أ. د.عبد الخالق عبد الجبار النقيب م. د. حازم منصور كوركيس الاء ماجد حمد
 كلية التقنيات الطبية والصحية كلية التربية- ابن الهيثم كلية التربية- ابن الهيثم

الملخص

في هذا البحث تم مقارنة مقدرات طرائق الأمكان الأعظم، الأمكان الأعظم المئوية وطريقة المربعات الصغرى الاعتيادية لكل من معلمتي القياس والشكل حيث كان الأنموذج الأحصائى هو الأسي وذلك بتوظيف أسلوب (Monte Carlo) في المحاكاة ولحجوم عينات مختلفة وبافتراض العديد من التقابلات عند القيم الأبتدائية للمعلمتين وقد تم أعتماد مؤشرين احصائيين من أجل المقارنة بين أفضلية المقدرات وهما متوسط مربعات الخطأ (MSE) ومتوسط النسبة المنوية للخطأ (MPE) ولقد تم التوصل في هذا ألبحث الى أن طريقة PCE هي الأفضل ولجميع حجوم العينات.

Abstract

The main aim of this paper is to study how the different estimators of the two unknown parameters (shape and scale parameter) of a generalized exponential distribution behave for different sample sizes and for different parameter values. In particular,

. Maximum Likelihood, Percentile and Ordinary Least Square estimators had been implemented for different sample sizes (small, medium, and large) and assumed several contrasts initial values for the two parameters. Two indicators of performance Mean Square Error and Mean Percentile Error were used and the comparisons were carried out between different methods of estimation by using monte carlo simulation technique .. It was observed from the results that the PCE method had a better performance than the other methods for different sample sizes.



1. Introduction

During the last century, vast activities have been observed in generalizing of the distributions. These distributions were formulated by statisticians, mathematicians, and engineers to mathematically model or represent certain behavior. Some of these distributions tend to better represent life data and are most commonly lifetime distributions such as exponential distribution.

In the recent years a new distribution, named as Generalized Exponential (GE) distribution has been introduced by Gupta and Kundu .They were considered a special case of the exponentiated Weibull model assuming the location parameter to be zero, and compared its performances with the two-parameter gamma family and the two-parameter Weibull family, mainly through data analysis and computer simulations [1]. It was observed that many properties of this new family are quite similar to those of a Weibull or a gamma family; therefore this distribution can be used quite effectively to analyze lifetime data in place of gamma, Weibull distributions [2]. Therefore all the three distributions, namely generalized exponential, Weibull and Gamma are all extensions / generalizations of the one-parameter exponential distribution in different ways [3].

(1)

Different methods of estimation of the parameters of the generalized exponential distribution are proposed to realize the best results for the estimation of the two parameters using different sample sizes.

In fact, there are certain references that are devoted exclusively to different types of statistical distributions and several papers have already appeared on the estimation of the unknown parameter/parameters of a GE distribution; behave for different sample sizes and for different parameter values [4], [5], [6], [7], [8], [9], and [10].

Now a day all the scientific calculators or computers have standard uniform random number generator, therefore, generalized exponential random deviates can be easily generated from a standard uniform random number generator [6].

The genesis of the model is provided by Gompertz and Verhulst during the first half of the nineteenth century which was used one of the cumulative distribution functions to compare known human mortality tables and to represent population growth as

follows [6]:-
$$G(t) = (1 - \rho e^{-\lambda t})^{\alpha}$$
; For $t > \frac{1}{\lambda} \ln \rho$, ----- (1-1)

Where ρ , α , and λ represent location, shape, and scale parameters respectively, and all are positive real numbers.

Ahuja and Nash (1967) proposed the generalized Gompertz- Verhulst family of distributions and used this model and some related model for growth curve mortality [3]. Gupta and Kundu (1999) observed that the exponentiated exponential distribution is a special case of Gompertz-Verhulst distribution function (1-1) when $\rho = 1$ [6]. Therefore, x is a two parameter generalized exponential random variable if it has the distribution function as



The two-parameter generalized exponential distribution is a particular member of the three-parameter exponentiated Weibull distribution [6], and it can be introduced as shown in figure (1).



Figure (1) the Block Diagram of General Distributions. Source: [Introduced by the researcher]

(2)

The density function of the generalized exponential (GE) distribution is defined as the following:-

 $f(t) = \alpha \lambda (1 - e^{-\lambda t})^{\alpha - 1} e^{-\lambda t}$ (1-3)

The paper is organized as follows: Section two contains the estimation procedures of the unknown parameters of the generalized exponential distribution. Section three contains the empirical work and the results of simulation as well as the conclusions and future *work*.

2. Statistical Inference of Generalized Exponential distribution

An estimator is statistic that specifies how to use the sample data to estimate an unknown parameter of the population [11]. In the following sections three estimations procedure are considered, the maximum likelihood estimators, the percentiles estimators, and least squares estimators, and compare their performances through numerical simulation for different sample sizes and for different parameters values.

A Maximum Likelihood Estimation (MLE) represent a very general method of point estimation which is applicable whether the regularity conditions are or are not satisfied [12]. Consider estimation of α and λ when both are unknown; let x_1, \dots, x_n be

$$\frac{\partial L}{\partial \lambda} = \frac{n}{\lambda} + (\alpha - 1) \sum_{i=1}^{n} \frac{x_i e^{-\lambda x_i}}{(1 - e^{-\lambda x_i})} - \sum_{i=1}^{n} x_i = 0$$
 (2-3)

____22

From (2-2), we obtain the MLE of α as a function of λ ; say $\hat{\alpha}$ (λ), where

Putting $\alpha'(\lambda)$ in (2-1), we obtain:

 $\mathbf{g}(\boldsymbol{\lambda}) = C - n \ln \sum_{i=1}^{n} (-\ln(1 - e^{-\lambda x_i})) + n \ln(\lambda) - \sum_{i=1}^{n} \ln(1 - e^{-\lambda x_i}) - \lambda \sum_{i=1}^{n} x_i \quad \dots \quad (2-5)$ where the value of the constant C is given as: $\mathbf{C} = \mathbf{nln}(\mathbf{n}) - \mathbf{n}$

Therefore, *MLE* of λ , say λ_{MLE} , can be obtained by maximizing (3-5) with respect to λ.

it was observed by Gupta and Kundu [4] that g (λ) is a unimodal function and that λ_{MLE} which maximizes (2-5) can be obtained from the fixed point solution of

----- (2-6) $h(\lambda) = \lambda$

Where

$$h(\lambda) = \left[\frac{\sum_{i=1}^{n} ((x_i e^{-\lambda x_i}) / (1 - e^{-\lambda x_i}))}{\sum_{i=1}^{n} \ln(1 - e^{-\lambda x_i})} + \frac{1}{n} \sum_{i=1}^{n} \frac{x_i}{(1 - e^{-\lambda x_i})}\right]^{-1} \quad \dots (2-7)$$

(3)

An iterative procedure can be used to find a solution of (2-6) and it works very well.

Once we obtain α_{MLE} , the MLE of α say α_{MLE} can be obtained from (2-4) as

 $\alpha_{MLE} = \alpha(\lambda_{MLE}).$

In percentile methods the generalized exponential distribution has the explicit distribution function, therefore in this case the unknown parameters α and λ_{\perp} can be estimated by equating the sample percentile points with the population percentile points and it is known as the percentile method [6].

Among the most easily obtained estimators of the parameters of the Weibull distribution are the graphical approximation to the best linear unbiased estimators. It can be obtained by fitting a straight line to the theoretical points obtained from the distribution function and the sample percentile points. In case of a GE distribution also it is possible to use the same concept to obtain the estimators of α and λ based on the percentiles, because of the structure of its distribution function, when both the parameters are unknown[4].

Since $F(x,\alpha,\lambda) = (1 - e^{-\lambda x})^{\alpha}$ then:(2_8) $\mathbf{x} = -\frac{1}{2} \ln\{1 - [F(x, \alpha, \lambda)]^{\frac{1}{\alpha}}\}$

If *Pi* denotes some estimate of **F** ($\mathcal{X}_{(i)}$; α , λ) then the estimate of α and λ can be obtained by minimizing

Where $P_i = \frac{i}{n+1}$ represent the studied formula And $E(F(x_i)) = \frac{i}{n+1}$ the expected value $\Rightarrow E(x_i) = F^{-1}(\frac{i}{n+1})$

F(x) represents *c.d.f* for distribution, and $E(x_i)$ named (inverse probability of the cumulative sampling distribution) [13].

The suggested formula for Pi will be [14]

$$P_i = \frac{i - 0.5}{n + 0.5}$$

2.1 Algorithms of the Suggested Methods

The cumulative distribution function of the generalized exponential distribution can be written in the form:-

 $\hat{t} = -\frac{1}{\lambda} \ln[1 - \{\hat{F}(t)\}^{1/\alpha}]$ (2_11)

Using uniform distribution and generating U where

$$U = \begin{cases} 1 & t \in [0,1] \\ 0 & \text{otherwise} \end{cases}$$
(4)

Since U = 1 - U in-case of generating continues uniform random variable, then $\hat{F}(t)^{1/\alpha} = 1 - \hat{F}(t)^{1/\alpha}$ (2_12)

Taking the logarithm, for the both side of Eq (2-11), then the following equation will produce:

Ln(
$$\hat{t}$$
)= ln($\frac{1}{\alpha\lambda}$) + ln[-ln $\hat{F}(t)$]------ (2-13)
Thus the model (2-13) is linear , if we let

 $y_i = \ln(\hat{t}_i), x_i = \ln[-\ln\hat{F}(t_i)] \text{ and } \hat{\beta}_0 = \ln(\frac{1}{\alpha\lambda})$



In equation (2_13) the slope is constant and equal to 1 Using simple linear regression equation then $y_i = \beta_0 + \beta_1 x_i + e_i$ with $\beta_1 = 1$

Employing the initial value of $\alpha \& \lambda$ in the right side of Eq (2-13) with substitution of the generating uniform values in $\hat{F}(t_i) = u_i$, to obtain the left side

$$\hat{t} = \exp(\ln(\hat{t}))$$

and if error is added to this model, then $t_i = t_i + e_i$

Since that E(e) = 0, where $e \sim exp(1)$, so the errors are independent and uncorrelated [15]. The Maximum Likelihood Estimators (*MLEs*) of α and λ in equations (2-4) & (2-7), will be as follows

$$\hat{\alpha}(\lambda) = \frac{-n}{\sum_{i=1}^{n} \ln(1 - e^{-\lambda(t_i + e_i)})} - (2-14)$$

$$h(\lambda) = \left[\frac{\sum_{i=1}^{n} (x_i e^{-\lambda(t_i + e_i)}) / (1 - e^{-\lambda(t_i + e_i)}))}{\sum_{i=1}^{n} \ln(1 - e^{-\lambda(t_i + e_i)})} + \frac{1}{n} \sum_{i=1}^{n} \frac{(t_i + e_i)}{(1 - e^{-\lambda(t_i + e_i)})}\right]^{-1} - (2-15)$$

In order to make a comparison between the three estimators methods, (*MLE, PCE, and LSE*), the same procedure in finding (2-13) will be repeated twice time and the

equation of the straight line will produce. Through out solving them the value of α & $\hat{\lambda}$ will be founded.

Hence

 $\ln(\hat{t}_i) + e_i = \ln(t_i) = y_i \& x_i = \ln[-\ln(u_i)] \text{ and } \overline{\beta_0} = \ln(\frac{1}{\alpha\lambda}) = -\ln\alpha - \ln\lambda$ 2.2 Percentile Estimator (suggested)

$$t_{1}' = -\ln \alpha - \ln \lambda + \ln \left[-\ln \left(\frac{i}{n+1} \right) \right]$$
$$t_{2}' = -\ln \alpha - \ln \lambda + \ln \left[-\ln \left(\frac{i}{n+1} \right) \right]$$

$$t_1' + t_2' = -2\ln\alpha - 2\ln\lambda + 2\ln\left[-\ln\left(\frac{i}{n+1}\right)\right]$$
$$\ln\alpha = \ln\left[-\ln\left(\frac{i}{n+1}\right)\right] - \ln\lambda - \frac{1}{2}(t_1' + t_2')$$

_22



(5)

$$\hat{\alpha}(\lambda) = \exp\left\{\ln\left[-\ln\left(\frac{i}{n+1}\right)\right] - \ln\lambda - \frac{1}{2}(t_1 + t_2)\right\}$$
 ------ (2-16)
Where:

$$t_1 = t_1 + e_1$$
$$t_2 = t_2 + e_2$$

and the same method is used to estimate $\hat{\lambda}(\alpha)$ and obtain:

$$\hat{\lambda}(\alpha) = \exp\left\{\ln\left[-\ln\left(\frac{i}{n+1}\right)\right] - \ln\alpha - \frac{1}{2}(t_1 + t_2)\right\}$$
(2-17)

 $\alpha(\lambda)$ & $\lambda(\alpha)$ in suggested formula of *Pi* will be as follows

$$\hat{\alpha}(\lambda) = \exp\left\{\ln\left[-\ln\left(\frac{i-0.5}{n+0.5}\right)\right] - \ln\lambda - \frac{1}{2}(t_1+t_2)\right\}$$
------ (2-18)
$$\hat{\lambda}(\alpha) = \exp\left\{\ln\left[-\ln\left(\frac{i-0.5}{n+0.5}\right)\right] - \ln\alpha - \frac{1}{2}(t_1+t_2)\right\}$$
------ (2-19)

2.3Least Square Estimator (suggested)

$$\overline{t_1' = -\ln \alpha - \ln \lambda + \ln \left[-\ln \left(1 - \frac{i}{n+1} \right) \right]}$$

$$t_2' = -\ln \alpha - \ln \lambda + \ln \left[-\ln \left(1 - \frac{i}{n+1} \right) \right]$$

$$t_1' + t_2' = -2\ln \alpha - 2\ln \lambda + 2\ln \left[-\ln \left(1 - \frac{i}{n+1} \right) \right]$$

$$\ln \alpha = \ln \left[-\ln \left(1 - \frac{i}{n+1} \right) \right] - \ln \lambda - \frac{1}{2} (t_1' + t_2')$$

$$\hat{\alpha}(\lambda) = \exp \left\{ \ln \left[-\ln \left(1 - \frac{i}{n+1} \right) \right] - \ln \lambda - \frac{1}{2} (t_1 + t_2) \right\}$$
Where: $t_1 = \hat{t}_1 + e_1$ and $t_2 = \hat{t}_2 + e_2$

$$(2-20)$$

The same method is used to estimate
$$\hat{\lambda}(\alpha)$$
, hence
 $\hat{\lambda}(\alpha) = \exp\left\{\ln\left[-\ln\left(1-\frac{i}{n+1}\right)\right] - \ln\alpha - \frac{1}{2}(t_1+t_2)\right\}$ ------- (2-21)
 $\hat{\alpha}(\lambda) \& \hat{\lambda}(\alpha)$ in suggested formula of *Pi* will be
 $\hat{\alpha}(\lambda) = \exp\left\{\ln\left[-\ln\left(1-\frac{i-0.5}{n+0.5}\right)\right] - \ln\lambda - \frac{1}{2}(t_1+t_2)\right\}$ ------- (2-22)
 $\hat{\lambda}(\alpha) = \exp\left\{\ln\left[-\ln\left(1-\frac{i-0.5}{n+0.5}\right)\right] - \ln\alpha - \frac{1}{2}(t_1+t_2)\right\}$ ------- (2-23)
(6)
3. Simulation and Empirical work

One of the most important applications of computer science is <u>Computer simulation</u>. It's an enormous area used all over in the sciences, economics, finance, and several applications [16]. Simulation often involves phenomena that are too complex for analytic characterization [17]. This analysis may be done, sometimes, through analytical or numerical methods, but the model may be too complex to be dealt with. Essentially, simulation process consists of building a computer model that describes the behavior of a system and experimenting with this computer model to reach conclusions that support decisions [18]. In order to make comparison of the three estimation methods of the parameters of generalized exponential distribution (*MLE*, *PCE*, and *OLS*) to specify the best method we make a simulation prototype provide assumption of many cases which it can be existed in real world and use the basic step process in any simulation experiment once we have estimated the corresponding simulation model.

✤ <u>Algorithms steps</u>

(I) First step:- specifying different sample sizes from generalized exponential distribution, such as small sample size (n=20) and medium sample size (n=50) and large sample size (n=100). Then assuming initial values for the two parameters (α, λ) such as $(\alpha, \lambda) = (1, 1) (1, 2) (2, 1) (2, 2)$

(II) Second step:- Generation of data which include :-

- Generated the random data which was taken from the uniform distribution in the interval (0, 1) using Excel, and SPSS, software computer package.

- The generation of errors for all data and in suggested methods the random errors have been generated using the standard exponential distribution instead of normal distribution which has been used in conventional methods introduced by Gupta and Kundu. (III) Third step:- This step contains the following:-

-Using the same value of t for three methods and applying the equation $\hat{t} = -\frac{1}{2}\ln(1-U^{\frac{1}{\alpha}})$

- finding (t) by using the equation

$$t_i = t_i + e_i$$
 Where *i*=1, -----, *n*

- The values of $\alpha \& \hat{\lambda}$ of the generalized exponential distribution can be determined according to the estimation formulas given in section 2.

(IV) Fourth step:- smoothing the obtained values

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

different error, so we obtain 100 value of $\hat{\alpha}$, and 100 value of $\hat{\lambda}$ for each contrast. Then the mean of each case will be calculated to find the estimated α , and λ .

(V) Fifth step:- In this step the following comparison indicator will be employed to make a compare between different methods by (Mean Square Error(MSE), and Mean Percentile Error (*MPE*).which can be determined by using the formulas:

$$MSE = \frac{1}{n} \sum_{i=1}^{n} [\hat{F}(t_i) - F(t_i)]^2 \qquad \dots \dots (3_1)$$
$$MPE = \sum_{l=1}^{n} \frac{|t_i - \hat{t}_i|}{t_i} \qquad \dots \dots (3_2)$$

(7)

Results and Conclusion

As a consequence for practical work and taking the mean square error and mean percentile error as the indicators of preference between the different estimation methods, the following results are obtained:-

1_Comparing the performances of all the methods it is clear that as far as the minimum MSE and minimum MPE are concerned, PCE works the best in almost all the cases considered for estimating both α and λ followed by OLS.

2_It can be mentioned that when the sample size increased the mean square error decreased.

3_In this work the white noise error is generated in exponential distribution and it was followed the distribution and gave best results.

Table 1 through table 4 represent the results of simulation (estimating the scale and shape parameters) of the generalized exponential distribution for different methods and different sample sizes at $(\alpha, \lambda) = (1, 1) (1, 2) (2, 1) (2, 2)$ respectively.



Method	Sample size	â	â	MSE	MPE
	20	3.228871	0.907844	1.104497	69.6079
MLE	50	3.11685	0.783125	1.039894	69.10407
	100	2.746391	0.89621	1.067277	72.78716
	20	0.175853	0.175853	1.00582	78.32911
PCE	50	0.211527	0.211527	1.034397	75.51366
	100	0.249629	0.249629	1.022964	74.06739
	20	0.224637	0.224637	1.09849	70.42788
OLS	50	0.179287	0.179287	1.071279	75.5493
	100	0.215326	0.215326	1.041934	70.17599

Table (1) : Estimates For $(\alpha, \lambda) = (1, 1)$



(8)

method	sample	â	λ	MSE	MPE
	20	1.7691	0.994711	1.088926	72.8922
MLE	50	2.034755	1.064445	1.084715	72.18949
	100	2.196437	1.038664	1.031213	72.08663
	20	0.202638	0.346885	1.014435	73.62835
PCE	50	0.1339	0.26885	1.043127	73.853
	100	0.132157	0.264314	1.01573	73.02496
	20	0.141439	0.247487	1.128787	71.34058
OLS	50	0.143163	0.289006	1.017047	71.07631
	100	0.137776	0.276699	1.042774	70.59355

Table(2)Estimatesfor $(\alpha,\lambda) = (1,2)$

Table (3) : Estimates for $(\alpha, \lambda) = (2, 1)$

method	Sample size	â	â	MSE	MPE
	20	1.923868	1.059507	1.123285	74.17211
MLE	50	2.047879	1.013519	1.01845	72.09521
	100	1.846939	1.020239	1.054538	72.05456
	20	0.352757	0.174889	1.089119	76.88577
PCE	50	0.298843	0.146531	1.019789	74.42378
	100	0.308093	0.147665	1.020181	73.57866
	20	0.262607	0.158162	1.102657	75.92325
OLS	50	0.274431	0.137216	1.031817	72.53302
	100	0.313378	0.156689	1.021473	69.39181

Table (4) : Estimates for $(\alpha, \lambda)=(2, 2)$

method	Sample size	â	λ	MSE	MPE
	20	1.91203	1.063802	1.114514	69.10407
MLE	50	1.569443	1.091228	1.092651	74.93182
	100	1.545476	1.076076	1.033517	74.93182
	20	0.128318	0.128318	0.98849	73.14346
РСЕ	50	0.167614	0.167614	0.999185	74.17297
	100	0.173671	0.173671	1.043462	73.13837
	20	0.160134	0.160134	1.015512	73.0145
OLS	50	0.178732	0.178732	1.056235	73.83068
	100	0.174178	0.174178	1.032273	72.16212



(9) References

- [1] R.D.Gupta and D.Kundu, "Generalized exponential distributions", Australian and New Zealand Journal of Statistics, vol. 41, no. 2, 173-188, 1999.
- [2] R. D. Gupta and, D. Kundu," Exponential Family: An Alternative to Gamma and Weibull distributions", Biometrical Journal, vol 43, no.1, 117-130, 2001.
- [3] D.Kundu, "Exponentiated Exponential Distribution", available at www.google.com,2004.
- [4] R. D. Gupta and, D. Kundu, "Generalize Exponential Distribution: Different methods of estimation ", Journal of Statistical Computations and Simulations, vol. 69, no. 4, 315-338, 2001.
- [5] M. Z. Raqab, and M. Ahsanallah ," Estimation of the Location and Scale Parameters of Generalized Exponential Distribution Based on Order Statistics" ,Journal of Statistical Computations and Simulations, 2001.
- [6] R. D. Gupta and, D. Kundu, "Generalized Exponential Distribution: Existing results and some recent developments ", Journal of Statistical Planning and Inference, 2007.
- [7] R. D. Gupta and, D. Kundu, "Generalized Exponential Distribution:Statistics Inferences", Journal of Statistical Theory and Applications, vol.1 no. 2, 101-118, 2002
- [8] R. D. Gupta and, D. Kundu," Discriminating between Weibull and Generalized Exponential Distributions", Computational Statistics and Data Analysis, vol. 43, 179 - 196, 2003.
- [9] R. D. Gupta and, D. Kundu," Closeness of Gamma and Generalized Exponential Distribution", Communications in Statistics - Theory and Methods, vol. 32, no. 4, 705- 721, 2003le.com, 2005.
- [10] R. D. Gupta and, D. Kundu," Discriminating between Gamma and Generalized Exponential Distributions", Journal of Statistical Computation and Simulation, vol. 74, no. 2, 107-121, 2004.
- [11] T. Mcclave, "Probability and Statistics for Engineers", 3rd Edition, Pws-Kent publishing company, Boston, 1990.
- [12] Robert. V. Hogg, "Introduction to Mathematical Statistics", 3rd edition.
- [13] A. Al –Naqeeb," Estimation of the Residual Variance in Linear Models within Cases Normal Dist. And Non-Normal dist.", College of Administration& Economics-University of Al- Mustansiriyah, 1990.
- [14] A. Al –Naqeeb," Estimation some non-Linear Type by Using Min Ch square method for Binary Data by simulation method", Research Journal of Aleppo University, Basic Science Series, 2007.
- [15] A. Al –Naqeeb,"Measures The Coincidence Power of the Two Random Variables (Dependant And Residual) In The Linear Regression Models By Using The Probability Plot Method", 5 th international conference of, Iraq league of the statistical sciences,1993.
- [16] "Simulation ", from 118wiki .PDF available at http:// 22cll8.cs.uiowa.edu
- [17] D. McFadden & etal," Statistical Simulation", Econometrics, Laboratory University of California, available at www.google.com, 2001.
- [18] D. R. Insua and etal," Simulation in Industrial Statistics ", Technical Report, available at <u>www.google.com</u>