International Journal of Scientific and Innovative Mathematical Research (IJSIMR)
Volume 3, Issue 11, November 2015, PP 54-58
ISSN 2347-307X (Print) & ISSN 2347-3142 (Online)
www.arcjournals.org

A Test Procedure to Discriminate Between Probability Models

G.Prasad
Department of Statistics
Aditya Degree College
Rajahmundry, India
Bhaskarprasad22@gmail.com

R.R.L.Kantam
Department of Statistics
Acharya Nagarjuna University
Nagarjuna Nagar, India
kantam_rrl@rediffmail.com

Abstract: *A new method of generating probability distribution on the bases of given two specified probability models are adopted using the well-known Pareto, Rayleigh distributions. The resulting model is considered as null population and a test statistic is suggested to discriminate the null population between two successive alternative populations Pareto, Rayleigh models. The critical values of the test statistic and the powers are evaluated. A comparative study is presented.*

Keywords: *Pareto-Rayleigh distribution, Pareto distribution, Rayleigh distribution, order statistics, population quantiles, power function.*

1. Introduction

Let $F(x)$ be the cumulative distribution function (CDF) of any random variable X and $r(t)$ be the probability density function (PDF)of a random variable, T, defined on $[0,\infty)$. The CDF of the T-X family of distributions defined by Alzaatreh, et al. (2012) is given by

$$G(x)=\int_0^{-\log[1-F(x)]} r(t)dt \tag{1.1}$$

Alzaatreh, et al. (2012) named this family of distributions the Transformed-Transformer family (or T-X family).

If a random variable T follows the Pareto distribution type IV with parameter α then

$$r(t)=\alpha(1+t)^{-(\alpha+1)}; \quad t>0, \alpha>1 \tag{1.2}$$

If a random variable X follows the Rayleigh distribution with parameter σ then

$$F(x)=1-e^{-x^2/2\sigma^2}; \qquad x>0, \sigma>0 \tag{1.3}$$

Using (1.1), (1.2) and (1.3), the CDF of Pareto-Rayleigh distribution (as a member of T-X family) is given by

$$G(x)=1-\left[1+\frac{x^2}{2\sigma^2}\right]^{-\alpha}; \quad x>0, \alpha>1, \sigma>0 \tag{1.4}$$

The probability density function (pdf) corresponding to (1.4) is

$$g(x)=\frac{\alpha}{\sigma^2}x\left[1+\frac{x^2}{2\sigma^2}\right]^{-\alpha-1}; \quad x>0, \alpha>1, \sigma>0$$

where α is shape parameter and σ is scale parameter.

It is known that, methods of point estimation in Pareto-Rayleigh distribution do not yield closed form expression as estimators of their parameters. We are therefore motivated to explore the possibility, if any other population probability model be taken as an alternative to Pareto-Rayleigh distribution with a reasonable power of in distinguishability between the chosen pairs of distributions. This motivation leads us to the formulation of a problem in the testing of statistical hypothesis say "A given sample of a specified size belongs to a Pareto-Rayleigh distribution". In this sense Pareto-Rayleigh becomes a null population for the sample under consideration with another suitable alternative. The procedure of identifying the given sample with Pareto-Rayleigh distribution or an alternative distribution with small risks and reasonably larger powers will be discussed in this paper. Such studies of discriminatory problems between probability models are made by Gupta *et al.* (2002), Gupta and Kundu (2003a), Gupta and Kundu (2003b), Kundu and Gupta (2004a, 2004b),Kundu and Manglick (2004), Kundu *et al.* (2005), Kundu and Manglick (2005), Kundu (2005), Kundu and Raqab (2007), Arabin and Kundu (2009), Arabin and Kundu (2010), Arabin and Kundu (2012a), Arabin and Kundu (2012b) and the references therein. Recently Sultan (2007) developed a test criterion to distinguish generalized exponential distribution from Weibull, Normal distributions using moments of order statistics in samples drawn from generalized exponential distribution. Also, Sultan (2007), Srinivasa Rao and Kantam (2013), Srinivasa Rao and Kanatam (2014), Kantam *et al.* (2014) are the articles in which similar hypotheses are tested based one test criteria using moments of order statistics and population quantiles.

In this paper we adopt the approach namely population quantiles test procedure to distinguish between Pareto-Rayleigh and Pareto, Pareto-Rayleigh and Rayleigh. A description of the test procedure of population quantiles and its application is presented in Section - 2.The quantile based approach is adapted to our chosen null and alternative populations to get the critical values and power of the test criteria in Section - 3. In all these cases the percentiles of respective test statistics and the powers of the test criteria values of various sample sizes evaluated numerically are tabulated in Section - 4. A comparative study of the proposed test criteria in the case of the respective chosen null and alternative populations is presented in Section - 5.

2. Quanitle Based Test Procedure

Let F_1, F_2 denote two probability distributions of continuous type to be regarded as null, alternative populations respectively. Let $x_1, x_2, x_3, ..., x_n$ be a complete ordered random sample of size n (ordering is mandatory as the method indicates). Here we test the Null hypothesis.

We want to test the null hypothesis

H_0: The sample has come from the population F_1 against the alternative hypothesis

H_1: The sample has come from the population F_2.

Sultan (2007) suggested a test statistic given by the formula

$$T = \frac{\sum_{i=1}^{n} \xi_i x_{(i)}}{\sqrt{\sum_{i=1}^{n} \xi_i^2 \sum_{i=1}^{n} x_i^2}} \tag{2.1}$$

where $x_{(i)} - i^{th}$ ordered observation in the sample and ξ_i - the expected value of - i^{th} standard order statistic in a sample of size n from null population.

If $x_1, x_2, x_3, ..., x_n$ is truly a sample from the null population the formula for *'T'* would serve as a test statistic to discriminate a null population and the corresponding alternative population with the help of its critical values.

Hence, the sampling distribution of *'T'* and its percentiles therefrom are essential to make use of the test statistic *'T'*. Since the sampling distribution of *'T'* involves non-linear functions of order statistics it cannot be generally tractable analytically in the case of all populations, necessitating one to go for empirical sampling distributions, empirical percentiles of *'T'* through simulation. Sultan (2007) did the same with generalised exponential distribution as null population.

The statistic *'T'* defined in (2.1) is based on moments of order statistics in samples from null population which may not always be available in all cases. This motivated Srinivasa Rao and Kantam (2013) to suggest the statistic *'T'* with ξ_i replaced by the corresponding population quantiles which are available analytically for all populations with invertible cumulative distribution functions. They further established T based on population quantiles is more powerful than that based on moments of order statistics in their selected models. We adopt the same in the present paper for Pareto-Rayleigh as null population, because moments of order statistics for these two null populations are not available either analytically or computationally for ready use.

3. Pareto-Rayleigh Distribution VS Pareto/Rayleigh Distributions using Quantile Based Test Procedure

Let $x_1, x_2, x_3, ..., x_n$ be a complete ordered random sample of size n. Here we test the null hypothesis.

H_0: The sample has come from Pareto-Rayleigh distribution (σ = 1, α = 2, 3, 4) against one of the two alternative hypotheses.

(i) H1: The sample has come from Pareto Distribution with σ = 1, α = 2, 3, 4.
(ii) H1: The sample has come from Rayleigh distribution with σ =1

Proceeding on the same lines as described earlier, our proposed test statistic is

$$T = \frac{\sum_{i=1}^{n} \xi_i x_{(i)}}{\sqrt{\sum_{i=1}^{n} \xi_i^2 \sum_{i=1}^{n} x_i^2}} \tag{3.1}$$

where $\xi_i = F^{-1}(p_i)_i = \sqrt{2\sigma^2\left[1-p_i^{\frac{-1}{\alpha}}-1\right]}$ *with* $p_i = \frac{i}{n+1}$ is the i^{th} population quantile in Pareto-Rayleigh distribution and $x_{(i)} - i^{th}$ ordered observation in the sample.

As mentioned earlier we have tabulated the percentiles of empirical sampling distribution of '*T*' of equation (3.1) for n = 5 (5) 25, σ = 1, α =2, 3, 4 through 10,000 Monte-Carlo simulation runs and are given in Table – 3.1.

The percentiles of *T* would serve as critical values to test null hypothesis that a given sample comes from Pareto-Rayleigh distribution. The power of this test statistic with Pareto (σ =1, α = 2, 3, 4) and Rayleigh distributions (σ = 1) as alternatives are evaluated accordingly and the values are given in Table 3.2 and Table 3.3 respectively for n=5(5)25.

4. Tables

The percentiles of 'T' values of Pareto-Rayleigh distribution, Power of statistic T with Pareto and Rayleigh distribution are evaluated and are presented in the following tables.

Table 3.1. The percentiles of 'T': Pareto –Rayleigh distribution with σ = 1, α =2, 3, 4

	Pi \ n	0.9987	0.995	0.99	0.975	0.95	0.90	0.10	0.05	0.025	0.01	0.005	0.00135
2	5	0.99951	0.99893	0.99858	0.99770	0.99667	0.99513	0.94678	0.92202	0.89741	0.85452	0.82775	0.79678
	10	0.99881	0.99837	0.99800	0.99732	0.99659	0.99542	0.95235	0.92535	0.89403	0.85608	0.82959	0.76414
	15	0.99880	0.99839	0.99810	0.99748	0.99690	0.99593	0.95570	0.92862	0.89344	0.84646	0.81545	0.73574
	20	0.99883	0.99846	0.99821	0.99771	0.99717	0.99633	0.95697	0.92951	0.89703	0.84973	0.80078	0.75067
	25	0.99891	0.99859	0.99830	0.99791	0.99742	0.99666	0.95953	0.93288	0.89921	0.84716	0.80330	0.74303
3	5	0.99953	0.99896	0.99857	0.99784	0.99690	0.99553	0.95408	0.93464	0.91464	0.88027	0.85417	0.82084
	10	0.99883	0.99846	0.99813	0.99756	0.99695	0.99596	0.96370	0.94604	0.92553	0.89703	0.87726	0.81940
	15	0.99895	0.99853	0.99826	0.99777	0.99728	0.99650	0.96903	0.95254	0.93082	0.90022	0.87473	0.81772

	20	0.99900	0.99868	0.99844	0.99800	0.99758	0.99692	0.97201	0.95554	0.93685	0.90821	0.87275	0.83488
	25	0.99910	0.99876	0.99855	0.99818	0.99781	0.99725	0.97416	0.95957	0.93975	0.91157	0.88474	0.83810
4	5	0.99950	0.99898	0.99862	0.99791	0.99706	0.99572	0.95733	0.94037	0.92158	0.89239	0.86766	0.83516
	10	0.99888	0.99850	0.99820	0.99769	0.99709	0.99619	0.96856	0.95380	0.93849	0.91512	0.90046	0.85141
	15	0.99902	0.99858	0.99836	0.99789	0.99747	0.99674	0.97369	0.96159	0.94562	0.92283	0.90456	0.86399
	20	0.99906	0.99873	0.99854	0.99816	0.99776	0.99716	0.97737	0.96558	0.95177	0.92994	0.90708	0.87521
	25	0.99913	0.99887	0.99864	0.99833	0.99800	0.99750	0.97968	0.96871	0.95562	0.93530	0.91846	0.87921

Table 3.2. *Power of T Statistic: Pareto-Rayleigh distribution Vs Pareto distribution*

	Level of Significance α								
	0.10			0.05			0.01		
n	α = 2	α = 3	α = 4	α = 2	α = 3	α = 4	α = 2	α = 3	α = 4
5	0.0020	0.0023	0.0021	0.0100	0.0106	0.0104	0.0191	0.0195	0.0197
10	0.0001	0.0003	0.0003	0.0017	0.0015	0.0013	0.0046	0.0035	0.0031
15	0.0000	0.0000	0.0000	0.0002	0.0001	0.0001	0.0007	0.0006	0.0005
20	0.0000	0.0000	0.0000	0.0002	0.0000	0.0000	0.0005	0.0004	0.0004
25	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Table 3.3. *Power of T Statistic: Pareto-Rayleigh distribution Vs Rayleigh distribution*

	Level of Significance α								
	0.10			0.05			0.01		
n	α = 2	α = 3	α = 4	α = 2	α = 3	α = 4	α = 2	α = 3	α = 4
5	0.0122	0.0162	0.0160	0.0632	0.0631	0.0618	0.1217	0.1187	0.1161
10	0.0114	0.0128	0.0118	0.0555	0.0559	0.0563	0.1141	0.1143	0.1143
15	0.0080	0.0109	0.0112	0.0416	0.0528	0.0560	0.0939	0.1139	0.1188
20	0.0069	0.0083	0.0097	0.0341	0.0518	0.0572	0.0783	0.1078	0.1184
25	0.0062	0.0106	0.0121	0.0283	0.0484	0.0564	0.0611	0.0988	0.1146

5. CONCLUSION

To test each of the above null hypotheses against the respective alternative the proposed test statistics need their percentiles to serve as critical values for accepting or otherwise the hypothesis. The percentiles of the statistic based on quantiles of null population do not require the specification of the alternative population as described in its procedure in Section – 2. Also, these percentiles are invariant of scale transformations of the population. Accordingly two tables are adequate for presenting them. These are given in Table 3.1 and the power of this test statistic for the chosen null Vs alternatives are given in Table 3.2 and Table 3.3.

In order to test Pareto-Rayleigh Vs Pareto and Pareto-Rayleigh Vs Rayleigh the quantile approach giving almost the same and poor powers as evidenced from Tables 3.2 and 3.3 (reproduced on the following page) ranging from 0.000 to 0.0197 of Pareto distribution and from point 0.0062 to 0.1217 of Rayleigh distribution. We may therefore conclude that Pareto distribution and Rayleigh distribution can be a reasonable alternative to Pareto-Rayleigh distribution. Thus we may arrive at the following broad conclusions on the basis of our proposed test.

REFERENCES

[1]. Alman Alzaatreh, Felix Famoye and Carl Lee (2013). Weibull-Pareto Distribution and its Applications. *Communications in Statistics – Theory and Methods 42:9*, Vol.13, 1673-1691.

[2]. Kundu, D., and Raqab, M. Z. (2005). Generalized Rayleigh Distribution: Different Methods of Estimation. *Computational Statistics & Data Analysis*, Vol. 49, 187-200.

[3]. R.D., D. Kundu, D., and Manglick, A. (2002). Probability of Correct Selection of Gamma Versus GE or Weibull vs. GE Models Based on Likelihood Ratio Test. *Recent Advances in Statistical*

Methods, Editor Y.P. Chaubey, Publisher World Scientific Publishing Company Inc., London, 147-156.

[4]. Alman Alzaatreh, Felix Famoye and Carl Lee, (2012): Gamma-Pareto Distribution and its Applications, *Journal of Modern Applied Statistical Methods*, Vol. 11, No. 1, 78-94.

[5]. Kantam, R.R.L., Priya, Ch.M. and Ravikumar, M.S. (2014). Discrimination between Linear Failure Rate Distribution and Rayleigh Distribution. *Journal of Reliability and Statistical Studies*, Vol.7, Special Issue (2014), 9-17.

[6]. Srinivasa Rao, B., and Kantam, R.R.L. (2014). Discriminating Between log-logistic and Rayleigh Distributions. *Pakistan Journal of Statistics and Operations Research*, Vol.X, No.1, 1-7.

[7]. Sultan, K.S. (2007). Order Statistics from the Generalised Exponential Distribution and Applications. *Communications in Statistics – Theory and Methods*, Vol. 36, No. 7, 1409-1418.

[8]. Kundu, D., and Manglick, A. (2005). Discriminating Between the gamma and log- normal distributions, *Journal of Applied Statistical Sciences*, Vol. 14, No. 1 & 2, 175-187.

[9]. Kundu, D. (2005). Discriminating between the Normal and Laplace Distributions. *Advances in Ranking and Selection, Multiple Comparisons, and Reliability*, N. Balakrishnan, H. N. Nagaraja and N. Kannan (eds.), Birkhauser, Boston, 65-78.

[10]. Gupta, R.D., and Kundu, D. (2003b). Closeness of Gamma and Generalized Exponential Distributions, *Communications in Statistics – Theory and Methods*, Vol. 32, No.4, 705- 721

AUTHOR'S BIOGRAPHY

Grandhi Prasad,

Head of the Department,

Department of Statistics,

Aditya Degree College.

Rajahmundry, A.P. India.

Dr. R.R.L.Kantam

Department of Statistics,

Acharya Nagarjuna University,

Nagarjuna Nagar, India.