IRAQI STATISTICIANS JOURNAL

https://isj.edu.iq/index.php/isj

ISSN: 3007-1658 (Online

Estimating Reliability Functions of a Weighted Pareto-Poisson Distribution with the Application

Inam A.R. Noaman[1], Duha K. Mahmoud [2]

[1] Department of statistics, College Administration and Economics - University of Diyala
[2] Mathematics department, College of Science - University of Diyala

ARTICLE INFO

Article history:

Received 20/11/2024
Revised 20/11/2024
Accepted 14/1/2025
Available online 15/5/2025

Keywords:

Weighted Pareto – Poisson distribution
reliability
Percentiles Estimators Method
Downhill simplex algorithm
Simulation

ABSTRACT

This research is concerned with studying the estimation of the parameters and reliability function of a weighted probability distribution consisting of a combination of the Pareto distribution and the Poisson distribution. It is called the weighted Pareto-Poisson distribution, which is a new probability distribution for life time.

The weighted distribution Pareto-Poisson is distribution with four parameters α, θ, λ, γ. In the research, the parameters and distribution reliability function were estimated in this research using two methods, namely Percentiles Estimators Method and Estimation with Downhill simplex algorithm. The mean square error measurement was used to compare the results of the estimates obtained from the two methods, in addition to give three sets of initial default values for the distribution parameters α, θ, λ, γ and three sizes for the samples used in the estimation 30, 60, 90. The sizes are considered small, medium, and large. The process was repeated. 1000 times. In addition, five failure times were taken to estimate the reliability function. The Python programming language was used to simulate this data distribution. As for the applied aspect, the life times of thalassemia patients were studied for a sample size of 60 people in the Specialized Center for Hematology in the Diyala Health Department, where the Downhill simplex method was used to estimate the parameters and the reliability function, as it was the best in simulation

1. Introduction

The Pareto-Poisson distribution, which was presented by Ahmed El-Shahhat in 2022, has wide importance in medical, industrial, engineering, biological, and other applications for testing life time [2]. This distribution has the probability function pdf and cumulative distribution cdf are given by equations (1) and (2):

$$f(\mathrm{x},\alpha,\lambda,\theta)=\frac{\alpha\theta\lambda^{\alpha}e^{\theta}}{x^{\alpha+1}\,(e^{\theta}-1)}\qquad \alpha,\theta>0\,,x\ge\lambda \qquad (1)$$

$$F(\mathrm{x},\alpha,\theta,\lambda)=1-\frac{e^{\theta}-e^{\theta\left(\frac{\lambda^{\alpha}}{x^{\alpha}}\right)}}{e^{\theta}-1}\qquad \alpha,\theta,\lambda>0\,,x\ge\lambda\quad ,\text{x is random varaible}\qquad (2)$$

The researcher made some modification and additions to produce a new distribution called the weighted Pareto-Poisson distribution using the Azzallini’s method which is a new way to find weighted distributions [1, 6], which is as follows, as in equation (3):

$$\mathrm{f_w}(x)=\frac{1}{\mathrm{P}(\mathrm{X_1}<\gamma \mathrm{x_2})}f(\mathrm{X_1}).F(\gamma\mathrm{x_2})\qquad (3);$$

$\mathrm{x_1}$ *and* $\mathrm{x_2}$ is random varaible

* Corresponding author.E-**mail** address: inaamsta@uodiyala.edu.iq
https://doi.org/10.62933/mmjkw787

Where the probability density function of the Weighted Pareto – Poisson distribution is:

$$f(x,\alpha,\theta,\lambda,\gamma) = \frac{(1+\gamma^{-\alpha})\,\alpha\theta\lambda^{\alpha}e^{\theta\lambda^{\alpha}x^{-\alpha}}\left[e^{\theta}-e^{\theta\lambda^{\alpha}(\gamma x)^{-\alpha}}\right]}{\left[(1+\gamma^{-\alpha})e^{2\theta}-e^{\theta+\theta\gamma^{-\alpha}}\right]x^{\alpha+1}},\theta > 0, x \geq \lambda \geq 1, \alpha > 0 \quad (4)$$

As for the cumulative distribution function and the reliability function is:

$$F(x,\alpha,\theta,\lambda,\gamma) = 1-\frac{(1+\gamma^{-\alpha})\,e^{\theta+\theta\lambda^{\alpha}x^{-\alpha}}-e^{(1+\gamma^{-\alpha})\theta\lambda^{\alpha}x^{-\alpha}}}{(1+\gamma^{-\alpha})e^{2\theta}-e^{\theta+\theta\gamma^{-\alpha}}}\ ,\theta > 0\ , x \geq \lambda \geq 1\ ,\ \alpha > 0 \quad (5)$$

As for the reliability function is:

$$R(x,\alpha,\theta,\lambda,\gamma) = \frac{(1+\gamma^{-\alpha})\,e^{\theta+\theta\lambda^{\alpha}x^{-\alpha}}-e^{(1+\gamma^{-\alpha})\theta\lambda^{\alpha}x^{-\alpha}}}{(1+\gamma^{-\alpha})e^{2\theta}-e^{\theta+\theta\gamma^{-\alpha}}}\ ,\theta > 0\ , x \geq \lambda \geq 1\ ,\alpha > 0 \quad (6)$$

2. The Theoretical Side

In this part, we will stady the estimation method, which are (Percentiles Estimators Method, Estimation with the Downhill simplex algorithm) to estimate the parameters of the above-mentioned distribution, which are as follows:

2.1. Percentiles Estimators Method [5]

This method depends on estimating the cumulative distribution function $F(x.\ \alpha.\ \theta.\ \lambda.\ \gamma)$ in a non-parametric way that was proposed by Kao, assuming that Gi represents the estimate of the cumulative function of the weighted Pareto-Poisson composite distribution. This method can be summarized as follows: -

The cumulative distribution function is given by

$$F_w(x.\ \alpha.\ \theta.\ \lambda.\ \gamma) = 1-\frac{(1+\gamma^{-\alpha})\,e^{\theta+\theta\lambda^{\alpha}x^{-\alpha}}-e^{(1+\gamma^{-\alpha})\theta\lambda^{\alpha}x^{-\alpha}}}{(1+\gamma^{-\alpha})e^{2\theta}-e^{\theta+\theta\gamma^{-\alpha}}}$$

Since Gi is a non-parametric quantity, it has a number of formulas, including:

$$Gi = \frac{i+0\cdot375}{n+0\cdot25}$$

Or

$$Gi = \frac{i}{n+1}$$

After equating Gi with the cumulative function, we have

$$Gi = 1-\frac{(1+\gamma^{-\alpha})\,e^{\theta+\theta\lambda^{\alpha}x^{-\alpha}}-e^{(1+\gamma^{-\alpha})\theta\lambda^{\alpha}x^{-\alpha}}}{(1+\gamma^{-\alpha})e^{2\theta}-e^{\theta+\theta\gamma^{-\alpha}}}$$

$$(1-Gi) = \frac{(1+\gamma^{-\alpha})\,e^{\theta+\theta\lambda^{\alpha}x^{-\alpha}}-e^{(\theta\lambda^{\alpha}x^{-\alpha}+\theta\lambda^{\alpha}x^{-\alpha}\gamma^{-\alpha})}}{(1+\gamma^{-\alpha})e^{2\theta}-e^{\theta+\theta\gamma^{-\alpha}}}$$

$$(1-Gi) = \frac{e^{\theta\lambda^{\alpha}x^{-\alpha}}\left[(1+\gamma^{-\alpha})\,e^{\theta}-e^{\theta\lambda^{\alpha}x^{-\alpha}\gamma^{-\alpha}}\right]}{e^{2\theta}+e^{2\theta}\gamma^{-\alpha}-e^{\theta+\theta\gamma^{-\alpha}}}$$

Applying the natural logarithm to both sides of the above equation, we consider the following

$$\ln(1-Gi) = \ln\left(e^{\theta\lambda^{\alpha}x^{-\alpha}}\left[(1+\gamma^{-\alpha})\,e^{\theta}-e^{\theta\lambda^{\alpha}x^{-\alpha}\gamma^{-\alpha}}\right]\right)-\ln\left(e^{2\theta}+e^{2\theta}\gamma^{-\alpha}-e^{\theta+\theta\gamma^{-\alpha}}\right)$$

$$\ln(1-Gi)-\ln\left(e^{\theta\lambda^{\alpha}x^{-\alpha}}\left[(1+\gamma^{-\alpha})\,e^{\theta}-e^{\theta\lambda^{\alpha}x^{-\alpha}\gamma^{-\alpha}}\right]\right)+\ln\left(e^{2\theta}+e^{2\theta}\gamma^{-\alpha}-e^{\theta+\theta\gamma^{-\alpha}}\right) = 0$$

$$\ln(1-Gi)-\theta\lambda^{\alpha}x^{-\alpha}-\ln\left(e^{\theta}+e^{\theta}\gamma^{-\alpha}-e^{\theta\lambda^{\alpha}x^{-\alpha}\gamma^{-\alpha}}\right)+\ln\left(e^{2\theta}+e^{2\theta}\gamma^{-\alpha}-e^{\theta+\theta\gamma^{-\alpha}}\right) = 0$$

Square the above equation and apply the summation to have

$$\sum_{i=1}^{n}\left[\ln(1-Gi)-\theta\lambda^{\alpha}x^{-\alpha}-\ln\left(e^{\theta}+e^{\theta}\gamma^{-\alpha}-e^{\theta\lambda^{\alpha}x^{-\alpha}\gamma^{-\alpha}}\right)+\ln\left(e^{2\theta}+e^{2\theta}\gamma^{-\alpha}-e^{\theta+\theta\gamma^{-\alpha}}\right)\right]^{2}$$

By differentiating Equation (6) with respect to the parameters and simplifying the derivative products, we obtain:

$$\frac{\partial}{\partial\alpha}\sum_{i=1}^{n}\left[\ln(1-Gi)-\theta\lambda^{\alpha}x^{-\alpha}-\ln\left(e^{\theta}+e^{\theta}\gamma^{-\alpha}-e^{\theta\lambda^{\alpha}x^{-\alpha}\gamma^{-\alpha}}\right)+\ln\left(e^{2\theta}+e^{2\theta}\gamma^{-\alpha}-e^{\theta+\theta\gamma^{-\alpha}}\right)\right]^{2} = 0 \qquad (7).$$

$$2\sum_{i=1}^{n}\left[\ln(1-Gi)-\theta\lambda^{\alpha}x^{-\alpha}-\ln\left(e^{\theta}+e^{\theta}\gamma^{-\alpha}-e^{\theta\lambda^{\alpha}x^{-\alpha}\gamma^{-\alpha}}\right)+\ln\left(e^{2\theta}+e^{2\theta}\gamma^{-\alpha}-\right.\right.$$

$$e^{\theta+\theta\gamma^{-\alpha}})\Big] \cdot \left[\frac{\theta\lambda^{\alpha}\ln x-\theta\lambda^{\alpha}\ln\lambda}{x^{\alpha}} + \frac{\gamma^{-\alpha}e^{\theta}\ln\gamma}{e^{\theta}+e^{\theta}\gamma^{-\alpha}-e^{\theta\lambda^{\alpha}x^{-\alpha}\gamma^{-\alpha}}} + \frac{e^{\theta\lambda^{\alpha}x^{-\alpha}\gamma^{-\alpha}}(-\theta x^{-\alpha}\gamma^{-\alpha}\ln\gamma-\theta\lambda^{\alpha}x^{-\alpha}\gamma^{-\alpha}\ln x+\theta\lambda^{\alpha}x^{-\alpha}\gamma^{-\alpha}\ln\lambda)}{e^{\theta}+e^{\theta}\gamma^{-\alpha}-e^{\theta\lambda^{\alpha}x^{-\alpha}\gamma^{-\alpha}}} + \frac{\gamma^{-\alpha}\theta e^{\theta+\theta\gamma^{-\alpha}}\ln\gamma-\gamma^{-\alpha}e^{2\theta}\ln\gamma}{e^{2\theta}+e^{2\theta}\gamma^{-\alpha}-e^{\theta+\theta\gamma^{-\alpha}}}\right] = 0 \qquad (8).$$

Let

$$K_1 = \left[\frac{\theta\lambda^{\alpha}\ln x-\theta\lambda^{\alpha}\ln\lambda}{x^{\alpha}} + \frac{\gamma^{-\alpha}e^{\theta}\ln\gamma}{e^{\theta}+e^{\theta}\gamma^{-\alpha}-e^{\theta\lambda^{\alpha}x^{-\alpha}\gamma^{-\alpha}}} + \frac{e^{\theta\lambda^{\alpha}x^{-\alpha}\gamma^{-\alpha}}(-\theta x^{-\alpha}\gamma^{-\alpha}\ln\gamma-\theta\lambda^{\alpha}x^{-\alpha}\gamma^{-\alpha}\ln x+\theta\lambda^{\alpha}x^{-\alpha}\gamma^{-\alpha}\ln\lambda)}{e^{\theta}+e^{\theta}\gamma^{-\alpha}-e^{\theta\lambda^{\alpha}x^{-\alpha}\gamma^{-\alpha}}} + \frac{\gamma^{-\alpha}\theta e^{\theta+\theta\gamma^{-\alpha}}\ln\gamma-\gamma^{-\alpha}e^{2\theta}\ln\gamma}{e^{2\theta}+e^{2\theta}\gamma^{-\alpha}-e^{\theta+\theta\gamma^{-\alpha}}}\right]$$

By substituting the value of K_1 into Equation (8), we consider

$$\therefore \sum_{i=1}^{n}\left[\ln(1-\mathrm{Gi})\,K_1 - \theta\lambda^{\alpha}x^{-\alpha}K_1 - \ln\left(e^{\theta}+e^{\theta}\gamma^{-\alpha} - e^{\theta\lambda^{\alpha}x^{-\alpha}\gamma^{-\alpha}}\right)K_1 + \ln\left(e^{2\theta}+e^{2\theta}\gamma^{-\alpha} - e^{\theta+\theta\gamma^{-\alpha}}\right)K_1\right] = 0 \quad (9).$$

$$\frac{\partial}{\partial\theta}\sum_{i=1}^{n}\left[\ln(1-\mathrm{Gi}) - \theta\lambda^{\alpha}x^{-\alpha} - \ln\left(e^{\theta}+e^{\theta}\gamma^{-\alpha} - e^{\theta\lambda^{\alpha}x^{-\alpha}\gamma^{-\alpha}}\right) + \ln\left(e^{2\theta}+e^{2\theta}\gamma^{-\alpha} - e^{\theta+\theta\gamma^{-\alpha}}\right)\right]^2 = 0$$

$$\sum_{i=1}^{n}\left[\ln(1-\mathrm{Gi}) - \theta\lambda^{\alpha}x^{-\alpha} - \ln\left((1+\gamma^{-\alpha})e^{\theta} - e^{\theta\lambda^{\alpha}x^{-\alpha}\gamma^{-\alpha}}\right) + \ln\left((1+\gamma^{-\alpha})e^{2\theta} - e^{\theta+\theta\gamma^{-\alpha}}\right)\right] \cdot \left[-\lambda^{\alpha}x^{-\alpha} - \frac{(1+\gamma^{-\alpha})e^{\theta}-\lambda^{\alpha}x^{-\alpha}\gamma^{-\alpha}e^{\theta\lambda^{\alpha}x^{-\alpha}\gamma^{-\alpha}}}{(1+\gamma^{-\alpha})e^{\theta}-e^{\theta\lambda^{\alpha}x^{-\alpha}\gamma^{-\alpha}}} + \frac{2(1+\gamma^{-\alpha})e^{2\theta}-(1+\gamma^{-\alpha})e^{\theta+\theta\gamma^{-\alpha}}}{(1+\gamma^{-\alpha})e^{2\theta}-e^{\theta+\theta\gamma^{-\alpha}}}\right] = 0 \quad (10).$$

Let

$$K_2 = -\lambda^{\alpha}x^{-\alpha} - \frac{(1+\gamma^{-\alpha})e^{\theta}-\lambda^{\alpha}x^{-\alpha}\gamma^{-\alpha}e^{\theta\lambda^{\alpha}x^{-\alpha}\gamma^{-\alpha}}}{(1+\gamma^{-\alpha})e^{\theta}-e^{\theta\lambda^{\alpha}x^{-\alpha}\gamma^{-\alpha}}} + \frac{2(1+\gamma^{-\alpha})e^{2\theta}-(1+\gamma^{-\alpha})e^{\theta+\theta\gamma^{-\alpha}}}{(1+\gamma^{-\alpha})e^{2\theta}-e^{\theta+\theta\gamma^{-\alpha}}}$$

By substituting the value of K_2 into Equation (10), we have

$$\therefore \sum_{i=1}^{n}\left[\ln(1-\mathrm{Gi})\,K_2 - \theta\lambda^{\alpha}x^{-\alpha}K_2 - \ln\left((1+\gamma^{-\alpha})e^{\theta} - e^{\theta\lambda^{\alpha}x^{-\alpha}\gamma^{-\alpha}}\right)K_2 + \ln\left((1+\gamma^{-\alpha})e^{2\theta} - e^{\theta+\theta\gamma^{-\alpha}}\right)K_2 = 0\right] \quad (11).$$

$$\frac{\partial}{\partial\lambda}\sum_{i=1}^{n}\left[\ln(1-\mathrm{Gi}) - \theta\lambda^{\alpha}x^{-\alpha} - \ln\left(e^{\theta}+e^{\theta}\gamma^{-\alpha} - e^{\theta\lambda^{\alpha}x^{-\alpha}\gamma^{-\alpha}}\right) - \ln\left(e^{2\theta}+e^{2\theta}\gamma^{-\alpha} - e^{\theta+\theta\gamma^{-\alpha}}\right)\right]^2 = 0$$

$$\sum_{i=1}^{n}\left[\ln(1-\mathrm{Gi}) - \theta\lambda^{\alpha}x^{-\alpha} - \ln\left(e^{\theta}+e^{\theta}\gamma^{-\alpha} - e^{\theta\lambda^{\alpha}x^{-\alpha}\gamma^{-\alpha}}\right) + \ln\left(e^{2\theta}+e^{2\theta}\gamma^{-\alpha} - e^{\theta+\theta\gamma^{-\alpha}}\right)\right] \cdot \left[-\theta\alpha\lambda^{\alpha-1}x^{-\alpha} + \frac{\theta\alpha\lambda^{\alpha-1}x^{-\alpha}\gamma^{-\alpha}e^{\theta\lambda^{\alpha}x^{-\alpha}\gamma^{-\alpha}}}{(1+\gamma^{-\alpha})e^{\theta}-e^{\theta\lambda^{\alpha}x^{-\alpha}\gamma^{-\alpha}}}\right] = 0 \quad (12).$$

Let

$$K_3 = \frac{\theta\alpha\lambda^{\alpha-1}x^{-\alpha}\gamma^{-\alpha}e^{\theta\lambda^{\alpha}x^{-\alpha}\gamma^{-\alpha}}}{(1+\gamma^{-\alpha})e^{\theta}-e^{\theta\lambda^{\alpha}x^{-\alpha}\gamma^{-\alpha}}} - \theta\alpha\lambda^{\alpha-1}x^{-\alpha}$$

By substituting the value of K_3 into Equation (12)

$$\therefore \sum_{i=1}^{n}\left[\ln(1-\mathrm{Gi})K_3 - \theta\lambda^{\alpha}x^{-\alpha}K_3 - \ln\left(e^{\theta}+e^{\theta}\gamma^{-\alpha} - e^{\theta\lambda^{\alpha}x^{-\alpha}\gamma^{-\alpha}}\right)K_3 + \ln\left(e^{2\theta}+e^{2\theta}\gamma^{-\alpha} - e^{\theta+\theta\gamma^{-\alpha}}\right)K_3\right] = 0 \quad (13).$$

$$\frac{\partial}{\partial\gamma}\sum_{i=1}^{n}\left[\ln(1-\mathrm{Gi}) - \theta\lambda^{\alpha}x^{-\alpha} - \ln\left(e^{\theta}+e^{\theta}\gamma^{-\alpha} - e^{\theta\lambda^{\alpha}x^{-\alpha}\gamma^{-\alpha}}\right) - \ln\left(e^{2\theta}+e^{2\theta}\gamma^{-\alpha} - e^{\theta+\theta\gamma^{-\alpha}}\right)\right]^2 = 0$$

$$\sum_{i=1}^{n}\left[\ln(1-\mathrm{Gi}) - \theta\lambda^{\alpha}x^{-\alpha} - \ln\left(e^{\theta}+e^{\theta}\gamma^{-\alpha} - e^{\theta\lambda^{\alpha}x^{-\alpha}\gamma^{-\alpha}}\right) + \ln\left(e^{2\theta}+e^{2\theta}\gamma^{-\alpha} - e^{\theta+\theta\gamma^{-\alpha}}\right)\right] \cdot \left[\frac{\alpha\gamma^{-\alpha-1}e^{\theta}-\alpha\theta\lambda^{\alpha}x^{-\alpha}\gamma^{-\alpha-1}e^{\theta\lambda^{\alpha}x^{-\alpha}\gamma^{-\alpha}}}{(1+\gamma^{-\alpha})e^{\theta}-e^{\theta\lambda^{\alpha}x^{-\alpha}\gamma^{-\alpha}}} - \frac{\alpha\gamma^{-\alpha-1}e^{2\theta}-\alpha\theta\gamma^{-\alpha-1}e^{\theta+\theta\gamma^{-\alpha}}}{(1+\gamma^{-\alpha})e^{2\theta}-e^{\theta+\theta\gamma^{-\alpha}}}\right] = 0 \quad (14).$$

Let

$$K_4 = \frac{\alpha\gamma^{-\alpha-1}e^{\theta}-\alpha\theta\lambda^{\alpha}x^{-\alpha}\gamma^{-\alpha-1}e^{\theta\lambda^{\alpha}x^{-\alpha}\gamma^{-\alpha}}}{(1+\gamma^{-\alpha})e^{\theta}-e^{\theta\lambda^{\alpha}x^{-\alpha}\gamma^{-\alpha}}} - \frac{\alpha\gamma^{-\alpha-1}e^{2\theta}-\alpha\theta\gamma^{-\alpha-1}e^{\theta+\theta\gamma^{-\alpha}}}{(1+\gamma^{-\alpha})e^{2\theta}-e^{\theta+\theta\gamma^{-\alpha}}}.$$ Then

$$\therefore \sum_{i=1}^{n}\Big[\ln(1-\mathrm{Gi})\,K_4 - \theta\lambda^{\alpha}x^{-\alpha}K_4 - \ln\left(e^{\theta}+e^{\theta}\gamma^{-\alpha} - e^{\theta\lambda^{\alpha}x^{-\alpha}\gamma^{-\alpha}}\right)K_4 + \ln\left(e^{2\theta}+e^{2\theta}\gamma^{-\alpha} - e^{\theta+\theta\gamma^{-\alpha}}\right)K_4\Big] = 0 \quad (15).$$

After derivation and equating Equations (9), (11), (13) and (15) with zero, non-linear equations appeared that could not be solved with ordinary analysis methods. Therefore, one of the numerical analysis methods must be used, as the Newton-Raphson method was used to find the estimated parameters using the Percentiles Method ($\hat{\alpha}.\ \hat{\theta}.\ \hat{\lambda}.\ \hat{\gamma}$). After finding the values of the estimators, these values are substituted into the equation (5) in order to obtain the Percentile estimators of the reliability function of the distribution.

(2-2) Estimation with the Downhill simplex algorithm [3, 4, 7]

The Downhill simplex algorithm is a geometric algorithm, also called Nelder-Mead, that was proposed by Nelder and Mead in 1965. It is used for numerical optimization, uses objective function values, and relies on geometric foundations (reflection, expansion, contraction and shrink). Many research studies have worked on using this algorithm as a numerical method to facilitate the process of estimating parameters in events. Therefore, there are equations that can only be solved using numerical methods, but in this study, the algorithm is used to estimate the four parameters and estimate the reliability function of the weighted (Pareto-Poisson) distribution directly. The steps of the algorithm are as follows:

1- Determine the objective function to find the estimators. The objective function is the probability density function of the weighted (Pareto-Poisson) distribution in Equation (4),

2- Define the algorithm parameters (σ_1, σ_2, σ_3, and σ_4) where (σ_1: Reflection, σ_2: Expansion, σ_3: Contraction, σ_4: Shrink) and specify the number of solutions (n). In most research that studied this algorithm, the values (σ_1=1, σ_2=2, σ_3=0.5, σ_4=0.5) these parameters are defined as follows:

$\sigma_1 > 0,\ \sigma_2 > 1,\ 0 < \sigma_3 < 1,\ 0 < \sigma_4 < 1$

3- configure array $(n \times 4)$

$$H = \begin{bmatrix} \alpha_1 & \theta_1 & \lambda_1 & \gamma_1 \\ \alpha_2 & \theta_2 & \lambda_2 & \gamma_2 \\ \vdots & \vdots & \vdots & \vdots \\ \alpha_n & \theta_n & \lambda_n & \gamma_n \end{bmatrix} \quad (16).$$

4- Find the value of the objective function by substituting the values of each row in the above matrix into the objective function

5- Find the mean square error (MSE) of the objective function for each row in the matrix (H)

$$\mathrm{MSE}\big(f(x)\big) = \frac{\sum_{i=1}^{n}\big(\hat{f}(x)-x_i\big)^2}{n} \quad (17).$$

x_i: Viewing data

$\hat{f}(x)$: is data obtained from estimating distribution parameters

6- Arranging the values of the mean square error (MSE) from the lowest value, which represents the best solution, to the highest value, which represents the worst solution

7- Find the mean of the solution matrix

$$M = \frac{1}{n}\sum_{i=1}^{n} H_i \quad (18).$$

$H_i = [\hat{\alpha}_i \quad \hat{\theta}_i \quad \hat{\lambda}_i \quad \hat{\gamma}_i]$

8- Finding a test point that represents the Reflection point (Rp) is found as follows:

$$Rp = M + \sigma_1(M - H_n) \quad (19).$$

And, calculate the objective function f(Rp). If it is $f(H_1) < f(Rp) < f(H_n)$, this means that the reversal point is located between the best and worst points, where the worst point (H_n) will be replaced by the point (Rp). If it is ($f(Rp) \le f(H_1)$), this means that the reversal is the best solution. Go to the next step.

9- Calculating a new test point that represents the Expansion point (Ep) and is calculated as follows:

$$Ep = Rp + \sigma_2(Rp - M) \quad (20).$$

Find the objective function of the point f(Ep)

- If $f(Ep) < f(Rp)$ the worst point H_n will be replaced by Expansion (Ep)
- If $f(Rp) \geq f(H_n)$ will move to the next step

10- Generate a new test point representing the Contraction point (Cp) where two states are met

a- Find the External Contraction point (ECp). If $f(Rp) < f(H_n)$ then the following formula is used to find

$$ECp = M + \sigma_3(Rp - M) \qquad (21).$$

The objective function f(ECp) is found. If it is $f(ECp) \leq f(Rp)$, we replace the worst point (H_n) with the external contraction point (ECp). Otherwise, we go to step (11).

b- If $f(Rp) \geq f(Hn)$ then the Internal Contraction point (ICp) is found by the following formula

$$ICp = M - \sigma_3(M - H_n) \qquad (22).$$

Find the objective function for the internal contraction point f(ICp). If $f(ICp) < f(H_n)$. Then replace the worst point (H_n) with an internal contraction ICp otherwise we go to step (11).

11- Calculate a new test point representing the Shrink point (Sp) through the following formula

$$Sp = Sp_1 + \sigma_4(Sp_i - Sp_1) \,.\, i = 2.3. \dots \dots . n + 1 \qquad (23).$$

12- When the stopping condition is met, the best solution found that achieves the lowest value of the objective function will be printed according to the following formula

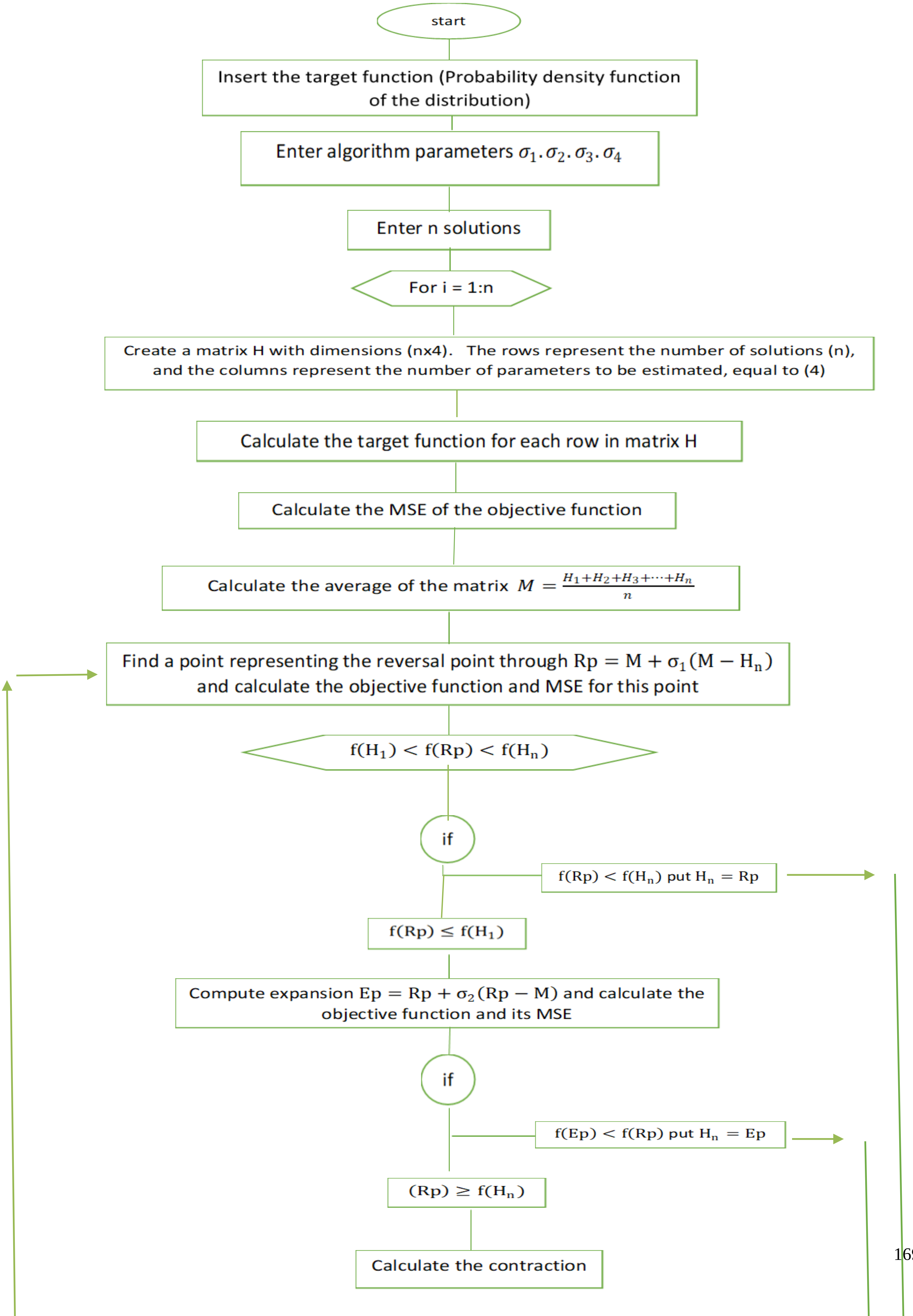
$$\left|\frac{\max(f) - \min(f)}{\max(f)}\right| < \in \qquad (24).$$

Where $\in$ is equal to $0 \cdot 5 \times 10^{-4}$. If this condition is achieved, go to the next step, otherwise return to point (8)

13- Print the optimal solution, which is the one that makes the function as small as possible which is the parameters that achieve the lowest value for the function

14- Stop repeating

The working of the algorithm can be illustrated in the following diagram


start
Insert the target function (Probability density function of the distribution)
Enter algorithm parameters $\sigma_1.\sigma_2.\sigma_3.\sigma_4$
Enter n solutions
For i = 1:n
Create a matrix H with dimensions (nx4). The rows represent the number of solutions (n), and the columns represent the number of parameters to be estimated, equal to (4)
Calculate the target function for each row in matrix H
Calculate the MSE of the objective function
Calculate the average of the matrix $M = \frac{H_1+H_2+H_3+\cdots+H_n}{n}$
Find a point representing the reversal point through $Rp = M + \sigma_1(M - H_n)$ and calculate the objective function and MSE for this point
$f(H_1) < f(Rp) < f(H_n)$
if
$f(Rp) < f(H_n)$ put $H_n = Rp$
$f(Rp) \le f(H_1)$
Compute expansion $Ep = Rp + \sigma_2(Rp - M)$ and calculate the objective function and its MSE
if
$f(Ep) < f(Rp)$ put $H_n = Ep$
$(Rp) \ge f(H_n)$
Calculate the contraction

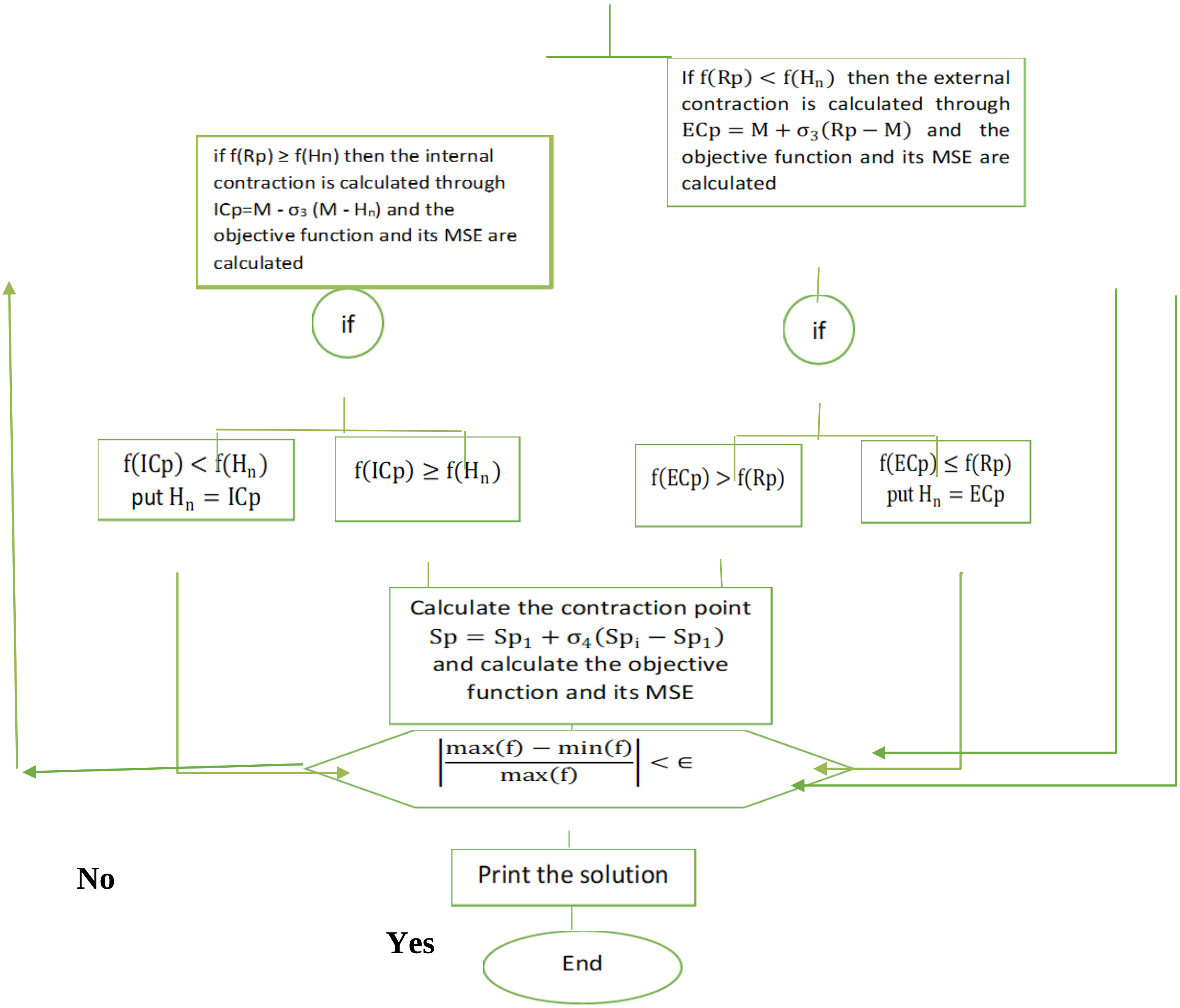


Diagram 1. The operation of the Downhill simplex algorithm in estimating parameters

3. Experimental Aspect

In this aspect, simulation was conducted in order to compare the estimator results of the methods used in estimation, and the simulation steps were as follows:

- Choose default values for complex distribution parameters (α,θ,λ,γ), Table 1 shows these values

Table 1. Default values for parameters

α	**θ**	**λ**	**γ**
0.5	0.8	0.1	0.5
0.5	0.8	0.2	0.6
0.5	0.4	0.1	0.5

- Three different sizes were chosen (small=30, medium=60, large=90) repeat the process 1000 times.
- Choose five failure times in order to estimate the reliability function.

Failure time	1.09	2.09	3	4	5

- Use the gamma distribution was used to generate the random variable
- Use the Mean Square Error (MSE) measure to compare and obtain the best estimator

$$\text{MSE} = \frac{1}{R}\sum_{i=1}^{R}(\varphi - \hat{\varphi})^2$$

Where

φ: The true value of the parameter.

$\hat{\varphi}$: Estimated value of the parameter.

R: Number of repetitions.

3. Result Simulation of Analysis

3-1 Parameter estimation

The results of the estimated parameters of the simulation experiment, which were written using a program written in Python, were presented and analyzed according to the following tables (2, 3, 4):

Table 2. Estimated values and MSE values for the estimation methods used
When $\alpha = 0.5, \theta = 0.8, \lambda = 0.1, \gamma = 0.5$

Parameters			**Per**	**D.S**
n=30	α	$\hat{\alpha}$	1.4146762450	0.1336148320
		MSE	0.0008366326	0.0001342380
	θ	$\hat{\theta}$	0.2385609277	0.0545056107
		MSE	0.0003152138	0.0005557618
	λ	$\hat{\lambda}$	0.4962250471	0.8926115677
		MSE	0.0001569942	0.0006282330
	γ	$\hat{\gamma}$	0.9738718845	0.9454559943
		MSE	0.0002245545	0.0001984310
	α	$\hat{\alpha}$	1.3341285061	0.8547592365
		MSE	0.0006957703	0.0001258547

n=60	θ	$\hat{\theta}$	0.0630553692	0.8845294946
		MSE	0.0005430873	7.14523545e-06
	λ	$\hat{\lambda}$	0.2937550611	0.1121420616
		MSE	3.75410237e-05	1.47429659e-07
	γ	$\hat{\gamma}$	0.7459153455	1.9706235482
		MSE	6.04743571e-05	0.0021627336
n=90	α	$\hat{\alpha}$	1.42282883108	0.2473931031
		MSE	0.0008516130	6.38102443e-05
	θ	$\hat{\theta}$	0.10523513754	0.3929594495
		MSE	0.0004826982	0.0001656820
	λ	$\hat{\lambda}$	0.36552498858	1.5983804680
		MSE	7.05035195e-05	0.0022451440
	γ	$\hat{\gamma}$	0.86242165151	0.8849337178
		MSE	0.0001313494	0.0001481739

From table (2), we note that:
The Downhill simplex method is the best among all estimation methods at sample sizes (30, 60, 90) because it has the lowest MSE.

Table 3. Estimated values and MSE values for the estimation methods used
When $\alpha = 0.5, \theta = 0.8, \lambda = 0.2, \gamma = 0.6$

Parameters			**Per**	**D.S**
n=30	α	$\hat{\alpha}$	1.3681793504	0.5414622080
		MSE	0.0007537353	1.71911469e-06
	θ	$\hat{\theta}$	0.1462652333	1.1492121831
		MSE	0.0004273691	0.0001219491
	λ	$\hat{\lambda}$	0.3891356936	0.1810708931

		MSE	3.57723105e-05	3.58311088e-07
	γ	$\hat{\gamma}$	0.8479971668	1.3471418602
		MSE	6.15025947e-05	0.0005582209
n=60	α	$\hat{\alpha}$	1.4341943205	0.3961635974
		MSE	0.0008727190	1.07819985e-05
	θ	$\hat{\theta}$	0.1749570692	0.3520626166
		MSE	0.0003906786	0.0002006478
	λ	$\hat{\lambda}$	0.4428259178	0.5010736742
		MSE	5.89644263e-05	9.06453572e-05
	γ	$\hat{\gamma}$	0.9468900841	1.6898706565
		MSE	0.0001203327	0.0011878180
n=90	α	$\hat{\alpha}$	1.3768613492	0.5939803040
		MSE	0.0007688858	8.83229753e-06
	θ	$\hat{\theta}$	0.0582849149	0.3075219648
		MSE	0.0005501412	0.0002425346
	λ	$\hat{\lambda}$	0.3043001970	0.1578025374
		MSE	1.08785310e-05	1.78062584e-06
	γ	$\hat{\gamma}$	0.7783323685	1.6650097454
		MSE	3.18024336e-05	0.0011342457

We observe from table (3):

The Downhill simplex method is the best among all estimation methods at sample sizes (30, 60, 90) because it has the lowest MSE.

Table 4. Estimated values and MSE values for the estimation methods used
When $\alpha = 0.5,\ \theta = 0.4,\ \lambda = 0.1,\ \gamma = 0.5$

Parameters			**Per**	D.S
n=30	α	$\hat{\alpha}$	1.4085419354	1.2029637077
		MSE	0.0008254484	0.0004941579
	θ	$\hat{\theta}$	0.0992952894	0.7081020392

		MSE	0.0004909870	8.44523519e-06
	λ	$\hat{\lambda}$	0.3657365453	0.2080369980
		MSE	2.74686024e-05	6.45933368e-08
	γ	$\hat{\gamma}$	0.8666315442	1.6417585531
		MSE	7.10923803e-05	0.0010852608
n=60	α	$\hat{\alpha}$	1.4064573756	0.4789738095
		MSE	0.0008216649	4.42100686e-07
	θ	$\hat{\theta}$	0.1774616457	0.0371225908
		MSE	0.0003875540	0.0005819819
	λ	$\hat{\lambda}$	0.4377818129	0.3407712259
		MSE	5.65401905e-05	1.98165380e-05
	γ	$\hat{\gamma}$	0.9167500397	2.0437253865
		MSE	0.0001003305	0.0020843429
n=90	α	$\hat{\alpha}$	1.2879092636	1.2617436351
		MSE	0.0006208010	0.0005802533
	θ	$\hat{\theta}$	0.0081566311	0.3618796687
		MSE	0.0006270159	0.0001919494
	λ	$\hat{\lambda}$	0.2225596698	1.8016675339
		MSE	5.08938701e-07	0.0025653388
	γ	$\hat{\gamma}$	0.6445909614	0.9985236997
		MSE	1.98835383e-06	0.0001588211

3-2 Estimating the reliability function [5]

In this paragraph, the reliability function of the distribution will be estimated and compared with the true value, as follows tables (5, 6, 7):

Table 5:. The values when $\alpha = 0.5, \theta = 0.8, \lambda = 0.1, \gamma = 0.5$

M	t	Real	Per	D.S
n=30	1.09	0.99850782	0.97999314	0.99999910
	MSE		0.0003427933	2.22391603e-06
	2.09	0.96362552	0.96709587	0.99997051
	MSE		1.20433291e-05	0.0013209582
	3	0.94789179	0.96326856	0.99994044
	MSE		0.0002364450	0.0027090619

	4	0.93705140	0.96129170	0.99991059
	MSE		0.0005875921	0.0039512777
	5	0.92957358	0.96021013	0.99988430
	MSE		0.0009385981	0.0049435973
n=60	1.09	0.99850782	0.999329169	0.59311069
	MSE		6.74614179e-07	0.1643468330
	2.09	0.96362552	0.9988113316	0.56855993
	MSE		0.001238041337	0.1560768204
	3	0.94789179	0.99865252	0.55997714
	MSE		0.0025766517	0.1504777756
	4	0.93705140	0.99856823	0.55485495
	MSE		0.0037843203	0.1460741263
	5	0.92957358	0.99852109	0.55168200
	MSE		0.0047537591	0.1428020462
n=90	1.09	0.99850782	0.99630727	0.99738819
	MSE		4.84242030e-06	1.25357133e-06
	2.09	0.96362552	0.99443732	0.99995773
	MSE		0.0009493670	0.0013200294
	3	0.94789179	0.99389164	0.99817769
	MSE		0.0021159862	0.0025286717
	4	0.93705140	0.99361100	0.99574575
	MSE		0.0031989883	0.0034450267
	5	0.92957358	0.99345787	0.99341924
	MSE		0.0040812025	0.0040762683

Note from table (5): The Downhill simplex method is the best among all estimation methods at sample sizes (30,90) because it has the lowest MSE, and the Percentiles Estimators method is the best at sample sizes (60).

Table 6. *The values when* $\alpha = 0.5, \theta = 0.8, \lambda = 0.2, \gamma = 0.6$

M	t	Real	Per	D.S
n=30	1.09	0.97920504	0.99423636	0.67580483
	MSE		0.0002259405	0.0920516874
	2.09	0.93552150	0.99048863	0.61429060
	MSE		0.0030213853	0.1031892911
	3	0.91510155	0.98934631	0.58931127
	MSE		0.0055122843	0.1061393065
	4	0.90087741	0.98874557	0.57297675
	MSE		0.0077208135	0.1075188428
	5	0.89101105	0.98841234	0.56210538
	MSE		0.0094870112	0.1081789397
	1.09	0.97920504	0.98819296	0.98194552

n=60	MSE		8.07827059e-05	7.51023063e-06
	2.09	0.93552150	0.98179818	0.96553606
	MSE		0.0021415311	0.0009008738
	3	0.91510155	0.97993388	0.95704455
	MSE		0.0042032310	0.0017592152
	4	0.90087741	0.97897825	0.95075816
	MSE		0.0060997412	0.0024880892
	5	0.89101105	0.97845829	0.94619431
	MSE		0.0076470197	0.0030451921
n=90	1.09	0.97920504	0.99922103	0.94233885
	MSE		0.0004006398	0.0013591159
	2.09	0.93552150	0.99875758	0.93069547
	MSE		0.0039988018	2.32905655e-05
	3	0.91510155	0.99861933	0.92576412
	MSE		0.0069752195	0.0001136903
	4	0.90087741	0.99854713	0.92248081
	MSE		0.0095393742	0.0004667068
	5	0.89101105	0.99850724	0.92027289
	MSE		0.0115554308	0.0008562552

In the table (6), The Percentiles Estimators method is the best among all estimation methods at sample sizes (30), and the Downhill simplex method is the best at a sample size of (60,90)

Table 7. The values when $\alpha = 0.5,\ \theta = 0.4,\ \lambda = 0.1,\ \gamma = 0.5$

M	t	Real	Per	D.S
n=30	1.09	0.99248442	0.99663735	0.67962378
	MSE		1.72468275e-05	0.0978817800
	2.09	0.98156066	0.99494237	0.67962378
	MSE		0.0001790701	0.1071959002
	3	0.97644055	0.99444332	0.64670454
	MSE		0.00032409972	0.1087258362
	4	0.97284371	0.99418530	0.64272545
	MSE		0.0004554634	0.1089780655
	5	0.97032975	0.99404394	0.64046671
	MSE		0.0005623628	0.1088096251
n=60	1.09	0.99248442	0.98936780	0.99939415
	MSE		9.71332022e-06	4.77443686e-05
	2.09	0.98156066	0.98300593	0.99907901
	MSE		2.08880537e-06	0.0003068925
	3	0.97644055	0.98111014	0.99892640

	MSE		2.18050707e-05	0.0005056134
	4	0.97284371	0.98012726	0.99881708
	MSE		5.30501006e-05	0.0006746159
	5	0.97032975	0.97958799	0.99873955
	MSE		8.57150078e-05	0.0008071167
n=90	1.09	0.99248442	0.99999960	0.85653309
	MSE		5.64779304e-05	0.0184827641
	2.09	0.98156066	0.99999324	0.99647989
	MSE		0.0003397600	0.0002225834
	3	0.97644055	0.99999126	0.97526693
	MSE		0.0005546359	1.37738390e-06
	4	0.97284371	0.99999020	0.95810281
	MSE		0.0007369319	0.0002172941
	5	0.97032975	0.99998960	0.94705345
	MSE		0.0008797067	0.0005417861

Note from table (7) The Percentiles Estimators method is the best among all estimation methods at sample sizes (30, 60), and the Downhill simplex method is the best at a sample size of (90)

All real data were collected and applied to the best estimation method used in this research, which represents the period of survival of a patient with thalassemia until death, which is a period calculated in years. The sample consider of 60 people. This data was obtained from the Thalassemia Center in the Diyala Health Department in Diyala Governorate on 21/5/2024, and the following table shows this data.

4. Experiment data

Table 8. Real sample data for patient survival time in years for thalassemia patients

Sequence	Times of failure (T)	Sequence	Times of failure (T)	Sequence	Times of failure (T)
1	0.6	21	3.9	41	7.8
2	0.7	22	4.2	42	7.8
3	0.8	23	4.2	43	7.9
4	1	24	4.3	44	8.1
5	1.1	25	4.3	45	8.4
6	1.11	26	4.6	46	8.5
7	1.11	27	4.6	47	9.1
8	1.3	28	4.7	48	10
9	1.8	29	5	49	10.1

10	1.9	30	5.1	50	10.3
11	2.1	31	5.4	51	11
12	2.3	32	5.4	52	12
13	2.5	33	5.5	53	12.9
14	2.5	34	5.7	54	14.2
15	2.9	35	5.9	55	14.6
16	3.1	36	5.9	56	14.7
17	3.1	37	6	57	15.5
18	3.1	38	6.1	58	15.9
19	3.5	39	6.4	59	17.7
20	3.8	40	7.7	60	27.3

5. Test data

In order to know whether the real data for the applied aspect follow the weighted composite (Pareto-Poisson) distribution with four parameters (α, θ, λ, γ), its parameters are estimated using simulation, Also, a goodness of fit test was conducted based on the best estimation method that appeared in the simulation. The Kolmogorov Smirnov Test and Chi-square test, were are conducted according to the following hypotheses.

Null hypothesis H0: The data follow the weighted (Pareto-Poisson) distribution

Alternative hypothesis H1: The data do not follow the weighted (Pareto-Poisson) distribution

The results are compared with the level of significance (0.05). All results for the two tests were calculated using the Python program. The table below shows the values of the goodness of fit tests that are mentioned.

Table (9): Results of testing the suitability of real data to a weighted Pareto-Poisson distribution

Test	**Statistic**	**P-value**
Kolmogorov Smirnov	0.167874247	0.06370857
Chi-Square	13.01334829	0.07178385

In the table above, the results of the tests show acceptance of the null hypothesis of the Kolmogorov Smirnov test for the weighted (Pareto-Poisson) distribution, where we notice that the p-values are greater than 0.05 in the two tests mentioned. The values reached (0.0637) for the Kolmogorov Smirnov test and (0.0717) for the Chi-square test. Thus we accept the alternative hypothesis of the chi-square test.

The results of the data for the applied side showed that the probability of the patient surviving for 6 months is (0.987878), the probability of surviving for 5 years is (0.720039), and the probability of surviving for 11 years is (0.693401). This means that the probability of survival decreases with the increase in the number of years, as it becomes (0.684258) when 27 years and 3 months have passed, As shown in the following table:

Table 10. Estimated values of the reliability function for real data

Times of failure (T)	$\hat{R}$	Times of failure (T)	$\hat{R}$	Times of failure (T)	$\hat{R}$
0.6	0.987878	3.9	0.736274	7.8	0.701551
0.7	0.975997	4.2	0.730883	7.8	0.701551
0.8	0.961391	4.2	0.730883	7.9	0.701172
1	0.929121	4.3	0.729274	8.1	0.700448
1.1	0.913178	4.3	0.729274	8.4	0.699439
1.11	0.911621	4.6	0.724921	8.5	0.699122
1.11	0.911621	4.6	0.724921	9.1	0.697396
1.3	0.883758	4.7	0.723610	10	0.695272
1.8	0.827646	5	0.720039	10.1	0.695064
1.9	0.819071	5.1	0.718957	10.3	0.694664
2.1	0.804002	5.4	0.715987	11	0.693401
2.3	0.791284	5.4	0.715987	12	0.691907
2.5	0.780481	5.5	0.715081	12.9	0.690802
2.5	0.780481	5.7	0.713380	14.2	0.689506
2.9	0.763283	5.9	0.711815	14.6	0.689164
3.1	0.756384	5.9	0.711815	14.7	0.689082
3.1	0.756384	6	0.711078	15.5	0.688474
3.1	0.745079	6.1	0.710370	15.9	0.688198
3.5	0.745079	6.4	0.708403	17.7	0.687145
3.8	0.738283	7.7	0.701942	27.3	0.684258

6. Conclusions

The most important conclusions reached by the researcher in the simulation experiments to estimate the parameters and the reliability function are:

The Downhill simplex method was the best of the two methods by which it was estimated because it has the lowest mean square error in all hypothetical data and for all sample sizes. The percentile estimator’s method was the best in estimating the reliability function because it had the least mean square error in all sample sizes.

7. Recommendations

1- The researcher recommends using other methods to estimate the parameters of the weighted composite (Pareto-Poisson) distribution other than the methods used by the researcher, including Bayesian methods.

2- Using the weighted Pareto-Poisson distribution in other applications, especially industrial, engineering, and others
3- Use other failure data and different sample sizes
4- Inserting weights into other composite distributions

References

[1] A. A. Azzallini's class of distributions which includes the normal ones. JSTOR.1985; 12(2): 171-178.
[2] A. El-Shahhat The Pareto–Poisson Distribution: Characteristics, Estimations and Engineering Applications. IJS. 2022 December; 85(1): 1-42
[3] A. Moraglio and C. Johnson, 'Geometric Generalization of the Nelder-Mead Algorithm' Proc. 10th Eur. Conf. Evol. Comput. in Combin. Optim. (EvoCOP 2010), Istanbul, Turkey, Apr. 7-9, 2010, pp. 131-142.
[4] A. Moraglio and J. Togelius, 'Geometric Nelder-Mead Algorithm for the permutation representation', in "Proceedings of the IEEE Congress on Evolutionary Computation" Aug. 2010, pp. 1–8.
[5] B. K. Fahad, 'Choose the Best Method for Estimation the Parameters of Probability Kappa Distribution With practical application', Thesis, University of Karbala, Karbala, 2018.
[6] E. Q. Abdel Rahman find better estimate of the parameters and reliability function of the compound distribution (Pareto – Truncated Exponential) With practical application. Al-Basra: Al-Basra University; 2022.
[7] L. A. J. D. A. Helfi, 'Comparison of the methods of estimating the parameters and reliability function for the lambda distribution with four parameters with practical application', Thesis, University of Baghdad, Baghdad, 2015.