Discriminating Between Generalized Exponential Distribution and Some Life Test Models Based on Population Quantiles

B. Srinivasa Rao
R. V. R. and J. C. College of Engineering, Guntur, India, boyapatisrinu@yahoo.com

R. R. L Kantam
Acharya Nagarjuna University, Guntur, India, kantam.rrl@gmail.com

Discriminating Between Generalized Exponential Distribution and Some Life Test Models Based on Population Quantiles

B. Srinivasa Rao
R.V.R & J.C College of Engineering
Guntur, India

R. R. L. Kantam
Acharya Nagarjuna University
Guntur, India

A test statistic based on population quantiles using sample order statistics is suggested. The quantiles of the test statistics are evaluated for generalized exponential distribution. Similar test statistic based on moments of sample order statistic is referred and the proposed test formula is compared with it. Between the pairs of the above models it is established that the test formula proposed by us is more effective and useful than the formula based on the moments of order statistics as developed by Sultan (2007).

Keywords: Population quantiles, generalized exponential distribution

Introduction

The three-parameter generalized exponential (GE) distribution has its probability density function (pdf) as

$$f\left(x\right) = \frac{\alpha}{\sigma}\left(1 - e^{-\left(\frac{x-\mu}{\sigma}\right)}\right)^{\alpha-1} e^{-\left(\frac{x-\mu}{\sigma}\right)}, x > \mu, \mu > 0, \alpha > 0, \sigma > 0 \qquad (1)$$

Its cumulative distribution function (cdf) is given by

$$F\left(x\right) = \frac{\alpha}{\sigma}\left(1 - e^{-\left(\frac{x-\mu}{\sigma}\right)}\right)^{\alpha}, x > \mu, \mu > 0, \alpha > 0, \sigma > 0 \qquad (2)$$

The two-parameter GE distribution has its pdf as

Dr. Rao is an associate professor of statistics in the Department of Mathematics and Statistics. Email her at: boyapatisrinu@yahoo.com. R. R. L. Kantam is in the Department of Statistics. Email at: kantam.rrl@gmail.com.

$$f(x) = \frac{\alpha}{\sigma}\left(1 - e^{-\left(\frac{x}{\sigma}\right)}\right)^{\alpha-1} e^{-\left(\frac{x}{\sigma}\right)}, x > 0, \alpha > 0, \sigma > 0 \qquad (3)$$

Its cdf is given by

$$F(x) = \frac{\alpha}{\sigma}\left(1 - e^{-\left(\frac{x}{\sigma}\right)}\right)^{\alpha}, x > 0, \alpha > 0, \sigma > 0 \qquad (4)$$

The GE distribution was introduced by Gupta and Kundu (1999). It is compared with the two-parameter Gamma and Weibull distributions in Gutpa and Kundu (2001a). Different models of estimations are discussed in Gutpa and Kundu (2001b). Raqab and Ahsanullah (2001) and Raqab (2002) studied the properties of order and record statistics from the two-parameter GE distribution respectively. Discriminating between gamma and GE distribution were studied by Gutpa and Kundu (2004). Discriminating between lognormal and GE distribution was given in Kundu et al (2005). The expected values of order statistics may not always be available in numerical form nor analytically simple beyond a given sample size. However if the distribution function is invertible analytically the population quantile for any 'n' can be easily obtained. Also moment of order statistics are conceptually similar to the population quantiles with an admissible measure of closeness. Therefore, quantiles are used to develop the test statistic and to distinguish the GE distribution from other well-known life testing models. The proposed work is similar to that of Sultan (2007) wherein moments of order statistics are used to develop the test statistic with GE distributions null population. The aim of this article is to explore the usefulness of analytical expressions of population quantiles of GE distribution. In section 2 the GE distribution and its quantiles are developed. In section 3 the goodness of fit tests of the two-parameter and three parameter GE distribution are developed. Section 4 deals with the power of the proposed test procedure in comparison with that of Sultan (2007) with the same alternative populations. In section 5 the performance of quantiles of GE distribution is tested, and Section 6 contains concluding remarks.

The GE distribution and its quantiles

The pth quantile of population is defined as the solution of the equation $F(x) = p_i$ where F(x) is the cdf given in (1.4). This is also called the standard population

quantile. If $x_1, x_2, ... x_n$ is an ordered sample of size n and $p_i = \dfrac{i}{n+1}$ then the solution of $F(x) = p_i$ is defined as i^{th} population quantile corresponding to its order statistic x_i. In the sample this is denoted by d_i i.e., $F(d_i) = \dfrac{i}{n+1}$. Expected value of its order statistic in the sample is denoted by μ_i, the theory of order statistics indicate that μ_i, d_i can be approximated by each other. If the distribution function of the population is in a closed form, d_i's sometimes can be obtained more easily than α's , moments of order statistics. This possibility is explored in developing the proposed test statistics of this article. Given the form of $F(x)$ and a natural number n, d_i's can be obtained by inverting the population distribution function. This was done for the GE distribution with the shape parameter $\alpha = 0.5, 2.0$ and sample size $n = 10, 20, 25$. To make use of them in the proposed test statistic, the details are given in the following sections.

Goodness-of-Fit Test using quantiles

Test for two-parameter case   Let $x_1, x_2, ... x_n$ denote a sample from two-parametric GE distribution. The correlation type goodness of fit test procedure in this case using quantiles can be formed as follows:

$H_0 : F$ is correct, that is $x_1, x_2, ... x_n$ have $GE(0, \sigma, \alpha)$ given in (4) versus $H_1 : F$ is not correct, that is $x_1, x_2, ... x_n$ have another cdf and the test statistic used to run the test is given by

$$T_1 = \frac{\sum_{i=1}^n x_i d_i}{\sqrt{\sum_{i=1}^n x_i^2 \sum_{i=1}^n d_i^2}} \qquad (5)$$

The test statistic T_1 is simulated through Monte-Carlo method based on 10,000 simulations. Table 1 represents the percentages points of T_1 for sample sizes $n = 10, 20, 25$ and $\alpha = 0.5, 2.0$. It can be seen from the Table 1, the percentage points of T_1 follow the naturally expected order.

Table 1. Percentage points of T_1

α	n	0.50%	1%	2%	2.50%	5%	10%	20%	30%	40%	50%
0.5	10	0.9947	0.9931	0.991	0.9901	0.9865	0.9808	0.9708	0.9611	0.9499	0.9374
	20	0.9921	0.9903	0.9879	0.9865	0.9818	0.9747	0.963	0.9506	0.9375	0.9221
	25	0.991	0.989	0.986	0.9847	0.9796	0.9717	0.959	0.9462	0.9334	0.9189
2	10	0.9954	0.9944	0.9933	0.9928	0.9908	0.9877	0.9829	0.9782	0.9734	0.9674
	20	0.9944	0.9932	0.9918	0.9912	0.9891	0.9863	0.9818	0.9773	0.9723	0.9665
	25	0.9936	0.9925	0.9911	0.9906	0.9886	0.9856	0.9812	0.977	0.9722	0.9669

Test for the three-parameter case   Let $x_1, x_2,...x_n$ denote a sample from three-parametric GE distribution and let $Z_i = X_i - X_1$ and $v_i = d_i - d_i, i = 1, 2,....n-1$. The correlation type goodness of fit test in this case using quantiles can be formed as follows:

$H_0 : F$ is correct, that is $x_1, x_2,...x_n$ have $GE(0, \sigma, \alpha)$ given in (2) versus $H_1 : F$ is not correct, that is $x_1, x_2,...x_n$ have another cdf and the test statistic used to run the test is given by

$$T_2 = \frac{\sum_{i=1}^{n} z_i v_i}{\sqrt{\sum_{i=1}^{n} z_i^2 \sum_{i=1}^{n} v_i^2}} \tag{6}$$

The statistic T_2 is simulated through Monte-Carlo method based on 10,000 simulations. Table 2 represents the percentages points of T_2 for sample sizes $n = 10, 20, 25$ and $\alpha = 0.5, 2.0$. It can be seen from the Table 2 the percentage points of T_2 follow the naturally expected order.

Table 2. Percentage points of T_2

α	n	0.50%	1%	2%	2.50%	5%	10%	20%	30%	40%	50%
0.5	10	0.9951	0.9934	0.9914	0.9905	0.9869	0.9813	0.9712	0.9615	0.9505	0.9382
	20	0.9923	0.9904	0.998	0.9865	0.9818	0.9748	0.963	0.507	0.9376	0.9222
	25	0.9912	0.9891	0.9861	0.9847	0.9797	0.9718	0.959	0.9462	0.9334	0.919
2	10	0.9968	0.9958	0.9947	0.9943	0.9928	0.9902	0.9857	0.9812	0.9767	0.971
	20	0.9953	0.9942	0.9929	0.9924	0.9905	0.9879	0.9834	0.979	0.974	0.9682
	25	0.9945	0.9935	0.9922	0.9917	0.9898	0.987	0.9826	0.9785	0.9737	0.9683

Power of the test

The power of the test is calculated by replacing $GE(\mu,\sigma,\alpha)$ random variates generator in the simulation program with generators from the alternative distributions including: normal, lognormal, Cauchy, Weibull and gamma. Based on different sample sizes and 10,000 simulations, the power is calculated to be

$$Power = \frac{\# \text{ of rejections of } H_0}{10,000}$$

Where H_0 is rejected if $T_1(T_2)$ greater than or equal to the corresponding percentage points given in Table 1 (Table 2 and $T_1(T_2)$ is evaluated from the alternative distributions. Table 3 and 4 represent the power of the test for the two-parameter and three-parameter cases, respectively. The different alternative distributions considered are: (i) normal distribution N(μ, σ), (ii) lognormal Ln(μ, σ), (iii) Weibull distribution with location parameter μ, scale parameter σ and shape parameter α, W(μ, σ, α), (iv) gamma distribution with location parameter μ, scale parameter σ and shape parameter k G(μ, σ, k) and (v0 Cauchy distribution with location parameter μ and scale parameter σ C(μ, σ). Table 3 and 4 indicate that the correlation test has good power to reject sample from the chosen alternative distributions.

Table 3. Power of the test of the two-parameter case $(\sigma = 1)$

α	n	N(0,1)		W(0,1,3)		G(0,1,7)	
		5%	10%	5%	10%	5%	10%
0.5	10	0.9227	0.745	1	1	0.9991	0.9946
	20	0.9999	0.9985	1	1	0.9845	0.9306
	25	0.9989	0.9986	1	1	0.9959	0.9857
2	10	0.9971	0.9999	0.9991	0.9947	1	0.9996
	20	0.9997	0.9996	1	0.998	0.9972	0.9944
	25	0.9998	0.9997	1	1	0.9995	0.9986

Table 4. Power of the test of the two-parameter case $\left(\mu = 0, \sigma = 1\right)$

		LN(1,5)		W(0,1,6)		C(0,1)	
α	n	5%	10%	5%	10%	5%	10%
0.5	10	0.975	0.937	1	1	1	1
	20	0.975	0.943	1	1	1	1
	25	0.982	0.959	1	1	1	1
2	10	1	1	1	1	1	1
	20	1	1	1	1	1	1
	25	1	1	1	1	1	1

Tables similar to that of 3 and 4 are available in Sultan (2007), evaluated using the moments of order statistics. By comparison, notice that the coverage probability given in the tables are uniformly larger than what are given in Sultan (2007). Therefore, the test statistic proposed based on the population quantiles is more powerful than that based on the moments of order statistics. Moreover, for a distributional GE distribution moments of order statistics are not available completely beyond a given sample size whereas population quantiles are available for any sample size provided the mathematical form of the cdf is analytically invertible. Therefore it can be concluded that the proposed test statistic T is more powerful than that of Sultan (2007).

Numerical Examples

In order to show the performances of the test of GE distribution in both cases (two-parameter and three-parameter), four sets of order statistics each of size 25 were simulated, they are

1. Sample from GE(0,1,2): two-parameter case of the GE distribution with scale parameter is equal to 1 and shape parameter is equal to 2

2. Sample from GE(1,1,2): three-parameter case of GE distribution with location parameter is equal to 1, scale parameter is equal to 1 and shape parameter is equal to 2.

3. Sample from G(0,2,2): gamma distribution with location parameter is equal to 0, scale parameter is equal to 2 and shape parameter is equal to 2.

4. Sample from GE(2,2,2): gamma distribution with location parameter is equal to 2, scale parameter is equal to 2 and shape parameter is equal to 2.

The above four order statistics samples with the analogous quantiles of order statistics from GE(0,1,2) are used to run the test. The results of the test at 5% significance level and at α=2 (whether accept (A) or reject (R) H_0) are given for different values in the following table.

Table 5. Results at 5% significance, $\alpha = 2$

Decision			
GE(0,1,2)	**GE(1,1,2)**	**G(0,2,2)**	**G(2,2,2)**
A	A	R	R

Conclusions

This article proposed a test formula parallel to the one developed by Sultan (2007). It was found to be simple and can be used for any sample size. Moreover, it is more effective with respect to power evaluation and coverage probabilities.

References

Gupta, D. R., & Kundu, D. (1999). Generalized exponential distribution. *Australia and New Zealand Journal of Statistics, 41*(2), 173-188.

Gupta, D. R., & Kundu, D. (2001a). Exponentiated exponential family:an alternative to gamma and Weibull distributions. *Biometrica Journal, 43*, 117-130.

Gupta, D. R., & Kundu, D. (2001b). Generalized exponential distributions: different methods of estimation. *Journal of Statistical Computation and Simulation, 69*, 315-338.

Gupta, D. R., & Kundu, D. (2004). Discriminating between gamma and generalized exponential distributions. *Journal of Statistical Computation and Simulation, 74*(2),107-121.

Kundu, D., Gupta, D. R., & Manglic, A. (2005). Discriminating between the lognormal and generalized exponential distributions. *Journal of Statistical Planning and Inference, 127*, 213-227.

Raqab, M. Z.(2002). Inferences for generalized exponential distribution based on record statistics. *Journal of Statistical Planning and Inference, 104*, 339-350.

Raqab, M. Z., & Ahsanullah, M. (2001). Estimation of the location and scale parameters of generalized exponential distribution based on record statistics. *Journal of Statistical Computation and Simulation, 69*,109-124.

Sultan, K. S. (2007). Order Statistics from the Generalized Exponential Distribution and Applications. *Communications in Statistics-Theory and Methods, 36*(7), 1409-1418.