Journal of Economics
and Administrative Sciences (JEAS)

Available online at http://jeasiq.uobaghdad.edu.iq

Comparing Some Methods of Estimating the Parameters and Survival Function of a Log-logistic Distribution with a Practical Application

Samah Sabah Hassan al-aameri

Prof.Dr.Entsar Arebe Fadam AL.Doori

College of Administration and Economics / University of Baghdad

entsar_arebe@coadec.uobaghdad.edu.iq

Samahsabah1993@gmail.com

Received: 4/10/2021 **Accepted:2/11/2021** **Published: March / 2022**

Abstract

The Log-Logistic distribution is one of the important statistical distributions as it can be applied in many fields and biological experiments and other experiments, and its importance comes from the importance of determining the survival function of those experiments. The research will be summarized in making a comparison between the method of maximum likelihood and the method of least squares and the method of weighted least squares to estimate the parameters and survival function of the log-logistic distribution using the comparison criteria MSE, MAPE, IMSE, and this research was applied to real data for breast cancer patients. The results showed that the method of Maximum likelihood best in the case of estimating the parameters of the distribution and survival function.

Keywords : Survival function, log-logistic distribution, Maximum likelihood method, least squares method, weighted least method .

1. Introduction

The log-logistic distribution still is an important statistical distribution that can be applied to many biological, life, chemical, and physical experiments. Therefore, the survival function of this distribution has acquired importance through the possibility of using it to expect survival for many life experiments. The research is in its study of the log-logistic distribution of survival functions for conducting a series of simulation experiments with sufficient repetitions with different parameter values and models with the adoption of different methods for estimating parameters and the survival function using statistical criteria. Several previous studies of this distribution were conducted in 1995 the researchers (Singh & Guo) made a study to estimate the parameters of the log-logistic distribution. They used the principle of maximum entropy (POME) to use a new method to estimate the parameters of the two-parameter log-logistic distribution. In (2005), researcher Fitzgerald analyzed heavy rainfall using the log-logistic distribution. He used intensive computer methods to examine t-values from the log-logistic distribution to the Irish annual maximum precipitation. In the study, he examined the properties of the L-moment using the traditional bootstrap in the data. In the year 2021, the researcher (Abd-Elmonem A. M) and others presented a relatively new heavy-tail statistical model using alpha power transformation and exponentiated log-logistic distribution, which was called alpha power exponentiated log-logistic distribution. Its statistical properties were derived mathematically such as moments, quantitative function, entropy, inequality curves, and order statistics. Six estimation methods were presented mathematically and the behavior of the proposed model parameters was verified through real data sets, some actuarial measures were deduced mathematically such as the value-at-risk tail .

2. Log-Logistic Distribution:

The log-logistic distribution is one of the continuous probability distributions that has attracted the interest of researchers Raghab A. and Green (1984). It has been used extensively in models of population, growth and biological problems and has some important applications in solving many practical problems, especially in survival data as a model of events, the log-logistic distribution is also known as the Fisk distribution in economics. Suppose $(t_1 \le t_2 \le \cdots \le t_n)$ an ordinal sample of size n drawn from a population that follows a log-logistic distribution whose probability function is as follows:

$$f(x, \alpha, \beta) = \frac{\frac{\beta}{\alpha}\left(\frac{x}{\alpha}\right)^{\beta-1}}{\left[1+\left(\frac{x}{\alpha}\right)^{\beta}\right]^{2}} \qquad x > 0 \quad \alpha, \beta > 0 \ , [0, \infty] \qquad (1)$$

Where : α is Scale Parameter , β is Shape Parameter

The cumulative distribution function is given as follows:

$$F(x) = P_r(X < x) = \int_0^t f(k)\, dk = \frac{1}{1+\left(\frac{x}{\alpha}\right)^{-\beta}} \qquad (2)$$

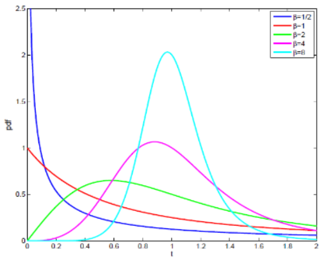

Figure (1) represents the pdf curve of the Log-logistic of different values of

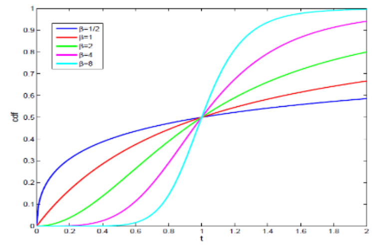

Figure (2) represents the CD curve of the Log- logistic distribution of different values of β

The Survival function:

The survival function is the probability of the patient surviving during a certain period under special conditions and factors. The survival function is denoted by S (t), and the survival function is expressed by the following mathematical equation:

$$S(x) = \Pr(X > x)$$
$$= \int_t^{\infty} f(y)dy$$
$$= 1 - F(x)$$
$$= \frac{1}{1+\left(\frac{x}{\alpha}\right)^{\beta}} \qquad (3)$$

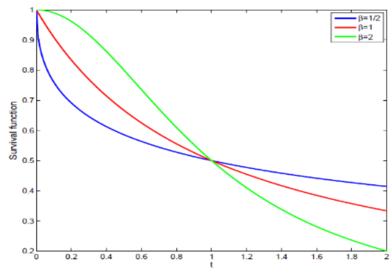

Figure (3) represents the survival curve of the log-logistic distribution of different values of β

3. Properties of a log-logistic distribution :

i. Central moment:

$$ET^k = \int_0^{\infty} T^k f(t)\, dt = \alpha^K B\left(1 - \frac{k}{\beta}, 1 + \frac{K}{B}\right); where\ K = 1, 2, \dots;$$

$$; B(a,b) = \int_0^1 t^{a-1}(1-t)^{b-1} dt$$

$$\alpha^k(k\pi / \beta)\, csc(k\pi / \beta) \quad if\ \beta > 1 \qquad \text{AshaDixit [7]}$$

ii. Mean :

$$M_t = \alpha(\pi / \beta)\, csc(\pi/\beta)$$

iii. Variance :

$$\sigma_t^2 = \alpha^2[(2\pi / \beta)\csc(2\pi / \beta) - (\pi/\beta)^2 \csc^2(\pi/\beta)]$$

iv-Median :

Me=α

v- Mode :

$$Mo = \alpha\left(\frac{\beta_{-1}}{\beta+1}\right)^{\frac{1}{\beta}} \quad \text{if } \beta > 1$$

4- Estimation methods:

4-1 Maximum likelihood estimation :

This method is considered one of the most important estimation methods for its good properties, including stability, high efficiency and consistency in some cases. Suppose we have a number many observations from the log-logistic distribution, denoted by $x_1 \le x_2 \le \ldots \le x_n$. The logarithmic probability function for α,β is as follows:

$$L(\alpha,\beta\,|x_i) = \beta^n\alpha^{-n\beta}\prod_{i=1}^{n} x_i^{\beta-1}\prod_{i=1}^{n}\left[1+\left(\frac{x_i}{\alpha}\right)^{\beta}\right]^{-2} \quad (4)$$

The Log-likelihood function can be written

$$Ln\,L(\alpha,\beta|x_i) = n\,ln\,\beta - n\beta ln\,\alpha + (\beta-1)\sum_{i=1}^{n} ln\,(x_i) - 2\sum_{i=1}^{n} ln\,\left(1+\left(\frac{x_i}{\alpha}\right)^{\beta}\right) \quad (5)$$

We derive equation (5) for the parameters α and β

$$\frac{\partial \ln L(\alpha,\beta|x_i)}{\partial\alpha} = -\frac{n\beta}{\alpha} + \frac{2\beta}{\alpha}\sum_{i=1}^{n}\frac{1}{1+\left(\frac{x_i}{\alpha}\right)^{-\beta}} \quad (6)$$

$$\frac{\partial \ln L(\alpha,\beta|x_i)}{\partial\beta} = \frac{n}{\beta} - n\ln(\alpha) + \sum_{i=1}^{n} ln\,(x_i) - 2\sum_{i=1}^{n}\frac{ln\left(\frac{x_i}{\alpha}\right)}{1+\left(\frac{x_i}{\alpha}\right)^{-\beta}} \quad (7)$$

And then we equalize equations (6) and (7) with zero and then solve them by one of the numerical methods to obtain the estimators of the maximum likelihood for the parameters α and β .

The estimator for the survival function becomes S(x) as x > 0 .

$$\hat{s}(x_i)_{ML} = \frac{1}{1+\left(\frac{x_i}{\hat{\alpha}_{ML}}\right)^{\hat{\beta}_{ML}}} \quad (8)$$

4.2 Least square ordinary method:

The method of least squares is one of the important methods for estimating parameters because it works to find estimates by minimizing the error squares between the cumulative distribution function (CDF) of the studied distribution and one of the nonparametric estimations of the cumulative distribution function. It can be calculated mathematically as follows:

$$LS = \sum_{i=1}^{n}\left(F(x_i) - \frac{i}{n+1}\right)^2 \quad (9)$$

$$LS = \sum_{i=1}^{n}\left(\frac{1}{1+\left(\frac{x_i}{\alpha}\right)^{-\beta}} - \frac{i}{n+1}\right)^2$$

We derive for the α and β parameters

$$\frac{\partial LS}{\partial \alpha} = 2\sum_{i=1}^{n}\left(\frac{1}{1+\left(\frac{x_i}{\alpha}\right)^{-\beta}} - \frac{i}{n+1}\right)\frac{-\beta}{\alpha}\left(\frac{x_i}{\alpha}\right)^{-\beta}\left(1+\left(\frac{x_i}{\alpha}\right)^{-\beta}\right)^{-2}$$

Now we set the derivative equal to zero and simplify, we get:

$$\Rightarrow \sum_{i=1}^{n} \left(\frac{x_i}{\hat{\alpha}}\right)^{-\hat{\beta}}\left(1+\left(\frac{x_i}{\hat{\alpha}}\right)^{-\hat{\beta}}\right)^{-3}$$

$$= \sum_{i=1}^{n} \frac{i}{n+1}\left(\frac{x_i}{\hat{\alpha}}\right)^{-\hat{\beta}}\left(1+\left(\frac{x_i}{\hat{\alpha}}\right)^{-\hat{\beta}}\right)^{-2} \qquad (10)$$

$$\frac{\partial LS}{\partial \beta} = 2\sum_{i=1}^{n}\left(\frac{1}{1+\left(\frac{x_i}{\alpha}\right)^{-\beta}} - \frac{i}{n+1}\right)\left(\frac{x_i}{\alpha}\right)^{-\beta} ln\left(\frac{x_i}{\alpha}\right)\left(1+\left(\frac{x_i}{\alpha}\right)^{-\beta}\right)^{-2}$$

Now we set the derivative equal to zero and simplify, we get:

$$\Rightarrow \sum_{i=i}^{n} \left(\frac{x_i}{\hat{\alpha}}\right)^{-\hat{\beta}}\left(1+\left(\frac{x_i}{\hat{\alpha}}\right)^{-\hat{\beta}}\right)^{-3} ln\left(\frac{x_i}{\hat{\alpha}}\right)$$

$$= \sum_{i=1}^{n} \left(\frac{x_i}{\hat{\alpha}}\right)^{-\hat{\beta}}\left(1+\left(\frac{x_i}{\hat{\alpha}}\right)^{-\hat{\beta}}\right)^{-2}\left(\frac{i}{n+1}\right) ln\left(\frac{x_i}{\hat{\alpha}}\right) \qquad (11)$$

And equations (10) and (11) cannot be solved by ordinary mathematical methods, so we use one of the numerical methods, as the Newton-Raphson iterative numerical method, to get the estimates of $\hat{\alpha}_{Ls}$ and $\hat{\beta}_{LS}$.

Therefore, the estimator of the survival function S(x) based on the least squares method for the ordered data of the distribution is as follows:

$$\hat{s}(x_i)_{LS} = \frac{1}{1+\left(\frac{x_i}{\hat{\alpha}_{LS}}\right)^{\hat{\beta}_{LS}}} \qquad (12)$$

<u>4.3 Weighted least squares method:</u>

The weighted least squares method is one of the important estimation methods, and this method is distinguished from the previous method by the presence of the weight factor (Wi). [Kumar et.al, 2019, pp17] [Majli , 2020]

$$WLS = \sum_{i=1}^{n} Wi\left(F(T_i) - \frac{i}{n+1}\right)^2 \qquad (13)$$

Where: $w_i = \frac{i(n-i+1)}{(n+1)^2(n+2)}$

$$WLS = \sum_{i=1}^{n} Wi\left(\frac{1}{1+\left(\frac{t_i}{\alpha}\right)^{-\beta}} - \frac{i}{n+1}\right)^2$$

We derive for the α and β parameters

$$\frac{\partial WLS}{\partial \alpha} = 2\sum_{i=1}^{n} Wi\left(\frac{1}{1+\left(\frac{t_i}{\alpha}\right)^{-\beta}} - \frac{i}{n+1}\right)\frac{-\beta}{\alpha}\left(\frac{t_i}{\alpha}\right)^{-\beta}\left(1+\left(\frac{t_i}{\alpha}\right)^{-\beta}\right)^{-2}$$

Now we set the derivative equal to zero and simplify, we get:

$$\Rightarrow \sum_{i=1}^{n} Wi\left(\frac{t_i}{\hat{\alpha}}\right)^{-\hat{\beta}}\left(1+\left(\frac{t_i}{\hat{\alpha}}\right)^{-\hat{\beta}}\right)^{-3} = \sum_{i=1}^{n} Wi\frac{i}{n+1}\left(\frac{t_i}{\hat{\alpha}}\right)^{-\hat{\beta}}\left(1+\left(\frac{t_i}{\hat{\alpha}}\right)^{-\hat{\beta}}\right)^{-2} \quad (14)$$

$$\frac{\partial LS}{\partial \beta} = 2\sum_{i=1}^{n} Wi\left(\frac{1}{1+\left(\frac{t_i}{\alpha}\right)^{-\beta}} - \frac{i}{n+1}\right)\left(\frac{t_i}{\alpha}\right)^{-\beta} ln\left(\frac{t_i}{\alpha}\right)\left(1+\left(\frac{t_i}{\alpha}\right)^{-\beta}\right)^{-2}$$

Now we set the derivative equal to zero and simplify, we get:

$$\Rightarrow \sum_{i=i}^{n} Wi\left(\frac{t_i}{\hat{\alpha}}\right)^{-\hat{\beta}}\left(1+\left(\frac{t_i}{\hat{\alpha}}\right)^{-\hat{\beta}}\right)^{-3} ln\left(\frac{t_i}{\hat{\alpha}}\right)$$

$$= \sum_{i=1}^{n} Wi\left(\frac{t_i}{\hat{\alpha}}\right)^{-\hat{\beta}}\left(1+\left(\frac{t_i}{\hat{\alpha}}\right)^{-\hat{\beta}}\right)^{-2}\left(\frac{i}{n+1}\right) ln\left(\frac{t_i}{\hat{\alpha}}\right) \quad (15)$$

And equations (14) and (15) cannot be solved by ordinary mathematical methods, so we use one of the numerical methods, as the Newton-Raphson iterative numerical method, to get the estimates of $\hat{\alpha}_{Ls}$ and $\hat{\beta}_{LS}$.

Therefore, the estimator of the survival function S(x) based on the weighted least squares method for the ordered data of the distribution is as follows:

$$\hat{s}(x_i)_{wLS} = \frac{1}{1+\left(\frac{x_i}{\hat{\alpha}_{wLS}}\right)^{\hat{\beta}_{wLS}}} \quad (16)$$

5- Results and Discussion

5.1 Simulation Experiments

The simulation method was used in the Monte Carlo method to compare between the different estimation methods, as this method is characterized by flexibility and saves a lot of costs by taking into account the different sample sizes and the different values of the distribution parameters and the repetition of the experiment each time. In this method, data is generated without resorting to real data and without prejudice to the required accuracy. This method is summarized in the following steps:

i. **Define the default values :**

Default values are used for the two parameters based on the estimated values of the practical side and the selection of three sample sizes, which are (50,100,250).

ii. **Data generation :**

If the random variable is generated by the inverse transformation method as follows:

$$x_i = \alpha\left(\frac{1-u}{u}\right)^{-\frac{1}{\beta}} i = 1,2,3,\dots,n \quad (17)$$

iii. **Solve the equations obtained by numerical methods.**

iv. **The best method is determined by the comparison scale (IMSE) in the case of estimating the survival function, and probability distribution function.**

$$IMSE(\hat{S}(x)) = \frac{1}{r}\sum_{i=1}^{r}\left[\frac{1}{n_t}\sum_{j=1}^{n_t}\left(\hat{s}_i(x_j) - s(x_j)\right)^2\right] \qquad (18)$$

$$IMSE(\hat{f}(x)) = \frac{1}{r}\sum_{i=1}^{r}\left[\frac{1}{n_t}\sum_{j=1}^{n_t}\left(\hat{f}_i(x_j) - f(x_j)\right)^2\right] \qquad (19)$$

where:

r: The number of repetitions of the experiment (1000)times

n_t: The number of data generated per sample

$\hat{f}_i(x_j), \hat{s}_i(x_j)$: Estimator the probability distribution function and survival function, respectively

$f(x_j)$, $S(x_j)$: Probability distribution function and survival function according to the initial values, respectively.

v. Calculate the mean squared error (MSE) and mean squared absolute error (MAPE) for each value of the variable (x_i) for the α and β distribution parameters.

$$MSE(\hat{\alpha}_i) = \frac{\sum_{i=1}^{r}(\hat{\alpha}_i - \alpha_i)^2}{r} \qquad (20)$$

$$MSE(\hat{\beta}_i) = \frac{\sum_{i=1}^{r}(\hat{\beta}_i - \beta_i)^2}{r} \qquad (21)$$

$$MAPE(\hat{\alpha}_i) = \frac{\sum_{i=1}^{r}|\hat{\alpha}_i - \alpha_i|}{r} \qquad (22)$$

$$MAPE(\hat{\beta}) = \frac{\sum_{i=1}^{r}|\hat{\beta}_i - \beta_i|}{r} \qquad (23)$$

where:

r: The number of repetitions of the experiment (1000) times

$\hat{\alpha}_i, \hat{\beta}_i$: Estimators of distribution parameters, respectively.

α, β : The parameters of the distribution according to the initial values respectively.

- **<u>The results are represented by the following tables:</u>**
- **Table (1) shows the default and estimated parameter values for α and β, sample sizes, and for all experiments**

		$\hat{\alpha}$			$\hat{\beta}$		
Model	**n**	**MLE**	**OLS**	**WLS**	**MLE**	**OLS**	**WLS**
α=2.0,β=2.0	**50**	**2.05412**	**2.000036**	**2.997847**	**2.021349**	**2.999514**	**2.998063**
	100	**1.09172**	**2.000018**	**2.000657**	**1.028554**	**1.998609**	**1.998956**
	250	**1.037706**	**1.001085**	**1.999924**	**1.040964**	**1.99913**	**1.997603**
α=2.5,β=2.5	**50**	**2.490443**	**2.499445**	**2.50228**	**2.49015**	**2.499176**	**2.501289**
	100	**2.399779**	**2.491107**	**2.500231**	**2.423485**	**2.499966**	**2.501379**
	250	**2.534005**	**2.500365**	**2.497524**	**2.524394**	**2.502364**	**2.497461**
α=2.7,β=4.6	**50**	**2.709721**	**2.702244**	**2.700921**	**4.691514**	**4.692622**	**4.599715**
	100	**2.682198**	**2.701333**	**2.698383**	**4.608248**	**4.680102**	**4.598125**
	250	**2.718995**	**2.700196**	**2.697727**	**4.735429**	**4.602452**	**4.601017**

- **Table (2) shows the MSE values for parameter estimation (α) and (β) by all methods, sample sizes and for all experiments**

		$\hat{\alpha}$			$\hat{\beta}$		
Model	**n**	**MLE**	**OLS**	**WLS**	**MLE**	**OLS**	**WLS**
α=2.0,β=2.0	**50**	**2.72E-02**	**4.68E-01**	**1.45E-01**	**1.51E-02**	**1.95E-02**	**1.94E-02**
	100	**2.06E-02**	**1.30E-01**	**1.14E-01**	**1.20E-02**	**1.25E-02**	**1.41E-02**
	250	**1.40E-03**	**1.77E-02**	**1.76E-02**	**0.0006247**	**1.85E-03**	**1.43E-03**
α=2.5,β=2.5	**50**	**4.08E-02**	**1.14E-01**	**1.68E-01**	**2.97E-02**	**9.77E-02**	**1.29E-01**
	100	**4.99E-03**	**1.12E-01**	**1.06E-01**	**1.70E-02**	**1.74E-02**	**1.03E-01**
	250	**.38E-031**	**1.83E-03**	**1.77E-03**	**1.00E-03**	**1.55E-03**	**6.17E-03**
α=2.7,β=4.6	**50**	**1.52E-02**	**1.65E-01**	**1.49E-01**	**0.04609**	**5.37E-01**	**1.55E-01**
	100	**6.29E-03**	**1.54E-01**	**1.48E-01**	**9.72E-03**	**6.89E-02**	**8.14E-02**
	250	**2.51E-04**	**1.65E-02**	**1.65E-03**	**8.07E-03**	**1.55E-02**	**1.46E-02**

We notice from Table No. (2) and in all experiments, and at a sample size of n = 50, 100, 250, the method of Maximum likelihood appeared to be best in estimating the parameters of the distribution α and β .

- **Table (3) shows the MAPE values for parameter estimation (α) and (β) by all methods, sample sizes and for all experiments**

		$\hat{\alpha}$			$\hat{\beta}$		
Model	**n**	**MLE**	**OLS**	**WLS**	**MLE**	**OLS**	**WLS**
α=2.0,β=2.0	**50**	**0.120694**	**1.83E-01**	**3.32E-01**	**0.196073**	**3.39E-01**	**3.72E-01**
	100	**0.110064**	**1.24E-01**	**2.65E-01**	**0.162713**	**1.85E-01**	**2.38E-01**
	250	**.85E-021**	**2.38E-02**	**2.05E-02**	**2.39E-02**	**3.59E-02**	**2.92E-02**
α=2.5,β=2.5	**50**	**0.153058**	**2.78E-01**	**3.42E-01**	**0.102067**	**2.44E-01**	**3.20E-01**
	100	**0.150764**	**2.30E-01**	**2.88E-01**	**0.019428**	**2.16E-01**	**3.17E-01**
	250	**1.56E-03**	**3.42E-03**	**2.61E-03**	**0.0023499**	**3.40E-03**	**2.98E-03**
α=2.7,β=4.6	**50**	**8.68E-02**	**3.51E-01**	**3.39E-01**	**0.628493**	**7.89E-01**	**7.12E-01**
	100	**5.59E-02**	**3.06E-01**	**3.14E-01**	**0.271041**	**2.72E-01**	**2.18E-01**
	250	**3.02E-03**	**3.14E-03**	**4.12E-03**	**0.013454**	**3.25E-02**	**3.03E-02**

We notice from Table No. (3) and in all experiments, and at a sample size of n = 50, 100, 250, the method of Maximum likelihood appeared to be best in estimating the parameters of the distribution α and β .

When estimating the parameter (β) and in the case of the third experiment, and with a sample size of n = 100, the least squares weighted with best appeared.

- **Table (4) shows the IMSE values for estimating the survival function by all methods, sample sizes and for all experiments.**

Model	**n**	**MLE**	**OLS**	**WLS**	**best**
α=2.0,β=2.0	**50**	**8.91E-03**	**1.13E-02**	**1.13E-02**	**MLE**
	100	**8.31E-03**	**1.12E-02**	**1.12E-02**	**MLE**
	250	**1.16E-03**	**1.01E-02**	**1.02E-02**	**MLE**
α=2.5,β=2.5	**50**	**1.19E-02**	**9.06E-02**	**9.11E-02**	**MLE**
	100	**1.17E-02**	**9.04E-02**	**9.07E-02**	**MLE**
	250	**8.77E-03**	**0.009283**	**9.35E-03**	**MLE**
α=2.7,β=4.6	**50**	**7.30E-03**	**9.13E-03**	**9.11E-03**	**MLE**
	100	**7.21E-03**	**9.07E-03**	**9.10E-03**	**MLE**
	250	**1.04E-02**	**9.00E-03**	**9.07E-03**	**MLE**

We notice from Table No. (4) and in all experiments, and at a sample size of n = 50, 100, 250, the method of Maximum likelihood appeared best in estimating the survival function.

- Table (5) shows the IMSE values for estimating the probability distribution function for all methods, sample sizes and for all experiments.

model	n	MLE	OLS	WLS	best
α=2.0,β=2.0	50	3.05E-02	3.16E-02	3.16E-02	MLE
	100	3.19E-03	0.0033067	3.30E-03	MLE
	250	4.44E-05	4.54E-04	4.54E-04	MLE
α=2.5,β=2.5	50	0.048072	0.049546	4.95E-02	MLE
	100	4.96E-03	5.09E-03	5.09E-03	MLE
	250	6.02E-04	6.03E-03	6.05E-04	MLE
α=2.7,β=4.6	50	4.89E-02	5.19E-02	5.19E-02	MLE
	100	3.67E-02	3.74E-02	3.74E-02	MLE
	250	3.60E-03	3.77E-03	3.77E-03	MLE

We notice from Table No. (5) and in all experiments and at a sample size of n = 50, 100, 250, the method of Maximum likelihood appeared best in estimating the probability distribution function.

5.2 Applied Real Data

Real data were collected for (50) breast cancer patients at Medical City Hospital - Baghdad for the year 2020. At the time of the patient's entry to the center until discharge was recorded, and that all of them were in a state of death upon discharge, and this data was considered complete data .

Table (6) shows the real data

t_i	2.1 7.2 1.2 3.4 8.6 2.1 6.4 3 8.12 5.4 5.2 3.4 2.1 2.1 3.4 6.4 6.4
t_i	3.2 3.2 3.2 6.2 12.4 1.9 5.6 4.1 2.1 3.4 9.8 3.1 2.1 2.1 12.4 11.4
t_i	3.4 3.4 6.5 3.2 12.1 5.4 6.1 6.1 6.1 2.4 8.4 10.5 10.5 8.5 9.1 9.1 5.3

Table No. (6) represents the failure times for breast cancer patients in months, and the fractional number represents the failure times in days

Data fitting test:

To show the suitability of the real data to the log-logistic distribution, a test was conducted for the above data using the Standard Easy Fit 5.5 program to test the following statistical hypothesis at the 0.05 level of significance.

Ho : The data follows a log-logistic distribution

H1 : The data does not follow a log-logistic distribution

The results of the test are as in the following table (7):

Table(7) shows the fitting tests for the application data

Test	Sig
Kolmogorov-smirnov	0.13623
Anderson Darling	0.9763
Chi-squared	8.977

We notice from Table (7) that the sig value for all tests is greater than 0.05, so we accept the null hypothesis Ho, meaning that the selected real data follows a log-logistic distribution.

Table (8) shows the values of estimator parameters log-logistic distribution of maximum likelihood method

$\hat{\alpha}_{MLE}$	3.279519
$\hat{\beta}_{MLE}$	4.889562

Table (9) shows the estimation of the survival and probability density function for the application data

t_i	$\hat{S}(t)_{MLE}$	$\hat{f}(t)_{MLE}$	t_i	$\hat{S}(t)_{MLE}$	$\hat{f}(t)_{MLE}$
1.2	0.217695	2.94E-02	5.4	8.58E-03	3.90E-02
1.9	0.107034	0.156025	5.4	8.58E-03	3.90E-02
2.1	8.89E-02	0.212528	5.6	7.67E-03	3.14E-02
2.1	8.89E-02	0.212528	6.1	5.87E-03	1.87E-02
2.1	8.89E-02	0.212528	6.1	5.87E-03	1.87E-02
2.1	8.89E-02	0.212528	6.1	5.87E-03	1.87E-02
2.1	8.89E-02	0.212528	6.2	5.57E-03	1.69E-02
2.1	8.89E-02	0.212528	6.4	5.03E-03	1.39E-02
2.1	8.89E-02	0.212528	6.4	5.03E-03	1.39E-02
2.4	6.82E-02	0.298726	6.4	5.03E-03	1.39E-02
3	4.17E-02	0.388731	6.5	4.78E-03	1.26E-02
3.1	3.86E-02	0.386944	7.2	3.41E-03	6.66E-03
3.2	3.58E-02	0.380625	8.12	2.26E-03	3.12E-03
3.2	3.58E-02	0.380625	8.4	2.01E-03	2.52E-03
3.2	3.58E-02	0.380625	8.5	1.93E-03	2.33E-03
3.2	3.58E-02	0.380625	8.6	1.85E-03	2.17E-03
3.4	3.09E-02	0.356744	9.1	1.51E-03	1.51E-03
3.4	3.09E-02	0.356744	9.1	1.51E-03	1.51E-03
3.4	3.09E-02	0.356744	9.8	1.15E-03	9.46E-04
3.4	3.09E-02	0.356744	10.5	8.95E-04	6.10E-04
3.4	3.09E-02	0.356744	10.5	8.95E-04	6.10E-04
3.4	3.09E-02	0.356744	11.4	6.57E-04	3.62E-04
4.1	1.90E-02	0.22437	12.1	5.23E-04	2.47E-04
5.2	9.61E-03	8.09E-02	12.1	5.23E-04	2.47E-04
5.3	9.08E-03	7.35E-02	12.4	4.76E-04	2.12E-04

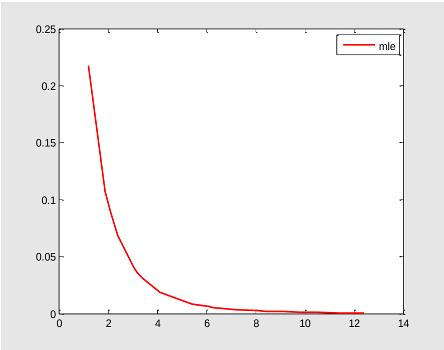

Figure (4) represents the behavior estimating the survival function

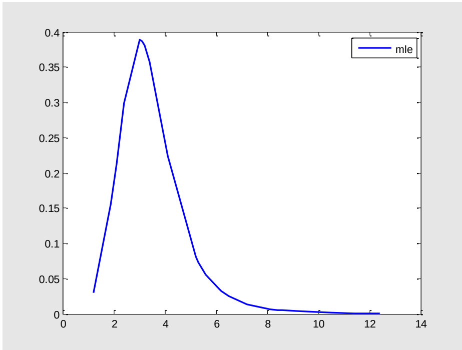

Figure (5) represents the behavior of of estimating the probability densityfunction

- We notice from Table No. (9) and Figure 4 that the survival function shows decreasing, and this corresponds to the theory functions.
- We notice from Table No. (9) and Figure 5 that the probability density function shows increasing and then gradually decreased .

6. Conclusions

i. The superiority of the Maximum likelihood method over the ordinary and weighted least squares method for the three models and at a sample size of n=50,100,250 in the process of estimating parameters , survival function and probability density function for log-logistic distribution as it had less MSE, MAPE, IMSE .

ii. The method of least squares and weighted gives weak estimates compared to the method of Maximum likelihood for all sample sizes and all experiments .

iii. From the conclusions of the practical side, it becomes clear that the survival function values are decreasing with the increase in the time of infection for a group of breast cancer patients under study, and this corresponds to the theoretical properties of this function as it is a monotonic decreasing function.

iv. The values of the probability distribution function show increase and then gradually decrease as the values of t increase.

7. Recommendations

i. The researcher recommends the adoption of the Maximum likelihood method in estimating the survival function for the two-parameters log-logistic distribution of the complete data.

ii. The researcher recommends the use of log-logistic regression in estimating the survival function of living organisms by using different estimation methods.

iii. The researcher recommends necessity of accrediting the Ministry of Health the estimators of the survival function and the probability density function in the simulation aspect of patients' breast cancer, and to benefit in the development of future treatment plans and follow of patients.

iv. The researcher recommends using other methods to estimate the parameters and survival function of the log-logistic distribution, such as the Mixed method and the elimination method for TOM.

References

1- Abd-Elmonem A.M. Teamah , et.al .(2021) "Heavy-Tailed Log-Logistic Distribution: Properties, Risk Measures and Applications" , *Department of Mathematics, Faculty of Science, Tanta University, Tanta 31527, Egypt*

2- Al-Majalli, Ahmed Ajel (2020) "estimation of the survival function for a compound probability distribution (Weibull-Rayleigh) with practical application" a master's thesis in statistics submitted to the College of Administration and Economics at the Holy University of Karbala .

3- Al-Saadi, Bashir Faisal (2010) "Some methods parametric and non-parametric for estimating the reliability function with practical application" a master's thesis in statistics submitted to the College of Administration and Economics at the University Of Baghdad.

4- Aryal, G. R. (2013). "Transmuted Log-Logistic Distribution", International journal of Statistics Applications and Probability, 1(2), 11-20.

5- Dixit, Asha (2008), "Exact Comparison of Hazard Rate Functions of Log-Logistic Survival Distributions" A Thesis Submitted to the Graduate Faculty of Auburn University

6-Fitzgerald, D. L. (2005). "Extreme Rainfall Using the Log-Logistic Distribution", Stochastic Environmental Research and Risk Assessment, 19, 249-257.

7-Kumar.D., Jain, N., and Abo-Kasem, O. (2019) "Parameter Estimation for the Exponentiated Kumaraswamy-Power Function Distribution Based on Order Statistics with Application " Journal of Statistical Research 16-19.

8-Ragab, A. and Green J. (1984). "On order statistics from the log-logistic distribution and their properties", Communications in Statistics-Theory and Methods, 13, 2713{2724.

9- Singh, V. P,and Guo, H. (1995). "Parameter Estimation For 2-parameter Log-logistic Distribution by Maximum Entropy", Civil Engineering System, 12, 343-357.

Journal of Economics and Administrative Sciences Vol.28 (NO. 131) 2022, pp. 180-192

مقارنة بعض طرائق تقدير المعلمات ودالة البقاء لتوزيع log-logistic مع تطبيق عملي

الباحث/ سماح صباح حسن
Samahsabah1993@gmail.com

أ.د. انتصار عريبي فدعم
كلية الادارة والاقتصاد / جامعة بغداد
entsar_arebe@coadec.uobaghdad.edu.iq

Received: 4/10/2021 Accepted:2/11/2021 Published: March / 2022

مستخلص البحث:

يعد توزيع **Log-Logistic** من التوزيعات الاحصائية المهمة اذ يمكن تطبيقه في العديد من التجارب الحياتية والبايلوجية وغيرها من التجارب واهميته تأتي من اهمية تحديد دالة البقاء لتلك التجارب يتلخص البحث على اجراء مقارنة بين طريقة الامكان الاعظم المربعات الصغرى وطريقة المربعات الصغرى الموزونة لتقدير معالم ودالة البقاء لتوزيع **log-logistic** بأستعمال معايير المقارنة **MSE, MAPE, IMSE** وتم تطبيق هذا البحث على بيانات حقيقة لمرضى سرطان الثدي . وقد أظهرت النتائج ان طريقة الامكان الاعظم تفوقت بالافضلية في حالة تقدير معالم التوزيع ودالة البقاء .

المصطلحات الرئيسة في البحث: دالة البقاء , توزيع **log-logistic** , طريقة الامكان الاعظم , طريقة المربعات الصغرى , طريقة المربعات الموزونة .

*البحث مستل من رسالة ماجستير