GOODNESS OF FIT TESTS BASED ON GAMMA DISTRIBUTION UNDER VARIOUS ALTERNATIVES

Tripakshi Borthakur
Assistant Professor,
Department of Statistics, D.R. College, Golaghat (Assam)
Email: borthakur.tripakshi06@gmail.com, Contact no: 6000429627

Abstract

The goodness of fit tests is an analytical procedure for testing the assumptions underlying the model which we are trying to fit to the data. These tests illustrate how well the distribution we selected fits to our data. The gamma distribution is widely used in the areas of engineering, economics, sciences, industry, business etc. It is quite important to study the assumption of gamma distribution as it has gained wide applicability in various fields. Various goodness of fit tests are available for testing the assumption of gamma distribution. Most commonly used tests are the tests based on EDF (Empirical Distribution Function) tests, such as , Anderson-Darling, Kolmogorov-Smirnov, Cramer-von Mises test. The aim of this article is to study the performance of some goodness of fit tests for gamma distribution by considering three different sample sizes against three different alternative hypotheses. Different tables are made which display the results obtained by Monte Carlo Simulation technique. Finally, conclusions are made.

Keywords: Anderson-Darling test, Goodness of Fit, Kolmogorov-Smirnov test, Monte-Carlo simulation, power, Watson test.

1. INTRODUCTION

The goodness of fit test is used to identify the appropriate distribution to be fitted for a given data. It describes how well it fits a set of observations. The best way to verify the compatibility of a set of observed sample values with a normal or any other distribution by the help of goodness of fit tests. It summarizes the discrepancy between the observed and expected values under a statistical model. The goodness of fit statistics are goodness of fit indices with known sampling distributions that are used in statistical hypothesis testing. It shows how well the distribution we selected fits to our data. These tests can be used to test whether the fitted model may have generated the observed data. In order to evaluate the performance of a goodness of fit test, it is necessary to study the power of the proposed method against some specific alternatives.

Goodness of fit testing may be regarded as a diagnostic procedure in order to test the assumptions underlying the model that one may try to fit to the data. The goodness of fit test has gained wide applicability in recent times in various fields like in finance, in decision making science, and in many other fields.

As an important part of data analysis, it is quite essential to study the goodness of fit technique which ensures that the data come from a particular family of distribution. Various goodness of fit tests have been found available to solve the problem of goodness of fit, out of which the most commonly used goodness of fit test is the Chi-Square test of goodness of fit which is mostly used to estimate the fit quality of binned data. But, it should not be considered as the ultimate solution for every goodness of fit problem as it is neither capable nor expected to detect fit inefficiencies for all possible problems. The K-S test proposed by Kolmogorov and Smirnov (1933), test given by Anderson- Darling (1952, 1954), Watson test (1961) etc. are another most commonly used goodness of fit tests which overcome the drawbacks of Chi-square test and hence can be applied in different situations which provide better results as compared to the Chi-Square test. While conducting a goodness of fit test, the researcher often wants to retain the null hypothesis. In other words, they want to demonstrate that a sample follows a specific probability distribution (e.g., a normal or gamma distribution). In employing the inferential tests, the researchers want to reject the null hypothesis, i.e., they want to exhibit that one or more samples do not come from the same population. It should be noted that, if the null hypothesis is rejected, the alternative hypothesis for a goodness of fit test does not specify an alternative distribution that would become the most likely distribution for the data. There are several tests of goodness of fit available in literature which can be applied in different situations proposed by various authors have been highlighted below:

- Test developed by Shapiro and Wilk (1965) for testing a complete sample for normality.
- Test given by Dahiya and Gurland (1972) based on generalized minimum chi-square for the Gamma and Exponential distributions.
- Locke (1976) developed a rank test for testing two parameter Gamma distribution against general alternatives.

- A test for exponentiality based on the Gini's index introduced by Gail and Gastwirth (1978).
- A test based on the empirical distribution of the correlation coefficient of the points on a probability plot with points found by using moment type estimates of the parameters proposed by Vogel and McMartin (1991).
- A modified Greenwood statistic for testing the gamma hypothesis proposed by Shapiro and Chen (2001).
- Wilding and Mudholkar (2008) considered a test for gamma hypothesis based on characteristic independence between the mean and coefficient of variation.
- Noughabi, Arghami and Borzadaran (2014) proposed a general goodness of fit test based on the estimated Gini index.
- Goodness of fit tests based on a new characterization of Pareto distribution proposed by Obradovic, Jovanovic and Milosevic (2015).

The gamma distribution which is also known as Type III Pearson's system of distribution has been widely used to model non-negative measurements that reveal unimodality and right skewness and can be applied in the areas of engineering, sciences, industry, economics and business. Earlier, its applications were mainly involved as a derived distribution and its relationship to the chi-squared distribution, but in recent times, it has gained its importance as a population model. Different authors gave its applications in different fields. Some of which were given by McDonald and Jensen (1979), Bougeault (1982), McCullagh and Lang (1984), Matis et al. (1992), Reiser and Rocke (1993), Yang (1994), Tan (1995), Yeh (1997) which comprise of geological studies, ecology, inventory control and queuing problems, economic studies, meteorology, reliability assessments, computer evaluations and biological studies. Besides these, it can be recommended as the failure time model in the field of econometrics. It is also used to model the size of insurance claims, rainfall and the errors in multi-level Poisson regression models. As a result, it is quite essential to assess the validity of the assumption of gamma distribution before applying it in any field using appropriate goodness of fit tests.

In this paper, some well established goodness of fit tests have been considered to study the validity of gamma hypothesis by the use of simulation technique. The description of each of the selected tests for gamma distribution, the effects on the power of the tests due to the sample sizes and the type of alternative distributions are also included and are given in different sections below.

2. GOODNESS OF FIT TESTS FOR GAMMA DISTRIBUTION:

Because of extensive use of the gamma models in the areas of econometrics, engineering, biomedical sciences and various other fields, the importance of goodness of fit for gamma distribution has also increased. But, only a few goodness of fit tests are available in literature that are used to test if the data indeed come from a gamma population. Many of the existing approaches for testing the gamma hypothesis usually involve the adaption of empirical distribution function tests of uniformity, using the probability integral transformation. The tests of uniformity that are used in this approach include the work of Kolmogorov (1933) and the members of Cramer-von Mises family such as the tests given by Anderson-Darling (1954). In this article, we have considered three such tests of goodness of fit which are defined below:

3. TEST PROCEDURES:

Let $X_1, X_2, .., X_n$ be independent and identically distributed random variables having the gamma density

$$f(x;\alpha,\beta) = \frac{1}{\alpha\ \Gamma\beta}\left(\frac{x}{\alpha}\right)^{\beta-1}\exp\left[\frac{-x}{\alpha}\right], x>0, \quad \alpha, \quad \beta>0$$

where α and β are respectively the scale parameter and the shape parameter. In this study, we have considered three empirical distribution function test statistics and the test procedures are given below:

3.1 Tests based on Empirical Distribution Function (EDF):

3.1.1 Anderson-Darling test:

The Anderson-Darling test for Gamma distribution (Stephens, 1986), based on the statistic

$$A^2 = -n - \frac{1}{n}\sum_{i=1}^{n}(2i-1)\left[\ln Z_i + \ln(1-Z_{n+1-i})\right] \tag{3.1}$$

where $Z_i = F\left(X_{(i)};\ \alpha,\ \beta\right)$ ($i=1,2,...,n$) with F denoting the c.d.f. of Gamma distribution. For testing the gamma hypothesis, the value of A^2 is compared with the empirical critical values. In case the value of A^2 exceeds the critical value at corresponding significance levels, the gamma hypothesis may be rejected.

3.1.2 Kolmogorov-Smirnov test:

The Kolmogorov- Smirnov statistic denoted by *D*, introduced by Kolmogorov (1933) is given by the formula

$$D = \max\left\{D^{+}, D^{-}\right\} \tag{3.2}$$

where
$$D^{+} = \max_{i=1,2,\ldots,n}\left(\frac{i}{n} - Z_i\right), \tag{3.3}$$

$$D^{-} = \max_{i=1,2,\ldots,n}\left(Z_i - \frac{(i-1)}{n}\right) \tag{3.4}$$

and $Z_i\,'s$ have the same meaning defined in section 3.1.1 above. The value of *D* is compared with its critical values for taking the decision about the null hypothesis. The null hypothesis is rejected when the value of *D* goes beyond the simulated critical value, at corresponding significance levels.

3.1.3 Watson test:

The Watson test for gamma distribution (1961) is given by the statistic

$$U^2 = W^2 - n(\overline{Z} - 0.5) \tag{3.5}$$

$$W^2 = \sum_{i=1}^{n}\left(Z_i - \frac{2i-1}{2n}\right)^2 + \frac{1}{12n} \tag{3.6}$$

where, W^2 is the Cramer-von Mises statistic defined in (3.6) and $\overline{Z}$ is the mean of $Z_i\,'s$ defined in section 3.1.1 above. If the value of U^2 exceeds the empirical critical value of the test statistic, then the gamma hypothesis is rejected at corresponding significance levels.

4. TYPES OF ALTERNATIVES:

In this article, we have considered three different statistical distributions, viz., Weibull, Lognormal and Rayleigh distribution, whose density functions are defined below:

4.1 Weibull Alternative:

The p.d.f. of Weibull distribution is given by

$$f(x) = \frac{\eta}{\sigma^{\eta}} x^{\eta-1} \exp\left[-\left(\frac{x}{\sigma}\right)^{\eta}\right], \; 0 \le x < \infty \;, \eta > 0, \; \sigma > 0$$

4.2 Lognormal Alternative:

The p.d.f. of Lognormal distribution is given by

$$f(x) = \frac{1}{x\sigma\sqrt{2\pi}} exp\left[-\frac{1}{2\sigma^2}(logx - \mu)^2\right], \; 0 < x < \infty, \sigma > 0$$

4.3 Rayleigh Distribution:

The p.d.f. of Rayleigh distribution is given by

$$f(x;\sigma) = \frac{x}{\sigma^2} e^{-\frac{x^2}{2\sigma^2}}, \; x \ge 0$$

5. SIMULATION STUDY:

In this article, Monte- Carlo simulation technique has been adopted in order to estimate the power of the selected tests for gamma distribution. For three different sample sizes (n=10, 25, 50) by considering α=0.05 level of significance against the selected alternative distributions mentioned in Section 4 above, the effects on the power of the tests due to the sample size and the type of the alternative distributions are also shown in different tables with the help of simulation method. Here, the Uniform (0,1) variates are generated by RND function using QBASIC and for the other distributions, method of inverse integral transformation of uniform variates is used. Normal deviates are generated by using Box-Muller (1958) formula which is used to generate lognormal variates. For each result 10,000 repetitions are made. The ratio of the value of the test statistic greater than the critical value divided by the total number of repetition gives the empirical level under null case and power of the test statistics under the alternative hypothesis.

6. RESULTS AND DISCUSSIONS:

Table 1: Empirical Power of the Tests under Weibull Alternative (α=0.05)

Sample Size(n)	(η,σ)	Test Statistics		
		AD **(** A^2 **)**	***KS(D)***	***Watson*** **(** U^2 **)**
10	**(1,1)**	.6175	.2665	.3903
	(2,1)	.0645	.0550	.0588
	(3,2)	.9835	.9591	.8363
	(4,2)	.9976	.9949	.9637
25	**(1,1)**	.8963	.5494	.7768
	(2,1)	.1117	.0813	.0675
	(3,2)	1.000	.9999	.9973
	(4,2)	1.000	1.000	1.000
50	**(1,1)**	.9935	.8377	.9766
	(2,1)	.1861	.1159	.0859
	(3,2)	1.000	1.000	1.000
	(4,2)	1.000	1.000	1.000

Table 2: Empirical Power of the Tests under Lognormal Alternative (α=0.05)

Sample Size(n)	(μ, σ)	Test Statistics		
		AD **(** A^2 **)**	***KS(D)***	***Watson*** **(** U^2 **)**
10	**(0, .2)**	.1569	.3854	.9669
	(0, .4)	.0596	.0959	.1949
	(0, .5)	.0892	.0826	.0830
	(0, .7)	.2252	.1083	.0842
	(0, .9)	.4320	.1660	.2213
	(0, 1)	.5442	.2000	.3118
25	**(0, .2)**	.9997	.9952	1.000
	(0, .4)	.2676	.2986	.4819
	(0, .5)	.2004	.1631	.1273
	(0, .7)	.3727	.1784	.1157

	(0, .9)	.7025	.3206	.4609
	(0, 1)	.8346	.4109	.6433
50	**(0, .2)**	1.000	1.000	1.000
	(0, .4)	.7690	.6743	.9209
	(0, .5)	.4424	.3255	.2326
	(0, .7)	.5995	.2929	.2038
	(0, .9)	.9265	.5666	.7804
	(0, 1)	.9830	.7015	.9267

Table 3: Empirical Power of the Tests under Rayleigh Alternative (α=0.05)

Sample Size(n)	σ^2	**Test Statistics**		
		AD (A^2)	***KS(D)***	***Watson*** (U^2)
10	**0.2**	.8375	.6402	.4822
	0.5	.0645	.0550	.0588
	0.6	.0426	.0477	.0568
	0.8	.1181	.1181	.0940
	1.0	.2880	.2462	.1639
25	**0.2**	.9993	.9798	.9262
	0.5	.1116	.0813	.0675
	0.6	.0500	.0494	.0639
	0.8	.2309	.2237	.1627
	1.0	.5711	.5007	.3415
50	**0.2**	1.000	1.000	.9998
	0.5	.1861	.1159	.0859
	0.6	.0542	.0467	.0795
	0.8	.3960	.4785	.2938
	1.0	.8504	.7804	.6385

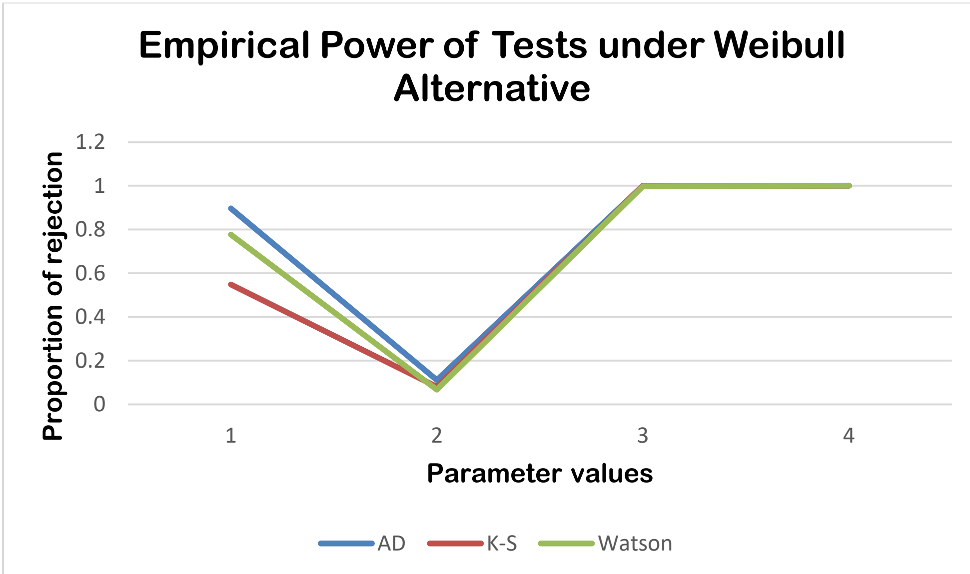

Figure 1: Empirical Power of the Tests under Weibull Alternative (n=25, α=0.05)

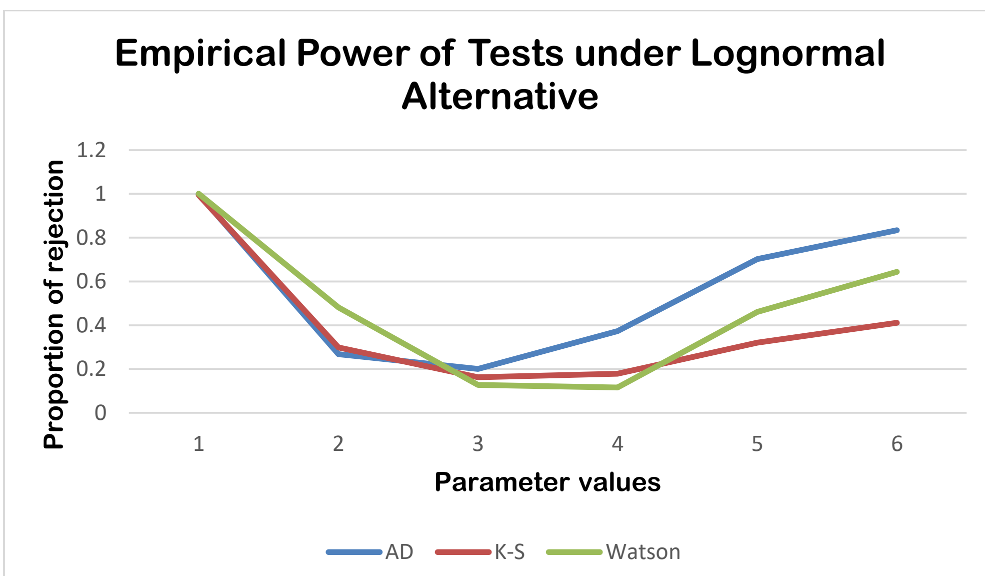

Figure 2: Empirical Power of the Tests under Lognormal Alternative (n=25, α=0.05)

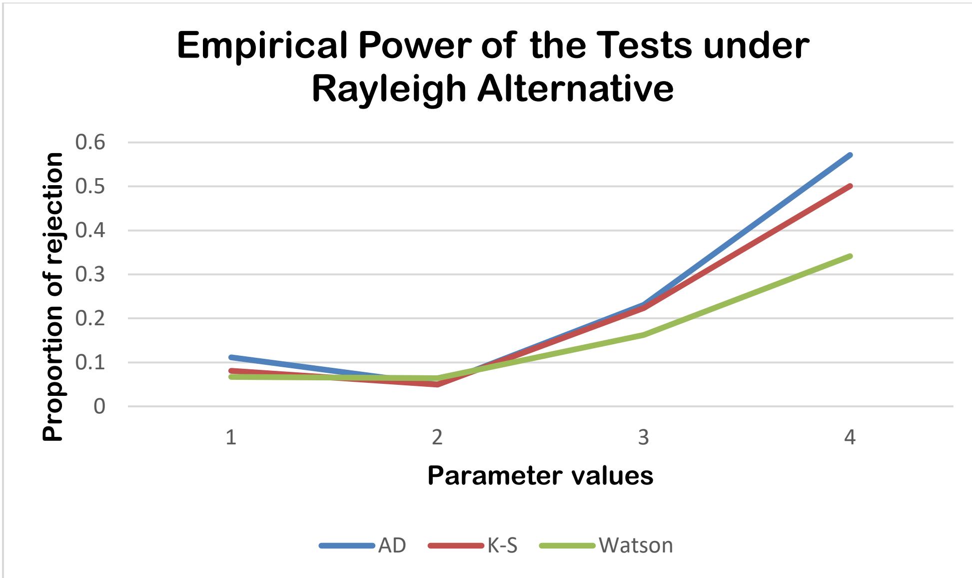

Figure 3: Empirical Power of the Tests under Rayleigh Alternative (n=25, α=0.05)

Table 1 depicts the empirical power of the selected tests under Weibull alternative for four different set of the parameters. It is being observed for this alternative that the power of all the tests increases with the increase in the sample sizes. Also, for the parameter set (2,1), the empirical power of all the tests become very low as compared to the other parameter sets viz., (1,1), (3,2) and (4,2) as this set of parameter give the shape of Weibull distribution like the gamma distribution. However, out of these three tests, the power of *AD* (A^2) test is seem to be more powerful than the power of the other two tests in all the situations. The power of Watson (U^2) and *K-S* (D) tests are found to be lower than *AD* (A^2) test for small values of the parameters and finally, for large values of the parameters as well as for large sample sizes, the power of these two tests also come closer to the *AD* (A^2) test.

Table 2 represents the empirical power of the three tests under the alternative of Lognormal distribution for six different sets of the parameters. Though, the power of all the tests increases with the increase in the sample sizes yet, an ups and down has been found for the empirical power of the tests with respect to the parameter values. For instance, the empirical power of the three tests for the parameter sets (0,0.4), (0,0.5) and (0,0.7) have been gradually decreasing as compared to the parameter set (0,0.2) and again it increases for the parameter sets (0,0.9) and (0,1). For this alternative, it is being observed that, for small values of the parameters, the Watson (U^2) test seems to be more powerful than the other tests. But as the sample size as well as the values of the parameter increases, the *AD* (A^2) test gives comparatively higher power.

Table 3 shows the empirical power of the three tests under Rayleigh alternative five different values of the parameter. It is being observed for this alternative that the power of all the tests increases with the increase in the sample sizes. Here also, an ups and down for the empirical power of all the three tests have been found for this alternative with respect to the parameter values. For instance, the power of the selected tests gradually decreases and becomes very low for the parameter values σ^2=0.5 and σ^2=0.6 than the power found for the parameter value σ^2=0.2. Again, the power of the tests increases for the parameter values σ^2=0.8 and σ^2=1.0 as compared to the above mentioned parameter values. However, the empirical power of *AD* (A^2) test is found to be higher than the power of the other tests followed by *K-S* (*D*) test, whereas the overall power of Watson test is found to be lower as compared to the others.

7. CONCLUSION:

From the above discussions, it may be concluded that the *AD* (A^2) test is more preferable as compared to the other two tests as it has shown greater power in all the situations. The power of *K-S* (*D*) test is also found to be reasonably good, whereas, the Watson test displays comparatively lower power for Rayleigh alternative in all the situations considered here. Finally, we arrive at the conclusion that the Anderson-Darling (A^2) test may be recommended for testing the goodness of fit for gamma distribution against the alternatives that have been considered for the study. In the absence of this test, we may prefer the *K-S* (*D*) test.

REFERENCES:

[1] Anderson, T.W. and Darling, D.A. (1952): Asymptotic theory of certain "goodness of fit" criteria based on stochastic processes, The Annals of Mathematical Statistics, 23(2), 193-212.

[2] Anderson, T.W. and Darling, D.A. (1954): A test of goodness of fit, Journal of the American Statistical Association, 49(268), 765-769.

[3] Bakouch, H., Saboor, A., Chesneau, C., and Saeed, A. (2017): A new lifetime model by mixing gamma and geometric distributions with hydrological data, https://hal.archives-ouvertes.fr/hal-01536094.

[4] Bougeault, P. (1982): Cloud-ensemble relations based on the gamma probability distribution for the higher-order models of the planetary boundary layer, Journal of the Atmospheric Sciences, 39(12), 2691-2700.

[5] Box, G.E.P and Muller, M.E.A. (1958): A note on generation of normal deviates, The Annals of Mathematical Statistics, 29, 610-611.

[6] Chen, G. and Balakrishnan, N. (1995): A general purpose approximate goodness of fit test, Journal of Quality Technology, 27(2), 154-160.

[7] Choi, S.C. and Wette, R. (1969): Maximum likelihood estimation of the parameters of the gamma distribution and their bias, Technometrics, 11(4), 685-690.

[8] Cohen, A.C. and Whitten, B.J. (1986): Modified moment estimation for the three parameter Gamma distribution, Journal of Quality Technology, 18(1), 53-62.

[9] Cohen, A.C. and Whitten, B.J. (1988): Parameter estimation in Reliability and life span Models, Marcel Dekker, Inc., New York.

[10] D'Agostino, R.B. and Stephens, M. (1986): Goodness of fit technique, Marcel Dekker, Inc., New York.

[11] Dahiya, R.C. and Gurland, J. (1972): Goodness of fit tests for the gamma and exponential distributions, Technometrics, 14(3), 791-801.

[12] DiCiccio, T.J. (1987): Approximate inference for the generalized gamma distribution, Technometrics, 29(1), 33-40.

[13] Filliben, J.J. (1975): The probability plot correlation coefficient test for normality, Technometrics, 17(1), 111-117.

[14] Gail, M. H., and Gastwirth, J. L. (1978): A scale-free goodness-of-fit test for the exponential distribution based on the Gini statistic, Journal of the Royal Statistical Society, Series B (Methodological), 40(3), 350-357.

[15] Kallioras, A.G., Koutrouvelis, I.A. and Canavos, G.C. (2006): Testing the fit of gamma distributions using the empirical moment generating function, Communication in Statistics-Theory and Methods, 35(3), 527-540.

[16] Kirimi, E., Ouko, A. and Kipkoech, C.W. (2014): Modified moment estimation for a two parameter gamma distribution, IOSR Journal of Mathematics, 10(6), 42-50.

[17] Koutrouvelis, I.A. and Canavos, G.C. (1997): Estimation in the three parameter gamma distribution based on the empirical moment generating function, Journal of Statistical Computation and Simulation, 59(1), 47-62.

[18] Kolmogorov, A.N. (1933): Sulla determinazione empirica di una legge di distribuzione, Giornale dell' Instituto Italiano degli Attuari, 4, 81-91.

[19] Lin, C.C. and Mudholkar, G.S. (1980): A simple test for normality against asymmetric alternatives, Biometrika, 67(2), 455-461.

[20] Lin, J., and Wong, S. K. M. (1990): A new directed divergence measure and its characterization, International Journal of General System, 17(1), 73-81.

[21] Locke, C. (1976): A test for the composite hypothesis that a population has a gamma distribution, Communication in Statistics- Theory and Methods, 5(4), 351-384.

[22] Lockhart, R.A. and Stephens, M.A. (1985): Goodness of fit tests for the gamma distribution, Technical Report 85-22, Department of Mathematics and Statistics, Simon Fraser University.

[23] Matis, J. H., Rubink, W. L., and Makela, M. (1992): Use of the gamma distribution for predicting arrival times of invading insect populations, Environmental Entomology, 21(3), 436-440.

[24] McCullagh, P., and Lang, P. (1984): Stochastic models for rock instability in tunnels, Journal of the Royal Statistical Society, Series B (Methodological), 344-352.

[25] McDonald, J. B., and Jensen, B. C. (1979): An analysis of some properties of alternative measures of income inequality based on the gamma distribution function, Journal of the American Statistical Association, 74(368), 856-860.

[26] Noughabi, H. A., Arghami, N. R., and Borzadaran, G. R. M. (2014): A test of goodness of fit based on Gini Index, Journal of the Turkish Statistical Association, 7(1), 23-32.

[27] Obradovic, M., Milan, J. and Milosevic, B. (2015): Goodness of fit tests for pareto distribution based on a characterization and their asymptotics, Journal of Theoretical and Applied Statistics, 49 (5), 1026-1041.

[28] Reiser, B., and Rocke, M. D. (1993): Inference for stress-strength problems under the gamma distribution, IAPQR transactions, 18, 1-22.

[29] Shapiro, S.S. and Wilk, M.B. (1965): An analysis of variance test for normality (complete samples), Biometrika, 52(3/4), 591-611.

[30] Shapiro, S.S. and Chen, L. (2001): Composite tests for the gamma distribution, Journal of Quality Technology, 33(1), 47-59.

[31] Stephens, M.A. (1986): Tests based on EDF statistics, In: D'Agostino, R.B.,Stephens, M.A., Eds., Goodness of fit techniques, Marcel Dekker, New York.

[32] Tan, W. Y. (1995): On the HIV incubation distribution under AZT treatment, Biometrical Journal, 37(3), 319-338.

[33] Von Mises, R. (1931): Wahrscheinlichkeitsrechnung und ihre Anwendung in der Statistik und theoretischen Physik, Leipzig and Wien, Franz Deuticke.

[34] Watson, G.S. (1961): Goodness-of-fit tests on a circle, Biometrika, 48(1/2), 109-114.

[35] Wilding,G.E. and Mudholkar, G.S. (2008): A gamma goodness of fit test based on characteristic independence of the mean and coefficient of variation, Journal of Statistical Planning and Inference, 138(12), 3813-3821.

[36] Yang, Z. (1994): Estimating the pattern of nucleotide substitution, Journal of Molecular Evolution, 39(1), 105-111.

[37] Yeh, Q. J. (1997): A practical implementation of gamma distribution to the reordering decision of an inventory control problem, Production and Inventory Management Journal, 38(1), 51-57.