Optimal Solutions of Parameters of the Modified Weibull Distribution using Least Squares and Maximum Likelihood Methods

Jamel Kenouza [a], Abderrahmane Belguerna [b]

[a,b] Assistant Professor of Mathematics, S.A University Center of Naama, Algeria.

Abstract: Parametric statistics are among the most important research that enables us to identify the advantages of the communities studied from the study of some of the samples taken from them. In this paper, we identify Weibwll's modified model with three parameters a. b. c, then we study the asymptotic properties of these parameters estimation, using the least squares and maximum likelihood methods by comparing estimated parameters with seized parameters by changing the size of the sample each time.

Keywords: Optimization, Least Squares, Maximum Likelihood, Modified Weibull Distribution, Numerical illustration.

1. Introduction

In parametric statistics, modelling, applied mathematics, computation and operations research the optimization theory and methods play a very important role, which has wide applications in sciences, engineering, management, technology and many others fields. The subject is involved in the optimal solution of problems that are mathematically presented. The optimal solution of a practical problem can be found from lots of schemes by means of scientific methods and tools. It involves the study of optimality conditions of the problems, the construction of model, the determination of algorithmic, the establishment of convergence of these algorithms, and numerical experiments with typical problems and real life problems. It did not become an independent subject until the late 1940s, when G.B. Dantzig presented the well-known simplex algorithm for linear programming. After the 1950s, when conjugate gradient methods and quasi-Newton methods were presented, the nonlinear programming developed greatly. Now various modern optimization methods can solve difficult and large scale optimization problems, and become an indispensable tool for solving problems in diverse fields. We suggest the mathematical form of the examples and the basic conditions for finding the optimal solution, whether smaller or larger. Second, we give the general format of the least squares method and the maximum likelihood method to find the three parameters for a modified weibull distribution by optimization. So we study and illustrate the properties of the three parameters estimate of this two methods. Finally, an applied and comparative study by extracting the right model for the modified Weibull distribution by stabilizing the solution, changing the sample size every time and from our existing model, taking the three parameters as optimal solutions, using the two methods and studying the rate of convergence.

The general form of optimization problems is:

$$\min_{x\in R^n} f(x) \qquad (1)$$

where $x \in X \subset R^n$ is a variable, f is a function, $X \subset R^n$

2. Least-Squares Problems

In this section we're talking about nonlinear least square :

$$\min_{x\in R^n} f(x) = \frac{1}{2} r(x)^T r(x) = \frac{1}{2}\sum_{i=1}^{m} [r_i(x)]^2 \ , m \geq n. \qquad (2)$$

Where $r: R^n \to R^n$ is a nonlinear function of x.

It can also be translated (2) as a non-linear wreck of equations as follows:

$$r_i(x) = 0 \ , i = 1, \dots, m \ . \qquad (3)$$

Where $r_i(x)$ is the residual function , while $m = n$

In this work we simulate the data that tracks the Modified Weibull distribution and then estimate the three parameters using this method.

We calculate $\hat{\theta}$ such that the function $F_{\hat{\theta}}(x_i)$ fits the data as well as possible, this conduct us to minimizing problem below:

$$\min_{\theta \in R^n} \sum_{i=1}^{m} [e_i(x)]^2 \qquad (4)$$

Where

$$e_i(x) = F_{\hat{\theta}}(x_i) - y_i \qquad i = 1, \dots, m$$

And $\sum_{i=1}^{m} [e_i(x)]^2$ is the sum of the squares of the error

We proposed use Newton's method to solve the nonlinear system with three unknown parameters and we give some effective and special methods for solving nonlinear least-squares problem with MWD.

The MWD distribution function (a, b, c) defined by:

$$F(x, \mathrm{a}{:}\,\mathrm{b}{:}\,c) = 1 - e^{-\mathrm{a}x - \beta x^c} \quad , x > 0 \qquad (5)$$

the Jacobean given by :

$$J(x) = \begin{pmatrix} \frac{\partial e_1}{\partial \alpha}(x) \frac{\partial e_1}{\partial \beta}(x) \frac{\partial e_1}{\partial \gamma}(x) \\ \frac{\partial e_2}{\partial \alpha}(x) \frac{\partial e_2}{\partial \beta}(x) \frac{\partial e_3}{\partial \gamma}(x) \\ \dots \quad \dots \quad \dots \\ \frac{\partial e_n}{\partial \alpha}(x) \frac{\partial e_n}{\partial \beta}(x) \frac{\partial e_n}{\partial \gamma}(x) \end{pmatrix} \qquad (6)$$

Then the gradient

$$G(x) = \sum_{i=1}^{m} (\nabla r_i(x) \nabla r_i(x)^T + r_i(x) \nabla^2 r_i(x)) = J(x)^T J(x) + S(x) \qquad (7)$$

Where

$$S(x) = \sum_{i=1}^{m} r_i(x) \nabla^2 r_i(x) \qquad (8)$$

Therefore, the quadratic model of the function f(x) is:

$$q^{(k)}(x) = f(x_k) + g(x_k)^T (x - x_k) + \frac{1}{2}(x - x_k)^T G(x)(x - x_k) \qquad (9)$$

$$= \frac{1}{2} r(x_k)^T r(x_k) + \left(J(x_k)^T r(x_k))^T (x - x_k) + \frac{1}{2}(x - x_k)^T (J(x_k)^T J(x_k) + S(x_k) \right) (x - x_k)$$

Then the Newton's method for (2) given by:

$$x_{k+1} = x_k - ((J(x_k)^T J(x_k) + S(x_k))^T J(x_k)^T r(x_k))^T \qquad (10)$$

Algorithm :

Step 0: *Given* $x_0 \geq 0$, $k := 0$.

Step 1: *If* $\|g_k\| \leq \varepsilon$, *stop.*

Step 2: solve

$J(x_k)^T J(x_k) s = -J(x_k)^T r(x_k)$ for s_k

Step 3: *Set* $x_{k+1} = x_k + s_k$, $k := k + 1$. *Go to Step 1.*

Theorem 01

Consider f: $R^n \to R$ and $f \in C^2$. We assume that $\dot{x}$ is the local minimized of the nonlinear least-squares problem (2) and $J(\dot{x})^T J(\dot{x})$ is positive definite. Also we assume that the sequence $\{x_k\}$ generated by Algorithm « I » converges to $\dot{x}$. Then, if $G(x)$ and $J(\dot{x})^T J(\dot{x})$ are Lipschitz continuous in the neighbourhood of $\dot{x}$, we have :

$$\|x_{k+1} - \dot{x}\| \leq \|J(\dot{x})^T J(\dot{x})\| \|S(\dot{x})\| \|x_k - \dot{x}\| O(\|x_k - \dot{x}\|) \qquad (11)$$

Theorem 02

Consider $f : D \subset R^n \to R$ and $f \in C^2(D)$, where D is an open convex set. Let $J(x)$ be Lipschitz continuous on D, i.e :

$$\|J(x) - J(y)\|_2 \leq \|x - y\|_2 \ \forall\, x, y \in R^2 \qquad (12)$$

Least Square Estimation:

Historically, the least-squares method was the first to be developed in the early 1800s by Adrien-Marie Legendre (1752–1833), a French mathematician, and Gauss (1777–1855), a German mathematician, as a *curve-fitting method* in the context of the *theory of errors*; see Stigler (1986) and Gorroochurn (2016).

let on consider $x_1, \ldots, x_n$ are the outputs of random variables $X_1, \ldots, X_n$ independent and same distribution F(x) We're using the least square method to find the parameters $(\hat{a}, \hat{b}, \hat{c})$ may be obtained by minimising the following quantity in relation to a, b, c:

$$\min_{\theta \in R^3} \sum_{i=1}^{m} \left[y_i - \hat{a}x_i - \hat{b}x_i^{\hat{c}}\right]^2 \quad where \ \ \theta = (\hat{a}, \hat{b}, \hat{c}) \quad (13)$$

And $y_i = -\log(z_i) =$

To obtain a0; b0; c0, we must solve the following non-linear system of:

$$\frac{\partial l}{\partial a} = \sum_{i=1}^{n} x_i y_i - a\sum_{i=1}^{n} x_i^2 - b\sum_{i=1}^{n} x_i^{c+1} = 0$$
$$\frac{\partial l}{\partial b} = \sum_{i=1}^{n} x_i^c y_i \log(x_i) - a\sum_{i=1}^{n} x_i^{c+1} - b\sum_{i=1}^{n} x_i^{2c} = 0 \qquad (14)$$
$$\frac{\partial l}{\partial c} = \sum_{i=1}^{n} x_i^c y_i \log(x_i) - a\sum_{i=1}^{n} x_i^{c+1}\log(x_i) - b\sum_{i=1}^{n} x_i^{2c}\log(x_i) = 0$$

In order to find the optimal solution for this system, we rely on the modified Newton numerical method, in which case we choose the primitive condition and identify the value of accuracy in order to choose convergence to the solution.

3. Modified Newton′s Method

The Newton–Raphson numerical method is an alternative search method. This method appears approximate the near optimal solution, more efficient and faster in term of convergence. It is an iterative root finding technique using the partial derivatives of the function as the new system of equations. The algorithm uses Cramer's direct method to find the solution of the system.

Newton's method for multivariable optimization searches is based on Newton's single variable algorithm for finding the roots and the Newton–Raphson method for finding roots of the first derivative, given a x_0, iterate $x_{n+1} = x_n - f'(x_n)/f''(x_n)$ until $|x_{n+1} - x_n|$ is less than some small tolerance.

We might also use Method 2 (see Meerschaert, 1993, p. 73). The algorithm is expanded to include partial derivatives with respect to each variable's dimension.

And in Miley, when you get a three-microscope non-linear equations system in both ways, and to find the optimal solution, we use the Newton-Raphson algorithm.

4. Model, Definitions and Assumptions:

let on consider $x_1, \ldots, x_n$ are the outputs of random variables $X_1, \ldots, X_n$ independent and same distribution

The MWD distribution function (a; b; c) defined by :

$$F(x, \alpha\!:\!\mathrm{b}\!:\!c) = 1 - e^{-\alpha x - \mathrm{b}x^c} \ , x > 0 \qquad (15)$$

Where c>0 a, b> 0 and a + b > 0, we have is a scale parameter, and a, b are form parameters.

after the MWD (a,b,c) the following distribution is obtained :

after the MWD(a,b,c) we have the following distribution

for c = 2 we have LFRD(a,b)

for a = 0 we have WD(a,b) distribution of weibull

for a = 0 and c = 2 we have RD(b)

for b = 0 we have ED(a) the exponential distribution.

the density function of MWD(a,b,c) :

$$f(x, \mathrm{a}\!:\!\mathrm{b}\!:\!c) = (\mathrm{a}x - \mathrm{bc}x^{c-1})e^{-\alpha x - \mathrm{b}x^c} \ , x > 0 \qquad (16)$$

the hazard function of MWD(a; b; c) :

$$h(x, \mathrm{a}\!:\!\mathrm{b}\!:\!c) = (\mathrm{a}x - \mathrm{bc}x^{c-1}) \qquad (17)$$

It is easy to verify from (17) that: the hazard function is constant when c = 1, when c < 1, the hazard function is decreasing, and (3) if c > 1 the increasing hazard.

The maximum likelihood method

let on consider $x_1,\dots,x_n$ are the outputs of random variables $X_1,\dots,X_n$ independent and same distribution,

the likelihood function *:*

for sample $x_1,\dots,x_n$, the function of the parameter θ :

$$l(\theta; x_1,\dots,x_n) = \prod_{i=1}^{n} P(X_i = x_i; \theta) \qquad (18)$$

When the likelihood function is considered as a function of θ dependent by observations $x_1,\dots,x_n$

Maximum-likelihood:

The maximum likelihood estimator (MVE) of is the corresponding random variable noted by θ_n can be defined :

$$\theta_n = \mathbf{arg}\ \underset{\theta}{\mathbf{max}}\, \mathbf{log} L(\theta; x_1,\dots,x_n) \qquad (\mathbf{19})$$

when $\theta = (\theta_1,\dots,\theta_d)$ and all of the following partial derivatives exist, θ_n is solution of the system of equations called likelihood equations

$$\frac{\partial}{\partial\theta_i} \mathbf{log} L(\theta; x_1,\dots,x_n) \ \ i \in \{1,..,d\}$$

the likelihood of MWD :

$$L(x, \mathrm{a}: \beta: c) = \prod_{i=1}^{n} (a - \beta c {x_i}^{c-1}) e^{-\alpha x_i - \beta {x_i}^{c}}$$

The log-likelihood function is :

$$l = \sum_{i=1}^{n} \log(\alpha x_i - \beta c {x_i}^{c-1}) - a\sum_{i=1}^{n} x_i - \sum_{i=1}^{n} x_i^{c-1}$$

Computing the first partial derivatives of l and setting the results equal zeros, we get the likelihood equations as in the following form :

$$\frac{\partial l}{\partial a} = \sum_{i=1}^{n} \frac{1}{a + bc{x_i}^{c-1}} - \sum_{i=1}^{n} x_i = 0$$
$$\frac{\partial l}{\partial b} = \sum_{i=1}^{n} \frac{c{x_i}^{c-1}}{a + bc{x_i}^{c-1}} - \sum_{i=1}^{n} {x_i}^{c} = 0 \qquad (20)$$
$$\frac{\partial l}{\partial c} = \sum_{i=1}^{n} \frac{x_i^{c-1}(c \log x_i)}{a + bc{x_i}^{c-1}} - \sum_{i=1}^{n} {x_i}^{c} \log x_i = 0$$

To find the solutions of the nonlinear system, we use the Newton method, which represents the discretionary solutions of the three parameters, we will mention the algorithm in what comes from this work.

Properties of maximum likelihood estimators:

To know the spin-offs of the estimates and to access the exact values, the information filter can be computed, the asymptotic distribution of the MLE is given by, Miller (1981),

$$\begin{pmatrix} \tilde{a} \\ \tilde{b} \\ \tilde{c} \end{pmatrix} \sim N \left[\begin{pmatrix} a \\ b \\ c \end{pmatrix}, \begin{pmatrix} V_{11} & V_{12} & V_{13} \\ V_{21} & V_{22} & V_{23} \\ V_{31} & V_{32} & V_{33} \end{pmatrix} \right] \qquad (21)$$

Then the observed information matrix is given by:

$$I = \begin{pmatrix} I_{11} & I_{12} & I_{13} \\ I_{21} & I_{22} & I_{23} \\ I_{31} & I_{32} & I_{33} \end{pmatrix}$$

so that the variance-covariance matrix may be approximated as :

$$V = \begin{pmatrix} V_{11} & V_{12} & V_{13} \\ V_{21} & V_{22} & V_{23} \\ V_{31} & V_{32} & V_{33} \end{pmatrix} = \begin{pmatrix} I_{11} & I_{12} & I_{13} \\ I_{21} & I_{22} & I_{23} \\ I_{31} & I_{32} & I_{33} \end{pmatrix}$$

We change $I_{ij} = \widehat{I_{ij}}$, when (a, b, c) replaces $(\hat{a}, \hat{b}, \hat{c})$ we get it

$$\hat{a} \pm z_{\theta/2}\sqrt{\widehat{V_{11}}} \qquad \hat{b} \pm z_{\theta/2}\sqrt{\widehat{V_{22}}} \qquad \hat{c} \pm z_{\theta/2}\sqrt{\widehat{V_{33}}},$$

where z_{θ} is the upper θ _the percentile of the standard normal distribution

θ_nConverges almost surely to θ when $n \to +\infty$

$$\sqrt{\boldsymbol{n}}(\theta_n - \theta) \to N(0,1)$$

$\theta_n$is asymptotically Gaussian, unbiased,

In this case, we use a numerical method that enables us to find the best solution: the Newton-Raphson method, in the end, we address an application.

5.Comparison between MV et LS

Application: Numerical study and simulation*:*

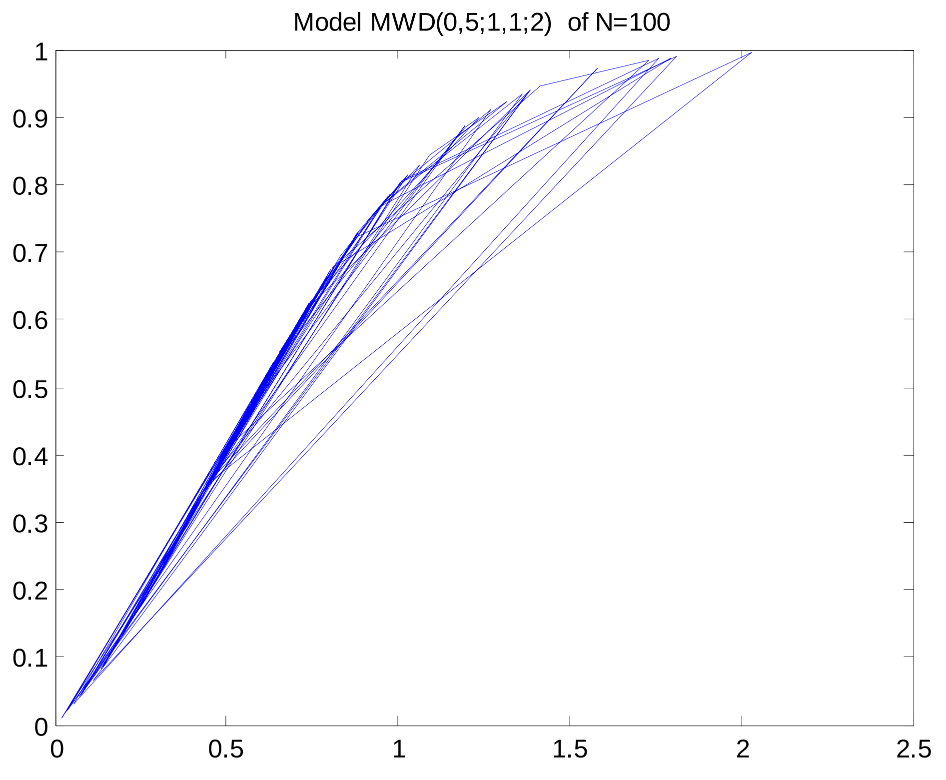
Simulation

Let's consider the distribution function to be modified so that we take the parameters(0.5,1.1,2) and extract the model (X, Y) to change each time the N=10, then N=25, and then N=50.

$$F(x, \text{a:b:}c) = 1 - e^{-0.5x - 1,1x^2} \;, x > 0$$

The following diagram shows the simulation of the modified Weibull distribution model by taking the optimal solution (0.5; 1.1 ;2). About 100 graphs.

Graph(01)



Estimation

Since in the modelling part we relied on optimal solutions (0.5; 1.1; 2). To test the effectiveness and properties of the two methods we take each time the length of the sample N=10, 25 and then 50 to study a convergence (a, b, c) to the optimal solution (0.5 ;1.1 ;2).

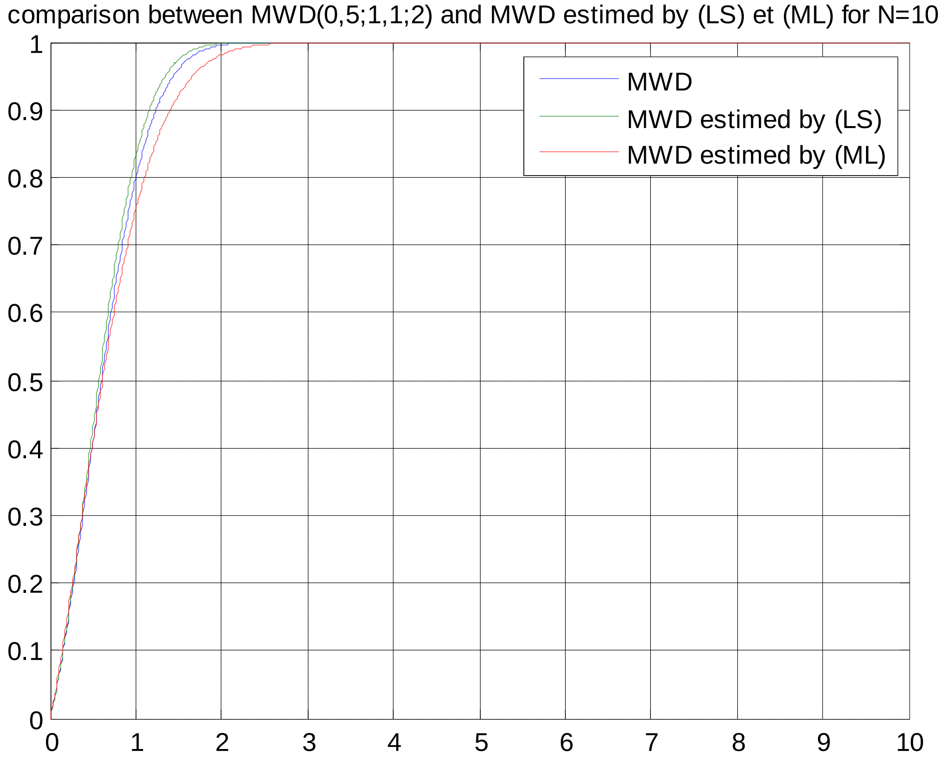
Table 1 below shows the average and standard deviation of estimated parameters from a sample size of N=10 with constant replication, using tow methods.

Table.1. average and standard deviation of estimated parameters by tow methods.

	the average of the parameters	**the variance of the parameters**
(LS)	(0.488 ;1,29 ;2,019)	(0.0068 ;0.032 ;0.699)
(ML)	(0.531 ;0.867 ;1.738)	(0.011 ;0.235 ;0.231)

From the first table we note that the average parameter of what sample size is equal to 10 approaches the optimal solution. The following diagram (2) shows the exact distribution function as compared to the function estimated using the minimization method. Also, graph (3) shows the comparison of errors in the tow methods, which shows that both take the normal distribution model as well as errors in the least squares method that are smaller than errors at the maximum likelihood , which explains that the better method in this case is the small squares. We also note that the standard deviation is approaching zero.

Graph(02)



Graph(03)

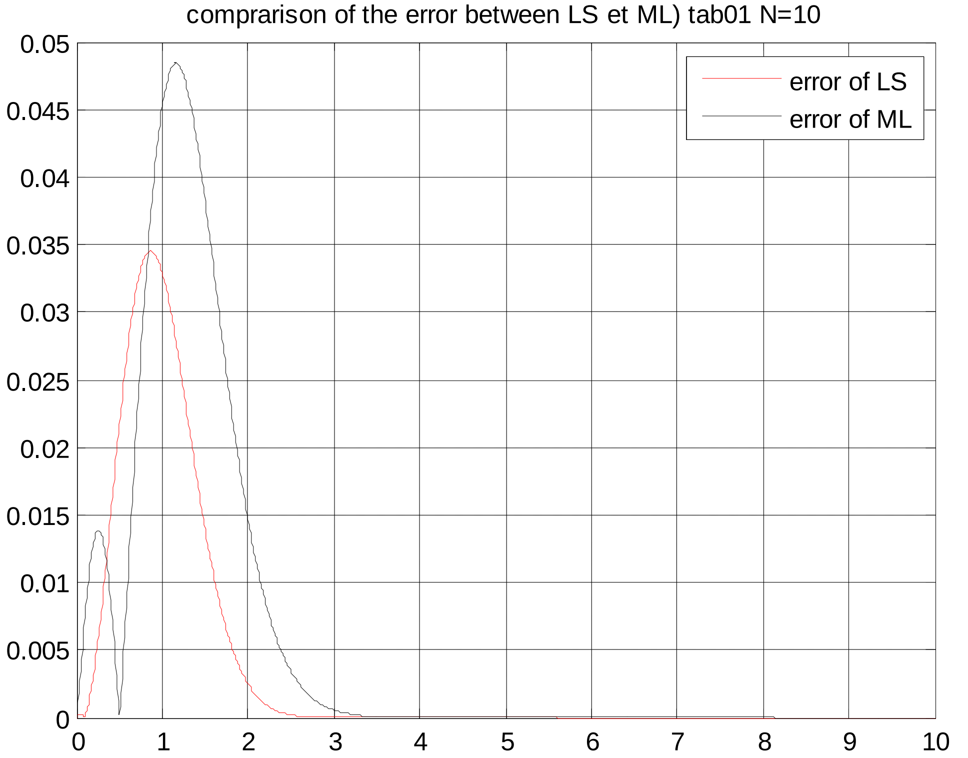
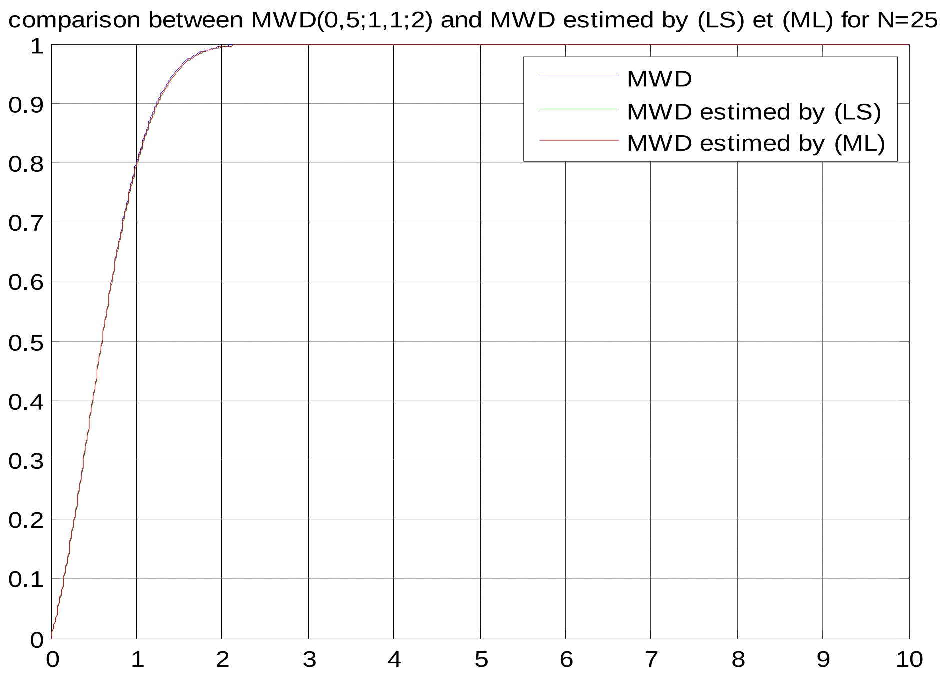


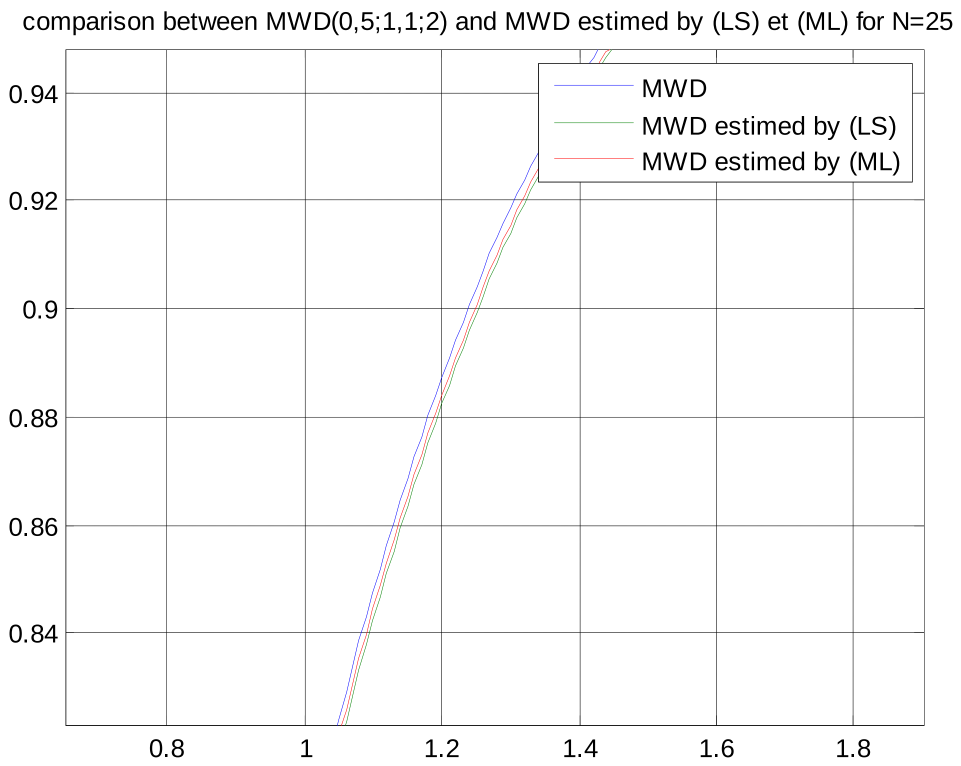
Table.2.

	the average of the parameters	the variance of the parameters
(LS)	(0.489 ;1,085 ;1,968)	(0.0031 ;0.0012 ;0.012)
(ML)	(0.509 ;1.077 ;1.977)	(0.004 ;0.024 ;0.052)

Graph(04)



Graph(05)


comparison between MWD(0,5;1,1;2) and MWD estimed by (LS) et (ML) for N=25
MWD
MWD estimed by (LS)
MWD estimed by (ML)
0.94
0.92
0.9
0.88
0.86
0.84
0.8
1
1.2
1.4
1.6
1.8

Graph(06)

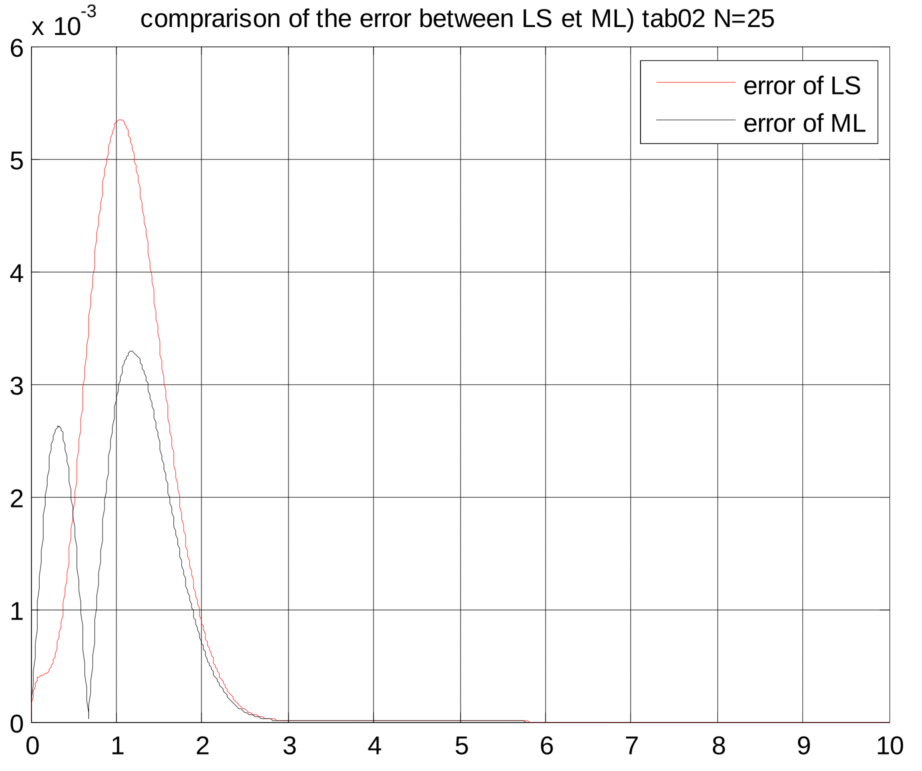
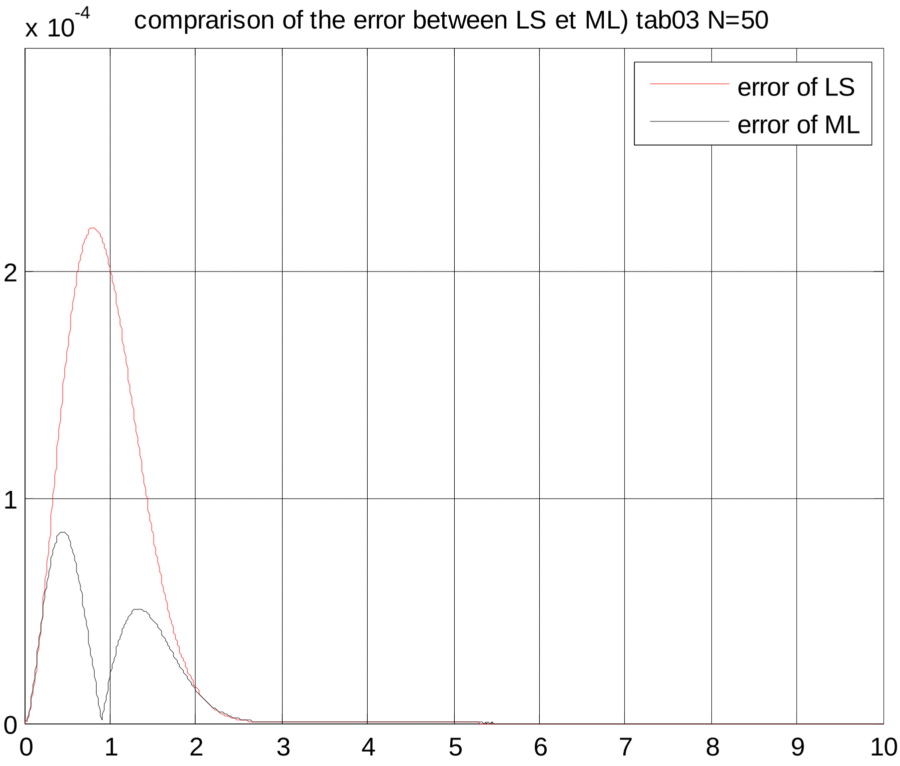

x 10^-3
comprarison of the error between LS et ML) tab02 N=25
error of LS
error of ML
6
5
4
3
2
1
0
0
1
2
3
4
5
6
7
8
9
10

Table. 3

	the average of the parameters	the variance of parameters
(LS)	(0.500081 ;1,09995 ;2,00015)	($2.15*10^{-8}$;$3.6*10^{-9}$;$6.18*10^{-8}$)
(ML)	(0.4999 ;1.10006 ;1.999)	($5.8*10^{-10}$;$3.8*10^{-9}$;$3.16*10^{-9}$)

Graph(07)



6.Comparison

Through the second and third tables and the funny graphics, so it's possible to observe Mille. The rate of parameters approaches the optimal solution, the larger the sample size as well as the standard deviation, the greater the sample size. For graphics the diagonal curve of the tow methods is close to the basic curve Errors take the normal distribution model and approach zero as the sample size increases And we can also say that a the maximum likelihood with N= 25 and 50 was the fastest approach to the optimal solution.

7.Conclusion

This study was designed to figure out the best digital means of reaching a solution, taking into account the basic conditions.

Based on the foregoing

The least squares method and the maximum likelihood method contribute greatly to finding optimal solutions to mathematical problems, making it easier for a lot of researchers in the field to estimate parameters, applied mathematics, or examples to figure out the way for them.

To solve the system of non-linear equations, the best way to reach the approximate solution is the iterative Newton-Raphson method with the primitive condition given.

To estimate Weibull's modified parameters, two special methods can be used to see what sample features follow for this distribution.

.

References (APA)

Ammar M. Sarhan and Mazen Zaindin (2009). Estimation for the parameters of the modified weibull distribution.. Internatinal journal of Basic et Applied Faculty of Science, 11, 131-135.

Antony, Waas, Joe Eisley (2011). Analysis of Structures_ An Introduction Including Numerical Methods -John Wiley & Sons .

Aris, Spanos (2019). Probability Theory And Statistical Inference_ Empirical Modeling With Observational Data-Cambridge University Press.

Aster, Richard, C. Borchers, Brian Thurber, Clifford H (2019). Parameter estimation and inverse problems-Elsevier.

Brezinski, C (2000). Numerical analysis Interpolation and extrapolation. Volume 2 , Springer Optimization and Its Applications.

Dimitrio,s ,C Kravvaritis_,Athanasios ,N ,Yannacopoulos (2019).Variational Methods in Nonlinear Analysis With Applications in Optimization and Partial Differential Equations-de Gruyter.

G. Muraleedharan, A.D. Rao a, P.G. Kurup, N. Unnikrishnan Nair, Mourani Sinha (2007). Modified Weibull distribution for maximum and significant waveheight simulation and prediction, International journal for coastal harbour and offshore engineers 54(8) ,631-635.

H. K. T. Ng (2005). Interval estimtion for the parameteres of the modied weibull distribution model with censored data a simulation study, Ieee Transactons on rellability, 54,374-376.

Horst Rinne (2009). The_Weibull_Distribution_A_Handbook(Bookos.org)_3 by Taylor & Francis Group,crc press.

M. Akram, A. Hayat (2014). Comparison of estimatore of the weibull Distribution . Financial Economentrics Series SWP (9-14).

Jean, Louis, Merrien (2007). Analyse numérique avec matlab. Dunod, Paris.

Wenyu, Sun, Ya-Xiang ,Yuan (2010). Optimization theory and methods_ Nonlinear programming-Springer (Springer Series in Computational Mathematics 35) Peter.

Deuflhard, A (2011). Newton Methods for Nonlinear Problems_ Affine Invariance and Adaptive Algorithms-Springer-Verlag Berlin Heidelberg.

William , P (2020). Nonlinear Optimization Models and Applications-CRC Press (Textbooks in Mathematics).

Steven, Chapra (2011). Applied Numerical Methods With MATLAB_ for Engineers & Scientists, 3rd Edition McGraw-Hill Science_Engineering_Math.