▨ 연구논문

A Bootstrap Test of Independence for an Absolutely Continuous Bivariate Exponential Model

In Suk Lee · Dal Ho Kim · Jang Sik Cho
Dept. of Statistics, Kyungpook National University

Abstract

In this paper, we consider the problem of testing independence in the absolutely continuous bivariate exponential distribution of Block and Basu(1974). We construct a bootstrap procedure for testing zero and non-zero values of the parameter λ_3 which measures the degree of dependence and compare the power of the bootstrap test with likelihood ratio test(LRT) by Gupta et al.(1984) and the test based on maximum likelihood estimator(MLE) $\hat{\lambda}_3$ by Hanagal and Kale(1991) for small and moderate sample sizes.

1. Introduction.

We consider a system consisting of two components (C_1, C_2) and let $\overline{F}(x,y)=P(X>x, Y>y)$ where X and Y denote failure times of C_1 and C_2 respectively. In general the component failure times X and Y may be dependent. Marshall and Olkin(1967) proposed a bivariate exponential distribution as a model for failure time distribution of a system with two components which can fail simultaneously. This distribution however is not absolutely continuous and so there are some situations when this model is not appropriate. As an alternative Block and Basu(1974) dropped the condition of exponential marginals and used the loss of memory property to propose an absolutely continuous bivariate exponential model. According to Block and Basu(1974) the failure times of the two components (C_1, C_2) are said to follow ACBVE (λ_1, λ_2, λ_3) if

$$\overline{F}(x,y) = P(X>x, Y>y)$$
$$= \frac{\lambda}{(\lambda_1+\lambda_2)} \exp\{-\lambda_1 x - \lambda_2 y - \lambda_3 \max(x,y)\}$$
$$- \frac{\lambda_3}{(\lambda_1+\lambda_2)} \exp\{-\lambda \max(x,y)\}, \qquad (1.1)$$

where $x,\ y > 0$; $\lambda_1, \lambda_2, \lambda_3 > 0$ and $\lambda = \lambda_1 + \lambda_2 + \lambda_3$.

In the above modification of Marshall-Olkin model given by Block and Basu(1974), the property that $\lambda_3 = 0$ if and only if (X, Y) are independent is preserved. The test of hypothesis $\lambda_3 = 0$ is therefore equivalent to testing for independence of (X, Y). The problem of tests for independence in Marshall-Olkin model has been studied in detail by Bemis, Bain and Higgins(1972) and Bhattacharyya and Johnson(1973). In the case of Block and Basu model, Gupta Mehrotra and Michalek(1984) obtained a likelihood ratio test(LRT) for $\lambda_3 = 0$ when $\lambda_1 = \lambda_2$ are unknown. They derived the exact distribution of the LRT statistic under H_0. But it is very difficult to find out the distribution of the LRT statistic under the alternatives. Hanagal and Kale(1991) derived a test using the asymptotic normal distribution based on maximum likelihood estimator(MLE) $\hat{\lambda}_3$ of λ_3 for $H_0 : \lambda_3 = 0$ against $H_1 : \lambda_3 > 0$ when λ_1 and λ_2 are unknown and unequal. But the power performance of the test based on MLE $\hat{\lambda}_3$ may be not good in small samples since the test statistic is often skewed and biased.

Efron(1979) initially introduced the bootstrap method to assign the accuracy for an estimator. General theory for bootstrap hypothesis testing is discussed briefly by Hinkley(1988) during a survey of bootstrap methods, and at greater length by Hinkely(1989). Beran(1988) discussed pivoting in the context of bootstrap hypothesis testing. Hall and Wilson(1991) and Becher(1993) illustrated the two guidelines of pivoting and sampling under null hypothesis by applying bootstrap tests to specific data sets.

In this paper, we construct the bootstrap procedure for testing independence using bootstrap pivoting in the ACBVE model and compare the powers with the LRT and the test based on the MLE $\hat{\lambda}_3$ via Monte Carlo simulation in small and moderate samples.

2. Tests of Independence

Let (x_1, y_1), (x_2, y_2), $\cdots$, (x_n, y_n) be a random sample from the ACBVE (λ_1, λ_2, λ_3) in (1.1). In this case, the marginal probability density function(pdf) of X is given by

$$f_1(x) = \frac{\lambda(\lambda_1+\lambda_3)}{\lambda_1+\lambda_2}\exp\{-(\lambda_1+\lambda_3)x\} - \frac{\lambda_3\lambda}{\lambda_1+\lambda_2}\exp\{-\lambda x\},\ x > 0. \tag{2.1}$$

Similarly the marginal pdf of Y is given by

$$f_2(y) = \frac{\lambda(\lambda_2+\lambda_3)}{\lambda_1+\lambda_2}\exp\{-(\lambda_2+\lambda_3)y\} - \frac{\lambda_3\lambda}{\lambda_1+\lambda_2}\exp\{-\lambda y\},\ y > 0. \tag{2.2}$$

Under the restriction of the identical marginals, $\lambda_1 = \lambda_2$ in (1.1), Gupta, Metrotra and Michalek(1984) obtained the likelihood ratio statistic $T = R(1-R)$ for

$$H_0 : \lambda_3 = 0,\ \text{where}\ R = \frac{\sum_{i=1}^{n} w_i}{2\sum_{i=1}^{n} v_i + \sum_{i=1}^{n} w_i},\quad w_i = |x_i - y_i|\ \text{and}\ v_i = \min(x_i, y_i).$$

They showed that the ratio R has a beta (n, n) distribution under $H_0 : \lambda_3 = 0$. The test function of Gupta, Metrotra and Michalek(1984) is given by

$$\phi_{GMM}(x, y) = \begin{cases} 1, & \text{if } T \le c \\ 0, & \text{otherwise} \end{cases} \tag{2.3}$$

where c, $0 \le c \le 1/4$, is chosen so that $P(T \le c \mid H_0) = \alpha$.

Suppose that in a sample of size n, n_1 observations are such that $x_i < y_i$ and $n_2 = n - n_1$ are such that $x_i \ge y_i$. Then the likelihood of the sample is given by

$$L(\lambda_1, \lambda_2, \lambda_3) = \left[\frac{\lambda\lambda_1(\lambda_2+\lambda_3)}{(\lambda_1+\lambda_2)}\right]^{n_1} \cdot \left[\frac{\lambda\lambda_2(\lambda_1+\lambda_3)}{(\lambda_1+\lambda_2)}\right]^{n_2} \cdot$$

$$\exp\{-\lambda_1\sum_{i=1}^{n} x_i - \lambda_2\sum_{i=1}^{n} y_i - \lambda_3\sum_{i=1}^{n}\max(x_i, y_i)\}. \tag{2.4}$$

Hence, the likelihood equations are as follows:

$$\frac{\partial \log L(\lambda_1, \lambda_2, \lambda_3)}{\partial \lambda_1} = -\sum_{i=1}^{n} x_i + n/\lambda - n/(\lambda_1+\lambda_2) + n_1/\lambda_1 + n_2/(\lambda_1+\lambda_3) = 0; \tag{2.5}$$

$$\frac{\partial \log L(\lambda_1, \lambda_2, \lambda_3)}{\partial \lambda_2} = -\sum_{i=1}^{n} y_i + n/\lambda - n/(\lambda_1+\lambda_2) + n_2/\lambda_2 + n_1/(\lambda_2+\lambda_3) = 0; \tag{2.6}$$

$$\frac{\partial \log L(\lambda_1, \lambda_2, \lambda_3)}{\partial \lambda_3} = -\sum_{i=1}^{n} \max(x_i, y_i) + n/\lambda + n_1/(\lambda_2+\lambda_3) + n_2/(\lambda_1+\lambda_3) = 0. \quad (2.7)$$

We can solve the likelihood equations by Newton-Raphson method and obtain MLE's ($\hat{\lambda}_1$, $\hat{\lambda}_2$, $\hat{\lambda}_3$). Hanagal and Kale(1991) obtained the Fisher information matrix $I(\lambda_1, \lambda_2, \lambda_3) = (I_{ij})$: $(i, j) = 1,2,3$ given as

$$I_{11} = \frac{1}{\lambda^2} - \frac{1}{(\lambda_1+\lambda_2)^2} + \frac{1}{\lambda_1(\lambda_1+\lambda_2)} + \frac{\lambda_2}{(\lambda_1+\lambda_2)(\lambda_1+\lambda_3)^2},$$

$$I_{22} = \frac{1}{\lambda^2} - \frac{1}{(\lambda_1+\lambda_2)^2} + \frac{1}{\lambda_2(\lambda_1+\lambda_2)} + \frac{\lambda_1}{(\lambda_1+\lambda_2)(\lambda_2+\lambda_3)^2},$$

$$I_{33} = \frac{1}{\lambda^2} + \frac{1}{(\lambda_1+\lambda_2)}\left[\frac{\lambda_1}{(\lambda_2+\lambda_3)^2} + \frac{\lambda_2}{(\lambda_1+\lambda_3)^2}\right], \quad I_{12} = \frac{1}{\lambda^2} - \frac{1}{(\lambda_1+\lambda_2)^2},$$

$$I_{13} = \frac{1}{\lambda^2} + \frac{\lambda_2}{(\lambda_1+\lambda_2)(\lambda_1+\lambda_3)^2}, \text{ and } I_{23} = \frac{1}{\lambda^2} + \frac{\lambda_1}{(\lambda_1+\lambda_2)(\lambda_2+\lambda_3)^2}.$$

The inverse of Fisher information matrix, $I^{-1}(\lambda_1, \lambda_2, \lambda_3) = (I^{ij})$: $(i, j) = 1,2,3$, can be obtained from $I(\lambda_1, \lambda_2, \lambda_3) = (I_{ij})$: $(i, j) = 1,2,3$. For H_0 : $\lambda_3 = 0$, they proposed the test statistic $\sqrt{n}\,\hat{\lambda}_3/\sqrt{\widehat{I^{33}}}$ which has asymptotic normal distribution with mean zero and variance one. For H_0 : $\lambda_3 = 0$ vs H_1 : $\lambda_3 > 0$, the test function of Hanagal and Kale(1991) is given by

$$\phi_{HK}(x, y) = \begin{cases} 1, & \text{if } \sqrt{n}\,\hat{\lambda}_3/\sqrt{\widehat{I^{33}}} > \Phi^{-1}(1-\alpha) \\ 0, & \text{otherwise} \end{cases}, \quad (2.8)$$

where $\widehat{I^{33}}$ is computed from I^{33} based on MLE's ($\hat{\lambda}_1$, $\hat{\lambda}_2$, $\hat{\lambda}_3$) and $\Phi(\cdot)$ is the cumulative distribution function of standard normal.

3. A Bootstrap Test of Independence

In this section we consider the bootstrap test for H_0 : $\lambda_3 = 0$ vs H_1 : $\lambda_3 > 0$ based on pivoting for which the asymptotic distribution does not depend on any unknown parameters. Since the asymptotic distribution of the statistic $\sqrt{n}\,\hat{\lambda}_3/\sqrt{\widehat{I^{33}}}$ under H_0 does not depend on any unknown parameter, we will use the statistic based on $\sqrt{n}\,\hat{\lambda}_3/\sqrt{\widehat{I^{33}}}$ as pivoting. The bootstrap testing procedure for H_0 : $\lambda_3 = 0$ vs H_1 : $\lambda_3 > 0$ can be described as follows:

(1) Obtain the MLE's $\hat{\lambda}_1$, $\hat{\lambda}_2$ and $\hat{\lambda}_3$ by solving equations (2.5), (2.6) and (2.7) simultaneously.

(2) Construct the sampling distribution function based on the MLE's, say, $ACBVE^*(\hat{\lambda}_1, \hat{\lambda}_2, \hat{\lambda}_3)$.

(3) Generate B random samples of size n from fixed $ACBVE^*(\hat{\lambda}_1, \hat{\lambda}_2, \hat{\lambda}_3)$. The corresponding samples called the *bootstrap samples* are denoted by $((x_1^{*b}, y_1^{*b}), (x_2^{*b}, y_2^{*b}), \cdots, (x_n^{*b}, y_n^{*b}))$, $b = 1, 2, \cdots, B$.

(4) Construct the likelihood equations (2.5), (2.6) and (2.7) based on bootstrap samples, that is, for $b = 1, 2, \cdots, B$,

$$-\sum_{i=1}^{n} x_i^{*b} + n/\hat{\lambda} - n/(\hat{\lambda}_1 + \hat{\lambda}_2) + n_1^{*b}/\hat{\lambda}_1 + n_2^{*b}/(\hat{\lambda}_1 + \hat{\lambda}_3) = 0 \quad (3.1)$$

$$-\sum_{i=1}^{n} y_i^{*b} + n/\hat{\lambda} - n/(\hat{\lambda}_1 + \hat{\lambda}_2) + n_2^{*b}/\hat{\lambda}_2 + n_1^{*b}/(\hat{\lambda}_2 + \hat{\lambda}_3) = 0 \quad (3.2)$$

$$-\sum_{i=1}^{n} \max(x_i^{*b}, y_i^{*b}) + n/\hat{\lambda} + n_1^{*b}/(\hat{\lambda}_2 + \hat{\lambda}_3) + n_2^{*b}/(\hat{\lambda}_1 + \hat{\lambda}_3) = 0, \quad (3.3)$$

where n_i^{*b} is bootstrap version of n_i, $i = 1, 2, 3$.

(5) Solve the equations (3.1), (3.2) and (3.3) simultaneously and obtain solutions $\hat{\lambda}_1^{*b}$, $\hat{\lambda}_2^{*b}$ and $\hat{\lambda}_3^{*b}$. And construct the Fisher information matrix based on $\hat{\lambda}_1^{*b}$, $\hat{\lambda}_2^{*b}$ and $\hat{\lambda}_3^{*b}$, say, $I(\hat{\lambda}_1^{*b}, \hat{\lambda}_2^{*b}, \hat{\lambda}_3^{*b}) = (\widehat{I_{ij}^{*b}})$: $(i, j) = 1,2,3$. Also, compute the inverse matrix of Fisher information, say, $I^{-1}(\hat{\lambda}_1^{*b}, \hat{\lambda}_2^{*b}, \hat{\lambda}_3^{*b}) = (\widehat{I^{ij*b}})$: $(i, j) = 1,2,3$, $b = 1, 2, \cdots, B$.

(6) Construct the bootstrap distribution function of $\sqrt{n}\,\hat{\lambda}_3/\sqrt{\widehat{I^{33}}}$ under H_0 : $\lambda_3 = 0$, say $\hat{H}^*$, as followings:

$$\hat{H}^*(s) = \frac{1}{B}\sum_{b=1}^{B} I\left(\sqrt{n}(\hat{\lambda}_3^{*b} - \hat{\lambda}_3)/\sqrt{\widehat{I^{33*b}}} \le s\right), \quad (3.4)$$

where $I(\cdot)$ is an indicator function and s is arbitrary real number.

Then we construct the bootstrap test function for H_0 : $\lambda_3 = 0$ vs H_1 : $\lambda_3 > 0$

as follows:

$$\phi_{BOOT}(x, y) = \begin{cases} 1, & \text{if } \sqrt{n}\,\hat{\lambda}_3/\sqrt{\hat{I}^{33}} > cp^* \\ 0, & \text{otherwise} \end{cases}, \tag{3.5}$$

where $cp^* = \hat{H}^{*-1}(1-\alpha) = \inf\{s : \hat{H}^*(s) \geq 1-\alpha\}$.

4. Monte Carlo Simulation Studies

In this section we compare the powers of the above three tests at the levels $\alpha = 0.05$ and 0.10 by simulating 1000 samples each of sizes $n = 5,\ 10,\ 15$ and 20 for the combination of values of the parameters $\lambda_1 = \lambda_2 = 0.05$ and 0.10 and $\lambda_3 = 0.05,\ 0.10$ and 0.15. These are presented in the Tables 1, 2 and 3.

We can summarize the following facts by inspection of Tables 1, 2 and 3.

(1) For the most cases, the test $\phi_{BOOT}(x, y)$ based on bootstrap procedure has the best power performance and the second best is a large sample test $\phi_{HK}(x, y)$ based on MLE. The worst case is the LRT $\phi_{GMM}(x, y)$.

(2) The powers of all tests tend to increase as sample size increase. In particular, the power of the test $\phi_{BOOT}(x, y)$ increases more quick than those of the tests $\phi_{GMM}(x, y)$ and $\phi_{HK}(x, y)$.

(3) For fixed $\lambda_1 = \lambda_2$ and fixed sample size, the powers of all tests tend to increase as the value of λ_3 increases.

(4) The powers of all tests tend to decrease as the common values of $\lambda_1 = \lambda_2$ increase.

Note that we used the restriction of marginal homogeneity for the comparison purpose. But the bootstrap test itself has no such restriction by construction.

<Table 1> The power of the tests $\phi_{GMM}(x,y)$, $\phi_{HK}(x,y)$ and $\phi_{BOOT}(x,y)$ for H_0 : $\lambda_3=0$ vs. H_1 : $\lambda_3=0.05$

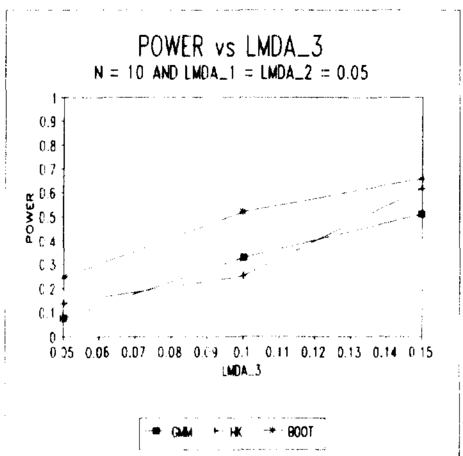
$\lambda_1=\lambda_2$	n	α	$\phi_{GMM}(x,y)$	$\phi_{HK}(x,y)$	$\phi_{BOOT}(x,y)$
0.05	5	0.05	0.0500	0.1370	0.2110
		0.10	0.0880	0.2490	0.2870
	10	0.05	0.0750	0.1900	0.2460
		0.10	0.1400	0.3240	0.3340
	15	0.05	0.1040	0.2320	0.2610
		0.10	0.1750	0.3570	0.3600
	20	0.05	0.1060	0.2420	0.2600
		0.10	0.1700	0.3750	0.3740
0.10	5	0.05	0.0180	0.0490	0.0540
		0.10	0.0530	0.1380	0.0910
	10	0.05	0.0290	0.0390	0.0490
		0.10	0.0590	0.0990	0.1020
	15	0.05	0.0370	0.0440	0.0800
		0.10	0.0790	0.0960	0.1260
	20	0.05	0.0350	0.0630	0.1130
		0.10	0.0820	0.1200	0.1640

<Table 2> The power of the tests $\phi_{GMM}(x,y)$, $\phi_{HK}(x,y)$ and $\phi_{BOOT}(x,y)$ for H_0 : $\lambda_3=0$ vs. H_1 : $\lambda_3=0.10$

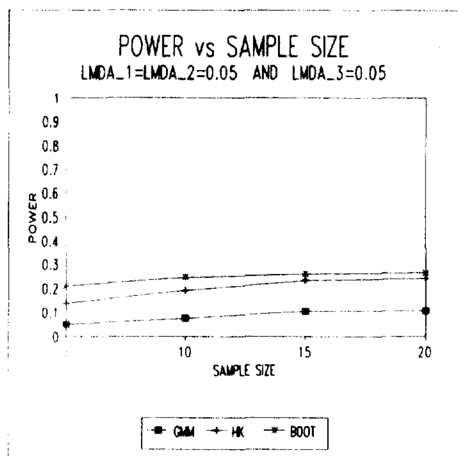
$\lambda_1=\lambda_2$	n	α	$\phi_{GMM}(x,y)$	$\phi_{HK}(x,y)$	$\phi_{BOOT}(x,y)$
0.05	5	0.05	0.1700	0.2500	0.3690
		0.10	0.2550	0.4150	0.4450
	10	0.05	0.3280	0.4650	0.5180
		0.10	0.4270	0.5860	0.5850
	15	0.05	0.5090	0.6230	0.6520
		0.10	0.5940	0.7220	0.7160
	20	0.05	0.6150	0.7110	0.7260
		0.10	0.6810	0.7900	0.7900
0.10	5	0.05	0.0520	0.1050	0.2140
		0.10	0.1000	0.2530	0.2820
	10	0.05	0.0750	0.1700	0.2160
		0.10	0.1230	0.2910	0.3000
	15	0.05	0.0910	0.2170	0.2570
		0.10	0.1490	0.3540	0.3580
	20	0.05	0.1100	0.2650	0.2940
		0.10	0.1900	0.3990	0.4020

〈Table 3〉 The power of the tests $\phi_{GMM}(x,y)$, $\phi_{HK}(x,y)$ and $\phi_{BOOT}(x,y)$ for H_0 : $\lambda_3=0$ vs. H_1 : $\lambda_3=0.15$

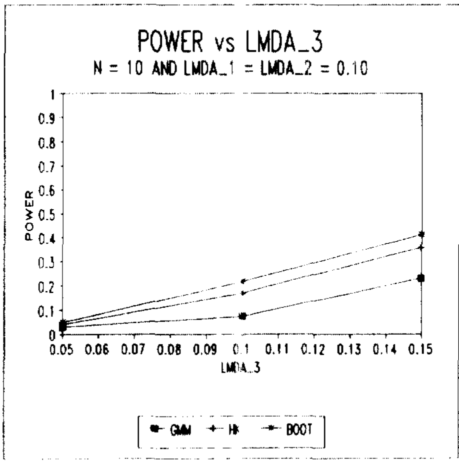
$\lambda_1=\lambda_2$	n	α	$\phi_{GMM}(x,y)$	$\phi_{HK}(x,y)$	$\phi_{BOOT}(x,y)$
0.05	5	0.05	0.2250	0.2840	0.4280
		0.10	0.3200	0.4550	0.4960
	10	0.05	0.5090	0.6160	0.6580
		0.10	0.6100	0.6860	0.6880
	15	0.05	0.6830	0.7390	0.7440
		0.10	0.7210	0.7800	0.7800
	20	0.05	0.7830	0.8140	0.8200
		0.10	0.8100	0.8350	0.8340
0.10	5	0.05	0.1130	0.1880	0.2980
		0.10	0.1840	0.3390	0.3760
	10	0.05	0.2300	0.3580	0.4120
		0.10	0.3100	0.4910	0.5030
	15	0.05	0.3100	0.4540	0.4900
		0.10	0.3840	0.5820	0.5870
	20	0.05	0.3480	0.5040	0.5230
		0.10	0.4420	0.6350	0.6330



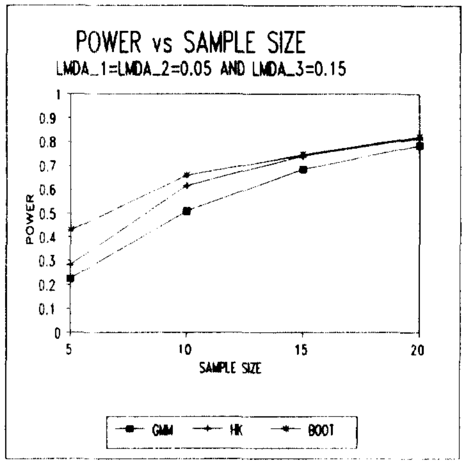
〈Figure 1〉 Plot of the power against λ_3 for $n=10$ and $\lambda_1=\lambda_2=0.05$



〈Figure 2〉 Plot of the power against n for $\lambda_1=\lambda_2=0.05$ and $\lambda_3=0.05$



〈Figure 3〉 Plot of the power against λ_3 for $n=10$ and $\lambda_1=\lambda_2=0.10$



〈Figure 4〉 Plot of the power against n for $\lambda_1=\lambda_2=0.05$ and $\lambda_3=0.15$

References

[1] Becher, H.(1993), "Reader reaction : Bootstrap hypothesis testing procedures," *Biometrics* 49, pp. 1268-1272.

[2] Bemis, B.M., Bain, L.J. and Higgins, J.J.(1972), "Estimation and hypothesis testing for the parameters of a bivariate exponential distribution," *Journal of the American Statistical Association*, Vol. 67, pp. 927-929.

[3] Beran, R.(1988), "Prepivoting test statistics : A bootstrap view of asymptotic refinements," *Journal of the American Statistical Association*, Vol. 83, pp. 687-697.

[4] Bhattacharyya, G.K. and Johnson, R.A.(1973), "On test of independence in abivariate exponential distribution," *Journal of the American Statistical Association*, Vol. 68, pp. 704-706.

[5] Block, H.W, and Basu, A.P.(1974), "A continuous bivariate exponential extension," *Journal of the American Statistical Association*, vol. 69, pp. 1031-1037.

[6] Efron, B.(1979), "Bootstap methods : Another look at the jacknife," *Annals of Statistics*, Vol. 7, pp. 1-26.

[7] Gupta, R.L., Metrotra, G.K. and Michalek, J.E.(1984), "A small sample test foran absolutely continuous bivariate exponential model," *Communications in*

Statistics, A13, pp. 1735-1740.

[8] Hall, P. and Wilson, S.R.(1991), "Two guidelines for bootstrap hypothesis testing," *Biometrics*, Vol. 47, pp. 757-762.

[9] Hanagal, D.D. and Kale, B.K.(1991), "Large sample tests of independence for absolutely continuous bivariate exponential distribution," *Communications in Statistics*, A20, pp. 1301-1313.

[10] Hinkely, D. V.(1988), "Bootstrap methods.(With discussion)," *Journal of the Royal Statistical Society*, Series B, Vol. 50, pp. 321-337.

[11] Hinkely, D. V.(1989), "Bootstrap Significance tests," *Proceeding 47th Session of the International Statistical Institute, Paris*, 29 August-6 September 1989, Vol 3, pp. 65-74.

[12] Marshall, A.W. and Olkin, I.(1967), "A multivariate exponential distribution, "*Journal of the American Statistical Association*, Vol. 62, pp. 30-44.