Suggested method of Estimation for the Two Parameters of Weibull Distribution Using Simulation Technique

Abdulkhaleq.A.ALNaqeeb[1*] and Alaa.M.Hamad[2]
[1]Department of Bio-Medical Statistics, College of Health Medical Technology, Baghdad, Iraq.
[2]Department of Mathematics, College of Education/Ibn Al-Haitham, University of Baghdad, Baghdad, Iraq.

Abstract

In this paper, suggested method as well as the conventional methods (probability plot-(p.p.) for estimations of the two-parameters (shape and scale) of the Weibull distribution had proposed and the estimators had been implemented for different sample sizes small, medium, and large of size 20, 50, and 100 respectively by simulation technique**.** The comparisons were carried out between different methods and sample sizes. It was observed from the results that suggested method which were performed for the first time (as far as we know), by using MSE indicator, the comparisons between the studied and suggested methods can be summarized through extremely asymptotic for indicator (MSE) results by generating random errors with W(1, 1) and with generating with **N(0,** σ^2**)** at the first contrast (1, 1).The suggested method was reported better performance through the sequentially status of ascending ordered for assumed initial contrast parameters, specially with generating of random error with W (1, 1) cooperation with generating of N (0,δ) in the simple linear regression equation**.**

Key words: Generalized Exponential, GE, and Inverse Prob. of cum. sampling dist.

طريقة مقترحة لتقدير معلمتي توزيع ويبل باستخدام تقنية المحاكاة

عبد الخالق النقيب*[1]، الآء ماجد حمد[2]
[1]قسم البصريات، الكلية التقنية الطبية، بغداد، العراق
[2]قسم الرياضيات ،كلية التربية /ابن الهيثم ،جامعة بغداد ،بغداد،العراق

الخلاصة

في هذا البحث تم اقتراح صيغة جديدة بالإضافة إلى الصيغة التقليدية لغرض تقدير معلمتي الشكل و القياس لتوزيع ويبل حيث تم اختيار طريقة الرسم البياني ولأحجام وعينات مختلفة صغيرة، متوسطة، كبيرة (٢٠و٥٠و١٠٠) مع عدة توليفات افتراضية للمعلمتين باستخدام تقنية المحاكاة.

وان عملية المقارنة بين الطرق المختلفة باستخدام مؤشر معدل مربعات الخطأ كمؤشر لأفضل أداء ومقارنة بين الصيغة التقليدية والمقترحة.

لقد لوحظ من خلال النتائج التي تم الحصول عليها بأفضلية الصيغة المقترحة (وتم استخدامها لأول مرة على حد علمنا) على الصيغة المدروسة التقليدية وخصوصا عند توليد الأخطاء العشوائية في الانحدار الخطي البسيط.

*Email:Abdulkhaliq.alnaqeeb@yahoo.com

Introduction

The Weibull distribution is named for Professor Waloddi Weibull whose papers led to the wide use of the distribution. He demonstrated that the Weibull distribution fit many different datasets and gave good results, even for small samples. The Weibull distribution has found wide use in industrial fields where it is used to model time to failure data [1]. As well as, Weibull distribution is one of the extreme-value distributions and plays an important role in reliability and maintainability analysis [2]. It was introduced by Bailey and Dell (1973) as model for diameter distributions, has been applied extensively in forestry due to its ability to describe a wide range of unimodal distributions including reversed-J shaped, exponential, and normal frequency distributions, the relative simplicity of parameter estimation, and its closed cumulative density functional form (e.g. Bailey, Dell 1973; Schreuder, Swank1974; Schreuder et al. 1979; Little 1983; Rennollset al. 1985; Mabvurira et al. 2002), and (4) its previous success in describing diameter frequency distributions within boreal stand types (e.g.Bailey, Dell 1973; Little 1983; Kilkki et al. 1989;Liu et al. 2004; Newton et al. 2004, 2005) [3] .

The three parameters Weibull distribution is indexed by a shape (a) , a scale (b), and a threshold (d) parameter. Using these symbols, the three parameter Weibull density function may be written as:

$$f\ (t,a,b)=\frac{a}{b}(\frac{t-d}{b})^{a-1}e^{-(t-d\ /b)^{a}}$$

$a>0,\ b>0, -\infty<d<\infty$

The symbol (t) represents the random variable (usually elapsed time). The threshold parameter (d) represents the minimum value of t that can occur. Setting the threshold to zero results in the common, two parameters Weibull distribution.

Shape Parameter – (a)

The shape (or power) parameter controls the overall shape of the density function. Typically, this value ranges between 0.5 and 8.0. The estimated standard errors and confidence limits displayed by the program are only valid when a > 2.0.

One of the reasons for the popularity of the Weibull distribution is that it includes other useful distributions as special cases or close approximations. For example, if

a = 1 The Weibull distribution is identical to the exponential distribution.

a = 2 The Weibull distribution is identical to the Rayleigh distribution.

a = 2.5 The Weibull distribution approximates the lognormal distribution.

a=3.6 The Weibull distribution approximates the normal distribution.

Cale Parameter – (b)

The scale parameter only changes the scale of the density function along the time axis. Hence, a change in this parameter has the same effect on the distribution as a change in the scale of time - for example, from days to months or from hours to days. However, it does not change the actual shape of the distribution.

(b) Is known as the characteristic life. No matter what the shape, 63.2% of the population fails by t = b + d.

Some authors use (1/b) instead of (b) as the scale parameter. Although this is arbitrary, we prefer dividing by the scale parameter since that is how you usually scale a set of numbers. For example, remember how you create a z-score when dealing with the normal data or create a percentage by dividing by the maximum.

The reliability function is one minus the cumulative distribution function. That is,

$F(t)=1-R(t)$

Cumulative distribution function C.D.F, F (t). This is the probability that an individual survives until time t and the Survival function, S (t) or Reliability function, R (t). This is the probability that an individual survives beyond time t. This is usually the first quantity that is studied.

Methods of estimating parameters:

1.Probability Plotting Method (P.P.E) – (The Conventional Methods):

This procedure uses the data from probability plot to estimate the parameters. The formula of estimation parameter depends on which option was selected for the probability plot model.

Probability plot model:

$F=A+B$ (time) ---------- (1)

The Cumulative Distribution Function (C.D.F) **:** F (t) is **:**

$F(t) = 1 - e^{-(\frac{t-d}{b})^a}$; (may be rearranged as (assuming D is zero))

$$F(t) = 1 - e^{-(\frac{t}{b})^a}$$

$\ln[-\ln(1-F(t))] = -a\ln(b) + a\ln(t)$ ----(2)

This is now in linear form. If we let:

$\ln[-\ln(1-F(t))] = y$ **&** $\ln(t) = x$ **;** the above equastion become:

$\hat{y} = -a[\ln(\hat{b})] + \hat{a}x$ ----(3)

Using simple linear reg ression , we can estimater . the intercept and slope . Using estimater,we obtan estimater the weibull parameter a and b as follows **:**

$$\hat{a} = slope$$

$$\hat{b} = (\frac{-\text{int}ercept}{a})$$

Now, probability plot model can be written as in equation (2) above**.**

$\frac{1}{a}\ln[-\ln(1-F(t))] + \ln(b) = \ln(t)$ -----(4)

This now in a linear form**.** If we let **:**

$x = \ln[-\ln(1-F(t))]$

$\hat{y} = \ln(t)$ **;** the above equation (3) become**:**

$\hat{y} = \frac{\hat{1}}{a}x + \ln(\hat{b})$ ----(5)

Using simple linear reg ression , we can estimater . the intercept and slope.

Using estimater,we obtan estimater of weibull band a as:

$$\hat{a} = \frac{1}{slope}$$

And

$$\hat{b} = \exp(\text{int}ercept)$$

Now**,** in adding the random error (e_i) to the eq**.** (3) and eq**.** (5)**,** then **;**

$y = -a[\ln(\hat{b})] + \hat{a}x + e_i$ and $y = \ln(\hat{b}) + \frac{\hat{1}}{a}x + e_i$

respectively**.**

A

nd was suggested that: e_i ~ W (1, 1) as well as e_i ~N$(0, \sigma^2)$**;** and the random error is independed and uneorrelated would be**.**

2-Simulation Probability Plotting Method (P.P.E) – (The Suggested Method) :

The following transformation has been employed:

Probability plot model:

$$F = A + B$$

The Cumulative Distribution Function (C.D.F): F (t) is:

$F(t) = 1 - e^{-(\frac{t-d}{b})^a}$;(may be rearranged as (assuming D is zero))

$$F(t) = 1 - e^{-(\frac{t}{b})^a}$$

$$1 - F(t) = e^{-(\frac{t}{b})^a}$$

$\ln(1-F(t)) = -(\frac{t}{b})^a$ ----- (6)

Using uniform distribution and generating

Uwhere; $U = \begin{cases} 1 & t \in [0,1] \\ 0 & otherwise \end{cases}$

Since U= 1-U ; $F(t) = 1 - F(t)$ ----- (7)

That is in the case of generating continuos uniform random variable,

Taking the natural logarithim,for the both side of Eq.(7)then the following equation will produce:

$$\ln[1 - F(t)] = \ln F(t)$$

$\ln F(t) = -\left(\frac{t}{b}\right)^a$ ------ (8)

Taking the natural logarithim of the two sided for equation(8) :

$$\ln[\ln[F(t)]] = -b\ln(\frac{t}{a})$$

$\ln[\ln F(t)] = a\ln(b) - a\ln(t)$ -----(9)

This is now in linear form. If we let:

$\ln[\ln F(t)] = y$, $a\ln(b) = \hat{\beta}_\phi$ (i.e. the intercept) , $\ln(t) = x$,

and $-a = \hat{B}_1$ (i.e. the slope)

Hence equation(9) will produce linear from:

$\hat{y}_i = \hat{B}_\phi + \hat{B}_1 x_i$ ------(10)

Now, in adding the random error (e_i) to the eq.(10),then;

$y_i = \hat{y}_i + e_i$ -----(11)

and was suggested that e_i ~ Weibull (1,1) not N(0, σ^2) and error is independed and uncorredated**.**

Empirical work:

First step: **-** specified the assumed values by choosing different sample size of Weibull distribution, such as small sample size (n=20) and medium sample size (n=50) and large sample size (n=100).And then choosing the values of assumption parameters (*a*, *b*) in each several contrasts and choosing for the initial values of the two parameters (shape and scale) parameters as shown in table (1).

Second step: - Generation of data which includes**:**

-Generated the random data which was taken from the uniform distribution in the interval (0,1).

-The generation of errors for all data and in suggested method the random errors have been generated using the standard Weibull distribution of (1, 1) and that introduced by the researchers as well as of normal distribution with $(0,\sigma^2)$ which has been used in conventional methods and that introduced by Gupta and Kundu.

Third step: **-** This step contains the following:-

Using the same value of $\hat{t}$ for three methods and applying the equation

$y_i = \hat{y}_i + e_i$ as mentioned in (11).

The values of (a) and (b) of the Weibull distribution can be determined according to the estimator of each method.

- Fourth step: **-** smoothing the obtained values

-In this step the iteration of data will be repeated 100 times to generate a new different error.

Fifth step:-In this step the following comparison indicator will be employed to make a compare between different methods by Mean Square Error *(MSE).*

Results and Conclusion

As a consequence for empirical work and taking the mean square error as the indicator of preference between the different estimator methods, the following results are obtained in the table 2-.

From table 2- as following obtained:-

1-For the conventional methods and for different sample sizes the following results are obtained:

(i) Small sample size (n=20)

a. For the assumed contrast parameters, the Conventional method-1 were given the best results at the 1st contrast (1, 1) and the suggested generating random errors with W(1, 1) is extremely asymptotic with generating with N(0, σ^2) **.**
b. For the assumed contrast parameters, the Conventional method-2 were given the best results at the 1st contrast (1, 1) and the suggested generating random errors with N(0, σ^2) is extremely asymptotic with generating with W(1, 1)**,** as well as the results of the last contrast (2, 2) were given the best results at the suggested generating random errors with W(1, 1) than with generating with **N(0,** σ^2 **) .**
c. For the assumed contrast parameters, the Suggested method were given the best results at the 1st contrast (1, 1) and the suggested generating random errors with W(1, 1) is extremely asymptotic with generating with N(0, σ^2) **.**

(ii) Medium sample size (n=50).

a. For the assumed contrast parameters, the Conventional method-1 were given the best results at the 1st contrast (1, 1) and the suggested generating random errors with W(1, 1) is extremely asymptotic with generating with N(0, σ^2) **.**
b. For the assumed contrast parameters, the Conventional method-2 were given the best results at the 1st contrast (1, 1) and the suggested generating random errors with N(0, σ^2) is extremely asymptotic with generating with W(1, 1)**,** as well as the results of the last contrast (2, 2) were given the best results at the suggested generating random errors with W(1, 1) than with generating with N(0, σ^2) **.**
c. For the assumed contrast parameters, the Suggested method were given the best results at the 1st contrast (1, 1) and the suggested generating random errors with W(1, 1) is extremely asymptotic with generating with **N(0,** σ^2 **) .**

(iii) Large sample size (n=100).

a. For the assumed contrast parameters, the Conventional method-1 were given the best results at the 1st contrast (1, 1) and the suggested generating random errors with

W(1, 1) is extremely asymptotic with generating with N(0, σ^2) **.**

b. For the assumed contrast parameters, the Conventional method-2 were given the best results at the 1st contrast (1, 1) and the suggested generating random errors with N(0, σ^2) is extremely asymptotic with generating with W(1, 1)**,** as well as the results of the last contrast (2, 2) were given the best results at the suggested generating random errors with W(1, 1) than with generating with **N(0,** σ^2 **) .**

c. For the assumed contrast parameters, the Suggested method were given the best results at the 1st contrast (1,1) and the suggested generating random errors with W(1,1) is extremely asymptotic with generating with **N(0,** σ^2 **) .**

The comparisons between the studied and suggested methods can be summarized through extremely asymptotic for indicator (MSE) results by generating random errors with W(1,1) and with generating with **N(0,** σ^2**)** at the first contrast (1, 1), and the suggested method were reported better performance through the sequentially status of ascending ordered for assumed initial contrast parameters, specially with generating of random error with W (1, 1) cooperation with generating of N (0,δ) in the simple linear regression equation**.**

Table 1- Assumed of the initial contrast parameters.

a	*b*
1	**1**
1	**2**
2	**1**
2	**2**

Table 2- Estimation of shape and scale parameters of weibull distribution for the conventional and suggested methods (the results of simulation technique)

Sample's sizes	Assumed contrasts		Generating the random error					
			W(1, 1)			N(0, σ^2)		
	a	b	Con. 1(*)	Con. 2(*)	Suggested	Con. 1(*)	Con. 2(*)	Suggested
20	1	1	1.6114214	1.2953770	1.5016968	1.5643693	0.899486	1.3052860
	1	2	5.1330121	5.9877661	4.8690923	4.7630717	4.7247900	4.9972023
	2	1	3.6056050	3.5543393	3.6345830	4.4272407	5.5242014	4.3505600
	2	2	8.1745000	1.8451684	3.9837000	17.629800	4.8885127	9.0761000
50	1	1	2.0107283	2.5863732	2.1510279	2.1311468	2.7293860	2.0391324
	1	2	4.6763127	5.2624995	4.0715202	4.7626475	3.3936200	4.8935799
	2	1	5.0943621	5.6724677	4.6320633	5.8903455	5.3630736	5.7602753
	2	2	4.6419503	1.2302989	4.6458346	7.7241503	4.3226739	7.7652185
100	1	1	1.4970435	1.9771494	1.5162744	1.5419518	2.0249568	1.5591135
	1	2	5.1507410	6.4251960	4.1203160	5.0345676	1.4043065	5.0674327
	2	1	4.8215457	5.6719330	4.8387642	5.5528370	5.3331900	5.6612127
	2	2	5.1648266	1.3226713	5.1689143	8.4141269	4.5338045	8.3070662

(*) **Con. 1** : Conventional method-1 ; **Con. 2** : Conventional method-2

Reference

1. Bartkut,V. and Sakalauskas,L.**2008**.The method of three-parameter Weibull distribution estimation.ACTA Journal of ET Commendations Universitatis Tartuensis De Mathematica,12.pp:65-78.

2.Lei,Y.,**2008**.Evaluation of three methods for estimating the weibull distribution parameters of chinese pine(pinus tabulaeformis)Journal Of Forest Science, *54*(12),pp:566–571.

3.Johnson,N. Kotz,S. and Balakrishnan,N. **1995**.Continuous Univariate Distributions, Wiley series in Probability the New York ,2