Discriminating Between The Normal and The Laplace Distributions

Debasis Kundu[1]

Abstract

Both normal and Laplace distributions can be used to analyze symmetric data. In this paper we consider the logarithm of the ratio of the maximized likelihoods to discriminate between the two distribution functions. We obtain the asymptotic distributions of the test statistics and it is observed that they are independent of the unknown parameters. If the underlying distribution is normal the asymptotic distribution works quite well even when the sample size is small. But if the underlying distribution is Laplace the asymptotic distribution does not work well for the small sample sizes. For the later case we propose a bias corrected asymptotic distribution and it works quite well even for small sample sizes. Based on the asymptotic distributions, minimum sample size needed to discriminate between the two distributing functions is obtained for a given probability of correct selection. Monte Carlo simulations are performed to examine how the asymptotic results work for small sizes and two data sets are analyzed for illustrative purposes.

Key Words and Phrases: Asymptotic distributions; Likelihood ratio tests; Probability of correct selection; Location Scale Family.

Address of Correspondence: [1] Department of Mathematics, Indian Institute of Technology Kanpur, Pin 208016, INDIA. e-mail: kundu@iitk.ac.in., Phone 91-512-2597141, Fax 91-512-2597500

1 Introduction

Suppose an experimenter has n observations and the elementary data analysis, say for example histogram plot, stem and leaf plot or the box-plot, suggests that it comes from a symmetric distribution. The experimenter wants to fit either the normal distribution or the Laplace distribution, which one is preferable?

It is well known that the normal distribution is used to analyze symmetric data with short tails, whereas the Laplace distribution is used for the long tails data. Although, these two distributions may provide similar data fit for moderate sample sizes, but it is still desirable to choose the correct or more nearly correct model, since the inferences often involve tail probabilities, where the distribution assumption is very important. Therefore, it is very important to make the best possible decision based on whatever data are available.

For a given data set, whether it follows one of the two given distribution functions, is a very well known problem. Discriminating between any two general probability distribution functions was studied by Cox [6], [7], Chambers and Cox [4], Atkinson [1], [2], Dyer [10] and Chen [5]. Dumonceaux and Antle [9] consider the problem of discriminating between the log-normal and Weibull distributions and Dumonceaux, Antle and Haas [8] consider the problem of discriminating between any two distribution functions from the location-scale families. In both [9] and [8] the authors propose the test statistics and compute the critical values based on the Monte Carlo simulations. Fearn and Nebenzahl [11] and Bain and Englehardt [3] consider the problem of discriminating between the gamma and Weibull distributions. Wiens [20], Kim, Sun and Tsutakawa [13], Firth [12] and Kundu and Manglick [14] consider different aspects of discriminating between the log-normal and gamma distributions.

In this paper we discriminate between the normal and Laplace distributions using the ratio of the maximized likelihoods (RML). It should be mentioned that Dumonceaux, Antle and Haas [8] also use an equivalent statistic to the RML to discriminate between the two distribution functions. They did not study any distributional property of the proposed statistic. In this paper, using the approach of White [18], [19] we obtain the asymptotic distribution of the logarithm of RML. It is observed that the asymptotic distribution of the logarithm of the RML is normal and it does not depend on the parameters of the underlying distribution function. Numerical study indicates that if the underlying distribution is normal the asymptotic distribution works quite well even for small sample sizes but the same is not true if the underlying distribution is Laplace. When the underlying distribution is Laplace, we propose a bias corrected asymptotic distribution and it works quite well for small sample sizes also.

The asymptotic distribution can be used to compute the probability of correct selection (PCS). We also obtain the minimum sample size necessary to discriminate between the two distributing functions for a given PCS.

The rest of the paper is organized as follows. At the end of the present section we provide all the notation we are going to use in this paper. We briefly describe the likelihood ratio method in section 2. Asymptotic distributions of the logarithm of RML statistics under null hypotheses are obtained in section 3. Sample size determination has been performed in section 4. Some numerical experiments are performed in section 5 and two real data sets are analyzed in section 6. Finally we conclude the paper in section 7.

We use the following notation for the rest of the paper. The density function of a normal random variable with location parameter $-\infty < \mu < \infty$ and scale parameter $\sigma > 0$, will be denoted by

$$f_N(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}; \quad \text{for} \quad -\infty < x < \infty. \tag{1}$$

A normal distribution with mean μ and variance σ^2 will be denoted by $N(\mu, \sigma^2)$. The density function of a Laplace distribution, with location parameter $-\infty < \eta < \infty$ and scale parameter $\theta > 0$, will be denoted by

$$f_L(x;\eta,\theta) = \frac{1}{2\theta} e^{-\frac{|x-\eta|}{\theta}}; \quad \text{for} \quad -\infty < x < \infty. \tag{2}$$

A Laplace distribution with the location and scale parameters η and θ respectively, will be denoted by $L(\eta,\theta)$.

Also, almost sure convergence will be denoted by $a.s.$ For any Borel measurable function, $h(.)$, $E_N(h(U))$, and $V_N(h(U))$ will denote the mean and variance of $h(U)$ under the assumption that U follows $N(\mu,\sigma^2)$. Similarly, we define, $E_L(h(U))$ and $V_L(h(U))$ as the mean and variance of $h(U)$ when U follows $L(\eta,\theta)$. Moreover, if $g(.)$ and $h(.)$ are two Borel measurable function, we define $cov_N(g(U),h(U)) = E_N(g(U)h(U)) - E_N(g(U))E_N(h(U))$ and $cov_L(g(U)h(U)) = E_L(g(U)h(U))$ - $E_L(g(U))E_L(h(U))$. Finally, we define

$$\begin{aligned} \text{median}\{a_1,\ldots,a_{2m+1}\} &= a_{m+1}, \\ \text{median}\{a_1,\ldots,a_{2m}\} &= \tfrac{a_m+a_{m+1}}{2} \end{aligned}$$

where $a_1 < a_2 < \ldots < a_{2m+1}$ are ordered real numbers.

2 Ratio of Maximized Likelihood

We have a sample $X_1,\ldots,X_n$, from one of the two distribution functions. The likelihood functions, assuming that the data follow $N(\mu,\sigma^2)$ or $L(\eta,\theta)$, are

$$l_N(\mu,\sigma) = \prod_{i=1}^{n} f_N(X_i;\mu,\sigma)$$

and

$$l_L(\eta,\theta) = \prod_{i=1}^{n} f_L(X_i;\eta,\theta),$$

respectively. The logarithm of RML is defined as

$$T = \ln\left[\frac{l_N(\hat{\mu},\hat{\sigma})}{l_L(\hat{\eta},\hat{\theta})}\right]. \tag{3}$$

Here $(\hat{\mu}, \hat{\sigma})$ and $(\hat{\eta}, \hat{\theta})$ are the maximum likelihood estimators (MLEs) of (μ, σ) and (η, θ) respectively based on the sample $X_1, \ldots, X_n$. Therefore T can be written as

$$T = \frac{n}{2}\ln 2 - \frac{n}{2}\ln \pi + n\ln \hat{\theta} - n\ln \hat{\sigma} + \frac{n}{2}, \tag{4}$$

where

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} X_i \quad \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \hat{\mu})^2 \quad \hat{\eta} = \text{median}\{X_1, \ldots, X_n\} \quad \text{and} \quad \hat{\theta} = \frac{1}{n}\sum_{i=1}^{n} |X_i - \hat{\eta}| . \tag{5}$$

The following discrimination procedure can be used. Choose the normal distribution if the test statistic $T > 0$, otherwise choose the Laplace distribution as the preferred model. Note that if the null distribution is $N(\mu, \sigma^2)$, then the distribution of T is independent of μ and σ. Similarly, if the null distribution is $L(\eta, \theta)$, then the distribution of T is independent of η and θ.

3 Asymptotic Properties of the Logarithm of RML

In this section we derive the asymptotic distributions of T.

Case 1: The Data Follow Normal Distribution

In this case we assume that X follows $N(\mu, \sigma^2)$. We have the following result.

Theorem 1: Under the assumption that the data follow $N(\mu, \sigma^2)$, the distribution of T is asymptotically normally distributed with mean $E_N(T)$ and $V_N(T)$.

To prove theorem 1, we need the following lemma.

Lemma 1: Suppose the data follow $N(\mu, \sigma^2)$, then as $n \to \infty$, we have

(1) $\hat{\eta} \to \tilde{\eta}$ a.s., and $\hat{\theta} \to \tilde{\theta}$ a.s., where

$$E_N\left(\ln(f_L(X; \tilde{\eta}, \tilde{\theta}))\right) = \max_{\eta,\theta} E_N\left(\ln(f_L(X; \eta, \theta))\right).$$

Note that $\tilde{\eta}$ and $\tilde{\theta}$ may depend on μ and σ, but we do not make it explicit for brevity. Let us denote;

$$T^* = \ln\left[\frac{l_N(\mu,\theta)}{l_L(\tilde{\eta},\tilde{\theta})}\right], \tag{6}$$

(2) $n^{-\frac{1}{2}}\left(T - E_{LN}(T)\right)$ is asymptotically equivalent to $n^{-\frac{1}{2}}\left(T^* - E_{LN}(T^*)\right)$.

Proof of Lemma 1: The proof follows using the similar argument of White ([19]; Theorem 1) and therefore, it is omitted.

Proof of Theorem 1: Using the Central Limit Theorem (CLT), it can be easily seen that $n^{-\frac{1}{2}}\left(T^* - E_{LN}(T^*)\right)$ is asymptotically normally distributed. Therefore, the proof immediately follows from part (2) of Lemma 1 and the CLT.

Comments: It should be mentioned here that the distance between $L(\tilde{\eta},\tilde{\theta})$ and $N(\mu,\sigma)$ is minimum in terms of the Kullback-Liebler information measure.

Now we compute $\tilde{\eta}$, $\tilde{\theta}$, $E_N(T)$ and $V_N(T)$. Note that;

$$E_N\left(\ln(f_L(X;\eta,\theta))\right) = -\ln 2 - \ln\theta - E\left|\frac{X-\eta}{\theta}\right|.$$

Since X follows $N(\mu,\sigma^2)$, therefore, it is clear that

$$\tilde{\eta} = \mu \qquad \text{and} \qquad \tilde{\theta} = E|X-\mu| = \sigma\sqrt{\frac{2}{\pi}}. \tag{7}$$

Now we provide the expression for $E_N(T)$ and $V_N(T)$. Note that $\lim_{n\to\infty}\frac{E_N(T)}{n}$ and $\lim_{n\to\infty}\frac{V_N(T)}{n}$ exist. Let us denote $\lim_{n\to\infty}\frac{E_N(T)}{n} = AM_N$ and $\lim_{n\to\infty}\frac{V_N(T)}{n} = AV_N$. Since the distribution of T is independent of μ and σ, therefore to compute $E_N(T)$ and $V_N(T)$, without loss of generality, we consider $\mu = 0$ and $\sigma = 1$, *i.e.* X follows $N(0,1)$. Therefore, for large n,

$$\begin{aligned}\frac{E_N(T)}{n} &\approx AM_N = E_N\left[\ln f_N(X;0,1) - \ln f_L(X;\tilde{\eta},\tilde{\theta})\right] \\ &= E_N\left[-\frac{1}{2}\ln(2\pi) - \frac{X^2}{2} + \ln 2 + \ln\tilde{\theta} + \left|\frac{X-\tilde{\eta}}{\tilde{\theta}}\right|\right] \\ &= -\frac{1}{2}\ln(2\pi) - E_N\left(\frac{X^2}{2}\right) + \ln 2 + \ln\tilde{\theta} + \sqrt{\frac{\pi}{2}}E_N|X|\end{aligned}$$

$$= \ln 2 + \frac{1}{2} - \ln \pi = 0.0484172. \tag{8}$$

We also have

$$\begin{aligned}\frac{V_N(T)}{n} &\approx AV_N = V_N\left(\ln f_N(X;0,1) - \ln f_L(X;\tilde{\eta},\tilde{\theta})\right)\\ &= V_N\left(-\frac{1}{2}X^2 + \sqrt{\frac{\pi}{2}}|X|\right)\\ &= \frac{1}{4}V_N(X^2) + \frac{\pi}{2}V_N(|X|) - \sqrt{\frac{\pi}{2}}cov_N(X^2,|X|)\\ &= \frac{1}{2} + \frac{\pi}{2}\left(1 - \frac{2}{\pi}\right) - 1 = 0.0707963. \qquad (9)\end{aligned}$$

Case 2: The Data Follow Laplace Distribution

In this case we assume that X follows $L(\eta,\theta)$. We have the following result.

Theorem 2: Under the assumption that the data follow $L(\eta,\theta)$, the distribution of T is asymptotically normally distributed with mean $E_L(T)$ and variance $V_L(T)$.

To prove Theorem 2, we need Lemma 2, similar to Lemma 1.

Lemma 2: Suppose the data follow $L(\eta,\theta)$, then as $n \to \infty$, we have

(1) $\hat{\mu} \to \tilde{\mu}$ a.s., $\hat{\sigma} \to \tilde{\sigma}$ a.s, where

$$E_L\left(\ln(f_N(X:\tilde{\mu},\tilde{\sigma}))\right) = \max_{\mu,\sigma} E_L\left(\ln(f_N(X;\mu,\sigma))\right).$$

Note that here also $\tilde{\mu}$ and $\tilde{\sigma}$ may depend on η and θ, but we do not make it explicit for brevity. If we denote

$$T_* = \ln\left[\frac{l_N(\tilde{\mu},\tilde{\sigma})}{l_L(\eta,\theta)}\right],$$

then

(2) $n^{-\frac{1}{2}}\left[T - E_L(T)\right]$ is asymptotically equivalent to $n^{-\frac{1}{2}}\left[T_* - E_L(T_*)\right]$.

Proof of Lemma 2: It also follows from Theorem 1 of White [19].

Proof of Theorem 2: It follows similarly as Theorem 1 of Case 1.

Now we discuss how to obtain $\tilde{\mu}$, $\tilde{\sigma}$, $E_L(T)$ and $V_L(T)$. Consider;

$$E_L\left(\ln(f_N(X;\mu,\sigma))\right) = E_L\left[-\frac{1}{2}\ln(2\pi) - \ln\sigma - \frac{(X-\mu)^2}{2\sigma^2}\right].$$

Since X follows $L(\eta,\theta)$, it is immediate that

$$\tilde{\mu} = \eta \quad \text{and} \quad \tilde{\sigma} = \sqrt{2}\,\theta. \tag{10}$$

Now we provide the expression for $E_L(T)$ and $V_L(T)$. Similarly as before, $\lim_{n\to\infty}\frac{E_L(T)}{n}$ and $\lim_{n\to\infty}\frac{V_L(T)}{n}$ exist and we denote $\lim_{n\to\infty}\frac{E_L(T)}{n} = AM_L$ and $\lim_{n\to\infty}\frac{V_L(T)}{n} = AV_L$. Note that as mentioned before, the distribution of T is independent of η and θ and we take $\eta = 0$ and $\theta = 1$ with out loss of generality, for the calculations of AM_L and AV_L. Therefore for large n

$$\begin{aligned}\frac{E_L(T)}{n} &\approx AM_L = E_L\left[\ln(f_N(X;\tilde{\mu},\tilde{\sigma})) - \ln(f_L(X;0,1))\right] \\ &= E_L\left[-\frac{1}{2}\ln(2\pi) - \ln\tilde{\sigma} - \frac{(X-\tilde{\mu})^2}{2\tilde{\sigma}^2} + \ln 2 + |X|\right] \\ &= -\frac{1}{2}\ln(2\pi) - \frac{1}{2}\ln 2 - \frac{1}{2} + \ln 2 + 1 = -0.0723649. \end{aligned} \tag{11}$$

Moreover,

$$\begin{aligned}\frac{V_L(T)}{n} &\approx AV_L = V_L\left[\ln(f_N(X;\tilde{\mu},\tilde{\sigma})) - \ln(f_L(X;0,1))\right] \\ &= V_L\left[-\frac{X^2}{4} + |X|\right] = \frac{1}{16}V_L(X^2) + V_L(|X|) - \frac{1}{2}cov_L(X^2,|X|) \\ &= 0.25. \end{aligned} \tag{12}$$

4 Determination of Sample Size and Testing

4.1 Minimum Sample Size Determination

In this subsection section we propose a method to determine the minimum sample size needed to discriminate between the normal and Laplace distributions for a given user specified probability of correct selection (PCS). It is expected that the user specifies the PCS before hand.

First we consider Case 1, *i.e.* the data follow $N(\mu, \sigma)$. Since T is asymptotically normally distributed with mean $E_N(T)$ and variance $V_N(T)$, therefore, the PCS is

$$PCS = P[T > 0] \approx \Phi\left(\frac{E_N(T)}{\sqrt{V_N(T)}}\right) = \Phi\left(\frac{n \times AM_N}{\sqrt{n \times AV_N}}\right). \tag{13}$$

Here Φ is the distribution function of the standard normal distribution. Now to determine the minimum sample size required to achieve at least p^* protection level (namely the PCS), we equate;

$$\Phi\left(\frac{n \times AM_N}{\sqrt{n \times AV_N}}\right) = p^* \tag{14}$$

and obtain

$$n = \frac{z_{p^*}^2 AV_N}{(AM_N)^2}. \tag{15}$$

Here z_{p^*} is the $100p^*$ percentile points of a standard normal distribution. For Case 2, *i.e.* when the data follow $L(\eta, \theta)$, we obtain

$$n = \frac{z_{p^*}^2 AV_L}{(AM_L)^2}. \tag{16}$$

Therefore, to achieve overall p^* protection level, we need at least

$$n = z_{p^*}^2 \max\{\frac{AV_L}{(AM_L)^2}, \frac{AV_L}{(AM_L)^2}\} = z_{p^*}^2 \max\{30.2, 47.7\} = z_{p^*}^2 48. \tag{17}$$

4.2 Testing of Hypotheses

Cox [6], [7], Dumonceaux and Antle [9] consider the above discrimination problem as the following two testing of hypotheses problems.

$$\text{Problem 1:} \quad H_0 : \text{Normal} \quad vs. \ H_1 : \text{Laplace}. \tag{18}$$

$$\text{Problem 2:} \quad H_0 : \text{Laplace} \quad vs. \ H_1 : \text{Normal}. \tag{19}$$

Dumonceaux and Antle [9] provided the exact critical regions and the powers of the likelihood ratio tests based on Monte Carlo simulations. Our asymptotic results derived in the previous section can be used for testing the above two hypotheses as follows:

Test 1: For Problem 1: Reject the null hypothesis H_0 at the α % level of significance, if $T < n \times 0.0484172 - z_\alpha \times \sqrt{n \times 0.0707963}$, and accept otherwise.

Test 2: For Problem 2: Reject the null hypothesis H_0 at the α % level of significance, if $T > -n \times 0.0723649 + z_\alpha \times \sqrt{n \times 0.25}$, and accept otherwise.

5 Numerical Experiments:

In this section we present some numerical experimental results performed to verify how the asymptotic results work for different sample sizes. All computations are performed at the Indian Institute of Technology Kanpur, using Pentium IV processor. We use the random deviate generator RAN2 of Press *et al.* [17] and all the programs are written in FORTRAN. They can be obtained from the author on request. We compute the PCS based on simulations and also based on the asymptotic normality results derived in section 3. Since the distribution (numerical value) of T is independent of the shape and scale parameters, we consider the location and scale parameters to be zero and one respectively in all cases. We consider different sample sizes namely n = 20, 40, 60, 80 and 100.

First we consider the case when the null distribution is normal and the alternative is Laplace. In this case we generate a random sample of size n from $N(0,1)$ and compute T and check whether T is positive or negative. We replicate the process 10,000 times and obtain an estimate of PCS. The results are reported in Table 1. Similarly, we obtain the results when the null distribution is Laplace and the alternative is normal, and the results are reported in Table 2.

n	20	40	60	80	100
MC	0.80	0.87	0.92	0.95	0.96
AR	0.79	0.88	0.92	0.95	0.96

Table 1: The probability of correct selection based on Monte Carlo simulations (MC) with 10,000 replications and also based on the asymptotic results (AR) when the null distribution is normal and the alternative is Laplace.

n	20	40	60	80	100
MC	0.65	0.79	0.86	0.90	0.93
AS	0.78	0.86	0.91	0.94	0.96
BC	0.68	0.79	0.85	0.89	0.92

Table 2: The probability of correct selection based on Monte Carlo simulations (MC), based on the asymptotic results (AS) and based on biased corrected asymptotic results (BC) with 10,000 replications and when the null distribution is Laplace and the alternative is normal are reported.

It is quite clear from Tables 1 and 2 that as sample size increases the PCS increases as expected. The asymptotic results in Table 1 are quite close to the corresponding simulated results. The same thing is not true for the Table 2. In Table 2 it is observed that for small sample sizes the asymptotic results do not match well with the simulated results. We study carefully the reason and it is observed that T is biased in this case. We propose the following biased corrected version for small sample sizes. We suggest to use AM_L = -0.0723649 + $\frac{.64}{n}$ and AV_L = 0.25 - $\frac{2.09}{n}$. The results, based on the biased corrected version, are also provided in Table 2. The biases are obtained by regressing the simulated mean and variance of T with the theoretical mean and variance respectively, over a wide range of sample sizes. Based on the bias correction, the value of n in (16) can be modified to

$$n = \frac{-(1.28AM_L - z_{p^*}^2 AV_L) + \sqrt{(1.28AM_L - z_{p^*}^2 AV_L)^2 - 4AM_L^2(0.64^2 + 2.09z_{p^*}^2)}}{2AM_L^2},$$

and therefore, equation (17) and Test 2 can also be modified accordingly.

Now we consider the discrimination problem as a testing of hypothesis problem as defined in the previous section. Let us define the rejection regions as $\{T < 0\}$ and $\{T > 0\}$ for problems 1 and 2 respectively. Therefore, it is immediate that P[Type I error] = 1 - PCS. From Tables 1 and 2, it is clear that P[Type I error] varies between 0.20 to 0.04 as the sample size varies between 20 to 100 when the null distribution is normal and when the null distribution is Laplace it varies between 0.35 to 0.07. Similarly in two cases the power of the tests vary between 0.65 to 0.93 and between 0.80 to 0.96 respectively as sample size varies between 20 to 100.

6 Data Analysis

In this section we analyze two data sets and use our method to discriminate between the two distribution functions.

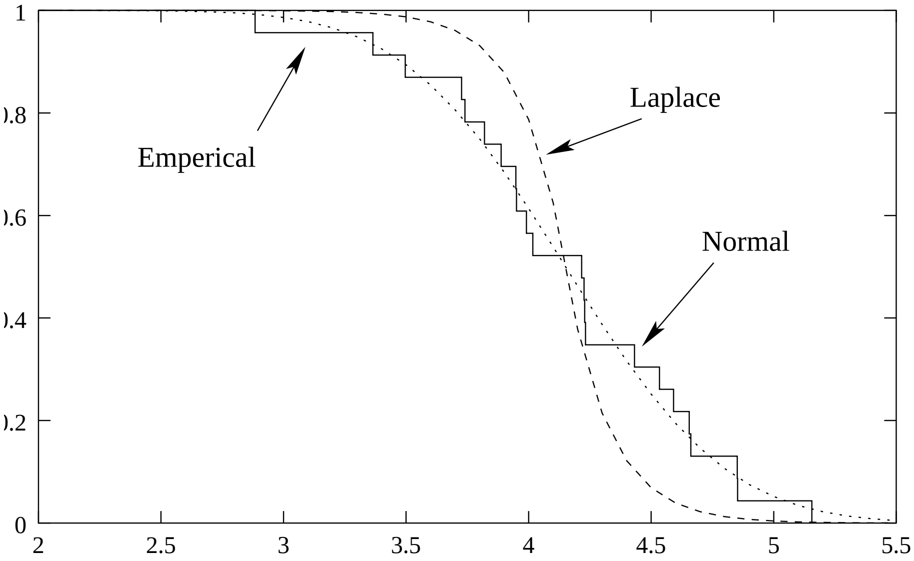

Figure 1: Emperical survival function, fitted normal and fitted Laplace survival functions for data set 1.

Data Set 1: The first data set is as follows; (Lawless; 1986, page 228). The data are the number of million revolutions before failure for each of the 23 ball bearings in the life tests and they are: 17.88, 28.92, 33.00, 41.52, 42.12, 45.60, 48.80, 51.84, 51.96, 54.12, 55.56, 67.80, 68.44, 68.64, 68.88, 84.12, 93.12, 98.64, 105.12, 105.84, 127.92, 128.04, 173.40.

Several authors analyzed this data set using gamma, Weibull, log-normal, generalized exponential models. We make the log transformation of the data and try to fit the normal and Laplace distributions. When we use the normal model, the MLEs of the different parameters are $\hat{\mu} = 4.1506$, $\hat{\sigma} = 0.5215$ and the corresponding log-likelihood (LL) value is -17.6615. The Kolmogorv-Smirnov (K-S) distance between the data and the fitted normal distribution function is 0.0899 and the corresponding p value is 0.99. Similarly when we fit the Laplace model, the MLEs are $\hat{\eta} = 4.1506$, $\hat{\theta} = 0.4200$ and the corresponding LL value is -18.9899. The non-parametric survival function, fitted normal survival function and fitted Laplace survival function are plotted in Figure 1.

The K-S distance between the data and the fitted Laplace distribution function is 0.2473

Intervals	Observed	Normal	Laplace
$< \ln 35$	3	2.92	2.79
ln 35 - ln 55	7	6.04	5.38
ln 55 - ln 80	5	6.48	8.20
ln 80 - ln 100	3	3.15	2.73
$> \ln 100$	5	4.41	3.90

Table 3: The observed and the expected frequencies due to the normal and Laplace distributions for data set 1

and the corresponding p value is 0.12. We also present the observed, expected frequencies for different intervals in Table 3.

The χ^2 values are 0.579 and 2.089 for the normal and Laplace distributions respectively. For data set 1, K-S distances, χ^2 values and Figure 1 indicate that for the transformed data, the normal distribution is a better fit compared to the Laplace distribution,

The logarithm of RML *i.e.*, $T = 1.3284 > 0$. It also indicates to choose the normal model. From (13), it is clear that if the data follow normal distribution then based on sample size 23, PCS = 0.81 and if the data follow Laplace distribution then the bias corrected PCS = 0.71. Therefore, in this case the PCS is at least min$\{0.71, 0.81\}$ = 0.71. Based on the assumption that the data follow normal distribution, the p-value = 0.57. Similarly based on the assumption that the data follow Laplace distribution, the p-value = 0.11. Comparing the two p values also we would prefer to choose normal distribution over Laplace distribution. Therefore, in this case, the LL values, χ^2 values, K-S distances and our proposed method indicate to choose the normal model and the probability of correct selection is at least 71%. If we consider the two testing of hypotheses problems (18) and (19), then based on the data we can not reject the null hypotheses in both cases even with the 10% level of significance.

Data Set 2: Now consider the following data set artificially created from the Laplace distribution, namely $L(0, 1)$ and truncated at the second decimal places. The data are as

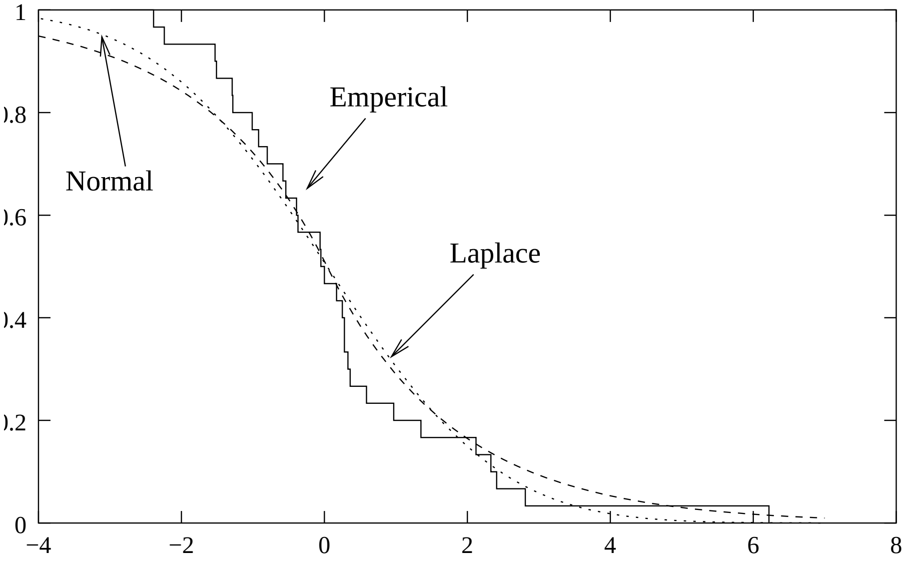

Figure 2: Emperical survival function, fitted normal and fitted Laplace survival functions for data set 2.

follows: -1.28, 0.36, -1.29, -0.80, 0.28, -0.06, -1.53, 0.28, -0.54, 0.17, 0.59, 6.22, 2.41, 0.33, -1.51, 0.25, 2.33, 2.81, -0.92, 2.12, -1.01, 1.35, -0.37, -0.39, -4.39, -2.39, 0.97, -0.58, -2.24, -0.05.

Using the normal model the MLEs of the different parameters are $\hat{\mu} = 0.0373$, $\hat{\sigma} = 1.8906$ and the corresponding log-likelihood (LL) value is -61.6744. The Kolmogorv-Smirnov (K-S) distance between the data and the fitted normal distribution function is 0.1656 and the corresponding p value is 0.3833. Similarly using the Laplace model, the MLEs are $\hat{\eta} = 0.0373$, $\hat{\theta} = 1.3298$ and the corresponding LL value is -59.3458. The non-parametric, fitted normal and fitted Laplace survival functions are plotted in Figure 2.

The K-S distance between the data and the fitted Laplace distribution function is 0.1499 and the corresponding p value is 0.5100. The observed and expected frequencies are presented in Table 4.

The χ^2 values are 1.356 and 1.035 for normal and Laplace distributions respectively. For

Intervals	Observed	Normal	Laplace
< -3.0	1	1.64	1.55
-3.0 – -1.5	4	4.42	3.22
-1.5 – 0.0	11	8.70	9.90
0.0 – 1.5	9	8.71	10.34
1.5 – 3.0	4	4.78	3.37
> 3.0	1	1.75	1.62

Table 4: The observed and the expected frequencies due to the normal and Laplace distributions for data set 2

data set 2, K-S distances, χ^2 values and Figure 2 indicate that the Laplace distribution is better fit compared to the normal distribution,

The logarithm of RML *i.e.*, T = -2.3286 < 0. Therefore, our proposed method suggests to choose the Laplace model. From (13), it is clear that if the data follow normal distribution then based on sample size 30, PCS = 0.84 and if the data follow Laplace distribution then the bias corrected PCS = 0.74. Therefore, in this case the PCS is at least min{0.74, 0.84} = 0.74. Based on the assumption that the data follow normal distribution, the p-value < 0.005. Similarly based on the assumption that the data follow Laplace distribution, the p-value > 0.62. Comparing the two p values also it is immediate that the Laplace distribution is preferable over normal distribution. Therefore, in this case, the LL values, χ^2 values, K-S distances and our proposed method indicate to choose the Laplace model and the probability of correct selection is at least 74%. If we consider the two testing of hypotheses problems (18) and (19), then based on the data we reject the null hypotheses for the problem (18) for any $\alpha < 0.005$.

7 Conclusions

In this paper we consider the problem of discriminating the two important families of distributions namely the normal and Laplace families. We consider the statistic based on the ratio of maximized likelihoods and obtain the asymptotic distributions of the test statistics under null hypotheses. It is observed that the asymptotic distributions are asymptotically normal and they are independent of the parameters of the null distribution. Our method can be used for discriminating between any two members of the different location and scale families. The exact mean and variance of the corresponding normal distribution needs to be derived in each case. Finally, we should mention that for a given data set it may happen that none of the two distribution functions provide good fit. It should be clear from the $K - S$ values and also from the χ^2 values. Those cases some other family members should be used.

References

[1] Atkinson, A. (1969), "A test of discriminating between models", *Biometrika*, vol. 56, 337-341.

[2] Atkinson, A. (1970), "A method for discriminating between models" (with discussions), *Journal of the Royal Statistical Society*, Ser. B, vol. 32, 323-353.

[3] Bain, L.J. and Englehardt, M. (1980), "Probability of correct selection of Weibull versus gamma based on likelihood ratio", *Communications in Statistics*, Ser. A., vol. 9, 375-381.

[4] Chambers, E.A. and Cox, D.R. (1967), "Discriminating between alternative binary response models", *Biometrika*, 54, 573-578.

[5] Chen, W.W. (1980), "On the tests of separate families of hypotheses with small sample size", *Journal of Statistical Computations and Simulations*, vol. 2, 183-187.

[6] Cox, D.R. (1961), "Tests of separate families of hypotheses", *Proceedings of the Fourth Berkeley Symposium in Mathematical Statistics and Probability*, Berkeley, University of California Press, 105-123.

[7] Cox, D.R. (1962), "Further results on tests of separate families of hypotheses", *Journal of the Royal Statistical Society*, Ser. B, vol. 24, 406-424.

[8] Dumonceaux, R., Antle, C.E. and Haas, G. (1973), "Likelihood ratio test for discriminating between two models with unknown location and scale parameters", *Technometrics*, vol. 15, 19-31.

[9] Dumonceaux, R. and Antle, C.E. (1973), "Discriminating between the log-normal and Weibull distribution", *Technometrics*, vol. 15, 923-926.

[10] Dyer, A.R. (1973), "Discrimination procedure for separate families of hypotheses", *Journal of the American Statistical Association*, vol. 68, 970-974.

[11] Fearn, D.H. and Nebenzahl, E. (1991), "On the maximum likelihood ratio method of deciding between the Weibull and Gamma distributions", *Communications in Statistics - Theory and Methods*, vol. 20, 579-593.

[12] Firth, D. (1988), "Multiplicative errors: log-normal or gamma?", *Journal of the Rpyal Statistical Society, Ser. B*, 266-268.

[13] Kim, H., Sun, D. and Tsutakawa, R. K. (2002), "Lognormal vs. Gamma: Extra Variations", *Biometrical Journal*, vol.44, 305-323.

[14] Kundu, D. and Manglick, A. (2004), "Discriminating between the Log-Normal and gamma distributions", to appear in the *Journal of the Applied Statistical Sciences.*

[15] Lawless, J.F. (1982), *Statistical Models and Methods for Lifetime Data*, John Wiley and Sons, New York.

[16] Linhardt, H. and Zucchini, W. (1986), *Model Selection*, Wiley, New York.

[17] Press, W.H., Flannery, B.P., Teukolsky, S.A. and Vetterling, W.T. (1993), *Numerical Recipes; The Art of Scientific Computing*, Cambridge University Press, Cambridge, U.K.

[18] White, H. (1982a), "Maximum likelihood estimation of mis-specified models", *Econometrica*, vol. 50, 1-25.

[19] White, H. (1982b), "Regularity conditions for Cox's test of non-nested hypotheses", *Journal of Econometrics*, vol. 19, 301-318.

[20] Wiens, B.L. (1999), "When log-normal and gamma models give different results: a case study", *American Statistician*, vol. 53, 89-93.