



Comparison of Some Methods for Estimating the Survival Function and Failure Rate for the Exponentiated Expanded Power Function Distribution

Maysam Sajit Khdair⁽¹⁾

University of Baghdad, College of Administration
and Economics, Baghdad, Iraq

Maysam.sajit1201a@coadec.uobaghdad.edu.iq

Emad Hazim Aboudi⁽²⁾

University of Baghdad, College of Administration
and Economics, Baghdad, Iraq

emadhazim@coadec.uobaghdad.edu.iq

Received: 10/8/2022

Accepted: 4/9/2022

Published: September / 2022



This work is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International \(CC BY-NC 4.0\)](https://creativecommons.org/licenses/by-nc/4.0/)

Abstract

We have presented the distribution of the exponentiated expanded power function (EPPF) with four parameters, where this distribution was created by the exponentiated expanded method created by the scientist Gupta to expand the exponential distribution by adding a new shape parameter to the cumulative function of the distribution, resulting in a new distribution, and this method is characterized by obtaining a distribution that belongs for the exponential family. We also obtained a function of survival rate and failure rate for this distribution, where some mathematical properties were derived, then we used the method of maximum likelihood (ML) and method least squares developed (LSD) to estimate the parameters and because of the nonlinear relationship between the parameters, numerical algorithms were used to find the estimates of the two methods. They are Newton-Raphson (NR) and Nelder mead (NM) algorithms to improve the estimators, and a Monte Carlo simulation experiment was conducted to evaluate the performance of the two algorithms' estimates, and the average integrated error criterion (IMSE) was used to compare the survival function estimates and the failure rate. The results showed the efficiency of the maximum likelihood method estimates and least squares developed using the two algorithms (NR, NM) where their results were close, and this shows the new distribution efficiency (EPPF) for modeling survival data.

Paper type :Research paper.

Keywords: Expanded Exponentiated Power function distribution, Survival and failure rate, Maximum likelihood, Developed least squares, Newton Raphson, Nelder mead.

(1)

(2)

1. Introduction

In this unexpected scientific world, the probability distributions are equivalent, which represent an inevitable role in explaining the real world phenomenon. In the theory of distribution so far, the distribution of the power function (PF) is considered one of the simplest and most easy life distributions, and it represents the inverse of the Pareto distribution and is also considered a special case of the beta distribution. The distribution of the power function is better than the exponential distribution, the logarithmic distribution, the Whipple distribution and other distributions in representing life data, and the recent developments have focused on creating new flexible expanded distributions from the classical distributions and the developed distributions to facilitate the construction of probabilistic models to represent survival data and data types. The other is better, and the exponentiated expansion method is one of the well-known methods where several new classes of distributions can be developed by adding one or more parameters to the probability distribution or set of distributions studied. Gupta and Kundu (2001) studied a new family distribution the Exponentiated Exponential (EE). Cordeiro, et al (2013) introduced proposed a new class of generalized exponentiated distributions that can be interpreted as a dual construct for Lehman substitutions. Ahmed, et al (2015) introduced a new distribution called the exponentially shifted Rayleigh distribution (ETGR). Pena-Ramirez, et al. (2018) proposed a new model called the Generalized Weibull Exponentiated Power Distribution (EPGW). Afify and Zayed (2018) introduced a new extension of the exponentiated distribution called the Half-Logistic Exponentiated Distribution. Andrade, et al (2019) studied a new generalization of the Extended Gompertz distribution called the Exponentiated Extended Gompertz Distribution (EGEG) that consists of a major extension of the Extended Gompertz. Aldahlan, et al (2020) presented the family of generalized Weibull power chain distributions (EPGWPS) with an exponentiated power, which was obtained by summing up the series Weibull distributions and the generalized power of the exponent. Zamani, et al (2022) introduced a new distribution called the expanded exponentiated Chin distribution (EE-C).

2. Exponentiated Expanded Power Function (EPPF)

Statistical models are commonly used to predict real-life data, although many univariate models are suitable for real life; extended models are more flexible to explain the real-life phenomenon in terms of failure rate and survival analysis. Expanded power (EPF) is commonly used to explain finite-rare datasets, and this distribution was constructed from an extended power function (PF) distribution, which is a special case of the beta distribution and is also the inverse of the Pareto distribution. And the probability density function (PDF) for the expanded power function distribution (EPF) is given by the following formula:

$$g(x, a, l, b) = \frac{lbx^{lb-1}}{a^{lb}} \quad 0 < x < a ; \quad a, l, b > 0 \quad (1)$$

Whereas (l, b) represent The Shape parameter and (a) is The Scale parameter. Also, the Cumulative distributive function (CDF) is given in the following:

$$G(x) = \left(\frac{x}{a}\right)^{lb} \quad 0 < x < a \quad (2)$$

A new distribution can be constructed using the method of exponentiated expansion by adding a new parameter of the form of the cumulative function of the exponentiated power function (EPF), and thus we can derive a new distribution that contains four parameters by the formula:

$$F(x) = [G(x)]^c \quad (3)$$

Recompense the formula (2) into (3), we get the following:

$$F(x) = \frac{x^{lbc}}{a^{lbc}} \quad 0 < x < a \quad (4)$$

By deriving the formula (4), we get the probability density function of the new distribution (EPPF), as follows:

$$f(x, a, l, b, c) = \frac{lbcx^{lbc-1}}{a^{lbc}} \quad 0 < x < a ; a, l, b, c > 0 \quad (5)$$

The pdf function of the exponentiated expanded power function distribution has four parameters that are (a) the scale parameter and the shape parameter (l, b, c). (Rather,2018)

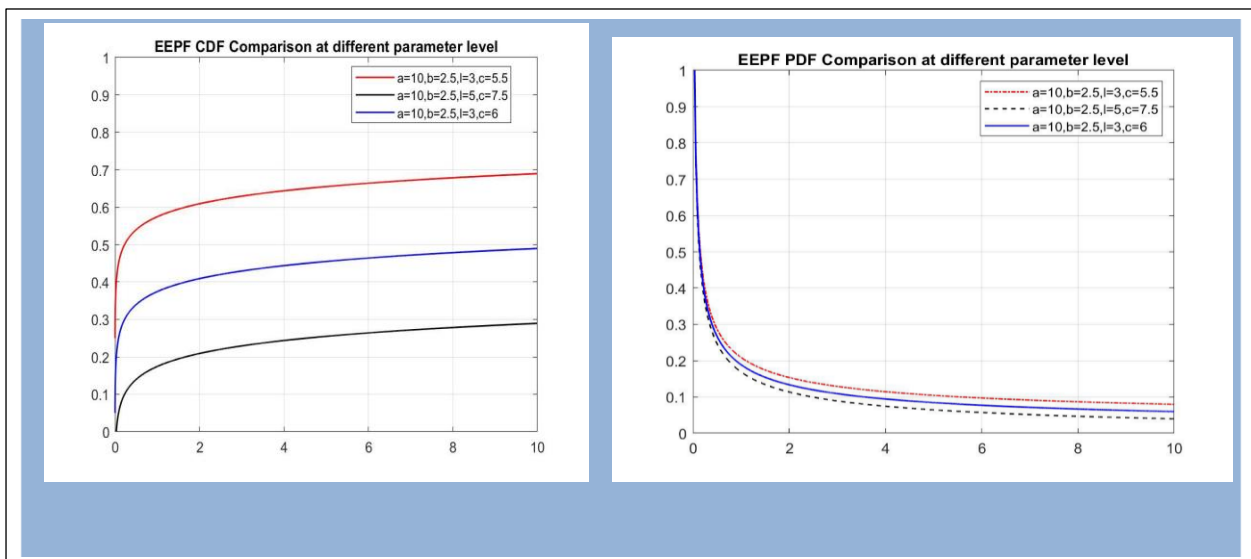


Figure 1: shows the CDF & PDF function of the distribution and the new exponentiated expanded power function (EPPF).

Some properties of the new probability distribution (EPPF) were derived as in the following table:-

Table 1: Represents the characteristics of the new probability distribution (EPPF)

Mean	$E(x) = \frac{lbc a}{lbc + 1}$
Variance	$Var(x) = \frac{lbc a^2}{(lbc + 1)^2 (lbc + 2)}$
Standard deviation	$\sigma = \frac{a \sqrt{lbc}}{(lbc + 1) \sqrt{lbc + 2}}$
Variation coefficient	$C \cdot V = \frac{\sqrt{lbc}}{lbc \sqrt{lbc + 2}} * 100$
Coefficient of Skewness	$C \cdot S = \frac{2(1 - lbc) \sqrt{lbc + 2}}{(lbc + 3) \sqrt{lbc}}$
Coefficient of Kurtosis	$C \cdot K = \frac{3(lbc + 2)(3l^2 b^2 c^2 - lbc + 2)}{lbc(lbc + 3)(lbc + 4)}$

2.1. Survival Function and Failure rate

It is the survival of the experimental unit for a period of not less than (x). In other words, if x is a random variable referring to the experimental unit, then S(x) represents the probability of survival of the experimental unit for the next period. The survival function can be expressed by the following formula:

$$S(x) = 1 - F(x) \quad (6)$$

And by recompensing formula (4) into formula (6), we get the general formula for the survival function of the new distribution (EPPF), which is:

$$S(x) = \frac{a^{lbc} - x^{lbc}}{a^{lbc}} \quad (7)$$

Failure rate, which is the probability of failure to occur in the subsequent period time, knowing that the item was in good condition, and is expressed as the following:

$$h(x) = \frac{f(x)}{S(x)} \quad (8)$$

Where (x) is the specific time between two-time intervals and substituting formula (5) and (7) into formula (8), we get the failure rate function for the new distribution which is:

$$h(x) = \frac{lbc x^{lbc-1}}{a^{lbc} - x^{lbc}} \quad (9)$$

3. Estimation

We will deal with the estimation of the parameters of the new distribution of the expanded exponential power function (EPPF) using two methods of estimation, which are the maximum likelihood method and the developed least squares method.

3.1. Maximum Likelihood (ML)

This method is considered one of the most important estimation methods because of its good properties, and one of the most important properties is the stability property. For this method, it is the parameter values that make the probability work at its maximum limit. If (x_1, x_2, \dots, x_n) represents the

vocabulary of a random sample of size (n), then the maximum-likelihood function is denoted by (L). (Abody and Nuimai, 2016)

$$L = f(x_1, \gamma) \cdot f(x_2, \gamma) \dots \dots \dots f(x_n, \gamma) \quad (10)$$

Where $(\gamma = a, l, b, c)$

function (PDF) The new distribution (EEPF) is

$$f(x, \gamma) = \frac{lbcx^{lbc-1}}{a^{lbc}}$$

$$L_f = \left(\frac{lbc}{a^{lbc}}\right)^n \prod_{i=1}^n x_i^{lbc-1} \quad (11)$$

Taking the logarithm of both sides, we get the following:

$$\ln L_f = n \ln l + n \ln b + n \ln c - n(lbc) \ln a + (lbc - 1) \sum_{i=1}^n \ln x_i \quad (12)$$

To find the estimators Maximum Likelihood for the unknown parameters, we derive the first derivative of the formula (12) as follows:

$$\frac{d \ln L_f}{da} = -\frac{nlbc}{a}$$

$$= 0$$

$$\frac{d \ln L_f}{dl} = \frac{n}{l} - nbc \ln a + bc \sum_{i=1}^n \ln x_i$$

$$= 0$$

$$\frac{d \ln L_f}{db} = \frac{n}{b} - nlc \ln a + \lambda \theta \sum_{i=1}^n \ln x_i$$

$$= 0$$

$$\frac{d \ln L_f}{dc} = \frac{n}{c} - nlb \ln a + lb \sum_{i=1}^n \ln x_i$$

$$= 0$$

When the above derivatives are equal to zero, it is not possible to obtain accurate estimates when solving them by analytical methods, so we will use numerical methods to find estimates for unknown parameters (a, l, b, c), and one of these methods is Newton-Raphson algorithm, and this method depends on the variance and covariance matrix, so we find the second derivative of the parameters and partial derivatives as follows:

$$\frac{d^2 \ln L_f}{d^2 a} = -\frac{nlbc}{a^2}$$

$$\frac{d^2 \ln L_f}{d^2 l} = -\frac{n}{l^2}$$

$$\frac{d^2 \ln L_f}{d^2 b} = -\frac{n}{b^2}$$

$$\frac{d^2 \ln L_f}{d^2 c} = -\frac{n}{c^2}$$

$$\frac{d^2 \ln L_f}{da dl} = -\frac{c}{nlc}$$

$$\frac{d^2 \ln L_f}{da db} = -\frac{a}{nlb}$$

$$\frac{d^2 \ln L_f}{da dc} = -\frac{a}{nlb}$$

$$\frac{d^2 \ln Lf}{dl db} = -nc \ln a + c \sum_{i=1}^n \ln x_i$$

$$\frac{d^2 \ln Lf}{dl dc} = -nb \ln a + b \sum_{i=1}^n \ln x_i$$

$$\frac{d^2 \ln Lf}{db dc} = -nl \ln a + l \sum_{i=1}^n \ln x_i$$

$$[f_n] = \begin{bmatrix} \frac{d \ln Lf}{da} \\ \frac{d \ln Lf}{dl} \\ \frac{db}{d \ln Lf} \\ \frac{dc}{d \ln Lf} \end{bmatrix}, \quad [X_n] = \begin{bmatrix} a \\ l \\ b \\ c \end{bmatrix}$$

$$J_n = \begin{bmatrix} \frac{d^2 \ln Lf}{d^2 a} & \frac{d^2 \ln Lf}{da dl} & \frac{d^2 \ln Lf}{da db} & \frac{d^2 \ln Lf}{da dc} \\ \frac{d^2 \ln Lf}{da dl} & \frac{d^2 \ln Lf}{d^2 l} & \frac{d^2 \ln Lf}{dl db} & \frac{d^2 \ln Lf}{dl dc} \\ \frac{d^2 \ln Lf}{da db} & \frac{d^2 \ln Lf}{dl db} & \frac{d^2 \ln Lf}{d^2 b} & \frac{d^2 \ln Lf}{db dc} \\ \frac{d^2 \ln Lf}{da dc} & \frac{d^2 \ln Lf}{dl dc} & \frac{d^2 \ln Lf}{db dc} & \frac{d^2 \ln Lf}{d^2 c} \end{bmatrix}$$

3.2. Least Square Developed (LSD)

It is a method used to estimate the parameters of the probabilistic models by converting these models into the sum of squares of deviations formula (i.e., the formula adopted in the ordinary least squares method) (Aboudi and Nuimai, 2016) where we assume that x_1, x_2, \dots, x_n represents a random sample with a specific distribution $F(x_i)$ and x_i represents the order statistics for the sample (Ashour and Eltehiwy, 2015).

$$E[F(x_i)] = E(P(X \leq x_i)) = \frac{i}{n + 1} \tag{13}$$

Using prediction on least squares (OLS) estimators, as follows:

$$\sum_{i=1}^n e_i^2 = \sum_{i=1}^n \left[F(x) - \frac{i}{n + 1} \right]^2 \tag{14}$$

Substituting the formula (4) for the distribution (EPPF) into formula (14) we get

$$\sum_{i=1}^n e_i^2 = \sum_{i=1}^n \left[\frac{x^{lbc}}{a^{lbc}} - \frac{i}{n + 1} \right]^2 \tag{15}$$

And by deriving formula (15) to find estimators of parameters for the new distribution (EPPF), these are:

$$\frac{d \sum_{i=1}^n e_i^2}{da} = 2 \sum_{i=1}^n \left[\frac{x^{lbc}}{a^{lbc}} - \frac{i}{n + 1} \right] \left[\frac{-lbc x^{lbc}}{a^{lbc+1}} \right]$$

$$\frac{d^2 \sum_{i=1}^n e_i^2}{da^2} = 2 \sum_{i=1}^n \left[\left(\frac{x^{lbc}}{a^{lbc}} - \frac{i}{n + 1} \right) \left(\frac{lbc x^{lbc} (lbc + 1)}{a^{lbc+2}} \right) + \left(\frac{l^2 b^2 c^2 x^{2lbc}}{a^{2lbc+2}} \right) \right]$$

$$\begin{aligned} \frac{d \sum_{i=1}^n e_i^2}{dl} &= 2 \sum_{i=1}^n \left[\frac{x^{lbc}}{a^{lbc}} - \frac{i}{n+1} \right] \left[\frac{bcx^{lbc}(\ln x - \ln a)}{a^{lbc}} \right] \\ \frac{d^2 \sum_{i=1}^n e_i^2}{dl^2} &= 2 \sum_{i=1}^n \left[\frac{b^2 c^2 x^{lbc} (\ln x - \ln a)^2 \left(\frac{2x^{lbc}}{a^{lbc}} - \frac{i}{n+1} \right)}{a^{lbc}} \right] \\ \frac{d \sum_{i=1}^n e_i^2}{db} &= 2 \sum_{i=1}^n \left[\frac{x^{lbc}}{a^{lbc}} - \frac{i}{n+1} \right] \left[\frac{lcx^{lbc}(\ln x - \ln a)}{a^{lbc}} \right] \\ \frac{d^2 \sum_{i=1}^n e_i^2}{db^2} &= 2 \sum_{i=1}^n \left[\frac{l^2 c^2 x^{lbc} (\ln x - \ln a)^2 \left(\frac{2x^{lbc}}{a^{lbc}} - \frac{i}{n+1} \right)}{a^{lbc}} \right] \\ \frac{d \sum_{i=1}^n e_i^2}{dc} &= 2 \sum_{i=1}^n \left[\frac{x^{lbc}}{a^{lbc}} - \frac{i}{n+1} \right] \left[\frac{lbx^{lbc}(\ln x - \ln a)}{a^{lbc}} \right] \\ \frac{d^2 \sum_{i=1}^n e_i^2}{dc^2} &= 2 \sum_{i=1}^n \left[\frac{l^2 b^2 x^{lbc} (\ln x - \ln a)^2 \left(\frac{2x^{lbc}}{a^{lbc}} - \frac{i}{n+1} \right)}{a^{lbc}} \right] \\ \frac{d^2 \sum_{i=1}^n e_i^2}{da dl} &= 2 \sum_{i=1}^n \left[\left(\frac{x^{lbc}}{a^{lbc}} - \frac{i}{n+1} \right) \left(\frac{-bcx^{lbc} - lb^2 c^2 (\ln x - \ln a)}{a^{lbc+1}} \right) \right. \\ &\quad \left. + \frac{lb^2 c^2 x^{2lbc} (\ln x - \ln a)}{a^{2lbc+1}} \right] \\ \frac{d^2 \sum_{i=1}^n e_i^2}{da db} &= 2 \sum_{i=1}^n \left[\left(\frac{x^{lbc}}{a^{lbc}} - \frac{i}{n+1} \right) \left(\frac{-lcx^{lbc} - bl^2 c^2 (\ln x - \ln a)}{a^{lbc+1}} \right) \right. \\ &\quad \left. + \frac{bl^2 c^2 x^{2lbc} (\ln x - \ln a)}{a^{2lbc+1}} \right] \\ \frac{d^2 \sum_{i=1}^n e_i^2}{da dc} &= 2 \sum_{i=1}^n \left[\left(\frac{x^{lbc}}{a^{lbc}} - \frac{i}{n+1} \right) \left(\frac{-lcx^{lbc} - cl^2 b^2 (\ln x - \ln a)}{a^{lbc+1}} \right) \right. \\ &\quad \left. + \frac{cl^2 b^2 x^{2lbc} (\ln x - \ln a)}{a^{2lbc+1}} \right] \\ \frac{d^2 \sum_{i=1}^n e_i^2}{dl db} &= 2 \sum_{i=1}^n \left[\left(\frac{x^{lbc}}{a^{lbc}} - \frac{i}{n+1} \right) \left(\frac{cx^{lbc}(\ln x - \ln a) + lbc^2 x^{lbc} (\ln x - \ln a)^2}{a^{lbc}} \right) \right. \\ &\quad \left. + \frac{lbc^2 x^{2lbc} (\ln x - \ln a)^2}{a^{2lbc}} \right] \\ \frac{d^2 \sum_{i=1}^n e_i^2}{dl dc} &= 2 \sum_{i=1}^n \left[\left(\frac{x^{lbc}}{a^{lbc}} - \frac{i}{n+1} \right) \left(\frac{bx^{lbc}(\ln x - \ln a) + lcb^2 x^{lbc} (\ln x - \ln a)^2}{a^{lbc}} \right) \right. \\ &\quad \left. + \frac{lcb^2 x^{2lbc} (\ln x - \ln a)^2}{a^{2lbc}} \right] \end{aligned}$$

$$\frac{d^2 \sum_{i=1}^n e_i^2}{db \, dc} = 2 \sum_{i=1}^n \left[\left(\frac{x^{lbc}}{a^{lbc}} - \frac{i}{n+1} \right) \left(\frac{l x^{lbc} (\ln x - \ln a) + bc \omega^2 x^{lbc} (\ln x - \ln a)^2}{a^{lbc}} \right) + \frac{bc l^2 x^{2lbc} (\ln x - \ln a)^2}{a^{2lbc}} \right]$$

The above equations represent a system of non-linear equations that can only be solved by using one of the numerical methods to obtain estimations (LSD), so we will use the Newton-Raphson iterative algorithm to estimate the required parameters, and the same matrices in a maximum likelihood method, we also use the numerical Nelder-Mead algorithm to improve the values Estimated parameters of the Newton-Raphson algorithm method.

4. Numerical Algorithms

4.1. Newton-Raphson's Algorithm (NR)

It is an algorithm used to get the best approximation of the zeros or roots of the function of the real values function, that is, it is used to solve linear equations and find a solution to the roots of non-linear equations that are complex and cannot be solved algebraically or linearly.

The main objective of the (NR) algorithm is to get the best estimate of the parameters of the (EPPF) distribution based on initial estimates and in an iterative manner. The matrix of equations of maximum likelihood (ML) and developed least squares (LSD) will be used, which are nonlinear equations that are difficult to solve by traditional methods (Al-Khafaji and Saleh, 2021).

The steps for calculating Newton-Raphson's algorithm are as follows:

- Determining the initial values of the parameters (a_0, l_0, b_0, c_0) by assuming them.
- Substitute the initial values of the vector $[f_n]$ and the matrix $[J_n]$.
- Calculate the estimates of the distribution parameters (EPPF) by repeating them for $(n+1)$ as shown below.

$$\begin{bmatrix} a_{(n+1)} \\ l_{(n+1)} \\ b_{(n+1)} \\ c_{(n+1)} \end{bmatrix} = \begin{bmatrix} a_{(n)} \\ l_{(n)} \\ b_{(n)} \\ c_{(n)} \end{bmatrix} - [J_n]^{-1} [f_n]$$

- The process is repeated, and in each iteration, the absolute difference between the estimates of the new parameters and the estimates of the previous parameters is calculated. If the absolute difference is very small, it is stopped (when applying the conditions, the algorithm stops).

$$|a_{(n+1)} - a_{(n)}| < e$$

$$|l_{(n+1)} - l_{(n)}| < e$$

$$|b_{(n+1)} - b_{(n)}| < e$$

$$|c_{(n+1)} - c_{(n)}| < e$$

Where (e) the value of a constant is very small.

4.2. **Nelder Mead Algorithm (NM)**

The Nelder Mead algorithm is classified in the general category of direct search methods. It has been used in many fields for numerical optimization. It is good at finding the optimal solution and very quickly it is characterized by ease of programming and the speed of its technique, one of its main advantages is to use only the function without the need for derivatives of the function that you are only guessing the number of points for each variable in the function and is considered more efficient than traditional alternative methods. This algorithm adopts the concept of (Simplex), which is optional or random values for the parameters of the distribution or function under study and has a geometric shape for n dimensions and (n + 1) points or values, and there are several geometric shapes, which depend largely on the initial assumption.

The mechanism of its work was done by arranging the points in ascending order ($\gamma_1, \gamma_2, \dots, \gamma_{n+1}$) where $\gamma = (a, l, b, c)$, and each point is compensated by the objective function under stud $[f(\gamma_1) < f(\gamma_2) < \dots < f(\gamma_{n+1})]$, where γ_1 denotes the value of the best point (the best value of the parameter γ) and $f(\gamma_1)$ represents the best value of the objective function, while γ_{n+1} is the value of the lowest point and $f(\gamma_{n+1})$ is the lowest value of the function The goal, and these points are tested by a process of continuous improvement in iterative by updating the worst point through the processes (reflection (r), expansion (e), contraction (c), and shrinkage (sh)) that is, replacing the worst point with a new, better one, and the repetition continues until it reaches the optimal minimum of the objective function, which is the optimal solution (Yalçınkaya and Yolcu, 2018).

The steps of calculating the Nelder Mead algorithm for the distribution of the Exponentiated Expanded Power Function (EETF) are as follows:

1. Determine the objective function that includes parameters (a, l, b, c), the goal of this algorithm is to make the objective function as little as possible (that is, it reaches the minimum optimum) here the objective function is the maximum likelihood function and the developed least squares function.

$$f(\gamma) = -\ln L(\gamma) \quad \gamma = (\alpha, \lambda, \beta, \theta)$$

$$\sum_{i=1}^n e_i^2 = \sum_{i=1}^n \left[\frac{x^{\lambda\beta\theta}}{\alpha^{\lambda\beta\theta}} - \frac{i}{n+1} \right]^2$$

2. Determining the values of the four parameters of the algorithm (NM) that are the reflection parameter a, the expansion parameter b, the contraction parameter c, and the contraction parameter d. In most of the research that dealt with this algorithm, the parameter values were (a=1,b=2,c=d=0.5).

3. Generating the initial solution matrix, whose dimensions are (5 * 4), which are test points for each parameter:

$$W = \begin{bmatrix} a_1 & l_1 & b_1 & c_1 \\ a_2 & l_2 & b_2 & c_2 \\ a_3 & l_3 & b_3 & c_3 \\ a_4 & l_4 & b_4 & c_4 \\ a_5 & l_5 & b_5 & c_5 \end{bmatrix}$$

The number of columns is equal to (4) as much as the number of parameters of the EEPF distribution, and the number of rows is equal to (5) from (the number of parameters +1). The four distribution parameters for each row can be expressed in γ , and the matrix becomes after modification as follows:

$$W = \begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_3 \\ \gamma_4 \\ \gamma_5 \end{bmatrix} ; \gamma = (\alpha, \lambda, \beta, \theta)$$

4. Calculation of the objective function $f(\gamma)$ for each point

$f(\gamma_1), f(\gamma_2), f(\gamma_3), f(\gamma_4), f(\gamma_5)$

5. Arranging the values of the results of the objective function (step 4) in ascending order from lowest value to highest value

$f(\gamma_1) < f(\gamma_2) < f(\gamma_3) < f(\gamma_4) < f(\gamma_5)$

6. Calculate the mean of the matrix (W), which represents the midpoint

$$\bar{\gamma} = \frac{\sum_{i=1}^m W_i}{m}$$

7. Reflection Point γ_r is calculated as follows:

$$\gamma_r = \bar{\gamma} + a(\bar{\gamma} - \gamma_5)$$

After that, the objective function is calculated for the point γ_r . If the value of the resulting objective function lies between $f(\gamma_1) \leq f(\gamma_r) \leq f(\gamma_5)$, we replace the worst point, which is (γ_5) with the point (γ_r), that is ($\gamma_r = \gamma_5$) and then move to Step 5. Otherwise, we complete the solution and move to step 8.

8. Calculation of the expansion point (Expansion point γ_e) and it is calculated according to the following law:

$$\gamma_e = \bar{\gamma} + b(\gamma_r - \bar{\gamma})$$

If the expansion point is better than the inflection point, i.e. $f(\gamma_e) < f(\gamma_r)$, then we replace (γ_5) with the point (γ_e) i.e. ($\gamma_e = \gamma_5$) and move to step 5. Otherwise, we complete the solution and move on to step 9.

9. We calculate the contraction Point (γ_c) according to the following formula:

$$\gamma_c = \bar{\gamma} + c(\gamma_5 - \bar{\gamma})$$

If the point of contraction is better than (γ_5), then we replace (γ_5) with the point (γ_c), meaning that $f(\gamma_c) < f(\gamma_5)$, then ($\gamma_c = \gamma_5$) and move to step 5. Otherwise, we complete the solution and move on to step 10.

10. We calculate the Shrink point γ_{sh} according to the following law:

$$\gamma_{sh} = \gamma_1 + d(\gamma_i - \gamma_1) \quad i = 1, 2, 3, 4, 5$$

The filter that represents the best solution is reduced, since $f(\gamma_{sh}) < f(\gamma_5)$, and (γ_{sh}) is (γ_1) and replace all points with all points except the best, which is (γ_1) and go to step 5.

11. If the condition to stop repetition is met, the optimal solution will be stopped and the optimal solution will be printed. This condition is met when the lowest value of the objective function is obtained, that is:

$$\left| \frac{\text{Max}(f) - \text{Min}(f)}{\text{Max}(f)} \right| < \epsilon$$

5. Discussion of Results

5.1. simulation

The simulation experiments included several stages that were applied to obtain the parameters and estimate the survival function and the failure or hazard rate of the new distribution of the extended exponentiated power function (EPPF). The Monte Carlo simulation method was used, which is considered the most common and most widely used in the research and analysis of parameter estimates in several ways for the model under study and comparison of the preference of the methods used using many samples of different sizes. The simulation process is characterized by flexibility as it gives the ability to experiment and test by repeating the process several times by interpreting the input to the estimation process each time. This method consists of the following steps:

1. Determining the initial values: Initial values were used for the four parameters, and three sample sizes were taken ($n=15,50,100$).

2. Data generation:

- The random variable x was generated using the inverse transformation of the cumulative function $F(x)$ for the new Expanded Exponentiated Power Function (EPPF) distribution.

- Estimating the parameters of the model for the new distribution by using the Maximum Likelihood Method (ML) and the developed least squares (LSD) using the Newton-Rapson numerical algorithm (NR) and an improvement on the Newton-Rapson estimators (NR) by the numerical algorithm (NM) Nelder Mead.

- Estimating the survival and failure rate of the new distribution of the expanded exponentiated power function (EPPF) using estimates (ML & LSD) by the two algorithms (NR,NM), and the mean integral error squares (IMSE) criterion is used, which is the integration of the total area of x_i and its reduction by one value that represents a year of the total time is calculated according to the following formula: (Hassan al-aameri & Doori,2021)

$$\text{IMSE}(\hat{S}(x)) = \frac{1}{r} \sum_{i=1}^r \left[\frac{1}{n_x} \sum_{j=1}^{n_x} (\hat{S}(x) - s(x_j))^2 \right]$$

$$\text{IMSE}(\hat{h}(x)) = \frac{1}{r} \sum_{i=1}^r \left[\frac{1}{n_x} \sum_{j=1}^{n_x} (\hat{h}(x) - h(x_j))^2 \right]$$

Whereas

r : The number of iterations of the experiment (1000) times.

n_x :The number of data generated for each sample.

$\hat{S}(x), \hat{h}(x)$: Estimated survival and failure functions, respectively.

$s(x), h(x_j)$: Survival and failure function according to the initial values, respectively.

Table 2:The simulation results for the estimated parameters, survival and failure rate and mean integral error of the distribution represent the exponentiated expanded power function (EPPF) of the elementary model (I)

N	Method		a = 10	b = 2.5	l = 3	c = 7.5	Survival rate	IMSE(SF)	Failure Rate	IMSE(HF)
			\hat{a}	\hat{b}	\hat{l}	\hat{c}				
15	ML	NR	9.5040	0.0006	3.6560	7.3862	0.9999	1.2408E-05	1.5662E-16	0.0005
		NM	9.6	3.4008	4.5115	7.2045	0.4961	3.7347E-07	1.1352	2.0942E-07
	LSD	NR	9.7126	1.7999	1.9999	5.9999	0.4943	3.4301E-08	1.4203	5.6212E-08
		NM	9.7126	1.8	2	6	0.49028	5.766E-08	1.4443	1.8054E-07
50	ML	NR	9.5040	0.0004	3.6544	7.3862E-13	0.9999	2.5954E-07	1.0944E-16	0.0006
		NM	9.6	3.4008	4.5115	7.2045	0.4995	4.0569E-09	0.9042	4.3729E-09
	LSD	NR	9.7516	1.9998	2.1998	6.2998	0.47759	8.3679e-09	1.12	1.9088e-08
		NM	9.7516	2	2.2	6.3	0.49084	1.7557e-09	1.0969	1.6083e-08
100	ML	NR	9.5040	0.0004	3.6544	7.3862E-13	0.9999	4.3369E-07	1.1678E-16	0.0017
		NM	9.6001	3.4008	4.5115	7.2045	0.4825	5.0009E-11	1.2417	1.4986E-09
	LSD	NR	9.38	2.3999	2.7999	6.4999	0.46186	1.0241E-09	1.4146	2.8658E-08
		NM	9.38	2.4	2.8	6.5	0.4538	2.3959E-10	1.5905	1.398E-07

Table 3:The simulation results for the estimated parameters, survival and failure rate and mean integral error of the distribution represent the exponentiated expanded power function (EPPF) of the elementary model (II)

N	Method		a = 10	b = 2.5	l = 4	c = 6	Survival rate	IMSE(SF)	Failure Rate	IMSE(HF)
			\hat{a}	\hat{b}	\hat{l}	\hat{c}				
15	ML	NR	9.0090	0.0001	2.5467	4.6287	0.9986	9.4499E-07	0.0002	0.0005
		NM	9.2031	2.0102	3.5086	5.0166	0.5614	1.9840E-09	0.8578	2.1351E-09
	LSD	NR	9.8215	1.8999	3.3999	5.1999	0.35637	2.4263E-07	3.1885	9.5464E-06
		NM	9.8215	1.9	3.4	5.2	0.35084	1.1611E-07	3.4883	5.8763E-06
50	ML	NR	9.0090	0.0006	2.5677	4.6413	0.9799	1.7224E-06	0.0008	0.0002
		NM	9.2031	2.0102	3.5086	5.0166	0.4457	1.2531E-09	1.5826	1.1938E-08
	LSD	NR	9.8197	2.1999	3.5999	5.4999	0.45643	2.6166E-10	1.7573	8.5503E-11
		NM	9.8197	2.2	3.6	5.5	0.44365	2.1867E-09	1.9243	1.7634E-07
100	ML	NR	9.0090	2.4339E-06	2.5463	4.6190	0.9998	3.3964E-07	2.9057E-06	0.0009
		NM	9.2031	2.0102	3.5086	5.0166	0.5035	5.6212E-11	0.8741	1.0706E-08
	LSD	NR	9.8312	2.2999	3.799	5.2999	0.45622	1.8026E-10	1.696	5.3338E-09
		NM	9.812	2.3	3.8	5.3	0.44924	3.4808E-11	1.5944	1.5245E-08

Table 4:The simulation results for the estimated parameters, survival and failure rate and mean integral error of the distribution represent the exponentiated expanded power function (EPPF) of the elementary model (III)

N	Method		a = 10	b = 3.5	l = 5	c = 7.5	Survival rate	IMSE(SF)	Failure Rate	IMSE(HF)
			\hat{a}	\hat{b}	\hat{l}	\hat{c}				
15	ML	NR	9.5049	0.0006	3.656	7.3862E-13	0.9999	1.2408E-05	1.566E-16	0.0005
		NM	9.6	3.4008	4.5115	7.2045	0.4961	3.7347E-07	1.1352	2.0942E-07
	LSD	NR	9.9065	2.4999	4.4999	5.9999	0.4759	2.5678E-07	1.0491	2.1014E-06
		NM	9.9065	2.5	4.5	6	0.46871	8.6993E-07	1.0723	8.8461E-07
50	ML	NR	9.5040	0.0004	3.6544	7.3862E-13	0.9999	2.5954E-07	1.0944E-16	0.0006
		NM	9.6	3.4008	4.5115	7.2044	0.4995	4.0569E-09	0.9042	4.3729E-09
	LSD	NR	9.9089	2.7999	4.5999	6.2999	0.39836	1.903E-10	2.2301	7.1398E-07
		NM	9.9089	2.8	4.6	6.3	0.4177	1.5672E-08	2.6949	4.9121E-09
100	ML	NR	9.5049	0.0004	3.6530	7.3862E-13	0.9999	4.3369E-07	1.1678E-16	0.0017
		NM	9.6001	3.4008	4.5115	7.2045	0.4825	5.0010E-11	1.2417	1.4986E-09
	LSD	NR	9.9206	2.9999	4.7999	6.8999	0.49519	1.5247E-09	1.0306	1.539E-09
		NM	9.9206	3	4.8	6.9	0.48681	3.2972E-11	1.0717	2.0263E-08

Table 5:The simulation results for the estimated parameters, survival and failure rate and mean integral error of the distribution represent the exponentiated expanded power function (EPPF) of the elementary model (IV)

N	Method		a = 11	b = 3	l = 3	c = 5.5	Survival rate	IMSE(SF)	Failure Rate	IMSE(HF)
			\hat{a}	\hat{b}	\hat{l}	\hat{c}				
15	ML	NR	7	1.5057	4.7802E-05	3.9944	0.9998	4.0672E-05	2.6741E-05	0.0001
		NM	10.2	2.318	2.547	5.0633	0.5146	8.6574E-07	0.95689	3.9153E-06
	LSD	NR	10.644	2.4	2.2	4.5	0.43788	4.3413E-09	1.6211	1.6663E-05
		NM	10.644	2.4	2.2	4.5	0.43665	6.1711E-11	1.6597	1.5811E-05
50	ML	NR	7	1.7126	0.0009	4.5735	0.9825	1.6124E-07	0.0007	0.0018
		NM	10.2	2.318	2.547	5.0633	0.4861	1.6699E-09	1.2785	3.3935E-09
	LSD	NR	10.726	2.2	2.6	5.1	0.43009	4.4965E-09	2.2334	1.9352E-07
		NM	10.726	2.2	2.6	5.1	0.42635	1.3497E-10	4.172	9.0505E-08
100	ML	NR	7	0.0009	2.9534	3.9807	0.9460	8.6831E-08	0.0011	0.0001
		NM	10.2	2.318	2.8151	5.0633	0.4515	1.2762E-11	1.4464	1.6276E-09
	LSD	NR	10.802	2.7999	2.8999	5.4999	0.4834	3.4105E-11	1.0545	6.9787E-08
		NM	10.802	2.8	2.9	5.5	0.4989	1.1555E-11	0.9957	7.5006E-11

Table 6:The simulation results for the estimated parameters, survival and failure rate and mean integral error of the distribution represent the exponentiated expanded power function (EPPF) of the elementary model (V)

N	Method		a = 11	b = 3	l = 4	c = 5.5	Survival rate	IMSE(SF)	Failure Rate	IMSE(HF)
			\hat{a}	\hat{b}	\hat{l}	\hat{c}				
15	ML	NR	7.5	2.969	8.6797E-05	3.2419	0.9994	3.2288E-05	7.695E-05	0.0022
		NM	10.6	2.6971	3.1018	5.0636	0.5385	4.0547E-07	1.2755	2.1351E-05
	LSD	NR	10.75	2.2	3	4.5	0.4410	1.6519E-08	1.6112	1.2533E-05
		NM	10.75	2.2	3	4.5	0.4406	1.6958E-08	1.7362	1.1846E-05
50	ML	NR	7.5	2.9525	0.0001	3.2379	0.9972	1.0705E-09	0.0001	0.0019
		NM	10.6	2.6775	3.0993	5.0602	0.43214	3.5877E-09	2.4778	1.3644E-07
	LSD	NR	10.802	2.4999	3.3999	4.9999	0.4297	3.5394E-10	1.1835	3.455E-10
		NM	10.802	2.5	3.4	5	0.4303	2.4759E-09	1.2435	1.571E-07
100	ML	NR	7.5	2.9685	0.0004	3.2493	0.9811	1.3502E-08	0.0004	0.0020
		NM	10.6	2.6636	3.0993	5.0602	0.4855	8.4184E-10	1.3761	3.5852E-09
	LSD	NR	10.84	2.5999	3.8999	5.1999	0.4155	2.1836E-09	1.2635	7.7612E-07
		NM	10.84	2.9	3.9	5.2	0.4369	6.7023E-12	1.2237	1.3032E-10

Table 7:The simulation results for the estimated parameters, survival and failure rate and mean integral error of the distribution represent the exponentiated expanded power function (EPPF) of the elementary model (VI)

N	Method		a = 11	b = 3.5	l = 5	c = 7.5	Survival rate	IMSE(SF)	Failure Rate	IMSE(HF)
			\hat{a}	\hat{b}	\hat{l}	\hat{c}				
15	ML	NR	7	0.0002	3.6108	7.0576E-13	0.9999	6.0436E-05	3.8624E-17	0.0012
		NM	10.2	3.0737	4.547	7.1506	0.50616	2.7044E-09	1.1836	6.2806E-09
	LSD	NR	10.895	2.7999	4.1999	6.4999	0.3630	1.6516E-08	2.2464	7.4826E-06
		NM	10.895	2.8	4.2	6.5	0.35435	2.0106E-07	2.3985	4.7242E-06
50	ML	NR	7.5	0.0006	3.5701	6.7201E-13	0.9999	8.8335E-10	1.3328E-16	0.0008
		NM	10.6	3.2383	4.5496	7.0602	0.44852	2.4944E-09	1.3474	2.2119E-08
	LSD	NR	10.916	2.9999	4.4999	6.9999	0.49876	2.7058E-10	0.87754	9.6111E-09
		NM	10.916	3	4.5	7	0.50723	1.6174E-08	0.74527	1.7905E-13
100	ML	NR	7.5	0.0007	3.5668	6.7201E-13	0.9999	2.3407E-08	1.5148E-16	0.0015698
		NM	10.6	3.2383	4.5496	7.0602	0.46125	9.6081E-10	1.4523	2.2316E-08
	LSD	NR	10.927	3.3999	4.7999	7.3999	0.43144	3.0764E-11	1.3386	9.1151E-09
		NM	10.927	3.4	4.8	7.4	0.44661	8.5684E-11	1.2708	4.39E-09

Tables (2), (3), (4), (5), (6), and (7) indicate the performance of the method maximum likelihood (ML) and the developed least squares LSD using the two algorithms (NR, NM) are very low for all sizes and this is because the value of the estimators of one of the shape parameters has a value less than the initial value, so we find that the performance of the NM algorithm has improved from these estimations. As for the LSD method, we find its results are close in all algorithms, and that these two algorithms at those sizes have the best The mean of the integral error squares relative to the survival rate and failure rate, and this indicates that the estimations of the parameters are good and that the (NM) algorithm for the two methods was efficient.

6. Conclusion

In this paper, a new distribution was constructed that is the expanded exponentiated power function (EPPF) distribution and the properties of this distribution were found and its coefficients were estimated using the maximum likelihood (ML) method, least squares developed (LSD), and it was concluded from the experimental results that the proposed new distribution of the expanded exponentiated power function is effective in representing the experimental data, the survival rates were at the appropriate levels, the failure rate was low to almost stable, and the Nelder Mead (NM) algorithm worked efficiently in improving the estimates of the algorithm Newton Raphson (NR), whose estimates were close to the initial values; we also noted the efficiency of the new distribution in representing small sizes, the survival rate of the two algorithms was good as well as the failure rate, and the mean integrated error criterion (IMSE) was low for both survival rate and failure rate, and this refers to the efficiency of the expanded exponentiated power function distribution estimates using maximum likelihood and developed least squares method estimated using two algorithms (NR, NM), and this demonstrates the efficiency and flexibility of the new distribution (EPPF) in data modelling.

References

1. Abody, E.H. and Nuimai, A.B., 2016. Comparison Between Two Approaches (MLE & DLS) to Estimate Frechet Poisson Lindley Distribution Compound by Using Simulation. *Ibn AL-Haitham Journal For Pure and Applied Sciences*, 29(3), pp.401-414.
2. Afify, A.Z., Zayed, M. and Ahsanullah, M., 2018. The Extended Exponential Distribution and Its Applications. *J. Stat. Theory Appl.*, 17(2), pp.213-229.
3. Ahmed, Z., Nofal, Z.M. and Abd El Hadi, N.E., 2015. Exponentiated transmuted generalized Raleigh distribution: A new four parameter Rayleigh distribution. *Pakistan journal of statistics and operation research*, pp.115-134.
4. Aldahlan, M.A., Jamal, F., Chesneau, C., Elbatal, I. and Elgarhy, M., 2020. Exponentiated power generalized Weibull power series family of distributions: Properties, estimation and applications. *PloS one*, 15(3), p.e0230004.
5. Alkhafaji, M.A. and Saleh, R.A., 2021, May. A Statistical Study on the Parameters of the Skew Normal Distribution Depending on the Use of the Genetic Algorithm Using the Simulation Method. In *Journal of Physics: Conference Series* (Vol. 1879, No. 3, p. 032017). IOP Publishing.
6. Ashour, S.K. and Eltehiwy, M.A., 2015. Exponentiated power Lindley distribution. *Journal of advanced research*, 6(6), pp.895-905.
7. Cordeiro, G.M., Ortega, E.M. and da Cunha, D.C., 2013. The exponentiated generalized class of distributions. *Journal of data science*, 11(1), pp.1-27.
8. De Andrade, T.A., Chakraborty, S., Handique, L. and Gomes-Silva, F., 2019. The exponentiated generalized extended Gompertz distribution. *Journal of Data Science*, 17(2), pp.299-330.
9. Gupta, R.D. and Kundu, D., 2001. Exponentiated exponential family: an alternative to gamma and Weibull distributions. *Biometrical Journal: Journal of Mathematical Methods in Biosciences*, 43(1), pp.117-130.

10. Hassan al-aameri, S.S. and Doori, E.A.F.A., 2021. Comparing Some Methods of Estimating the Parameters and Survival Function of a Log-logistic Distribution with a Practical Application.
11. Yalçinkaya, A., Şenoğlu, B. and Yolcu, U., 2018. Maximum likelihood estimation for the parameters of skew normal distribution using genetic algorithm. *Swarm and Evolutionary Computation*, 38, pp.127-138.
12. Peña-Ramírez, F.A., Guerra, R.R., Cordeiro, G.M. and Marinho, P.R., 2018. The exponentiated power generalized Weibull: Properties and applications. *Anais da Academia Brasileira de Ciências*, 90, pp.2553-2577.
13. Rather, A.A. and Subramanian, C., 2018. Exponentiated Mukherjee-Islam Distribution. *Journal of Statistics Applications & Probability*, 7(2), pp.357-361.
14. Zamani, Z., Afshari, M., Karamikabir, H., Alizadeh, M. and Ali, M.M., 2022. Extended Exponentiated Chen Distribution: Mathematical Properties and Applications. *Statistics, Optimization & Information Computing*, 10(2), pp.606-626.

مقارنة بين بعض طرائق تقدير دالة البقاء ومعدل الفشل لتوزيع دالة القوة الموسع الاسي

أ.م.د. عماد حازم عبودي⁽²⁾الباحث/ ميسم ساجت خضير⁽¹⁾

جامعة بغداد، كلية الإدارة والاقتصاد، بغداد، العراق

جامعة بغداد، كلية الإدارة والاقتصاد، بغداد، العراق

emadhazim@coadec.uobaghdad.edu.iqMaysam.sajit1201a@coadec.uobaghdad.edu.iq

Received: 10/8/2022

Accepted: 4/9/2022

Published: September / 2022

هذا العمل مرخص تحت اتفاقية المشاع الإبداعي نَسب المُصنَّف - غير تجاري - الترخيص العمومي الدولي 4.0

[Attribution-NonCommercial 4.0 International \(CC BY-NC 4.0\)](https://creativecommons.org/licenses/by-nc-sa/4.0/)

مستخلص البحث

لقد قدمنا توزيع دالة القوة الموسع الاسي (EPPF) بأربعة معلمات ، حيث تم انشاء هذا التوزيع بواسطة طريقة التوسعة الاسية التي اوجدها العالم جوبتا للتوسعة التوزيع الاسي وذلك بإضافة معلمة شكل جديدة لدالة التراكمية للتوزيع فينتج عنها توزيع جديد ، وتتميز هذه الطريقة في الحصول على توزيع ينتمي للعائلة الاسية ، كما حصلنا على دالة معدل البقاء ومعدل الفشل لهذا التوزيع ، حيث تم اشتقاق بعض الخصائص الرياضية ، ثم استعملنا طريقة الإمكان الاعظم (ML) وطريقة المربعات الصغرى المطورة (LSD) لتقدير المعلمات وبسبب العلاقة الملاحظة بين المعلمات تم استعمال الخوارزميات العددية لإيجاد مقدرات الطريقتين وهي خوارزميتي نيوتن رافسون (NR) و Nelder mead (NM) لتحسين المقدرات، وأجريت تجربة محاكاة مونت كارلو للتقييم أداء تقديرات الخوارزميتين ، وتم استعمال متوسط معيار الخطأ المتكامل (IMSE) لمقارنة مقدرات دالة البقاء ومعدل الفشل، وقد اظهرت النتائج كفاءة تقديرات طريقة الإمكان الاعظم والمربعات الصغرى المطورة باستعمال الخوارزميتين (NR ، NM) حيث كانت نتائجهم متقاربة ، وهذا يوضح كفاءة التوزيع الجديدة (EPPF) للنمذجة بيانات البقاء.

نوع البحث: ورقة بحثية

المصطلحات الرئيسية للبحث: توزيع دالة القوة الموسع الاسي ، البقاء ومعدل الفشل ، الإمكان الأعظم ، المربعات الصغرى المطورة ، نيوتن رافسون ، Nelder maed.

(1) بحث مستل من رسالة ماجستير