Journal of Probability and Statistical Science
14(2), 141-147, Aug. 2016

Discriminating between Generalized Exponential and Gamma Distributions

Orawan Supapueng Kamon Budsaba Andrei I. Volodin Pranee Nilkorn
Thammasat University *University of Regina* *Silpakorn University*

ABSTRACT Generalized Exponential and Gamma distributions are the most popular in analyzing skewed lifetime data. They have many similar properties. Nevertheless they have some different properties, especially when the lifetime data analysis emphasizes the tail of the probabilities. We can observe that it will be more efficient if we can select the correct distribution for a given data. Therefore in this article, we investigate the asymptotic method for distinguishing these two distributions. It is observed that the asymptotic distribution is independent of a nuisance parameter. We perform some numerical experiments to observed that the asymptotic method works for different sample sizes.

Keywords Exponential distribution; Generalized gamma distribution; Generalized invariant property; Lifetime distribution; Separate hypothesis; Statistical testing.

1. Introduction

Nowadays we can observe a dramatic increase of production in different areas of activities. Regardless of what is produced, it is crucial to pay more attention to the reliability of products. Reliability of a device or a product, it is an important indication of its quality. The significant matter about reliability theory is the concept of lifetime distributions. There are many different lifetime distributions because every product will provide different information about its' lifetime so that we should be careful and critical in selecting a lifetime distribution to describe lifetime data from a representative sample of units.

Received May 2015, revised July 2015, in final form August 2016.
Orawan Supapueng (email: aorstat@gmail.com) and Kamon Budsaba are affiliated to the Department of Mathematics and Statistics at Thammasat University, Rangsit Center, Pathum Thani 12121, Thailand. Andrei Volodin is a Professor in the Department of Mathematical and Statistics at University of Regina, Regina, Saskatchewan, Canada; email: andrei@uregina.ca. Pranee Nilkorn was affiliated to Silpakorn University, Bangkok 10710, Thailand.
The authors gratefully acknowledge the financial support provided by Thammasat University Research Fund under the TU Research Scholar, Contract No. TN45/2558.

 JPSS: ISSN 1726-3328

2. Methods

This research desired to develope statistical tests for discriminating between models Gamma and Generalized Exponential as the following:

$$\begin{aligned}
\text{Case 1:} \quad & H_0: \quad X \sim Gamma \\
& H_1: \quad X \sim Generalized\ Exponential \\
\text{Case 2:} \quad & H_0: \quad X \sim Generalized\ Exponential \\
& H_1: \quad X \sim Gamma
\end{aligned}$$

which these distributions are the special case of the Generalized Gamma distribution. Therefore we will consider a random sample, $X_1, X_2, ..., X_n$, from Generalized Gamma distribution and then specified the parameters to each distribution. Stacy [2] proposed the Generalized Gamma distribution (ggd.) with probability density

$$\frac{1}{\theta} f\left(\frac{x}{\theta} \mathcal{I}\, \lambda, \beta\right) = \frac{1+\beta}{\Gamma\left(\frac{(1+\lambda)}{(1+\beta)}\right)} \theta^{\left(\frac{1+\lambda}{1+\beta}\right)} x^{\lambda} exp\left(-\frac{x^{1+\beta}}{\theta}\right) \quad \mathcal{I}\ x > 0, \ \ \theta > 0, \ \lambda, \ \beta > -1 \quad (1)$$

which is a popular model of reliability theory in the study of the lifetime distribution. The Generalized Gamma distribution is very complicated function. It will be easier if we use invariant property statistic under the transformation $x' = \theta x$ and also by the simple idea that we do not test this parameter. The invariant tests are a function of the maximal invariant $T = \{Y_1, Y_2, ..., Y_{n-1}\}$ where $Y_i = \frac{x_i}{x_n}, i = 1, 2, ..., n-1$ ([1], Chapter 6). So once again, it is simple to take $\theta = 1$ and also there are no information lose.

Therefore we use the Generalized Gamma distribution with $\theta = 1$ where the probability density function as the follow

$$f\left(x \mathcal{I}\, \lambda, \beta\right) = \frac{1+\beta}{\Gamma\left(\frac{(1+\lambda)}{(1+\beta)}\right)} x^{\lambda} exp\left(-x^{1+\beta}\right) \quad \mathcal{I}\ x > 0, \quad \lambda, \ \beta > -1 \qquad (2)$$

Note that

we assign $\quad \beta = 0 \quad$ for the Gamma distribution

$\lambda = 0 \quad$ for the Generalized Exponential distribution

In [3] and [4] Volodin discussed the statistic $\mathcal{T} = c_1 \mathcal{T}_1 + c_2 \mathcal{T}_2$. This research will focus on the statistic: $\mathcal{T} = \mathcal{T}_1 / \mathcal{T}_2$ where

$$\mathcal{T}_1 = \frac{1}{n} \sum \ln X_i - \ln\left(\frac{1}{n} \sum X_i\right), \quad \mathcal{T}_2 = \ln\left(\frac{1}{n} \sum X_i\right) - \frac{\frac{1}{n} \sum X_i \ln X_i}{\frac{1}{n} \sum X_i}.$$

Which is

$$\mathcal{T} = \frac{\frac{1}{n} \sum_{i=1}^{n-1} lnY_i - ln\left(\frac{1}{n}\left(\sum_{i=1}^{n-1} Y_i + 1\right)\right)}{ln\left(\frac{1}{n}\left(\sum_{i=1}^{n-1} Y_i + 1\right)\right) - \frac{\frac{1}{n} \sum_{i=1}^{n-1} Y_i lnY_i}{\frac{1}{n}\left(\sum_{i=1}^{n-1} Y_i + 1\right)}}$$

Statistics $\mathcal{T}_1$ and $\mathcal{T}_2$ depend on three simple statistics:

$$T_1 = \frac{1}{n}\sum_1^n X_i \quad , T_2 = \frac{1}{n}\sum_1^n \ln\left(X_i\right), \quad T_3 = \frac{1}{n}\sum_1^n X_i \ln(X_i) \ .$$

Then in this notation, statistics $\mathcal{T}_1$ and $\mathcal{T}_2$ can be written as

$$\mathcal{T}_1 = T_2 - \ln T_1 \ , \quad \mathcal{T}_2 = \ln T_1 - \frac{T_3}{T_1} \ .$$

The main test statistic takes the form

$$\mathcal{T} = \frac{T_2 - lnT_1}{lnT_1 - \frac{T_3}{T_1}}.$$

The test statistic $\mathcal{T} = g(T_1, T_2, T_3)$ is a function of the three main statistics. The Asymptotic distribution of $\mathcal{T}$ is found with the help of the Delta method, which is a procedure of the stochastic representation of $\mathcal{T}$ with the accuracy $O_P(1/\sqrt{n})$. We expand function g in Taylor series by the powers of $T_i - \mu_i, i = 1, 2, 3$, where $\mu_i = E(T_i)$ and the mathematical expectation is taken under the assumption that one of hypotheses is true (null or alternative).

These two distributions are reduced to the ordinal Exponential distribution E^0 when parameters $\lambda = \tau = \beta = 0$. Because of that we can interpret the Exponential distribution E^0 as a boundary that separates null hypotheses and the alternative. Based on this observation we can define the critical constant for all tests. For Case 1 used critical constant $C1$, for the test statistic $\mathcal{T} > C1$, is the $(1 - \alpha)100$ percentile of the statistic when the sample is taken from the Exponential distribution E^0. And for the case 2 used the critical constant $C2$, for the test statistic $\mathcal{T} < C2$, is the $(\alpha)100$ percentile of the statistic when the sample is taken from the E^0.

3. Theoretical Results

In order to find the mean, variance and covariance of the statistics T_i, $i = 1, 2, 3$ for the Generalized Gamma distribution . It is neccessary to calculate $E[X^k(lnX)^j]$ for $k, j = 0, 1, 2$. Therefore our task is to calculate these all moments.

Lemma 1. *Let random variable X has the Generalized Gamma distribution with $\theta = 1$ then*

$$\mu_{k,j} = E[X^k(lnX)^j] = \frac{\Gamma^{(j)}\left(\frac{\lambda+k+1}{1+\beta}\right)}{(1+\beta)^j\Gamma\left(\frac{1+\lambda}{1+\beta}\right)}, \tag{3}$$

where $\Gamma^{(j)}(a) = \frac{d^j}{dx^j}\Gamma(x)|_{x=a}$

Proof. We have $\mu_{k,j} = \int_0^\infty \frac{1+\beta}{\Gamma\left(\frac{1+\lambda}{1+\beta}\right)} x^{\lambda+k}(lnx)^j \, exp\{-x^{1+\beta}\}dx$. Consider transform variable, let $t = x^{1+\beta}$. Then we get $x = t^{\frac{1}{1+\beta}}$ and $dx = \left(\frac{1}{1+\beta}\right) t^{-\frac{\beta}{1+\beta}} dt$. Hence

$$\mu_{k,j} = \frac{1}{\Gamma\left(\frac{1+\lambda}{1+\beta}\right)} \int_0^\infty t^{\frac{\lambda+k}{1+\beta} - \frac{\beta}{1+\beta}} (ln\, t^{\frac{1}{1+\beta}})^j \, e^{-t} dt = \frac{(1+\beta)^{-j}}{\Gamma\left(\frac{1+\lambda}{1+\beta}\right)} \int_0^\infty t^{\frac{\lambda+k-\beta}{1+\beta}} (\, ln\, t)^j \, e^{-t} dt.$$

The integral $\int_0^\infty t^a \, (\ln\, t)^j \, e^{-t}\, dt = \frac{d^j}{da^j} \int_0^\infty t^a \, e^{-t} dt = \frac{d^j}{da^j} \Gamma\,(a+1) = \Gamma^{(j)}\,(a+1)$. Hence we obtain $\mu_{k,j} = \Gamma^{(j)}\left(\frac{\lambda+k+1}{1+\beta}\right) / \left[(1+\beta)^j \, \Gamma\left(\frac{1+\lambda}{1+\beta}\right)\right]$. □

For this we use the following supporting function and formulas. The derivative of the $ln\Gamma(x)$ is called Digamma Euler function which denoted as $\psi(x) = \Gamma'(x)/\Gamma(x)$. and Trigamma Euler function is in the from $\psi'(x) = \Gamma''(x)/\Gamma(x) - (\Gamma'(x)/\Gamma(x))^2$. Hence, $\Gamma'(x) = \Gamma(x)\psi(x)$ and $\Gamma''(x) = \Gamma(x)[\psi'(x) + (\psi(x))^2]$. Since $g(T)$ is differentiable function So we expand $\mathcal{T} = g(T) = (T_2 - lnT_1)/(lnT_1 - T_3/T_1)$ by the first-order Taylor series is

$$\mathcal{T} = g(T) = g(\mu) + \sum_{i=1}^{3} g_i'(\mu)(T_i - \mu_i)$$

where $\mathbf{T} = (T_1, T_2, T_3)$ and $\mu = (\mu_1, \mu_2, \mu_3)$

Lemma 2. *Suppose X has Generalized Gamma Distribution. The asymptotic mean and variance of the test statistic is as following:*

$$\mu_{\mathcal{T}} = E[\mathcal{T}] = g(\mu) = \frac{\mu_2 - ln\mu_1}{ln\mu_1 - \frac{\mu_3}{\mu_1}},$$

$$Var[\mathcal{T}] = E[(g(T) - g(\mu))^2] = E\left[\Big(\sum_{i=1}^{3} g_i'(\mu)(T_i - \mu_i)\Big)^2\right]$$

$$= \sum_{i=1}^{3} (g_i'(\mu))^2 \sigma_i^2 + 2\sum_{ij}\sum g_i'(\mu) g_j'(\mu) \eta_{ij},$$

$$\text{where} \quad g_1'(\mu) = \frac{\mu_3 \left(\frac{1-\mu_2+ln\mu_1}{\mu_1^2}\right) - \frac{\mu_2}{\mu_1}}{\left(ln\mu_1 - \frac{\mu_3}{\mu_1}\right)^2}, \quad g_2'(\mu) = \frac{1}{ln\mu_1 - \frac{\mu_3}{\mu_1}}, \quad g_3'(\mu) = \frac{\frac{\mu_2}{\mu_1} - \frac{ln\mu_1}{\mu_1}}{\left(ln\mu_1 - \frac{\mu_3}{\mu_1}\right)^2}.$$

We derived the asymptotic probability distribution of the test statistic by central limit theorem. Let $\mathcal{T}_n$ be a sequence of random variables such that

$$\sqrt{n}(\mathcal{T}_n - \mu_{\mathcal{T}}) \xrightarrow{d} N(0, \sigma_{\mathcal{T}}^2).$$

Therefore the asymptotic probability distribution of the test statistic is
For $X \sim G[\lambda]$:

$$\sqrt{n}(\mathcal{T} - \mu_G) \xrightarrow{d} N(0, \sigma_G^2).$$

which

$$\mu_G = \frac{\psi(1+\lambda) - ln(1+\lambda)}{ln(1+\lambda) - \psi(2+\lambda)},$$

$$\sigma_G^2 = \sum_{i=1}^{3}(g'_{Gi}(\mu))^2\sigma_{Gi}^2 + 2\sum_{ij}\sum g'_{Gi}(\mu)g'_{Gj}(\mu)\eta_{Gij},$$

$$\text{where} \quad g'_{G1}(\mu) = \frac{\psi(2+\lambda) - \psi(2+\lambda)\psi(1+\lambda) + \psi(2+\lambda)ln(1+\lambda) - \psi(1+\lambda)}{(1+\lambda)(ln(1+\lambda) - \psi(2+\lambda))^2},$$

$$g'_{G2}(\mu) = \frac{1}{ln(1+\lambda) - \psi(2+\lambda)}, \quad g'_{G3}(\mu) = \frac{\frac{\psi(1+\lambda)}{1+\lambda} - \frac{ln(1+\lambda)}{1+\lambda}}{(ln(1+\lambda) - \psi(2+\lambda))^2}.$$

For $X \sim GE[\beta]$:

$$\sqrt{n}(\mathcal{T} - \mu_E) \xrightarrow{d} N(0, \sigma_E^2)$$

which

$$\mu_E = \frac{\frac{\psi\left(\frac{1}{1+\beta}\right)}{1+\beta} - ln\Gamma\left(\frac{2}{1+\beta}\right) + ln\Gamma\left(\frac{1}{1+\beta}\right)}{ln\Gamma\left(\frac{2}{1+\beta}\right) - ln\Gamma\left(\frac{1}{1+\beta}\right) - \frac{\psi\left(\frac{2}{1+\beta}\right)}{1+\beta}},$$

$$\sigma_E^2 = \sum_{i=1}^{3}(g'_{Ei}(\mu))^2\sigma_{Ei}^2 + 2\sum_{ij}\sum g'_{Ei}(\mu)g'_{Ej}(\mu)\eta_{Eij},$$

$$\text{where} \quad g'_{E1}(\mu) = \frac{\frac{\Gamma\left(\frac{1}{1+\beta}\right)\psi\left(\frac{2}{1+\beta}\right)}{\Gamma\left(\frac{2}{1+\beta}\right)(1+\beta)}\left(1 - \frac{\psi\left(\frac{1}{1+\beta}\right)}{1+\beta} + ln\Gamma\left(\frac{2}{1+\beta}\right) - ln\Gamma\left(\frac{1}{1+\beta}\right)\right) - \frac{\psi\left(\frac{1}{1+\beta}\right)\Gamma\left(\frac{1}{1+\beta}\right)}{(1+\beta)\Gamma\left(\frac{2}{1+\beta}\right)}}{\left(ln\Gamma\left(\frac{2}{1+\beta}\right) - ln\Gamma\left(\frac{1}{1+\beta}\right) - \frac{\psi\left(\frac{2}{1+\beta}\right)}{1+\beta}\right)^2},$$

$$g'_{E2}(\mu) = \frac{1}{ln\Gamma\left(\frac{2}{1+\beta}\right) - ln\Gamma\left(\frac{1}{1+\beta}\right) - \frac{\psi\left(\frac{2}{1+\beta}\right)}{1+\beta}},$$

$$g'_{E3}(\mu) = \frac{\frac{\psi\left(\frac{1}{1+\beta}\right)\Gamma\left(\frac{1}{1+\beta}\right)}{(1+\beta)\Gamma\left(\frac{2}{1+\beta}\right)} - \frac{\Gamma\left(\frac{1}{1+\beta}\right)}{\Gamma\left(\frac{2}{1+\beta}\right)}\left(ln\Gamma\left(\frac{2}{1+\beta}\right) - ln\Gamma\left(\frac{1}{1+\beta}\right)\right)}{\left(ln\Gamma\left(\frac{2}{1+\beta}\right) - ln\Gamma\left(\frac{1}{1+\beta}\right) - \frac{\psi\left(\frac{2}{1+\beta}\right)}{1+\beta}\right)^2}.$$

4. Simulation Results

The objective of this study is a development of statistical tests for discriminating between the popular lifetime distribution which concludes 4 cases (see method). In order to obtain numerical information to demonstrate the effectiveness of tests, computer simulations will be performed. The measurement capability of the tests based on both size of the test and empirical power.

In this work used the number of simulation $N = 10,000$ for each of the three parameters λ, τ and β including form -0.9 to 11 with increment 0.1 (start from -0.9, -0.8, -0.7, -0.6, etc. until 11). However the result were present in this paper only case 2 with $n = 50$ as the figure 1.

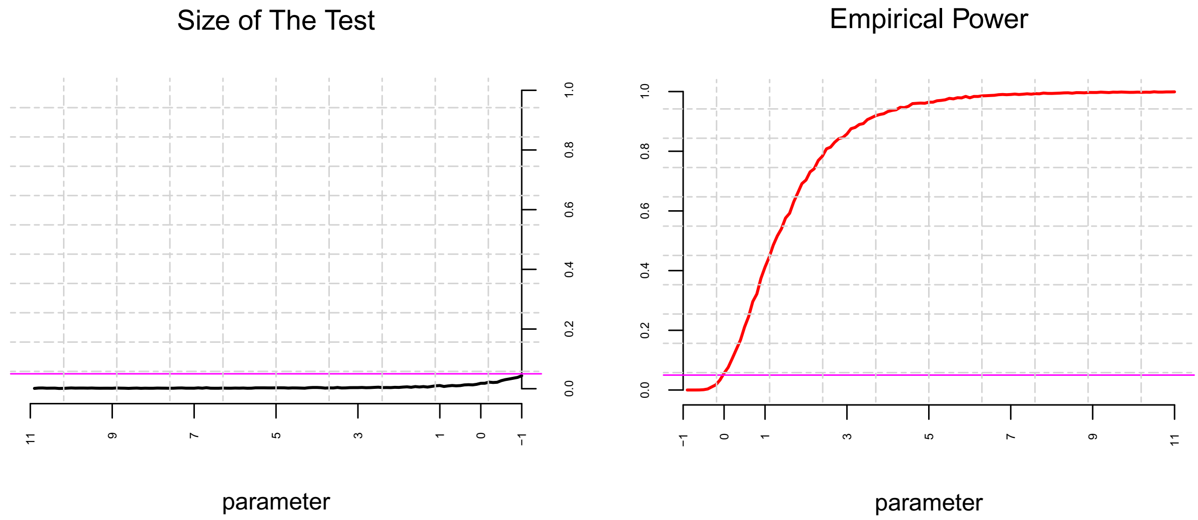

Figure 1 Interpretation from simulation studies for Generalized Exponential vs Gamma test with $n = 50$

Comment in figure 1:

- Size of the test decreases starting about assigned significant level and continuous nearly 0.
- Empirical power starts at 0 for the smallest parameter and rapidly increases to the assigned significant level when parameter increases to 0. After that it increases to 0.8 when parameter increases to 2.5. Afterwards it slightly increases approximate 0.98 for parameter increases to 6. Subsequently it is steady about 0.98.

Discussion

Computer simulation was used to assess the proposed tests. 10^4 sample size was generated and the statistics τ calculated for every individual test. Type I error and power of each test calculated and it was observed from the simulation results that all constructed four tests are powerful unbiased test when parameter is greater than zero. All tests are much more powerful when parameter and sample size are more valuable.

References

[1] Lehmann, E. L. and Romano, J. P. (2005). Testing Statistical Hypotheses, 3rd ed.,

Springer, New York.

[2] Stacy, E. W. (1962). A generalization of the gamma distribution, *The Annals of Mathematical Statistics*, **33**, 1187-1192.

[3] Volodin, I. N. (1981). On discriminating between types connected with the generalized gamma distribution, *Selected Translations in Mathematical Statistics and Probability*, **15**, 213-227.

[4] Volodin, I. N. (1974). The discrimination of the gamma and weibull distributions (Russian), *TeoriyaVeroyatnostei ieePrimeneniya*, **19**, 398-404.