method (PEM):

This method was used effectively for generalized exponential distribution and Weibull distribution.

Where it was first discovered by Kao (1958,1959).[13]

Since $G(x;q,w,\alpha)=\frac{1-e^{-\frac{1}{2q^2}(\frac{x}{w})^{\alpha}}}{1-e^{-\frac{1}{2q^2}(\frac{x_s}{w})^{\alpha}}}$, therfor

$$ln[G(x;q,w,\alpha)]=ln\left[\frac{1-e^{-\frac{1}{2q^2}(\frac{x}{w})^{\alpha}}}{1-e^{-\frac{1}{2q^2}(\frac{x_s}{w})^{\alpha}}}\right] \quad \textbf{(54)}$$

Suppose $X_{(j)}$ is the j-th order statistic , i.e $X_{(1)}<X_{(2)}<\cdots<X_{(n)}$. If P_i indicates some estimate of $G(x_{(j)};q,w,\alpha)$, then the estimate of p, b and α can be obtained by minimizing .

$$\sum_{j=0}^{n}\left[ln(P_j)-ln\left(\frac{1-e^{-\frac{1}{2q^2}(\frac{x}{w})^{\alpha}}}{1-e^{-\frac{1}{2q^2}(\frac{x_s}{w})^{\alpha}}}\right)\right]^2$$

We have

$$\sum_{j=0}^{n}\left[ln(P_j)-ln\left(1-e^{-\frac{1}{2q^2}(\frac{x}{w})^{\alpha}}\right)+\left(1-e^{-\frac{1}{2q^2}(\frac{x_s}{w})^{\alpha}}\right)\right]^2 \quad \textbf{(55)}$$

We mainly consider $P_j=\frac{j}{n+1}$

The $\hat{q}_{PCE}$ can be found by derive the equation (55) with respect to q, it is given.

$$\sum_{j=1}^{n}2\left[\ln(P_j)-ln\left(1-e^{-\frac{1}{2\hat{q}^2}(\frac{x}{w})^{\alpha}}\right)+ln\left(1-e^{-\frac{1}{2\hat{q}^2}(\frac{x_s}{w})^{\alpha}}\right)\right]\left[\left(-\frac{\frac{1}{\hat{q}^3}(\frac{x}{w})^{\alpha}e^{-\frac{1}{2\hat{q}^2}(\frac{x}{w})^{\alpha}}}{1-e^{-\frac{1}{2q^2}(\frac{x}{w})^{\alpha}}}\right)+\left(\frac{\frac{1}{\hat{q}^3}(\frac{x_s}{w})^{\alpha}e^{-\frac{1}{2\hat{q}^2}(\frac{x_s}{w})^{\alpha}}}{1-e^{-\frac{1}{2\hat{q}^2}(\frac{x_s}{w})^{\alpha}}}\right)\right]=0$$

Then

$$\sum_{j=1}^{n}[\ln (P_j) - ln\left(1 - e^{-\frac{1}{2\hat{q}^2}(\frac{x}{w})^{\alpha}}\right) + ln\left(1 - e^{-\frac{1}{2\hat{q}^2}(\frac{x_0}{w})^{\alpha}}\right)][\left(-\frac{\frac{1}{\hat{q}^3}(\frac{x}{w})^{\alpha} e^{-\frac{1}{2\hat{q}^2}(\frac{x}{w})^{\alpha}}}{1 - e^{-\frac{1}{2q^2}(\frac{x}{w})^{\alpha}}}\right) + \left(\frac{\frac{1}{\hat{q}^3}(\frac{x_0}{w})^{\alpha} e^{-\frac{1}{2\hat{q}^2}(\frac{x_0}{w})^{\alpha}}}{1 - e^{-\frac{1}{2\hat{q}^2}(\frac{x_0}{w})^{\alpha}}}\right)] = 0$$

(56)

The $\hat{w}_{PCE}$ can be found by derive the equation (55) with respect to w, it is given

$$\sum_{j=0}^{n} 2[ln(P_j) - ln\left(1 - e^{-\frac{1}{2q^2}(\frac{x}{\hat{w}})^{\alpha}}\right) + ln\left(1 - e^{-\frac{1}{2q^2}(\frac{x_0}{\hat{w}})^{\alpha}}\right)][\left(-\frac{\frac{\alpha}{2q^2}\left(\frac{x^{\alpha}}{\hat{w}^{\alpha+1}}\right) e^{-\frac{1}{2q^2}(\frac{x}{\hat{w}})^{\alpha}}}{1 - e^{-\frac{1}{2q^2}(\frac{x}{\hat{w}})^{\alpha}}}\right) + \left(\frac{\frac{\alpha}{2q^2}\left(\frac{x_0{}^{\alpha}}{\hat{w}^{\alpha+1}}\right) e^{-\frac{1}{2q^2}(\frac{x_0}{\hat{w}})^{\alpha}}}{1 - e^{-\frac{1}{2q^2}(\frac{x_0}{\hat{w}})^{\alpha}}}\right)] = 0$$

Then

$$\sum_{j=0}^{n}[ln(P_j) - ln\left(1 - e^{-\frac{1}{2q^2}(\frac{x}{\hat{w}})^{\alpha}}\right) + ln\left(1 - e^{-\frac{1}{2q^2}(\frac{x_0}{\hat{w}})^{\alpha}}\right)][\left(-\frac{\frac{\alpha}{2q^2}\left(\frac{x^{\alpha}}{\hat{w}^{\alpha+1}}\right) e^{-\frac{1}{2q^2}(\frac{x}{\hat{w}})^{\alpha}}}{1 - e^{-\frac{1}{2q^2}(\frac{x}{\hat{w}})^{\alpha}}}\right) + \left(\frac{\frac{\alpha}{2q^2}\left(\frac{x_0{}^{\alpha}}{\hat{w}^{\alpha+1}}\right) e^{-\frac{1}{2q^2}(\frac{x_0}{\hat{w}})^{\alpha}}}{1 - e^{-\frac{1}{2q^2}(\frac{x_0}{\hat{w}})^{\alpha}}}\right)] = 0 \quad (57)$$

The $\hat{\alpha}_{PCE}$ can be found by derive the equation (55) with respect to b, it is given.

$$\sum_{j=0}^{n} 2[ln(P_j) - ln\left(1 - e^{-\frac{1}{2q^2}(\frac{x}{w})^{\hat{\alpha}}}\right) + ln\left(1 - e^{-\frac{1}{2q^2}(\frac{x_0}{w})^{\hat{\alpha}}}\right)][\left(-\frac{\frac{1}{2q^2}\left(\frac{x}{w}\right)^{\hat{\alpha}} ln(\frac{x}{w}) e^{-\frac{1}{2q^2}(\frac{x}{w})^{\hat{\alpha}}}}{1 - e^{-\frac{1}{2q^2}(\frac{x}{w})^{\hat{\alpha}}}}\right) + \left(\frac{-\frac{1}{2q^2}\left(\frac{x_0}{w}\right)^{\hat{\alpha}} ln(\frac{x}{w}) e^{-\frac{1}{2q^2}(\frac{x_0}{w})^{\hat{\alpha}}}}{1 - e^{-\frac{1}{2q^2}(\frac{x_0}{w})^{\hat{\alpha}}}}\right)] = 0$$

Then

$$\sum_{j=0}^{n}[ln(P_j) - ln\left(1 - e^{-\frac{1}{2q^2}(\frac{x}{w})^{\hat{\alpha}}}\right) + ln\left(1 - e^{-\frac{1}{2q^2}(\frac{x}{w})^{\hat{\alpha}}}\right)][\left(-\frac{\frac{1}{2q^2}(\frac{x}{w})^{\hat{\alpha}} ln(\frac{x}{w}) e^{-\frac{1}{2q^2}(\frac{x}{w})^{\hat{\alpha}}}}{1 - e^{-\frac{1}{2q^2}(\frac{x}{w})^{\hat{\alpha}}}}\right) + \left(\frac{-\frac{1}{2q^2}(\frac{x_\circ}{w})^{\hat{\alpha}} ln(\frac{x}{w}) e^{-\frac{1}{2q^2}(\frac{x_\circ}{w})^{\hat{\alpha}}}}{1 - e^{-\frac{1}{2q^2}(\frac{x_\circ}{w})^{\hat{\alpha}}}}\right)] = 0$$

(58)

The solution to these equations from (56) – (58) is by numerical method.

***7.4. Moments Method (W):*[14]**

Let $X_1, X_2, \dots \dots, X_3$ be a random sample from $X \sim TRPD(x; q, w, \alpha)$ This method can be found through .

$$E(X^r) = \frac{1}{n}\sum_{i=1}^{n} X_i^r \quad \textbf{(59)}$$

Where $E(X^r)$ is the r-th moment about origin .

If r = 1 then the equation (59) becomes as follows

$$E(X) = \sum_{j=0}^{\infty} \frac{\alpha(-1)^j x_\circ^{(j+1)\alpha+1}}{j!\,(2q^2)^{j+1} w^{\alpha(j+1)} [1 - e^{-\frac{1}{2q^2}(\frac{x_\circ}{w})^{\alpha}}]\,(j+1)\alpha+1} = \sum_{i=1}^{n} \frac{1}{n} X_i = \bar{X}$$

(60)

If r =2 then the equation (59) becomes as follows

$$E(X^2) = \sum_{j=0}^{\infty} \frac{\alpha(-1)^j x_\circ^{(j+1)\alpha+2}}{j!\,(2q^2)^{j+1} w^{\alpha(j+1)} [1 - e^{-\frac{1}{2q^2}(\frac{x_\circ}{w})^{\alpha}}]\,(j+1)\alpha+2} = \sum_{i=1}^{n} \frac{1}{n} X_i^{\,2}$$

(61)

If r =3 then the equation (59) becomes as follows

$$E(X^3) = \sum_{j=0}^{\infty} \frac{\alpha(-1)^j x_\circ^{(j+1)\alpha+3}}{j!(2q^2)^{j+1} w^{\alpha(j+1)} [1 - e^{-\frac{1}{2q^2}(\frac{x_\circ}{w})^{\alpha}}]\,(j+1)\alpha+3} = \sum_{i=1}^{n} \frac{1}{n} X_i^{\,3} \quad \textbf{(62)}$$

To find the estimates of the parameters p, b, α we solve these equation from(60) – (62) by a numerical path such as Newton Raphson Method.

Application

These are dataset agree to remissful (in months) of a random sample of (128) for sick bladder cancer [9].

[0.08,2.09,3.48,4.87,6.94,8.66,13.11,23.63,0.20,2.23,3.52,4.98,6.97,9.02,13.29,0.40,2.26,3.57,5.06,7.09,9.22,13.80,25.74,0.50,2.46,3.64,5.09,7.26,9.47,14.24,25.82,0.51,2.54,3.70,5.17,7.28,9.74,14.76,26.31,0.81,2.62,3.82,5.32,7.32,10.06,14.77,32.15,2.64,3.88,5.32,7.39,10.34,14.83,34.26,0.90,2.69,4.18,5.34,7.59,10.66,15.96,36.66,1.05,2.2.69,4.23,5.41,7.62,10.75,16.62,43.01,1.19,2.75,4.26,5.41,7.63,17.12,46.12,1.26,2.83,4.33,5.49,7.66,11.25,17.14,79.05,1.35,2.87,5.62,7.87,11.64,17.36,1.40,3.02,4.34,5.71,7.93,11.79,18.10,1.46,4.40,5.85,8.26,11.98,19.13,1.76,3.25,4.50,6.25,8.37,12.02,2.02,3.31,4.51,6.54,8.53,12.03,20.28,2.02,3.36,6.76,12.07,21.73,2.07,3.36,6.93,8.65,12.63,22.69]

We have installed the truncated Rayleigh Pareto distribution on the dataset using (MLE), and compare the proposal with the Rayleigh Lomax distribution(ROD) and Lomax distribution. The model selection is carried out using the AIC, the BIC, The CAIC, and the HQIC .

$$AIC = -2\hat{\ell} + 2q \,, \tag{63}$$

$$BIC = -2\hat{\ell} + q\log(n), \tag{64}$$

$$HQIC = -2\hat{\ell} + 2q\log(\log(n)), \tag{65}$$

$$CAIC = -2\hat{\ell} + \frac{2qn}{n-q-1}, \tag{67}$$

Whereas $(\hat{\ell})$ represent the log-likelihood function evaluate at the maximum likelihood estimates, (n) the sample size, and (q)are the number of parameters,

The MLEs of the pattern parameters for the data are given in Table (1), and the numerical values of the model selection statistics $\hat{\ell}$, AIC,BIC,CAIC and HQIC are listed in Table (2).

From Table (2) we see that the (TRPD) model gives the smallest values for the criteria AIC , BIC , CAIC and HQIC so it represents the data set better than the other chosen models.

Table 1. Parameters Estimates for the Data

Model	Parameters Estimates		
TRPD(x;α,p,b)	$\hat{\alpha}$=0.823	$\hat{q}$=2.658	$\hat{w}$=1.802
R-LD(x;β,θ,b)	$\hat{\beta}$=1.107	$\hat{\theta}$=4.454	$\hat{b}$=5.798
LD(x;α,λ)	$\hat{\alpha}$=2.889	$\hat{\lambda}$=19.614	
R-PD(x;p,b,α)	$\hat{\alpha}$=1.048	$\hat{b}$=1.262	$\hat{p}$=3.164

Table 2. the values of Statistics $\hat{\ell}$, AIC, BIC, CAIC and HQIC for the Data set.

Model	$\hat{\ell}$	AIC	BIC	CAIC	HQIC
TRPD(x;α,q,w)	-391.8296	789.6593	803.0674	789.8528	793.1357
R-LD(x;β,θ,b)	-474.3170	954.6339	968.0421	954.8275	958.1103
LD(x;α,λ)	-420.8238	847.6477	861.0558	847.8412	851.1240

R-PD(x;p,b,α)	-414.0869	834.1738	847.5819	834.3673	837.6502

References

[1]. Wen-liany Hung and ching- yichen. (2004***). "Approximmate MLE of the scale parameter of the truncated Rayleigh distribution under the first – censored data"*** Department of Statistics Tamkang University Tamsui, Taipei Taiwan, 25137 R.O.C. Department of Mathematics Education National Hsinchu Teachers College Hsin-Chu Taiwan R.O.C.

[2]. David R-clark, FCAS. (2013). ***"A note on the upper- truncated Pareto distribution"***, Presented at the. 2013 Enterprise Risk Management SymposiumApril 22-24, 2013.

[3]. Taylor and francis. (2014). ***"Parameter Estimation for the Truncated Pareto Distribution"***, Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK

[4]. Mathias Reachke. (2012). ***"Inference for the truncated exponential distribution"***, Published online: 5 February 2011.

[5].AL_ Yasseri , Ahmad Yassin Taqi. (2014). "Estimation of parameter and Reliability function for truncated Logistic distribution" , M.CS. Thesis. AL-Mustansiriya University.

[6].Dibagay, Dler Mustafa. (2007). ***"Numerical Estimation of the Parameters of the Gamma and Truncated Exponential Distribution"***, M.Sc. Thesis, University of Mosul.

[7].Hormuz. Amir Hanna. (1990). ***"Mathematical Statistics",***Directorate of Dar Al-Kutub for Printing and Publishing, University of Mosul.

[8] Blumenthal & Goel. (1988). ***"Estimation With Truncated Data"*** OHIO State Univresity For columus VOL.4

[9]Al_ Sultani ,Bnadheer Dhea'a Mohmmad. (2017). ***" New Distribution on Proposed Classes"*** M.cs. Thesis. University of Babylon.

[10].Hassan, Habe Ali. (2015). ***"Issues related with uniform distribution"***, M.CS. Thesis. AL-Mustansiriya University.

[11]AL_Ghalib ,Sabah Muhammad Khudair. (2019). ***"On some Special oge distributions"*** , M.CS. Thesis. University of Pabylon.

[12].AL- Sultani , B.D. (2017). ***"New Distributions on proposed classes",*** M.Sc.Thesis, University of Babylon.

[13].Hanaa H. Abu – Zinadah. (2014). ***" Six Method of Estimations for the Shape Parameter of Exponentiated Gompertz Distribution"*** Department of Statistics, sciences Faculty for Girls King Abaulaziz University, P.O. BOX 3269 Jededah 21436, Saudi Arabia.

Generalized Odd Generalized Exponential

Weibull Distribution, ***Sabah Muhammad Khudair Hussain Al-Ghalib**

sabahalghalb@gmail.com

**** Kareema Abed AL_Kadim**

Dept. of Mathematics,College of Education for Pure Sciences,

University of Babylon,Hilla, Iraq

kareema.kadim@yahoo.com

Abstract: We introduce five parameters continuous distribution called generalized odd generalized exponential Weibull (GOGE-W-E) distribution for modeling life time data. We introduce an explicit expressions for the studying. We present and study some of its properties, the cdf , pdf, functions of reliability like moment, quantile and median, the moment generating function, Rényi entropy and order statistics. The five parameters of the suggested are estimated by the maximum likelihood estimation method . We illustrate its usefulness by means of an application to a real data sets

1 Introduction

In a statistical analysis, lifetime distributions such as the exponential distribution, Weibull distribution, normal distribution, and gamma distribution play an important role in many fields of the real life such as reliability, survival analysis, ecology, medicine, and social sciences. There are continuous motivations to develop these lifetime distributions to become more flexible or more fitting for specific real data sets. Thus, in recent years many different families of distributions have been developed by generalizing the common families of continuous distributions such as Weibull distribution and exponential distribution by adding one or more than one additional parameter(s) to the baseline model. Among these, the exponentiated Weibull family, Mudholkar, and Srivaastav, (1993), generalized exponential (GE) distribution, (Gupta and kundu, 2007), modified Weibull distribution, Lai *et al.*(2009) , Sarhan and Zaindin (2003)25], Beta-Weibull distribution, Famoye et al.(2005) , A flexible Weibull extension (Bebbington *et al.(*2007), Beta modified Weibull distribution, (Silva, *et al*.2010 , Nadarajah, *et al.* 2011), Beta generalized Weibull distribution (Singla, *et al.* 2012) and a new modified Weibull distribution, [Almalki and Yuan, 2013) among others. An important generalization (Gupta and Kundu, 2007),he has been suggested of the exponential distribution known as generalized exponential (GE). The cumulative distribution function (cdf) of GE is given by

$$F(x;\alpha,\lambda) = \left[1 - e^{-\lambda x}\right]^{\alpha}, x > 0, \alpha > 0, \lambda > 0. \quad (1)$$

Recently, a new family of continuous distributions called the odd generalized exponential (OGE) family has been introduced in (El-Damcese, *et al.* 2015, Tahir, *et al.*2015). This family is flexible because of hazard rate shapes could be decreasing, increasing, bathtub and upside down bathtub. Many special OGE distributions have been introduced such as the odd generalized exponential Weibull (OGE-W) distribution, odd generalized exponential normal (OGE-N) distribution by Tahir, *et al.* (2015) , odd generalized exponential generalized linear exponential (OGE-GLE) distribution by Luguterah and Nasiru (2017) and odd generalized exponential flexible Weibull extension (OGE-FWE) distribution by Mustafa, *et al.* (2018).The pdf and cdf Tahir, *et al.* (2015) of the odd generalized exponential (OGE) family are defined as follows:

If $G(x), x > 0$ is cumulative distribution function (cdf) of a random variable X , then its probability density function (pdf) is $g(x)$ and the survival function is $\bar{G}(x) = 1 - G(x)$, then we define the cdf of the OGE family by replacing x in equation (1) by $\frac{G(x)}{\bar{G}(x)}$ leading to **:**

$$F(x;\alpha,\lambda) = \left[1 - e^{-\lambda\frac{G(X)}{\bar{G}(x)}}\right]^{\alpha}, x > 0, \alpha > 0, \lambda > 0. \quad \textbf{(2)}$$

The pdf corresponding to equation (2) is given by

$$f(x;\alpha,\lambda) = \frac{\alpha\lambda g(x)}{\bar{G}(x)^2}\; e^{-\lambda\frac{G(X)}{\bar{G}(x)}}\left[1 - e^{-\lambda\frac{G(X)}{\bar{G}(x)}}\right]^{\alpha-1} \; x > 0, \alpha > 0, \lambda > 0. \quad \textbf{(3)}$$

2 The cdf and the pdf of GOGE-W Distribution

In this section, five parameters generalized odd generalized exponential Weibull GOGE-W $(\alpha, \lambda, a, b, k)$ distribution are introduced and studied. We define the cdf of this distribution as follows:

Adding the shape parameter (k) to the (2) to get the generalized odd generalized exponential family (GOGE) as follows:

$$F(x; \alpha, \lambda, k) = \left[1 - e^{-\lambda\left[\frac{G(x)}{\bar{G}(x)}\right]^k}\right]^\alpha \qquad (4)$$

Taking $G(x) and \bar{G}(x)$ equal to the cdf and survival function of Weibull distribution, respectively which are

$$G(x; a, b) = 1 - e^{-ax^b}\ , \quad x > 0, a > 0, b > 0, \qquad (5)$$

$$\bar{G}(x; a, b) = e^{-ax^b},\ \ x > 0, a, b > 0 \qquad (6)$$

Then the cdf and pdf of the GOGE-W are defined by as follows respectively:

$$F(x; \alpha, \lambda, a, b, k) = \left[1 - e^{-\lambda(e^{ax^b}-1)^k}\right]^\alpha, x > 0, \alpha, \lambda, a, b, k > 0\ , \qquad (7)$$

$$f(x; \alpha, \lambda, a, b, k) = \ \alpha\lambda abk\, x^{b-1}\, e^{ax^b} e^{-\lambda(e^{ax^b}-1)^k}\ (e^{ax^b} - 1)^{k-1}\left[1 - e^{-\lambda(e^{ax^b}-1)^k}\right]^{\alpha-1}$$
(8)

Where α , b , $k > 0$ are shape parameters and a , $\lambda > 0$ are scale parameters.

2.1 The Limit of the pdf and cdf of GOGE-W Distribution

The limit of the **pdf** is given as follows:

$$\lim_{x\to 0} f(x; \alpha, \lambda, a, b, k) = = 0 \qquad (9)$$

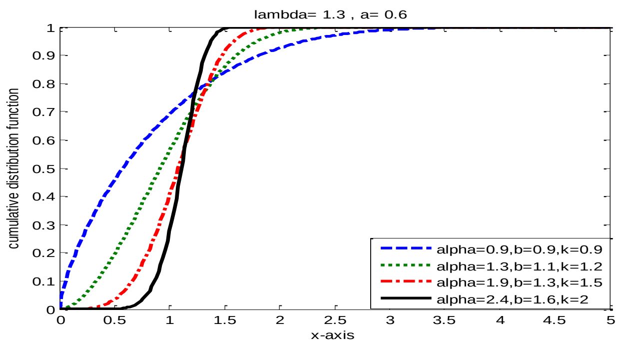
$$lim_{x\to\infty} f(x; \alpha, \lambda, a, b, k) = NAN \qquad (10)$$

The limit of the cdf is given as follows:

$$\lim_{x\to 0} F(x; \alpha, \lambda, a, b, k) = \lim_{x\to 0}\left[1 - e^{-\lambda(e^{ax^b}-1)^k}\right]^\alpha = 0, \qquad (11)$$

$$\lim_{x\to\infty} F(x; \alpha, \lambda, a, b, k) = \lim_{x\to\infty}\left[1 - e^{-\lambda(e^{ax^b}-1)^k}\right]^\alpha = 1 \qquad (12)$$

The plots of cdf and pdf of the GOGE-W for different parameters are given by the following figures:



Figure(1):The cdf of GOGE-W distribution with the parameters
$a = 0.6$; $\lambda = 1.3$; $\alpha = (0.9, 1.3, 1.9, 2.4)$; $b = (0.9, 1.1, 1.3, 1.6)$ $and\ k = (0.9, 1.2, 1.5, 2)$

We can see from the Figure **(1),** that the cdf of the GOGE-W distribution is non decreasing with increasing x *and the parameters* α, λ, a, b, k .

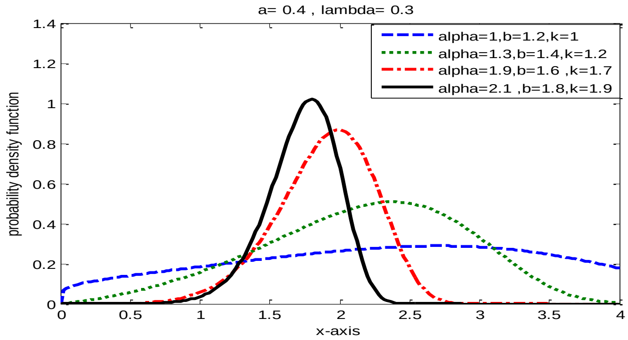


Figure (2) :The pdf of GOGE-W distribution with the parameters $a = 0.4$; $\lambda = 0.3$; $\alpha = (1, 1.3, 1.9, 2.1)$; $b = (1.2, 1.4, 1.6, 1.8)$ $and\ k = (1, 1.2, 1.7, 1.9)$.

2.2 Reliability Analysis

In this section, we introduce the reliability(survival) function $\bar{F}(x)$, hazard rate function $h(x)$, reversed hazard function $r(x)$ and cumulative hazard rate function $H(x)$ of $X\sim$ GOGE-W $(\alpha, \lambda, a, b, k)$.

2.2.1 Reliability Function

The reliability (survivor) function of $X\sim$ GOGE-W $(\alpha, \lambda, a, b, k)$ is defined as follows :

$$\bar{F}(x; \alpha, \lambda, a, b, k) = 1 - F(x; \alpha, \lambda, a, b, k) = 1 - \left[1 - e^{-\lambda (e^{a x^b} - 1)^k}\right]^{\alpha} \qquad (13)$$

The plots of $\bar{\boldsymbol{F}}(\boldsymbol{x}; \boldsymbol{\alpha}, \boldsymbol{\lambda}, \boldsymbol{a}, \boldsymbol{b}, \boldsymbol{k})$ are given by the following figures:

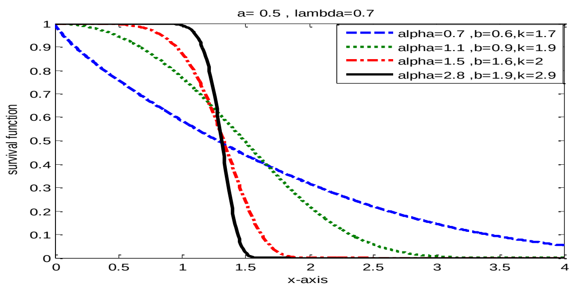


Figure (3) : The $\bar{F}(x)$ of GOGE-W distribution with the parameters $a = 0.5$;
$\lambda = 0.7$; $\alpha = (0.7, 1.1, 1.5, 2.8)$; $b = (0.6, 0.9, 1.6, 1.9)$ $and\ k = (1.7, 1.9, 2, 2.9)$.

Figure **(3)** indicates that the $\bar{F}(x)$ of GOGE-W is non increasing function and we can see that :

(1) $\bar{F}(\infty) = \mathbf{0}$

(2) $\bar{F}(0) = \mathbf{1}$

2.2.2 Hazard Function

The hazard function of $X \sim$ GOGE-W $(\alpha, \lambda, a, b, k)$ is defined as follows :

$$h(x;\alpha,\lambda,a,b,k) = \frac{f(x;\alpha,\lambda,a,b,k)}{\bar{F}(x;\alpha,\lambda,a,b,k)} = \frac{\alpha\lambda abk\, x^{b-1} e^{ax^b} e^{-\lambda(e^{ax^b}-1)^k} (e^{ax^b}-1)^{k-1}}{1-\left[1-e^{-\lambda(e^{ax^b}-1)^k}\right]^{\alpha}} \times \left[1-e^{-\lambda(e^{ax^b}-1)^k}\right]^{\alpha-1} \quad (14)$$

The plots of hazard function of GOGE-W are given by the following figures:

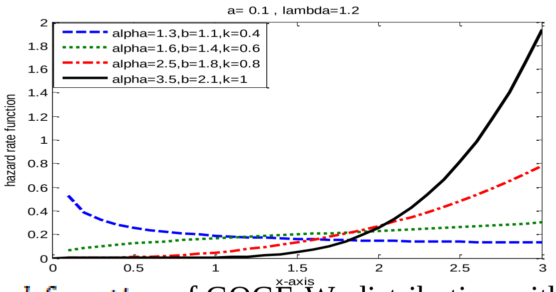


Figure (4) : The *hazard function* of GOGE-W distribution with the parameters $\lambda = 1.2$; $a = 0.1$; $\alpha = (1.3, 1.6, 2.5, 3.5)$; $b = (1.1, 1.4, 1.8, 2.1)$, $k = (0.4, 0.6, 0.8, 1)$.

2.2.3 The Reverse Hazard Function

The reverse hazard function of $X \sim G$OGE-W $(\alpha, \lambda, a, b, k)$ is defined as follows:

$$r(x;\alpha,\lambda,a,b,k) = \frac{f(x;\alpha,\lambda,a,b,k)}{F(x;\alpha,\lambda,a,b,k)} = \frac{\alpha\lambda abk\, x^{b-1} e^{ax^b} e^{-\lambda(e^{ax^b}-1)^k} (e^{ax^b}-1)^{k-1}}{\left[1-e^{-\lambda(e^{ax^b}-1)^k}\right]} \quad (15)$$

The plots of reverse hazard function of OGE-W-E are given by the following figures:

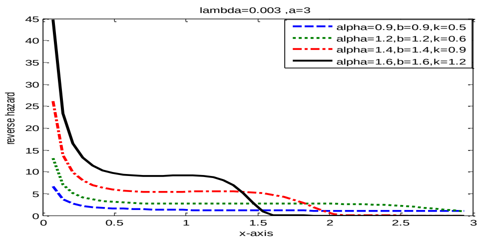


Figure (5) : the reverse hazard function of GOGE-W distribution with the parameters $a = 3$; $\lambda = 0.003$; $\alpha = (0.9, 1.2, 1.4, 1.6)$; $b = (0.9, 1.2, 1.4, 1.6)$ and $k = (0.5, 0.6, 0.9, 1.2)$.

2.2.4 The Cumulative Hazard Function

The cumulative hazard function of $X\sim$ GOGE-W (α,λ,a,b,k) is defined as: $H(x;\alpha,\lambda,a,b,k) = -ln[\,1- F(x;\alpha,\lambda,a,b,k)]$

$$= -\ln\left[1-\left(1-e^{-\lambda(e^{ax^b}-1)^k}\right)^{\alpha}\right], \qquad (16)$$

The plots of cumulative hazard function of GOGE-W are given by the following figures:

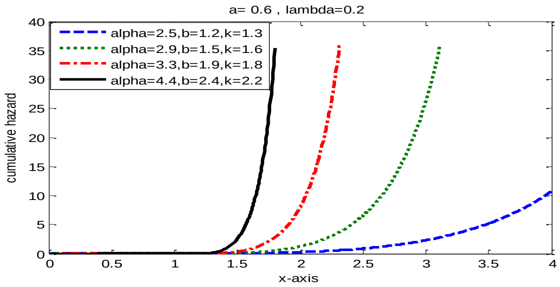


Figure (6) : The cumulative hazard of GOGE-W distribution with the parameter $a = 0.6;$ $\lambda = 0.2;\ \alpha = (2.5, 2.9, 3.3, 4.4); b = (1.2, 1.5, 1.9, 2.4)\ and\ k = (1.3, 1.6, 1.8, 2.2).$

2.3.Quantile and Median

The quantile x_q of the GOGE-W distribution is given by:

$$F(x_q) = P(x_q \le q) = q, 0 < q < 1, \qquad (`7)$$

From equation (7), we obtain the following equation:

$$\left[1-e^{-\lambda(e^{a x_q^{\,b}}-1)^k}\right]^{\alpha} = q \quad , \qquad (18)$$

We can obtain x_q by solving the following equation:

$$x_q = [\tfrac{1}{a} ln|\ ((-\tfrac{1}{\lambda} ln\ |\left(1-q^{\frac{1}{\alpha}}\right)|)^{\frac{1}{k}}+1)\ |]^{\frac{1}{b}} \qquad (19)$$

And we can obtain the median of GOGE-W by setting q=0.5 in Eq. **(19)** and solve this equation.

2.4. The Moments and Coefficient of Skewness , Kurtosis and Variation

In this section we introduce the r^{th} moment about the origin , rth moment about the mean and Coefficient of Skewness , Kurtosis and Variation for the GOGE-W distribution.

Theorem 1 : The r^{th} moment about the origin for GOGE-W is given by:

$$\mu'_r = \frac{\alpha\lambda K}{a^{\frac{r}{b}}}\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{m=0}^{\infty}\frac{\binom{\alpha-1}{i}\binom{k(j+1)-1}{m}(-1)^{i+j+m}(i+1)^j\,\lambda^j\,\Gamma(\frac{r}{b}+1)}{j!\,(-k(j+1)+m)^{1+\frac{r}{b}}},$$

Proposition 2: The mean (μ) for a random variable $X\sim G$OGE-W is given by:

$$\mu = \frac{\alpha\lambda K}{a^{\frac{1}{b}}} \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \sum_{m=0}^{\infty} \frac{\binom{\alpha-1}{i}\binom{k(j+1)-1}{m}(-1)^{i+j+m}(i+1)^j \lambda^j \Gamma\left(\frac{1}{b}+1\right)}{j!\left(-k(j+1)+m\right)^{1+\frac{1}{b}}}, \qquad \textbf{(20)}$$

2.4.1 Coefficient of Skewness , Kurtosis and Variation

In this subsection we derive the skewness, kurtosis and variation of GOGE-W distribution based on the moment as in the following theorem:

Theorem 3 The skewness, kurtosis and variation for a random variable $X \sim G$OGE-W $(\alpha, \lambda, a, b, k)$ are given in **1** , **2** and **3,** respectively as follows:

1.The coefficient of skewness (CS) of GOGE-W distribution is given by

$$CS = \frac{A}{B} \qquad (21)$$

where

$$A = \alpha\lambda K \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \sum_{m=0}^{\infty} \sum_{n=0}^{3} \frac{\binom{\alpha-1}{i}\binom{k(j+1)-1}{m}\binom{3}{n}\mu^{3-n}(-1)^{i+j+m+3-n}(i+1)^j \lambda^j \Gamma\left(\frac{n}{b}+1\right)}{j!\ a^{\frac{n}{b}}\left(-k(j+1)+m\right)^{1+\frac{n}{b}}}$$

and

$$B = \left[\alpha\lambda K \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \sum_{m=0}^{\infty} \sum_{n=0}^{2} \frac{\binom{\alpha-1}{i}\binom{k(j+1)-1}{m}\binom{2}{n}\mu^{2-n}(-1)^{i+j+m+2-n}(i+1)^j \lambda^j \Gamma\left(\frac{n}{b}+1\right)}{j!\ a^{\frac{n}{b}}\left(-k(j+1)+m\right)^{1+\frac{n}{b}}}\right]^{\frac{3}{2}}$$

2.The coefficient of kurtosis (Ck) of GOGE-W distribution is given by:

$$CK = \frac{C}{D} \qquad (22)$$

where

$$C = \alpha\lambda K \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \sum_{m=0}^{\infty} \sum_{n=0}^{4} \frac{\binom{\alpha-1}{i}\binom{k(j+1)-1}{m}\binom{4}{n}\mu^{4-n}(-1)^{i+j+m+4-n}(i+1)^j \lambda^j \Gamma\left(\frac{n}{b}+1\right)}{j!\, a^{\frac{n}{b}}\left(-k(j+1)+m\right)^{1+\frac{n}{b}}} \text{ and}$$

$$D = \left(\alpha\lambda K \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \sum_{m=0}^{\infty} \sum_{n=0}^{2} \frac{\binom{\alpha-1}{i}\binom{k(j+1)-1}{m}\binom{2}{n}\mu^{2-n}(-1)^{i+j+m+2-n}(i+1)^j \lambda^j \Gamma\left(\frac{n}{b}+1\right)}{j!\ a^{\frac{n}{b}}\left(-k(j+1)+m\right)^{1+\frac{n}{b}}}\right)^2$$

3.The coefficient of variation (CV) of GOGE-W distribution is given by:

$$CV = \frac{E}{F} \qquad (23)$$

where

$$E = \left(\alpha\lambda K \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \sum_{m=0}^{\infty} \sum_{n=0}^{2} \frac{\binom{\alpha-1}{i}\binom{k(j+1)-1}{m}\binom{2}{n}\mu^{2-n}(-1)^{i+j+m+2-n}(i+1)^j \lambda^j \Gamma\left(\frac{n}{b}+1\right)}{j!\ a^{\frac{n}{b}}\left(-k(j+1)+m\right)^{1+\frac{n}{b}}}\right)^{\frac{1}{2}}$$

and

$$F=\frac{\alpha\lambda K}{a^{\frac{1}{b}}}\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{m=0}^{\infty}\frac{\binom{\alpha-1}{i}\binom{k(j+1)-1}{m}(-1)^{i+j+m}(i+1)^{j}\,\lambda^{j}\,\Gamma(\frac{1}{b}+1)}{j!(-k(j+1)+m)^{1+\frac{1}{b}}}$$

2.4.2 The Moment Generating Function

The moment generating function of GOGE-W distribution is given by the following theorem:

Theorem 3.6 The moment generating function $M_X(t)$ of GOGE-W distribution is given by:

$$M_X(t)=\alpha\lambda K\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{m=0}^{\infty}\sum_{r=0}^{\infty}\frac{\binom{\alpha-1}{i}\binom{k(j+1)-1}{m}t^{r}(-1)^{i+j+m}\,(i+1)^{j}\,\lambda^{j}\,\Gamma(\frac{r}{b}+1)}{r!\,j!\,a^{\frac{r}{b}}\,(-k(j+1)+m)^{1+\frac{r}{b}}}$$

2.5 Order Statistics

In this section, the pdf of the j^{th} order statistic and the pdf of the smallest and largest order statistics of GOGE-W distribution are derived. Let $X_1, X_2, \ldots, X_n$ be a random sample from an GOGE-W distribution and $X_{1:n} \le X_{2:n} \le \cdots \le X_{n:n}$ denote the order statistics obtained from this sample, then the pdf of $X_{j:n}$ is given by:

$$f_{j:n}(x;\Phi)=\frac{1}{B(j,n-j+1)}[F(x;\Phi)]^{j-1}[1-F(x;\Phi)]^{n-j}f(x;\Phi) \qquad (24)$$

Where $f(x;\Phi)$ is the pdf of GOGE-W distribution given by equation (3.4), $F(x;\Phi)$ is the cdf of GOGE-W distribution given by equation (7) , $\Phi=(\alpha,\lambda,a,b,k)$ and B(., .) which is the Beta function. Then the pdf for the jth order statistic is as follows:

$$f_{j:n}(x;\Phi)=\alpha\lambda abk\sum_{i=0}^{n-j}\frac{(-1)^{i}n!}{i!\,(j-1)!\,(n-j-i)!}x^{b-1}e^{ax^{b}}e^{-\lambda\left(e^{ax^{b}}-1\right)^{k}}\left(e^{ax^{b}}-1\right)^{k-1}$$

$$\times\left[1-e^{-\lambda(e^{ax^{b}}-1)^{k}}\right]^{\alpha(j+i)-1} \qquad (25)$$

Then we can find the pdf of the smallest order statistics , say $f_{1:n}(x;\Phi)$ and the largest order statistics, say $f_{n:n}(x;\Phi)$ as follows:

$$f_{1:n}(x;\Phi)=\alpha\lambda abk\sum_{i=0}^{n-1}\frac{(-1)^{i}n!}{i!\,(n-1-i)!}x^{b-1}e^{ax^{b}}e^{-\lambda\left(e^{ax^{b}}-1\right)^{k}}\left(e^{ax^{b}}-1\right)^{k-1}$$

$$\times\left[1-e^{-\lambda\left(e^{ax^{b}}-1\right)^{k}}\right]^{\alpha(1+i)-1} \qquad (26)$$

$$f_{n:n}(x;\Phi)=n\alpha\lambda abk\,x^{b-1}e^{ax^{b}}e^{-\lambda\left(e^{ax^{b}}-1\right)^{k}}\left(e^{ax^{b}}-1\right)^{k-1}\left[1-e^{-\lambda(e^{ax^{b}}-1)^{k}}\right]^{\alpha n-1} \qquad (27)$$

2.6. Rényi Entropy [22]

The Rényi entropy of a random variable X with probability density function $f(x)$ is defined by:

$$I_R(\delta)=\frac{1}{1-\delta}\log\left(\int_0^{\infty}f^{\delta}(x)dx\right), where\ \delta>0\ and\ \delta\neq 1 \qquad (28)$$

Proposition 3.8.1. If X a random variable has a GOGE-W distribution ,then the Rényi entropy of X is given by:

$$I_R(\delta) = \frac{1}{1-\delta}\log\left[\alpha^\delta \lambda^\delta b^\delta k^\delta \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{m=0}^{\infty} \frac{\binom{\delta\alpha-\delta}{i}\binom{k(j+\delta)-\delta}{m}(-1)^{i+j+m}(i+\delta)^j\,\lambda^j\,\Gamma\left(\frac{\delta b-\delta+1}{b}\right)}{j!\,a^{\frac{1-\delta}{b}}(-k(j+\delta)+m)^{\frac{\delta b-\delta+1}{b}}}\right] \quad (29)$$

2.7. Parameters Estimation of GOGE-W Distribution

In this section, the considered estimation methods (the maximum likelihood estimation and the moment method) have been illustrated to estimate the five parameters of GOGE-W distribution .

2.7.1. Maximum Likelihood Estimation

If $x_1, x_2, \dots, x_n$ denotes a random sample from the GOGE-W distribution ,then the Likelihood function is given by:

$L = \prod_{i=1}^{n} f(x_i; \alpha, \lambda, a, b, k)$, (*)

By substituting the Eq. (8) into Eq. (*), we obtain:

$$L = \prod_{i=1}^{n} \alpha\lambda abk\, x_i^{\,b-1} e^{a x_i^b} e^{-\lambda\left(e^{a x_i^b}-1\right)^k} \left(e^{a x_i^b}-1\right)^{k-1}\left[1-e^{-\lambda\left(e^{a x_i^b}-1\right)^k}\right]^{\alpha-1}, (30)$$

the log-likelihood function is:

$$\ell = n\ln(\alpha) + n\ln(\lambda) + n\ln(a) + n\ln(b) + n\ln(k) + (b-1)\sum_{i=1}^{n}\ln(x_i) + a\sum_{i=1}^{n} x_i^{\,b}$$
$$-\lambda\sum_{i=1}^{n}\left(e^{a x_i^b}-1\right)^k + (k-1)\sum_{i=1}^{n}\ln\left(e^{a x_i^b}-1\right) + (\alpha-1)\sum_{i=0}^{n}\ln\left(1-e^{-\lambda\left(e^{a x_i^b}-1\right)^k}\right), \quad (31)$$

By taking the partial derivatives of ℓ with respect to the parameters α, λ, a, b and k and setting the results equal zeros, we get the following equations:

$$\frac{\partial\ell}{\partial\alpha} = \frac{n}{\alpha} + \sum_{i=1}^{n}\ln\left[1-e^{-\lambda\left(e^{a x_i^b}-1\right)^k}\right] = 0 \quad (32)$$

$$\frac{\partial\ell}{\partial\lambda} = \frac{n}{\lambda} - \sum_{i=1}^{n}\left(e^{a x_i^b}-1\right)^k + (\alpha-1)\sum_{i=1}^{n}\frac{\left(e^{a x_i^b}-1\right)^k e^{-\lambda\left(e^{a x_i^b}-1\right)^k}}{\left[1-e^{-\lambda\left(e^{a x_i^b}-1\right)^k}\right]} = 0, \quad (33)$$

$$\frac{\partial\ell}{\partial a} = \frac{n}{a} - \lambda k\sum_{i=1}^{n}\left(e^{a x_i^b}-1\right)^{k-1} e^{a x_i^b} x_i^{\,b} + \sum_{i=1}^{n} x_i^{\,b} + (k-1)\sum_{i=0}^{n}\frac{e^{a x_i^b} x_i^{\,b}}{\left(e^{a x_i^b}-1\right)}$$

$$+\lambda k(\alpha-1)\sum_{i=1}^{n}\frac{\left(e^{ax_i^b}-1\right)^{k-1} e^{-\lambda\left(e^{ax_i^b}-1\right)^k} e^{ax_i^b} x_i^b}{\left[1-e^{-\lambda\left(e^{ax_i^b}-1\right)^k}\right]}=0 \qquad (34)$$

$$\frac{\partial \ell}{\partial b}=\frac{n}{b}+\sum_{i=1}^{n} ln\,(x_i)-\lambda ak\sum_{i=1}^{n}\left(e^{ax_i^b}-1\right)^{k-1} e^{ax_i^b} x_i^b ln(x_i)+a\sum_{i=1}^{n} x_i^b ln(x_i)$$

$$+a(k-1)\sum_{i=1}^{n}\frac{ln(x_i)\, e^{ax_i^b} x_i^b}{\left(e^{ax_i^b}-1\right)}+a\lambda k(\alpha-1)\sum_{i=0}^{n}\frac{ln(x_i)\left(e^{ax_i^b}-1\right)^{k-1} e^{-\lambda\left(e^{ax_i^b}-1\right)^k} e^{ax_i^b} x_i^b}{\left[1-e^{-\lambda\left(e^{ax_i^b}-1\right)^k}\right]}=0\,. \qquad (35)$$

$$\frac{\partial \ell}{\partial k}=\frac{n}{k}-\lambda\sum_{i=1}^{n}\left(e^{ax_i^b}-1\right)^{k} ln(e^{ax_i^b}-1)+\sum_{i=1}^{n} ln(e^{ax_i^b}-1)$$

$$+\lambda(\alpha-1)\sum_{i=1}^{n}\frac{\left(e^{ax_i^b}-1\right)^{k} ln(e^{ax_i^b}-1)}{\left[1-e^{-\lambda\left(e^{ax_i^b}-1\right)^k}\right]}=0 \qquad (36)$$

The MLEs estimators of the parameters α, λ, a, b *and* k can be obtained by solving these equations numerically.

2.7.2. Method of Moment Estimators

Let $X_1, X_2, X_3, \ldots, X_4$ be a random sample from $X \sim$GOGE-W $(\alpha, \lambda, a, b, k)$, the method of moments of GOGE-W distribution is defined by the following equation:

$$E(X^r)=\sum_{s=1}^{n}\frac{1}{n}X_s^{\;r} \qquad (37)$$

Where $E(X^r)$ is the rth moment about the origin given in equation (37).

For the case r=1, equation (37) becomes as follows:

$$\frac{\alpha\lambda K}{a^{\frac{1}{b}}}\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{m=0}^{\infty}\frac{\binom{\alpha-1}{i}\binom{k(j+1)-1}{m}(-1)^{i+j+m}(i+1)^j \lambda^j\, \Gamma(\frac{1}{b}+1)}{j!(-k(j+1)+m)^{1+\frac{1}{b}}}=\overline{X}, \qquad (38)$$

For the case r=2, equation (37) becomes as follows:

$$\frac{\alpha\lambda K}{a^{\frac{2}{b}}}\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{m=0}^{\infty}\frac{\binom{\alpha-1}{i}\binom{k(j+1)-1}{m}(-1)^{i+j+m}(i+1)^j \lambda^j\, \Gamma(\frac{2}{b}+1)}{j!\,(-k(j+1)+m)^{1+\frac{2}{b}}}=\sum_{s=1}^{n}\frac{1}{n}X_s^{\;2} \qquad (39)$$

For the case r=3, equation (37) becomes as follows:

$$\frac{\alpha\lambda K}{a^{\frac{3}{b}}}\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{m=0}^{\infty}\frac{\binom{\alpha-1}{i}\binom{k(j+1)-1}{m}(-1)^{i+j+m}(i+1)^j \lambda^j\, \Gamma(\frac{3}{b}+1)}{j!(-k(j+1)+m)^{1+\frac{3}{b}}}=\sum_{s=1}^{n}\frac{1}{n}X_s^{\;3} \qquad (40)$$

For the case r=4, equation (37) becomes as follows:

$$\frac{\alpha\lambda K}{a^{\frac{4}{b}}}\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{m=0}^{\infty}\frac{\binom{\alpha-1}{i}\binom{k(j+1)-1}{m}(-1)^{i+j+m}(i+1)^j\,\lambda^j\,\Gamma(\frac{4}{b}+1)}{j!(-k(j+1)+m)^{1+\frac{4}{b}}} = \sum_{s=1}^{n}\frac{1}{n}X_s{}^4 \qquad (41)$$

For the case r=5, equation (37) becomes as follows:

$$\frac{\alpha\lambda K}{a^{\frac{5}{b}}}\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{m=0}^{\infty}\frac{\binom{\alpha-1}{i}\binom{k(j+1)-1}{m}(-1)^{i+j+m}(i+1)^j\,\lambda^j\,\Gamma(\frac{5}{b}+1)}{j!(-k(j+1)+m)^{1+\frac{5}{b}}} = \sum_{s=1}^{n}\frac{1}{n}X_s{}^5 \qquad (42)$$

By using the numerical methods such as Newton Raphson method, we can get the estimators for the parameters α, λ, a, b *and* k by solving the equations (38)$-$ (42) for α, λ, a, b *and* k.

2.8. Applications

In this section, we provide two applications to real data to demonstrate the importance of the GOGE-W distribution and we will compare GOGE-W distribution with the following distributions:

Odd Generalized Exponential Weibull (OGE-W) with cdf is

$$F(x;\alpha,\lambda,\beta,\theta) = [1 - e^{-\lambda\left(e^{\theta x^{\beta}}-1\right)}]^{\alpha-1}, x > 0.$$

Modified Weibull distribution (MWD) with cdf $F(x;\alpha,\beta,\gamma) = 1 - e^{-\alpha x-\beta x^{\gamma}}, x > 0.$

Flexible Weibull (FW) with cdf *is* $F(x;\alpha,\gamma,\beta,\theta) = 1 - e^{-e^{(\beta x^{\gamma}+\theta x^{\alpha})}}, x > 0.$

In order to compare the GOGE-W distribution with the above distributions, the measures of goodness-of-fit including the Akaike Information Criterion (AIC), Hannan-Quinn Information Criterion (HQIC), Consistent Akaike Information Criterion (CAIC), and Bayesian Information Criterion (BIC) are used .

First Data Set

We have a real data set from Colorado Climate Center, Colorado State University (http://ulysses.atmos.colostate.edu). These data consist of 100 annual maximum precipitation (inches) for one rain gauge in Fort Collins, Colorado, from 1900 through 1999. The data set are : 239, 232, 434, 85, 302, 174, 170, 121, 193, 168, 148, 116, 132, 132, 144, 183, 223, 96, 298, 97, 116, 146, 84, 230, 138, 170, 117, 115, 132, 125, 156, 124, 189, 193, 71, 176, 105, 93, 354, 60, 151, 160, 219, 142, 117, 87, 223, 215, 108, 354, 213, 306, 169, 184, 71, 98, 96, 218, 176, 121, 161, 321, 102, 269, 98, 271, 95, 212, 151, 136, 240, 162, 71, 110, 285, 215, 103, 443, 185, 199, 115, 134, 297, 187, 203, 146, 94, 129, 162, 112, 348, 95, 249, 103, 181, 152, 135, 463, 183, 241. The MLEs of the model parameters for the first data are given in Table (1) and the numerical values of the model selection statistics $\hat{\ell}$, AIC, HQIC, CAIC and BIC are listed in Table (2). We can see from Table (2) that the GOGE-W model gives the smallest values for the criteria AIC, HQIC, CAIC and BIC so it represents the first data set better than the other selected models .

Table 1. Parameters Estimates for the First Data .

Model	Parameters estimates				
GOGE-W (α,λ,a,b,k)	$\hat{\alpha}$ =7.231	$\hat{\lambda}$=0.004	$\hat{a}$ =1.375	$\hat{b}$ =0.143	$\hat{k}$ = 2.336

OGE-W $(\alpha,\lambda,\beta,\theta)$		$\hat{\alpha}$ =3.395	$\hat{\lambda}$=0.104	$\hat{\beta}$=0.418	
MWD (α,β,γ)	$\hat{\alpha}$=0.006,	$\hat{\beta}$=0.002	$\hat{\gamma}$=0.001		
(FW) $(\alpha,\gamma,\beta,\theta)$	$\hat{\alpha}$ =0.514	$\hat{\gamma}$=0.065	$\hat{\beta}$=0.223	$\hat{\theta}$=0.043	

Table 2 The Values of the Statistics $\hat{\ell}$, AIC, HQIC, CAIC and BIC for the First Data Set.

model	$\hat{\ell}$	AIC	HQIC	CAIC	BIC
GOGE-W (α,λ,a,b,k)	-571.2944	1152.6	1157.9	1153.2	1165.6
OGE-W $(\alpha,\lambda,\beta,\theta)$	-574.506	1157	1161.2	1157.4	1167.4
MWD (α,β,γ)	-617.2024	1240.4	1243.6	1240.7	1248.2
(FW) $(\alpha,\gamma,\beta,\theta)$	-780.9602	1569.9	1574.1	1570.3	1580.3

.

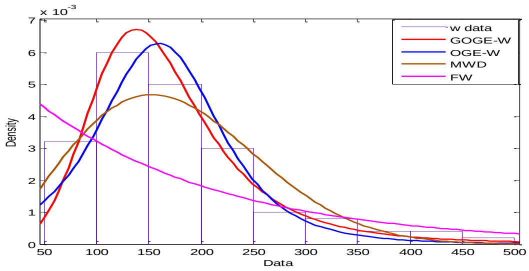


Figure (7) : Estimated densities of selection distributions for the first data set

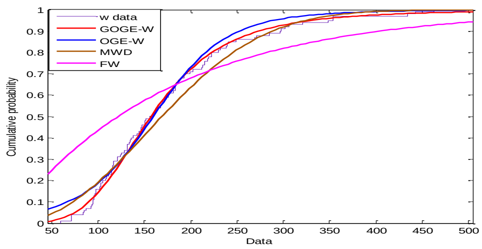


Figure (8) : Estimated cdfs of selection distributions for the first data set.

Second Data Set [1]

The second data set represents the life time of 50 devices . The data set are: 0.1, 0.2, 1, 1, 1, 1, 1, 2, 3, 6, 7, 1, 1, 12, 18, 18, 18, 18, 18, 21, 32, 36, 40, 45, 46, 47, 50, 55, 60, 63, 63, 67, 67, 67, 67, 72, 75, 79, 82, 82, 83, 84, 84, 84, 85, 85, 85, 85, 85, 86, 86. The MLEs of the model parameters for the second data are given in Table (3) and the numerical values of the model selection statistics $\hat{\ell}$, AIC, HQIC, CAIC and BIC are listed in Table (4). We can see from table (4) that the GOGE-W model gives the smallest values for the criteria AIC, HQIC, CAIC and BIC so it represents the second data set better than the other selected

models .

Table 3Parameters Estimates for the Second Data

Model	$\hat{\ell}$	AIC	HQIC	CAIC	BIC

GOGE-W $(\alpha, \lambda, a, b, k)$	-219.3052	448.6104	452.251	449.9741	458.1705
OGE-W $(\alpha, \lambda, \beta, \theta)$	-228.9607	465.9214	468.8339	466.8103	473.5695
MWD (α, β, γ)	-242.1313	490.2626	492.4469	490.7843	495.9986
(FW) $(\alpha, \gamma, \beta, \theta)$	-286.467	581.0881	584.0006	581.977	588.7362

Table 4 The Values of Statistics $\hat{\ell}$, AIC, HQIC, CAIC and BIC for the Second Data Set.

Model	Parameters estimates				
GOGE-W $(\alpha, \lambda, a, b, k)$	$\hat{\alpha}$ =0.409	$\hat{\lambda}$=0.006	$\hat{a}$ =0.171	$\hat{b}$ =0.83	$\hat{k}$ = 0.816
OGE-W $(\alpha, \lambda, \beta, \theta)$	$\hat{\alpha}$ =0.825	$\hat{\lambda}$=0.061	$\hat{\beta}$=0.512	$\hat{\theta}$=0.385	
MWD (α, β, γ)	$\hat{\alpha}$=0.018,	$\hat{\beta}$=0.054	$\hat{\gamma}$=0.388	-	-
(FW) $(\alpha, \gamma, \beta, \theta)$	$\hat{\alpha}$ =1.032	$\hat{\gamma}$=0.42	$\hat{\beta}$=0.047	$\hat{\theta}$=0.008	

References

[1] Aarset, M. V. How to identify bathtub hazard rate. IEEE Transactions on Reliability,36 (1987) (106-108).

[2] Almalki, S. J. and Yuan, J. The new modified Weibull distribution, Reliability Engineering and System Safety 111 (2013) 164-170.

[3] Bebbington, M. S., Lai C. D and Zitikis, R., A flexible Weibull extension, Reliability Engineering and System Safety **92(6)** (2007), 26-719.

[4] El-Damcese, M. A., Mustafa, A., El-Desouky, B. S. and Mustafa, M.E.,The odd generalized exponential gompertz, arXiv preprint arXiv :1507.06400 (2015).

[5] Famoye, F., Lee ,C. and Olumolade, O., The beta-Weibull distribution, Journal of Statistical Theory and Application **4(2)** (2005) 36-121.

[6] Gupta ,R.D. and kundu, D., Generalized exponential distribution: Existing results and some recent Developments, Journal of Statistics Planning and Inference **137(11)** (1999) 3537-3547

[7] Lai,C.D., Xie, M. and Murthy,D. N., A modified Weibull distribution, IEEE *Transaction on Reliability **52(1)** (2003), 7-33*

[8] Luguterah A. and Nasiru, S., The Odd Generalized Exponential Generalized Linear Exponential (OGE-GLE) distribution, Journal of Statistics Application & Probability, **6(1)** (2017) 139-148.

[9] Mudholkar, G.S. and Srivastava, D.K., Exponentiated Weibull family for analyzing bathtub failure- rate data, IEEE Transactions on Reliability **42(2)** (1993) 299-302.

[10] Mustafa, A., El-Desouky, B. S. and AL-Garash, S., Odd Generalized Exponential Flexible Weibull Extension (OGE-FEW) distribution , Journal of Statistical Theory and Applications, 17(1) (2018)(77-90).

[11] Nadarajah, S., Cordeiro, G.M, and Ortega, E.M.M., General results for the beta-modified Weibull distribution, Journal of Statistical Computation and Simulation **81(10)** (2011).

[12] Rényi, A . (1961). On measures of entropy and information. Proceeding of the Fourth Berkeley Symposium on Mathematics, Statistics and Probability 1, University of California Press, Berkeley, 547-561

[13] Sarhan, A.M. and Zaindin, M., Modified Weibull distribution, Applied Sciences **11** (2009) 123-136.

[14] Silva, G.O. Ortega, E.M. and Cordeiro, G.M. The beta modified Weibull distribution, Lifetime Data Analysis **16** (2010), 30-409.

[15] Singla, N. Jain, K. and Kumar, S.S. The beta generalized Weibull distribution:properties and applications, Reliability Engineering and System Safety **102**(2012), 5-15.

[16] Tahir, M. H., Cordeiro, G. M., Alizadeh, M., Mansoor, M. Zubair, M. and Hamedani, G.G.,The odd generalized exponential family of distribution with Application, Journal of Statistical Distribution and Applications **2(1)** (2015).

Lomax-Rayleigh Distribution: traditional and heuristic methods of Estimation

Nashaat Jasim Mohammed Al-Anber

Technical College of Management- Baghdad
Middle Technical University Baghdad
Dr.nashaat@mtu.edu.iq

Abstract: We introduce a continuous distribution called the Lomax-Rayleigh (L-R) distribution that extends the Lomax distribution. The generalization of the probability density function and cumulative distribution function of this distribution, the expression for the r^{th} moment and moment generating function was established.

We considered a traditional methods of estimation such as the maximum likelihood method and nonlinear least square estimation method to estimate the parameters and utilizing the Artificial Intelligence Algorithms such as Genetics Algorithm and pattern search method in estimation process. A comparison study among that methods is carried out through simulation experiments.

We concluded that pattern search method is more efficient than other methods depending upon mean square error criteria.

Keywords: Rayleigh distribution, Lomax distribution, hazard rate function, Genetics algorithm, pattern search method.

1. Introduction:

Probability distribution has many applications in describing real world situations. There are lot of researches have shown that some real life data that cannot be modeled adequately by traditional statistical distributions, because the complexity found in it. Recently there is rapid grown direction toward generalization , mixing, Transmuting and exponentiation of existing distributions, so some families and new general formulas of distributions is appeared in papers that dealing with skewed data and data drawn from non-homogenous populations. Some of the earlier works include those Gupta and Kundu, 1999[12] , Eugene et al., 2004[9] , Famoye et al.,2005[10] , Akinsete et al., 2008[4], Miroslav and Balakrishnan., 2012[14] , Alzaatreh et al.,2013[3] , Adeleke et al., 2013[5]; Akarawak et al., 2013[6] and Akarawak et al., 2015[6], Ghosh and Hamedani 2015[11] and Khaleel et al.,2016[13].

The main idea was that any parametric family of distributions can be incorporated into larger families through an application of the probability integral transform. Then the number of parameters and complexity of new families is increased, so almost in most cases numerical techniques to get estimates of parameters are used, since closed form of estimators of that parameters are difficult to derive it.

Recently Venegas et al. 2019[17] introduced the two parameter Lomax-Rayleigh distribution as compound between Lomax and Rayleigh distribution. Özsoy, 2020[16] used the heuristic optimization approaches such as Genetic Algorithms , Differential Evolution, Particle Swarm Optimization, and Simulated Annealing to estimate the parameter of generalized gamma distribution: comparison among maximum likelihood method and heuristic optimization approaches are proved that later have nice properties in estimation of parameters.

The rest of the paper is organized as follows. In Section 2 we present the probability density function and cumulative distribution function, reliability and hazard function of the Lomax-Rayleigh model with its general formula of moments. In Section 3 we discuss traditional methods such as: maximum likelihood estimation and non-linear least square estimation method and we introduce heuristic optimization approaches such as Genetic Algorithms and pattern search method. In section 4 we present a comparison among different estimation methods via simulation experiments. Finally, in Section 5,6 we report the discussion and final conclusions

2. Lomix-Rayleigh Distribution (L-R)

$F(x) = \int_0^x f(u)du$ ……. (1)

Let $x = -\log(1-G)$

Where G is a cumulative distribution function of another distribution.

$F(x) = \int_0^{-\log(1-G)} f(u)du$ …….. (2)

By differentiate the equation(2) for the base distribution , we get:

$f(x) = f(-\log(1-G))\frac{g}{(1-G)}$ ………. (3)

Equation(3) represent a transformation for mixing two distributions through inserting the cumulative distribution function of any distribution for the base distribution so we called it Lomax-X family .

2.2 Derivation of Lomax- Rayleigh Distribution (L-R)

2.2.1 Derivation of the pdf and cdf of R-L

The pdf and cdf of the Lomax - Rayleigh distribution is derived in this section as a class of Lomax-X family of generalized distributions.

Theorem 2.1:

Let the pdf of a Rayleigh distribution which is abase variable be:

$f(x) = 2\lambda x e^{-\lambda x^2}$, $x > 0, \lambda > 0$ …….. (4)

And the pdf Lomax distributed random variable be:

$f(y) = \frac{\alpha}{\beta}\left(1+\frac{y}{\beta}\right)^{-(\alpha+1)}$, y> 0 , $\alpha, \beta > 0$ ……. (5)

Then the pdf of the Rayleigh-Lomax distribution is given by:

$f(x) = 2\alpha\theta x(1+\theta x^2)^{-(\alpha+1)}$ $x > 0,\ \alpha, \theta > 0$ ……. (6)

Proof

The pdf of the Lomax-X family of distribution is given by:

$f(x) = f(-\log(1-G))\frac{g}{(1-G)}$ ……. (7)

Where g and G are the pdf and cdf of Rayleigh distribution.

By substituting the pdf and cdf or Rayleigh distribution in equation (4) , we get:

$f(x) = \frac{2\alpha\lambda}{\beta} x\left(1+\frac{\lambda x^2}{\beta}\right)^{-(\alpha+1)}$, $x > 0, \alpha, \beta, \lambda > 0$ …… (8)

Let: $\frac{\lambda}{\beta} = \theta$, then the Lomax-Rayleigh distribution be:

$f(x) = 2\alpha\theta x(1+\theta x^2)^{-(\alpha+1)}$, $x > 0, \alpha, \theta > 0$

Where is α shape parameter and θ is scale parameter. The Lomax – Rayleigh distribution is right skewed as show in figure (1).

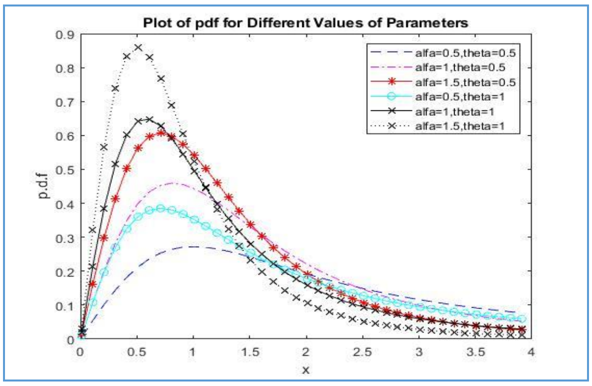


Figure (1):Plot of pdf for different values of parameters

Corollary 2.1: The function of Lomax-Rayleigh is pdf function.

Proof:

The integral of function must equal one. That is:

$$\int_0^\infty f(x)dx = 1$$

Then

$$\int_0^\infty f(x)dx = \int_0^\infty 2\alpha\theta x(1+\theta x^2)^{-(\alpha+1)}dx$$

$$= 2\alpha\theta\int_0^\infty xe^{-(\alpha+1)\log(1+\theta x^2)}dx$$

Let:

$$y = (\alpha+1)\log(1+\theta x^2)$$

Then:

$$x = \left(\frac{1}{\theta}\right)^{\frac{1}{2}}\left(1-e^{\frac{y}{(\alpha+1)}}\right)^{\frac{1}{2}}$$

$$dx = \left(\frac{1}{\theta}\right)^{\frac{1}{2}}\frac{1}{2}\left(1-e^{\frac{y}{(\alpha+1)}}\right)^{\frac{-1}{2}}\left(-e^{\frac{y}{(\alpha+1)}}\frac{dy}{(\alpha+1)}\right)$$

then

$$\int_0^\infty f(x)dx = 2\alpha\theta\int_0^\infty\left(\frac{1}{\theta}\right)^{\frac{1}{2}}\left(1-e^{\frac{y}{(\alpha+1)}}\right)^{\frac{1}{2}}e^{-y}\left(\frac{1}{\theta}\right)^{\frac{1}{2}}\frac{1}{2}\left(1-e^{\frac{y}{(\alpha+1)}}\right)^{\frac{-1}{2}}\frac{e^{\frac{y}{(\alpha+1)}}}{(\alpha+1)}dy$$

$$= \frac{\alpha}{(\alpha+1)}\int_0^\infty e^{-y\left(1-\frac{1}{(\alpha+1)}\right)}dy$$

Let:

$$z = \left(\frac{\alpha}{(\alpha+1)}\right)y$$

$$y = \left(\frac{(\alpha+1)}{\alpha}\right)z$$

$$dy = \left(\frac{(\alpha+1)}{\alpha}\right)dz$$

$$\int_0^\infty f(x)dx = \frac{\alpha}{(\alpha+1)}\int_0^\infty e^{-y\left(1-\frac{1}{(\alpha+1)}\right)}dy = \frac{\alpha}{(\alpha+1)}\int_0^\infty e^{-z}\frac{(\alpha+1)}{\alpha}dy = 1$$

Corollary 2.2: The cdf of Lomax-Rayleigh is:

$F(x) = 1 - e^{-\alpha \log(1+\theta x^2)}$

Proof:

$$F(x) = \Pr(X \le x) = \int_0^x f(u)du$$

Then:

$$F(x) = \int_0^x 2\alpha\theta u e^{-(\alpha+1)\log(1+\theta u^2)} du$$

Let:

$y = (\alpha+1)\log(1+\theta u^2)$

Then:

$$u = \left(\frac{1}{\theta}\right)^{\frac{1}{2}} \left(1 - e^{\frac{y}{(\alpha+1)}}\right)^{\frac{1}{2}}$$

$$du = \left(\frac{1}{\theta}\right)^{\frac{1}{2}} \frac{1}{2} \left(1 - e^{\frac{y}{(\alpha+1)}}\right)^{\frac{-1}{2}} \left(-e^{\frac{y}{(\alpha+1)}} \frac{dy}{(\alpha+1)}\right)$$

$$F(x) = \frac{\alpha}{(\alpha+1)} \int_0^{(\alpha+1)\log(1+\theta x^2)} e^{-y\,(1-\frac{1}{\alpha+1})} dy$$

Let:

$$z = \left(\frac{\alpha}{(\alpha+1)}\right) y$$

$$y = \left(\frac{(\alpha+1)}{\alpha}\right) z$$

$$dy = \left(\frac{(\alpha+1)}{\alpha}\right) dz$$

$$F(x) = \frac{\alpha}{(\alpha+1)} \int_0^{\alpha\log(1+\theta x^2)} e^{-z} \frac{(\alpha+1)}{\alpha} dz = 1 - e^{-\alpha\log(1+\theta x^2)} \qquad \text{........ (9)}$$

The survival function and hazard function are as follow:

$$S(x) = 1 - F(x) = e^{-\alpha\log(1+\theta x^2)} \qquad \text{...... (10)}$$

$$h(x) = \frac{f(x)}{S(x)} = \frac{2\theta}{1+\theta x^2} \qquad \text{...... (11)}$$

Moment generating function:

$$\mu_x(t) = E(e^{xt}) = \int_0^\infty e^{xt} 2\theta\alpha x (1+\theta x^2)^{-(\alpha+1)} dx$$

$$= 2\theta\alpha \int_0^\infty x e^{xt} e^{-(\alpha+1)\log(1+\theta x^2)} dx$$

Since: $e^y = \sum_{i=0}^{\infty} \frac{y^i}{i!}$

$$= 2\alpha\theta \int_0^\infty x \sum_{j=0}^{\infty} \frac{(xt)^j}{j!} e^{-(\alpha+1)\log(1+\theta x^2)} dx$$

$$= 2\alpha\theta \sum_{j=0}^{\infty} \frac{(t)^j}{j!} \int_0^\infty x^{j+1} e^{-(\alpha+1)\log(1+\theta x^2)} dx$$

Let:
$y = (\alpha + 1)\log(1 + \theta x^2)$
Then:
$$x = \left(e^{\frac{y}{\alpha+1}} - 1\right)^{1/2} \left(\frac{1}{\theta}\right)^{1/2}$$

$$dx = \frac{1}{2}\left(e^{\frac{y}{\alpha+1}} - 1\right)^{-\frac{1}{2}} \left(e^{\frac{y}{\alpha+1}} - \frac{1}{\alpha+1}\right)\left(\frac{1}{\theta}\right)^{1/2}$$

Then:
$$\mu_x(t) = \frac{\alpha\theta}{\alpha+1}\int_0^\infty \left(e^{\frac{y}{\alpha+1}} - 1\right)^{j/2} e^{-y\left(\frac{\alpha}{\alpha+1}\right)} dy$$
$$= \frac{\alpha\theta}{\alpha+1}\sum_{i=0}^{j/2} C_i^{j/2}(-1)^{\frac{j}{2}-i}\int_0^\infty e^{-y\left(\frac{\alpha-i}{\alpha+1}\right)} dy$$

Let:
$$z = y\left(\frac{\alpha - i}{\alpha + 1}\right)$$
Then:
$$y = z\left(\frac{\alpha + 1}{\alpha - i}\right)$$
$$dy = \left(\frac{\alpha + 1}{\alpha - i}\right) dz$$
$$= \frac{\alpha\theta}{\alpha+1}\sum_{i=0}^{j/2} C_i^{j/2}(-1)^{\frac{j}{2}-i}\int_0^\infty e^{-z}\left(\frac{\alpha+1}{\alpha-i}\right) dy$$
$$\mu_x(t) = \alpha\sqrt{\theta}\sum_{j=0}^{\infty}\frac{t^j}{j!}\sum_{i=0}^{j/2}\frac{C_i^{j/2}(-1)^{\frac{j}{2}-i}}{(\alpha-i)} \quad \text{……… (12)}$$

3 Estimation Methods

In this section we will derive the estimators of the unknown parameters by using traditional methods (Maximum likelihood and Nonlinear Least Square) and heuristic method (Genetic algorithm and pattern search)

3.1 Maximum Likelihood Method MLE:

The principle of this method is to find the values of the parameters that maximize the likelihood function, where the likelihood function is a function of data and unknown parameters only, so it represent the information of sample.
Then:
$$Likelihood\ Function = L(x, \alpha, \theta) = 2^n \alpha^n \theta^n \prod_{i=1}^{n} x_i\, e^{-(\alpha+1)\sum_{i=1}^{n}\log(1+\theta x^2)} \text{ .(13)}$$
Equation (13) is monotonic, so the values of parameters that maximize it is same as that maximize log of it.
Then:
$$LogL(\alpha,\theta) = nlog(2) + nlog(\alpha) + nlog(\theta) + \sum_{i=1}^{n}\log(\theta) - (\alpha+1)\sum_{i=1}^{n}\log(1+\theta x_i^2) \text{ ... (14)}$$
By taking partial derivatives of equation (14) and equating them to zero ,we get normal equation . then:

$\frac{n}{\hat{\alpha}} - \sum_{i=1}^{n} \log\left(1 + \hat{\theta} x_i^2\right) = 0$ ....(15)

$\frac{n}{\hat{\theta}} - (\hat{\alpha} + 1) \sum_{i=1}^{n} \frac{x_i^2}{1+\hat{\theta} x_i^2} = 0$ ....(16)

The solution of equation (15) and (16) , represent the estimate of parameters, by substitute equation(15) in equation(16) , we get

$\frac{n}{\hat{\theta}} - \left(\frac{n}{\sum_{i=1}^{n} \log(1+\hat{\theta} x_i^2)} + 1\right) \sum_{i=1}^{n} \frac{x_i^2}{1+\hat{\theta} x_i^2} = 0$ ……….. (17)

The equation (17) is highly nonlinear, so we use newton Raphson method to get the solution of the parameters θ, Let the $\hat{\theta}_{mle}$ be the maximum likelihood estimate of $\hat{\theta}_{mle}$ then from equation (17) the estimator of parameters α will be:

$\hat{\alpha}_{mle} = \frac{n}{\sum_{i=1}^{n} \log(1+\hat{\theta} x_i^2)}$ ……(18)

3.2 Non-Linear Least Square Method NLLSM:

The principle of this method is to find the values of the parameters that minimize the square sum of errors,

Let:

$\hat{F} = \frac{i}{n+1}$, be non-parametric estimator od cdf ,where

n:sample size.

$i = 1,2, \dots, n$

Then by equating the non-parametric estimator by cdf of distribution, we get:

$\hat{F} = 1 - e^{-\alpha \log(1+\theta x_i^2)}$

And

$(1 - \hat{F}) = e^{-\alpha \log(1+\theta x_i^2)}$

If we let: $y = (1 - \hat{F})$, then:

$y_i = e^{-\alpha \log(1+\theta x_i^2)} + \epsilon_i$ …….. (19)

Is nonlinear model of parameters α and. We can use nonlinear least square method that minimize the squared sun square of errors to get estimate of parameters. Where sum of square of error is:

$Q = \sum_{i=1}^{n} \left(y_i - e^{-\alpha \log(1+\theta x_i^2)}\right)^2$ ……. (20)

3.3 Genetic Algorithm GA:[8]

Genetic Algorithm is one of the most powerful stochastic optimization technique, its base idea from Charles Darwin's theory of natural evolution "survival of the fittest". It is very useful in estimation of nonlinear models, particularly in cases where the function cannot be solved in more traditional ways so it is more efficient of obtaining global optimum solution which is represent of parameter estimates. This algorithm reflects the method of natural selection where the fittest individuals are selected for reproduction in order to select better offspring from the parent population.

GA has its basic steps from genetics artificially to construct search algorithms that are robust and require minimal problem information with small overall computational time .It has three main operators which is selection, crossover and mutation, that making it an important tool for optimization.

The process of natural selection starts with the selection of fittest individuals from a population. So it works with a population of solutions instead of a single solution. They produce offspring which inherit the characteristics of the parents and will be added to the next generation. If parents have better fitness, their offspring will be better than parents and have a better chance at surviving. This process continually will find a generation with the best individuals.

GA is a population-based algorithm, where an individual in the population, representing a solution, which is called a chromosome. It is basically a binary vector, where each item in the vector is called a gene. Since the individuals are represented in binary, it is important to choose a proper encoding of the solution. The fitness is assigned to each chromosome and new generation of chromosomes is created in the reproduction process. Parent chromosomes, from which new chromosomes are created, are chosen quasi-randomly, so the better the fitness will have higher probability for the chromosome to be chosen. Next, the genetic operators are used to create descendant of parents. These include:

1. Initial population or start Generate random population of n chromosomes which is more suitable solutions for the problem.
2. Fitness function Evaluate the fitness $f(x)$ of each chromosome x in the population
3. Selection of two parent chromosomes from a population according to their fitness so the best fitness will be with higher probability to be selected.
4. Crossover: with a crossover probability cross over the parents to form a new offspring (children). If no crossover was performed, offspring is an exact copy of parents
5. Mutation: with a mutation probability mutate new offspring at each local which is position in chromosome.
6. Place new offspring in a new population
7. Replace: use new generated population for a further run of algorithm
8. Testing: if the end condition is satisfied, stop, and return the best solution in current population
9. Loop: Go to step 2

GA works on a population consisting of some solutions where the population size is the number of solutions. Each solution is called individual. Each individual solution has a chromosome. The chromosome is represented as a set of parameters (features) that defines the individual. Each chromosome has a set of genes. Each gene is represented by somehow such as being represented as a string of 0 and 1 ,the string which represent parameter solution is evaluated in terms of its fitness or its objective
Function which is often represent sum of square of residual.

The important part in this Algorithm is formulation fitness or objective function to be minimized, which will take the form:

$$\sum_{i=1}^{n}\left(y_i - e^{-\alpha\log(1+\theta x_i^2)}\right)^2 \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots.. (21)$$

Where:

$y_i = 1 - \frac{i}{n+1} \quad i = 1,2,\ldots,n$

And by applying the GA Algorithm function in MatLab program, that required the x and y values, fitness function, population size, crossover probability, mutation
Probability for minimization of equation (21) and number of generation, We will get the estimate of the parameters.

3.4 Pattern Search Method (PSM):[2,7,15]

Pattern search (PS) algorithm is one class of direct search evolutionary algorithms used to solve constrained optimization problems. While it was first formally proposed in early 1960, it has

popularity with users due to their simplicity and their practical success on a wide range of optimization problems.
This method do not require any information about the gradient of the objective function at hand, while searching for an optimum solution , so it is directional method that make use of a finite number of directions with appropriate descent properties. It is suitable for situation where the first and second derivatives of fitness or objective function are not exist, so that do not make explicit use of derivatives. It need only some values of objective function for some values of variable to run, therefore for this reason it named derivative free algorithm.
This algorithm calculates objective or fitness function values of the pattern and then try to find a minimizer. If it finds a new minimum, then it changes the center of pattern and iterates. This search continues until the search step gets sufficiently small, thus ensuring convergence to a local minimum.
The algorithm required:

1- Starting points or Initialization.
2- Value of acceleration factor.
3- Initial perturbation factor.
4- Perturbation tolerance factor.

The important part in this Algorithm is fitness function that take the form:

$$\sum_{i=1}^{n}\left(y_i - e^{-\alpha\log(1+\theta x_i^2)}\right)^2 \qquad \ldots\ldots\ldots\ldots\ldots (22)$$

Where:

$$y_i = 1 - \frac{i}{n+1} \quad , i = 1,2,\ldots,n$$

And by applying the PSM algorithm in MatLab , that required the x and y values and fitness function ,we will get the estimate of the parameters.

4 Simulation Experiments:

In order to compare among methods, a simulation experiments were carried. A range of sample sizes that represent small, moderate and high are used. A simulated data are generated according to invers cdf method as in formula (23):

$$x = (\theta)^{\frac{1}{2}}\left(e^{\frac{-1}{\alpha}\log(1-U)} - 1\right)^{\frac{1}{2}} \qquad \ldots\ldots (23)$$

Where
U: is a uniform random variant.
The GA parameters is set as: population size =600 , crossover probability=0.9 , mutation probability for minimization =0.01 and number of generations as 100. The terminated with accuracy level is equal to 0.001.
For different values of parameters that represent small and large range of values of parameters the results of mean square error of estimates are listed in tables (1) to table(6).

Table (1): Mean Square Error Values for $(\alpha = 1, \theta = 1)$

n	Parameters	MLE	NLLSM	GA	PSM	best
10	α	0.667988	0.056365	0.195805	0.056506	NLLSM
	θ	18.544	8.020789	8.193958	5.096595	p
15	α	0.649208	0.048823	0.127806	0.088396	NLLSM
	θ	16.39391	0.161808	0.49311	0.161498	NLLSM
25	α	0.595117	0.020371	0.125151	0.01738	p
	θ	0.008116	0.169578	1.656273	0.156868	MLE
50	α	0.44214	0.012686	0.118044	0.012252	NLLSM
	θ	12.22169	0.140998	1.434298	1.005031	NLLSM

100	α	**0.263562**	**0.012388**	**0.10213**	**0.011835**	**PSM**
	θ	**11.71317**	**0.148881**	**0.250246**	**0.131913**	**PSM**
150	α	**0.165628**	**0.011762**	**0.064163**	**0.010663**	**PSM**
	θ	**9.160459**	**0.126539**	**0.474766**	**0.244446**	**NLLSM**
200	α	**0.141976**	**0.011443**	**0.085911**	**0.011381**	**PSM**
	θ	**4.539029**	**0.118452**	**0.134299**	**0.109151**	**PSM**

Table (2): Mean Square Error Values for $(\alpha = 1, \theta = 2)$

n	parameters	MLE	NLLSM	GA	PSM	best
10	α	0.157305	0.194045	0.189704	0.148948	PSM
	θ	0.05785	14.27066	13.97386	7.980625	MLE
15	α	0.179088	0.094797	0.050026	0.071181	NLLSM
	θ	0.045906	7.769966	3.461991	5.29719	MLE
25	α	0.119167	0.017216	0.030205	0.014241	PSM
	θ	0.043018	3.082504	0.089634	2.797779	MLE
50	α	0.132248	0.014909	0.032651	0.011055	PSM
	θ	0.023253	1.534542	0.062939	1.443489	MLE
100	α	0.115316	0.01349	0.01979	0.078092	NLLSM
	θ	0.03802	0.761876	0.059276	0.675418	MLE
150	α	0.023624	0.012765	0.014229	0.012517	PSM
	θ	0.019973	0.151739	0.047618	0.147541	MLE
200	α	0.01028	0.011309	0.009544	0.00272	PSM
	θ	0.017866	0.168252	0.037297	0.145366	MLE

Table (3): Mean Square Error Values for $(\alpha = 1, \theta = 3)$

n	parameters	MLE	NLLSM	GA	PSM	best
10	α	0.377612	3.085701	13.22014	0.372275	PSM
	θ	2.056876	19.66302	16.64826	12.73178	MLE
15	α	0.365532	3.228973	1.721176	0.168565	PSM
	θ	2.363185	17.81436	12.375	11.27879	MLE
25	α	0.35708	2.052149	0.065441	0.034932	PSM
	θ	1.3637	4.172735	8.307877	3.67963	MLE
50	α	0.334262	1.674159	0.634387	0.071647	PSM
	θ	1.063835	3.691743	10.8264	3.452363	MLE
100	α	0.012047	0.104377	0.68987	0.094123	MLE
	θ	0.060599	0.658005	5.123502	0.58999	MLE
150	α	0.01251	0.006508	0.01276	0.006175	PSM
	θ	0.046318	0.044234	0.072274	0.040933	PSM
200	α	0.049992	0.003836	0.011342	0.003654	PSM
	θ	0.032683	0.035722	0.029568	0.034731	GA

Table (4): Mean Square Error Values for $(\alpha = 2, \theta = 1)$

n	parameters	MLE	NLLSM	GA	PSM	best
10	α	3.914358	9.866198	1.505866	0.711264	PSM
	θ	3.615759	0.618972	8.545414	1.248075	NLLSM
15	α	3.267255	0.192999	1.206978	0.093107	PSM
	θ	2.165781	0.313614	0.510557	1.214682	NLLSM

25	α	1.710191	0.778096	1.051513	0.054885	PSM
	θ	1.241398	0.105136	0.029204	1.591178	GA
50	α	0.49134	0.333224	0.040609	0.091582	GA
	θ	1.112081	0.938523	0.936618	2.354406	GA
100	α	0.536402	0.030253	0.017792	0.04454	GA
	θ	0.907246	0.450256	0.492739	0.490337	NLLSM
150	α	0.536402	0.030253	0.017792	0.04454	GA
	θ	0.907246	0.450256	0.492739	0.490337	NLLSM
200	α	0.002725	0.008118	0.005135	0.040473	MLE
	θ	0.059195	0.545362	0.006545	0.167839	GA

Table (5): Mean Square Error Values for $(\alpha = 3, \theta = 1)$

n	parameters	MLE	NLLSM	GA	PSM	best
10	α	33.47505	6.299863	30.2549	2.668031	PSM
	θ	10.42354	107.5047	73.35895	43.72544	MLE
15	α	5.224674	4.663007	29.33486	1.550078	PSM
	θ	10.94425	2.892324	4.251695	1.337027	PSM
25	α	6.14677	3.44457	17.20558	0.913376	PSM
	θ	6.248712	1.51389	9.35474	0.900419	PSM
50	α	2.320209	3.66052	4.682172	1.231651	PSM
	θ	4.888494	1.043221	9.924652	0.731079	PSM
100	α	1.466885	0.150727	0.391447	0.077133	PSM
	θ	0.012384	0.068221	0.079565	0.497864	MLE
150	α	1.944818	0.060009	0.209343	0.065627	NLLSM
	θ	0.01148	0.052933	0.06477	0.604872	MLE
200	α	0.922491	0.01921	0.494312	0.06126	NLLSM
	θ	0.011136	0.033581	0.010293	0.054012	GA

Table (6): Mean Square Error Values for $(\alpha = 3, \theta = 3)$

n	Parameters	MLE	NLLSM	GA	PSM	best
10	α	3.652754	73.15335	58.53988	1.74371	PSM
	θ	41.21292	352.9747	4.26432	11.16823	GA
15	α	3.332197	23.20017	4.481787	0.242188	PSM
	θ	24.97281	6.581023	2.744208	3.43716	GA
25	α	1.805887	1.223486	2.358017	0.252733	PSM
	θ	21.41936	6.795088	2.962349	3.270942	GA
50	α	1.46855	1.381749	2.765233	0.073792	PSM
	θ	16.02572	5.68269	1.060771	2.964387	GA
100	α	0.112281	1.296906	1.350315	0.087335	PSM
	θ	0.045603	1.389105	1.21282	0.387443	MLE
150	α	0.117031	1.469199	1.104577	0.096361	PSM
	θ	0.012876	1.446183	0.341098	0.081406	MLE
200	α	0.062984	1.454363	0.959521	0.029787	PSM
	θ	0.011285	0.862902	0.108012	0.053908	MLE

5 Discussion

The minimum mean square error of parameters for the four method are marked in last column for each case. It is shown that PSM method attained the first, since it reach minimum in 44% of cases. The MLE attained the second, since it get 25% of cases , NLLSM 17% and GA 14%.as illustrated in table(7).

Table (7): Mean Square Error Values for $(\alpha = 3, \theta = 1)$

method	counts	Percentage
MLE	21	25%
NLLSM	14	17%
GA	12	14%
PSM	37	44%
total	84	100%

6 Conclusion

This paper introduce new probability distribution that is mixed between Lomax distribution and Rayleigh distribution, we get closed form for the pdf and cdf, the since the theoretical mean square error was difficult to find for estimation method, so we used simulation experiments. We concluded that pattern search method is more efficient than the rest methods.

References

[1] 1-Adeleke, I. A, Akarawak, E. E. E, and Okafor, R. O. (2013). Investigating the Distribution of the Ratio of Independent Beta and Weibull Random Variables. Journal of Mathematics

[2] and Technology, 4(1): 16-22.

[3] 2-Al-Sumaita J. S. ,AL-Othmanb A.K. and Sykulskia J.K.(2007) Application of pattern search method to power system valve-point economic load dispatch International Journal of Electrical Power & Energy Systems Volume 29, Issue 10, December 2007, Pages 720-730

[4] doi.org/10.1016/j.ijepes.2007.06.016

[5] 3-Alzaatreh, A., Lee, C. and Famoye, F. (2013a). A New Method for Generating Families of

[6] Continuous Distributions.Metron:International Journal of Statistics, 71(1):63-79.

[7] Akinsete, A., Famoye, F. and Lee, C. (2008). The beta-Pareto distribution.Statistics, 42:547-

[8] 563.

[9] 5-Akarawak, E. E. E., Adeleke, I. A. and Okafor, R. O. (2013). The Weibull-Rayleigh Distribution

[10] and its Properties. Journal of Engineering Research, 18(1): 56-67.

[11] 6-Akarawak, E. E. E., Adeleke, I. A. and Okafor, R. O. (2015). On the Distribution of the Ratio of

[12] Independent Gamma and Rayleigh Random Variables. Journal of Scientific

[13] Research and Developments, 15(1): 54-63.

[14] 7-DennisJ. and Torczon V. (1994), Derivative-free pattern search methods for multidisciplinary design problems., paper AIAA-94-4349 in Proceedings of the 5th AIAA/ USAF/NASA/ISSMO Symposium on Multidisciplinary Analysis and Optimization,Panama City, FL, Sept. 7-9, pp. 922-932.

[15] 8-Eiben, Agoston E., and James E. Smith.(1992). Introduction to evolutionary computing. Vol. 53. Heidelberg: springer, 2003.John H. Holland 'Genetic Algorithms', Scientific

[16] American Journal, July.

[17] 9-Eugene, N., Lee, C. and Famoye, F. (2002). Beta normal distribution and its applications.Communication in Statistics- Theory and Methods, 31(4): 497 – 512.

[18] 10-Famoye, F., Lee, C. and Olugbenga, O. (2005).The Beta-Weibull distribution. Journal of

[19] Statistical Theory and Applications, 4(2):121-138.
[20] 11-Ghosh,I and Hamedani,G.(2015) .The Gamma-Kumaraswamy Distribution: An Alternative to Gamma Distribution Communications in Statistics – Theory and Methods
[21] DOI: 10.1080/03610926.2015.1122055
[22] 12-Gupta, R. D. and Kundu, D. (1999). Generalized Exponential Distribution. Australian and
[23] New Zealand Journal of Statistics, 41: 173– 188
[24] 13-Khaleel,M. A., Ibrahim N. A., Shitan,M.and Merovci,F.(2016). Some properties of Gamma Burr type X distribution with application AIP Conference Proceedings 1739, 020087
[25] doi: 10.1063/1.4952567
[26] 14- Miroslav ,R.M. and Balakrishnan,N.,(2012).The gamma-exponentiated exponential distribution. Journal of Statistical Computation and Simulation 82, 1191–1206
[27] 15-Michinari, MommaMichinari, MommaKristin P., BennettKristin P.and Bennett(2002)
[28] A Pattern Search Method for Model Selection of Support Vector Regression
[29] Conference: Proceedings of the Second SIAM International Conference on Data Mining, Arlington, VA, USA, April 11-13,DOI: 10.1137/1.9781611972726.16
[30] 16- Özsoy1 V. S., Ünsal M. G. and Örkcü H. H.,(2020) Use of the heuristic optimization in the parameter estimation of generalized gamma distribution: comparison of GA, DE, PSO and SA methods. Computational Statistics doi.org/10.1007/s00180-020-00966-4
[31] 17- Venegas O., Iriarte Y. A., Astorga J. M. and G´omez H.W.,(2019) Lomax-Rayleigh Distribution with an Application.Appl. Math. Inf. Sci. 13, No. 5, 741-748 (2019) 741
[32] http://dx.doi.org/10.18576/amis/130506

Spatial Analysis of Female Breast Cancer Incidence in Iraq during 2000-2015

M M AL-Hashimi[1] and E M Nouri[2]

[1]Department of Statistics and Informatics, College of Computer science & Mathematics, University of Mosul, Mosul, Iraq
[2]Department of Operations Research and Intelligent Techniques, College of Computer science & Mathematics, University of Mosul, Mosul
Email. muzahim63@gmail.com

Abstract: Breast cancer in females is the most common cancers diagnosed worldwide, and the leading cause of cancer death. This study explored the spatial distribution pattern of female BC in different districts in Iraq between 2000 and 2015. Data were obtained from the Iraqi Cancer Registry. The age standardized incidence rate (ASIR) were calculated according to provinces and geographical district for three periods (2000-2004, 2005-2009, 2010-2015). spatial statistical tools were applied to evaluate hotspots, coldspots, spatial clustering and outliers. Results showed a spatial correlation with hotspots, coldspots, and detecting spatial outlier. This study identified 10 districts as high-risk areas for BC, including AL-Kadhimiyah, Al-Karkh, Al-Adhamia, Al-Rissafa, Al-Sadir, and Abu-Graib in Baghdad province, Bakooba district in Diyala province, Al-Hindia district in Karbala province, and Shatt Al-Arab district in Basrah province, and we have evidenced an increase of breast cancer incidence rates during 2010-2015. More researches are needed to investigate the reasons for the geographic and temporal variations of breast cancer incidence in Iraq.

10. Introduction

Breast cancer in females is the most common cancers diagnosed worldwide, and the leading cause of cancer death. Globally, there were about 2.1 (11.6% of all sites) million newly diagnosed cases of breast cancer and about 0.6 (6.6% of all sites) million deaths in females in 2018 [1]. A 2018 GLOBOCAN report highlighted the large geographical diversity of major cancers observed in 185 countries [1]. The highest incidence rates were in Australia/New Zealand and Western Europe, and the lowest incidences were in middle Africa and south-central Asia, and the females in very high Human Development Index (HDI) countries have a higher breast cancer ASIR (75.2/100,000) than females in either low HDI countries (32.8/100,000) [1]. The substantial variations in breast cancer rates may reflect changes of risk factors associated with economic development and availability of early detection and timely treatment [2].

According to Iraqi Cancer Registry, breast cancer ranked first for both incidence and mortality. While the incidence of breast cancer among females in Iraq was relatively less than that in developed countries, we witnessed a substantially increasing incidence during the period 2010-2015.

The Iraqi Cancer Registry was established in 1974. It is located in the Institute of Radiology and Nuclear Medicine and supervised directly by the Iraqi Cancer Board. The information of every new cancer patient was collected from Baghdad province until 1989, then the registry services were extended to cover all provinces of Iraq. Cases registered were reported from inpatients and outpatients of government hospitals and cases attending the institute of Radiology and Nuclear Medicine. In addition, private hospitals, pathological and hematological laboratories requested to report to the registry all cases of cancer that come to their attention. The Iraqi Cancer Registry Annual Report provides information about cancer in Iraq including new cases of cancer and mortality [3].

This study was conducted in Iraq, a country in the south-west of Asia, which consists of 18 provinces according to Iraqi classification of administrative regions, three of which are located in the Kurdish region. It covers an area of 437,072 Km^2 [4] and the population according to the World Population Review was over 39 million. 'Figure 1' shows the location of Iraq.

Figure 1. The location of Iraq

Spatial analysis has been commonly used in health studies, such as epidemiology [5]. Spatial epidemiology is “the description and analysis of the geographical distribution of disease” [6]. Understanding the geographical distribution of disease in a population can provide important insight about the causes and controls of disease. Disease maps provide visual representations and effective tool to show huge amount of geographical information, it can be used to display patterns of diseases for defined geographical area [7].

The aim of this paper is to explore the spatial distribution pattern of female breast cancer in different provinces in Iraq during (2000-2015).

11. Materials and Methods

The incidences of female breast cancer (code C50- based on the International Classification of Diseases (ICD-10)) between 2000 and 2015 were obtained from Iraqi Cancer Registry. Annual female population by 5-year age groups, gender and provinces were obtained from the Ministry of Planning/Central Organization for Statistics. This data covers all provinces with 83 districts in the country, with the exception of three provinces in the Kurdish region (Erbil, Duhok and Al-Sulaymaniyah, for which the data is incomplete).

To evaluate the incidence of female breast cancer during the study period 2000-2015, we split the data according to geographical district into three periods (2000-2004, 2005-2009, and 2010-2015). The ASIR were calculated using the world standard population. Having obtained estimates ASIR, two spatial statistical tools were applied to evaluate spatial clustering and outliers, the global spatial autocorrelation (i.e., Moran’s I) and local indicators of spatial association (i.e., Anselin local Moran’s I and Getis-Ord Gi*). The descriptive analysis was carried out using SAS statistical software, version 9.4. As for visualizing the geographical district differences in the incidences, we constructed maps of the breast cancer using ArcMap 10.6.

The Moran’s Global statistic (Moran’s I) was used to estimate the spatial autocorrelation. This statistic measures the similarity of the overall area with respect to ASIR. The values of Moran’s I ranged from −1 to 1, where large and positive values indicate presence significant spatial autocorrelation $(p < 0.05)$. Low values of Moran’s I indicate presence significant regularity, while negative values suggest the clustering of dissimilar values.

The Global Moran's I do not suggest where clusters of low or high ASIR might appear. Thus, local tests for spatial autocorrelation such as Getis-Ord Gi* and Anselin Local Moran’s I are needed [8].

We used two local measures of spatial association within ArcGIS 10.6.1 to examine the clusters of geographical districts with high or low ASIR’s or outliers are located and in addition, what type of spatial correlation is more important [9]. The local measure of Moran's I allowed us to identify spatial clusters of districts and spatial local outliers that are different from the neighboring districts.

based on the Moran's I statistic and corresponding (p −value). A high positive Moran's I value (hotspot) refer to high ASIR and surrounded by a high ASIR (high-high), low positive Moran's I value (coldpot) refer to low ASIR and surrounded by a low ASIR (low-low). A negative Moran's I value (high-low) refer to high ASIR and surrounded by a low ASIR, low negative Moran's I value (low-high) refer to low ASIR and surrounded by a high ASIR. Additionally, local Getis-Ord statistic [10] was applied to examine spatial clustering. Getis-Ord statistic generates a z-value and associated (p −value) for each district, where high z −value ($z > 1.96$) indicate a significant hotspot and low z −value ($z < -1.96$) indicate a significant coldspot ($p < 0.05$).

12. Results

Between 2000 and 2015, 44496 female breast cancer cases were reported in Iraq, which accounted for 34% of the cancer cases among females, and it ranked as the first most common type of female cancer in Iraq. The number of newly diagnosed Iraqi female breast cancer increased from 1653 in 2000 to 4720 in 2015, which corresponds to a 2.86-fold increase. ASIR (all-age) for female breast cancer in Iraq for the period 2000 to 2015 was 32.807 per 100,000. The national ASIR increased from 22.0679 in 2000-2004 to 30.1152 in 2005-2009 and to 41.0729 per 100,000 female population in 2010-2015 'Figure 2'. The crude incidence rate (all-age) of breast cancer for the period 2000 to 2015 was 22.21 per 100,000. Crude incidence rates increased from 17.34 in 2000-2004 to 19.84 in 2005-2009 and to 26.92 per 100,000 female population in 2010-2015.

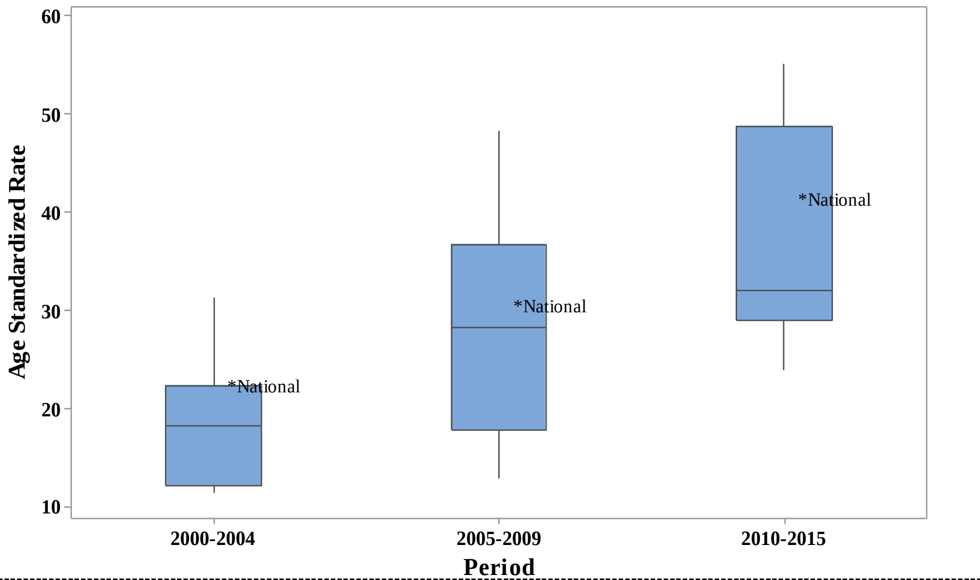


Figure 2. Age-Standardized incidence rate of breast cancer in Iraq over time

The geographical distribution of the breast cancer across provinces of Iraq is shown in 'Figure 3', 'Figure 4', and 'Figure 5'. The ASIR of breast cancer was 11.45-31.33 (31.33 per 100 000 females was derived in Baghdad) in 2000-2004, 12.82-48.42 (48.42 per 100 000 females was derived in Al-Najaf) in 2005-2009, and 23.90-55.273 (55.273 per 100 000 females was derived in Kirkuk) during the period 2010-2015. Baghdad and Kirkuk provinces stands out with the highest ASIR of BC for the three study periods. The provinces Al-Najaf and Al-Basrah had the highest ASIR for the period 2005-2009 along with Baghdad and Kirkuk provinces. However, the provinces Nineveh, Karbala, Babil, Al-Diwaniyah, Salah Al-Dean and Al-Muthana also considered high-risk areas. The provinces Al-Anbar and Thi-Qar had the lowest ASIR for the three study periods.

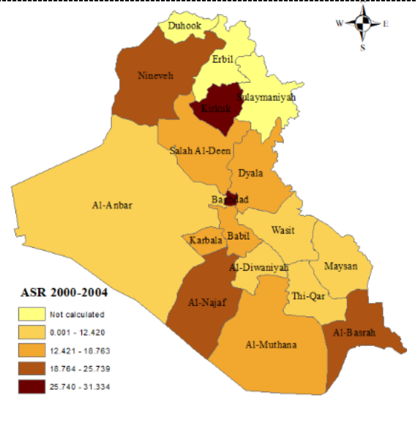


Figure 3. Age-standardized incidence (per 100,000 females at risk) by provinces during, 2000 to 2004

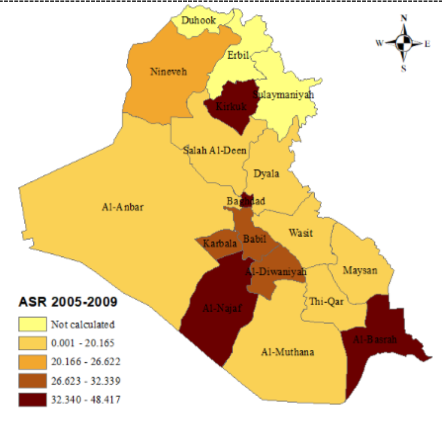


Figure 4. Age-standardized incidence (per 100,000 females at risk) by provinces during, 2005 to 2009

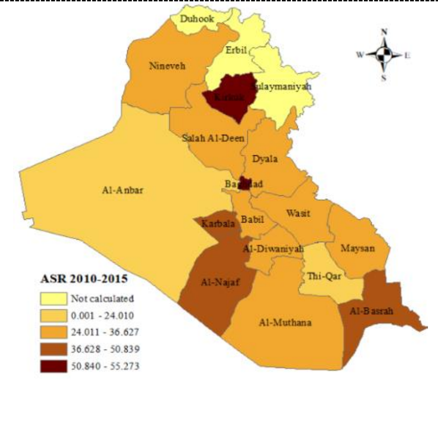


Figure 5. Age-standardized incidence (per 100,000 females at risk) by provinces during, 2010 to 2015

The geographical distribution of the breast cancer across districts of Iraq is shown in ‘Figure 6’, ‘Figure 7’, and ‘Figure 8’. The ASIR of breast cancer was 0.00-56.452 (56.452 per 100 000 females was derived in Al-Karkh district in Baghdad province) in 2000-2004, 0.00-65.3638 (65.3638 per 100 000 females was derived in Al-Najaf city in Al-Najaf province) in 2005-2009, and 1.50-99.3727 (99.3727 per 100 000 females was derived in Al-Karkh district in Baghdad province) in 2010-2015.

The global Moran’s I test statistics and the associated z-value of the ASIR in geographical districts were calculated for each of the three time periods. The results across the time periods were different, with insignificant positive spatial autocorrelation (Moran’s I index = $0.112\ (z = 1.886, P < 0.059)$) for the period 2000-2004, and significant positive spatial autocorrelation (Moran’s I index = $0.211(z = 3.358,\ P < 0.0078)$ for the period 2005-2009, and (Moran’s I index = $0.259\ (z = 4.120,\ P < 0.001)$) for the period 2010-2015. In other words, across Iraq, districts with similar ASIR of the breast cancer exhibits spatial clustering during the two periods 2005-2009 and 2010-2015.

The results of Anselin's Local Moran's I (cluster and outlier analysis) for the period 2005-2009 suggest the presence of a hotspots of breast cancer across districts, Shatt Al-Arab (Basrah province) in the south. Kuffa (Al-Najaf province), Al-Hindia (Karbala province), and Al-Sadir, Al-Rissafa and Al-Adhamia (Baghdad province) in the middle of Iraq ‘Figure 9’. These patterns were changed during the period 2010-2015, Shatt Al-Arab, Kuffa, and Al-Hindia, which were no longer a hotspot. Additionally, hotspots districts appeared in Bakooba, AL-Kadhimiyah, and Al-Karkh ‘Figure 10’. In 2005-2009, five districts were identified as a coldspots, Al-Mdainah (Basrah province), Al-Jabaish (Thi-Qar province) in the south, Balad (Salah-Al-Deen province) in the middle of Iraq, Rawa (Al-Anbar province) and Al-Baaj (Nineveh province) in the west of Iraq. These patterns remained the same during (2010-2015) with the exception of Rawa and Balad which were no longer a coldspots, and Al-Hattra (Nineveh province) appeared as a coldspot. There were three districts categorized as (high-low), included Al-Ammarah (Maysan province), Al-Nassiryah (Thi-Qar province) and Al-Samawa (Muthana province), where these districts showed high breast cancer incidence, but surrounded by low incidence districts ‘Figure 9’. This can be seen also over the period (2010-2015) with the exception of Al-Ammarah which no longer represent (high-low) type. In addition, (high-low) appeared in Mosul (North of Iraq). In 2005-2009, (low-high) district was located in Dooz (Salah Al-Deen province). In 2010-2015, the (low-high) districts were appeared in Al-Hindia, Al-Maddain and Abu-Graib with the exception of Dooz, which no longer identified as (low-high) type ‘Figure 10’.

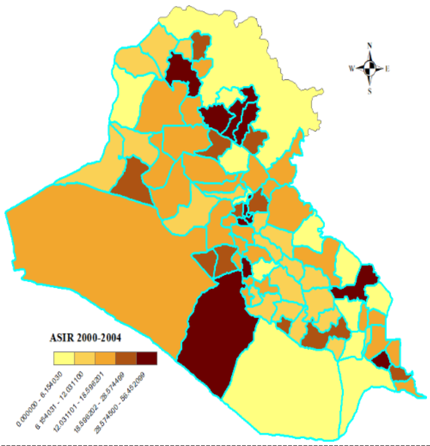


Figure 6. Age-standardized incidence (per 100,000 females at risk) by districts during, 2000 to 2004

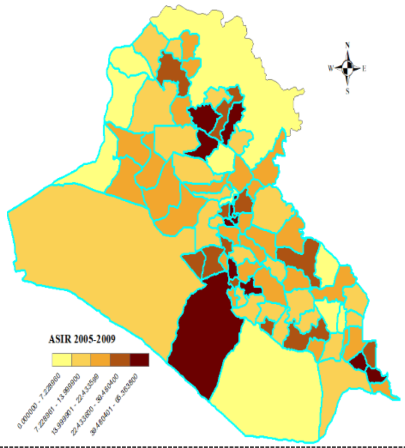


Figure 7. Age-standardized incidence (per 100,000 females at risk) by provinces during, 2005 to 2009

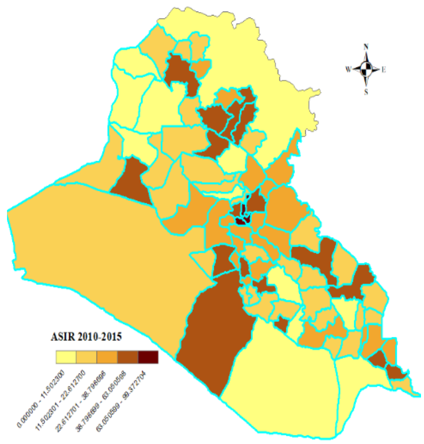


Figure 8. Age-standardized incidence (per 100,000 females at risk) by provinces during, 2010 to 2015

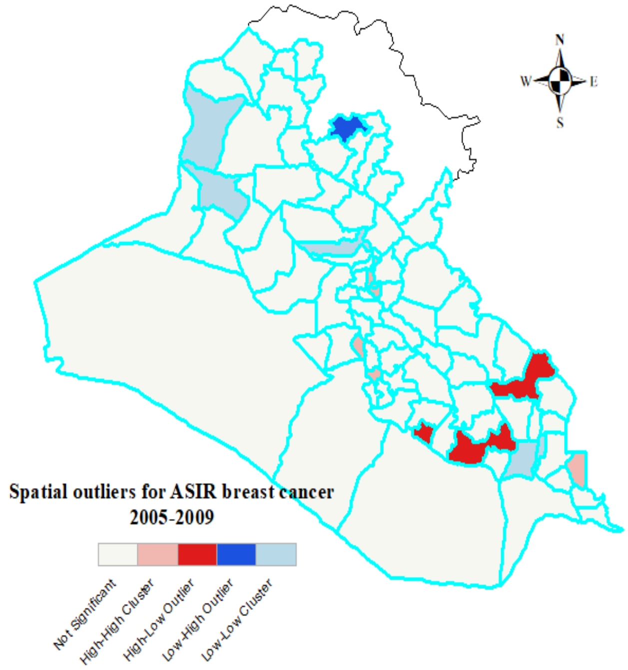


Figure 9. Hotspots and coldspots for female breast cancer incidence in Iraq, 2005–2009.

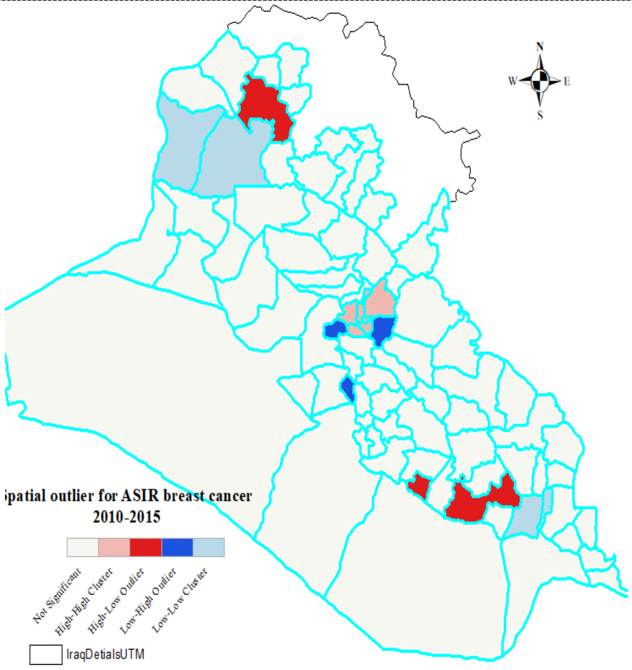


Figure 10. Clusters and spatial outliers for female breast cancer incidence in Iraq, 2010–2015.

The results of Getis-Ord Gi* can be seen in ‘Figure 11’ and ‘Figure 12’. During the period (2005-2009), 12 districts were identified as a hotspot, and 6 districts were identified a coldspot. During the period (2010-2015), 10 districts were identified as a hotspot, and 5 districts were identified as a coldspot.

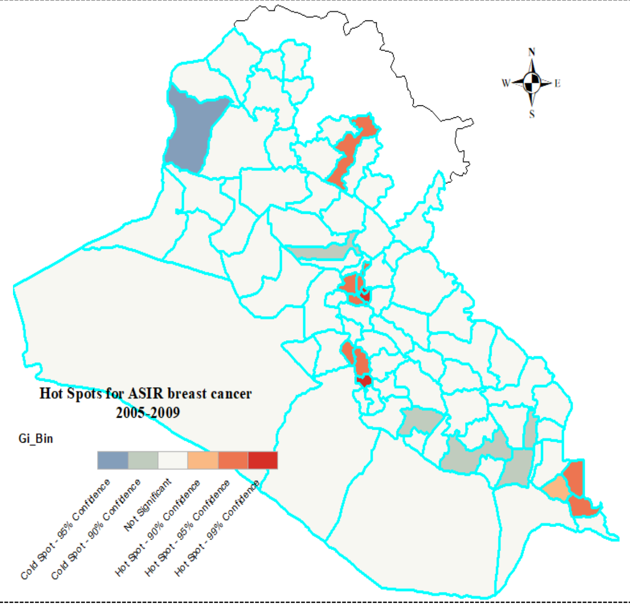


Figure 11. Hotspots and coldspots for female breast cancer incidence in Iraq, 2005–2009.

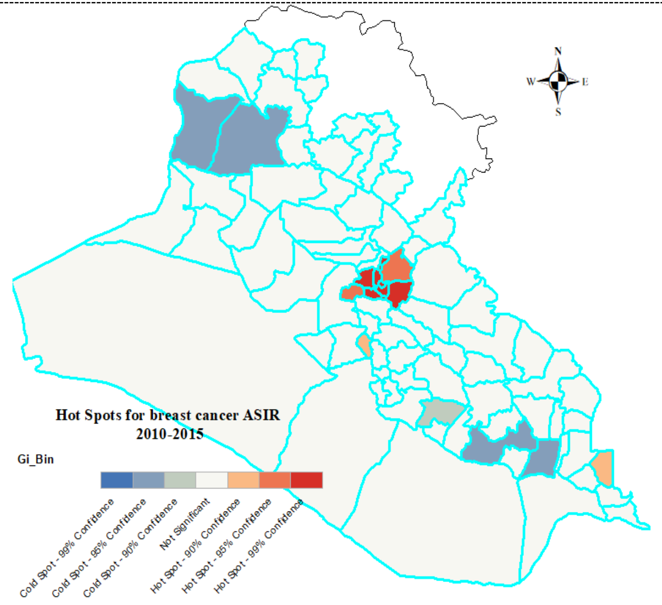


Figure 12. Hotspots and coldspots for female breast cancer incidence in Iraq, 2010–2015.

13. Discussion

This paper is the first to describe province and district-level variation in breast cancer incidence rates over the entire Iraq (except Kurdistan region), using spatial statistical tools to determine statistically significant spatial clusters of hotspots, coldspots, and spatial outliers in breast cancer incidence rates.

There is geographical variation in Iraq breast cancer incidences according to provinces and district-level. This study showed that the breast cancer by province has increased from 2000-2004 to 2010-2015, there was a clear trend in cancer incidence. The highest ASIR of breast cancer were observed in Baghdad and Kirkuk provinces in (2000-2004). In (2005-2009), the highest ASIR of breast cancer were observed in the Al-Najaf, Al-Basrah, Baghdad and Kirkuk provinces, while in (2010-2015), the highest ASIR of breast cancer were observed in Baghdad and Kirkuk Provinces. Highest incidence was observed in the Al-Karkh district in Baghdad province during the period (2010-2015).

A study in Kurdish region in Iraq (North of Iraq), reported that the (ASIR) of breast cancer in Kurdish women was (17.9/100000) during the period (2011-2013) [11], which is lower than the (ASIR) for all provinces bordering Kurdish area (Nineveh (32.2086/100000), Salah Al-Deen (27.7052/100000), Diyala (31.8138/100000), and Kirkuk (55.2729/100000) during the period (2010-2015). The (ASIR) of breast cancer in Sulaimaniyah women in Kurdish region was (36/100000) in 2012 [11]. This result higher than Salah Al-Deen and Diyala provinces, but lower than the (ASIR) in Kirkuk during the period (2010-2015).

A study in Basrah, reported that the ASIR of female breast cancer was 34.86 per 100000 during the period (2009-2012) [12], which is lower than the ASIR (38.1516 per 100 000 during the period (2005-2009), and 49.916 per 100 000 females during the period (2010-2015)) in the present study. Our finding is not consistent with the previous result. This inconsistency may be due to the difference in the study periods, or the difference between the methods of calculating (ASIR). This paper adopted world standard population to adjust age effects on a geographical scale to calculate (ASIR), while the previous study was not referred to the method of calculation.

The (ASIR) in Iraq during the period (2010-2015) was higher than that in Saudi Arabia (22.1/100000) [13], Oman (22.1/100000) [14], Iran (29.83/100000) [15]. It less than that in Lebanon (96.5/100000) [16], Jordan (45.7/100000) [17], Bahrain (47.4/100000) [18], Qatar (48·2 /100000) [19] and Kuwait (46.7 /100000) [20], but not lower than the ASIR in Basrah Province (49.9160 / 100

000 during the period 2000-2015), which is located in southern Iraq, bordering Kuwait. The (ASIR) in Iraq was lower than the ASIR in Korea (49.5/100000) [21], Eastern Europe (54.5), UK (93.6/100000), USA (84.9/100000) and Worldwide (46.3/100000) [1].
The results of global Moran's I indicated a spatial dependence of female breast cancer across Iraq districts. In other words, districts with high ASIR tend to cluster together. The results of Getis Ord Gi * identified the hotspots and coldspots across Iraq that revealed the existence of high-risk areas for breast cancer in AL-Kadhimiyah, Al-Karkh, Al-Adhamia, Al-Rissafa, Al-Sadir, and Abu-Graib district in Baghdad province, Bakooba district in Diyala province, Al-Hindia district in Karbala province, and Shatt Al-Arab district in Basrah province during 2010-2015.
However, it is beyond the aim of this paper to know the reasons of geographical variation of breast cancer incidence in Iraq. Recent studies indicate that exposure to environmental risk factors, such as certain chemicals during development stages of breast tissue prior to birth and until menopause, may increase risk of breast cancer later in life [22]. The substantial variations in breast cancer rates may reflect changes of risk factors associated with economic development and availability of early detection and timely treatment [2]. The variation in cancer incidence rates between the provinces in Iraq might be explained to environmental risk factors, such as chemicals exposures, and occupational exposures. Another reason of higher rates in some provinces might be that the oncology centers are located in these provinces and the cases occurring in the province were registered, while might be the other provinces do not register the cases completely. However, the reasons for the higher incidence rates in some districts merits further investigation.

14. Conclusion

In summary, there is a geographic variation in the ASIR of breast cancer across Iraq. districts with low and high breast cancer ASIR tend to cluster together. This study identified 10 districts as high-risk areas for BC, including AL-Kadhimiyah, Al-Karkh, Al-Adhamia, Al-Rissafa, Al-Sadir, and Abu-Graib in Baghdad province, Bakooba district in Diyala province, Al-Hindia district in Karbala province, and Shatt Al-Arab district in Basrah province, and we have evidenced an increase of breast cancer incidence rates during 2010-2015. More researches are needed to investigate the reasons for the geographic and temporal variations of breast cancer incidence in Iraq.

Acknowledgments

We would like to express our thanks to the staff of Iraqi Cancer Board-Ministry of Health for their effort for providing data access. We are grateful to university of Mosul, for supporting our work.

Conflict of interest

We declare that we have no conflicts of interest to disclose regarding this manuscript.

Funding This manuscript received no specific grant from any funding agency in the public or commercial sectors.

15. References

[1] Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A 2018 GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries *Ca Cancer J Clin*. **68** 394-424.
[2] DeSantis CE, Bray F, Ferlay J, Lortet-Tieulent J, Anderson BO, Jemal A 2015 International variation in female breast cancer incidence and mortality rates *Cancer Epidemiology and Prevention Biomarkers*. **24** 1495-506.
[3] Iraqi Cancer Registry *Annual Report 2018*.
[4] Ibp I 2012 Iraq Country Study Guide Volume 1 Strategic Information and Developments.

[5] Su SC, Kanarek N, Fox MG, Guseynova A, Crow S, Piantadosi S 2010 Spatial analyses identify the geographic source of patients at a National Cancer Institute Comprehensive Cancer Center *Clinical Cancer Research*. **16** 1065-72.
[6] Lawson AB 2013 *Statistical methods in spatial epidemiology*. John Wiley & Sons.
[7] Boscoe FP, Ward MH, Reynolds P 2004 Current practices in spatial analysis of cancer data: data characteristics and data sources for geographic studies of cancer *International journal of health geographics*. **3** 28.
[8] Khan D, Rossen LM, Hamilton BE, He Y, Wei R, Dienes E 2017 Hot spots, cluster detection and spatial outlier analysis of teen birth rates in the US, 2003–2012 *Spatial and spatio- temporal epidemiology*. **21** 67-75.
[9] Anselin L, Sridharan S, Gholston S 2007 Using exploratory spatial data analysis to leverage social indicator databases: the discovery of interesting patterns *Social Indicators Research*. **82** 287-309.
[10] Getis A, ORD J 1992 The analysis of spatial association by use of distance statistics Geographical Analysis **24** 189–206.
[11] Karim SA, Ghalib HH, Mohammed SA, Fattah FH 2015 The incidence, age at diagnosis of breast cancer in the Iraqi Kurdish population and comparison to some other countries of Middle-East and West *International Journal of Surgery*. **13** 71-5.
[12] Nasr N 2016 Breast cancer in Basrah governorate: Pattern of geographical distribution Iraq *Journal of Community Medicine*. **29** 1-4.
[13] Baslaim M, Baroum IH, Salman BA, Baghlaf BS, Al-Farsi MA 2016 Breast Cancer Screening Program in Jeddah, Saudi Arabia: Is There a Need for a National Program *Int J Womens Health Wellness*. **2** 042.
[14] Mehdi I, Monem EA, Al Bahrani BJ, Al Kharusi S, Nada AM, Al Lawati J, Al Lawati N 2014 Age at diagnosis of female breast cancer in Oman: Issues and implications *South Asian journal of cancer*. **2** 101.
[15] Ahmadi A, Ramazani R, Rezagholi T, Yavari P 2018 Incidence pattern and spatial analysis of breast cancer in Iranian women: Geographical Information System Applications. *Eastern Mediterranean Health Journal*. 24 360-7.
[16] Fares MY, Salhab HA, Khachfe HH, Khachfe HM 2019 Breast cancer epidemiology among Lebanese women: an 11-year analysis *Medicina*. **55** 463.
[17] Abdel-Razeq H, Attiga F, Mansour A 2015 Cancer care in Jordan. *Hematology/oncology and stem cell therapy*. **8** 64-70.
[18] Al Awadhi MA, Abulfateh NM, Abu-Hassan F, Fikree MA, Janahi E, Carlo R 2016 Cancer incidence and mortality in the Kingdom of Bahrain statistics and trends. *Bahrain Medical Bulletin*. **158** 1-5.
[19] Chouchane L, Boussen H, Sastry KS Breast cancer in Arab populations: molecular characteristics and disease management implications *The lancet oncology*. **14** 417-24.
[20] Al Ramadhan MA 2017 Eradicating breast cancer: longevity impact on Kuwaiti women *Asian Pacific journal of cancer prevention* **18** 803.
[21] Won YJ, Jung KW, Oh CM, Park EH, Kong HJ, Lee DH, Lee KH 2018 Geographical Variations and Trends in Major Cancer Incidences throughout Korea during 1999-2013 *Cancer research and treatment: official journal of Korean Cancer Association*. **50** 1281.
[21] National Institute of Environmental Health Sciences 2012 *Breast Cancer Risk and Environmental Factors*.

Quality testing algorithms reduce overall completion time and overall delay in machine scheduling

Fathalh A. Cheachan

Department of Statistic,College of Economics and Admistration
University of Mustansiriyah,Baghdad,
aiiiii_f@yahoo.com

Abstract :In this paper we shed light on the issue of scheduling a single machine for foggy delay time and foggy work time. For the purpose of reducing the value of maximizing blur-delay function. A comparison was made and tested between several local methods ((TA), (TS), (GA), (ACO) and (MA)). The results of the selection reached 1500 works. Through the results, it was found that ACO gives the best approximate solutions .

1 Introduction

The solutions to most of the problems that a person faces in life are inaccurate or uncertain, and that uncertainty results from a shortage of data collected from a specific problem and this data is vague, inaccurate or incomplete, which leads to a lack of information, as it was used in The beginning is mathematical laws to address uncertainty, and then statistical methods have been used, and then mathematical laws and statistical methods have been combined, and this combination has led to the birth of statistical theory, which is a real goal for addressing uncertainty, but the potential theory based on the usual set does not address every h No uncertainty, and therefore two methods were used (the first is the honesty function, and the second is fuzzy logic) in dealing with uncertainties that are based on the theory of fuzzy group. The purpose of this paper is to introduce the fuzziness of scheduling information into the classical single machine scheduling problem and to propose mathematical model using the modified S-curve affiliation function to solve the problem of single machine scheduling . In a real life situation, it can often be observed that some kinds of information, such as due dates, processing time, and technological constraints, are not necessarily deterministic (or crisp). Recently, some scheduling models with fuzzy due date and fuzzy processing time have been studied. In this paper we are interested in the direct generalization of the traditional objective function measure with the due date and completion time being fuzzy numbers with using the modified S-curve.

2 Preliminaries

In this section we specify the context of this study and recall basic definitions that will be used later. We also present the principles underlying the main approaches for defining fuzzy distances.

i-Support Set

It is a regular subset of a comprehensive set (X) whose elements have a degree of affiliation greater than zero.

$$S(\widetilde{A}) = \{ M_{\widetilde{A}}(x) > 0 \ , \ x \in X \}$$

ii-Normal Fuzzy Set

It is a subset of a comprehensive set (X) containing an element whose affiliation is equal to one.

$$S(\tilde{A}) = \{ M_{\tilde{A}}(x) = 1 \ , \ x \in X \}$$

iii-Crossover Point

It is an element (xi) in the fuzzy set ($\tilde{A}$), its affiliation score is 0.5

$$M_{\tilde{A}}(x_i) = 0.5$$

iv- Height of Fuzzy Set

It is the highest value of an affiliation function that element (x) has in the fuzzy set ($\tilde{A}$)

$$Hgt(\tilde{A}) = \underset{x}{Sup} \ M_{\tilde{A}}(x)$$

v- **□-cut Set**

t is a regular set containing elements of the universal set (X) whose affiliation ranks in A is greater or equal to certain values of □

$$A_{\alpha} = \{ M_{\tilde{A}}(x) \geq \alpha \ , \ x \in X \}$$

There are special types of pieces, including □- **cut :**

a- strong □-cut

It is a regular set (A_{α})containing elements of the universal set (X) whose affiliation ranks are greater than certain values of □

$$A_{\alpha} = \{ M_{\tilde{A}}(x) > \alpha \ , \ x \in X \}$$

b- □- Cut =١

It is a regular group that contains elements whose ranks $(\tilde{A})$ are equal to one

$$A_1 = \{ M_{\tilde{A}}(x) = 1 \ , \ x \in X \}$$

3-Curve function S

It is a non-linear logistic function written as

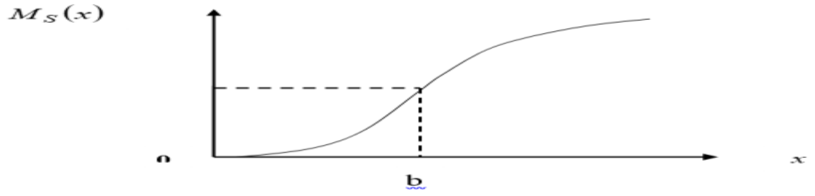
$$M_S : X \rightarrow [0,1]$$

$$M_S(x) = \frac{1}{1 + e^{-a(x-b)}}$$

Where

b: - the point of the inversion

A: - Curved inclination



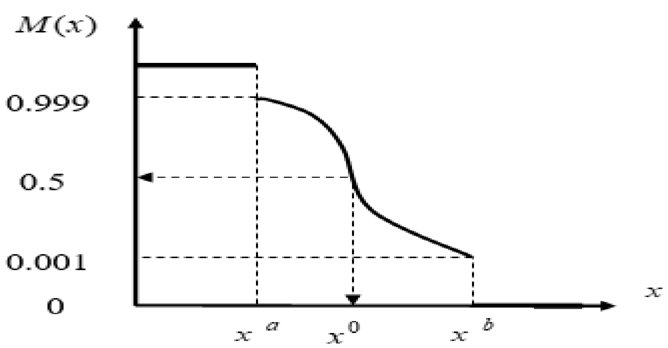
(1) Figure

S-curve function

3-modified S -curve function

The modified S-curve affiliation function is a special case of the logistic function at specific values for whose value does not range and the reason for this is that in the production system the required work capacity is not always 100% and at the same time it is not 0% and :can be expressed in the following formula

$$M(x) = \begin{cases} 1 & x_j < x_j^l \\ 0.999 & x_j = x_j^l \\ \dfrac{w}{1 + u e^{\alpha\left[\frac{x_j - x_j^l}{x_j^u - x_j^l}\right]}} & \\ 0.001 & x_j = x_j^u \\ 0 & x_j > x_j^u \end{cases}$$



(٢) Figure

Adjusted S-curve affiliation function

And to find the values of the constants of the function

Imposing it $x^a = 0$, $x^b = 1$ and so on

$$0.999 = \frac{w}{1 + u e^{\alpha(0)}}$$

$w = 0.999(1+u)$ .... (1)

$$0.001 = \frac{w}{1 + u e^{\alpha}}$$

In the above equation, we substitute for w by the equivalent in the equation(1)

$$0.001 = \frac{0.999\,(1+u)}{1 + u e^{\alpha}}$$

Taking the logarithm of the two sides, we simply get

$$\alpha = Ln \frac{1}{0.001}\left(\frac{0.998}{u} + 0.999\right) \qquad \text{...........(2)}$$

According to my properties of the logistic function

$$M\left(x^0\right) = 0.5 \qquad , \qquad x^0 = \frac{x^a + x^b}{2}$$

$$\frac{w}{1 + u e^{\frac{\alpha}{2}}} = 0.5$$

And by taking a logarithm of the two sides, we simply get

$$\alpha = 2Ln\left(\frac{2w - 1}{u}\right) \qquad \text{.............(3)}$$

And by substituting for (w) with its equal in equation (1) and with its equivalent in the equation(2) we obtain

$$2Ln\left(\frac{2(0.999)(1+u) - 1}{u}\right) = Ln \frac{1}{0.001}\left(\frac{0.998}{u} + 0.999\right)$$

$$(0.998 + 1.998\,u)^2 = u\,(998 + 999\,u)$$

:By solving the above equation we obtain

u =0.001001001

And when we substitute it in an equation, we get
w =1

α=13.81350

4 Problem Formulation

Assume that there are n free employments to be handled on a solitary machine. Each activity $j = 1,2, \dots, n$ requires fuzzy handling time p_j and fuzzy due date d_j which are an Adjusted S-curve fuzzy number. The machine can process all things considered employment at once, . The membership functions $\tilde{P}_j(x)$ and $\tilde{D}_j(x)$ are characterized as far as two numbers $[p_j^l, p_j^u]$ and $[d_j^l, d_j^u]$ as follows**:**

$$\tilde{p}_j = \begin{cases} 1 & p_j < p_j^l \\ 0.999 & p_j = p_j^l \\ \dfrac{w}{1 + ue^{\alpha\left[\frac{p_j - p_j^l}{p_j^u - p_j^l}\right]}} & \\ 0.001 & p_j = p_j^u \\ 0 & p_j > p_j^u \end{cases}$$

and

$$\tilde{D}_j = \begin{cases} 1 & d_j < d_j^l \\ 0.999 & d_j = d_j^l \\ \dfrac{w}{1 + ue^{\alpha\left[\frac{d_j - d_j^l}{d_j^u - d_j^l}\right]}} & \\ 0.001 & d_j = d_j^u \\ 0 & d_j > d_j^u \end{cases}$$

The possible range of the fuzzy processing time $\tilde{P}_j$ and due date $\tilde{D}_j$ are $[p_j^l, p_j^u]$ and $[d_j^l, d_j^u]$,Assuming the processing time p_j and due date d_j are crispy numbers then the cost function we are interested to study has the following form $\mathbf{L_{max}}$. If the processing time and due date are a fuzzy numbers, then $\tilde{L}_{max}$ is a hazy function, we denote the problem formulated in this sections as $1|\tilde{P}_j = \text{Adjusted S} - \text{curve}, \tilde{D}_j = \text{Adjusted S} - \text{curve}|\tilde{L}_{max}(\tilde{P}_j, \tilde{D}_j)$

5 Model Development

In this paper, the belonging functions for delay time and work time were formed and then the values of fuzzy coefficients were calculated, where one of the used logistic affiliation functions known as the modified S-curve affiliation function that depends on two important factors, the first factor is the level of acceptance (Mu) whose value ranges between (0.001 - 0.999), and the second factor is the blur factor (α) whose value has been determined between (41 - 1) Then calculate the delay and working time coefficients relative to Mu and the blurring factor

To extract fuzzy due date

$$M_{\tilde{D}_j} = \frac{w}{1 + ue^{\alpha\left[\frac{d_j - d_j^l}{d_j^u - d_j^l}\right]}}$$

$$e^{\alpha\left[\frac{d_j - d_j^l}{d_j^u - d_j^l}\right]} = \frac{1}{u}\left[\frac{w}{M_{d_j}} - 1\right]$$

$$\alpha\left[\frac{d_j - d_j^l}{d_j^u - d_j^l}\right] = ln\frac{1}{u}\left[\frac{w}{M_{d_j}} - 1\right]$$

$$\tilde{d}_j = d_j^l + \left[\frac{d_j^u - d_j^l}{\alpha}\right] ln\frac{1}{u}\left[\frac{w}{M_{d_j}} - 1\right]$$

To extrac fuzzy processing time

$$M_{\tilde{P}_j} = \frac{w}{1 + ue^{\alpha\left[\frac{p_j - p_j^l}{p_j^u - p_j^l}\right]}}$$

$$e^{\alpha\left[\frac{p_j - p_j^l}{p_j^u - p_j^l}\right]} = \frac{1}{u}\left[\frac{w}{M_{p_j}} - 1\right]$$

$$\alpha\left[\frac{p_j - p_j^l}{p_j^u - p_j^l}\right] = ln\frac{1}{u}\left[\frac{w}{M_{p_j}} - 1\right]$$

$$\tilde{p}_j = p_j^l + \left[\frac{p_j^u - p_j^l}{\alpha}\right] ln\frac{1}{u}\left[\frac{w}{M_{p_j}} - 1\right]$$

6 Methodology

6.1 Local Search Techniques

Right now study nearby inquiry procedures which are valuable instruments for taking care $1|\tilde{P}_j = TFN, \tilde{D}_j = TFN|\tilde{L}_{max}(\tilde{C}_j, \tilde{D}_j)$ of single machine planning issue

Nearby pursuit is an iterative calculation that moves starting with one solution s then onto the next s' as per some local structure

6.2 Tabu search

The utilization of the tabu hunt was spearheaded by Glover who from 1985 onwards has distributed numerous articles talking about its various applications. Others rushed to embrace the system which has been utilized for such purposes as sequencing, planning, oil investigation and steering. The properties of the tabu pursuit can be utilized to improve other system by forestalling them getting stuck in the areas of neighborhood minima. The tabu pursuit uses memory to keep the inquiry from coming back to a formerly investigated area of the arrangement space too rapidly. This is accomplished by holding a rundown of potential arrangements that have been recently experienced. These arrangements are viewed as tabu-henceforth the name of the method. The size of the tabu rundown is one of the parameters of the tabu inquiry.

The tabu quest likewise contains instrument for controlling the inquiry. The tabu rundown guarantees that some arrangement will be unsuitable; notwithstanding, the limitation gave by the tabu rundown may turn out to be excessively constraining at times making the calculation become caught at a locally ideal arrangement. The tabu hunt presents the thought of desire criteria so as to defeat this issue. The goal criteria supersede the tabu limitations making it conceivable to widen the quest for the worldwide ideal.

An underlying arrangement is created (generally haphazardly). The tabu rundown is introduced with the underlying arrangement. Various cycles are performed which endeavor to refresh the present arrangement with a superior one, subject to the limitation of the tabu rundown. A rundown of applicant arrangement is proposed in each emphasis. The most permissible arrangement is chosen from the up-and-comer list. The present arrangement is refreshed with the most permissible one and the new present arrangements added to the tabu rundown. The calculation stops after a fixed number of emphasess or when a superior arrangement has been found for various cycles. Figure 4 shows the nonexclusive usage of tabu hunt.

```
S = Generate Initial Solution()
Initialize Tabu List (TL1, ..., TLr)
K = 0
While (termination condition in not satisfied) do
        Allowed Set (S,K) = {z ∈ N(s)| no tabu condition is violated or at least
        one Aspiration criterion is satisfied}
        S = Best Improvement (S, Allowed Set(S,K))
        Update Tabu List and Aspiration Condition()
        K = K+1
End while
```

Figure 4: A generic tabu search

6.3 Memetic algorithm

The memetic calculations can be seen as a marriage between a populace based worldwide method and a neighborhood search made by every one of the people. They are a unique sort of hereditary calculation with a nearby slope climbing. Like hereditary calculation, memetic Algorithms are a populace based methodology. They have demonstrated that they are requests of extent quicker than conventional hereditary calculation for some issue areas. In a memetic calculation the populace is instated aimlessly or utilizing a heuristic. At that point, every individual makes nearby pursuit to improve its wellness. To shape another populace for the people to come, more excellent people are chosen. The choice stage is indistinguishable illuminate to that utilized in the traditional tabu inquiry choice stage. When two guardians have been chosen, their chromosomes are joined and the traditional administrators of hybrid are applied to produce new people. The last are improved utilizing a neighborhood search strategy. The job of nearby hunt in memetic calculations is to find the neighborhood ideal all the more effectively then the tabu pursuit. Figure 5 clarifies the conventional execution of memetic calculation.

```
Encode solution space
    Set pop_size, max_gen, gen=0
    set cross_rate, mutate_rate;
    initialize population
    while(gen < gensize)
        Apply generic GA
        Apply local search
    end while
    Apply final local search to best chromosome
```

Figure 5: The memetic algorithm

6.4 Ant Colony Optimization

Ant Colony Optimization (ACO) has been presented by Dorigo and associates as another streamlining worldview which is roused by the path following conduct of genuine subterranean insect settlements. Algorithmic usage of this metaheuristic have demonstrated exceptionally encouraging outcomes for the notable Traveling Salesman Problem [4] and are as of now among the best accessible calculations for other difficult issues like the Quadratic Assignment Problem, the Sequential Ordering Problem, Vehicle Routing Problems, Flowshop Scheduling Problem , and directing issues in profoundly powerful situations .

```
Procedure ACO metaheuristic for static combinatorial problems
    Set parameters, initialize pheromone trails
    While (termination condition in not satisfied) do
        Construct Solutions
        Apply Local Search % optional
        Update Trails
    End while
```

Figure 6: Algorithmic skeleton for the ACO.

In ACO calculations a province of (counterfeit) ants iteratively develops answers for the issue viable utilizing (fake) pheromone trails which are related with properly characterized arrangement parts and heuristic data. The ants just impart in a roundabout way by altering the pheromone trails during the calculation's execution. Since the developed arrangements need not be locally ideal regarding little changes, in a considerable lot of the best performing ACO calculations the ants moreover improve their answers by applying a nearby hunt calculation. Consequently, most ACO calculations for static combinatorial advancement issues follow the specific algorithmic plan given in Figure 6.

6.5 Threshold acceptance method (TH)

A variation of reproduced toughening is the limit acknowledgment strategy . It contrasts from mimicked toughening just by the acknowledgment rule for the arbitrarily produced arrangement $s' \in N.\ s'$. is acknowledged whether the thing that matters$F(s') - F(s)$ is littler than some nonnegative limit t. t is a positive control parameter which is bit by bit decreased. Figure 7 shows the conventional execution of Threshold acknowledgment structure.

```
While (termination condition in not satisfied) do
        New solution ← neighbors(best solution);
        If new solution is better than actual solution then
                Best solution ← actual solution
        Elseif difference between old and new solution less than control
                parameter t then
                Best solution ← actual solution
        End if
End while
```

Figure 7: Threshold acceptance structure

The limit acknowledgment strategy has the bit of leeway that they can leave a neighborhood least. They have the drawback that it is conceivable to return to arrangements previously visited. Hence swaying around neighborhood minima is conceivable and this may prompt a circumstance where much computational time is spent on a little piece of the arrangement set.

6. Computational results

Local search methods were tested by coding then in Matlab R2010b and runs on a Pentium IV at 2.00GHz, 2.92GB computer. The tested problem instances are generated as follows:

For n = 10, 20, 30, 50, 100, 200, 500 , 1000 & 1500 and integer p_j for $j \in N = \{1,2,...,n\}$ is generated by randomly selecting integers from interval [1,10] in our experiments, problem instances of 5-10 jobs were randomly. The fuzzy processing times were generated uniformly with the support in the range [10, 30].

In the following table (1) show the efficiency local search heuristic methods (Threshold accepted (TA), Tabu search (TS), Ant colony optimization (ACO) and Memetic algorithm (MA)) have been approached in terms of comparable rate of value. ACO gives the best solution for all iterations but it cannot calculate the 1000 jobs because it took very long time, therefore for 1500 jobs MA took good solution.

Table (1) Compares of local search methods

	ACO	MA	TS	TH
10	219.562	219.9	219.577	219.575
20	421.887	431.4	427.679	413.77
30	727.45	775.15	747.2	756.336
50	2533.17	2941.9	2900.15	2937.3
100	20173.02	25582.2	25610.98	14845.43
200	41553.9	50254.9	61361.30	60623.30
500	270079.9	370048.9	385831.8	379110.0
1000	*****	293879	2018637	2016663
1500	*****	303888	3118632	3038719

In the following table (2) show the efficiency local search heuristic methods ((TA), (TS), (ACO) and (MA)) have been approached in terms of comparable rate of times. TH gives the best times for all jobs.

Table (2) Compares times of local search methods

	ACO	MA	TS	TH
10	2.24120	1.224046	1.170383	1.128396
20	3.043797	1.270907	1.194606	1.132170
30	7.300873	1.332569	1.226729	1.128742
50	12.40130	1.445514	1.329673	1.128829
100	47.94082	1.746360	1.881342	1.120645
200	243.9809	2.430228	3.125784	1.133338
500	3227.284	3.663263	6.437874	1.146149
1000	*****	23.46840	21.99403	1.166653
1500	*****	24.57851	22.00514	2.255542

7. Concluding remarks

We have built up another model to define the circumstance where occupations with fluffy due dates are to be planned on a solitary machine. The neighborhood search strategies used to take care of all the huge issues the outcome show the vigor and adaptability of nearby pursuit heuristics.

The theoretical study showed the forms of affiliation functions (linear and non-linear) characterizing the modified S-curve affiliation function by using the blurring factor in constructing or forming affiliation functions for fuzzy coefficients when resolving scheduling problems due to the high flexibility of this function in describing uncertainty in transactions Fuzzyness of scheduling problem, as this function gives decision makers a number of decisions that can be taken equal to the number of fuzzy factor values multiplied by the number of levels of acceptance degree and on this basis there is an opportunity for decision makers to make their decisions based on the degree of blurring in the parameters of the studied model and

8. Future work:

١- In this research, a single target function was used, but scheduling can be applied Blur instead of a modified S-curve belonging to problems with more than one goal.

٢- Scheduling can be applied with other non-linear belonging functions as a function Exponential affiliation to the same problem and a comparison of its results with the results obtained From using scheduling with a logistics affiliation

9. References:

[1] Brucker P. (2007), *Scheduling Algorithms*, Springer Berlin Heidelberg New York, Fifth Edition.

[2] Chyu, C.-C., & Chang, W.-S. (2011). Optimizing fuzzy makespan and tardiness for unrelated parallel machine scheduling with archived metaheuristics. The International Journal of Advanced Manufacturing Technology, 57(5), 763. doi:10.1007/s00170-011-3317-3

[3] Han, S., Ishii, H., & Fujii, S. (1994). One machine scheduling problem with fuzzy due dates. European Journal of Operational Research, 79(1), 1-12. doi:http://dx.doi.org/10.1016/0377-2217(94)90391-3

[4] Fattahi, P., Hajipour, V., & Nobari, A. (2015). A bi-objective continuous review inventory control model: Pareto-based meta-heuristic algorithms. Applied Soft Computing, 32, 211-223

[5] Pinedo, M. L. (2016). Scheduling: Theory, Algorithms, and Systems. New York, USA: Prentice Hall

Estimation Parameters Of The Multiple Regression Using Baysian Approach Based On The Normal Conjugate Function

Adel Abbood Najm[1] **Ahmed Hamza Abood[2]** **Shrook A.S. AL-Sabbah[3]**

adel.abood6969@gmail.com **ahmed.eabod@uokerbala.edu.iq** **shorouq.a@uokerbala.edu.iq**

(1) College of Administration and Economics / University of Sumer/ Iraq .

(2) Department of Anatomy, Histology and Embryology/Factuality of Medicine/Karbala University/Iraq.

(3)Statistics Department / Administration And Economics College / Kerbala University / Iraq .

Abstract: In this paper, we have been used Bayes Technique depending on the normal conjugate function to estimate parameters of the multiple regression model, and we have been tested significance of this model. The test showed in the application that the mean square error (MSE) for the used model was decreasing, also it showed that the determinant coefficient is increasing highly. In the same time, value of the computed F-test was significant, according to the above, we can consider that the model is significant.

Keywords: Bayes Approach, Conjugate Probability densty function, multiple regression, Parameters estimation prior and posterior distribution, Normal dist.

1- Introduction: The regression analysis is considered one of the most important statistical techniqes which are used by the researchers to analyze the data in their fields such as the industry, biology, social, production, etc. for the sake of reaching to the best results, and this issue is done by forming a correct formula for the relationship between the different phenomena, which are represented by the variables, and these variables are subjected to the regression formula in the its different forms. The regression model formulas is very useful in case of knowing direction of the explanatory variables which are dealing with it by the researcher, also knowing the effect range which is showed by these variables on the response variables, besides the interpretation ratio of the regression model contribution in explaining the relationship between the response variable and the explanatory variables and all of that is done by process of estimation parameters of the model.

Our goal in this paper is to estimate the multiple linear regression model by using Baysian approach based on the normal conjugate prior function, and then showing the model significance using F-test and we have been showed that by using an applicable real data.

3- Concept of the regression analysis Bayes Approach:

Regression analysis is dealing with studing and estimation a phenomenon by quantity way through collection and analysis the data and determination the relationship figure between these data, where the estimation and prediction for future for this phenomenon are achieved according to certain statistical methods after getting an equation or a curve explain the mathematical relationship between the response variables Y and an explanatory variable X_i and this so- called the simple linear regression.

If the relationship between the response (dependent) variable Y with several explanatory variables $(X_1, X_2, \dots, X_n)$, in this case we call it multiple linear regression, and we can express the relationship between the dependent variable and one variable or more than one variable mathematicaliy by the following regression equation: $Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_n X_{in} \dots \dots (1)$

Note that eq.(1) shows that the dependent variable Y is explained by several independent (explanatory) variables, but if we want to explain this relationship correctly, it must be adding the error term to eq.(1), where this term represents or denotes to the information which the model doesn't involve it:

$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_n X_{in} + U_i \ldots\ldots (2)$

an easy way for representing the multiple regression model in (2) is to re-write it using the matrices form and as follows: $Y = X\beta + U \ldots\ldots (3)$

Where Y: is a vector of $n \times 1$ for the observations of response variable.

β: is the parameters vector of $n \times p$ (which must be estimeled).

X: a matrix of order $n \times p$ for observations of explanatory variables.

U: error term of $n \times 1$.

To estimate the regression model, it must be and first of all, the model should have some characteristics depending on a certain assumptions to increase sobriety of the regression model, and if these assumptions is not available, then the model will suffer from the problems which are making the estimation process very weak and imprecise, hence the estimated parameters will be inefficient.

So, these assumptions are:

1) $E(u_i) = 0$

2) $V(u_i) = \sigma_u^2$

3) $u_i \sim N(0, \sigma_u^2)$

4) $cov(u_i, u_j) = 0 \quad \forall\, i \neq j,\ i = 1,2,\ldots,n$

I.Bayes Aproach: The idea of Baysian approach is dealing with the unknow parameters as a random variable having a distribution function and these parameters have a prior information differ quantively and qualitively depending on size of the available information which the researcher had through the experiences, the pervious experiments or which to be identical or close to the work. The difficulty of this approach lies in collection the information about the unknown parameters and determine its prior probability distribution precisely because the difficulty of getting the prior or previous information or lack of accuracy for these information, however, these information are formulated in a prior probability distribution, which so- called prior probability density function $P(\theta)$ or prior p.d.f where this function is the difference point between Bayes approach and the other classical approaches.

The prior p.d.f is combined with the maximam likelihood $P(Y/\theta)$ for the current observations of Y by using Bayes Inversion Formula to get a good information or close to the acotual information about the unknown parameter ,all of these information are putting in a probability distribution form which is called the posterior probability density function $P(\theta/Y)$ or (posterior p.d.f) where this distribution is supposed to be a good description about the unknown parameter together with existence of the sample information

$$P\left(\theta/Y_i\right) = \frac{[\prod_{i=1}^{n} P(Y_i/\theta)].P(\theta)}{\int_{R\theta} [\prod_{i=1}^{n} P(Y_i/\theta)].P(\theta)\, d\theta} \ldots\ldots (4)$$

Bayes estimation approach needs to use so- called the loss function where the Bayes estimator yields by minimzing the expected loss function for the posterior distribution of the unknown parameter θ , given that the sample data Y is known, and the loss function must satisfying the following two conditions:

1. $L(\hat{\theta}, \theta) \geq 0 \quad \forall\, \hat{\theta}, \theta$

2. $L(\hat{\theta}, \theta) = 0 \quad \forall\, \hat{\theta} = \theta$

There are different kinds for the loss function and according to this difference and to the kinds of the prior distribution of the parameter θ , Bayes estimators will be different too, but our goal is to get a Baysian estimator $(\hat{\theta}_{Bayes})$, which at it the posterior expected loss will be as small as possible, so we will adopt the weighted squared Error Loss Function.

II. Prior Probability Density Functions: In Bayes Technique, we must have an information about the unknown parameters where these information is considered the main support point for this technique. We have two ways to get these information, the first one is the scientific way because these information are camed from the same used data attribute or quality and in this case we call it Data Based on prior p.d.f , the second one is achieved by collecting the information using the causal observations and the theoretical assumptions, in other words, not from the used or real data, or these information is coming from sources have not any relationship with the data, here we call it Non data Based on prior p.d.f.

Sometimes we get the information by using a mixture between the two ways. However, these information which we get it and whatever its source have a remarkable rule in choosing the prior pdf. There are four priors p.d.f, uninformative prior, information prior, prior p.d.f depending on previous sample, and a normal conjugate prior p.d.f. In our paper, we depend on the last one.

III. Normal Conjugate Prior p.d.f :This function has a good properities comparing with the other functions mentioned above which is making it more widely using because it considered a p.d.f of well – known parameters besides it is being explicit, determined and proper, note that the prior p.d.f is built on the likelihood function of the current sample observations by considering it as a function in the parameter θ. And it is worthly to mention that the prior p.d.f $P(\theta)$, likelihood function $P(^{Y}/_{\theta})$, and the posterior p.d.f $P(^{\theta}/_{Y})$ are characterized that it have the same functional formulas, but with new different parameters. This type of functions is preffered to use instead of the uninformative prior p.d.f because it is being improper function, also we preffer it because it treats the multicollinearity problem. One of the famous normal conjugate prior p.d.f is the prior joint p.d.f of Normal – Gamma.

Here, we have two unknown parameters (β, σ) and each parameter has a certain distribution according to a certain conditions and rules, that is when we do not have information about these parameters, then we follow what Jeffery is said in determination the prior p.d.f.. He said if interval of the parameter which we want to estimate was in $(-\infty, \infty)$, then the prior p.d.f. will be a uniform distribution:

$P(\beta)\ d\beta\ \alpha\ d\beta,\ -\infty < \beta < \infty \quad or\ P(\beta)\ d\beta\ \alpha\ constant,\ -\infty < \beta < \infty\ \ldots\ldots(5)$ and if the parameter interval was in $(0, \infty)$ then the prior p.d.f will be a log – uniform distribution: $P(\sigma)\ d\sigma\ \alpha\ \frac{1}{\sigma}\ d\sigma \ldots\ldots(6)$

As a result, the vector (β) has a multivariate normal distribution given by the following p.d.f.

$$P(\beta/\sigma) = \frac{1}{\sigma^{m}}\exp[-\frac{1}{2\sigma^2}\{(\beta-\bar{\beta}_p)'\dot{Q}(\beta-\bar{\beta}_p)\}] \ldots\ldots(7)$$

Where $-\infty < \beta < \infty$

$\bar{\beta}$: denotes to mean of the prior distribution.

$\sigma^2 Q^{-1}$: denotes to variance – covariance matrix and furthermore, Q is a positive definite matrix. To get the joint p.d.f. of (β, σ), we use the following equation:

$$P(\beta, \sigma) = P(\sigma)P(\beta/\sigma) \ldots\ldots.(8)$$

$$P(\beta, \sigma)\ \alpha\ \frac{1}{\sigma^{m+1}}\exp[-\frac{1}{2\sigma^2}\{(\beta-\bar{\beta}_p)'\dot{Q}(\beta-\bar{\beta}_p)\}] \quad \ldots\ldots(9)$$

Where P= number of parameters.

The likelihood function for the observations is given by:

$$P(Y/\beta,\sigma)\ \alpha\ \frac{1}{\sigma^n}\exp[-\frac{1}{2\sigma^2}\{(Y-X\beta)'(Y-X\beta)\}] \quad \dots\dots(10)$$

combining Equation (9) which represents the prior p.d.f for the two parameters with eq(10) yields:

$$P(\beta,\sigma/Y)\ \alpha\ \frac{1}{\sigma^{n+m+1}}\exp[-\frac{1}{2\sigma^2}\{(Y-X\beta)'(Y-X\beta)+(\beta-\bar{\beta}_p)'Q(\beta-\bar{\beta}_p)\}] \dots\dots (11)$$

$$\text{or}\quad \alpha\ \frac{1}{\sigma^{n+m+1}}\exp[-\frac{1}{2\sigma^2}\left[\begin{Bmatrix} Y-X\beta \\ Q^{\frac{1}{2}}\bar{\beta}_p-Q^{\frac{1}{2}}\beta\end{Bmatrix}'\begin{Bmatrix} Y-X\beta \\ Q^{\frac{1}{2}}\bar{\beta}_p-Q^{\frac{1}{2}}\beta\end{Bmatrix}\right]$$

Now, by letting:

$$w=\begin{bmatrix} Y \\ Q^{\frac{1}{2}}\bar{\beta}_p\end{bmatrix},\ Z=\begin{bmatrix} X \\ Q^{\frac{1}{2}}\end{bmatrix}$$

then, the posterior distribution formula becomes:

$$P(\beta,\sigma/Y)\ \alpha\ \frac{1}{\sigma^{n+m+1}}\exp[-\frac{1}{2\sigma^2}\{(W-Z\beta)'(W-Z\beta)\}] \quad \dots\dots(12)$$

and by letting: $\bar{\beta}_{\beta c}=(\acute{Z}Z)^{-1}\acute{Z}W \quad \dots\dots(13)$

where: $\bar{\beta}_{\beta c}$ = Bayes estimator based on a normal conjugate prior function then,

the posterior p.d.f. becomes:

$$P(\beta,\sigma/Y)\ \alpha\ \frac{1}{\sigma^{n+m+1}}\exp[-\frac{1}{2\sigma^2}\{(W-Z\bar{\beta}_{\beta c})'(W-Z\bar{\beta}_{\beta c})+(\beta-\bar{\beta}_{\beta c})'Z'Z(\beta-\bar{\beta}_{\beta c})\}] \quad \dots(14)$$

Now, integrate eq.(14) with respect to σ, then we'll get the marginal p.d.f for the parameters vector (β).

$$P(\beta/Y)\ \alpha\ [(\beta-\bar{\beta}_{\beta c})'Z'Z(\beta-\bar{\beta}_{\beta c})]^{\frac{-n+m}{2}} \quad \dots\dots(15)$$

The fovmula (15) represents a p.d.f of m- variate t distribution with mean equals to $(\bar{\beta}_{\beta c})$ given by the formula (16) which represents Bayes estimator for the parameters vector (β) based on a normal conjugate prior function, that is, Bayes estimator is given by:

$$\bar{\beta}_{\beta c}=\left[\begin{pmatrix} X \\ Q^{\frac{1}{2}}\end{pmatrix}'\begin{pmatrix} X \\ Q^{\frac{1}{2}}\end{pmatrix}\right]^{-1}\left[\begin{pmatrix} X \\ Q^{\frac{1}{2}}\end{pmatrix}'\begin{pmatrix} Y \\ Q^{\frac{1}{2}}\bar{\beta}_p\end{pmatrix}\right] \dots\dots(16)$$

4.The Practical side: In this section we present an actual experiment about one of electrical energy production stations in Iraq – karbala station – here, we use eight variable one of them is the response (dependent) variable Y and the remaining variables represent the explanatory (independent) variables for (30) months in accordance with the following multiple linear regression model:

$$Y_i=\beta_0+\beta_1X_{i1}+\beta_2X_{i2}+\beta_3X_{i3}+\beta_4X_{i4}+\beta_5X_{i5}+\beta_6X_{i6}+\beta_7X_{i7}+u_i$$

Where: $Y_i = Electrical\ energy\ production\ (\ mega/watt)$

X_1: *Heavy fuel (black oil)(litre)*

X_2: *Light fuel (kerosence)(litre)*

X_3: *Lubricating oil (litre)*

X_4: *Chimical materials (litre)*

X_5: *Operating hours average (hour/month)*

X_6: *Overtime wages (in Dinar)*

X_7: *Monthly tempreture average.*

The statistical program (MATLAb 2105a) has been used to estimate parameters of the previous model.

Table(1) shows estimation of multiple model parameters based on Bayes approach

parameters	Bases NC
β_0 β_1 β_2 β_3 β_4 β_5 β_6 β_7	0.0060 0.2259 0.0983 -0.1821 0.1745 0.3117 -0.4789 0.0032
MSE	0.0011
R^2	0.9769

From table (1), we find that mean square error (MSE) is decreasing value and very low, while value of the adjusted coefficient of determination (R^2) seems to be very high.

Table(2) represent ANOVA for regression model

source	Sum of squares	d.f.	MS	F_c	F_{table}
Regression	43143.0581	7	6163.2940	3.4923	2.4638
Residuals	39485.2401	22	1764.7836		
Total	82628.2982	29			

By comparing the computed F_c with value of the tabulated F(7,22,0.05)=2.46, we see that $F_c > F_t$, and that means there is at least one explanatory variable affects on the response variable Y.

6. Conclusions:

1-The results have been showed that mean square error (MSE) was very low with high value of coefficient of determination R^2 , and this is a good indicator.

2- The concluded results showed that the model is significance according to F- test ,and that means there is a strong relationship between the dependent variable Y and the other independent variables.

References

1- Baldi P. and Long A., "A Bayesian framework for the analysis of microarray expression data: regularized t-test and statistic inferences of gene changes" Bioinformatics, 17(6): 509-519,2001.

2- Bishop C., "Pattern recognition and mechine learning" Springer, 2006.

3- Christopher M. Strichland and Clair L.Alston (2013), "Bayesian analysis of the Normal linear Regression model", First Edition, university of Technology Brisbane, Australia, John Wiley& Sons, Ltd.

4-Demichelis F., P. Magni, p. piergiorgi, M. Rubin, and R. Bellazi. "A hierarchical Naïve Bayes models for handling sample heterogeneity in classification problems: an application to tissue microarrays". BMS Bioinformatics, 7:514, 2016.

5- Gordon, N., J., D. J. Salmond, and A.F.M. Smith. "Novel Approach to Nonlinear Non- Gausian Bayesian state Estimation" IEEE Proceedings F on Radar and singal Processing. Vol.140, 1993, pp. 107-113.

6- Iswari A A I A C, IW Sumerjaya and IGAM Srinadi (2014), Analisis Regresi Bayes Linear Sederhana dengan prior Noninformatif E- Journal Mathematika, 3(2) 38-44.

7- Kevin P, Murphy, (2007),"Conjugate Bayesian analysis of the Gaussian distribution", murphy K.
8- Park, T., and G.Casella. "The Bayesian Lasso". Journal of the American statistical Association. Vol.103, No. 482, 2008, pp.681-686.
9- Raftery, E., Maigan, D. & Hoeting, J.A. (1997), " Bayesian model averaging for linear regression model", Jassa. Vol. 75, No, 327, p. 801-816.
10- Michael J. zyphur and Frederick L. Oswald, " Bayesian Estimation and inference: A user's Guide". Journal of Management 41(2): 390-420, 2013.
11- Sinay M. S. and J. S. J. Hsu (2014), Baysian Inferensi of a multivariate regression model (Hindawi Publishing Corporation).
12- Williams P. M. (1995), Bayesian regularisation and Pruning using a Laplace prior. Neural Computation, 7(1), 117-143.

Estimating Shrinkage Parameter of Generalized Liu Estimator in Logistic Regression Model

Najlaa Saad Alsharabi
Department of Statistics and Informatics, University of Mosul, Mosul, Iraq
E-mail: najlaa.s.a@uomosul.edu.iq , mob: 07703880964

Abstract: The logistic regression model is one of the modern statistical methods developed to predict the set of quantitative variables (nominal or monotonous), and it is considered as an alternative test for the simple and multiple linear regression equation as well as it is subject to the model concepts in terms of the possibility of testing the effect of the overall pattern of the group of independent variables on the dependent variable and in terms of its use For concepts of standard matching criteria, and in some cases there is a correlation between the explanatory variables which leads to contrast variation and this problem is called the problem of Multicollinearity. In this study a generalized Liu estimator was introduced to combat the multicollinearity in the logistic regression model. The generalized liu parameter (shrinkage coefficient) was estimated by different methods and a comparison was made between these methods and the ML method using the mean square error standard. Simulation results showed that the proposed generalized liu estimator possesses less (MSE) compared to the ML method in the case of a multicollinearity. Selecting the shrinkage coefficient based on work done by Akdeniz et el (1999) i.e (D4) is more efficient than other methods.

Key words logistic regression, multicollinearity, mean square error, ridge estimator, liu estimator

1. Introduction

The logistic regression model is an important statistical model in analyzing binary data (0 or 1) as the primary goal of most studies is to analyze and evaluate relationships between a set of variables to obtain a formula by which we describe the model and uses the logistic regression model to describe the relationship between the response variable of the discontinuous type and the explanatory variables, prediction, estimation and control of the values of the dependent variable according to the changes in the values of the variable with interpretation (Farhood, 2014)

One of the characteristics of the binary response logistic regression is that the dependent variable (Y) of the response variable follows the Bernoulli distribution taking the value (1) with a probability of (π) probability of success, and a value (0) with a probability (1- π) of failure probability (Qasim,2011). As we work in linear regression whose independent and dependent variables take continuous values, the model that correlation the variables is as follows:

$$Y = \beta_0 + \beta_1 X + e \qquad (1)$$

Whereas (Y): represents a continuous observational variable and assuming that the average values of (Y) observation or actual at a given value of the variable x which is E(Y) and that the variable (e) represents a random error, then the model can be written (1) as follows:

$$E\left(Y \middle| X\right) = \beta_0 + \beta_1 X \qquad (2)$$

It is known in regression that the right side of these models takes values ($-\infty, +\infty$), but when the variable (y) is binary, the above model is not appropriate because:

$$E(Y|X) = P_r(Y=1) = \pi \qquad (3)$$

Thus, the value of the right side is confined between the two numbers (0.1), and thus the model is not applicable from the regression point of view, and one of the methods of solving this problem is to enter an appropriate mathematical transformation on the dependent variable (Y). Since ($0 \le \pi \le 1$), then the ratio ($\pi / (1-\pi)$) is a positive amount confined between (0, ∞) i.e. ($0 \le \pi / (1-\pi) \le \infty$) and taking the natural logarithm for the base (e) of the amount ($\pi / (1-\pi)$) the value domain becomes between (-∞, + ∞) and is (($-\infty \le \log_e (\pi / (1-\pi)) \le \infty$). Therefore, the regression model can be written in the case of one explanatory variable as follows:

$$\log_e\left(\frac{\pi}{1-\pi}\right) = \beta_0 + \beta_1 X \qquad (4)$$

But if we have more than one explanatory variable, then the model is formulated as follows:

$$\log_e\left(\frac{\pi}{1-\pi}\right) = \beta_0 + \sum_{j=1}^{p} \beta_j X_{ij} \qquad (5)$$

As: i = 1,2,3, ..., n .$\beta_1, \beta_2, \dots \dots \dots, \beta_p$: Vector parameters required estimatexplanatory are e X_{ij}:. variables.

As for ($\pi / (1-\pi)$) odds of success rate or preference ratio for the desired event and its mathematical formula are as follows:

$$\frac{P(Y=1)}{1-P(Y=1)} = e^{\beta_0 + \sum_{j=1}^{p} \beta_j x_{ij}} \qquad (6)$$

The probability formula for the logistic regression model is written as follows:

$$\pi = \frac{e^{X\beta}}{1+e^{X\beta}} \qquad (7)$$

And the amount Loge($\pi / (1-\pi)$) is called the logs odds of success logarithm.

Logistic regression does not require many assumptions. It only requires that there is no correlation between the explanatory variables and that the volume of observations is large in each group that is assumed to be greater than five times the number of parameters used in the final model (Demosthenes, 2006).

The estimation of the parameters of the logistic regression model is carried out using the Maximum Likelihood Method (ML), which is one of the most famous estimation methods in statistics. Assuming that the observations are independent, the logarithmic likelihood function is defined by the following formula: (Hosmer and Lemeshow, 2000)

$$L = \sum_{i=1}^{n} Y_i \log(\pi_i) + (1-Y_i)\log(1-\pi_i) \qquad (8)$$

By maximizing the likelihood function (L) and taking the derivative with respect to the parameters (β) and equating the result of the equation with zero, the possibility function is given as:

$$0=\sum_{i=1}^{n}X_i\left(Y_i-\pi_i\right) \qquad (9)$$

Since equation (9) is a nonlinear parameter, some special methods should be used to obtain the appropriate solutions. Therefore, Iteratively Re-Weighted Lest Squares (IRLS) can be applied to obtain appropriate solutions. The maximum likelihood estimator (MLE) of the parameters (β) can be found using the IRLS algorithm as follows:

$$\hat{\beta}_{MLE}=S^{-1}X'\hat{W}\hat{Z} \qquad (10)$$

As $\hat{Z}_i=\log(\hat{\pi}_i)$, $\hat{W}=\mathrm{diag}\left(\hat{\pi}_i(1-\hat{\pi}_i)\right)$, $S=\acute{X}\hat{W}X$

One disadvantage of using MLE is that MSE becomes bulky when explanatory variables are Linear dependent, which is called the problem of multicollinearity. A condition number (CN) has been developed to test the existence of the problem of multicollinearity between the variables known as the following formula:

$$CN=\left(\frac{\lambda_{\max}}{\lambda_{\min}}\right)^{1/2} \qquad (11)$$

As: $\lambda_{\max}$, $\lambda_{\min}$ They represent the largest and smallest eigenvalue roots of the matrix (S), if the value of CN <10 this means there is no problem of multicollinearity between the explanatory variables and if it is 10< CN <30 then there is a problem of moderate multicollinearity between the explanatory variables and if the value CN> 30 This means that there is a strong multicollinearity problem between the explanatory variables (Inan and Erdogan, 2013;Algamal, 2018) Also when the eigenvalue root values of the matrix (S) are close to zero, this indicates that there is a problem of multicollinearity between the variables and this will lead to an increase in the value of (MSE) .

The value of the mean square error of equation (10) is found according to the following formula: [Siray et al. 2015]

$$MSE(\hat{\beta}_{ML})=\sum_{j=1}^{p}\frac{1}{\lambda_j} \qquad (12)$$

As: λ_i represent the eigenvalue roots of the matrix (S).

When there is multicollinearity, the maximum likelihood estimator method (ML) suffer from inflation in the variations of the estimated parameters and the occurrence of instability, and this inflation is represented by the diagonal elements of the matrix (S). To solve this problem, (Schaefer et al., 1984) suggested a logistic ridge estimator (LRE) that was first introduced by 1970 (Horal & Kennard), and used it to estimate the parameters for the Multiple Linear Regression Model. This method is summarized by adding a small positive constant quantity (k) whose value falls between zero and one (0≤ k ≤1) to the diagonal elements of the information matrix (S) to obtain more accurate estimator, and this method works to decouple the links between the explanatory variables and the logistic character estimator is defined according to the formula next: (Månsson and Shukur, 2011; Alanaza and Algamal, 2018)

$$\hat{\beta}_{LRE} = (S + kI)^{-1} \hat{X} \hat{W} \hat{Z} \qquad (13)$$

The estimator (ML) can be considered a special case of equation (13) when the value of (k = 0).The value of k in logistic regression models is found according to the formulas $k = \dfrac{1}{\hat{\beta}'_{ML}\hat{\beta}_{ML}}$ (Schaefer et al., 1984).

2. Generalized Liu Estimator (GL)

The researcher Liu proposed in 1993 a new estimator to address the problem of multicollinearity. It combined the features Stein estimator in 1956 and Ordinary Ridge Regression estimator (ORR). where it is estimated:

$$\hat{\alpha}_{RR} = \left(X^{*\prime}WX^{*} + KI\right)^{-1} X^{*\prime}WZ \qquad (14)$$

Where $X^{*} = XP$ and P is represents a "perpendicular matrix, whose columns represent distinctive vectors corresponding to the characteristic roots of the matrix of information $\left(X'WX\right)$ and $P'P = PP' = I$. This model called the Canonical Linear Model or Uncorrelated components model, and the estimation MLE of the (α) is given:

$$\hat{\alpha}_{MLE} = \left(\chi^{*\prime}W\chi^{*}\right)^{-1} \chi^{*\prime}WZ \qquad (15)$$

It has advantages and an advantage. It is advantageous in the practical application but it is a complex function of (K).(Algamala,2018)

Akdeniz & Kaciranlar proposed in 1995 a new estimator named (GL). It is the general state of estimator (LE) there is a special advantage to estimating (LE) overcomes the estimator (ORR) where (LE) is a linear function with a bias parameter (d). So it is easier to calculate then the character parameter k for estimator (ORR). The character is also estimated as a decreasing function in k while Liu is estimating an increasing function in (d). The general Liu is indicated by:

$$\hat{\alpha}_{GL} = (\Lambda + I)^{-1}\left(X^{*\prime}WZ + D\hat{\alpha}_{MLE}\right)$$

$$\hat{\alpha}_{GL} = (\Lambda + I)^{-1}\left(\Lambda\hat{\alpha}_{MLE} + D\hat{\alpha}_{MLE}\right)$$

$$\hat{\alpha}_{GL} = (\Lambda + I)^{-1}(\Lambda + D)\hat{\alpha}_{MLE}$$

Which can be written as follows:

$$\hat{\alpha}_{GL} = \left(I - (\Lambda + I)^{-1}(\Lambda - D)\right) \qquad (16)$$

$D = diag(d_i), 0 < d_i < 1$ represents a diagonal matrix with bias parameters (d_i) and $\Lambda = X^{*'}WX^{*}$ (Algamala and Asar, 2018).

The forecast for the estimate $\hat{\alpha}_{GL}$ as follows:

$$E\hat{\alpha}_{GL} = \left(I - (\Lambda + I)^{-1}(\mathrm{I} - D)\right)E\hat{\alpha}_{MLE}$$

$$E\hat{\alpha}_{GL} = \left(I - (\Lambda + I)^{-1}(\mathrm{I} - D)\right)\hat{\alpha}_{MLE} \qquad (17)$$

The estimator of (GL) is biased for parameter (α) and the biased estimated is:

$$Bias(\hat{\alpha}_{GL}) = E(\hat{\alpha}_{GL} - \alpha)$$
$$= -(\Lambda + I)^{-1}(I - D)\alpha \qquad (18)$$

The variance matrix of the (GL) is estimated as follows:

$$Var(\hat{\alpha}_{GL}) = (I - (\Lambda + I)^{-1}(I - D))Var(\hat{\alpha}_{MLE})(I - (\Lambda + I)^{-1}(I - D))'$$

$$= \hat{\sigma}^2 (I - M)\Lambda^{-1}(I - M)' \qquad (19)$$

where :

$$M = (\Lambda + I)^{-1}\ (I - D)$$

The matrix of average error squares to (GL) estimator are as follows:

$$MSE(\hat{\alpha}_{GL}) = Var(\hat{\alpha}_{GL}) + (Bias(\hat{\alpha}_{GL}))^2$$
$$= \hat{\sigma}^2 (I - M)\Lambda^{-1}\ (I - M)' + M\alpha\alpha' M' \qquad (20)$$

3. Estimating the shrinkage parameter

In order to estimate the optimal value of (D) in Eq.(16), several methods will be proposed. The idea behind these proposed estimators are obtained from the work of Hoerl and Kennard (1970), Kibria (2003) and Khalaf and Shukur (2005).Where several different methods of estimating the shrinkage parameter for linear ridge regression have been proposed. The first estimator which is based on the work by Hoerl and Kennard (1970) is the following:

$$D_1 = \frac{\hat{\alpha}_j^2 - 1}{\frac{1}{\hat{\lambda}_j} + \hat{\alpha}_j^2} \qquad (21)$$

$$D_2 = Max\left(0, \frac{\hat{\alpha}_{max}^2 - 1}{\frac{1}{\hat{\lambda}_{\max}} + \hat{\alpha}_{max}^2}\right) \qquad (22)$$

Where we define $\hat{\alpha}^2_{\max}$ and $\hat{\lambda}_{\max}$ to be the maximum element of $\hat{\alpha}^2_j$ and $\hat{\lambda}_j$ respectively. Furthermore, the following estimators, which are based on the ideas in Akdeniz and Kacıranlar (1995),are proposed:

$$D_3 = \frac{\hat{\lambda}_j\left(\hat{\alpha}_j^2 - \hat{\sigma}^2\right)}{\left(\hat{\lambda}_j\hat{\alpha}_j^2 + \hat{\sigma}^2\right)} \qquad (23)$$

Akdeniz et el proposed method in 1999 (Alheety and Kibria, 2009) as following:

$$D_4 = \left(1 - \sqrt{\frac{\hat{\sigma}^2\left(\hat{\lambda}_j + 1\right)^2}{\hat{\lambda}_j\hat{\alpha}_j^2 + \hat{\sigma}^2}}\right) \qquad (24)$$

The following estimators, which are based on the ideas in Kibria (2003),are proposed:

$$D_5 = Max\left(0, Median\left(\frac{\hat{\alpha}_j^2 - 1}{\frac{1}{\hat{\lambda}_j} + \hat{\alpha}_j^2}\right)\right) \qquad (25)$$

Using the average value and median is very common when estimating the shrinkage parameter for the ridge regression . Finally, the following estimators are proposed:

$$D_6 = Max\left(0, Max\left(\frac{\hat{\alpha}_j^2 - 1}{\frac{1}{\hat{\lambda}_j} + \hat{\alpha}_j^2}\right)\right) \qquad (26)$$

For these estimators other quintiles than the median is used which was successfully applied by Khalaf and Shukur (2005).

4. Monte Carlo simulation study

In this section, a comprehensive simulation study was conducted to evaluate the performance of the Estimating the shrinkage parameter (D) of Liu estimator. The explanatory variables $x_i^T = (x_{i1}, x_{i2}, ..., x_{in})$ have been generated from the following formula:

$$x_{ij} = (1-\rho^2)^{1/2} w_{ij} + \rho w_{ip} \qquad i = 1, 2, ..., n \;\&\; j = 1, 2, ..., p \qquad (27)$$

where ρ represents the correlation between the explanatory variables, p represents the number of explanatory variables, and w_{ij} are independent standard normal pseudo-random numbers and w_{ip}: represents the values of the last column of the variables generated. The response variable for (n) of observations was found according to the formula of the logistic regression model:

$$Y \approx B(\frac{\exp(X\beta)}{1+\exp(X\beta)}) \qquad (28)$$

and $\beta = \beta_1 + \beta_2 + \beta_3 + \ldots + \beta_p$ with $\sum_{j=1}^{p} \beta_j = 1$ and $\beta_1 = \beta_2 = \beta_3 = \ldots = \beta_p$ (Kibria, 2003; Månsson and Shukur, 2011). Because the sample size has direct impact on the prediction accuracy three representative values of the sample size are considered: 50, 75 and 150. In addition, the number of the explanatory variables are considered as $p = 5$ and $p = 8$ Further, because we are interested in the effect of multicollinearity, in which the degrees of correlation considered more important, three values of the pairwise correlation are considered with $\rho = (0.90,\ 0.95,\ 0.99)$. The experiment was repeated (1000) times. And the mean square error (MSE) is calculated according to the following formula:

$$MSE(\hat{\beta}_r) = \frac{1}{1000} \sum_{r=1}^{1000} (\hat{\beta}_r - \beta)^T (\hat{\beta}_r - \beta) \qquad (29)$$

where $\hat{\beta}_r$ is the obtained liu estimator with different shrinkage parameter D_1, D_2, D_3, D_4, D_5 *and* D_6. We conclude from the results of Table (1) The lowest value for MSE when $n = 150,\ p = 5$ *and* $\rho = 0.90$, the MSE of the D_5 was about 0.7747. As the correlation coefficient value increases, the MSE value increases when taking all the probabilities of the number of explanatory variables (p) and the sample size (n). In addition, the estimated performance (D_5) is better than the rest of the estimators. The more the number of explanatory variables (p) increases, the value of (MSE) increases, and this increase affects the quantity of estimators. However, the estimated performance (D_5) is better than the rest of the estimators. As the sample size increases, the value of MSE decreases when taking different values for each correlation coefficient and the number of explanatory variables. The best performance is performance shrinkage parameter D_5 for liu estimator. The performance of the parameter (D_3) of Liu estimation was the worst for having the highest values of the MSE.

Table 1: Average MSE values for different values of ρ, n and **p .**

	n	ρ	D_1	D_2	D_3	D_4	D_5	D_6
p=5	٥٠	0.90	2.8794	1.7449	3.6189	3.2373	1.2587	1.6581
		0.95	5.3993	3.3228	5.5453	5.7579	1.5086	2.8787
		0.99	31.6187	17.5280	39.7088	26.5289	3.3218	11.8269
	75	0.90	2.3207	1.2486	1.8853	2.7669	1.0973	1.2277
		0.95	4.2766	2.0539	3.5611	5.1903	1.3191	1.9051
		0.99	21.3497	11.0621	24.1213	20.3833	2.4319	7.9190
	150	0.90	1.4638	0.7814	1.1045	1.9021	0.7747*	0.7808
		0.95	2.8528	1.1400	2.0094	3.9486	1.0342	1.1257
		0.99	14.4391	4.8206	11.4733	18.2836	1.7055	3.8633

p=8	٥٠	0.90	5.9541	2.9551	5.3456	6.6659	1.8113	2.7295
		0.95	12.1249	5.9670	13.1731	11.7802	2.1132	5.0327
		0.99	98.7982	33.3329	130.7653**	78.6804	4.7248	23.0994
	75	0.90	4.3456	1.8381	3.3859	5.2627	1.5262	1.7887
		0.95	8.4137	3.5607	7.3510	9.7505	1.9428	3.2305
		0.99	53.4659	20.4732	65.5540	45.7253	3.6323	14.8509
	150	0.90	2.5463	1.0351	1.7324	3.4226	1.0185	1.0334
		0.95	5.1152	1.7471	3.4946	6.9885	1.5340	1.7172
		0.99	29.5347	8.5429	26.6366	34.1068	2.3168	6.8949

* The lowest value for MSE.

**The largest value for MAE.

5- Conclusion

In this paper, a compare of different shrinkage parameter selection of the liu regression model. Simulation results demonstrate that shrinkage parameter selection based on the work by Kibria (2003) ie (D_2) is more efficient than D_1, D_3, D_4, D_5 *and* D_6 methods when $\rho \geq 0.90$." As the sample size increases, the value of (MSE) decreases when taking different values for each correlation coefficient and number of explanatory variables.

References

1- Akdeniz, F., and Kacıranlar, S.(1995)," On the almost unbiased generalized Liu estimator and unbiased estimation of the bias and MSE". Common Statist, Theory, Meth. 24:1789–1797.

2- Alanaza, Mazin M. and Algamal, Zakariya Y. Proposed methods in estimating the ridge regression parameter in Poisson regression model. EJASA, 2018. Vol. 11, Issue 02, 506-515.

3- Alheety, M.I., and Kibria, B. M. G.(2009)," On the Liu and almost unbiased Liu estimators in the presence of multicollinearity with heteroscedastic or correlated errors". Surveys in Mathematics and its Applications. 4, 155-167.

4- Algamal, Zakariya Y.(2018).' biased estimators in poisson regression model in the presence of multicollinearity:a subject review'. QJAE. Vol. 20, Issue 01, 2312-9883.

5- Algamal, Zakariya Y. and Asar, Yasin,(2018)," Liu-type estimator for the gamma regression mode". Communications in Statistics - Theory and Methods Taylor & Francis Group, LLC..

6- Demosthenes B. Panagiotakos (2006) ," A comparison between Logistic Regression and Linear Discriminant Analysis for the Prediction of Categorical Health Outcomes", International Journal of Statistical Sciences, Number 5, pp (73-84).

7- Farhood, Suhaila Hammoud Abdullah (2014), “Using Logistic Regression to Study the Factors Affecting Stock Performance (An Applied Study on the Kuwait Stock Exchange), The Public Authority for Applied Education and Training, The State of Kuwait, Statistics Department, Al-Azhar Magazine, No. 16, Pg. 47-68.

8- Hosmer, D. D. and Lemeshow, S. (2000). Applied Logistic Regression: John Wiley and Sons.

9- Hoerl, A. E., and R. W. Kennard.,(1970)," Ridge regression: Biased estimation for nonorthogonal problems". Technometrics. 12 (1):55–67.

10- Inan, D., and Erdogan, B. E. (2013). Liu-type logistic estimator. Comm. Statist. Sim. Comp., 42(7), 1578-1586.

11- Kibria, B. M. G,(2003),"Performance of some new ridge regression estimators". Communications in Statistics—Theory and Methods. 32:419–435.

12- Khalaf, G., and Shukur, G.(2005)," Choosing ridge parameters for regression problems". Communications in Statistics—Theory and Methods.. 34:1177–1182.

13- Mansson, K. and Shukur, G. (2011). On ridge parameters in logistic regression. Comm. Statist. Theo. Meth., 40(18), 3366-3381.

14- Qasem, Bahaa Abdul-Razzaq (2011), "Analysis of the effect of some variables on the incidence of periodontal disease using the logistic regression model, Journal of Statistical Sciences, University of Basra, No. 27, pp. 139-164.

15- Schaefer, R. L., Roi, L. D. and Wolfe, R. A. (1984). A ridge logistic estimator. Comm. Statist. Theo. Meth., 13(1), 99-113.

A New Revised Efficient of VAM to Find the Initial Solution for the Transportation Problem

Hussein Ali Hussein[1]   Mushtak A.K. Shiker[2*]   Mohammed Shakir Mahdi Zabiba[3]

1,2Department of Mathematics, College of Education for Pure Sciences,University of Babylon, Babil-
3Department of Mathematics, Faculty of Education for Women, University of Kufa
1hussainalseaidi@gmail.com   2mmttmmhh@yahoo.com   3mohammedsh.mahdi@uokufa.edu.iq

Abstract: Transportation Problem (TP) is a very important problem which has been vastly studied in Operations Research domain. There are some classical methods to find the initial basic feasible solution (IBFS) which minimize the total shipping cost of (TP) such as north-west corner method (NWCM), minimum cost method (MCM) and Vogel's approximation method (VAM) which the best one of them. In this paper, we suggest a new amendment to (VAM) to find (IBFS) of (TP), which is an iterative method and the results will be near the optimal solution and in some cases equal to the optimal solution. In the numerical experiences we compare the results of the new approach with other classical methods to verify the efficiency of the new method. The proposed method is very effective and well-suited for use in solving these problems of various sizes.

Keywords: Liner Programming, Transportation Problems, Vogel's approximation method, Initial Basic Feasible Solution.

*** Corresponding Author**

Introduction

The transportation problem (TP) have been extensively discussed and studied in Operations Research. They are usually concerned with how to reduce the transportation cost of homogeneous commodities (or get the optimal time for transportation) that has a number of sources (such as factories) and a number of destinations (such as warehouses) whereas meeting supply and request constraints [3, 5]. Transportation modules act as a paramount tool of supply chains and logistics administration to obtain the optimum cost and improve services. We can find the initial basic feasible solution (IBFS) of the (TP) using single of known classic approaches, such as the north-west corner method (NWCM), minimum cost method (MCM), Vogel’s approximation method (VAM) [9]. In addition to the classical methods, many researchers have presented other methods to get (IBFS) to the TP.

In 1781, French mathematician Gaspard Monge, in cooperation with the army of Napoleon Bonaparte, published a mathematical model dealing with transporting soil at the lowest possible cost between different construction sites for the purpose of building forts and military ways. Although Monge laid a theoretical basis for solving (TP), no algorithm was advanced till 1941 when American mathematician Frank L. Hitchcock disseminated his solution of Monge’s problem [4]. Since that time, especially in last years, many methods have been presented to find (IBFS), for example, in (2015) Abdul S. Soomro et.al. suggested a modified Vogel’s approximate method for solving (TP) [2]. In (2016) Mollah M. A. et.al. presented a new approach to solve (TP) [7]. Also in (2016) Neetu M. Sharma and Ashok P. Bhadane developed an alternative method to north-west corner method for solving (TP) [8]. In (2017) Abul Kalam S. and Bellel H. discussed a new method for solving (TP) considering average penalty [1]. In (2018) Palanivel M. and Suganya M. provided a new method to solve (TP)- harmonic mean approach [10]. In (2019) Sourav P. clarified a new proposition to compute an (IBFS) of (TP) [11]. Also in (2019) Kenan K. and Yusuf S. gave an approximation method to obtain (IBFS) of (TP) [6].

The cost of (TP) has a major impact on the cost and prices of goods, so researchers in this field are seeking to provide the best method to minimize the cost of (TP). In this work we present a new approach to find (IBFS) of (TP) which minimizes its cost by making an amendment to the (VAM). The numerical results showed the efficiency of the new method by comparing its solution results with the solution of the three classical methods (NWCM, MCM and VAM). The (IBFS) of the new approach is better than the results of the three classic methods and near the optimal solution. Moreover, it is in some cases give us the optimal solution.

Transportation Model

In (TP) there exist m origins of supply $S_1, S_2, \dots, S_m$ and n destinations of demand $D_1, D_2, \dots, D_n$, each one of them is specially represented. The brackets symbolize the tracks connecting the origins and the destinations. Bracket (i, j) connection from origin i (i = 1,2, … , m) into destination j (j = 1,2, … , n), where c_{ij} represents the transportation cost for each unit of products, while x_{ij} is the charged quantity. The available units at supply in origin i represented by a_i, while the available units at request in destination j represented by b_j .

$$\text{Minimize} \quad Z = \sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij}\, x_{ij} \qquad \text{(gross cost)} \quad \dots\dots.. \qquad (1)$$

Subject

$$\sum_{j=1}^{n} x_{ij} = a_i \quad , \quad i = 1,2, \dots \dots m \text{ (supply restrictions)} \dots\dots\dots \qquad (2)$$

$$\sum_{i=1}^{m} x_{ij} = b_j \quad , \quad j = 1,2, \dots \dots n \text{ (demand restrictions)} \dots\dots\dots \qquad (3)$$

Where $x_{ij} \geq 0$ for every *i* and *j* ……… (4)

Table ❶: Creek of Transportation Problem

Origin (i)	Destination (j)				Supply (a_i)
	D_1	D_2	…	D_n	
S_1	x_{11} c_{11}	x_{12} c_{12}	…	x_{1n} c_{1n}	a_1
S_2	x_{21} c_{21}	x_{22} c_{22}	…	x_{2n} c_{2n}	a_2
⋮	⋮	⋮	x_{ij} c_{ij}	⋮	⋮
S_m	x_{m1} c_{m1}	x_{m2} c_{m2}	…	x_{mn} c_{mn}	a_m
Demand (b_j)	b_1	b_2	…	b_n	$\sum a_i = \sum b_j$

Types of Transportation Problems:

❶ **Balanced (TP):**

The (TP) is called balanced when the quantum of products in the apportionment pivot is equivalent to the quantum of products desired by the request pivot i.e. $\sum_{i=1}^{m} a_i = \sum_{j=1}^{n} b_j$.

❷ Unbalanced (TP):

The (TP) is called unbalanced when the quantum of products ready in the apportionment pivot is not equivalent to the quantum of products desired by the request pivot i.e. $\sum_{i=1}^{m} a_i \neq \sum_{j=1}^{n} b_j$.

The New Algorithm (Al-Saeedi's Approach)

We depict the proceedings for getting (IBFS) for (TP) by the new approach in the following steps:

Step 1: Building the TP cost matrix. Examination whether aggregate supply equals the aggregate demand, if not, the transportation problem must be balanced.

Step 2: For each row of the transportation cost matrix, we specify the two lower costs ready. We subtract these two costs (called penalty) and position it to the oath of that row in a novel column made from the extension of the table on the right. For each column of the transportation cost matrix, we specify the highest cell cost and lowest cell cost. We subtract these two costs (called penalty) and position it beneath that column in a novel row was formed for this purpose by extending the table below.

Step 3: Among these distinct penalty exhibit, in Step 2و we select the great value (the largest difference).

Step 4: We, allocate the maximum possible units to the smaller cost cell in the chosen row (or column). when an equality happens among the largest differences, the choice may be taken for that row (or column), which has the smaller cost. when an equality happens in such lowest cost as well, option may be taking than that row or column by which extreme possible needs are fill up. If an equality happens in such allocating extreme possible needs then chosen cell in the row (or column) through extreme supply (demand). The chosen cell is allotment to that quantity and the congruous consume row (or column) is deleted than moreover study.

Step 5: Cancel out the row (or column) that is satisfied.

Step 6: Repeat Steps (2 – 5) until all columns and rows are satisfied.

Step 7: In the end, calculate the aggregate (TP) cost by applying the next equation:

$$Z = \sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij} x_{ij}$$

Numerical Examples

Example①: Consider the following transportation cost problem in Table ❷.

Table ❷

Origin	Destination				Supply
	D_1	D_2	D_3	D_4	
S_1	19	30	50	10	7
S_2	70	30	40	60	9
S_3	40	8	70	20	18
Demand	5	8	7	14	34

The presented transportation problem is balanced because gross supply = gross request = 34. According to algorithm of (Al-Saeedi's Approach), the above table acquired as follows:

Table ❸

Origin	Destination								Supply	Penalty			
	D_1		D_2		D_3		D_4						
S_1	5	19	0	30	0	50	2	10	7	9	20	-	-
S_2	0	70	2	30	7	40	0	60	9	10	10	10	10
S_3	0	40	6	8	0	70	12	20	18	12	12	12	62
Demand	5		8		7		14		34				
Penalty	51		22		30		50						
	-		22		30		50						
	-		22		30		40						
	-		22		30		-						

The transportation cost total is $Z = \sum_{i=1}^{3} \sum_{j=1}^{4} c_{ij}\, x_{ij}$

$\therefore\ Z = (5 \times 19) + (2 \times 10) + (2 \times 30) + (7 \times 40) + (6 \times 8) + (12 \times 20) \ = \ 743$ units

Example②: Consider the following transportation problem matrix in Table ❹.

Table ❹

Origin	Destination						Supply
	D_1		D_2		D_3		
S_1		10		12		5	17
S_2		6		9		11	30
S_3		3		8		4	53
Demand	22		30		16		Unbalanced

The presented transportation problem is unbalanced because the gross supply is not equal to the gross request.

Table ❺

Origin	Destination								Supply
	D_1		D_2		D_3		D_4		
S_1		10		12		5		0	17
S_2		6		9		11		0	30
S_3		3		8		4		0	53
Demand	22		30		16		32		100

Now transportation problem matrix is balanced because gross supply = gross request = 100.
According to the new Algorithm, we get the next table:

Table ❻

Origin	Destination								Supply	Penalty			
	D_1		D_2		D_3		D_4						
S_1	0	10	0	12	15	5	2	0	17	5	5	5	7
S_2	0	6	0	9	0	11	30	0	30	6	9	-	-
S_3	22	3	30	8	1	4	0	0	53	3	4	4	4
Demand	22		30		16		32		100				
Penalty	7		4		7		0						
	-		4		7		0						
	-		4		1		0						
	-		4		1		-						

The transportation cost total is $Z = \sum_{i=1}^{3} \sum_{j=1}^{4} c_{ij}\, x_{ij}$

$$\therefore\ Z = (15 \times 5) + (2 \times 0) + (30 \times 0) + (22 \times 3) + (30 \times 8) + (1 \times 4) \ = \ 385 \text{ units}$$

Comparison

Table ❼

Name	NWCM	MCM	VAM	(Al-Saeedi's Approach)
Example①	1015	814	779	743
Example②	529	422	385	385

As shown in Table ❼, the new approach gives a relatively better initial basic feasible solution (IBFS) than the results obtained by the other algorithms.

Conclusion

In the current highly competitive market, pressure is growing rapidly for institutions to specify the best methods to shipping commodity to customers. This is the reason why different institutions want to provide products to customers in the best method in terms of cost or time, so this transportation model provides a strong framework to meet this problem. In this paper, we have proposed a new amendment to the Vogel's Approximation Method to solve (TP). Through numerical examples, we saw that the transportation cost total using the new algorithm (Al-Saeedi's Approach) is best by comparing with the classic solution methods and sometimes we can get the optimal solution when using this method. Therefore, we can conclude that the new algorithm can be used efficiently to find the initial basic feasible solution (IBFS). We have to remarkable that our suggested approach is easy, clear and allows desirable results as required by the decision makers which could be an attractive alternative method for solving the transportation problems. Furthermore, it can be used for balanced and unbalanced transportation problems.

References

[1] Abul Kalam S.M.A. and Bellel H.M.D. *A New Method for Solving Transportation Problems Considering Average Penalty*. IOSR Journal of Mathematics (IOSR-JM), Volume 13, Issue 1 Ver. IV, PP:40-43, (2017).

[2] Abdul S. S.; Muhammad J. and Gurudeo A. T. *Modified Vogel's Approximation Method for Solving Transportation Problems*. Mathematical Theory and Modeling, Vol.5, No.4, PP:32-42, (2015).

[3] Bilqis A., Chastine F. and Erma S. *Total Opportunity Cost Matrix – Minimal Total: A New Approach to Determine Initial Basic Feasible Solution of a Transportation Problem*. Egyptian Informatics Journal, PP:1-11, (2019).

[4] Frank L. H. *The Distribution of a Product from Several Sources to Numerous Localities*. Journal of Mathematics and Physics, Vol. 20, PP:224-230, (1941).

[5] Hamdy A. T. *Operations Research an Introduction*. Malaysia, Tenth Edition, PP:207-219, (2017).

[6] Kenan K. and Yusuf S. *A Novel Approximation Method to Obtain Initial Basic Feasible Solution of Transportation Problem*. Journal of King Saud University – Engineering Sciences, PP:1-8, (2019).

[7] Mollah M. A., Aminur R. K., Sharif U.M. and Faruque A. *A New Approach to Solve Transportation Problems*. Open Journal of Optimization, 5, PP:22-30, (2016).

[8] Neetu M. S. and Ashok P. *An Alternative Method to North-West Corner Method for Solving Transportation Problem*. International Journal for Research in Engineering Application & Management (IJREAM), Vol-01, Issue 12, PP:1-3, (2016).

[9] Niluka R. and Lashika R. *Mathematical Model and a Case Study for Multi-Commodity Transportation Problem*. International Journal of Theoretical and Applied Mathematics, 4(1), PP:1-7, (2018).

[10] Palanivel M. and Suganya M. *A New Method to Solve Transportation Problem - Harmonic Mean Approach*. Engineering Technology Open Access Journal, Volume 2, Issue 3, PP:1-3, (2018).

[11] Sourav P. *A New Proposition to Compute an Initial Basic Feasible Solution of Transportation Problem*. Proceedings of the 5thInternational Conference on Engineering Research, Innovation and Education, Paper ID 159, PP:871-877, (2019).

A New Projection Technique for Developing a Liu-Storey Method to Solve Nonlinear Systems of Monotone Equations

Mohammed Maad **Mushtak A. K.**

of Mathematics, College of Education for Pure Science, University of Babylon,
ohmath44@gmail.com, **mttmmhh@yahoo.com**

Abstract: The projection technique is a very important method and efficient for solving unconstraint optimization and nonlinear equations. In this study, we developed a Liu-Story (LS) algorithm for solving a monotone equations of nonlinear systems. The new algorithm satisfies the sufficient descent condition and it's a suitable method of large scale equations for its limited memory. We established a global convergence of suggest method under the mild conditions. Numerical results proved that the new algorithm works well and promising.

Keyword: Projection Algorithm, Monotone Equations, Nonlinear Systems, Conjugate Gradient Method and Line Search Method.
*** Corresponding Author**

Mathematics Subject Classification 90C30.

1. Introduction

As we know, the projection approach is a very simple iterative methods to find a solution vector x^* of nonlinear systems:

$$F(x) = 0, \qquad x \in R^n, \qquad (1.1)$$

s.t. $F: R^n \to R^n$ is continuous mapping and monotonicity condition hold, i.e. $\langle F(x) - F(y),\ x - y\rangle \geq 0, \quad \forall\, x, y \in R^n$. The problem arise in various application in applied mathematics, power engineering, economics and chemical systems. For example, the variational inequality [16], the problems in proximal algorithm with Bregman distances [7], the problems of economic equilibrium [2] can be reformulated as (1.1). For solving systems of equation, there exist a many of numerical method include the Newton method [3], trust region method [14], quasi-Newton method [4], the Levenberg-Marquardt method [8], derivative-free method [10,9,13], projection technique based gradient direction [1,11,15]. Many from that approaches are iterative process begin in x_k, the next iterate is $x_{k+1} = x_k + \alpha_k d_k,\ k \in N,$ where d_k is called search direction while α_k denote step length. We mixed the conjugate gradient method [3,7] with projection algorithm [2,12,15] to be a suitable to deal with a large-scale equations. In 1964 Goldstein [5] introduced the first projection technique for convex programming in Hilbert spaces. It was then Solodov and Svaiters[15] extended Goldstein method and constructed a hyperplane $\wp_k$ that strictly separates x_k from the solution set of (1.1) i.e.

$\wp_k = \{x \in R^n \mid F(z_k)^T(x - z_k) = 0\}$

s.t. $z_k = x_k + \alpha_k d_k$ is created by employing a line search condition with the direction d_k s.t.

$F(z_k)^T(x_k - z_k) > 0.$

With the monotonicity of F, we own that for each x^* s.t. $F(x^*) = 0$.

$F(z_k)^T(x^* - z_k) \leq 0.$

Yet, by Solodov and Svaiters [15] the following approximation x_{k+1} is constructed by projecting x_k onto $\wp_k$ i.e.

$$x_{k+1} = P_{\Omega}[\, x_k - \mathfrak{H}\, F(z_k)], \qquad (1.2)$$

$$\mathfrak{H} = \frac{F(z_k)^T (x_k - z_k)}{\|F(z_k)\|^2}.$$

Where $\Omega \subseteq R^n$. be a closed convex .

Recently, the Liu-Storey(LS) conjugate gradient formula [12] is

$$d_{k+1} = \begin{cases} -F_{k+1} + \dfrac{-F_{k+1}^T y_k d_k}{-d_k^T F_k}, & k \geq 1 \\ -F_{k+1} & k = 0 \end{cases}$$

Where $y_k = F_k - F_{k-1}$. Dependent on the last method, we will suggest the developed (LS) form for solving nonlinear systems (1.1) such that

$$d_{k+1} = \begin{cases} \lambda_k F_k + \dfrac{F_k^T w_k d_k - d_k^T F_k w_k}{\max\{\theta \|d_k\| \|w_k\|, d_k^T w_k, -d_k^T F_k\}}, & k \geq 1 \\ -F_k & k = 0 \end{cases} \qquad (1.3)$$

Where $y_k = F_k - F_{k-1} + r * s_k$, $\quad s_k = z_k - x_k$, $\quad r > 0$. $\lambda, \theta > 0$ is a fixed. It is simply to offer that the our formula can be reduced to a criterion Liu-Story technique if applied exact line search. The new line search of the proposed algorithm:

$$-F(x_k + \alpha_k d_k)^T d_k \geq \frac{\tau}{2} \gamma \alpha_k \|d_k\|^2, \quad s.t.\, \alpha_k = \rho\, ℶ. \qquad (1.4)$$

Where $\rho, \tau, ℶ$ and $\gamma \in (0,1)$. We win adopt the projection based technique to present a developed (LS) gradient method with projection approach to solve monotone equations of nonlinear systems .

In our work, we discuss a developed (LS) projection algorithm for nonlinear systems (1.1). in the next section, we introduce the new algorithm with some assumptions and analysis it's the global convergence. Finally, Some numerical experiments and conclusion are presented in the last section.

2. Projection Based Method

Know, projection gradient technique is another efficient algorithm to solve large scale unconstrained minimization problem:

$$\min_{x \in R^n} f(x), \qquad (2.1)$$

where $F: R^n \rightarrow R$ is smooth nonlinear function, because of its simplicity and low storage requirements [5]. The steps of a new algorithm are stated as follows:

(2.1) New Projection Algorithm (MOH2):

1. Given initial point $x_0 \in R^n$, and parameters $s, \theta > 0$ and $\rho, \tau, \varepsilon, r, ℶ, \beta \in (0\,,1)$.

Let $k = 0$;

2. If $\|F(z_k)\| \leq \varepsilon$ break. Otherwise find d_k by (2.5).

3. determine $\alpha_k = \rho\, ℶ$, such that

$$-F(x_k + \alpha_k d_k)^T d_k \geq \frac{\tau}{2} \gamma \alpha_k \|d_k\|^2$$

4. Let $z_k = x_k + \alpha_k d_k$.

5. If $\|F(z_k)\| \leq \varepsilon$, break and set $x_{k+1} = z_k$. Otherwise compute x_{k+1} by (6)

6. Let $k = k + 1$, and return to stage (2).

3. Global Convergence Test

In this part, we investigate the global convergence of the offer approach and we need some necessary assumptions:

(B1) The solution set of nonlinear equations (1.1) is nonempty.
(B2) The mapping $F(x)$ is monotone and Lipschitz continuous of R^n, i.e., $\exists\, L > 0$, s.t.

$$\|F(x) - F(y)\| \le L\,(x - y), \quad \text{for all } x, y \in R^n. \tag{3.1}$$

In the next lemma we display that the new algorithm (MOH2) has a sufficient descent condition.

Lemma 3.1: For each $k \ge 0$, we have

$$F(x_k)^T d_k = -\lambda_k \|F_k\|^2 \tag{3.2}$$

And

$$\|d_k\| \le \left(\lambda + \frac{2}{\theta}\right) \|F_k\| \tag{3.3}$$

Proof:- When $k = 0$, then (3.2) and (3,3) holds, since $d_0 = -F(x_0)$, from the definition of d_k in (1.3), we have

$$d_{k+1}^T F_k(x_k) = -\lambda \|F(x_{k+1})\|^2 + \left[\frac{F_{k+1}^T w_k d_k - d_k^T F_{k+1} w_k}{\max\{\theta \|d_k\| \|w_k\|, d_k^T w_k, -d_k^T F_k\}}\right]^T F_{k+1} = -\lambda \|F(x_{k+1})\|^2.$$

Thus (3.2) hold for all $k \ge 1$, and

$$\|d_{k+1}\| = \left\| -\lambda F(x_{k+1}) + \frac{F_{k+1}^T w_k d_k - d_k^T F_{k+1} w_k}{\max\{\theta \|d_k\| \|w_k\|, d_k^T w_k, -d_k^T F_k\}} \right\|$$

$$\le \lambda\, \|F(x_{k+1})\| + \frac{\|F_{k+1}^T\| \|w_k\| \|d_k\| + \|d_k^T\| \|F_{k+1}\| \|w_k\|}{\max\{\theta \|d_k\| \|w_k\|, d_k^T w_k, -d_k^T F_k\}}$$

$$\|d_{k+1}\| \le \left(\lambda + \frac{2}{\theta}\right) \|F(x_{k+1})\|.$$

And the inequality follows form

$$\max\{\theta \|d_k\| \|w_k\|, d_k^T w_k, -d_k^T F(x_k)\} \ge \theta \|d_k\| \|w_k\|$$

Then (3.3) is hold.

Now, we derive some properties of algorithm(2.1) and show that the line search is well defined.
Lemma 3.2: Let assumptions(B1,B2) satisfies, impels algorithm (2.1) will produce an approximation $z_k = x_k + \alpha_k d_k$ in a limited number of backtracking procedure.

Proof:- Assume $\|F(x_k)\| \to 0$ not satisfy, or algorithm (2.1) breaks. Then $\exists\, \epsilon > 0$ satisfing $\|F(x_k)\| > \epsilon$ for all $k \ge 0$.

We used a contradiction to establish this lemma, assume that the property (1.4) not satisfy for several iteration indexes k_*.

Set $\alpha_{k_*} = \rho$ ℶ. It can be concluded

$$-F(x_{k*} + \alpha_{k*}\, d_{k*})^T d_{k*} < \frac{\tau}{2} \gamma\, \alpha_{k*}\, \|(d_{k*})\|^2,$$

By assumption B1,B2 and (3.2) in lemma (3.1), we have

$\|F(x_k)\|^2 = -F(x_k)^T d_k$

$= [F(x_{k*} + \alpha_{k*}\, d_{k*}) - F(x_k)]^T\, d_k - F(x_{k*} + \alpha_{k*}\, d_{k*})^T d_{k*}$

$= L(\alpha_{k*} d_k) d_{k*} + \frac{\tau}{2}\, \gamma\, \alpha_{k*} \|(d_k)\|^2$

$< L\, \alpha_{k*} \|(d_{k*})\|^2 + \frac{\tau}{2}\, \gamma\, \alpha_{k*} \|(d_k)\|^2$

$\|F(x_k)\|^2 = \alpha_{k*} \|(d_{k*})\|^2 \left[L + \gamma \frac{\tau}{2}\right]$

by assumption B2, we conclude that $\exists\, M > 0\ s.t.$

$$\|F(x_k)\| \leq M. \qquad (3.4)$$

Thus, we have

$$\alpha_{k*} > \frac{\|F(x_k)\|^2}{\left[L + \gamma \frac{\tau}{2}\right] \|(d_{k*})\|^2} > \frac{\epsilon^2}{\left[L + \gamma \frac{\tau}{2}\right] \left[\lambda M + \frac{2M}{\theta}\right]^2} > 0.$$

This impels contradicts with definition of α_{k_*}. So the line search (1.4) can hold a nonnegative step length α_k in a limited number of backtracking steps its well defined.

Know, the next lemma like to lemma (3.1) in Solodov and Svaiter [14], so we omit the proof.

<u>Lemma(3.3):</u> Suppose that (B1, B2) satisfies and the sequence $\{x_k, z_k\}$ is produced by algorithm (2.1), for all x^* is a solution of (1.1) $s.t.\, F(x^*) = 0,$ then

$$\|x_{k+1} - x^*\| \leq \|x_k - x^*\|^2 - \sigma^2 \|x_k - z_k\|^4. \qquad (3.5)$$

And, the sequences $\{x_k, z_k\}$ are bounded, and

$$\lim_{k\to\infty} \|x_k - z_k\| = 0. \qquad (3.6)$$

$$\lim_{k\to\infty} \|x_{k+1} - x_k\| = 0. \qquad (3.7)$$

<u>Theorem 2.4:</u> Suppose that (B1,B2) satisfies and the sequence $\{x_k\}$ be produced by algorithm (2.1), then

$$\lim_{k\to\infty} \inf \|F_k\| = 0 \qquad (3.8)$$

Proof:- Assume that (3.8) is not hold. Set a fixed $\epsilon > 0$ s.t. $\|F_k\| \geq \epsilon$ this with (3.2) implies that

$$\|d_k\| \geq \lambda \|F_k\| \geq \lambda\epsilon, \quad \forall k \geq 0 \qquad (3.9)$$

By the relation of (3.5), (3.6) and (3.7), we obtain

$$\|d_{k+1}\| \leq \left(\lambda + \frac{2}{\theta}\right) \|F_{k+1}\| \leq \left(\lambda + \frac{2}{\theta}\right) M, \forall k \geq 0 \qquad (3.10)$$

This mean a sequence $\{\|d_k\|\}$ is bounded. Then $\exists\, N_1$ infinite index and an a limit point d holds

$$\lim_{k\to\infty} \|d_k\| = d, \quad k \in N_1$$

Via the boundedness of $\{x_k\}$ in lemma (3.3), we conclude that there exists a limit point $\bar{x}$ and there is set $N_2 \subset N_1$ be infinite index holds

$$\underset{k\to\infty}{Lim} x_k = \bar{x}, k \in N_2.$$

From lemma (3.2) and lemma (3.3), we have

$$\alpha_k \|d_k\| \to 0,\ k \to \infty,$$

this together with (3.10), we get $\underset{k\to\infty}{Lim}\alpha_k = 0.$

By (1.4) we have

$$-F(x_{k*} + \alpha_{k*}\ d_{k*})^T d_k < \frac{\tau}{2}\gamma\alpha_k'\|d_k\|^2 . \qquad (3.11)$$

Where $\alpha_k' = \frac{\alpha_k}{\rho}$

Therefore, taking the lim as $k \to \infty$ in both sides of (3.11) for all $k \in N_2$ generates $F(\bar{x})^T \bar{d} > 0.$

In the other term, by making the lim as $k \to \infty$ in both terms of (3.2) for all $k \in N_2$, such that $F(\bar{x})^T \bar{d} \leq 0.$ Which generates a contradiction.

4. Numerical Experiments

In the present section, we shown the our results of numerical experiments to analyze the performance of (MOH2) and compare it with the three famous methods, a scaled derivative-free projection method (SDA)[9], a modified Liu-Story conjugate projection method (MLS)[6] and a projection based technique (DFBP)[13].

In the suggested algorithm, we used the parameters: $\rho = 0.7, \sigma = 0.1, \tau = 0.4,\ \theta = 2,$ $\beta = 0.5$ and $\epsilon = 10^{-4}$. The parameters in the LS, PDFB and UU come from [6], [13] and [9] respectively. We take on the similar finish condition for each the forth algorithms i.e. we break its when the upper number of approximation override 500000 or the inequality $\|F(x_k)\| \leq 10^{-4}$ is satisfied. All algorithms written in MATLAB program R2014a and turn on a PC(win8) , CPU 2.30 GHz and 4 GB RAM where all these method applied in the same computer.

We solved 7 constraint test problem see Awwal et.al. [2] and using 8 initial different starting point similar to the problems in [3,12], such that

$$x_0 = (10,10,\ldots,10)^T,\ \ x_1 = (-10,-10,\ldots,-10)^T, x_2 = (1,1,\ldots,1)^T,\ x_3 = (-1,-1,\ldots,-1)^T$$
$$x_4 = (1,\tfrac{1}{2},\tfrac{2}{3}\ldots,\tfrac{1}{n})^T,\ \ x_5 = (0.1,0.1,\ldots,0.1)^T, x_6 = (\tfrac{1}{n},\tfrac{2}{n},\ldots,1)^T,\ \ x_7 = (1-\tfrac{1}{n},1-\tfrac{2}{n},\ldots,0)^T.$$

The preliminary numerical experiments are reported in tables(1-2),(1-3) for number of iteration (Ni), number of function evaluations(Nf), CPU time(CPU),probability (Prob) and dimension(Dim).

Fig (1-3),(1-4) and (1-5) presented the performance profile referring to Ni, Nf and CPU time respective .It can be observed from the fig (1-3) that our proposed MOH2 method wins higher percentage of the numerical experiments. Numerical results listed in tables(2-7) show that the new method is efficient for solving problem (1.1).

The present performance profile of number of iteration in fig(1), performance number of evolutions in fig(2) and performance of CPU time in fig(3). The performance of suggested method (MOH2) it is obvious from these figures much best than LS, DFBP and UU methods. whenever, the introduced methods are efficiently and there for promising.

Table 4.1: Numerical Results

P.	Dim.	S.P	New		LS		DFBP		UU	
			Ni	Nf	Ni	Nf	Ni	Nf	Ni	Nf
P_1	20000	x_0	74	150	30	322	56	171	261	523
	20000	x_1	74	150	30	322	56	171	261	523
	20000	x_2	62	126	5	22	33	94	269	539
	20000	x_3	62	126	5	22	33	94	269	539
	20000	x_4	40	82	43	175	10	22	170	341
	20000	x_5	52	106	3	12	7	16	228	457
	20000	x_6	60	122	70	283	31	85	260	521
	20000	x_7	60	122	70	283	31	85	260	521
P_2	50000	x_0	74	150	30	322	29	60	304	611
	50000	x_1	72	146	33	368	135	412	89	180
	50000	x_2	62	126	5	22	28	58	268	538
	50000	x_3	63	128	13	101	123	372	92	187
	50000	x_4	40	82	58	239	18	38	169	340
	50000	x_5	52	106	3	12	24	50	227	456
	50000	x_6	60	122	92	374	27	56	259	520
	50000	x_7	60	122	92	374	27	56	259	520
P_3	50000	x_0	122083	964783	140299	1199292	50582	140852	80563	161129
	50000	x_1	135515	1073878	152995	1309110	55873	156143	89469	178941
	50000	x_2	99605	789574	118372	1009956	41605	115103	65675	131353
	50000	x_3	117988	935984	136851	1168990	49004	135936	77670	155343
	50000	x_4	101372	806584	120253	1025846	42163	116584	66492	132987
	50000	x_5	98007	780850	116540	993990	40850	112790	64229	128461
	50000	x_6	41822	326437	50627	432073	17769	49279	28215	56433
	50000	x_7	42438	334940	50630	432073	17702	49121	28210	56423
P_4	10000	x_0	62	220	88	665	114	322	140	283
	10000	x_1	113	546	351	2583	165	457	188	379
	10000	x_2	73	315	50	386	103	289	95	193
	10000	x_3	84	365	200	1454	137	379	148	299
	10000	x_4	61	236	157	1148	120	333	129	261
	10000	x_5	74	332	200	1457	128	354	133	269
	10000	x_6	79	362	154	1128	118	329	126	255
	10000	x_7	82	381	154	1128	119	331	126	255

Table 4.1: Numerical Results - continued

P.	Dim.	S.P	New		LS		DFBP		UU	
			Ni	Nf	Ni	Nf	Ni	Nf	Ni	Nf
P_5	10000	x_0	161	641	1218	21600	386	1057	753	1509
	10000	x_1	151	561	1255	21712	385	1055	754	1511
	10000	x_2	150	558	1220	21610	387	1059	753	1509
	10000	x_3	146	533	1218	21598	385	1055	753	1509
	10000	x_4	156	599	1218	21600	387	1059	753	1509
	10000	x_5	147	534	1218	21599	385	1055	753	1509
	10000	x_6	160	636	1218	21599	388	1061	753	1509
	10000	x_7	158	617	1218	21599	390	1065	753	1509
P_6	5000	x_0	70	142	21	178	58	169	311	624
	5000	x_1	72	146	36	359	64	187	322	646
	5000	x_2	64	130	6	30	40	115	283	568
	5000	x_3	67	136	12	80	50	145	298	598
	5000	x_4	66	134	9	55	46	133	292	586
	5000	x_5	66	134	9	53	45	130	291	584
	5000	x_6	65	132	8	45	43	124	288	578
	5000	x_7	65	132	8	45	43	124	288	578
P_7	50000	x_0	493	3130	11113	22417	19422	38854	310748	621498
	50000	x_1	64	132	30	289	408	1685	174	350
	50000	x_2	487	3120	11089	22180	19411	38824	310670	621342
	50000	x_3	63	128	6	38	224	825	122	246
	50000	x_4	49	130	636	1274	1119	2240	17977	35956
	50000	x_5	453	3045	10880	21762	19042	38086	304721	609444
	50000	x_6	475	3052	10938	21878	19147	38296	306459	612920
	50000	x_7	475	3052	10938	21878	19147	38296	306457	612916

Table 4.2: Numerical results (CPU time) - continued

P.	Dim.	S.P	CPU time			
			New	LS	DFBP	UU
P_5	10000	x_0	0.28120	11.25000	1.45312	2.45312
	10000	x_1	0.21875	9.04687	1.06250	1.50000
	10000	x_2	0.21875	8.46870	0.85937	1.37500
	10000	x_3	0.18750	5.84370	0.79687	1.20312
	10000	x_4	0.2187	5.76562	0.78125	1.00010
	10000	x_5	0.18750	5.84370	0.73437	0.85937
	10000	x_6	0.25000	5.71870	0.76562	0.78125
	10000	x_7	0.21875	5.73437	0.73437	0.78125
P_6	5000	x_0	1.07812	3.35937	3.18750	8.87500
	5000	x_1	1.07812	4.82812	2.35937	5.07812
	5000	x_2	0.92187	0.32812	1.48437	4.18750
	5000	x_3	1.03125	0.82812	1.64062	4.18750
	5000	x_4	0.79687	0.50000	1.60937	4.09375
	5000	x_5	0.93750	0.53125	1.59375	4.09375
	5000	x_6	0.98437	0.43750	1.39062	4.09375
	5000	x_7	0.87500	0.34375	1.32812	4.03125
P_7	50000	x_0	8.15625	92.17187	142.21875	1802.67180
	50000	x_1	0.54687	1.51562	6.45312	1.12500
	50000	x_2	8.03125	84.0468	136.01562	1627.96870
	50000	x_3	0.50000	0.09375	2.53125	0.67187
	50000	x_4	0.43750	4.98437	7.48437	93.81250
	50000	x_5	7.64062	82.20312	131.87500	1588.15620
	50000	x_6	7.73437	82.43750	133.81250	1661.25000
	50000	x_7	7.53125	82.26560	134.34375	1887.56250

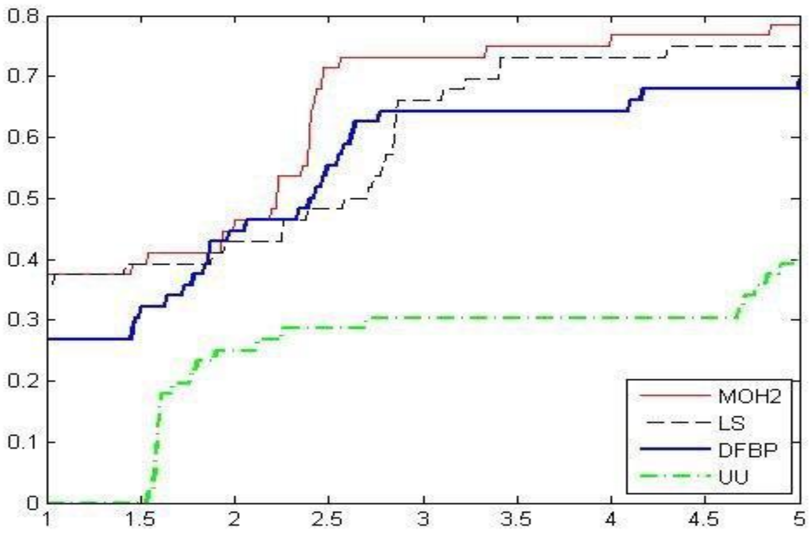


Fig1: Performance of iteration number

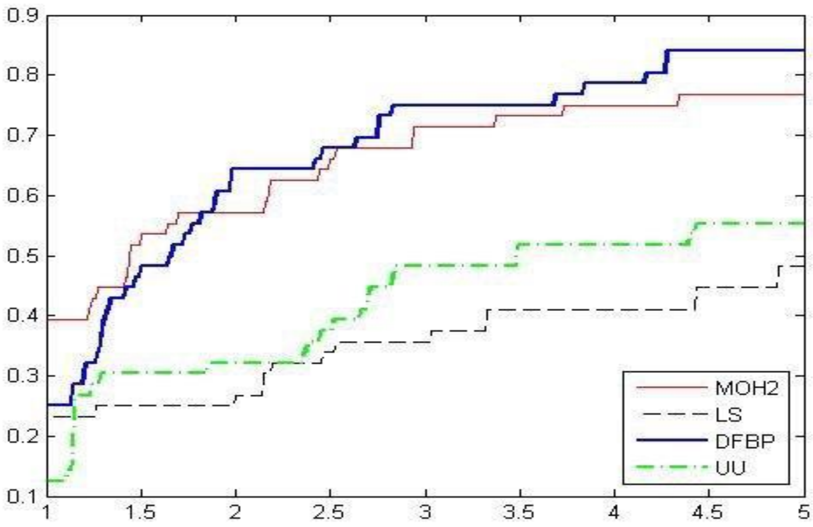


Fig2: Performance of the function evaluations

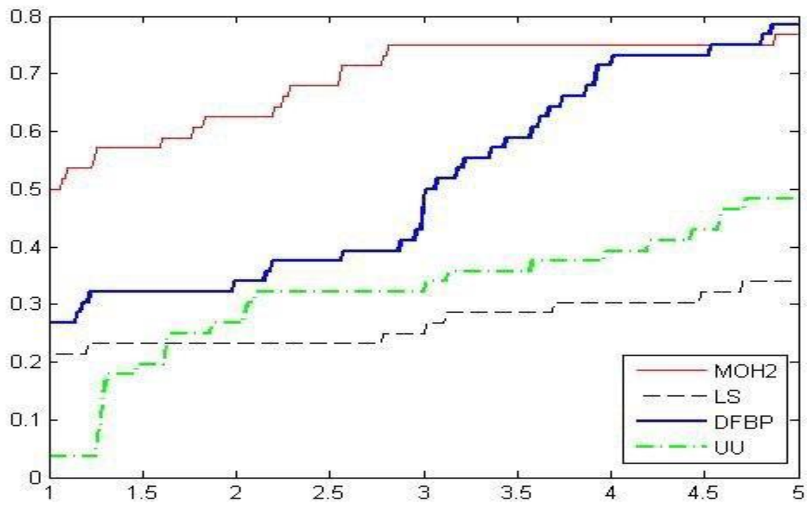


Fig3: Performance of the CPU time

5. Conclusions

In the present paper, we introduce a developed Liu-Story (LS) projection type based gradient algorithm to solve the nonlinear systems of monotone equations. The new algorithm is a suitable method of large scale equations due to its low memory requirements. The proposed method were shown to satisfying the sufficient descent condition and the global convergence was investigated with some suitable assumptions. The numerical experiments indicate that the proposed technique is efficient and very competitive.

6. References

[1] Ahookhosh, M., Amini, K., Bahrami, S. "Two derivative-free projection approaches for systems of large-scale nonlinear monotone equations". Numer. Algor. 64, 21–42 (2013).

[2] Awwal, A. M., Kumam, P., Abubakar A. B. and Wakili, A. "A projection Hestenes-Stiefel like method for monotone nonlinear equations with convex constraints", Thai Journal of Mathematics, 16, p.181-199, (2018).

[3] Barzilai, J. and Borwein, J. M., "Two-point step size gradient methods". IMA Journal of Numerical Analysis, 8(1):141{148, (1988).

[4] Dennis, J. E., More, J. J., "Quasi-Newton method, motivation and theory". SIAM Rev. 19, 46–89 (1977).

[5]. Goldstein, A.A., "Convex programming in Hilbert space", Amer. Math. Soc. 70, 709–710 (1964).

[6] Hu, Y., Wei, Z., "A Modified Liu-Storey Conjugate Gradient Projection Algorithm for Nonlinear Monotone Equations", International Mathematical Forum, Vol. 9, no. 36, 1767-1777, (2014).

[7] Iusem, A.N. and Solodov, M.V. "Newton-type methods with generalized distances for constrained optimization", Optim., 41, 257-278 (1997).

[8]. Kanzow, C., Yamashita, N. and Fukushima, M., "Levenberg-Marquardt methods for constrained nonlinear equations with strong local convergence properties", J. Comput. Appl. Math. 172, 375–397 (2004).

[9] Koorapetse, M., Kaelo1, P. and Offen1, E. R. "A Scaled Derivative-Free Projection Method for Solving Nonlinear Monotone Equations", Bulletin of the Iranian Mathematical Society, Vol. 45, p755-770 (2019).

[10]. Li, Q.N. and Li, D.H., "A class of derivative-free methods for large-scale nonlinear monotone equations". IMA J. Numer. Anal. 31, 1625–1635 (2011).

[11]. Liu, J.K. and Li, S.J., "A projection method for convex constrained monotone nonlinear equations with applications", Compt. Math. Appl. 70, 2442–2453 (2015).
[12] Liu, Y. and Storey, C. "E_cient generalized conjugate gradient algorithms", Part 1: Theory, J. Optim. Theory Appl., 69(1991), 129-137.
[13] Shiker, M. A. K. and Amini, K., "A new projection-based algorithm for solving a large scale nonlinear system of monotone equations", Croatian operational research review, crorr. 9, 63-73 (2018).
[14] Shiker, M.A.K. and Z. Sahib, "A modified technique for solving unconstrained optimization". J. Eng. Applied Sci., 13: 9667-9671 (2018).
[15] Solodov, M. V. and Svaiter, B. F., "A globally convergent inexact Newton method for systems of monotone equations", in: M. Fukushima, L. Qi (Eds.), Reformulation: Nonsmooth, Piecewise Smooth, Semismooth and Smoothing Methods, Kluwer Academic Publishers, 355-369 (1998).
[16] Zhang, L. and Zhou, W. J. "Spectral gradient projection method for solving nonlinear monotone equations", J. Comput. Appl. Math., 196, p. 478–484, (2006).

Three Terms of Derivative Free Projection Technique for Solving Nonlinear Monotone Equations

Mohammed Maad **Mushtak A. K.**

of Mathematics, College of Education for Pure Sciences, University of Babylon ohmath44@gmail.com, mttmmhh@yahoo.com

Abstract: The derivative-free projection technique is one of the efficient methods for solving nonlinear monotone equations. In this study, three terms of the derivative-free projection method with a monotone line search technique is presented. This method based on extension of a conjugate gradient descent and a developed gradient projection method to solve the nonlinear system of monotone equations. The proposed method can be used for a large scale equations due to limited memory requirement. We investigated the global convergence of the suggested approach without requiring differentiability and also the equation is Lipschitz continuous. The numerical results showed that the new algorithm is efficient and promised. **Keywords:** Projection Algorithm, Derivative-free Method, Monotone Equations, Conjugate Gradient Method and Line Search Method.

*** Corresponding Author**

Mathematics Subject Classification 90C30.

7. Introduction

We consider a derivative-free projection techniques are the most effective line search methods to solve the following nonlinear system of equations:

$$F(x) = 0, \qquad x \in R^n, \qquad (1.1)$$

s.t. $F: \Omega \subset R^n \to R^n$ be continuous and nonlinear monotone function, $\Omega \neq \emptyset$ closed convex, the monotonicity means

$$\langle F(x) - F(y), x - y \rangle \geq 0, \text{ for all } x, y \in R^n.$$

The gradient projection techniques are efficient to find the solution of large scale unconstraint optimization due to their simplicity and limited memory. A lot of computation methods have been proposed to solve unconstraint nonlinear problems. For example, Newton method, quasi newton method and Levenberg-Marquardt type method [3, 7]. A good property of the derivative-free for solving the monotone equation is that competitive with conjugate gradient descent [1, 6, 8, 11]. In this work, we developed a derivative-free projection to three terms of a derivative-free with a monotone line search technique. Also, motivated by the idea of Yuan [14], we construct a new projection method of three terms derivative-free for solving large scale systems of equations. The proposed approach used to solve a large scale systems of equations because it inherit nice properties of conjugate gradient descent such as the limited memory require and high efficient. The organized of this paper as: in section one we showed the conjugate gradient projection algorithm, in section two, we presented our algorithm with a new line search, in section three some lemma and global convergence are established and in section four we introduced the numerical experiments.

The conjugate gradient descent (CGD) is one of the important methods for solving unconstraint optimization problems and nonlinear equations. It is search direction as follows:

$$d_k = \begin{cases} -g_k & \text{if } k = 0, \\ -g_k + \beta_k \, d_{k-1} & \text{if } k \geq 1, \end{cases}$$

with g_k is a cost function f at x, $\beta_k = \{\beta_k^{HZ}, \ \eta_k\}$, here

$$\beta_k^{HZ} = \frac{1}{d_{k-1}^T y_{k-1}} \left[y_{k-1} - t d_{k-1} \frac{\|y_{k-1}\|^2}{d_{k-1}^T y_{k-1}} \right] g_k$$

$$\eta_k = \frac{-\lambda}{\|d_k\| \min\{\eta, \|g_k\|\}}, \quad y_{k-1} = g_k - g_{k-1}, \quad t = 2, \quad \lambda, \eta > 0.$$

Liu and Li [9] considered a conjugate gradient technique of Hager-Zhang [5] and suggested that the conjugate gradient descent method with $t = 1$ is high competitive than with $t = 2$. Also, Dai and Kou [4] applied the spectral technique to analyses the conjugate gradient descent method and showed that the CGD method with $t = 1$ is best than with $t = 2$. In our method, we choose the parameter $t = 1$ in the proposed algorithm.

The projection operator is a mapping $P_\Omega : R^n \to \Omega$ for all $x, y \in R^n$ it holds that [10]

$$\|P_\Omega(x) - P_\Omega(y)\| \leq \|x - y\|, \qquad (1.2)$$

where $P_\Omega(x) = argmin\{\|x - y\| | y \in \Omega\}$.

2. Algorithm:

Given an initial point x_0 an iterative scheme for (1.1) generates a sequence $\{x_k\}$ by $x_{k+1} = x_k + \alpha_k d_k$, $k \in N$ which a line search procedure employs alone the direction d_k to calculate step size α_k. Let $z_k = x_k + \alpha_k d_k$, by monotonicity of F, the hyperplane

$H_k = \{x \in R^n | \ F(z_k)^T (x - z_k) = 0\}$,

strictly separates x_k from the solution of the problem (1.1). Based on Solodov and Svaiters [12] advised that the other iteration point x_{k+1} is constructed by projecting x_k onto H_k that is x_{k+1} is determined by:

$$x_{k+1} = x_k - \frac{F(z_k)^T (x_k - z_k)}{\|F(z_k)\|^2} F(z_k). \qquad (2.1)$$

We assume that F holds some assumption as follows:

(B1) The solution set of (1.1) is nonempty.

(B2) The mapping $F(x)$ is monotone and Lipschitz continuous i.e., $\exists L > 0$, such that

$$\|F(x) - F(y)\| \leq L \, (x - y), \quad \text{for all } x, y \in R^n. \qquad (2.2)$$

We propose the following new direction formula for nonlinear monotone equations (1.1)

$$d_{k+1} = \begin{cases} -\varrho \, F_k + \beta_k^{MO} \lambda_k + \delta^T \, \beta, & k \geq 1 \\ -F_k & k = 0 \end{cases}, \qquad (2.3)$$

where $\tau = r y_k + m\lambda$, $\lambda_k = x_{k+1} - x_k$, $r, m > 0$, $y_k = F_k - F_{k-1}$, $\delta = \frac{\eta}{L + \lambda y}$,

$$\eta = \frac{\|F(z_k)\|}{1 + \|F(z_k)\|^2}, \quad \beta = \frac{\|F(x_k)\| \, y_k}{\|F(z_k)\|}, \quad \varrho = \frac{\lambda^T \lambda}{\lambda^T \tau}, \quad \beta_k^{MO} = \frac{m \, r \, F_k \, \tau}{\lambda^T \tau} - \frac{\lambda \|\tau\|^2 F_k}{(\lambda^T \tau)^2}, \quad \forall \, mr > 0.$$

Algorithm (2.1): New Projection Algorithm (MOH3):

1. The initial point $x_0 \in R^n$ is given, and parameters $\sigma, m, r > 0$ and $\rho, \gamma \in (0, 1)$.

Set $d_0 = -F_0$, $k = 0$;

2. Let $\alpha_k = \max \{\psi^i, i = 0,1,2, \ldots\}$, generated by a new line search

$$-F(x_k + \alpha_k d_k)^T d_k \geq \frac{\gamma \alpha_k \|F(z_k)\| \|d_k\|^2}{1 + \|F(z_k)\|^2}. \qquad (2.4)$$

3. Compute $z_k = x_k + \alpha_k d_k$.

4. If $\|F(x_k)\| \leq \varepsilon$ break. Otherwise calculate d_{k+1} by (2.3)

5. If $\|F(z_k)\| \leq \varepsilon$, break. Otherwise calculate x_{k+1} by

$$x_{k+1} = P_{\Omega}[\, x_k - \mathfrak{H}\, F(z_k)], \text{where} \quad \mathfrak{H} = \frac{F(z_k)^T(x_k - z_k)}{\|F(z_k)\|^2}.$$

6. Put $k = k+1$, and return to (2).

3. Global Convergence of the New Method

Remark (i): We conclude that by definitions of τ and λ that

$\lambda^T \tau = \lambda^T r y_k + m\lambda^T\lambda \geq m\|\lambda\|^2 \geq 0.$

This inequality based on the monotonicity of a mapping F, and always the divisors of β_k^{MO} and ϱ_k are greater than zero before the algorithm breaks.

The sufficiently descent property of Algorithm (2.1) showed in the next lemma.

Lemma(3.1): Suppose $\{d_k\}$ is the sequence of the search direction, $\{f_k\}$ be generated by algorithm (2.1). Then $\forall k \geq 0, \exists\, c > 0$ such that

$$F_K^T d_K \leq -c\,\|F_K\|^2 \qquad (3.1)$$

Proof: By the definition of τ and (2.2) we have

$$\lambda^T\tau \leq \|\lambda\|\|\tau\|$$

$$\leq \|\lambda\|(r\|F_k - F_{k-1}\| + m\|\lambda\|)$$

$$\leq (rL + m)\|\lambda\|^2,$$

so, we have

$$\varrho_k \geq \frac{1}{rL + m}, \qquad k \geq 1 \qquad (3.2)$$

By taking inner product (3.1) with F_k, and from (3.2) we get

$$F_K^T d_K = -\varrho_k\|F_K\|^2 + \beta_k^{New}\lambda_k F_K^T + \delta\beta F_K^T$$

$$= \left(\frac{-1}{rL + m}\right)\|F_K\|^2 + \frac{F_K^T\tau\lambda^T\tau F_K^T\lambda - \|\tau\|^2(F_K^T\lambda)^2}{(\lambda^T\tau)^2} + \left(\frac{\|F(z_k)\|}{L + \lambda\gamma}\right)\frac{\|F(x_k)\|^2}{\|F(z_k)\|}$$

$$F_K^T d_k \leq \left(\frac{-1}{rL + m} + \frac{1}{L + \lambda\gamma}\right)\|F_K\|^2 + \frac{F_K^T\tau\lambda^T\tau F_K^T\lambda - \|\tau\|^2(F_K^T\lambda)^2}{(\lambda^T\tau)^2}$$

Let $= (\lambda^T\tau)/\sqrt{2}\, F_K^T, \quad b = \sqrt{2}\,(F_K^T\lambda)\tau.$

And from $a.b \leq \frac{1}{2}(a^2 + b^2)$, we get

$$F_K^T d_k \leq \left(\frac{-1}{rL + m} + \frac{1}{L + \lambda\gamma}\right)\|F_K\|^2 + \frac{1}{4}\|F_K\|^2.$$

For $k=0$, $F_0^T d_0 = -\|F_0\|^2$

Thus (3.1) holds. □

Now, the next lemma shows that the line search of the proposed algorithms is well-defined.

Lemma(3.2): Let assumptions (B1,B2) satisfied, then there exists a step size α_k holds the line search (2.4) $\forall k \geq 0$.

Proof: Suppose that $\exists\, k_0 > 0,\ k_0$ scalar for which (2.4) is not true for all positive integer i such that:

$$-\langle F(x_{k0} + \psi\, d_{k0}, d_{k0})\rangle < \eta\,\gamma\,\psi\ \|d_{k0}\|^2,$$

by the Lipschitz continuity of F, set $i \to \infty$, then

$$-\langle F(x_{k0}),\, d_{k0}\rangle < 0. \qquad (3.3)$$

And from (3.1) we have

$$-F(x_{k0})^T\, d_{k0} \geq 0. \qquad (3.4)$$

This mean a contradiction between (3.3) and (3.4), this impels that (2.4) is well- defined. □

Lemma(3.3): Let assumptions (B1, B2) hold. The sequence $\{x_k, z_k\}$ be generated by algorithm (2.1), then for any $\bar{x}$ is a solution of (1.1) the following relation is satisfied

$$\|x_{k+1}-\bar{x}\|^2 \leq \|x_k-\bar{x}\|^2 - \gamma^2\eta^2\|x_k-z_k\|^4. \qquad (3.5)$$

Proof: By the monotonicity of *F*, we get

$$(F(z_k)-F(\bar{x}), z_k-\bar{x}) \geq 0.$$

Then

$$(F(z_k)-F(\bar{x}), x_k-\bar{x}) \geq -\langle F(z_k)-F(\bar{x}), z_k-x_k\rangle,$$

from the definition of z_k and (2.4)

$$\langle F(z_k), x_k-\bar{x}\rangle \geq -\langle F(z_k), z_k-x_k\rangle$$

$$= \alpha_k^2\eta\,\gamma\|d_k\|^2 \geq 0. \qquad (3.6)$$

From (2.1) we have

$$\|x_{k+1}-\bar{x}\|^2 = \|P_\Omega[\,x_k - \mathfrak{H}\,F(z_k)] - P_\Omega[\bar{x}]\|^2$$

$$\leq \|x_k - \mathfrak{H}F(z_k) - \bar{x}\|^2$$

$$\leq \|x_k-\bar{x}\|^2 - 2\mathfrak{H}F(z_k)^T(x_k-\bar{x}) + \mathfrak{H}^2\|F(z_k)\|^2$$

$$\leq \|x_k-\bar{x}\|^2 - \frac{F(z_k)^T(x_k-z_k).F(z_k)^T(x_k-z_k)}{\|F(z_k)\|^2}$$

$$\leq \|x_k-\bar{x}\|^2 - \alpha_k^4\eta^2\,\gamma^2\|d_k\|^4$$

$$\leq \|x_k-\bar{x}\|^2 - \eta^2\,\gamma^2\|x_k-z_k\|^4,$$

where the last three inequalities are followed from (1.2), (3.6) and (2.4) respectively. □

Remark(ii): By (3.5) we have

$$\eta^2\,\gamma^2\|x_k-z_k\|^4 \leq \|x_k-\bar{x}\|^2 -\leq \|x_{k+1}-\bar{x}\|^2.$$

It's not difficult to show that

$$\sum_{k=0}^{+\infty}\eta^2\,\gamma^2\|x_k-z_k\|^4 \leq \|x_0-\bar{x}\|^2 < +\infty.$$

Which means that

$$\lim_{k\to\infty}\|x_k-z_k\| = 0. \qquad (3.7)$$

Theorem(2.4): Let assumptions (B1,B2) satisfied and the sequence $\{F_k\}$ be determined by algorithm (2.1), then

$$\lim_{k\to\infty} \inf \|F_k\| = 0 \quad (3.8)$$

Proof: Assume that (3.8) is not hold. Let a constant $M > 0$ satisfies

$$\|F_k\| \geq M,\ \forall k \geq 0. \quad (3.9)$$

It follows from the definition of τ and (2.2) that

$$\|\tau_{k-1}\| \leq r\|F_k - F_{k-1}\| + m\lambda \leq (rl+m)\|\lambda\|.$$

From remark (i) and the definition of β_k^{MO}, we get

$$|\beta_k^{MO}| \leq \frac{mr\|\tau\|\,\|F_k\|}{\lambda^T\tau} + \frac{\|\tau\|^2\|F_k\|\|\lambda\|}{(\lambda^T\tau)^2}$$

$$\leq \frac{mr(rL+m)\|\lambda\|\,\|F_k\|}{m\|\lambda\|^2} + \frac{(rL+m)^2\|\lambda\|^2\|\lambda\|\,\|F_k\|}{m^2\|\lambda\|^4}, \forall k \geq 0$$

$$\frac{mr(rL+m)\|F_k\|}{m\|\lambda\|} + \frac{(rL+m)^2\|\lambda\|^2\|F_k\|}{m^2\|\lambda\|}$$

$$(r^2L + rm)\frac{\|F_k\|}{\|\lambda\|} + \frac{(rL+m)^2\|F_k\|}{m^2\|\lambda\|}$$

$$\left(1 + r^2L + rm + \frac{2rL}{m} + \frac{r^2L^2}{m^2}\right)\frac{\|F_k\|}{\|\lambda\|}.$$

By remark (i), we have

$$\|d_K\| \leq \frac{1}{m}\|F_K\| + |\beta_k^{MO}|\|\lambda\| + \|\delta\|\beta \quad (3.10)$$

$$\leq \left(1 + r^2L + rm + \frac{2rL}{m} + \frac{r^2L^2}{m^2}\right)\frac{\|F_k\|}{\|\lambda\|}\|\lambda\| + \frac{\|F(z_k)\|}{L + \lambda\gamma(1 + \|F(z_k)\|^2)})\frac{\|F_k\|}{\|F(z_k)\|}$$

$$= V\|F_k\| + \frac{1}{L + \lambda\gamma(1 + \|F(z_k)\|^2)}\|F_k\|$$

$$\leq \left(V + \frac{1}{L + \lambda\gamma(1 + \|F(z_k)\|^2)}\right)\|F_k\|,$$

where $V = 1 + (r^2L + rm) + \frac{2rL}{m} + \frac{r^2L^2}{m^2}$.

It follows from (2.4) that

$$-F(\bar{x}_k + \psi^{-1}\alpha_k d_k)^T d_k < \eta\,\gamma\,\psi^{-1}\alpha_k\|d_k\|^2,$$

where $\bar{z}_k = \bar{x}_k + \psi^{-1}\alpha_k d_k$.

From (3.1) and (3.2) we have

$$c\,\|F_K\|^2 \leq -F_k^T d_k$$

$$= [F(\bar{x}_k + \psi^{-1}\alpha_k d_k) - F(\bar{x}_k)]^T\, d_k - F(\bar{x}_k + \psi^{-1}\alpha_k d_k)^T d_k$$

$$\leq L(\psi^{-1}\alpha_k)\|d_k\|^2 + \psi^{-1}\eta\,\gamma\,\alpha_k\|d_k\|^2$$

i.e.

$$\alpha_k \|d_k\| \geq \frac{c\psi \|F_K\|^2}{(L+\eta\,\gamma)\|d_k\|}$$

From (3.9) and (3.10) we have

$$\alpha_k \|d_k\| \geq \frac{c\,\psi\,M}{\left(V + \frac{1}{L+\lambda\gamma(1+\|F(z_k)\|^2)}\right)(L+\eta\,\gamma)} > 0. \qquad (3.11)$$

So, by (3.5) and the definition of z_k, we get

$$\lim_{k\to\infty} \alpha_k \|d_k\| = 0$$

This implies a contradiction with (3.11), so, the assumption does not satisfied, and (3.8) holds. □

4. Numerical Experiments

Numerical results are used to assess the efficiency of the new approach (MOH3). We compare it with three famous algorithms:

(GC) which is introduced by Yan et al. [13].

(HS) which is introduced by Liu and Li [9].

(SP) which is introduced by Awwal et al. [2].

The parameter of suggested algorithm set as follows: $\gamma = 0.8,\ \psi = 0.7, m = 0.5, r = 0.001$ and $L = 0.5$. The parameter of the other methods comes from [13, 9, 2]. All Algorithms are terminated whenever $\|F_k\| \leq 10^{-6}$. The total number of iteration exceeds 500000. Our computations were carried using MATLAB R2014a and run PC with 4GH, CPU2.30-Windows8 operation system. We test the performance of the algorithm (2.1) with different initial starting points and various dimensions. Similar in [9], we check the test problem when the variables number n=5000, 10000, … with the following starting points

$x_0 = (10,10,\dots,10)^T,\ x_1 = (-10,-10,\dots,-10)^T, x_2 = (1,1,\dots,1)^T,\ x_3 = (-1,-1,\dots,-1)^T$

$x_4 = (1,\frac{1}{2},\frac{2}{3}\dots,\frac{1}{n})^T$, $x_5 = (0.1,0.1,\dots,0.1)^T, x_6 = (\frac{1}{n},\frac{2}{n},\dots,1)^T,\ x_7 = (1-\frac{1}{n},1-\frac{2}{n},\dots,0)^T$.

We compare the suggested method with the other method for number of iteration (NI), number of function evaluations (NF) and CPU time (CPU). From the fig (1, 2, 3) it's clear to show that the MOH3 is better than the other methods.

Table 4.1: Numerical results

P.	Dim.	S.P	MOH3		GC		HS		SP	
			Ni	Nf	Ni	Nf	Ni	Nf	Ni	Nf
P_1	20000	x_0	27	93	419	2546	121	244	56	114
	20000	x_1	27	93	419	2546	121	244	56	114
	20000	x_2	25	84	63	193	112	226	52	106
	20000	x_3	26	87	63	193	112	226	52	106
	20000	x_4	14	44	21	44	118	303	33	68
	20000	x_5	19	59	28	58	95	192	44	90
	20000	x_6	25	83	47	126	86	200	50	102
	20000	x_7	25	83	47	126	78	189	50	102
P_2	50000	x_0	36	150	419	2546	121	244	56	114
	50000	x_1	31	121	392	2430	41	84	21	44
	50000	x_2	34	133	63	193	112	226	52	106
	50000	x_3	28	112	61	242	33	68	18	38
	50000	x_4	22	90	21	44	120	321	33	68
	50000	x_5	32	122	28	58	95	192	44	90
	50000	x_6	35	132	47	126	91	209	50	102
	50000	x_7	35	132	47	126	91	209	50	102
P_3	50000	x_0	94564	495544	47063	126245	32933	66433	48536	121399
	50000	x_1	103235	541469	51588	138866	36768	74001	53822	134614
	50000	x_2	82718	431787	39343	104873	26378	53108	39699	99306
	50000	x_3	91101	477482	46037	123041	32357	64741	46814	117094
	50000	x_4	83186	434412	40130	106799	27718	55520	40221	100618
	50000	x_5	81715	426319	38760	103132	26736	53501	38837	97151
	50000	x_6	34923	182490	16855	44971	11456	23043	17045	42671
	50000	x_7	34925	182500	16826	44911	11567	23243	17043	42666
P_4	10000	x_0	50	244	108	418	62	127	99	273
	10000	x_1	81	388	261	1340	90	183	124	333
	10000	x_2	51	248	65	194	55	113	90	250
	10000	x_3	63	303	90	273	76	155	104	282
	10000	x_4	59	284	66	175	68	139	91	247
	10000	x_5	65	312	70	183	72	147	92	248
	10000	x_6	58	280	67	182	66	157	95	259
	10000	x_7	58	280	67	182	65	156	95	259

Table 4.1: Numerical results - continued

P.	Dim.	S.P	MOH3		GC		HS		SP	
			Ni	Nf	Ni	Nf	Ni	Nf	Ni	Nf
P_5	10000	x_0	61	247	21094	209916	2344	11705	133	268
	10000	x_1	68	276	20987	208704	2386	11804	133	268
	10000	x_2	66	268	21084	209805	2498	12442	133	268
	10000	x_3	67	272	21067	209609	2377	11912	133	268
	10000	x_4	67	272	21075	209709	2534	12640	133	268
	10000	x_5	67	272	21076	209719	2529	12634	133	268
	10000	x_6	67	272	21082	209786	2509	12550	133	268
	10000	x_7	66	268	21077	209728	2444	12208	133	268
P_6	5000	x_0	25	97	301	1550	128	258	59	120
	5000	x_1	26	99	550	3263	133	268	61	124
	5000	x_2	24	87	95	350	117	236	54	110
	5000	x_3	24	79	174	788	123	248	57	116
	5000	x_4	25	96	125	500	121	244	56	114
	5000	x_5	25	89	122	486	120	242	55	112
	5000	x_6	27	102	112	435	119	240	55	112
	5000	x_7	24	93	112	435	119	240	55	112
P_7	50000	x_0	15550	31145	44711	90690	129477	258956	118747	237496
	50000	x_1	789	1960	471	2844	71	144	36	74
	50000	x_2	15528	31058	44380	88766	129445	258892	118732	237466
	50000	x_3	647	1553	48	192	54	110	29	60
	50000	x_4	896	1794	2565	5132	7490	14982	3190	6382
	50000	x_5	15234	30470	43530	87062	126968	253938	117676	235354
	50000	x_6	15330	30662	43775	87552	128794	257609	117633	235268
	50000	x_7	15330	30662	43775	87552	128634	257289	117632	235266

Table 4.2: Numerical results (CPU time)

P.	Dim.	S. P	CPU time			
			MOH3	GC	HS	SP
P_1	20000	x_0	1.54687	16.59375	2.89062	0.50000
	20000	x_1	1.10937	16.46875	2.35937	0.39062
	20000	x_2	1.09375	0.84375	1.60937	0.32812
	20000	x_3	0.96875	0.92187	1.43750	0.35937
	20000	x_4	0.54687	0.21875	1.81250	0.23437
	20000	x_5	0.59375	0.25000	1.42187	0.29687
	20000	x_6	1.00001	0.65625	1.14062	0.37500
	20000	x_7	1.10937	0.59375	0.98437	0.28125
P_2	50000	x_0	0.57812	16.37500	2.79687	0.48437
	50000	x_1	0.40625	15.32812	0.82812	0.14062
	50000	x_2	0.48437	0.95312	2.39062	0.35937
	50000	x_3	0.31250	1.07812	0.64062	0.12500
	50000	x_4	0.34375	0.23437	2.35937	0.23437
	50000	x_5	0.48437	0.26562	2.06250	0.29687
	50000	x_6	0.51562	0.60937	2.20312	0.20312
	50000	x_7	0.48437	0.60937	2.15625	0.32812
P_3	50000	x_0	2118.89062	494.59375	286.39062	394.71875
	50000	x_1	2352.82812	452.89062	345.68750	439.18750
	50000	x_2	1863.10937	336.79687	242.79687	319.79687
	50000	x_3	2064.51562	396.03125	297.95312	376.59375
	50000	x_4	1871.04687	341.31250	255.04687	327.46875
	50000	x_5	1835.53125	331.34375	246.32812	313.59375
	50000	x_6	783.84375	144.21875	103.96875	137.81250
	50000	x_7	781.85937	144.04687	109.07812	181.32812
P_4	10000	x_0	0.14062	0.34375	0.65620	0.20312
	10000	x_1	0.25000	1.43750	0.93750	0.25000
	10000	x_2	0.15625	0.20312	0.57812	0.15625
	10000	x_3	0.18750	0.32812	0.20312	0.17187
	10000	x_4	0.15625	0.17187	0.34375	0.15625
	10000	x_5	0.18750	0.23437	0.26562	0.17187
	10000	x_6	0.17187	0.18750	0.20312	0.18750
	10000	x_7	0.10937	0.25000	0.31250	0.20312

Table 4.2: Numerical results (CPU time) - continued

P.	Dim.	S.P	CPU time			
			MOH3	GC	HS	SP
P_5	10000	x_0	0.17187	80.87500	12.76562	0.62500
	10000	x_1	0.10937	79.59375	12.84375	0.35937
	10000	x_2	0.14062	80.92187	20.25000	0.37500
	10000	x_3	0.10937	79.78125	23.26562	0.43750
	10000	x_4	0.09375	80.65625	24.42187	0.35937
	10000	x_5	0.10937	79.14062	12.0312	0.35937
	10000	x_6	0.10937	81.15625	12.23437	0.35937
	10000	x_7	0.07812	80.25000	11.59375	0.32812
P_6	5000	x_0	0.76562	12.10937	3.53125	2.85937
	5000	x_1	0.62500	24.17187	3.53125	2.03125
	5000	x_2	0.57812	2.54687	3.04687	1.57812
	5000	x_3	0.53125	6.29687	3.28125	1.53125
	5000	x_4	0.67187	3.96875	3.01562	1.23437
	5000	x_5	0.56250	3.62500	3.43750	1.04687
	5000	x_6	0.73437	3.39062	3.62500	1.01562
	5000	x_7	0.68750	3.57812	3.07812	0.90625
P_7	50000	x_0	151.26562	392.12500	1651.37500	15330.68750
	50000	x_1	8.32812	16.64062	1.07812	4.82812
	50000	x_2	148.37500	367.04687	1682.43750	15549.17187
	50000	x_3	6.62500	0.70312	0.70312	3.82812
	50000	x_4	8.39062	20.78125	97.00001	418.82812
	50000	x_5	147.18750	362.89062	1624.60937	15456.39062
	50000	x_6	148.45312	357.81250	1672.00001	15450.40625
	50000	x_7	150.73437	359.92187	1676.07812	15431.67187

5. Conclusions

We have suggested a new class of three terms of derivative-free projection technique for solving unconstraint optimization. The suggested approach is appropriate for large scale equations, because it has a nice property which is the low memory requirement. The global convergence of our method is established. The numerical results showed that our method is efficient and working better than the three other algorithms that the new algorithm is compared with.

6. References

[1] Amini, K., Shiker M.A.K., Kimiaei, M. "A line search trust-region algorithm with nonmonotone adaptive radius for a system of nonlinear equations". 4 OR- Journal of operation research. 2016; 14(2) : 133-152.

[2] Awwal, A. M., Kumam, P., Abubakar A. B. and Wakili, A. "A projection Hestenes-Stiefel like method for monotone nonlinear equations with convex constraints", Thai Journal of Mathematics, 16, p.181-199, (2018).

[3] Cheng, W. Y., Li, D H. 'Spectral scaling BFGS method', J. Optim. Theor. Appl. 146, 305–19 (2010).

[4] Dreeb, N. K., Hashim, K. H., Mahdi, M. M., Wasi, H. A., Dwail, H. H., Shiker, M. A. K. and Hussein, H. A. " Solving a Large-Scale Nonlinear System of Monotone Equations by using a Projection Technique". Journal of Engineering and Applied Sciences, 14: 10102-10108 (2019).

[5] Hager, W. W. and Zhang, H. A new conjugate gradient method with guaranteed descent and an efficient line search. SIAM J. Optim. 16, 170–92 (2005).

[6] Hashim, K. H., Dreeb, N. K., Dwail, H. H., Mahdi, M. M., Wasi, H. A., Shiker, M. A. K. and Hussein, H. A. "A New Line Search Method to Solve the Nonlinear Systems of Monotone Equations". Journal of Engineering and Applied Sciences, 14: 10080-10086 (2019).

[7] Li, Q.N. and Li, D.H., "A class of derivative-free methods for large-scale nonlinear monotone equations". IMA J. Numer. Anal. 31, 1625–1635 (2011).

[8] Liu, J. K. "Derivative-free spectral PRP projection method for solving nonlinear monotone equations with convex constraints". Math Numer Sin 38, 113–24 (2016).[in Chinese].

[9] Liu, J.K. and Li, S.J., "Spectral DY-Type projection method for nonlinear monotone systems of equations ", JCM. no 4, 341–354 (2015).

[10] Shiker, M. A. K. and Amini, K., "A new projection-based algorithm for solving a large scale nonlinear system of monotone equations", Croatian operational research review, crorr. 9, 63-73 (2018).

[11] Shiker, M. A. K. and Z. Sahib, "A modified technique for solving unconstrained optimization". J. Eng. Applied Sci., 13: 9667-9671 (2018).

[12] Solodov, M. V. and Svaiter, B. F., "A globally convergent inexact Newton method for systems of monotone equations", in: M. Fukushima, L. Qi (Eds.), Reformulation: Nonsmooth, Piecewise Smooth, Semismooth and Smoothing Methods, Kluwer Academic Publishers, 355-369 (1999).

[13] Yan, Q. R., Peng, X. Z. and Li, D. H. "A globally convergent derivative-free method for solving large-scale nonlinear monotone equations", J. Comput. Appl. Math., 234, p. 649-657, (2010).

[14] Yuan, N. A, "A derivative-free projection method for solving convex constrained monotone equations", Science Asia, 43, p. 195-200, (2017).

A Modification to Vogel's Approximation Method to Solve the Transportation Problems

Hussein Ali Hussein[1] and Mushtak A.K. Shiker[2*]
1,2Department of Mathematics, College of Education for Pure Sciences,
University of Babylon, Babil-Iraq
1hussainalseaidi@gmail.com, 2mmttmmhh@yahoo.com

Abstract: Transportation Problem (TP) is one of the paradigms in the Linear Programming Problem (LPP). The TP in Operations Research represent vastly applied optimization. (TP) has some goals, like reducing transportation costs or reducing transportation time, etc. Whereas meeting both supply level and request level requirements. Transportation problem plays a major role in industry, trade, logistics, etc. To get the most possible profit, organizations are always looking for better ways to reduce cost and improve revenue. To solve the transportation problem, it is always required to find an initial basic feasible solution (IBFS) for get the optimal solution. The Vogel's Approximation Method (VAM) is the important known traditional methods for obtaining an IBFS of TP. In this work, we introduce a new modification to the VAM to obtain an IBFS for the transportation problems almost nearer to the optimal solution. Proposed modification is illustrated with solved numerical examples. A comparison study was also conducted with the results of classic methods. This modified approach most of times give better solution and very close to the optimal solution, furthermore, sometimes gives the optimal solution. This method is clear, easy to comprehend.

Keywords: Operation Research, Liner Programming, Initial Basic Feasible Solution, Vogel's Approximation Method.

*** Corresponding Author**

1. Introduction

Transportation Problems (TP) are single of oldest and ultimate remarkable implementations of the linear programming problems (LPP). TP is a special category of linear programming, whose relates of daily effectiveness in our lives [1, 3, 6]. In the nature of public life, a nominated quantities of homogeneous goods is obtainable in a numeral of provenance, and a stationary quantities is desired to satisfy the request in every of consumption venues. Transportation models focus fundamentally on the optimal method that is the product from several plants or factories (it is supply assets) be transferred into a several of storehouse or clients (it is demand destinations) [2, 8, 11]. In this problem, the main goal is to find an optimal schedule for shipping the commodity while satisfying the demands in every destination. The TP was first introduced in 1941by Frank L. Hitchcock [4]. In 1947, Tjalling C. Koopmans, offered his paper to solve (TP) [14]. The two mentioned studies are the main achievements in the advancement of the different techniques for solving the transportation model. The transportation problems can also be solved by expressing it as an LP model via simplex technique provided by Dantzig G.B. in 1951 however it includes a considerable volume of variables and restrictions, and solve them by simplex technique requires lot of effort and a protracted time. Many investigators have expanded substitutional methods to find an IBFS that holds transportation costs at consideration. An IBFS for the TP be acquired using one of three classic techniques, north-west corner method (NWCM), minimum cost method (MCM), vogel approximation method (VAM). At those three approaches, Vogel's Approximation Method (VAM) method is preferable deem to the studies. We can improve results of initial solution to get the optimal solution to the transportation problem in one of two ways, namely the Modified Distribution Method (MODI) or Stepping Stone Method. Essentially, these methods differ mainly in terms of the best solutions in the beginning, and a good solution that has a beginning will produce a smaller objective value. The transportation problem is divided into two kinds, the balanced transportation problem and the unbalanced transportation problem. If the numeral of sources is equivalent to the numeral of requests, we say that is a balanced transportation problem. If not, we say the problem of unbalanced transportation. In recent years, several methods have been proposed to find IBFS for the transportation model. Of Implied Cost Method (ICM), Md. Ashraful Babu and others (2014) have shown that their method is better or similar to VAM. Abdul Sattar Soomro and others (2015) in their paper, modified (VAM). In (2017), Mollah Mesbahuddin Ahmed and others gave an innovative approach to obtain an (IBFS) for the (TP) [9]. In (2018) , Ravi Kumar and others proposed a new approach to find (IBFS) for the (TP) [11]. Also in (2018), Lakhveer kaur and others discussed an improvement in maximum difference method to find (IBFS) for the (TP). In (2019), S. C. Zelibe and C. P. Ugwuanyi developed a new solution of the transportation problem. In this work, we introduce a new modification for (VAM) in which the resulting of objective function is almost ideal and better or equal to the results of the solution according to the (VAM), but in any case the results of our proposed method are much better than the results of the North-West Corner Method and Minimum Cost Method.

2. Preliminaries

2.1 Formula of (TP)

Notations:

We will present the mathematical model. After clarifying the meaning of the following symbols:

m: Sources.

S_i , $(i = 1,2, \dots, m)$: Supply.

a_i , $(i = 1,2, \dots, m)$: Available quantities of each m capacity.

n: Destinations.

D_j , $(j = 1,2, \dots, n)$: Demand.

b_j , $(j = 1,2, \dots, n)$: Available quantities for each n requirements.

c_{ij}: The cost of transshipme one quantity of goods than origin i to destination j at every path.

x_{ij}: The numeral of quantities shipped in every path than origin i to destination j.

Model:

We can express the mathematical model as follows:

$$\text{Minimize } Z = \sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij} \, x_{ij} \qquad \text{(Overall cost) (1)}$$

Subject:

$$\sum_{j=1}^{n} x_{ij} = a_i \quad , \quad i = 1,2, \dots \dots m \text{ (supply constraints)} \quad (2)$$

$$\sum_{i=1}^{m} x_{ij} = b_j \quad , \quad j = 1,2, \dots \dots n \text{ (demand constraints)} \quad (3)$$

Where $x_{ij} \geq 0 \quad , \forall\, i \text{ and } j \qquad (4)$

2.2 Representation of Transportation Problems (TP)

Transportation problems (TP) represent a particular model of linear programming problems (LPP). Transportation problem is shipping different quantities of homogeneous goods than various sources (e.g. factories) to various destinations (e.g. warehouse) so that the total transportation cost (or time) is minimized.

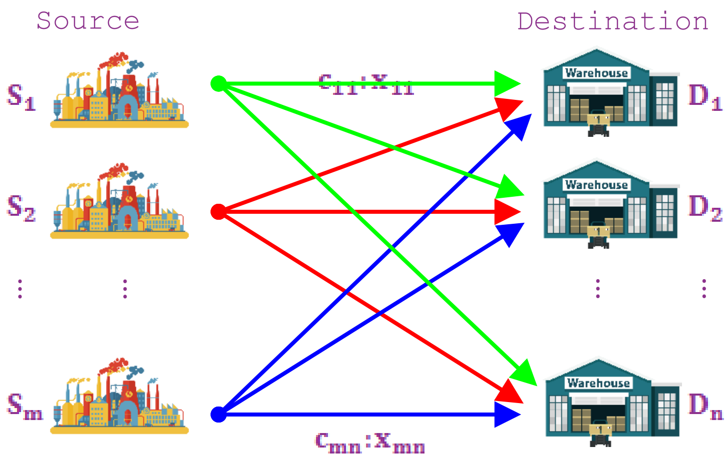


Fig. 1: The graph structure of TP

Table ①: Schedule of TP

Sources (i)	Destinations (j)				Supply (a_i)
	D_1	D_2	...	D_n	
S_1	x_{11} c_{11}	x_{12} c_{12}	...	x_{1n} c_{1n}	a_1
S_2	x_{21} c_{21}	x_{22} c_{22}	...	x_{2n} c_{2n}	a_2
⋮	⋮	⋮	x_{ij} c_{ij}	⋮	⋮
S_m	x_{m1} c_{m1}	x_{m2} c_{m2}	...	x_{mn} c_{mn}	a_m
Demand (b_j)	b_1	b_2	...	b_n	$\sum a_i = \sum b_j$

2.3 Some Definitions

① **Feasible solution (F.S.)**

A feasible solution to transportation problem is a collection of non negative assigned quantities $x_{ij} \geq 0$ that cater for restrictions (supply constraints and demand restrictions in transportation problem).

② **Basic Feasible Solution (B.F.S.)**

We say that a feasible solution of transportation problems is basic feasible solution (BFS) when it includes no most from (m+n-1) positive assigned quantities, such that m represent the numeral of rows while n represent the numeral of columns of TP.

3. General Procedure to Solve a Transportation Problem

Phase1: Mathematical formulation and table of TP.

Phase2: Find an IBFS.

Phase3: Modified the initial basic feasible solution (IBFS) obtained in Phase2 to find the optimal solution.

4. Methods for Finding IBFS of TP

The initial solve that we obtain from the traditional solution approaches or the new solution approaches should be distinguished by the following:

1) The solve should be feasible.

2) It must satisfy the non-passivity restriction.

3) The solution should be basic.

The classic method that used to find the initial solution of TP is:

🅐 North-West Corner Method (NWCM).

🅑 Minimum Cost Method (MCM).

🅒 Vogel's Approximation Method (VAM).

5. The new Algorithm to find (IBFS):

We can find the initial basic feasible solution (IBFS) to the transportation problem (TP), using the new suggestion method (Al-Saeedi's Method) according to the following solution steps:

Step 1: Build the transportation tableau depending on the given TP. Checking whether the overall supply equivalent the overall request, if not, introduce an unreal row (or column) (dummy supplier or add dummy demander) and every transportation cost c_{ij} of that row / column is equal to zero, i.e. balance the transportation problem.

Step 2: For every row of the transportation table, we determine the two lowest costs available. We find the difference between these two (called penalty) and put it to the right of that row in a new column. For every column of the transportation table, we determine the two highest costs available. We find the difference between these two (called penalty) and put it below that column in a new row.

Step 3: Of all the difference values shown in the two penalties, in step 2, we choose the largest value (the biggest difference).

Step 4: We assign the extreme possible quantities for the lower cost cell of the specified row (or column). In event of an equalize between the biggest differences, this row (or column), whose include the lowest cost, may be chosen. If there is an equalize at lowest cost also, a choice can be made from that row (or column) where maximum requirements are exhausted. when an equalize at allocating maximum requirements we take cell of row (or column) has maximum supply (or demand). The selected cell is assigned to that unit and the corresponding depleted row (or column) is removed from further study.

Step 5: We exclude the row (or column) that satisfied supply or request.

Step 6: Iterate steps 2 - 5 till every columns and rows are contented.

Step 7: Compute the overall cost of the TP. The overall cost of the TP can be calculated by applying the following equation $Z = \sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij} x_{ij}$.

6. Numerical Examples

Ex①

Depending to the data of the following table:

Table ②

Sources	Destinations				Supply
	D_1	D_2	D_3	D_4	
S_1	7	5	9	11	30
S_2	4	3	8	6	25
S_3	3	8	10	5	20
S_4	2	6	7	3	15
Demand	30	30	20	10	90

The given transportation table is balanced since total supply = total demand = 90.

According to the new Algorithm we get:

Table ③

Sources	D_1		D_2		D_3		D_4		Supply	Penalty				
	allocation	cost	allocation	cost	allocation	cost	allocation	cost						
S_1	0	7	10	5	20	9	0	11	30	2	2	2	2	4
S_2	5	4	20	3	0	8	0	6	25	1	1	1	1	5
S_3	20	3	0	8	0	10	0	5	20	2	5	-	-	-
S_4	5	2	0	6	0	7	10	3	15	1	4	4	-	-
Demand	30		30		20		10		90					
Penalty	3		2		1		5							
	3		2		1		-							
	3		1		1		-							
	3		2		1		-							
	-		2		1		-							

The total cost of transportation is $Z = \sum_{i=1}^{4}\sum_{j=1}^{4} c_{ij}\, x_{ij}$

$\therefore\ Z = (10\times 5) + (20\times 9) + (5\times 4) + (20\times 3) + (20\times 3) + (5\times 2) + (10\times 3) \ = \ 410$ units

Ex②

For another example, depending to the data of the following table:

Table ④

Sources	D_1	D_2	D_3	Supply
S_1	3	6	2	30
S_2	7	8	10	50
S_3	12	1	4	20
S_4	2	5	9	60
Demand	65	40	55	160

The given transportation table is balanced since total supply = total demand = 160.
According to algorithm of H. Al-Saeedi's First Method we get:

Table ⑤

Sources	D_1		D_2		D_3		Supply	Penalty		
	Destinations									
S_1	0	3	0	6	30	2	30	1	1	-
S_2	5	7	20	8	25	10	50	1	1	1
S_3	0	12	20	1	0	4	20	3	3	3
S_4	60	2	0	5	0	9	60	3	-	-
Demand	65		40		55		160			
Penalty	5		2		1					
	5		2		6					
	5		7		6					

The total cost of transportation is $Z = \sum_{i=1}^{4} \sum_{j=1}^{3} c_{ij}\, x_{ij}$

$\therefore\ Z = (30 \times 2) + (5 \times 7) + (20 \times 8) + (25 \times 10) + (20 \times 1) + (60 \times 2) \ = \ 645$ units

7. Results Analysis

Table ⑥

Name	NWCM	LCM	VAM	Al-Saeedi's Method
Ex①	540	435	415	410
Ex②	995	645	645	645

It is clear from Table ⑥ that the new method (Al-Saeedi's Method) gives the best IBFS comparing with the classical three methods.

8. Conclusion

It is extremely important to have an initial basic feasible solution (IBFS) for a balanced transportation problem. It is known that the Vogel's Approximation Method is the best. We have modified an efficient heuristic procedure over VAM for solving TP. The proposed algorithm gives comparatively a better initial basic feasible solution to the transportation problem compared to solutions obtained by classical algorithms that are either the optimal solve or almost the optimal solve. The method proposed in this paper can achieve a great deal of success in solving transport problems, and it is effective for both large and small sizes. This method is very profitable for decision makers whose deal with supply chain and logistics matters. The proposed method (Al-Saeedi's Method) is very easy to understand, involves simple calculation and thus saves time. Hence, this method may be preferred over the other existing methods.

9. References

[1] Abdul Sattar Soomro; Muhammad Junaid, and Gurudeo Anand Tularam. *Modified Vogel's Approximation Method for Solving Transportation Problems*. Mathematical Theory and Modeling, Vol.5, No.4, PP:32-42, (2015).

[2] Bilqis Amaliah; Chastine Fatichah, and Erma Suryani. *Total Opportunity Cost Matrix – Minimal Total: A New Approach to Determine Initial Basic Feasible Solution of a Transportation Problem*. Egyptian Informatics Journal, PP:1-11, (2019).

[3] Dantzig G.B. *Application of the simplex method to a transportation problem, in Activity Analysis of Production and Allocation* (T.C. Koopmans, ed.) Wiley, New York, PP:359-373, (1951).

[4] Frank L. Hitchcock. *The Distribution of a Product from Several Sources to Numerous Localities*. Journal of Mathematics and Physics, Vol. 20, PP:224-230, (1941).

[5] K. Dhurai and A. Karpagam. *To Obtain Initial Basic Feasible Solution Physical Distribution Problems*. Global Journal of Pure and Applied Mathematics, Volume 13, Number 9, PP:4671-4676, (2017).

[6] Lakhveer kaur; Madhuchanda Rakshit, and Sandeep Singh. *An Improvement in Maximum Difference Method to Find Initial Basic Feasible Solution for Transportation Problem*. International Journal of Computer Sciences and Engineering, Vol. 6, Issue 9, PP: 533-535, (2018).

[7] Madhavi.Malireddy. *A New Algorithm for Initial Basic Feasible Solution of Transportation Problem*. International Journal of Engineering Science Invention (IJESI), Volume 7, Issue 8, Ver IV, PP:41-43, (2018).

[8] Md. Ashraful Babu; Md. Abu Helal; Mohammad Sazzad Hasan, and Utpal Kanti Das. *Implied Cost Method (ICM): An Alternative Approach to Find the Feasible Solution of Transportation Problem*. Global Journal of Science Frontier Research: F Mathematics and Decision Sciences, Volume 14, Issue 1, Version 1.0, PP:4-13, (2014).

[9] Mollah Mesbahuddin Ahmed; Aminur Rahman Khan; Md. Sharif Uddin, and Faruque Ahmed. *A New Approach to Solve Transportation Problems*. Open Journal of Optimization, 5, PP:22-30, (2016).

[10] Mollah Mesbahuddin Ahmed; Nahid Sultana; Aminur Rahman Khan, and Md. Sharif Uddin. *An Innovative Approach to Obtain an Initial Basic Feasible Solution for The Transportation Problems*. Journal of Physical Sciences, Vol. 22, PP:23-42, (2017).

[11] Ravi Kumar R; Radha Gupta, and Karthiyayini O. *A New Approach to Find the Initial Basic Feasible Solution of a Transportation Problem*. International Journal of Research – GRANTHAALAYAH, Vol. 6, Iss. 5, PP:321-325, (2018).

[12] Reena G. Patel; Bhavin S. Patel, and P. H. Bhathawala. *On Optimal Solution of a Transportation Problem*. Global Journal of Pure and Applied Mathematics, Volume 13, Number 9, PP:6201-6208, (2017).

[13] S. C. Zelibe and C. P. Ugwuanyi. *On A New Solution of The Transportation Problem*. Journal of the Nigerian Mathematical Society, Vol. 38, Issue 2, PP:271-291, (2019).

[14] Tjalling C. Koopmans. *Optimum Utilization of the Transportation System*. Proceeding of the International Statistical Conference, Washington D.C. (1947).

Employ the Principle Components in the Detection of Feedback

Dr.Heyam A.A Hayawi

College of Computer Sciences and Mathematic/Dept.of Statistics and Informatics University of Mosul/Iraq

Abstract: In linear dynamic feedback models ,the relationship between two input series input and output shows that the approach is a propose to discover a feedback in linear dynamic system through examining autocorrelation function and partial autocorrelation function for principle components by using few tests of time series identification and exact using Ljung-Box test depending on Simulation approach to show the efficiency of the suggested approach and application on linear and nonlinear models within and without feedback, The obtained result is good and encourage.

Keyword: principle Component, State Space, Feedback

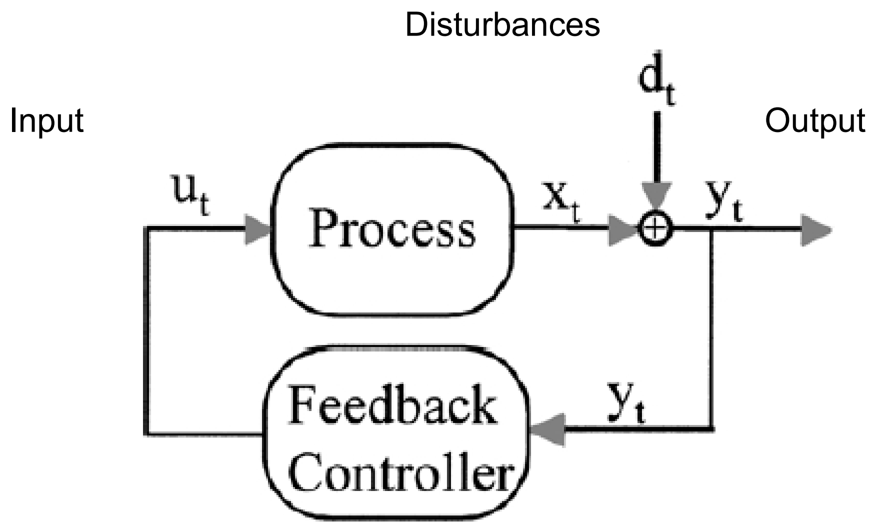
1.Introcuction

A model is defined as a hypothetical description, an easy representation, or a description designed for a particular process entity. Scientifically, the model is a mathematical or logical representation of a system of entities, phenomena, or processes. The system is influenced by external influences processed by the user and called inputs or triggers, including non-controlled signals known as disturbances or disturbances and the relationship between them is determined by the conversion function that reflects the change in the Acetate to turn into outputs (Khayat 2010).

The control system is described as open-loop. If the control work is independent of the output, otherwise it is closed-loop. In the case of closed-circuit control systems, the inputs to the system are based on a synchronous basis with output observations such as autopilot and aircraft control in the air.

Modeling is the process of building models that summarize the conceptual or graphical phenomena of a particular entity or system. Useful, reliable and insightful and able to distinguish whether the model reflects the truth and deals with deviations between theory and data (Al-Khayat, 2011).

Examining the data to ensure the presence of feedback between more than one time series, ie the input chain and the output chain, is necessary to know the appropriate kinetic model that should be used to obtain the best estimates of the parameters of the model. This measure is then used to identify the inputs that subsequently reduce the error and then improve the performance. Undoubtedly, feedback is a way of describing and understanding the indirect interference found in specific physical systems. The idea of reverse feeding can be illustrated in Figures (1) and (2) as follows:



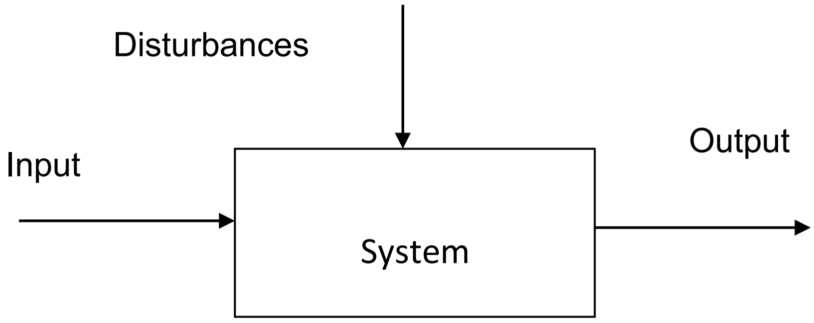


Figure (2): Lack of a feedback system in a motor system

In t ... ive self-correlation analysis were used for the output data series as well as the data of the first main component to detect the presence of the feedback or not by adopting the simulation method with linear and non-linear models.

1. AutoCorrelation Function(ACF)

The self-correlation coefficient is the statistical key in the analysis of the time series because it represents a measure of the correlation strength between the observations of the same random variable at k time periods and is calculated as shown in equation (1) as follows:(Makridakis,et al.,1998):

$$r_k = \frac{\sum_{t=k+1}^{n}\left(y_t - \bar{y}\right)\left(y_{t-k} - \bar{y}\right)}{\sum_{t=1}^{n}\left(y_t - \bar{y}\right)^2} \qquad , k = 0,1,2,\ldots,(n-1) \qquad \ldots(1)$$

2. Partial AutoCorrelation Function (PACF)

It is used to measure the correlation between Yt and Yt-k string values with the Yt value constant for the rest of the periods, ie to measure the degree of correlation between Yt and Yt-k by determining the effect of other values at other displacements and is calculated as given in equation (2) as follows:(Makridakis,et al.,1998):

$$\left.\begin{aligned} \phi_{kk} &= \frac{\rho_k - \sum_{j=1}^{k-1}\phi_{k-1,j}\ \rho_{k-j}}{1 - \sum_{j=1}^{k-1}\phi_{k-1,j}\ \rho_j} \\ \phi_{k,j} &= \phi_{k-1,j} - \phi_{kk}\ \phi_{k-1,k-j} \qquad , j = 1,\cdots,k-1 \end{aligned}\right\} \qquad \ldots(2)$$

٣. Principal Components Analysis

The method of analyzing the main components is an exploratory method and is preferred to simplify the description of a group of related variables that are treated equally. Be the number of original variables So that each major component is a linear structure of the original variables, and these components are eligible to explain most of the total variance, so they are arranged in descending order according to their variability, that is, what the first major component explains more than the second main component and the information that the second main component explains is greater than the main component. Third and so on to the rest of the other major components. (Afifi ,1984),(Al-Rawi,1987)

Key Component Analysis is a tool for coordinate Axis Transformation and Dimensionality Reduction, in which a new set of coordinates is calculated by maximizing the contrast of the sample data points with those coordinates.(Nelles ,2001)

In order to clarify and simplify the concept of the main components, two variables are taken and let X1 and X2 be present with N of observations, as the data of the two variables are converted to standard data by subtracting the sample mean from each view to convert to data with zero mean and variance respectively: - (Afifi, 1984)

$$\left.\begin{aligned} x_1 &= X_1 - \bar{X}_1 \\ x_2 &= X_2 - \bar{X}_2 \end{aligned}\right\} \quad \dots (3)$$

The basic idea of creating the main components is to obtain two new variables, C1 and C2, and each of them are linear functions of the two variables respectively, that is:-

$$\left.\begin{aligned} C_1 &= a_{11}x_1 + a_{12}x_2 \\ C_2 &= a_{21}x_1 + a_{22}x_2 \end{aligned}\right\} \quad \dots (4)$$

As:

The coefficients for the main components, i.e. the eigenvalues, represent Egien Values, and the mean of the two main components C1 and C2 is: -

$$mean\,C_1 = mean\,C_2 = 0 \quad \dots (5)$$

While contrasting

$$\left.\begin{aligned} Var\ C_1 &= a_{11}^2 S_1^2 + a_{12}^2 S_2^2 + 2a_{11}a_{12}rS_1S_2 \\ Var\ C_2 &= a_{21}^2 S_1^2 + a_{22}^2 S_2^2 + 2a_{21}a_{22}rS_1S_2 \end{aligned}\right\} \quad \dots (6)$$

As:

$S_1^2 = Var\ X_1 \quad , S_2^2 = Var\ X_2$

The main component parameters are chosen to fulfill three requirements: -

1- Variation of the first component, C1, is as large as possible.
2- The values of observations in the major components C1 and C2 are not related.
3- $a_{11}^2 + a_{12}^2 = a_{21}^2 + a_{22}^2 = 1$.

4. **Methods of detection of feedback**

There are many methods for detecting feedback, including:

1. Cross Correlation Function: The cross correlation function is an important analytical tool that shows the strength of the link between two observations, each of which belongs to a separate time series from the other, the first is called the input chain and the second is called the output chain, and the cross-link measures the relationship between the current values For the output chain and between the past and current values of the input chain, when modeling kinematic systems the input and output chains are refined and then the cross-link between the two purified series is observed and the values of the cross-link function are observed, so if the order function appears The cross-over between the two purified series is at least one significant value at a given displacement, indicating a reverse between the two series (Wei, 1990).
2. **ARMA Vector Operations**
ARMA model vector operations are an extension of single variable ARIMA models because they contain feedback between two or more series. To illustrate these models, we assume we have two series and that each series is a linear equation for its previous values at displacement 1 and previous values for another series at displacement 1 also as well as errors Randomness is written as shown in equation (8) as follows (Pankratz, 1991):

$$\begin{aligned} Y_{1,t} &= \phi_{11} Y_{1,t-1} + \phi_{12} Y_{2,ti1} + a_{1,t} \\ Y_{2,t} &= \phi_{21} Y_{1,t-1} + \phi_{22} Y_{2,ti1} + a_{2,t} \end{aligned} \qquad \ldots (8)$$

Since: ϕ The coefficients of the current values of $Y_{1,t}$ andof the two and with the previous values $Y_{2,t}$ series (a_1, t) and (a_2, t) are represented, and independent random errors with a mean of zero and the variance (σ_1, t)-, respectively, and equation (8) represents the operations of the self (σ_2, t) و regression vector (AR). Any multiple AR model is used to detect Feedback.

5. Diagnostic tests

One of the most important diagnostic tests is the portmanteau test, where it is known that the test statistic is used to determine whether the residual series of the model is white noise (WN) or not, and is called the portmanteau statistic (portmanteau statistic) and it tests the following hypothesis: (Tsay, 2002)

$$\left.\begin{aligned} H_{\circ} &: \rho_1 = \ldots = \rho_m = 0 \\ H_a &: \rho_k \neq 0 \qquad ; k \in \{1, \ldots, m\} \end{aligned}\right\} \qquad \ldots (٩)$$

In general, model adequacy tests based mainly on this statistic are called portmanteau tests (Davies & Davies, 1979). Initial tests of this nature, such as the B-P and the L-B, have been shown to be ineffective (Arranz, 2005). It was mentioned (Makridakis, et al., 1998) that these two tests sometimes fail to reject poorly-matched models of data, and care must be taken to accept a model based on portmanteau tests only. Because the portmanteau test statistic is affected by the number of self-correlation coefficients (m) used in its calculation, an increase in the number of self-correlation coefficients can lead to the acceptance of the null hypothesis (Vandel, 1992) (Al-Hamoushi, 2011).

Ljung-Box Test (L-B)

In 1978, researchers Ljung & Box made a simple modification to the B-P test, based on what was observed by Box & Pierce that rk is distributed according to the normal distribution with a mean of zero and the variance of its amount $(n-k)/\{n(n+2)\}$, i.e .: (Ljung & Box, 1978).

$$r_k \sim N\left(0, \frac{(n-k)}{n(n+2)}\right) \quad (١٠)$$

As:

$$\sqrt{\frac{n(n+2)}{n-k}}\, r_k \sim N(0,1) \quad (١١)$$

By squaring equation (11), we get:

$$\frac{n(n+2)}{(n-k)} r_k^2 \sim \chi_1^2 \quad (١٢)$$

Taking the sum of m from the self-correlations after estimating the parameters q, p of the model, then:

$$\left[n(n+2)\sum_{k=1}^{m}\frac{r_k^2}{(n-k)}\right] \sim \chi_{m-p-q}^2 \quad (١٣)$$

This is known as the L-B or QL-B statistic:

$$Q_{L-B} = n(n+2)\sum_{k=1}^{m}\frac{r_k^2}{n-k} \sim \chi_{m-p-q}^2 \quad (١٤)$$

Researchers Ljung & Box found that this modification of BP test gave them a statistic closer to the $\chi^{٢}$ distribution of BP stats because their mean levels were close to theoretical values (Davies & Newbold, 1979), and it was observed to give better results at small sample sizes (Yaffee & McGee, 2000). Where the calculated values Q_{L-B} are compared with the values of the tabular $\chi^{٢}$ with a degree of freedom (m-p-q) in the case of studying the rest of the model. If the calculated value is less than the tabular value, this leads to the conclusion that the remaining series are white noise, and vice versa.

6. Simulation experiences

The simulation method has wide uses in all fields, and its importance has emerged after the rapid development in providing applications (ready software). Simulation can be defined as a process of simulating the actual reality of real models, whether this simulation was done manually or by computer and it is a mirror of some aspects of the real world and appears dependent on random processes (Stochastic) (Al-Khayat, 2011).

In this research, simulation experiments were conducted on two models, one Linear and the other nonlinear, as observations were generated with four different volumes of observations for both models (n=50,100,250,50) view and the Ut model inputs were random signals generated from the standard random Gaussian signals distribution " rgs ". As for jamming, random signals were generated that follow the normal distribution. The following linear model was used without back-up: -

$$y_t = 0.7 * u(t-1) + e(t) \qquad \dots(15)$$

And the following linear model is feedback:

$$y_t = 0.7 * u(t-1) + 0.3 * y(t-1) + e(t) \qquad \dots(16)$$

While using the non-linear model and without feedback as follows: -

$$y_t = (1 + 0.5 * \exp(-0.6 * u(t-2)\text{^}2)) + e_t \qquad \dots(17)$$

The following non-linear model is feedback by:

$$y_t = (1 + 0.5 * \exp(-0.6 * u(t-2)\text{^}2) * 0.8 * y(t-1) + e_t \qquad \dots(18)$$

The experiment was repeated 1000 times and using the ready program MINITAB and using the macros (MACROS). The self-link and partial self-function of the original series of outputs were examined and then found the main components of the two series as well as examining the two functions above for the first main component using the LB test at a level of significance of 0.05 and the number of ACF and PACF equal For a quarter of the sample size and since it is expected that as long as the first major component explains most of the variance and reflects the strength of the link between the original variables, then in the case of a reverse feed between the original variables (the input and output chains), it is expected that the self-linkages and sequences S partial self In this case, whether for the output chain data or the first main component data, these associations are significant, leading to rejection H_0 in equation (9) and acceptance H_1 which reveals the existence of the feedback and in return is accepted H_0 in the absence of a feedback.

7. Simulation results

Applied to the output chain of the linear model and with the presence of the feedback in the model and when (n = 50) it was found that the Ljung-Box test statistic revealed the significance of the self-correlations by 99.8%, i.e. rejecting the null hypothesis which states that the values of the self-correlations are not significant as well as revealed the LB statistic The significance of the self-correlations of the first major component is 95%, which indicates that it is possible to detect the feedback of the original data series, which is better than it is when finding the main components of the data However, when the sample size increases, i.e. (n = 100,250,500 views), it turns out that the LB statistic revealed the significance of the self-correlations of the first component by 100%, while the ratio with the original output chain at the above sizes was (93.7%, 95%, 95). Respectively, these results provide a probability for the first major component when employed to detect the presence of feedback. This indicates that revealing the feedback using the first main component of the main components is better than the original data series.

Tables (1) and (2) show the percentages of the significance of the correlation functions of the linear model and the non-linear model and with the presence of the opposite or not as follows:

Table (1): Self-correlation significance percentages *

Sample Size	N=50		N=100		N=250		N=500	
The nature of data	Output series	The first major component	Output series	The first major component	Output series	The first major component	Output series	The first major component
Linear model without feedback	6.9%	7.1%	13.9%	14.1%	22.9%	22.2%	23.0%	22.2%
Linear model with feedback	99.8%	95.0%	93.7%	100%	95.0%	100%	95.0%	100%
Non-Linear model without feedback	7.6%	8.1%	15.1%	14.5%	22.5%	22.0%	22.0%	22.5%
Non-Linear model with feedback	98.5%	72.0%	100%	97.7%	98.5%	72.5%	98.5%	72.5

By examining the results in table (1) above, it is clear that:

1. Relying on the L-B test in examining the significance of the self-correlation function, whether for the original output chain data or for the data of the first major component, gave very encouraging results in its ability to distinguish between the models used in terms of the presence of the feedback in its movement or not.
2. On the other hand, it is also noted that there is a strong indication that the use of the first major component with the linear model in the presence of the feedback gave very strong results in the detection of the presence of the feedback as the size of the sample increased using the proposed method than with the non-linear model with the feedback.

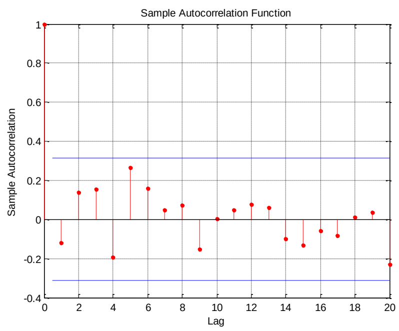
Table (2): The percentages of the significance of the partial self-correlation function

Sample size	N=50		N=100		N=250		N=500	
The nature of data	Output series	The first major componen t	Output series	The first major component	Output series	The first major component	Output series	The first major componen t
Linear model without feedback	7.5%	9.2%	12.5%	13.2%	22.5%	22.7%	22.0%	24.2%
Linear model with feedback	95.8%	93.0%	89.7%	95%	93.0%	95%	92.0%	93%
Non-Linear model without feedbac	8.6%	10.1%	17.1%	18.5%	25.5%	28.0%	21.0%	27.5%
Non-Linear model with feedback	95.5%	70.0%	95%	97.7%	95.5%	75.5%	93.5%	70.5

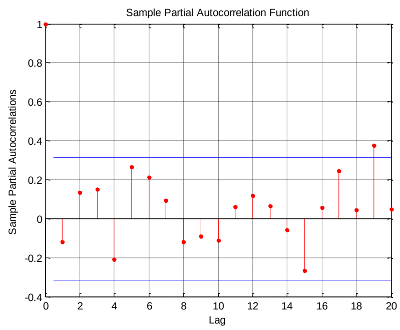
From the results of the above table, it is noted:

1. The employment of the L-B test in examining the significance of the partial self-correlation function gave encouraging results in its ability to distinguish between the models used in terms of whether or not the feedback was in its movement and whether the application was with the data of the original output chain or the main component.
2. The proposed method gave very encouraging and close results despite the difference in the sample sizes in detecting the feedback in the linear model, whether by using it for the output chain data or data for the first major component.
3. Although the proposed method is presented for good results in revealing the feedback using the data of the first major component of the non-linear model of different sample sizes, the use of the output chain data for the non-linear model with the reverse feeding gave strong results in revealing the presence of the feedback.

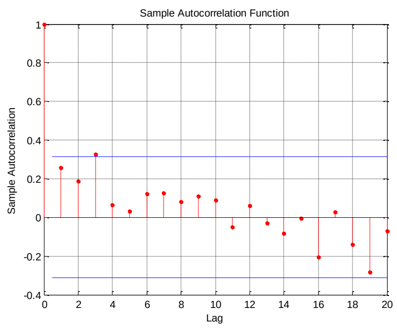
The behavior of the linear model and the non-linear model can be observed through the self-linking and partially self-correlating functions of one of the iterations through Figures (3) and (4) as follows:



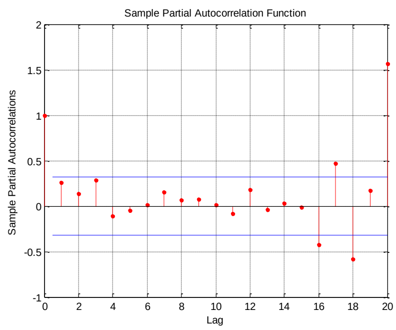
Self-correlation function for the linear model's original output string data



Partial self-correlation function for the linear model's original output chain data

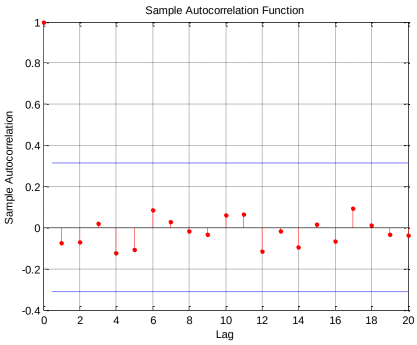


The self-correlation function of the first major component of the linear model

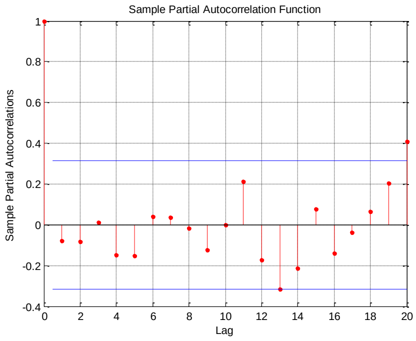


The partial self-correlation function of the first major component For the linear model

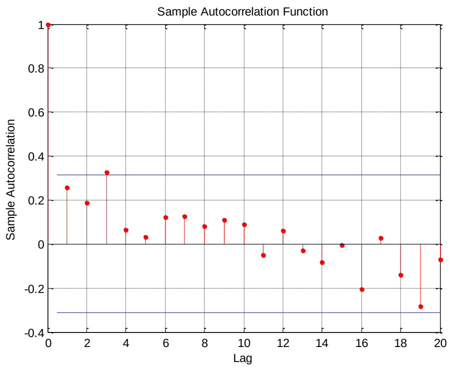
Figure (3): Behavior of the partial and self-linking functions of the linear model of one of the occurrences of the presence of the feedback when the sample size is n = 50.



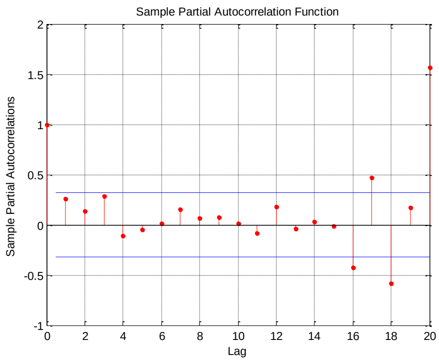
Self-correlation function for the non-linear model's original output string data



Partial self-correlation function for the original non-linear model output chain data



The self-correlation function of the first major component For the non-linear model



The partial self-correlation function of the first major component For the non-linear model

Figure (4): Behavior of the self and partial self-linking functions of one of the iterations when there is no reverse feed when the sample size is n = 50.

٨. Conclusions

The research reached some conclusions, including:

1. The proposed method based on the L-B test gave very encouraging results in distinguishing between the models used in terms of the presence of the feedback in its mobility or not, whether it was adopted in applying the proposed method to the self-correlation coefficients or partial self-correlation coefficients.
2. The process of applying the proposed method based on the self-correlation coefficients of the first major component led to very strong results, especially with the linear model with the opposite feedback in addition to obtaining encouraging results when using the non-linear model.
3. When applying the proposed method relying on the self-correlation coefficients for the first major component, the results were very encouraging and converging with those obtained when using the original output chain data for the linear model while the results obtained with the non-linear model were more robust with string data The original outputs with the main main component data.
4. The increase in the sample size in general played a positive role in raising the percentage of the ability to detect the presence of adverse reactions in the kinetic models used.

9. Reference

[1] Afifi,A.A.,(1984),:"Computer Aided Multivariate Analysis", Lifetime Learning Publications Belmont,California U.S.A.

[2] Alhamoshy,Wasaan Raad Thanoon,(2011):" Simulation study of the Schwartz standard of information behavior and some time series diagnostic tests", Master Thesis, College of Computer and Mathematic Sciences,Unversity of Mosul, Iraq.

[3] AlKayat,Bassel Thanoon,: (2011)," Introduction to Computer Static Simulation and Modeling Using MATLAB,Dar Iben Alhaytham For printing and publishing,Unversity of Mosul.

[4] AlKayat,Bassel Thanoon,: (2011)," Markovian modeling with practical applications,Dar Iben Alhaytham For printing and publishing,Unversity of Mosul.

[5] Al-Rawi,Kasha Mahmood, (1987),:" Introduction to regression analysis ", Dar Al-Kutub for Printing and Publishing, University of Mosul, Iraq.

[6] Arranz,M.A.,(2005),:"Pormanteau Test Statistics in Time Series",http://www.tol-project.Org/.

[7] Davies,N., & Newbold, P.,(1979),:"Some power studies of a Pormanteau Test of Time Series Model Specification",Biometrika,66,pp.153-155.

[8] Fleming,W.H.,(1989),:"Future Directios in Control Theory A Mathematical Perspective",Report,Society for Industrial and Applied Mathematics,Philadelphia,U.S.A.

[9] Kanjilal,P.P.,(1995):"Adaptive Prediction and Predctive Control",Peter Peregrines Ltd.London.

[10] Ljung,G.M.,Box,G.E.P.,(1978),:"On A Measure of Lack of Fit in Time series Models",Biometrika,65,pp.297-303.

[11] Makridakis, S., Wheelwright, S.,& Hyndman, R.,(1998),:"Forecasting Methods and Applications",3ed., John Wiley & Sons, New York.

[12] Nelles,O.,(2001),:"Nonlinear System Identification from Classical Approach to Neural Network and Fuzzy Models",Spring Verlag Belin Heidelberg Germany.

[13] Pankratz, Alan,(1991),:"Forecasting with Dynamic Regression Models",John Wiley and Sons,New York,U.S.A.

[14] Tsay,R.S.,(2002),:"Analysis of Financial Time Series:Financial Econometrics", John Wiley & Sons ,New York.

[15] Vandel,Walter,(1992),:" Time series from an applied point of view and Box Jenkins models ", Abd al-Mardi Azzam, Arabization, Mars Publishing House, Saudi Arabia.

[16] Wei,W.W.S.,(1990),:"Time series Analysis Univariate Multivariate Methods",Addison-Wesley Publising Company, Inc., The Advanced Book Program,California.U.S.A.

[17] Yaffee,R. A., & McGee,M.,(2000),:"Introduction to Time Series Analysis and Forecasting: With Applications of SAS and SPSS",Academic Press,San Diego.

استعمال التحليل التمييزي في تصنيف الاطفال حديثي الولادة الى طبيعيين وخدج (دراسة تطبيقية)

أ.م. د. سعد صبر محمد
جامعة /واسط
قسم الاحصاء /كلية الادارة والاقتصاد
Saadsabir455@Gmail.com

المستخلص : يعتبر أسلوب التحليل التمييزي احد الاساليب الاحصائية في تصنيف البيانات الى مجموعات بحسب المتغيرات المستقلة المؤثرة على متغير الاستجابة النوعي (او التصنيفي) , ويشمل مرحلة التصنيف و مرحلة التمييز. وان الهدف الأساسي من أسلوب التحليل التمييزي هو ايجاد دالة تعرف بالدالة التمييزية (او التصنيفية) , اذ يتم بموجبها تصنيف المشاهدات بحسب الصفات التي تحملها لكل متغير مستقل الى مجموعتين أو أكثر بحيث نستطيع الحكم باستعمال هذه الدالة على عائدية (او تمييز) اية مشاهدة جديدة الى احدى هذه المجموعات . وتوصل البحث الى أن المتغيرات المستقلة كوزن الطفل بالكغم ومدة الحمل بالاسبوع ذات تأثير معنوي .اما عمر الام بالسنة فليس له تأثير معنوي, وكانت نسبة كفاءة التصنيف الصحيح للأطفال الطبيعيين ١٠٠ % ونسبة كفاءة التصنيف الصحيح للأطفال الخدج ٩٦,٧ % , وان نسبة٩٨,٣٪ من الحالات المجمعة الأصلية المصنفة بشكل صحيح وهذه النسب تدعم أسلوب التحليل التمييز في التنبؤ.

الكلمات المفتاحية: التحليل التمييزي , الدالة التمييزية الخطية , اختبارات معنوية الدالة التمييزية الخطية ,احتمال خطأ التصنيف

The use of discriminant analysis in classifying new borns into normal and preterm infants (applied study)

Assist.Prof. Dr. Saad Sabir Mohameed
University / Wasit
Statistics Department / College of Administration and Economics
Saadsabir455@Gmail.com

Abstract: The discriminant analysis method is one of the statistical methods in classifying data into groups according to the independent variables affecting the specific response variable (or classification), and it includes the classification stage and the discrimination stage . The primary goal of the discriminant analysis method is to find a function known as the discriminant (or classification) function, whereby observations are classified according to the characteristics they carry for each independent variable into two or more groups so that we can judge using this function on the return of any new view to one These groups. The research found that the independent variables, such as the weight of a child in kg and the duration of pregnancy per week, had a significant effect. As for the mother's age in the year, it had no significant effect. The correct classification efficiency ratio for normal children was 100%, and the correct classification efficiency ratio for premature infants was 96.7%, and 98.3% of the original grouped cases were correctly categorized, and these percentages support the differential analysis method in prediction.

Key words: discriminant analysis, linear discriminant function, linear discriminant significance tests , Probability of Misclassification

١,١ المقدمة Introduction

يعتبر التحليل التمييزي Discriminant Analysis من اساليب متعدد المتغيرات التي تهتم بفصل مجموعات مختلفة من المفردات والتي تكون متشابهة في كثير من الصفات على أساس عدة متغيرات من خلال استخدام الدالة التمييزية , والتي هي عبارة عن تركيب خطي للمتغيرات المستقلة . وان اول من استعمل التحليل التمييزي هو كارل بيرسون Karl Pearson عام ١٩٢١ عندما اقترح اسلوب احصائي أطلق علية معاملات التشابه للأشياء المتماثلة . وفي عام ١٩٣١ أوجد فيشر Fisher دالة خطية لتصنيف المفردة الى احدى المجموعتين مع تساوي التباينات واطلق عليها اسم الدالة المميزة الخطية Linear Discriminant Function .
ومن هنا جاءت فكرة استخدام الدالة المميزة الخطية للمجتمعات متعددة المتغيرات ، فالفكرة الأساسية من التمييز هي التفرقة بين مجتمعات متداخلة أو متشابكة لها نفس الخصائص أو الصفات ، حيث أن وظيفة التحليل التمييزي هو ايجاد دالة يمكن بواسطتها تصنيف أو تمييز المشاهدات الجديدة الى مجتمعاتها الأصلية . وتعد الدالة التمييزية والتي تسمى بدالة فيشر طريقة فعالة ومناسبة فيما لو تحققت شروطها الخاصة وهي التوزيع الطبيعي للمتغيرات التوضيحية وتساوي مصفوفات التباين والتباين المشترك ، ويستخدم التحليل التمييزي في عملية التوقع , إذ يأتي الباحث بعدة متغيرات توضيحية مؤثرة في المتغير النوعي (او الوصفي) التابع بحيث يتوقع أن يميز (يصنف) بين مجتمعين في المستقبل ، ولكي نحصل على دالة تمييزية تستخدم في تصنيف المشاهدات بين مجتمعين في المستقبل، وأيضا للحصول على أعلى تمييز بين المجموعات على أن تكون نسبة التباين بالنسبة إلى التباين داخل المجموعات كبيرا.

عليه يستعمل التحليل التمييزي (**Discriminant Analysis**) لايجاد مجموعة من المعادلات المتوقعة المبنية على متغيرات مستقلة تستخدم لتصنيف الأفراد إلى مجموعات. ففي مجالات كثيرة ، يتشابه التحليل التمييزي بتحليل الانحدار المتعدد , الا ان الفرق هو ان المتغير التابع منفصل . ويأتي التطبيق الرئيسي للتحليل التمييزي في الطب فهو يساعد على تخمين الحالة الحرجة للمريض والإنذار بنتيجة المرض. فعلى سبيل المثال ، دراسة نتائج التحاليل المعملية والسريرية لغرض اكتشاف المتغيرات المختلفة إحصائيا في المجموعات قيد الدراسة. وباستخدام تلك المتغيرات يتم بناء دوال التمييز التي تساعد في تصنيف الأمراض بشكل فعال للمرضى المستقبليين إلى بسيط ومتوسط وخطير .

٢,١ الهدف من البحث The Aim of Research

يهدف البحث الى ايجاد معادلة تنبؤية لتصنيف الاطفال الجدد حديثي الولادة من الاطفال الخدج والاطفال الطبيعيين وكذلك تفسير المعادلة التنبؤية لفهم العلاقات التي قد توجد بين المتغيرات كوزن الطفل ومدة الحمل وعمر الام...الخ بشكل أفضل .

٣,١ مشكلة البحث The Problem of Research

يواجه احيانا في المستشفيات والردهات الطبية صعوبة التمييز بين الاطفال الخدج والاطفال الطبيعيين مما قد يؤدي سوء العزل الى مشاكل بسبب الخطأ في التصنيف قد تؤدي احيانا الى حرمان الطفل من الرعاية التي يكون الطفل بحاجة ماسة اليها , لذلك فان الاساليب الاحصائية باستعمال الدالة التمييزية في تصنيف الاطفال قد تساهم بشكل كبير في تصنيف المواليد وتوفير الرعاية الطبية لهم .

٤,١ فرضيات البحث Assumptions of Research

يعتمد البحث على الفرضيات الاتية :

من أجل القيام بالتحليل التمييزي يجب توافر الفروض الاتية :

١. أن المجتمعات موضوع الدراسة منفصلة Discrete وقابلة للتحديد وان كانت هذه المجتمعات تتداخل فيما بينها بدرجات متداخلة معينة.

٢. أن كل مفردة في كل مجتمع يمكن وصفها وتحديدها بمجموعة من المقاييس أو المتغيرات المستقلة، وأن تكون جميع متغيرات دالة التمييز مقاسة بقيم محددة.

٣. أن المجتمعات موضوع البحث او الدراسة تختلف بالنظر إلى أوساطها، أي أن متجهات أوساط المتغيرات للمجتمعات تكون غير متساوية.

٤. أن البيانات المستخدمة في التحليل التمييزي تحتوي على عينة عشوائية من أعضاء كل مجتمع من المجتمعات قيد البحث او الدراسة، بحيث تعد هذه العينات ممثلة للمجتمعات موضوع البحث او الدراسة.

٥. عدم وجود ارتباط بين المتغيرات المستقلة الداخلة في تكوين دالة التمييز حتى تفسير النتائج ، وتحديد المساهمة النسبية لكل متغير في القوة التمييزية الكلية .

٥,١ عينة البحث A sample of Research

لقد تم جمع البيانات الخاصة بالبحث من خلال الزيارات الميدانية لمستشفى الصويرة العام والاطلاع على السجلات بخصوص الاطفال حديثي الولادة , فقد تم جمع البيانات بواقع ٣٠ طفل حديثي الولادة ضمن الاطفال الطبيعيين وجمعت البيانات عن وزن الطفل ومدة الحمل فيه بالأسبوع وعمر الام بالسنة فكانت تلك المجموعة الاولى . ثم تم اختيار عينة من الاطفال حديثي الولادة ضمن الاطفال الخدج بواقع (٣٠)طفل ايضا وكانت المجموعة الثانية , وتم جمع البيانات لنفس المتغيرات اعلاه , وقد تم العمل لأجل ذلك للفترة من ٢٠١٩/٦/١ ولغاية٢٠١٩/٧/١ .

الفصل الثاني

التحليل التمييزي Discriminant Analysis

١,٢ المقدمة Introduction

يعد التحليل التمييزي أسلوب إحصائي لتحليل البيانات متعددة المتغيرات ، حيث يهتم بمسألة التمييز بين مجموعتين أو اكثر والتي تكون متشابهة في كثير من الصفات على أساس عدة متغيرات من خلال استخدام الدالة المميزة والتي هي عبارة عن تركيب خطي للمتغيرات المستقلة ، ويختلف التحليل التمييزي عن التحليل العنقودي في أن فكرة التحليل العنقودي تبدأ دون توافر معرفة مسبقة بعدد المجاميع أو أي من المفردات التي تنتمي لهذه المجموعة أو تلك ، كما ان التحليل التمييزي يختلف عن تحليل الانحدار في أن المتغير التابع في التحليل التمييزي هو متغير إسمي وهو من المتغيرات النوعية بينما المتغير التابع في تحليل الانحدار هو في الغالب متغير مستمر وهو من المتغيرات الكمية , أما عملية التصنيف (Classification) فهي العملية اللاحقة بعد تكوين الدالة المميزة حيث يتم الاعتماد على هذه الدالة بالتنبؤ وتصنيف المفردة الجديدة لإحدى المجموعات قيد الدراسة بأقل خطأ تصنيف ممكن ، ويشترط تساوي التباينات للمجموعات قيد البحث او الدراسة , وهناك تمييز خطي في حالة مجموعتين ، وتمييز خطي في حالة أكثر من مجموعتين ، أما التمييز غير الخطي فيستخدم في حالة عدم تساوي التباينات.[10,9,8,7,6,5]

٢.٢ الدالة المميزة الخطية في حالة مجموعتين[4,3,2,1]

Two Groups–Linear Discriminant Function

إن دالة التمييز هي انموذج يمكن صياغته اعتمادا على مؤشرات العينة التي تم اختيار مفرداتها ووضعت في مجموعتين مختلفتين ، وبواسطة هذه الدالة نستطيع أن نختبر المفردة ونحدد عائديتها إلى أي مجموعة . فلو فرضنا أن مجال العينة هو W سوف يقسم إلى قسمين (R) يعود إلى المجموعة الأولى و (W-R) يعود إلى المجموعة الثانية , أما الحد الفاصل بين المجموعتين فيمكن أن يعود إلى أية مجموعة من هاتين المجموعتين.

وبافتراض أن المجتمعين المراد مقارنتهما لهما نفس مصفوفة التغاير Σ ولكن بمتوسطي المتجهين التمييزين U_1 و U_2. فعند اخذ العينات(x_{11} , x_{12} , . . ., x_{1n1}) و (x_{21}, x_{22} , . . . ، x_{2n2}) من المجتمعين. كالعادة كل متجه يتكون من X_{ij} من القياسات ل p من المتغيرات. عليه فان الدالة التمييزية هي توفيق خطي لهذه المتغيرات p التي تعظم المسافة بين متوسطي متجهي المجموعتين. والتوفيق الخطي يساوي $Y = B'X$, ويكتب بالصيغة الاتية :

$$Y_{1i} = B'x_{1i} = Y_i = B_1X_{1i1} + B_2x_{1i2} + ... + B_pX_{1ip}\ ,\ i = 1, 2, . . . , n_1$$

$$.Y_{2i} = B'x_{2i} = B_1X_{2i1} + B_2x_{2i2} + ... + B_pX_{2iP}\ ,\quad i = 1, 2, . . . , n_2$$

اذ ان

$$\bar{Y}_2 = \frac{\sum y_{2i}}{n_2} \quad و \quad \bar{Y}_1 = \frac{\sum y_{1i}}{n_1}$$

وحيث ان:

P : عدد المتغيرات الداخلة في الدالة .

β : معاملات الدالة المميزة المعيارية.

r=2 : عدد الدوال المميزة.

ولتحديد الاختلافات بين المجموعتين فإنه من المناسب استخراج الأوساط الحسابية لهاتين المجموعتين . إن عملية التقدير للمعلمات (β,S) والتي تجعل الدالة تعطي أفضل تمييز بين المجموعتين لابد أن يتم من خلال جعل مربع الفرق بين متوسطي المجموعتين إلى التباين المشترك للمجموعتين أكبر ما يمكن , أي ان:

$$Q=\frac{\left(\bar{Y}_1-\bar{Y}_2\right)^2}{\Sigma_{i=1}^{2}\Sigma_{j=1}^{ni}\left(Y_{ij}-\bar{Y}_i\right)^2}=\frac{\left[B'\left(\bar{X}_1-\bar{X}_2\right)\right]^2}{B'SB}$$

$$=\frac{B'\left(\bar{X}_1-\bar{X}_2\right)\left(\bar{X}_1-\bar{X}_2\right)'B}{B'SB} \qquad \text{... ٢,٢}$$

حيث نقدر معلمات الدالة المميزة من خلال تعظيم النسبة Q باشتقاقها جزئيا ومساواتها بالصفر وكما يلي [15,14,13,12,11]:

$$\frac{\partial Q}{\partial B}=\frac{\left(\bar{X}_1-\bar{X}_2\right)}{\sqrt{\hat{B}^t S_{pooled}\hat{B}}}-\frac{1}{2}\hat{B}^t\left(\bar{X}_1-\bar{X}_2\right)\frac{2S_{pooled}\hat{B}}{\left(\hat{B}^t S_{pooled}\hat{B}\right)^{3/2}}=0$$

ومزيدا من التبسيط يعطي :

$$\bar{X}_1-\bar{X}_2=\left[\frac{\hat{B}^t\left(\bar{X}_1-\bar{X}_2\right)}{\hat{B}^t S_{pooled}\hat{B}}\right]S_{pooled}\hat{B}$$

وبالضرب بمعكوس المصفوفة S_{pooled} لطرفي المعادلة نحصل :

$$S_{pooled}^{-1}\left(\bar{X}_1-\bar{X}_2\right)=\left[\frac{\hat{B}^t\left(\bar{X}_1-\bar{X}_2\right)}{\hat{B}^t S_{pooled}\hat{B}}\right]\hat{B},$$

وبما ان المقدار ادناه عدد حقيقي ثابت وليكن C اي :

$$\frac{\hat{B}^t(\bar{X}_1-\bar{X}_2)}{\hat{B}^t S_{pooled}\hat{B}}=C$$

$$\hat{B}=CS^{-1}\left(\bar{X}_1-\bar{X}_2\right)' \qquad \text{... ٣,٢}$$

حيث ان

$$S = S_{pooled} = \frac{(n_1-1)S_1+(n_2-1)S_2}{n_1+n_2-2}, S_1 = \frac{\sum_{i=1}^{n_1}\left(X_i-\overline{X}_1\right)\left(X_i-\overline{X}_1\right)^t}{n_1-1}$$

$$S_2 = \frac{\sum_{i=n_1+1}^{n_1+n_2}\left(X_i-\overline{X}_2\right)\left(X_i-\overline{X}_2\right)^t}{n_2-1}$$

وان

$$\overline{X}_2 = \frac{\sum_{i=n_1+1}^{n_1+n_2} X_i}{n_2} \quad \overline{X}_1 = \frac{\sum_{i=1}^{n_1} X_i}{n_1}$$

$$\overline{Y}_1 = \frac{\sum_{i=1}^{n_1} Y_i}{n_1} = \frac{\sum_{i=1}^{n_1} \hat{B}^t X_i}{n_1} = \hat{B}^t \left(\frac{\sum_{i=1}^{n_1} X_i}{n_1} \right) = \hat{B}^t \overline{X}_1$$

ونفس الشيء $\overline{Y}_2 = \hat{B}^t \overline{X}_2$

وكذلك فان

$$\sum_{i=1}^{n_1}(Y_i-\overline{Y}_1)^2 = \sum_{i=1}^{n_1}\left(\hat{B}^t X_i - \hat{B}^t \overline{X}_1\right)^2 = \sum_{i=1}^{n_1}\left(\hat{B}^t X_i - \hat{B}^t \overline{X}_1\right)\left(\hat{B}^t X_i - \hat{B}^t \overline{X}_1\right)^t$$

$$= \sum_{i=1}^{n_1} \hat{B}^t \left(X_i-\overline{X}_1\right)\left(X_i-\overline{X}_1\right)^t \hat{B} = \hat{B}^t \left[\sum_{i=1}^{n_1}\left(X_i-\overline{X}_1\right)\left(X_i-\overline{X}_1\right)^t\right]\hat{B}.$$

ونفس الشيء

$$\sum_{i=n_1+1}^{n_1+n_2}\left(Y_i-\overline{Y}_2\right)^2 = \hat{B}^t\left[\sum_{i=n_1+1}^{n_1+n_2}\left(X_i-\overline{X}_2\right)\left(X_i-\overline{X}_2\right)^t\right]\hat{B}$$

لذلك

$$S_Y^2 = \frac{\sum_{i=1}^{n_1}\left(Y_i-\overline{Y}_1\right)^2 + \sum_{i=n_1+1}^{n_1+n_2}\left(Y_i-\overline{Y}_2\right)^2}{n_1+n_2-2}$$

$$= \frac{\hat{B}^t\left[\sum_{i=1}^{n_1}(X_i-\bar{X}_1)(X_i-\bar{X}_1)^t\right]\hat{B}+\hat{B}^t\left[\sum_{i=n_1+1}^{n_1+n_2}(X_i-\bar{X}_2)(X_i-\bar{X}_2)^t\right]\hat{B}}{n_1+n_2-2}$$

$$= \hat{B}^t\left[\frac{\sum_{i=1}^{n_1}(X_i-\bar{X}_1)(X_i-\bar{X}_1)^t+\sum_{i=n_1+1}^{n_1+n_2}(X_i-\bar{X}_2)(X_i-\bar{X}_2)^t}{n_1+n_2-2}\right]\hat{B}$$

$$= \hat{B}^t\left[\frac{(n_1-1)S_1+(n_2-1)S_2}{n_1+n_2-2}\right]\hat{B} = \hat{B}^t S_{pooled}\hat{B}$$

علية فان

$$Q = \frac{\bar{Y}_1-\bar{Y}_2}{S_Y} = \frac{\hat{B}^t(\bar{X}_1-\bar{X}_2)}{\sqrt{\hat{B}^t S_{pooled}\hat{B}}}$$

وفي حالة وجود مجموعتين يكون لدينا دالة تمييز واحدة فقط ، وفي حالة وجود ثلاث مجاميع يكون لدينا دالتين تمييزيتين . وبعد استخراج المعاملات β^ تصنف المفردة إلى إحدى المجموعتين بالاعتماد على نقطة متوسط المجموعتين (L) التي تجعل احتمال التصنيف الخاطئ أقل ما يمكن :

$$L = \frac{\bar{Y}_1+\bar{Y}_2}{2} \qquad ......٤,٢$$

فإذا كانت $\hat{Y}>L$ تصنف المشاهدة إلى المجموعة الأولى

فإذا كانت $\hat{Y}<L$ تصنف المشاهدة إلى المجموعة الثانية

وإذا كانت $\hat{Y}=L$ تصنف المشاهدة عشوائي إلى المجموعة الأولى أو الثانية

حيث ان :

$$\hat{Y} = (\bar{X}_1-\bar{X}_2)' S^{-1}X \qquad .... \; 5.2$$

٣,٢ اختبارات معنوية الدالة التمييزية الخطية TESTS OF SIGNIFICANCE

عندما يراد التمييز بين مجموعتين ، فإنه يمكننا أن نختبر الفرضية التي تنص على تساوي متوسطات المجموعتين (18,17,16,15):

$$H_0 : u_1 = u_2$$
$$H_1 : u_1 \neq u_2$$

اولا: مقياس (Hotelling - T٢)

إحصاءة الاختبار المستخدمة في حالة التمييز بين مجموعتين هي (Hotelling - T٢) وصيغته كما يلي:

$$T^2 = \frac{n_1 n_2}{n_1 + n_2} D^2 \qquad \ldots 6.3$$

حيث : D^2 تمثل مسافة مهالنوبيس (Mahalanobis Distance) وصيغتها كما يلي:

$$D^2 = \left(\bar{X}_1 - \bar{X}_2\right)' S^{-1} \left(\bar{X}_1 - \bar{X}_2\right) \qquad \ldots 7.2$$

ويستخدم اختبار (F) وصيغته كما يلي :

$$F = \frac{n_1 + n_2 - p - 1}{(n_1 + n_2 - 2)p} T^2 \qquad \ldots 8.2$$

وبمقارنتها مع قيمة F الجدولية بدرجة حرية ($P,n1+n2-p-1$) ومستوى المعنوية α، فاننا نرفض H_0 بمستوى معنوية α اذا كانت:

$$F_{cal} > F_{\alpha,(P,n1+n2-p-1)}$$

ونقبل H_1 , وهذا يدل على أن متوسطات المجموعات غير متساوية وأنه توجد فروق معنوية بين المجموعتين وهذا معناه أن الدالة المميزة الخطية قابلة للتمييز بدرجة عالية ، ايضا يمكن

ثانيا : مقياس ولكس (Wilks-Criteria)

ويكون على وفق الصيغة الاتية [13,12,11,10]:

$$\Lambda = \frac{|W|}{|T|} \qquad \ldots.. 9.2$$

حيث ان T: : مصفوفة التباين والتغاير الكلي للمجموعات.

W: مصفوفة التباين والتغاير داخل المجموعات.

وتتراوح قيمة Λ بين الصفر والواحد ، فإذا كانت قريبة أو مساوية للواحد فان ذلك يشير إلى أن متوسطات المجموعات متساوية وبذلك لا يوجد تمييز بين المجموعات ، اما اذا كانت قيمتها قريبة من الصفر فان ذلك يدل على قوة التمييز.

ثالثا: مقياس مربع كاي X^2

و يمكن ايضا استخدام مقياس اخر الا وهو مقياس مربع كاي X^2 , ويعد هذا المقياس أكثر دقة من مقياس Λ وصيغته كما يلي :

$$X^2 = -Log(\Lambda) \qquad \ldots.10.2$$

وبمقارنتها مع القيمة الجدولية لمربع كاي بدرجة حرية بدرجة حرية $P(K-1)$، حيث P عدد المتغيرات ، K عدد المجموعات. وعند مستوى المعنوية α، فاننا نرفض H_0 بمستوى معنوية α اذا كانت القيمة المحسوبة اكبر من القيمة الجدولية وبخلافة نقبل H_0 .

٣,٢ احتمال خطأ التصنيف[16,15,14] Probability of Misclassification

هناك نوعان من احتمال خطأ التصنيف هما :

١ـ احتمال خطأ التصنيف P_{12} وهو احتمال تصنيف المفردة إلى المجموعة الثانية وهي أصلا تعود إلى المجموعة الأولى.

٢ احتمال خطأ التصنيف P_{21} وهو احتمال تصنيف المفردة الى المجموعة الأولى وهي أصلا تعود إلى المجموعة الثانية.

وبذلك سوف يكون تقدير احتمال التصنيف كما يلي:

$$P_{12} = P_{21} = \phi(-\frac{D}{2}) \quad ....١1.2$$

تمثل ϕ: دالة التوزيع الطبيعي القياسي.

D : هو جذر مقياس مهالنوبيس(Mahalanobis Distance) وتساوي:

$$D^2 = \left(\bar{X}_1 - \bar{X}_2\right)' S^{-1}\left(\bar{X}_1 - \bar{X}_2\right)$$

وان P_{12}, P_{21} يتم ايجادها من جداول التوزيع الطبيعي المعياري .

ويعد خطأ التصنيف بانه عامل مهم لإثبات كفاءة الدالة المميزة ,حيث أن الدالة المميزة التي تعطي أقل خطأ تصنيف هي الدالة الأكثر كفاءة وتكون الأفضل من بين دوال التمييز.

الفصل الثالث

الجانب التطبيقي

١,٣ المقدمة

لقد تم جمع البيانات الاحصائية عن حالات الاطفال حديثي الولادة بنوعيه الطبيعيين والخدج من خلال مراجعة مستشفى الصويرة العام للفترة من ٢٠١٩//٦/١ ولغاية ٢٠١٩/٧/١ وبالاستعانة بالسجلات الخاصة بالمستشفى تم جمع البيانات عن ٣٠ طفل طبيعي بالولادة و٣٠ طفل من الخدج , فكان حجم العينة من المجموعتين ٦٠ طفلا . وباستعمال الحزمة الاحصائية SPSS V(19) , تم تحليل واستخراج النتائج .

٢,٣ وصف البيانات

لقد تم تسجيل البيانات عن المتغيرات لكل حالة وهي بحسب الجدول الاتي:

جدول (١) وصف المتغيرات

رمز المتغير	وصف المتغير
Y	حالة الطفل اما طبيعي ويرمز له(٠) او خدج (١)
X1	وزن الطفل بالكغم
X2	مدة الحمل بالأسبوع
X3	عمر الام بالسنة

جدول رقم (٢)حجم العينة والقيم المفقودة

جدول رقم (٢)حجم العينة

		حالة الطفل	وزن الطفل بالكغم	مدة الحمل بالاسبوع	عمر الام بالسنة
N	Valid	60	60	60	60
	Missing	0	0	0	0

جدول رقم(٤) حالة الطفلفيما اذا كان طبيعي او خدج

جدول رقم (٤) **حالة الطفل**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	طبيعي	30	50.0	50.0	50.0
	خدج	30	50.0	50.0	100.0
	Total	60	100.0	100.0	

جدول رقم (٥)الاحصاءات الوصفية عن متغيرات الوزن والعمر ومدة الحمل لكل طفل

جدول رقم (٥) الاحصاءات الوصفية للوزن والعمر ومدة الحمل لكل طفل

	N	Minimum	Maximum	Mean		Std. Deviation
	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic
وزن الطفل بالكغم	60	1.20	4.50	2.7762	.13672	1.05899
مدة الحمل بالاسبوع	60	31.00	40.00	36.2000	.30606	2.37073
عمر الام بالسنة	60	19.00	39.00	26.1000	.48869	3.78534
Valid N (listwise)	60					

٣,٣ التحليل الاحصائي للبيانات

جدول رقم (٦) ملخص حالة التحليل الاحصائي

جدول رقم (٦) ملخص حالة التحليل الاحصائي

Unweighted Cases		N	Percent
Valid		60	100.0
Excluded	Missing or out-of-range group codes	0	.0
	At least one missing discriminating variable	0	.0
	Both missing or out-of-range group codes and at least one missing discriminating variable	0	.0
	Total	0	.0
Total		60	100.0

جدول رقم (٧) احصاءات كل مجموعة

جدول رقم (٧) احصاءات كل مجموعة

حالة الطفل	Mean	Std. Deviation	Valid N (listwise)	
			Unweighted	Weighted

طبيعي	وزن الطفل بالكغم	3.7500	.40210	30	30.000
	مدة الحمل بالاسبوع	38.1000	1.24152	30	30.000
	عمر الام بالسنة	26.5000	4.17505	30	30.000
خدج	وزن الطفل بالكغم	1.8023	.39725	30	30.000
	مدة الحمل بالاسبوع	34.3000	1.55696	30	30.000
	عمر الام بالسنة	25.7000	3.37486	30	30.000
Total	وزن الطفل بالكغم	2.7762	1.05899	60	60.000
	مدة الحمل بالاسبوع	36.2000	2.37073	60	60.000
	عمر الام بالسنة	26.1000	3.78534	60	60.000

جدول رقم (٨) اختبارات تساوي متوسطات المجموعة

جدول رقم (٨) اختبارات تساوي متوسطات المجموعة

	Wilks' Lambda	F	df1	df2	Sig.
وزن الطفل بالكغم	.140	356.202	1	58	.000
مدة الحمل بالاسبوع	.347	109.242	1	58	.000
عمر الام بالسنة	.989	.666	1	58	.418

ويتبين من الجدول وجود فروق معنوية في وزن الطفل ومدة الحمل (أي عدم تساوي متوسطات المجموعات) بمعنى رفض فرضية العدم وعدم وجود فروق معنوية في عمر الام (أي تساوي متوسطات المجموعات) بمعنى قبول فرضية العدم . فكلما كانت قيمة المقياس صغيرة جدا دل على اهمية المتغير المستقل في الدالة التمييزية .

جدول رقم (٩)مصفوفات التغاير والارتباط داخل المجموعات(Pooled Within-Groups Matrices)

جدول رقم (٩)مصفوفات التغاير والارتباط داخل المجموعات

	عمر الام بالسنة	مدة الحمل بالاسبوع	وزن الطفل بالكغم
Covariance *وزن الطفل بالكغم*	.160	.230	.031
مدة الحمل بالاسبوع	.230	1.983	.538
عمر الام بالسنة	.031	.538	14.410
Correlation وزن الطفل بالكغم	1.000	.409	.020
مدة الحمل بالاسبوع	.409	1.000	.101
عمر الام بالسنة	.020	.101	1.000

. a. The covariance matrix has 58 degrees of freed

جدول رقم (١٠)مصفوفات التغاير للأطفال الطبيعيين والخدج

جدول رقم (١٠)مصفوفات التغاير للأطفال الطبيعيين والخدج

حالة الطفل		وزن الطفل بالكغم	مدة الحمل بالاسبوع	عمر الام بالسنة
طبيعي	وزن الطفل بالكغم	.162	.093	.057

	مدة الحمل بالاسبوع	.093	1.541	.776
	عمر الام بالسنة	.057	.776	17.431
خدج	وزن الطفل بالكغم	.158	.367	.004
	مدة الحمل بالاسبوع	.367	2.424	.300
	عمر الام بالسنة	.004	.300	11.390
Total	وزن الطفل بالكغم	1.121	2.108	.426
	مدة الحمل بالاسبوع	2.108	5.620	1.302
	عمر الام بالسنة	.426	1.302	14.329

a. The total covariance matrix has 59 degrees of freedom.

جدول رقم (١١) اختبار بوكس **Box's** للتساوي لمصفوفات التغاير

Log Determinants جدول رقم (١١) اختبار بوكس **Box's** للتساوي لمصفوفات التغاير

حالة الطفل	Rank	Log Determinant
طبيعي	3	1.411
خدج	3	1.032
Pooled within-groups	3	1.325

الرتب واللوغاريتمات الطبيعية للمحددات المطبوعة هي تلك الخاصة بمصفوفات مجموعة التباين المشترك

.

بالنسبة لجدول قيم لوغاريتم المحددات لمستويات المتغير المعتمد حيث كلما كانت قيمته كبيره دل ذلك على ان مصفوفة التباين والتباين المشترك لتلك المجموعة تختلف عن الباقي. وبما ان لوغاريتم المحددات صغيرة ومتساوية نسبيا فأنها تدل على التجانس .

جدول رقم (١٢) اختبار النتائج

جدول رقم (١٢) اختبار النتائج

Box's M		5.972
F	Approx.	.939
	df1	6
	df2	24373.132
	Sig.	.465

.

ويشير الجدول الى عدم معنوية الاختبار وقبول فرضية العدم بمعنى ان مصفوفة التباين والتباين المشتركة متجانسة .

- **اختبارات فرضية العدم لمصفوفات التباين والتغاير لتساوي المجتمع.**

جدول رقم (١٣) القيم الخاصة لدالة التمييز .

جدول رقم (١٣) القيم الخاصة

Function	Eigenvalue	% of Variance	Cumulative %	Canonical Correlation
1	6.296[a]	100.0	100.0	.929

استعملت الدوال التمييزية الاولى في التحليل

جدول رقم (١٤)اختبار لمدا ووكWilks' Lambda للدالة التمييزية

جدول رقم (١٤)اختبار لمدا ووكWilks' Lambda

Test of Function(s)	Wilks' Lambda	Chi-square	df	Sig.
1	.137	112.287	3	.000

ويؤكد الجدول (١٤) معنوية الاختبار واهمية المتغيرات المستقلة في الدالة التمييزية , اذ يلاحظ من الجدولين (١٣) و(١٤) بان الدالة التمييزية التي تقابلها القيمة الذاتية (٦,٢٩٦) بنسبة ارتباط قانوني (٠,٩٢٩) لكل المتغيرات التمييزية , وقد فسرت الدالة التمييزية ١٠٠٪ من التباين , وان قيمة الارتباط القانوني وتساوي (٠,٩٢٩) العالية جدا تعكس قوة العلاقة بين المتغيرات الداخلة في التحليل , ولغرض معرفة جودة التمييز للدالة نلاحظ نتائج كل من احصاءة Wilks Lambda وتساوي (٠,١٣٧) وهي قريبة من الصفر واختبار كاي Chi-square (١١٢,٢٨٧) وبمستوى دلالة ٠,٠٥ , ومعنوية الاختبار(sig=0.00) , وهذا يشير الى جودة التمييز بين المجموعتين , وبذلك فان الاختلاف بين المجموعتين جوهري لا يعود الى الصدفة . وفي ادناه جدول رقم (١٥) معاملات دالة التمييز المعيارية

جدول رقم (١٥) معاملات دالة التمييز المعيارية

	Function
	1
وزن الطفل بالكغم	.918
مدة الحمل بالاسبوع	.171
عمر الام بالسنة	.007

والجدول رقم (١٧) المصفوفة الهيكلية للدالة التمييزية

جدول رقم (١٧) المصفوفة الهيكلية

	Function
	1
وزن الطفل بالكغم	.988
مدة الحمل بالاسبوع	.547
عمر الام بالسنة	.043

بحسب الارتباطات المجمعة داخل المجموعات بين المتغيرات التمييزية والدوال التمييزية المعيارية متغيرات مرتبة تنازليا بحسب الحجم المطلق للارتباط داخل الدالة . اذ يأتي وزن الطفل بالدرجة الاولى من حيث قوة العلاقة (٠,٩٨٨) بالدالة التمييزية ثم بعدها مدة الحمل بالأسبوع (٠,٥٤٧) واخير واقل ارتباط عمر الام السنه (٠,٠٤٣) .

والجدول رقم (١٨) يبين معاملات الدالة التمييزية القانونية

جدول رقم (١٨) معاملات الدالة التمييزية القانونية

	Function
	1
وزن الطفل بالكغم	2.296
مدة الحمل بالاسبوع	.121
عمر الام بالسنة	.002
(Constant)	-10.818-

وهي معاملات غير معيارية .

والجدول رقم (١٩) يبين دوال في النقاط الوسطى للمجموعة

جدول رقم (١٩) دوال في النقاط الوسطى للمجموعة

حالة الطفل	Function 1
طبيعي	2.467
خدج	-2.467-

اذ يتم تقييم الدوال التمييزية القانونية غير القياسية في اوساط المجموعة

- **احصاءات التصنيف وكالاتي :**

الجدول رقم (٢٠) يبين الاحتمالات المسبقة للمجاميع

جدول رقم (٢٠) الاحتمالات المسبقة للمجاميع

حالة الطفل	Prior	Cases Used in Analysis	
		Unweighted	Weighted
طبيعي	.500	30	30.000
خدج	.500	30	30.000
Total	1.000	60	60.000

والجدول رقم (٢١) يبين معاملات دالة التصنيف

جدول رقم (٢١)معاملات دالة التصنيف

	حالة الطفل	
	طبيعي	خدج
وزن الطفل بالكغم	-4.787-	-16.114-
مدة الحمل بالاسبوع	19.467	18.868
عمر الام بالسنة	1.122	1.113
(Constant)	-377.432-	-324.056-

دوال فيشر الخطية التمييزية

دالة التصنيف للمولود الطبيعي **Y1=-4.787X1+19.467X2+1.122X3-377.432**

دالة التصنيف للمولود الخدج **Y2=-16.114X1+18.868X2+1.113X3-324.056**

والجدول رقم (٢٢) يبين نتائج التصنيف

جدول رقم (٢٢)نتائج التصنيف

		حالة الطفل	Predicted Group Membership		Total
			طبيعي	خدج	
Original	Count	طبيعي	30	0	30
		خدج	1	29	30
	%	طبيعي	100.0	.0	100.0
		خدج	3.3	96.7	100.0

98.3٪ من الحالات المجمعة الأصلية المصنفة بشكل صحيح

الفصل الرابع

الاستنتاجات والتوصيات

اولا : الاستنتاجات

١.اظهرت نتائج البحث معنوية المتغيرات X1 وزن الطفل بالكغم وX2 مدة الحمل بالاسبوع في المعادلة التمييزية .

٢. كانت نسبة كفاءة التصنيف الصحيح للأطفال الطبيعيين ١٠٠ % ونسبة كفاءة التصنيف الصحيح للأطفال الخدج ٩٦,٧ % , وان نسبة٩٨,٣٪ من الحالات المجمعة الأصلية المصنفة بشكل صحيح وهذه النسب تدعم أسلوب التحليل التميز في التنبؤ.

٣. تبين من خلال التطبيق ان هناك امكانية في تطبيق اسلوب التحليل التمييزي على حالات

ومجاميع مشابهة والتي لها كفاءة عالية في التنبوء.

ثانيا: التوصيات

١. باستعمال الدالة التمييزية (او التصنيفية) وتطبيقها في المستشفيات في تصنيف الاطفال حديثي الولادة الى اطفال طبيعيين واطفال خدج لغرض تقديم الرعاية والاسعاف لهم.

٢. استعمال هذا الأسلوب في الدراسات الاخرى للتنبؤ بانماذج تساعد في تصنيف المشاهدات الى مجاميعها.

٣.استعمال اساليب اخرى في التصنيف ومقارنتها بالدالة التمييزية وتحديد كفاءتها.

المصادر

١. الجبوري ، د. شلال حبيب وعبد, صلاح حمزة (٢٠٠٠), "تحليل متعدد المتغيرات"، دار الكتب للطباعة والنشر , جامعة بغداد، العراق.

٢.الجاعوني، فريد و غانم، عدنان (٢٠٠١),"التحليل الإحصائي متعدد المتغيرات (التحليل التمييزي) في توصيف وتوزيع الأسر داخل الهيكل الاقتصادي الاجتماعي في المجتمع"، مجلة جامعة دمشق للعلوم الاقتصادية والقانونية ،المجلد (٢٣) ،العدد(٢)، ص ص(٣١٣-٣٣١)

٣.المخلافي, فؤاد عبده اسماعيل ,"تصنيف وتمييز المحافظات اليمنية بحسب مصادر الدخل الفردي باستخدام أسلوبي التحليل العنقودي والتحليل التمييزي"، جامعة الناصر

٤.ابو علام, د. رجاء محمود," التحليل الاحصائي للبيانات باستخدام برنامج SPSS , القاهره .

٥. جودة - , محفوظ (٢٠٠٨), "التحليل الاحصائي المتقدم باستخدام " SPSS ، دار وائل للنشر، الطبعة الأولى ، عمان، الاردن .

٦.شاهين, حمزه اسماعيل و جباره , ازهار كاظم," مقارنة بين التحليل التمييزي واحتمال الاستجابة في تصنيف البيانات"، مجلة الادارة والاقتصاد ، العدد ١٠

٧.شاهين, حمزه اسماعيل (٢٠١٤), "مقارنه بين بعض طرائق التصنيف الخطية مع تطبيق عملي: , مجلة العلوم الاقتصادية والادارية , المجلد (٢٠) , العدد(٨٠)

٨. ذنون ، يونس ذنون (٢٠١٢)"استخدام طريقتي التحليل العنقودي والتحليل التمييزي في التصنيف مع تطبيق على نتائج الدرجات " ، مجلة تكريت للعلوم الإدارية والاقتصادية ، جامعة تكريت، المجلد(٨)،العدد (٢٥) ، ص ص(١٦٠-١٧٢)

٩. عزام ، عبد المرضي حامد (١٩٩٨), "التحليل الاحصائي للمتغيرات المتعددة من الوجهة التطبيقية" كتاب مترجم، دار المريخ - للنشر، الرياض ،المملكة العربية السعودية.

١٠.علي, مالك صالح و احسان , نازك(٢٠١١)," استخدام الدالة المميزة للتنبؤ بنتيجة الطالب" . المجلة العراقية , المجلد(٣), العدد(٥).

١١.يعقوب, م. م. اسماء ايوب (٢٠١٧), " التحليل العنقودي والتمييزي في دراسة تطبيقية على بعض المصارف العراقية" , مجلة الاقتصادي الخليجي, العدد(٣١)اذار .

12.Anderson,T.W.(1918)"An introduction to multivariate statistical analysis" by John Wiley &Sons ,Inc.

13. Dillon ,W.R and Goldstein ,M(1984)"Multivariate Analysis Methods and Application". John Wily &Sons ,Inc ,New York ,USA.

14.Hardle W., Simar L. (2003) , " Applied multivariate statistical analysis ", Berlin and Louvain-la-Neuve, Germany.

15. Johnson, R.A. and Wichern D.W. (2002): Applied Multivariate Statistical Analysis. New Jersey: Prentice Hall .

16. Neil H.Timm (2002) , " Applied multivariate analysis ",Springer verlag New York, Inc .

17..Rencher , Alvin C. (2002) , "Methed of Multivariate Analysis" ,2nd edit, by John Wiley & Sons , Inc . Published simultaneously in Canada

18.Pohar, Maja & Others(2004) ," Comparison of Logistic Regression and Linear Discriminant Analysis: A Simulation Study" , Metodološki zvezki, Vol. 1, No. 1, 2004, 143-161

A Comparison between Some Bayesian Estimation Methods for Parameter Exponential Distribution by Using Data Type II Censoring

Nabeel Ali Hussein [1], Sameer Annon Abbas [2], Ahmed Hadi Hussain[3,*]
and Abbas Musleh Salman[4]

Iraqi Ministry of Education 1, 2, 4
The Directorate of Education of the Governorate Babylon

University of Babylon, College of Engineering Al-musayab3
Department of Energy Engineering
Iraq

1E-mail: **nabeelali0086@gmail.com** 2E-mail: **sameer.annon@bab.epedu.gov.iq**
3E-mail: **met.ahmed.hadi@uobabylon.edu.iq** 4E-mail: **abbas.mmmm2019@gmail.com**

Corresponding Author: Ahmed Hadi Hussain

Abstract: Due to the importance of exponential distribution, especially in the modeling and analysis of various data in several fields in mathematics and statistics, including the applied field (experiments) and the pure field (theories). Some tests do not provide sufficient information on the trial data, so they are called “censored data”. The aim of this paper is to compare some Bayesian estimation methods for the exponential distribution of parameters using data type II censoring. We provided an estimate of the scale parameter (ESP) for the exponential distribution under data the control of type II by using the proposed Bayesian method and maximum probability. We also compared these methods using Mean square error (MSE). This study was conducted using simulation with different parameter values (θ) and different sample sizes (n = 10, 20, 50, 100). The calculation results showed that the best method of estimation is the proposed Bayesian method (BAY2), which uses the distribution of Chi-Square (n) in the previous information.

Keywords: Bayes Estimation, exponential distribution, Jeffery prior information, Maximum likelihood estimates.

I. Introduction:

The exponential distribution is the most important mathematical models, which has very many applications in several life’s areas in general and in mathematics in particular for both pure and applied where used in the modelling and analysis of data, in previous years, many researchers have reached to several results using this distribution see [1],[2] and [3].Usually, when the researcher performs some tests, the researcher may lack the ability to monitor and identify all the elements that have been selected on the basis of which the success of the experiment depends on its failure, the most important reasons: Some of them are related to the temporal factor of the experiment and some of them are related to constraints (such as financial cost and others) which negatively effect on the results. Therefore, when these effects are directly low with the number of failure observed states, the controlled experience of failure is better, more effective, and achieves less time and effort than the time-controlled experiment. After that, the researchers identified the experiment, which was controlled by failure, as surveillance type II. In this type of experiment, the test is terminated as soon as the number failures that pre-determined (r) is the number of units (n) that is tested. The researchers [4], [5], [6], [7],[8] and [9]presented the definition of predictive distribution in the life's areas and reached several results in this area. The methods of estimating Bayesian are studied in the classical exponential distribution (see [11]) and reached to the best methods of estimating through this distribution. In 2007, the researcher [12] also studied the Bayes estimation of the exponential distribution under double control of type II.Then the researchers used a set of previous distributions to estimate the parameter by exponential distribution. See [10]. The aim of this research is to compare some of the theoretical estimation methods for the exponential distribution of the parameters by using data type control II. We provided an estimate of the measurement parameter (ESP) of the exponential distribution under the data, and control of the second type that is using the proposed Bayes method and maximum probability. We also compared these methods using MSE (Mean Square Error). This study was conducted using simulation with different parameter values (θ) and different sample sizes (n= 10, 20, 50, 100).The calculated results showed that the best method of estimation is the proposed Bayesian method (BAY2), which uses the chi-Squared distribution (n) in the previous information.

II. Theoretical side

2.1 Exponential Distribution (E.D)

This distribution is commonly used to model waiting times between occurrences of rare events, lifetimes of electrical or mechanical devices. A continuous random variable X is said to have an Exponential distribution (E.D) with parameter (θ) if it has probability density function (p.d.f) and the cumulative distribution function (c.d.f) of (E.D) are given as follows respectively

$$f(x;\theta) = \frac{1}{\theta} e^{-\frac{x}{\theta}}, \quad x > 0 \,, \; \theta > 0 \qquad (1)$$

$$F(x;\theta) = 1 - e^{-\frac{x}{\theta}} \qquad (2)$$

2.2 Reliability function

Let X be a continuous random variable of distribution has parameter set $\theta = \{\theta_1, \theta_2, \dots, \theta_n\}$ and let $f(x;\theta)$ be its probability density function, where the survival time has parameter θ. The cdf of X is then

$$F(x;\theta) = p(X \leq x) = \int_{-\infty}^{x} f(t;\theta)\, dt$$

The survivor function or reliability function is defined as

$$R(x;\theta) = p(X > x) = 1 - F(x;\theta) \qquad (3)$$

In other words, the survivor function is the probability of survival beyond time.

2.3 Type II Censoring Data

Using this type of data mainly for clinical situations and the idea here is to select (m) the units so that m < n, n represents the size of the sample being studied. And the possible function of this data category is defined by the following formula in ascending order [3]:

$$L = (\theta | t_1, t_2, \dots, t_m, m) = \frac{n!}{(n-m)!} \prod_{i=1}^{m} f(t_i) \, [1 - F(t_0)]^{n-m} \qquad (4)$$

Such that

$1 - F(t_0) = R(t_0)$ represents the reliability function at time t_0

$f(t_i)$ represents density function of failure

$(n-m)$ the number of units is a failed after time t_0

For exponential distribution and by using (1) and (2) then equation (4) becomes as follows

$$L = (\theta | t_1, t_2, \dots, t_m, m) = \frac{n!}{(n-m)!} \prod_{i=1}^{m} \left(\frac{1}{\theta}\right) e^{-e^{-\frac{t}{\theta}}} \left[1 - \left(1 - e^{-\frac{t_0}{\theta}}\right)\right]^{n-m}$$

$$= \frac{n!}{(n-m)!} (\theta)^{-m} e^{-\frac{\sum_{i=1}^{m} t_i}{\theta}} \left[e^{-\frac{t_0}{\theta}}\right]^{n-m} \qquad (5)$$

When we take Log the two parties and from derivative second we get maximum likelihood estimator

$$\therefore \hat{\theta}_{MLE} = \frac{\sum_{i=1}^{m} t_i + (n-m) t_0}{m} \qquad (6)$$

III. Bayesian Estimation Methods

3.1 Standard Bayes Method (BAY1)

In this method the Bayes standard estimator is obtained in the case of the data under the control of type II. Then joint distribution function (pdf) as follows.

$$f(t_1, t_2, \dots, t_m, \theta) = g(t_1, t_2, \dots, t_m | \theta) \, g(\theta) \qquad (7)$$

By using Jeffery Prior information then density function of the posterior distribution as follows.

$$g(\theta) \propto k \sqrt{I(\theta)} \qquad (8)$$

I (θ) represent Fisher information, k constant

$$g(\theta) = \frac{k\sqrt{n}}{\theta} \; , \; \theta > 0 \qquad (9)$$

Then joint density function of T and θ described in equation (7) will be as follows

$$f(t_1, t_2, \dots, t_m, \theta) = \frac{n! k \sqrt{n}}{(n-m)! \, \theta^{m+1}} \; e^{-\frac{\left(\sum_{i=1}^{m} t_i + t_0 (n-m)\right)}{\theta}} \qquad (10)$$

Hence from (10) we find marginal density function of ***T*** is given by

$$f(t_1, t_2, \dots, t_m) = \int_0^{\infty} f(\underline{t}, \theta) \; d\theta$$

$$= \frac{n!\,k\sqrt{n}}{(n-m)!} \int_0^\infty \theta^{-m-1} e^{-\frac{\left(\sum_{i=1}^m t_i + t_0(n-m)\right)}{\theta}} d\theta \qquad (11)$$

By using transformation

$$u = \frac{\sum_{i=1}^m t_i + t_0(n-m)}{\theta} \Rightarrow \theta = \frac{\sum_{i=1}^m t_i + t_0(n-m)}{u} \Rightarrow d\theta = \frac{\sum_{i=1}^m t_i + t_0(n-m)\,du}{u^2}$$

Hence $f(t_1, t_2, \dots t_m) = \frac{n!\,k\sqrt{n}}{(n-m)!\left(\sum_{i=1}^m t_i + t_0(n-m)\right)^m} \int_0^\infty u^{m-1} e^{-u}\, du$

$$\therefore f(t_1, t_2, \dots t_m) = \frac{k\sqrt{n}\;\; n!\;\; \Gamma(m)}{(n-m)!\left(\sum_{i=1}^m t_i + t_0(n-m)\right)^m} \qquad (12)$$

Hence, density function of the posterior distribution of θ is given by

$$H^*(\theta\,|t_1, t_2, \dots, t_m) = \frac{f(t_1, t_2, \dots t_m, \theta_1)}{f(t_1, t_2, \dots t_m)} = \frac{\left(\sum_{i=1}^m t_i + t_0(n-m)\right)^m}{\theta^{m+1}\;\Gamma(m)} e^{-\frac{\left(\sum_{i=1}^m t_i + t_0(n-m)\right)}{\theta}} \qquad (13)$$

Let $\delta = \sum_{i=1}^m t_i + t_0(n-m)$, then

$$H^*(\theta\,|t_1, t_2, \dots, t_m) = \frac{\delta^m}{\theta^{m+1}\;\Gamma(m)} e^{-\frac{\delta}{\theta}} \qquad (14)$$

By using the quadratic loss function $c(\hat{\theta} - \theta)^2$, Then Bayes' estimator will be the estimator that minimizes the posterior risk given by

$$Risk\ (\theta) = E[c\,(\hat{\theta}, \theta)] = \int_0^\infty (\hat{\theta} - \theta)^2\; H^*(\theta\,|t_i)\; d\theta_1$$

$$Let\ \frac{\partial\ Risk\ (\theta)}{\partial\theta} = 0$$

Then $\hat{\theta}_{BAY1} = \int_0^\infty \theta\;\; H^*(\theta\,|t_i)\; d\theta = \int_0^\infty \frac{\delta^m}{\theta^m\;\Gamma(m)} e^{-\frac{\delta}{\theta}}\, d\theta$

By using transformation $u = \frac{\delta}{\theta} \Rightarrow \theta = \frac{\delta}{u} \Rightarrow d\theta = \frac{\delta\, du}{u^2}$

$$\therefore \hat{\theta}_{BAY1} = \frac{\sum_{i=1}^m t_i + (n-m)\,t_0}{m-1} \qquad (15)$$

3.2 Bayes method proposed (BAY2)

In this method, a Bayes estimator for the measurement parameter will be found using Natural Conjugate Prior and let this distribution is Gamma distribution(1 , μ) as Prior information where pdf of the gamma distribution as follows [6 , 14] .

$$f(\theta) = \frac{\mu}{\theta^2} e^{-\frac{\mu}{\theta}} ,\theta > 0 \qquad (16)$$

Then joint density function of T and θ and by use equation (4) and (7) will be as follows

$$f(t_1,t_2,..,t_m,\theta)=\frac{n!}{(n-m)!\,\theta^m}\; e^{-\frac{\left(\sum_{i=1}^{m} t_i+t_0(n-m)\right)}{\theta}}\;\frac{\mu}{\theta^2}\; e^{-\frac{\mu}{\theta}} \qquad (17)$$

Let $\delta=\sum_{i=1}^{m} t_i+t_0(n-m)$, then

$$f(t_1,t_2,..,t_m,\theta)=\frac{n!\,\mu}{(n-m)!\,\theta^{m+1}}\; e^{-\frac{(\delta+\mu)}{\theta}} \qquad (18)$$

Hence from (18) we find marginal density function of ***T*** is given by

$$f(t_1,\dots t_m)=\int_0^\infty f(\underline{t},\theta)\,d\theta=\frac{n!\,\mu}{(n-m)!}\int_0^\infty \frac{1}{\theta^{m+1}}e^{-\frac{(\delta+\mu)}{\theta}}d\theta$$

By using transformation $u=\frac{\delta+\mu}{\theta}\Rightarrow\theta=\frac{\delta+\mu}{u}\Rightarrow d\theta=\frac{\delta+\mu}{u^2}\,du$

$$f(t_1,t_2,\dots\dots t_m)=\frac{n!\,\mu}{(n-m)!}\int_0^\infty \frac{u^{m+1}}{(\delta+\mu)^{m+1}}e^{-u}\,\frac{\delta+\mu}{u^2}\,du$$

$$\therefore f(t_1,t_2,\dots\dots t_m)=\frac{n!\,\mu\;\Gamma(m)}{(n-m)!\,(\delta+\mu)^m} \qquad (19)$$

Hence, density function of the posterior distribution of $\boldsymbol{\theta}$ is given by

$$H^*(\theta\,|t_1,t_2,..,t_m)=\frac{f(t_1,t_2,\dots t_m,\theta)}{f(t_1,t_2,\dots t_m)}=\frac{(\delta+\mu)^m}{\theta^{m+1}\,\Gamma(m)}\;e^{-\frac{(\delta+\mu)}{\theta}} \qquad (20)$$

By using the quadratic loss function $c(\hat{\theta}-\theta)^2$, Then Bayes' estimator will be the estimator that minimizes the posterior risk given by

$$\hat{\theta}_{BAY2}=\int_0^\infty \theta\;\;H^*(\theta\,|t_i)\;d\theta=\int_0^\infty \frac{(\delta+\mu)^m}{\theta^m\,\Gamma(m)}\;e^{-\frac{(\delta+\mu)}{\theta}}\;d\theta$$

by using transformation $u=\frac{\delta+\mu}{\theta}\Rightarrow\theta=\frac{\delta+\mu}{u}\Rightarrow d\theta=\frac{\delta+\mu}{u^2}\,du$

$$\therefore\hat{\theta}_{BAY2}=\frac{\delta+\mu}{m-1}=\frac{\sum_{i=1}^{m} t_i+t_0(n-m)+\mu}{m-1} \qquad (21)$$

3.3 Bayes method proposed (BAY3)

Using Chi – square (n) distribution in Prior information at degree freedom (n) where the density function of Chi – square distribution as follows :

$$f(\theta)=\frac{\mu^{\frac{n}{2}-1}}{2^{\frac{n}{2}}\,\Gamma\left(\frac{n}{2}\right)}\;e^{-\frac{\mu}{\theta}}\;,I(0,\infty) \qquad (22)$$

Thus the joint density function of T and θ will be as follows

$$f(t_1,..,t_m,\theta)=\frac{n!}{(n-m)!\,\theta^m}\; e^{-\frac{\left(\sum_{i=1}^{m} t_i+t_0(n-m)\right)}{\theta}}\;\frac{\mu^{\frac{n}{2}-1}}{2^{\frac{n}{2}}\,\Gamma\left(\frac{n}{2}\right)}\; e^{-\frac{\mu}{\theta}} \qquad (23)$$

Let $\delta=\sum_{i=1}^{m} t_i+t_0(n-m)$, then

$$f(t_1,t_2,\ldots,t_m,\theta) = \frac{n!\ \mu^{\frac{n}{2}-1}}{(n-m)!\ \theta^m\ 2^{\frac{n}{2}}\ \Gamma\left(\frac{n}{2}\right)}\ e^{-\frac{(\delta+\mu)}{\theta}} \qquad (24)$$

Hence from (24) we find marginal density function of ***T*** is given by

$$f(t_1,\ldots t_m) = \int_0^\infty f(\underline{t},\theta)\ d\theta = \frac{n!\ \mu^{\frac{n}{2}-1}}{(n-m)!\ 2^{\frac{n}{2}}\ \Gamma\left(\frac{n}{2}\right)} \int_0^\infty \frac{1}{\theta^m}\ e^{-\frac{(\delta+\mu)}{\theta}} d\theta$$

By using transformation $u = \frac{\delta+\mu}{\theta} \Rightarrow \theta = \frac{\delta+\mu}{u} \Rightarrow d\theta = \frac{\delta+\mu}{u^2}\ du$

$$f(t_1,t_2,\ldots\ldots t_m) = \frac{n!\ \mu^{\frac{n}{2}-1}}{(n-m)!\ 2^{\frac{n}{2}}\ \Gamma\left(\frac{n}{2}\right)} \int_0^\infty \frac{u^m}{(\delta+\mu)^m}\ e^{-u}\ \frac{\delta+\mu}{u^2}\ du$$

$$\therefore f(t_1,t_2,\ldots\ldots t_m) = \frac{n!\ t^{\frac{n}{2}-1}\ \Gamma(m-1)}{(n-m)!\ 2^{\frac{n}{2}}\ \Gamma\left(\frac{n}{2}\right)\ (\delta+t)^{m-1}} \qquad (25)$$

Hence, density function of the posterior distribution of $\boldsymbol{\theta}$ is given by

$$H^*(\theta\ |t_1,t_2,\ldots,t_m) = \frac{f(t_1,t_2,\ldots t_m,\theta)}{f(t_1,t_2,\ldots t_m)} = \frac{(\delta+\mu\ )^{m-1}}{\theta^m\ \Gamma(m-1)}\ e^{-\frac{(\delta+\mu\ )}{\theta}} \qquad (26)$$

By using the quadratic loss function $c(\ \hat{\theta}-\theta\ )^2$. Then Bayes' estimator will be the estimator that minimizes the posterior risk given by

$$\hat{\theta}_{BAY3} = \int_0^\infty \theta\ H^*(\theta\ |t_i)\ d\theta = \int_0^\infty \frac{(\delta+\mu\ )^{m-1}}{\theta^{m-1}\ \Gamma\ (m-1)}\ e^{-\frac{(\delta+\mu\ )}{\theta}}\ d\theta$$

And by using the same transformation the previous

$$\therefore\ \hat{\theta}_{BAY3} = \frac{(\delta+\mu)\ \Gamma\ (m-2)}{\Gamma\ (m-1)} = \frac{(\delta+\mu)}{m-2} \qquad (27)$$

IV. Practical Aspect (Simulation):

Formulation of a model simulation includes the following essential and important steps for estimation of the scale parameter of exponential distribution that are respectively:

(P1) The initial values for the parameter θ

This step is important upon which later steps depend. Then we assume the initial values (θ =5.5, 6 , 6.5) for scale parameter θ of the of the exponential distribution.

(P2) Selected sample size (n)

We chose different sizes of the sample proportionally to the effect of sample size on the accuracy and efficiency of the results obtained from the estimation methods used, so we take the sizes (10, 30, 50, and 100).

(P3) The initial values for the time estimation of reliability function (t_0)

We take three values of the time $t_0 = 1 , 2 , 3$

(P4) Select values for the constants in the estimators

We take the value parameter $\mu = 1$.

(P5) Step of Data Generation:

In this step, the generation of weighted exponential distribution data using the inverse method is as follows

$$F_w\ (t_i;\beta) = U_i = 1 - e^{-t_i(\theta-1)} \qquad (28)$$

And the $t_i = \frac{-\log\ (1-U_i)}{\theta}$ (29)

Where $F_w\ (t_i;\theta)$ = The distribution function given in (2)

U_i =Uniformly distributed random variable on (0,1)

(P6) Measure comparison :

We adopt the mean square error $MSE(\ \hat{\theta}\ ) = \frac{\sum_{i=1}^{R}(\ \hat{\theta}_i - \theta)^2}{R}$

Where R=1000 is the number of replications .The tables below show the results of the estimation using the simulation. Program the simulation written by using (Matlab – 2011a)

V. Explanation Results (**Conclusions**)

The results of tables (1) to (4) show the following:

- Bayes method proposed estimation (BAY3) is the best, in all the size samples because it has the lowest (MSE).
- The second estimator (BAY2) is the best in the small size samples because it has the second lowest mean square error (MSE).
- We noted decreasing values of (MES) with increasing sample size of all cases , which corresponds with the statistical theory.
- (MLE) is in the last place by comparing with the other methods because it has achieved the highest level of (MSE) in all the sample sizes.

VI. Recommendations

- Applying this study by using other convenient continuous distributions as (Gamma, Beta, Weibull, ...).
- We recommend developing Bayes formulas that have been studied to other formats and under different loss functions.
- Testing the hypotheses theories for this study in the industrial fields and modeling machines failure times.

Table (1): Mean squared error for $\tilde{\theta}$ where n=10

θ	t_0	*MLE*	*BAY1*	*BAY2*	*BAY3*	*BEST*
٥,٥	١	٠,٠٢٨٢٣٤٧٦٠٥٩	٠,٠٢٨١٩٣٧٦٠٠٦	٠,٠٢٧٩٧٧٣٤٢٨٣	٠,٠٢٧٩٣٠٣٣٣٣٩	BAY3
	٢	٠,٠٢٨٢٠٩٣٢٨٧٧	٠,٠٢٨٣٧١٧٦٧٢٨	٠,٠٢٨١٥٤٦٦٦٦٠	٠,٠٢٨١١١٢٠٥٤٦	BAY3
	٣	٠,٠٢٧٧١١١٤٠٠٧	٠,٠٢٧٦٥٩٨٩٤٣٣	٠,٠٢٧٤٤٥٥٣٩٨٥	٠,٠٢٧٣٨٧٩٦٣٩٢	BAY3
٦	١	٠,٠٣٤٣٢٢٤٥٨٤٧	٠,٠٣٤٢٨٧٩٢٢١١	٠,٠٣٤٠٤٩٢١٥٨٠	٠,٠٣٤٠٠٨٤٢٢٧٦	BAY3
	٢	٠,٠٣٤٠٥٢٢٤١٤٨	٠,٠٣٤٠١٢٣٤٠٦٥	٠,٠٣٣٧٧٤٥٩٧٢٢	٠,٠٣٣٧٢٨٢٦٣١١	BAY3

Table (2): Mean squared error for $\tilde{\theta}$ where n=30

θ	t_0	*MLE*	*BAY1*	*BAY2*	*BAY3*	*BEST*
٥,٥	١	٠,٠٢٨٩٩٨٢٥٧٨٧	٠,٠٢٨٥٢٥٠٩٧٤٢	٠,٠٢٨٣٤٩٩٩٧٣٠	٠,٠٢٨١٩٤١٩٤٨	BAY3
	٢	٠,٠٢٨٢٩٠٤٥٦٦٤	٠,٠٢٨٢٦٢٤٧٠٨٧	٠,٠٢٧٨٩٧٠٣٩٣٦	٠,٠٢٧٧٨٨٢٢٢٤٤	BAY3
	٣	٠,٠٢٨٣٨٥٨٦٥٢٠	٠,٠٢٨١٧٩١٥١٧٦	٠,٠٢٧٠١١٢٧٤٢٢	٠,٠٢٦٦١٥٧٧٤٣١	BAY3
٦	١	٠,٠٣٤٢٧٧٨٢٧٩٥	٠,٠٣٤٠٨٥١٨٥٧٦	٠,٠٣٢٧٩٩٥٠٥٣٠	٠,٠٣٢٤٠٥٦٤٤٧٧	BAY3
	٢	٠,٠٣٣٦٦٩٢٧٤٤٩	٠,٠٣٣٤١١٢٦٢٩٥	٠,٠٣٢١٣٨٤٧٨٧٧	٠,٠٣١٦٦٦٩٦٤٢٩	BAY3
	٣	٠,٠٣٤٢٥٩٣٦٩١٧	٠,٠٣٤٠٦٤٧٣٤٠٤	٠,٠٣٢٧٧٩٤٤٣٠٥	٠,٠٣٢٣٨٣٢١١١٢	BAY3
٦,٥	١	٠,٠٣٩٩٦٩٧٨٩٥٨	٠,٠٣٩٧١٥٧٨٢١٣	٠,٠٣٨٣٢٦٩٨٤٧٨	٠,٠٣٧٨٤٥٠٤٣١٢	BAY3
	٢	٠,٠٤٠٣٥٣٠٦٥٩٤	٠,٠٤٠١٤٠٤٠٥٥٥	٠,٠٣٨٧٤٤١٣٧٨٤	٠,٠٣٨٣١١٥٤٧١٦	BAY3
	٣	٠,٠٤٠٤٦٦٧٧٧٠٣	٠,٠٤٠٤٠٤٥٤٦٥٢	٠,٠٣٩٩٦٧١٤١٠٧	٠,٠٣٩٨٨٥٣١٥٧٣	BAY٣

Table (2): Mean squared error for $\tilde{\theta}$ where n=30

θ	t_0	*MLE*	*BAY1*	*BAY2*	*BAY3*	*BEST*
٥,٥	١	٠,٠٢٨٨٠٨٠٨٥٦٩	٠,٠٢٨٧٥٥٩٧٤٣٩	٠,٠٢٨٣٨٩١٤٢٤٤	٠,٠٢٨٣٢٢٧٢٩٥٢	BAY3
	٢	٠,٠٢٨٢٨٥٠١٧٩١	٠,٠٢٨٠٢٧٧٦١٠٣	٠,٠٢٧٨٦٢٣٥٥٤٩	٠,٠٢٧٤٧٧٨٦١٥٤	BAY3
	٣	٠,٠٢٧٩٥٤٧٩٣٩٧	٠,٠٢٧٨٧٦٢٦١٧٥	٠,٠٢٧٥١٣١٤٥٨٣	٠,٠٢٧٤١٦٧٧١٣٩	BAY3
	١	٠,٠٣٣١٢٧٠٠٣٨١	٠,٠٣٣٠٢٨٨٧٥٦٤	٠,٠٣٢٦٣٣٥١٧٥٤	٠,٠٣٢٥١٥١١٩٠٣	BAY3

٦	٢	٠,٠٣٤٢٥٠٨٣٤٠٤	٠,٠٣٤١٩٠٠٨٤٢٠	٠,٠٣٣٧٨٧٨١٥٥١	٠,٠٣٣٧٠٨٨٦١١٠	BAY3
	٣	٠,٠٣٤١١١٠٦٢٢٢	٠,٠٣٤٠٧٢٣٢٧١٨	٠,٠٣٣٨٣٤٣٧٣٨٣	٠,٠٣٣٧٨٩٢٤٣٧٧	BAY3
٦,٥	١	٠,٠٤٠٣١٨٤٦٤٣٠	٠,٠٤٠٢٥١٢٤٢٥٥	٠,٠٣٩٨١٤٦٦٩٩٥	٠,٠٣٩٧٢٧٥٦٦٥٠	BAY3
	٢	٠,٠٤٠٣٣٨٦٢١٧٨	٠,٠٤٠٢٧٢٠٧٧٨١	٠,٠٣٩٨٣٥٣٩١٩٣	٠,٠٣٩٧٤٩٠٠٥١٧	BAY3
	٣	٠,٠٤٠١١٧١٦٥١٧	٠,٠٤٠٠٧٣٣٧٢٧١	٠,٠٣٩٨١٥٢٧٨٥٦	٠,٠٣٩٧٦٤٤٦٨٩١	BAY٣

Table (3): Mean squared error for $\hat{\theta}$ where n=50

θ	$t_\circ$	MLE	BAY1	BAY2	BAY3	BEST
٥,٥	١	٠,٠٢٨٠٧٥٣٨١٦٧	٠,٠٢٨٠٥٣٥٣٥١٢	٠,٠٢٧٩٤٦٥٨٢٥٣	٠,٠٢٧٩٢٣٢٥١٥٦	BAY3
	٢	٠,٠٢٨١٠٩٩٤٦٩٥	٠,٠٢٨٠٨٨٤٣٦٠٨	٠,٠٢٧٩٨١٤١٦٩١	٠,٠٢٧٩٥٨٤٢٦٨٢	BAY3
	٣	٠,٠٢٧٢١٤٣٢٢٠٨	٠,٠٢٧١٩٣٨٢٦١٨	٠,٠٢٧١٨٦٦٠٦٢٤	٠,٠٢٧١٥٢٩١٩٦٧	BAY3
٦	١	٠,٠٣٤٢٥٢٣١١٦٦	٠,٠٣٤٢٣٤٥٢٥٢١	٠,٠٣٤١١٦٣٦٥٦١	٠,٠٣٤٠٩٧٠٤٧٩٤	BAY3
	٢	٠,٠٣٣٧١٠٥٩٨٦٣	٠,٠٣٣٦٨٧٥٠٥٠١	٠,٠٣٣٥٧٠٢٩٤٠٥	٠,٠٣٣٥٤٥٥٨٣٧٥	BAY3
	٣	٠,٠٣٣٧٥٠٣١٤٠٦	٠,٠٣٣٧٢٧٦٠٧٩٨	٠,٠٣٣٦١٠٣٢٧٢٢	٠,٠٣٣٥٨٦٠١٠٦٦	BAY3
٦,٥	١	٠,٠٤٠٠١٠٤٦٤٤٧	٠,٠٣٩٩٦٤٥٥٥٤٠	٠,٠٣٩٧٠٦٨١٢٤٨	٠,٠٣٩٦٥٣٨١٥٤١	BAY3
	٢	٠,٠٣٩٨٣٢٨٠٢١٢	٠,٠٣٩٨٠٨٣٣٤٦٣	٠,٠٣٩٦٨٠٩١٠٦٧	٠,٠٣٩٦٥٤٦٩٢٨٠	BAY3
	٣	٠,٠٣٨٧٨٩١٣١٦٩	٠,٠٣٨٤٠٩٢٣٨٩٤	٠,٠٣٧٠٤٣٦٨١٤٥	٠,٠٣٦٤١٢٠٤٦٥٥	BAY٣

References

[١] Epstein, B.1954. Truncated life tests in the exponential case. *The Annals of Mathematical Statistics*, 25(3):555-564.

[٢] A. E. Luis and Q. M. William.1999. Statistical Prediction Based on Censored Life Data. *Statistics Preprints*, 11(1999):113-123.

[٣] Epstein, B.1960. Tests for the validity of the assumption that underlying distribution of life is exponential. Part I. *American Statistical Association and American Society for Quality*, 2(1):83-101.

[٤] Epstein, B.1960. Tests for the validity of the assumption that the underlying distribution of life is exponential. Part II. *Technometrics*, 2(1):167-183.

[٥] Al-metwally E.M. & Mubarak A .S. 2018. Bayesian and Maximum likelihood Estimation for Weibull Generalized Exponential Distribution. *Pakistan Journal of Statistics and Operation Research,* XIV(4):853-868.

[٦] Epstein, B., and Sobel, M.1954. Some problems relevant to life testing from an exponential distribution. *The Annals of Mathematical Statistics*, 25(2):373-381.

[٧] Hahn, G. J.1975. A Simultaneous prediction limit on the mean of future samples from an exponential

distribution. *Technometrics*, *17*(3):341-345.
[٨] Al-Kutubi, Hs, and Ibrahim , Na.2009. Bays Estimation for Exponetial Distribution with Extention of Jeffery Prior Information. *Malaysian Journal of Mathematical Sciences,*3(2):297-313.
[٩] Dunsmore, I. R..1974. The Baysian Predictive Distribution In Life Testing. *Technometrics,*16(3):455-466.
[1٠] Haq, A., and Dey, S.2011. Bayesian Estimation of Erlang Distribution Under different Prior Distributions. *Journal of Reliability and Statistical Studies*, *4*(1): 1-30.
[1١] Singh S.K.& Singh U. and Yadav A.S.2014. Bayesian Reliability Estimation of Inverted Exponential Distribution Under Progressive Type II Censored Data. *Journal of statistics Applications* & *probability An international Journal*,3(3):317-333.
[1٢] Singh U, and Kumar A.2007. Bayesian Estimation of The Exponential Parameter Under a Multiply Type-II Censoring. *Austrian Journal of Statistics*,36(3):227-238.

Estimate survival function by using Dagum distribution

Lecturer/ Hind Jawad Kadhim Al-Bderi
College of Administration and Economics / University of Al-Qadisiyah
hind.kadhum@qu.edu.iq
hind_satia@yahoo.com
http://orcid.org/0000-0003-1525

Abstract: This paper intends to estimate the unlabeled three parameters for Dagum distribution modle depend on censored samples type two ; employing the maximum likelihood estimator method to obtain the derivation of the point estimators for all unlabeled parameters depending on iterative techniques , as Newton – Raphson method ; then to derive ordinary least squares estimator method. Applying all these methods to estimate related probability functions; death density function, cumulative distribution function, survival function and hazard function (rate function).

When examining the numerical results for probability survival function by employing ‘mean squares error measure and mean absolute percentage measure’ , this may lead to work on the best method in modeling a set of real data; and this method is “Maximum likelihood estimator method for real censored type II sample”.

Keywords:-

Censored type two sample, Maximum Likelihood Estimator Method, Ordinary least squares estimator method, Survival Function, .

1- Introduction:-

The Dagum distribution was proposed in a series of papers in the seventies by Camilo Dagum. It is applied to the analysis of income distributions which are greatly related to personal income; for example, in 2015 in Poland, researchers constructed a confidence interval to measure the economic inequality through empirical analysis.(1)

Another study of modeling air quality and estimate potential air pollution over Pekanbaru city, Indonesia for seven years data, from 2009 to 2015; this study showed the factors that spoil the urban air quality by the use of the three-parameters Dagum distribution. The maximum likelihood method and the L moment estimator were compared. Probability density function (PDF) , (CDF) and (AIC) were applied to check the best criteria of the distribution. (2)

Frequently, a study of five years data, from 2011-2015 over Durban in south Africa was made to analize the significant parameters of rain attenuation for telecommunication links at heigher frequencies. This study deals with the tree-parameters Dagum distribution of the rain height variations each month, each season and each year; and again by using (ML) method of the distribution which showed a decent modeling of rain height over Durban city with the Root Mean Square Error of 0.26. (3)

Some other paper presented a lifetime model under the name the Weibull-Dagum distribution. The suggested modle determined an ordinary conversion of the Weibull random variable. The density function of this modle is very pliable, left-skewed and right-skewed. Maximum likelihood is employed to estimate the model parameters and show the ability of the new model by means of simulation. Almost the suggested model works better than beta-Dagum and McDonald-Dagum. (4)

This paper is divided as follows:- The objective of this paper, theoretical section, practical section, results and conclusions.

2- The objective of this paper:-

This paper aims to examine censored sample type two blood cancer by using MLEM and OLSEM ; and by comparing the two methods .

3- Definition and properties:-

The p.d.f for Dagum distribution is: $f_p(t;\chi,\delta,\phi) = \chi\delta\phi\ t^{-\phi-1}(1+\delta t^{-\phi})^{-\chi-1}$ ....(1) $t > 0$

$$\Omega = \left\{ \left(t;\chi,\delta,\phi\right) ;\ \chi\ ,\ \phi > 0\ ,\ \delta > 0 \right\}$$

Where δ : is scale (real) parameter

χ : is shape (real) parameter

ϕ : is shape (real) parameter

The cumulative distribution function for this distribution is: $F(t;\chi,\delta,\phi) = (1+\delta t^{-\phi})^{-\chi}$ ....(2)

Its survival function is given by: $S(t;\chi,\delta,\phi) = 1-(1+\delta t^{-\phi})^{-\chi}$ ....(3)

Its hazard rate function is given by: $h(t;\chi,\delta,\phi) = \dfrac{\chi\delta\phi\ t^{-\phi-1}(1+\delta t^{-\phi})^{-\chi-1}}{1-(1+\delta t^{-\phi})^{-\chi}}$ ....(4)

4- Maximum likelihood estimator method(MLEM) for censored type II sample:-

The MLM is the most common procedure to estimate the parameter λ which specifies a p.f. $f(t:\lambda)$, based on the observations $t_1, t_2, ..., t_n$ which were independently. sample from the distribution. .

$$L = \frac{m!}{(m-r)!}[\prod_{i=1}^{r} f(t_i)][1-F(T_0)]^{n-r} \qquad 0 \le t_{(1)} \le t_{(2)} \le ..... \le t_{(m)} \le t_{(0)} \ ....(5)$$

Let $\dfrac{m!}{(m-r)!} = c$

$$L = c[\prod_{i=1}^{r} \chi\delta\phi\ t^{-(1+\phi)}(1+\delta t^{-\phi})^{-(1+\chi)}][1-(1+\delta T_0^{-\phi})^{-\chi}]^{n-r} \ ....(6)$$

Applying the logarithm for the likelihood function:

$$\ln L = \ln c + \sum_{i=1}^{r} \ln \chi_i + \sum_{i=1}^{r} \ln \delta_i + \sum_{i=1}^{r} \ln \phi_i - (1+\phi)\sum_{i=1}^{r} \ln t_i - (1+\chi)\sum_{i=1}^{r} \ln(1+\delta t_i^{-\phi}) + (n-r)\ln(1)$$
$$+(n-r)\chi \ln(1+\delta T_0^{-\phi}) \ldots\ldots\ldots\ldots\ldots\ldots(7)$$

$$\frac{\partial \ln L}{\partial \chi} = \sum_{i=1}^{r} \frac{1}{\chi_i} - \sum_{i=1}^{r} \ln(1+\delta t_i^{-\phi}) + (n-r)\ln(1+\delta T_0^{-\phi}) \ldots\ldots\ldots\ldots\ldots\ldots(8)$$

$$\frac{\partial \ln L}{\partial \chi} = 0 \quad ; \quad \sum_{i=1}^{r} \frac{1}{\chi_i} - \sum_{i=1}^{r} \ln(1+\delta t_i^{-\phi}) + (n-r)\ln(1+\delta T_0^{-\phi}) = 0 \ldots\ldots\ldots\ldots\ldots\ldots(9)$$

$$\frac{\partial \ln L}{\partial \delta} = \sum_{i=1}^{r} \frac{1}{\delta_i} - (1+\chi)\sum_{i=1}^{r} \frac{t_i^{-\phi}}{(1+\delta t_i^{-\phi})} + (n-r)\chi \frac{T_0^{-\phi}}{(1+\delta T_0^{-\phi})} \ldots\ldots\ldots\ldots\ldots\ldots(10)$$

$$\frac{\partial \ln L}{\partial \delta} = 0 \quad ; \quad \sum_{i=1}^{r} \frac{1}{\delta_i} - (1+\chi)\sum_{i=1}^{r} \frac{t_i^{-\phi}}{(1+\delta t_i^{-\phi})} + (n-r)\chi \frac{T_0^{-\phi}}{(1+\delta T_0^{-\phi})} = 0 \ldots\ldots\ldots\ldots\ldots\ldots(11)$$

$$\frac{\partial \ln L}{\partial \phi} = \sum_{i=1}^{r} \frac{1}{\phi_i} - \sum_{i=1}^{r} \ln t_i + (1+\chi)\sum_{i=1}^{r} \frac{\delta t_i^{-\phi} \ln t_i}{(1+\delta t_i^{-\phi})} - (n-r)\chi \frac{\delta T_0^{-\phi} \ln T_0}{(1+\delta T_0^{-\phi})} \ldots\ldots\ldots\ldots\ldots\ldots(12)$$

$$\frac{\partial \ln L}{\partial \phi} = 0 \quad ; \quad \sum_{i=1}^{r} \frac{1}{\phi_i} - \sum_{i=1}^{r} \ln t_i + (1+\chi)\sum_{i=1}^{r} \frac{\delta t_i^{-\phi} \ln t_i}{(1+\delta t_i^{-\phi})} - (n-r)\chi \frac{\delta T_0^{-\phi} \ln T_0}{(1+\delta T_0^{-\phi})} = 0 \ldots\ldots\ldots\ldots\ldots\ldots(13)$$

There is no chance to find the estimators for the parameters (χ, δ, ϕ) , and it is kind of difficulty to process the nonlinear equations thus, it is better to make use of iterative methods in numerical analysis as Newton–Raphson method which is the best way to get the estimate values and number of iteration.

The Newton–Raphson method requires an initial value of each unknown parameters (χ, δ, ϕ) .

$$\begin{bmatrix} \chi_{i+1} \\ \delta_{i+1} \\ \phi_{i+1} \end{bmatrix} = \begin{bmatrix} \chi_i \\ \delta_i \\ \phi_i \end{bmatrix} - J_i^{-1} \begin{bmatrix} f(\chi) \\ g(\delta) \\ z(\phi) \end{bmatrix} \ldots\ldots\ldots\ldots\ldots\ldots(14)$$

$$f(\chi) = \sum_{i=1}^{r} \frac{1}{\chi_i} - \sum_{i=1}^{r} \ln(1+\delta t_i^{-\phi}) + (n-r)\ln(1+\delta T_0^{-\phi}) \ldots\ldots\ldots\ldots\ldots\ldots(15)$$

$$g(\delta) = \sum_{i=1}^{r} \frac{1}{\delta_i} - (1+\chi)\sum_{i=1}^{r} \frac{t_i^{-\phi}}{(1+\delta t_i^{-\phi})} + (n-r)\chi \frac{T_0^{-\phi}}{(1+\delta T_0^{-\phi})} \ldots\ldots\ldots\ldots\ldots\ldots(16)$$

$$z(\phi) = \sum_{i=1}^{r} \frac{1}{\phi_i} - \sum_{i=1}^{r} \ln t_i + (1+\chi)\sum_{i=1}^{r} \frac{\delta t_i^{-\phi} \ln t_i}{(1+\delta t_i^{-\phi})} - (n-r)\chi \frac{\delta T_0^{-\phi} \ln T_0}{(1+\delta T_0^{-\phi})} \ldots\ldots\ldots\ldots\ldots\ldots(17)$$

$$J_i^{-1} = \begin{bmatrix} \frac{\partial f(\chi)}{\partial \chi} & \frac{\partial f(\chi)}{\partial \delta} & \frac{\partial f(\chi)}{\partial \phi} \\ \frac{\partial g(\delta)}{\partial \chi} & \frac{\partial g(\delta)}{\partial \delta} & \frac{\partial g(\delta)}{\partial \phi} \\ \frac{\partial z(\phi)}{\partial \chi} & \frac{\partial z(\phi)}{\partial \delta} & \frac{\partial z(\phi)}{\partial \phi} \end{bmatrix} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots(18)$$

$$\frac{\partial f(\chi)}{\partial \chi} = -\sum_{i=1}^{r} \frac{1}{\chi_i^2} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots(19)$$

$$\frac{\partial f(\chi)}{\partial \delta} = -\sum_{i=1}^{r} \frac{t_i^{-\phi}}{1+\delta t_i^{-\phi}} + (n-r)\frac{T_0^{-\phi}}{1+\delta T_0^{-\phi}} \quad \ldots\ldots\ldots\ldots\ldots(20)$$

$$\frac{\partial f(\chi)}{\partial \phi} = \sum_{i=1}^{r} \frac{\delta t_i^{-\phi} \ln t_i}{1+\delta t_i^{-\phi}} - (n-r)\frac{\delta T_0^{-\phi} \ln T_0}{1+\delta T_0^{-\phi}} \quad \ldots\ldots\ldots\ldots\ldots(21)$$

$$\frac{\partial g(\delta)}{\partial \chi} = -\sum_{i=1}^{r} \frac{t_i^{-\phi}}{1+\delta t_i^{-\phi}} + (n-r)\frac{T_0^{-\phi}}{1+\delta T_0^{-\phi}} \quad \ldots\ldots\ldots\ldots\ldots(22)$$

$$\frac{\partial g(\delta)}{\partial \delta} = -\sum_{i=1}^{r} \frac{1}{\delta_i^2} + (1+\chi)\sum_{i=1}^{r} \frac{(t_i^{-\phi})^2}{(1+\delta t_i^{-\phi})^2} - (n-r)\chi\frac{(T_0^{-\phi})^2}{(1+\delta T_0^{-\phi})^2} \quad \ldots\ldots\ldots\ldots\ldots(23)$$

$$\frac{\partial g(\delta)}{\partial \phi} = (1+\chi)\sum_{i=1}^{r} \frac{t_i^{-\phi} \ln t_i}{(1+\delta t_i^{-\phi})^2} - (n-r)\chi\frac{T_0^{-\phi} \ln T_0}{(1+\delta T_0^{-\phi})^2} \quad \ldots\ldots\ldots\ldots\ldots(24)$$

$$\frac{\partial z(\phi)}{\partial \chi} = \sum_{i=1}^{r} \frac{\delta t_i^{-\phi} \ln t_i}{1+\delta t_i^{-\phi}} - (n-r)\frac{\delta T_0^{-\phi} \ln T_0}{1+\delta T_0^{-\phi}} \quad \ldots\ldots\ldots\ldots\ldots(25)$$

$$\frac{\partial z(\phi)}{\partial \delta} = (1+\chi)\sum_{i=1}^{r} \frac{t_i^{-\phi} \ln t_i}{(1+\delta t_i^{-\phi})^2} - (n-r)\chi\frac{T_0^{-\phi} \ln T_0}{(1+\delta T_0^{-\phi})^2} \quad \ldots\ldots\ldots\ldots\ldots(26)$$

$$\frac{\partial z(\phi)}{\partial \phi} = -\sum_{i=1}^{r} \frac{1}{\phi^2} - (1+\chi)\sum_{i=1}^{r} \frac{\delta t_i^{-\phi} (\ln t_i)^2}{(1+\delta t_i^{-\phi})^2} + (n-r)\chi\frac{\delta T_0^{-\phi} (\ln T_0)^2}{(1+\delta T_0^{-\phi})^2} \quad \ldots\ldots\ldots\ldots\ldots(27)$$

$$\begin{bmatrix} \varepsilon(\chi)_{i+1} \\ \varepsilon(\delta)_{i+1} \\ \varepsilon(\phi)_{i+1} \end{bmatrix} = \left\| \begin{bmatrix} \chi_{i+1} \\ \delta_{i+1} \\ \phi_{i+1} \end{bmatrix} - \begin{bmatrix} \chi_i \\ \delta_i \\ \phi_i \end{bmatrix} \right\| \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots(28)$$

5- <u>Ordinary Least Squares Estimator Method (OLSEM)</u>:-

The OLSEM is the most used way to estimate parameters in linear or nonlinear model. Researchers make use of this method to lessen the sum squares differences concerning observed sample values and expected estimated values by linear approximation. (5,6)

$$Y = \alpha_0 + \alpha_1 x + \varepsilon \;\; ....(29)$$

$$\varepsilon = Y_i - \hat{\alpha}_0 - \hat{\alpha}_1 x_i \;\; ....(30)$$

$$F(t_i) = [1 + \delta T^{-\phi}]^{-\chi} \;\; ....(31)$$

$$\sum_{i=1}^{n} \varepsilon_i^2 = \sum_{i=1}^{n} [y_i - \hat{y}_i]^2 \;\; ....(32)$$

$$\sum_{i=1}^{n} \varepsilon_i^2 = \sum_{i=1}^{n} [y_i - \hat{\alpha}_0 - \hat{\alpha}_1 x_i]^2 \;\; ....(33)$$

$$[F(t_i)]^{-\frac{1}{\chi}} = 1 + \delta T^{-\phi} \;\; ....(34)$$

$$\varepsilon = [F(t_i)]^{-\frac{1}{\chi}} - 1 - \delta T^{-\phi} \;\; ....(35)$$

$$\sum_{i=1}^{n} \varepsilon_i^2 = \sum_{i=1}^{n} [\{F(t_i)\}^{-\frac{1}{\chi}} - 1 - \delta T^{-\phi}] \;\; ....(36)$$

Let $\sum_{i=1}^{n} \varepsilon_i^2 = B(\chi, \delta, \phi)$

$$B(\chi, \delta, \phi) = \sum_{i=1}^{n} [\{F(t_i)\}^{-\frac{1}{\chi}} - 1 - \delta T^{-\phi}] \;\; ....(37)$$

$$\frac{\partial B}{\partial \chi} = \sum_{i=1}^{n} [F(t_i)]^{-\frac{1}{\chi}} \ln F(t_i) \frac{1}{\chi^2} \;\; ....(38)$$

$$\frac{\partial B}{\partial \chi} = 0 \quad ; \quad \sum_{i=1}^{n} [F(t_i)]^{-\frac{1}{\chi}} \ln F(t_i) \frac{1}{\chi^2} = 0 \;\; ....(39)$$

$$\frac{\partial B}{\partial \delta} = -\sum_{i=1}^{n} T_i^{-\phi} \;\; ....(40)$$

$$\frac{\partial B}{\partial \delta} = 0 \quad ; \quad -\sum_{i=1}^{n} T_i^{-\phi} = 0 \;\; ....(41)$$

$$\frac{\partial B}{\partial \phi} = \sum_{i=1}^{n} [\delta T_i^{-\phi} \ln T_i] \;\; ....(42)$$

$$\frac{\partial B}{\partial \phi} = 0 \quad ; \quad \sum_{i=1}^{n} [\delta T_i^{-\phi} \ln T_i] = 0 \; ....(43)$$

There is no chance to find the estimators for the parameters (χ, δ, ϕ) , and it is kind of difficulty to process the nonlinear equations thus, it is better to make use of iterative methods in numerical analysis as Newton–Raphson method which is the best way to get the estimate values and number of iteration.

The Newton–Raphson method requires an initial value of each unknown parameters (χ, δ, ϕ) .

$$\begin{bmatrix} \chi_{i+1} \\ \delta_{i+1} \\ \phi_{i+1} \end{bmatrix} = \begin{bmatrix} \chi_i \\ \delta_i \\ \phi_i \end{bmatrix} - J_i^{-1} \begin{bmatrix} z_1(\chi) \\ z_2(\delta) \\ z_3(\phi) \end{bmatrix} \; ....(44)$$

$$z_1(\chi) = \sum_{i=1}^{n} [F(t_i)]^{-\frac{1}{\chi}} \ln F(t_i) \frac{1}{\chi^2} \; ....(45)$$

$$z_2(\delta) = -\sum_{i=1}^{n} T_i^{-\phi} \; ....(46)$$

$$z_3(\phi) = \sum_{i=1}^{n} [\delta T_i^{-\phi} \ln T_i] \; ....(47)$$

$$J_i^{-1} = \begin{bmatrix} \frac{\partial z_1(\chi)}{\partial \chi} & \frac{\partial z_1(\chi)}{\partial \delta} & \frac{\partial z_1(\chi)}{\partial \phi} \\ \frac{\partial z_2(\delta)}{\partial \chi} & \frac{\partial z_2(\delta)}{\partial \delta} & \frac{\partial z_2(\delta)}{\partial \phi} \\ \frac{\partial z_3(\phi)}{\partial \chi} & \frac{\partial z_3(\phi)}{\partial \delta} & \frac{\partial z_3(\phi)}{\partial \phi} \end{bmatrix} \; ....(48)$$

$$\frac{\partial z_1(\chi)}{\partial \chi} = \sum_{i=1}^{n} \frac{1}{\chi^3} [F(t_i)]^{-\frac{1}{\chi}} \ln F(t_i) [\frac{1}{\chi} - 2] \; ....(49)$$

$$\frac{\partial z_1(\chi)}{\partial \delta} = 0 \; ....(50) \qquad \frac{\partial z_1(\chi)}{\partial \phi} = 0 \; ....(51)$$

$$\frac{\partial z_2(\delta)}{\partial \chi} = 0 \; ....(52) \qquad \frac{\partial z_2(\delta)}{\partial \delta} = 0 \; ....(53)$$

$$\frac{\partial z_2(\delta)}{\partial \phi} = \sum_{i=1}^{n} T_i^{-\phi} \ln T_i \; ....(54)$$

$$\frac{\partial z_3(\phi)}{\partial \chi} = 0 \quad \ldots\ldots\ldots\ldots\ldots(55)$$

$$\frac{\partial z_3(\phi)}{\partial \delta} = \sum_{i=1}^{n} T_i^{-\phi} \ln T_i \quad \ldots\ldots\ldots\ldots\ldots(56)$$

$$\frac{\partial z_3(\phi)}{\partial \phi} = -\sum_{i=1}^{n} \delta T_i^{-\phi} (\ln T_i)^2 \quad \ldots\ldots\ldots\ldots\ldots(57)$$

$$\begin{bmatrix} \varepsilon(\chi)_{i+1} \\ \varepsilon(\delta)_{i+1} \\ \varepsilon(\phi)_{i+1} \end{bmatrix} = \left\| \begin{bmatrix} \chi_{i+1} \\ \delta_{i+1} \\ \phi_{i+1} \end{bmatrix} - \begin{bmatrix} \chi_i \\ \delta_i \\ \phi_i \end{bmatrix} \right\| \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots(58)$$

6- Results and Discussion

The Educational Hospital in Diwaniya province was the place from which the data was gathered.

Keeping in mind that this work relies on data taken from real life, it is reached to select this kind of cancer (blood cancer) because it is remarkable widespread and deadly in Iraq; this disease has failure time (death time) which is phenomenon in this paper.

The study of this paper covers a period of six months; it begins from 30/6/2019 untill 31/12/2019; it is an experiment that includes(12) patients. (10) patients were dead and (2) patients remain alive .

When applying the test statistic (Kolmogorov-Smirnov) depending upon statistical programming (EasyFit 5.5 Professional) in order to fit Dagum distribution data , it is discovered that the calculated value is (0.12294) , this means data is distributed according to Dagum distribution .

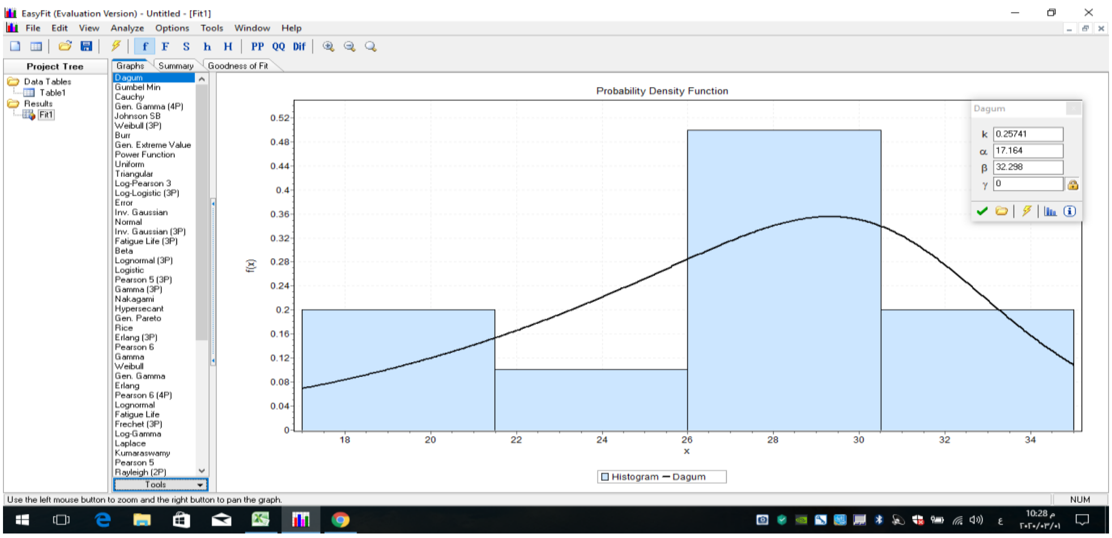
The null and alternative hypotheses are as follows :

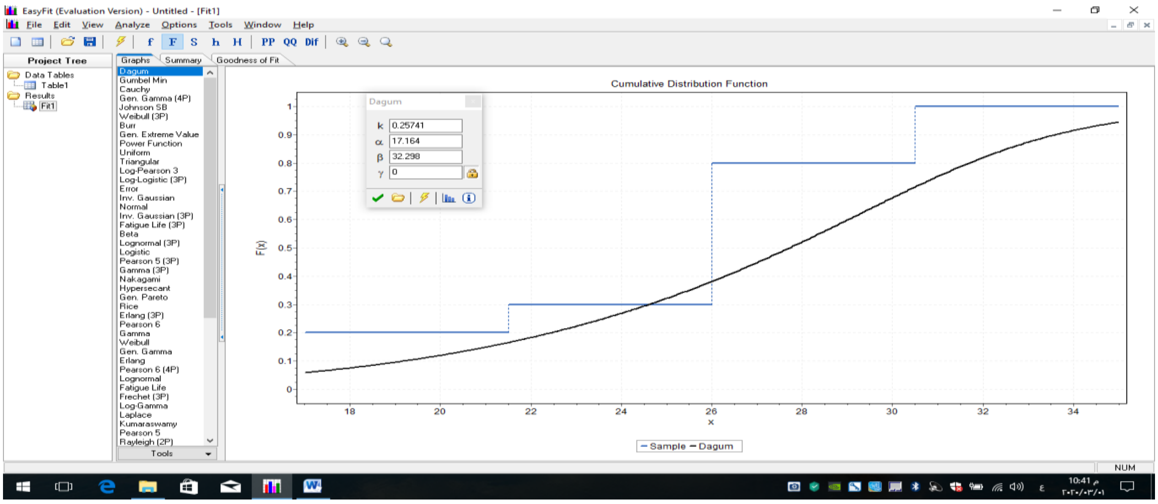
H_0 : The survival time data is distributed as Dagum.

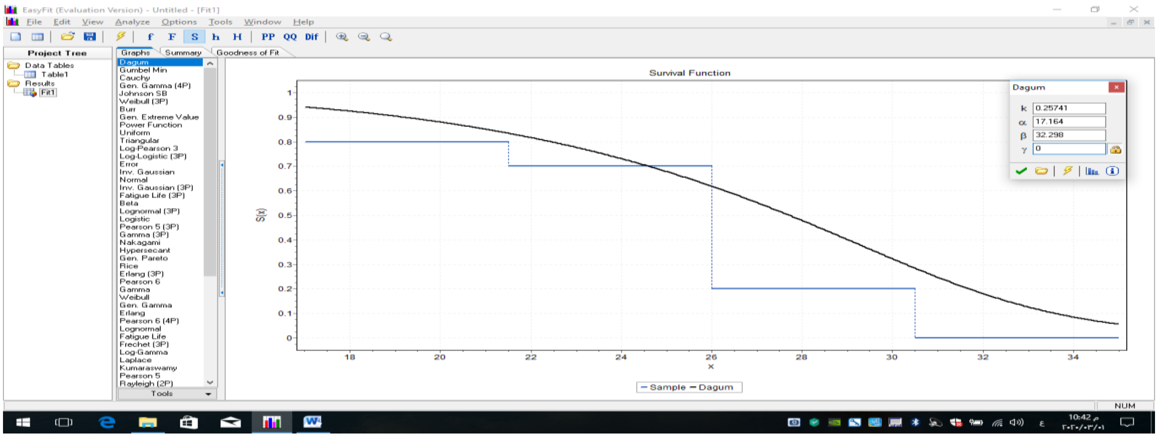
H_1 : The survival time data is not distributed as Dagum.

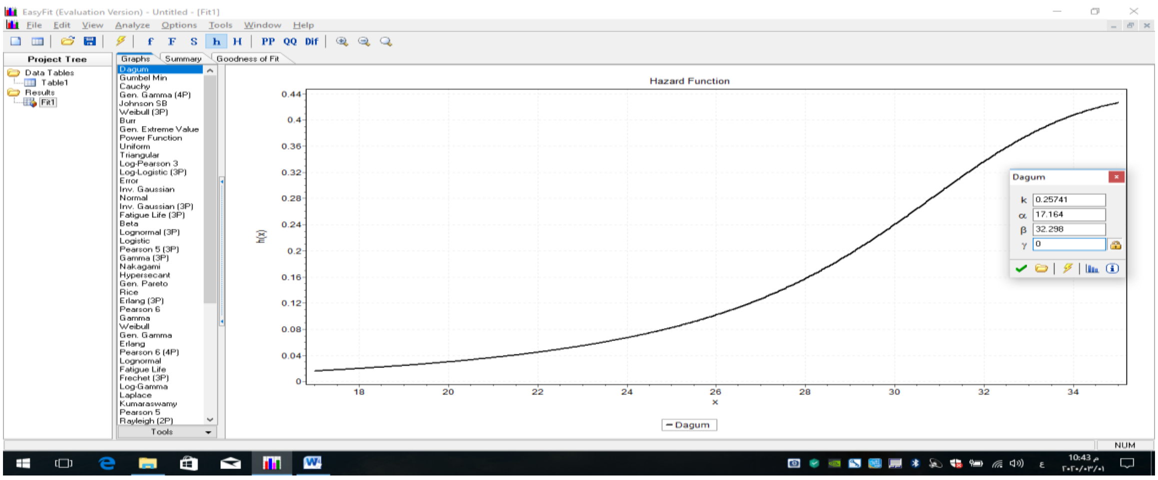
'Figure(1)'

Fit the datafor blood cancer from 'Educational Hospital Diwanyia'









When applying MATAB(R014a) , the estimated parameters results are as follows :

$\chi_0 = 0.92$; $\delta_0 = 0.2$; $\phi_0 = 0.5$

The assumed initial values for two-parameters are as follows:

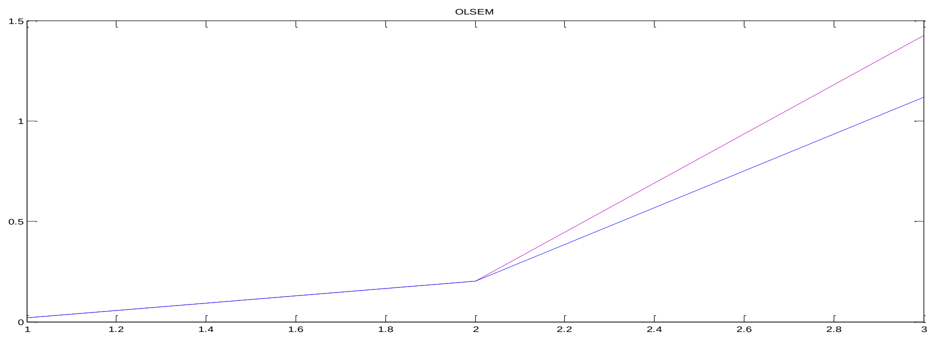
Table (1) : Estimated values for the parameters by MLEM for censored data type II

Estimate values	No. of iteration	Errors for all parameters
$\hat{\chi} = 28.2537$	33	0.8229
$\hat{\delta} = 0.8369$		0.2631
$\hat{\phi} = 1.5000$		0.0001

Table (2) : Estimated values for the parameters by OLSEM

Estimate values	No. of iteration	Errors for all parameters
$\hat{\chi} = 0.0212$	3	0.0004
$\hat{\delta} = 0.2031$		0.0010
$\hat{\phi} = 1.4262$		0.3091

After that , using ths estimated values for tree-parameters in Dagum dis. to find the numerical values for $f(t)$, $F(t)$, $s(t)$ and $h(t)$.



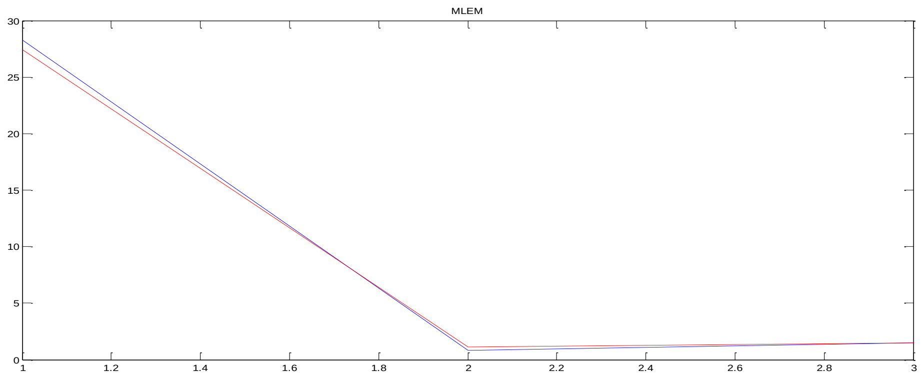


Table (3) : Estimated values for functions $f(t)$, $F(t)$, $s(t)$, $h(t)$ by MLEM

Failure Time/day	$\hat{f}(t)$	$\hat{F}(t)$	$\hat{s}(t)$	$\hat{h}(t)$
17	**0.021034074**	**0.715089254**	**0.995650654**	**0.004368346**
19	**0.016796441**	**0.752710014**	**0.995348166**	**0.004673575**
23	**0.011207044**	**0.807705732**	**0.993798547**	**0.006240151**
26	**0.008559691**	**0.83711328**	**0.993302026**	**0.00674314**
28	**0.007250765**	**0.852874195**	**0.992749235**	**0.007303722**
28	**0.007250765**	**0.852874195**	**0.992749235**	**0.007303722**
29	**0.006697974**	**0.859843552**	**0.991440309**	**0.008633592**
30	**0.006201453**	**0.866288886**	**0.988792956**	**0.011334066**
34	**0.004651834**	**0.887791134**	**0.983203559**	**0.017083381**
35	**0.004349346**	**0.892289384**	**0.978965926**	**0.021486012**

$$MSE[\hat{s}(t_i)] = \frac{1}{n}\sum_{i=1}^{n}[\hat{s}(t_i) - s(t_i)]^2 = 0.279313179$$

$$MAPE[\hat{s}(t_i)] = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{\hat{s}(t_i) - s(t_i)}{s(t_i)}\right| = 1.913506175$$

Table (4) : Estimated values for functions $f(t)$, $F(t)$, $s(t)$, $h(t)$ by OLSEM

Failure Time/day	$\hat{f}(t)$	$\hat{F}(t)$	$\hat{s}(t)$	$\hat{h}(t)$
17	**6.32892E-06**	**0.999924423**	**0.9999989**	**1.10013E-06**
19	**4.83465E-06**	**0.999935492**	**0.99999882**	**1.18022E-06**
23	**3.04351E-06**	**0.99995086**	**0.999998401**	**1.59857E-06**
26	**2.26128E-06**	**0.999958735**	**0.999998265**	**1.73547E-06**
28	**1.88953E-06**	**0.99996287**	**0.99999811**	**1.88954E-06**
28	**1.88953E-06**	**0.99996287**	**0.99999811**	**1.88954E-06**
29	**1.73547E-06**	**0.999964681**	**0.999997739**	**2.26129E-06**
30	**1.59857E-06**	**0.999966347**	**0.999996956**	**3.04351E-06**
34	**1.18022E-06**	**0.999971845**	**0.999995165**	**4.83467E-06**
35	**1.10013E-06**	**0.999972984**	**0.999993671**	**6.32896E-06**

$$MSE[\hat{s}(t_i)] = \frac{1}{n}\sum_{i=1}^{n}[\hat{s}(t_i) - s(t_i)]^2 = 0.284998487$$

$$MAPE[\hat{s}(t_i)] = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{\hat{s}(t_i) - s(t_i)}{s(t_i)}\right| = 1.928964296$$

7- Conclusions:

1- We notice in both methods that the estimated values of the probability survival function decrease with increasing failure times (an inverse relationship between them).

2- We notice in both methods that the estimated values of the potential risk function increase with increasing times of failure (a direct relationship between them).

3- It is recommended to use (MLEM) of Dagum distribution of blood cancer by employing MSE criterion.

References:

(1) Alina Jdrzejczak, Dorota Pekasiewicz and Wojciech Zieliski. “ Confidence Interval for Quantile Ratio of the Dagum Distribution ”. arXiv:1903.04223v1 , 201٩ Mar 11; PP. 1-13.

(2) Evi Febriantikasari, Arisman Adnan, Rado Yendra and M. N. Muhaijir. “ Using Dagum Distribution to Simulated Concentration PM10 in Pekanbaru City, Indonesia ”. Applied Mathematical Sciences, Vol. 13, 2019, No. 9, PP 439 – 448.

(3) E. O. Olurotimi, O. Sokoya, J. S. Ojo and P. A. Owolawi.“ Distribution of rain height over subtropical region: Durban, South Africa for satellite communication systems ”. Materials Science and Engineering. (2018). PP 1-4.

(4) M. H. TAHIR, GAUSS M , CORDEIRO, M. MANSOOR M. ZUBAIR AND MORAD ALIZADEH. “ The Weibull-Dagum Distribution: Properties and Applications ” . Journal Communications in Statistics - Theory and Methods Volume 45, 13 Sep 2016 PP. 7376-7398.

(5) Douglas C. M. and George C. R. “ Applied Statistics and Probability for Engineering ” . Third Edition , John Wiley and Sons , Inc , 2003 PP. 1-976.

(6) Hutcheson, G. D. and Sofroniou , N. “ The Multivariate Social Scientist ” , London : Sage Publications , 1999 PP. 1-228 .

Estimate the parameters of Weibull distribution by using nonlinear membership function by Gaussian function

Suhaila N. Abdullh[1] Iden Hasan Huessian[2]
1 ALkARKH University of science, college of science, Baghdad , Iraq
Sss333iraq@yahoo.com
2 University of Baghdad, college of science for women, Baghdad, Iraq
Iden_ALkanani@yahoo.com

Abstract: The main aim of the presented study is estimating the parameters of Weibull distribution by utilizing simulation to generated the samples size when n=10, 50,100. Considering in the current study the parameters estimator of Weibull membership function, then using the nonlinear membership function for Gaussian function to find the fuzzy number for these parameters estimator. After that utilizing the ranking function to transform the fuzzy number to crisp number
Keywords: Weibull distribution, simulation technique, nonlinear membership function, the ranking function.

Introduction

In 1951, Waloddi Weibull (Swedish mathematician) was the first to describe Weibull distribution. With regard to the probability theory and statistics, Weibull distribution is specified as the continuous probability distribution.
Weibull distribution has been a useful density function in reliability and survival analysis, which is a famous distribution in medical and industry.
In the case when certitudes happen, individuals will usually look back to the acquired data and attempt to evaluate future events. Usually, a major methodology; that is commonly applied as the probability theory has fulfilled such necessities to handle im-precision and un-certainty. Yet, with regard to the fulluncertainty conditions, the probability theory could not be adequate, also there must be integration between fuzzy logic and probability theory for enhancing the robustness. Full-uncertainty could be defined as nobody has the data on the incidence of likely conditions furthermore, in certain conditions cases no one knows anything regarding such becoming true possible events.
Zadeh in 1968 was first structure probability measures in fuzzy sets. Kwakernaak defined fuzzy random variables (FRVs), also the authors indicated many concepts related to the independent fuzzy variables for the first time in his paper [3]. After one year, he added algorithms related to the fuzzy random variable, also the author provided examples related to discrete case [4].
Pak et al. in (2013) [5] developed inferential procedures according to the fuzzy environment with Weibull distribution.
. Pak et al. in (2014) [6] developed inferential procedures according to the exponential distribution according to fuzzy data. Shafiq and Viertl in (2014) [9] calculated characterizing function related to fuzzy parameter estimate for Weibull parameters based on censored fuzzy data. Pak et al. in (2016) [7] applied the algorithm of Newton-Raphson for determining the maximum probability estimate of shape parameter of lognormal distribution when the observations are fuzzy. Al-Sultany in ;(2016)[10]; evaluated the performance of moment maximum likelihood and Bayes estimation method for estimating the unknown parameters and reliability function of inverse Weibull distribution according to the fuzzy data. In this paper, generating the data of Weibull distributions by utilizing the Monte carol technique through using simulation method, with different sample size when n=10, 50, 100 respectively. After that finding the unknown parameters of Weibull distribution by using the nonlinear membership function for Gaussian function to get the lower and the upper

fuzzy parameters number for inconsistent sample size utilizer, this method enables the construction of normal fuzzy number which con be adapted to have Gaussian shape.
Finally, using the ranking function to transform the fuzzy parameters number to crisp for various sample size.

Fuzzy set

Fuzzy sets have been a main field of focus in many ways starting from its initiation in the year 1965.
Some applications related to Fuzzy sets can be seen in robotics, operation research, logic, decision theory, medicine, AI, computer science, control engineering, expert systems management science, and pattern recognition.
The development in mathematics increased largely in the past as well as huge advances recently. In the presented study, an elementary mathematical system related to the Fuzzy set is going to be discussed, in addition to the major applications of Fuzzy sets to the other methods and concepts. Soft computing or computational intelligence was the nme used to describe Fuzzy sets, neural nets theory, and evolutionary programming since the year 1992.
The associated between such extents has become mainly close. In the presented study, the Fuzzy sets will be the main focus. The applications related to the fuzzy sets to real problems abound. Certain references are going to be provided. To describe even a part of them will definitely exceed the extent of the presented study; John Wiley & Sons [2].

I. GAUSSIAN FUNCTION

The membership definition for a Gaussian function G: → [0, 1] is given by two parameters as: Where α is the midpoint and K reflects the slop value. Note that K must be positive and that the function never reaches zero. The Gaussian function can also be extended to have different left and right slops. We then have three parameters in Where and are, respectively, left and right slopes.
I.1- G-membership function
Theorem 1:
Let by fuzzy number with Gaussian membership function as
Let and x=u(x) then
by taken root
If then
If then
These are a fuzzy number in a parametric form where .
If we inspection that the membership function is Gaussian function substituting then the fuzzy number of Gaussian membership function becomes as follow:
Let and x=u(x) then
By taken root
If then
If then
These are a fuzzy number in a parametric form where .

I.2 membership function e^[-(x-1)]^2
IV. Ranking function
Ranking fuzzy numbers are of high importance in data analysis, they are applied in forecasting in addition to the decision-making optimization. The approach of ranking has been initially suggested via Jain (1976). Yager (1981) suggested 4 indices which could be used to order the fuzzy quantities in [0,1].
Ghen and Ghen (2007)[1] suggested an approach for ranking the generalized fuzzy trapezoidal fuzzy number generalized fuzzy numbers. Compos and Gonzales (1989) suggested a subjective method for

ranking fuzzy numbers. Ramil and Mohamad [8] suggested a complete survey related to various approaches for ranking of fuzzy numbers.

Ranking of fuzzy is of high importance in forecasting, making optimizations, and risk analysis decision.

Fuzzy sets have been developed via Zadeh, that is a tool of high power to handle real-life cases. From the time when Yager suggested centroid system in the ranging method.

The advantage of Ranking function which applies in operation research, statistics and mathematics is to transform the fuzzy number to crisp number.

Where Then the Ranking function becomes as follow:

Where is left shape function of Is left shape function of Then = and = is weight function then and GMF equal to Gaussian membership function Theorem: if is a GMF in a parametric form then

Proof: the parametric form in Gaussian membership function is as follow:

Recall that

V. PARAMETER ESTIMATION BY MLE

The method of Maximum Likelihood Estimation as this technique gives a simpler Estimation as compared to the Method of Moments and the Local frequency ratio method of estimation. Now we are estimating the parameter of the Weibull distribution from which the sample comes. Let be a random sample of n observations from the Weibull population with pdf Let

I. PRACTICAL SIDE:

In this section, we will create a simulated environment to find the fuzzy parameters of Weibull distribution as described in the following steps.

1- Based on we can write the Matlab program where then

, And the Initial value =0.9 , =0.1 we got,

- n=10 then
- n=50 then
- n=100 then

2- to calculate the lower and upper fuzzy parameters using the eq(1),eq(2) Where n=10 , then

Then

K=34.602, in the same way

Where n= 50→ K=0.149, and Where n=100→ K=0.427

Table I-1The lower and upper fuzzy parameters

n=10

3- To determine The Ranking value by using the eq(3)

Where n=10 , 50,100, and let k=1 and $\alpha \geq 0.5$ we get :

Table I-2 Ranking value for parameters

n=10 n=50 n=100

0.5 0.217 32.684 3.367 0.0331 1.954 0.1422

0.6 0.394 32.86 3.544 0.21 2.131 0.16

0.7 0.571 33.04 3.721 0.387 2.308 0.496

0.8 0.748 433.26 3.898 0.564 2.485 0.673

0.9 0.925 33.39 4.075 0.741 2.662 0.85

CONCLUSION:

We propose a nonlinear membership function for Gaussian function in statistic for first time in Iraq. In table (1) the lower fuzzy parameters number are increasing and the upper fuzzy parameters number are decreasing. In table (2) the ranking function of parameters number are increasing we take $\alpha \geq 0.5$ because the value of α greater than and equal to (0.5) it is normal concave.

Reference:

[1] Chan S.J and Chan S.M.; (2007);"Fuzzy Risk analysis based on the ranking of generalized Trapezoidal fuzzy numbers;Applied intelligence; vol.1;pp1-11.

[2] John Wiley & Sons, (2010)," Fuzzy set theory", vol 2 , issue 3,pp 317-332.

[3] Kwakernaak, H., (1978)," Fuzzy random variables-I. Definitions and theorems." Inf, Sci. 15, 1–29.

[4] Kwakernaak, H., (1979)," Fuzzy random variables-II. Algorithms and examples for the discretecase", Inf. Sci. 17, 253–278.

[5] Pak A., Parham G.A and Saraj M., (2013),"Inference for the weibull distribution based on fuzzy data" Revista colombiana deEstadistica , VOL 36, NO.2,PP339-358.

[6] Pak A., Parham G.A and Saraj M., (2014),"Inference on the competing Risk Reliability problem for exponential distribution based on fuzzy data" IEEE Transactions on Reliability, VOL .63, NO.1,PP 2-12.

[7] Pak A., Parham G.A and Saraj M., (2016),"Inference for the shape parameter of lognormal distribution in presence of fuzzy data", Pakistan Journal of statistic and operation Research, VOL.12, NO.1, PP 89-99.

[8] Ramli N. and Mohamad D.; (2009);"A Comparative analysis of centroid method in ranking fuzzy numbers "; European Journal of scientific research;vol. 28;pp492-501.

[9] Shafiq M. and Viertl R., (2014),"Maximum likelihood Estimation for weibull distribution in case of censored fuzzy lifetime datd", Vienna university of Fechnology, pp1-17.

[10] Al-Sultany S.A-K., (2016),"Estimate the parameters and Reliability function of Inverse weibull distribution based on linear membership, Ph.D, Thesis of submitted to college of science at the University of Mustansiryah.