New Modified Anderson Darling Goodness of Fit Test for Lognormal and Gamma distributions

Jutaporn Neamvonk and Bumrungsak Phuenaree*

Department of Mathematics, Faculty of Science,
Burapha University, Chonburi, 20131, Thailand

**Corresponding author's email: bumrungsak [AT] buu.ac.th*

ABSTRACT— ***The purpose of this study is to present the new modified Anderson-Darling goodness of fit test, and compare to the efficiency of three tests; Kolmogorov Smirnov test, Anderson-Darling test and Zhang (2002) test. A simulation study is used to estimate the critical values at a significance level of 0.05. The type I error rate and test power are calculated using Monte Carlo simulation with 10,000 replicates. The data are generated from the specified distribution; i.e., Lognormal and Gamma distributions with sample size of 10, 20, 30, 50, 100 and 200. The results demonstrate that every test has control over the type I error probability. The new test has the highest power for two alternative hypotheses; Loglogistic and Logistic distributions. Moreover, when the alternative distribution is Normal distribution and the sample size is small, the new test has the highest power.***

Keywords— Goodness of fit test, Anderson-Darling test, Kolmogorov Smirnov test, Modified Anderson-Darling test

---

1. INTRODUCTION

In order to describe the characteristic of important data such as wind speed data, air pollution data, rainfall and so on, we must understand the probability distribution of these data. The first step is looking for the suitable distribution using histogram or density plot, and then determine whether these sample data come from an assumed distribution or not. This statistical procedure is called goodness of fit test.

Let $X_{(1)}, X_{(2)}, ..., X_{(n)}$ denote an ordered sample randomly chosen from the distribution function, $F(x)$, with size n. The statistical hypotheses are as follow

$$H_0 : F(x) = F_0(x), \quad \text{for all } x \in (-\infty, \infty)$$
$$H_1 : F(x) \neq F_0(x), \quad \text{for some } x \in (-\infty, \infty), \tag{1}$$

where $F_0(x)$ is the assumed distribution function. Zhang [1] noted that the hypotheses

$$H_0 : \bigcap_{t \in (-\infty,\infty)} H_{0t}$$
$$H_1 : \bigcup_{t \in (-\infty,\infty)} H_{1t},$$

with $H_{0t} : F(t) = F_0(t)$ and $H_{1t} : F(t) \neq F_0(t)$. Testing H_0 versus H_1 or H_{0t} versus H_{1t} are equivalent. Cressie and Read [2] presented the family of divergence statistic indexed by a real parameter λ for testing the above hypotheses

$$2nI^{\lambda} = \frac{2n}{\lambda(\lambda+1)} \left[F_n(t) \left\{ \frac{F_n(t)}{F_0(t)} \right\}^{\lambda} + \{1 - F_n(t)\} \left\{ \frac{1 - F_n(t)}{1 - F_0(t)} \right\}^{\lambda} - 1 \right]. \tag{2}$$

Replacing $\lambda = 1$ in equation 2, the statistic is called the Pearson's χ_t^2 in (3),

$$\chi_t^2 = \frac{n\{F_n(t) - F_0(t)\}^2}{F_0(t)\{1 - F_0(t)\}}. \tag{3}$$

If we replace $\lambda = 0$ in equation 2, the statistic is called the log-likelihood ratio G_t^2 in (4),

$$G_t^2 = 2n\left[F_n(t)\log\left\{\frac{F_n(t)}{F_0(t)}\right\} + \{1-F_n\}\log\left\{\frac{1-F_n(t)}{1-F_0(t)}\right\}\right], \tag{4}$$

where $F_0(t)$ is the assumed distribution function and $F_n(t)$ is the empirical distribution function. However, there are some other special cases, such as the Freeman-Tukey statistic ($\lambda = -1/2$), the modified log-likelihood ratio statistic ($\lambda = -1$) and the Neyman modified χ^2 ($\lambda = -2$).

For testing the hypotheses (1), the test statistic, Z , can be defined as follows;

$$Z = \int_{-\infty}^{\infty} Z_t dw(t), \tag{5}$$

$$Z_{\max} = \sup_{t\in(-\infty,\infty)} \{Z_t w(t)\}, \tag{6}$$

where the Z_t can be replaced by divergence statistic with a specified λ and the $w(t)$ is weight function.

However, two widely used goodness of fit tests; the Anderson-Darling test and the Kolmogorov-Smirnov test, are derived from (5) and (6) respectively. These tests are developed by different λ for Z_t and weight function. Then, in order to acquire the new goodness of fit test, it is interesting to identify a new test statistic.

2. MATERIALS AND METHODS

In this work, we suggest a new goodness of fit test and assess its effectiveness against the three existing tests; Kolmogorov-Smirnov, Anderson-Darling and Zhang's modified Anderson-Darling test

2.1 Kolmogorov-Smirnov test (KS)

We start at equation (6) by replacing Z_t with χ_t^2 in equation (3), and $w(t) = \frac{1}{n}F_0(t)\{1-F_0(t)\}$. We have

$$K_S^2 = \sup_{t\in(-\infty,\infty)} \left[\frac{n\{F_n(t)-F_0(t)\}^2}{F_0(t)\{1-F_0(t)\}}\right]\left[\frac{1}{n}F_0(t)\{1-F_0(t)\}\right].$$

By using the Kolmogorov-Smirnov test [1],

$$K_S = \sup_{t\in(-\infty,\infty)} |F_n(t)-F_0(t)|. \tag{7}$$

2.2 Anderson-Darling test (A^2)

We start at equation (5) by replacing Z_t with χ_t^2 in equation (3), and $dw(t) = dF_0(t)$. Then, we have

$$A^2 = \int_{-\infty}^{\infty} \frac{n\{F_n(t)-F_0(t)\}^2}{F_0(t)\{1-F_0(t)\}} dF_0(t).$$

The Anderson-Darling test [3-4] is given by

$$A^2 = -\frac{2}{n}\sum_{i=1}^{n}\left[\left(i-\frac{1}{2}\right)\log\left\{F_0\left(X_{(i)}\right)\right\} + \left(n-i+\frac{1}{2}\right)\log\left\{1-F_0\left(X_{(i)}\right)\right\}\right] - n. \tag{8}$$

2.3 Zhang Modified Anderson-Darling test (Z_A)

Zhang [1] modified the Anderson-Darling test by replacing Z_t of the equation (5) with G_t^2 in equation (4), and $dw(t) = \frac{1}{F_n(t)\{1-F_n(t)\}} dF_n(t)$. We have

$$Z = \int_{-\infty}^{\infty} G_t^2 \frac{1}{F_n(t)\{1-F_n(t)\}} dF_n(t) \cdot$$

The Zhang Modified Anderson-Darling test [1] is presented below

$$Z_A = 2\sum_{i=1}^{n}\left[\frac{n}{n-i+\frac{1}{2}}\log\left\{\frac{i-\frac{1}{2}}{nF_0\left(X_{(i)}\right)}\right\}+\frac{n}{i-\frac{1}{2}}\log\left\{\frac{n-i+\frac{1}{2}}{n\left\{1-F_0\left(X_{(i)}\right)\right\}}\right\}\right] \qquad (9)$$

2.4 The proposed test statistic

Yodsima, Pongsakchat, Phuenaree and Neamvonk [5] proposed the new goodness of fit test. From the equation 2, we replace $\lambda = -1$ (modified log likelihood ratio statistic). Then we have

$$2nI^{\lambda} = n\left[F_n(t)\left\{\frac{F_n(t)}{F_0(t)}\right\}^{-1} + \{1-F_n(t)\}\left\{\frac{1-F_n(t)}{1-F_0(t)}\right\}^{-1} - 1\right] \qquad (10)$$

In the next step, we modified the new test by replacing Z_t of the equation (5) with equation (10) and $dw(t) = F_0(t)^{-1}\{1-F_0(t)\}^{-1}dF_0(t)$. Finally we have the new test statistic M_A as follow

$$M_A = -2\sum_{i=1}^{n}\left[\frac{1}{1-F_0(x_{(i)})}\log\left\{\frac{i-\frac{1}{2}}{nF_0(x_{(i)})}\right\} + \frac{1}{F_0(x_{(i)})}\log\left\{\frac{n-i-\frac{1}{2}}{n\left(1-F_0(x_{(i)})\right)}\right\}\right] \qquad (11)$$

The previous work presented only the test statistic and to determine whether the test statistic's distribution has changed as a result of the distribution of the data. In this work we would like to study efficiency of the new test compared with other tests. As well as, the probability distribution used in this research are Lognormal and Gamma distribution since these distributions are the best suited for describing important data in many regions, such as wind speed [6], rainfall [7-8] and solar radiation data [9-10].

3. RESULTS AND DISCUSSION

There are three parts discussing in this section; the critical values, the type I error rates and the test power.

3.1 The critical values

The Monte Carlo simulation with 100,000 replicates is used to estimate the critical values for the four test statistics. The Lognormal and Gamma distributed data are generated for sample size of 10, 20, 30, 50, 100 and 200, and calculate the K_S, A^2, Z_A and M_A. The 95th quantile of sampling distributions for all test statistics is defined as the critical value at 0.05 significant level. This procedure is repeated with different sets of distribution parameter values and sample sizes. Then we average the critical values of different distribution parameter of the same distribution for each sample sizes and report in Table 1.

The critical values for testing Lognormal and Gamma distribution at significance level of 0.05 are shown in Table 1. It can be clearly seen that as sample size increase, the critical values for K_S and Z_A test statistic decrease. But, as sample size increase, the critical values for A^2 and M_A test statistics increase. The values for Lognormal Gamma distribution are quite closed for each test. This implies that the test statistic's sampling distribution is similar for these distributions.

Table 1 Critical values for testing Lognormal and Gamma distribution at significance level of 0.05.

n	Lognormal distribution				Gamma distribution			
	K_S	A^2	Z_A	M_A	K_S	A^2	Z_A	M_A
10	0.2145	0.724	3.505	9.454	0.2166	0.730	3.504	9.265
20	0.1672	0.736	3.451	15.574	0.1691	0.744	3.450	15.181
30	0.1419	0.743	3.421	19.610	0.1436	0.749	3.419	19.245
50	0.1141	0.745	3.386	24.749	0.1153	0.752	3.389	24.439
100	0.0836	0.750	3.351	31.499	0.0845	0.756	3.350	31.195
200	0.0605	0.751	3.327	37.524	0.0612	0.757	3.327	37.206

3.2 Type I error rates

In this part we consider the type I error rates by generating the Lognormal and Gamma distributions from 10,000 replicates with four different sets of distribution parameter values. All tests are conducted using the critical values in Table 1. The rejection rate is counting number of rejecting H_0 and divided by number of replications, known as the type I error rate shown in Figure 1-2. It is notice that the A^2, K_S, M_A and Z_A for testing Lognormal and Gamma distribution remain within Bradley's criteria of robustness; i.e., (0.025, 0.075) throughout all sample sizes.

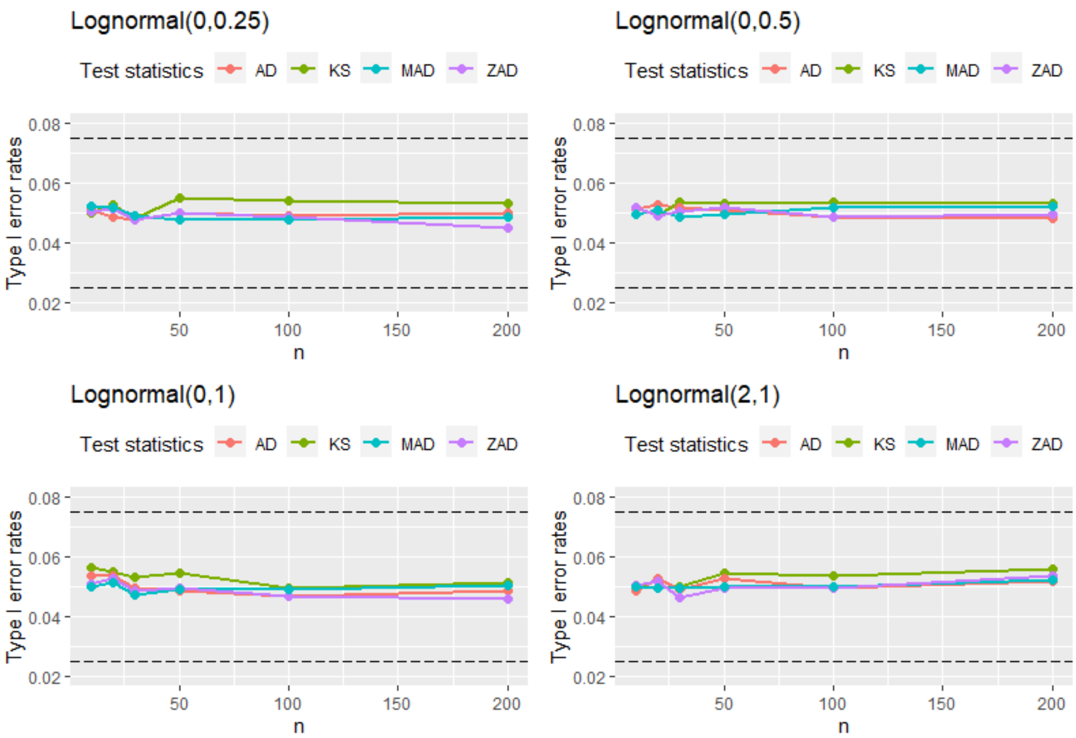

Figure 1: Type I error rates for testing Lognormal distribution

3.3 Power of the test

In this part, we test Lognormal or Gamma distributions as the null distribution in order to compute the test's power against the alternative distributions that are Loglogistic, Logistic, Normal, Lognormal (for testing Gamma distribution) and Gamma (for testing Lognormal distribution) distributions. The data are generated via an alternative distribution with 10,000 iterations. The power of the test is calculated as the proportion of the rejected null hypothesis. These results vary with the sample size of 10, 20, 30, 50, 100 and 200.

Figure 3 shows the power of the tests for testing Lognormal distribution. We note that the M_A test is the most effective test when the generated data are Loglogistic and Logistic distributions. Moreover, this test has higher power than any other tests when the generated data are Normal distribution with the sample size of 10. But, the Z_A test performs well for Gamma distribution and for Normal distribution with sample size greater than 10.

For testing Gamma distribution in Figure 4, the M_A test performs better than M_A, K_S, A^2 for Loglogistic and Logistic distributions. But the Z_A test performs well for Gamma distribution. Considering the Normal distribution as an alternative distribution, the M_A is the most effective test with sample size are 10 and 20, whereas the Z_A has the highest power when the sample size is over 20.

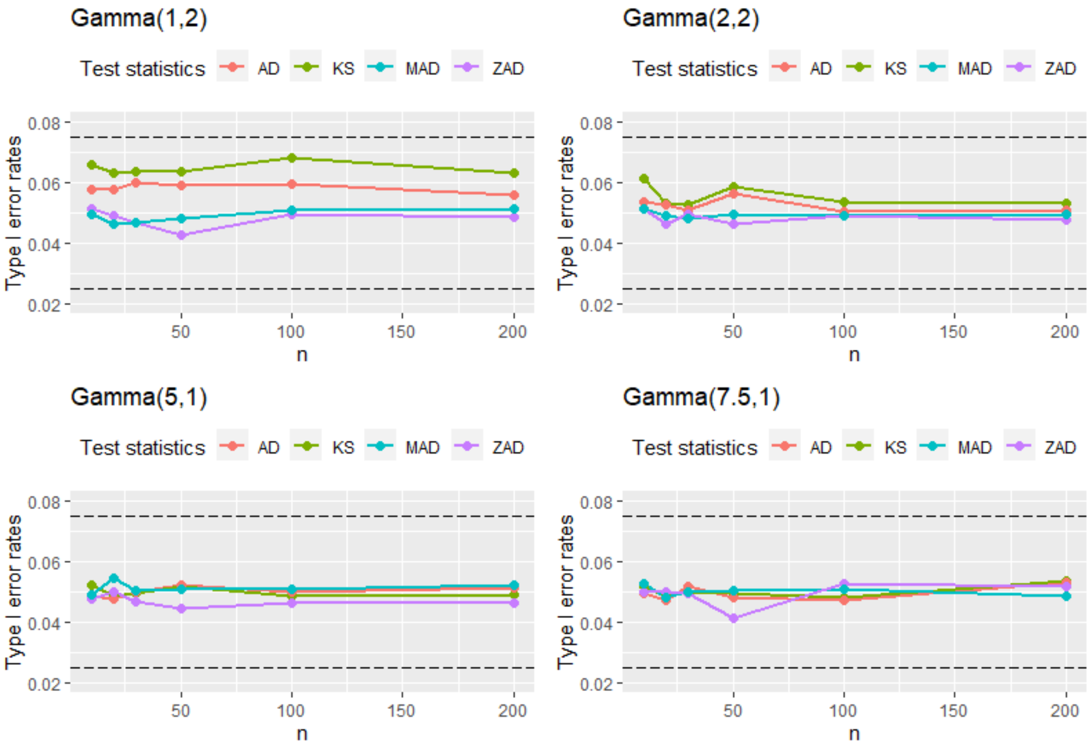

Figure 2: Type I error rates for testing Gamma distribution

4. CONCLUSIONS

In conclusion, we have suggested the new test statistics for goodness of fit test and compared to several tests. From the results, the new test M_A is the most effective test when the alternative distribution is Loglogistic and Logistic distribution for testing Lognormal and Gamma distributions, or sample size is small as 10.

5. ACKNOWLEDGEMENTS

This work was supported by the research grant of Faculty of Science, Burapha University (Grant no. 046/2559; 176/2558)

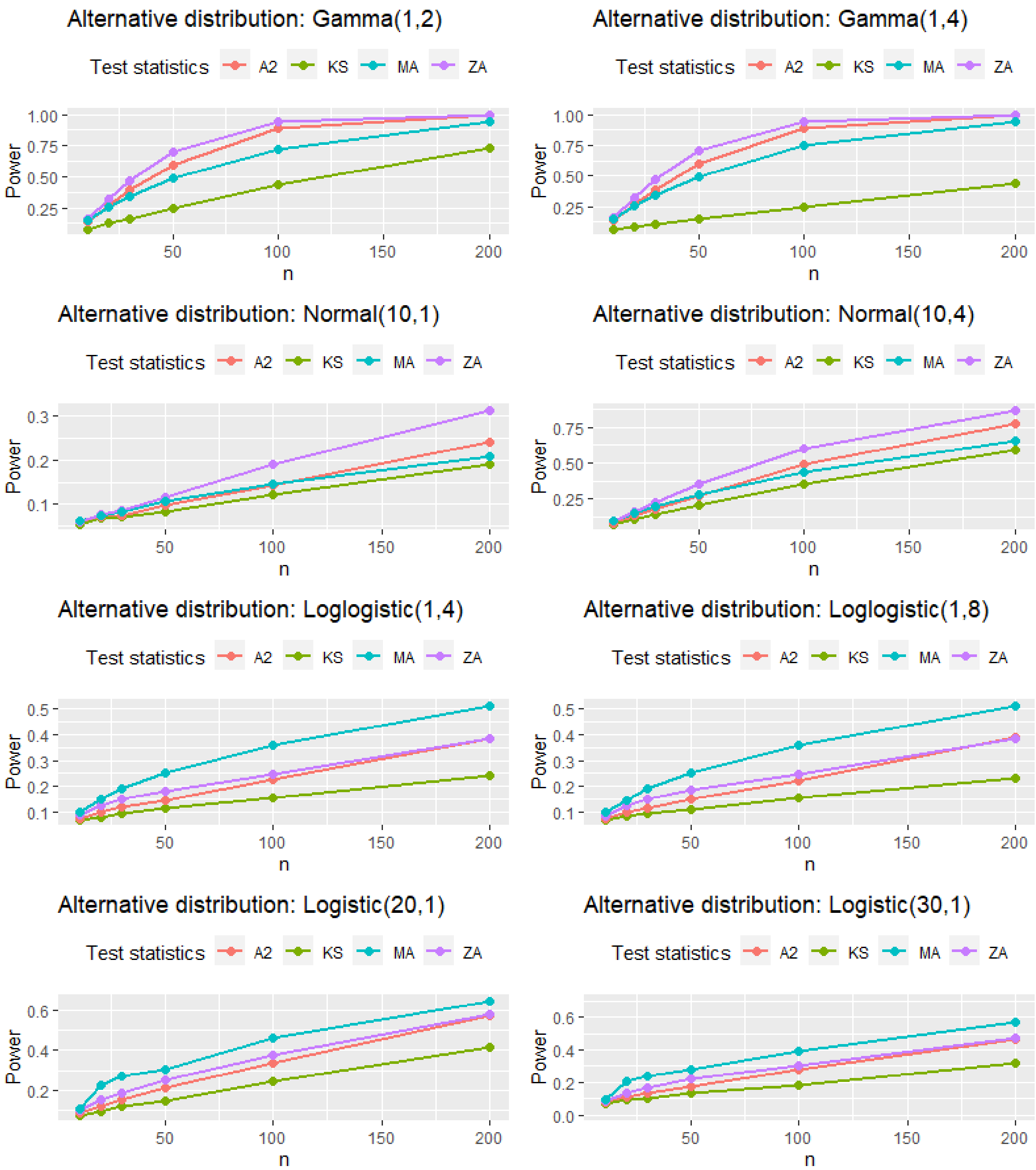

Figure 3: The power of the test for testing Lognormal distribution

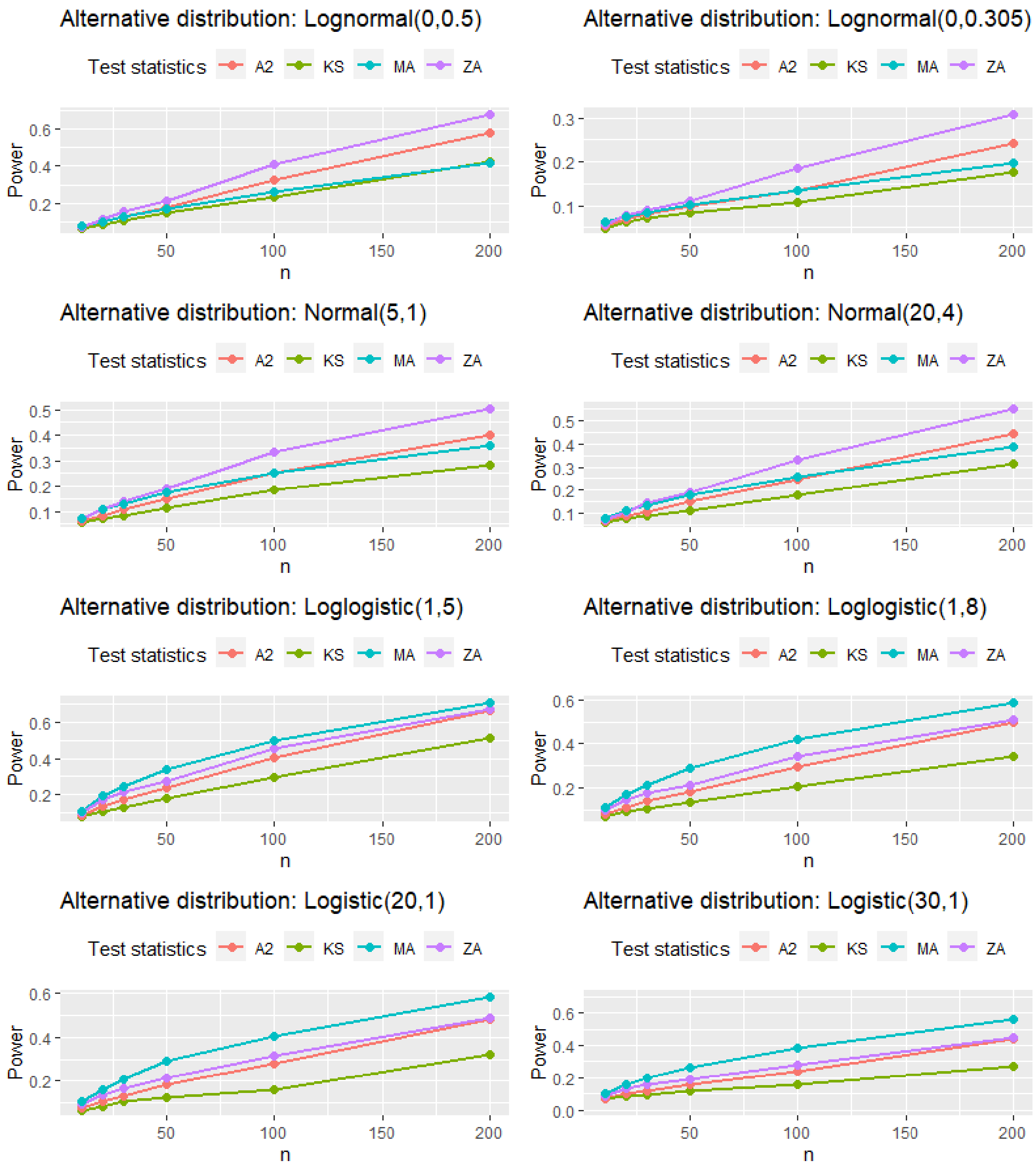

Figure 4: The power of the tests for testing Gamma distribution

6. REFERENCES

[1] Zhang, J. (2002). Powerful Goodness-of-Fit Tests Based on the Likelihood Ratio. *Journal of the Royal Statistical Society Series B, 64*(2), 281-294. DOI: 10.1111/1467-9868.00337

[2] Cressie, N. & Read, T.M.C. (1984). Multinomial Goodness-of-Fit Tests. *Journal of the Royal Statistical Society Series B, 46*(3), 440-464. Available at**:** https://www.jstor.org/stable/2345686

[3] Anderson, T.W. & 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.

[4] Anderson, T.W. & Darling, D.A. (1954). A Test of Goodness of Fit. *Journal of the American Statistical Association, 49*(268), 765-769.

[5] Yodsima, R., Pongsakchat, V., Phuenaree, B. & Neamvonk, J. (2016). A study on the distributions of goodness of fit test statistics. *Proceedings The 8th Science Research Conference.* 135-141.

[6] Morgan, E.C., Lackner, M., Vogel, R.M., & Baise, L.G. (2011). Probability distributions for offshore wind speeds. *Energy Conversion and Management, 52*(1), 15-26. doi:10.1016/j.enconman.2010.06.015

[7] Dikko, H.G., David, I.J., & Bakari, H.R. (2013). Modeling the Distribution of Rainfall Intensity using Quarterly Data. *IOSR Journal of Mathematics, 9*(1), 11-16. DOI:10.9790/5728-0911116

[8] Ximenes, P.S.M.P., Silva, A.S.A., Ashkar, F., & Stosic, T.(2021). *Best-fit probability distribution models for monthly rainfall of Northeastern Brazil, 84*(6), 1541-1556. DOI: 10.2166/wst.2021.304

[9] Arthur, Y.D., Gyamfi, K.B. & Appiah, S.K. (2013). Probability Distributional Analysis of Hourly Solar Irradiation in Kumasi-Ghana. *International Journal of Business and Social Research3*(3), 63-75. DOI:10.18533/ijbsr.v3i3.57

[10] Magenuka, T.K.M., Musasa, K. & Akindeji, K.T. (2020). Kernel Density Estimation of Solar Radiation and Wind Speed for South Africa. *Proceedings of the 5th NA International*