The Graphically Analysis of the Power of Hypothesis Parameters on Discrete (Binomial) and Continuous (Gamma) Distributions

Budi Pratikno, Annisa Kurniasih, Mashuri, and Nailatul Azizah.
Department of Mathematics, Faculty of Mathematics and Natural Sciences,
Universitas Jenderal Soedirman, Purwokerto, Indonesia
budi.pratikno@unsoed.ac.id, kurniasihannisa0@gmail.com , mashuri_unsoed@yahoo.com
nailatulazizah98@gmail.com

Abdul Talib Bon
Department of Production and Operations,
University Tun Hussein Onn Malaysia, Malaysia
talibon@gmail.com

Abstract

Many researches about the power of the hypothesis testing (one-side test) on parameter shape of some distributions have been studied by a lot of authors, such as Pratikno et al. (2017, 2019, 2020) and Pratikno (2012). Here, some of them (the powers) are used with non-sample prior information (NSPI) to improve the parameter estimation of the population. In this research, we then studied the power of hypothesis testing on one-side test of the parameter shape on of Binomial (discrete) and Gamma (continuous) distributions. To do this, we first derived the rejection area using uniformly most powerful test (UMPT) to find the formula of the Power. The *R-code* is then used to figure and graphically analysis of the graphs. The result showed that the curves are depended on the parameter shape, sample size (small *n*), and κ_0.

Keywords: *The Curve, power, parameter shape, and R-code*

1. Introduction

In the hypothesis testing, we have a probability error type I (α), a probability error type II (β) and a **power**. Later, power has significantly contributed to draw the improving population parameter. The power is then defined as a probability to reject H_0 *under* H_1 in testing hypothesis, $H_0 : \theta = \theta_0$ versus $H_1 : \theta \neq \theta_0$, for parameter θ (Wackerly, et al., 2008). We then noted that one of the statistical techniques to investigate the population inference, namely power and size. Here, we noted that the power and size now begin to be used to conclude the population inference, especially in parameter testing. Following Wackerly, et al. (2008) and Casella and Berger (2002), we the got the more detail of the definition of the power and size, that are, power is the probability to reject H_0 under H_1 and the size is the probability to reject H_0 under H_0.

Many authors already studied the improving inference population using power and size such as Pratikno et al. (2017, 2019, 2020), Pratikno (2012), Khan and Pratikno (2013), Khan (2003), Khan and Saleh (1995), Khan and Hoque (2003), Saleh (2006), Yunus (2010), and Yunus and Khan (2007). Here, Pratikno [2012], Khan and Pratikno [2013] and Khan [2003] studied about a testing intercept using non-sample prior information (NSPI). Moreover, Pratikno [2012] and Khan et al. [2016] used the power ans size to compute the cdf of the bivariate noncentral *F (*BNCF*)* distribution of the pre-test test (PTT) with NSPI in some regression models. The computational of the probability integral of the probability distribution function (pdf) and cdf of the BNCF distribution are very complicated and hard, so the R code is used.

In the context of the power on the distribution, Pratikno et al. (2017, 2019, 2020), are already presented the formula and value of the power on discrete and continuous distributions. Here, we then focused to analysis (research) about the power and size on Binomial (discrete) and Gamma (continuous) distribution. There are several steps to compute the power of the distribution: (1) determine the sufficiently statistics, (2) compute the rejection area using *uniformly most powerful test* (UMPT), (3) derive the formula of the power of the distributions in testing hypothesis, and (4) plot the graphs of the power of their distributions.

In this paper, Section 1 presented the introduction. The methodology research is given in Section 2. A simulation data (graph of the power and size) is obtained in Section 3, and Section 4 described the conclusion of the research.

2. Research Methodology

Step 1. We studied Binomial and Gamma Distribution.
Step 2. We then refer to previous research to derived the formula of the power and size .
Step 3. Following step 2, we then produced and graphically analysis of the graph of the power
Step 4. Finally, we draw the conclusion by choosing the maximum power and minimum size for both distributions

3. Discussion Results

3.1 The Power on the Binomial Distribution

In this section, we follow Pratikno et al. (2020) in discussing the power of the Binomial distribution. Here, we let X_i Bernoulli case with parameter θ, $n = 20$, and $Y = \sum_{j=1}^{n=20} X_j$, then *Y* is written as $Y : \ Bin(n,\theta)$ follows Binomial distribution with $n = 20$ and θ, as independent random variable with *probability mass function* (PMF) is given as

$$f(x_i) = \binom{n}{x} \theta^x (1-\theta)^{n-x} \quad . \tag{1}$$

For testing the two-side hypothesis, $H_0 : \theta = 0.1$ versus $H_1 : \theta \neq 0.1$ on rejection area $\{(x_1,...,x_{12}) : \mathrm{Y} \leq 6\}$, we then derive the formula of the power as given below

$$\begin{aligned}
\pi(\theta) &= P(\text{Reject } H_0 \,|\text{under } H_1) = \sum_{y=0}^{6} \binom{20}{y} \theta^y (1-\theta)^{20-y} \\
&= \binom{20}{0} \theta^0 (1-\theta)^{20-0} + \binom{20}{1} \theta^1 (1-\theta)^{20-1} + \text{L} \ + \binom{20}{6} \theta^6 (1-\theta)^{20-6} \\
&= \left(\frac{20!}{0!(20-0)!}\right) 1(1-\theta)^{20} + \left(\frac{20!}{1!(20-1)!}\right) \theta^1 (1-\theta)^{19} + \text{L} \ + \left(\frac{20!}{6!(20-6)!}\right) \theta^6 (1-\theta)^{14} \\
&= (1-\theta)^{20} + (20)\theta^1 (1-\theta)^{19} + \text{L} \ + (38760)\theta^6 (1-\theta)^{14}.
\end{aligned} \tag{2}$$

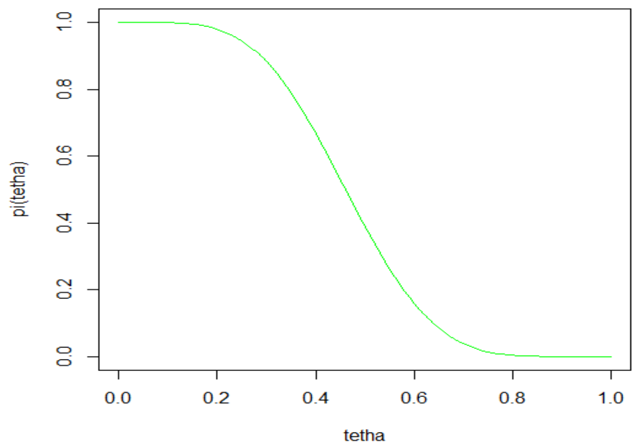
The size in testing the two-side hypothesis, $H_0 : \theta = 0.1$ versus $H_1 : \theta \neq 0.1$ on rejection area $\{(x_1,...,x_{12}) : \mathrm{Y} \leq 6\}$, is then given as

$$\alpha(\theta) = P(\text{Reject } H_0 \,|\, \text{under } H_0) = \sum_{y=0}^{6} \binom{20}{y} \theta^{y} (1-\theta)^{20-y} \Big|_{\theta=0.1}$$

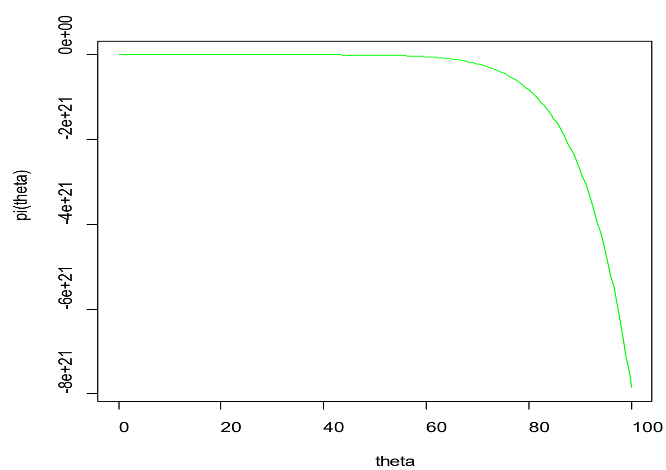
$$= (1-0.1)^{20} + (20)(0.1)^{1}(1-0.1)^{19} + \cdots + (38760)(0.1)^{6}(1-0.1)^{14} \tag{3}$$

From the equation (3), it is clear that the size is constant, so the graph is not depended on parameter shape, but it will be straight line.

Using the equation (2) and following Pratikno et al. (2020), we then produce the graph of power of the Binomial distribution. The graph is given in Figure 1.



(a)



(b)

Figure 1. The Power of the Binomial Distribution

From Figure 1., we see that the curves are similar (sigmoid and ½ sigmoid), the curve on the graph (a) tend to be one in short period on the parameter, that is $0 \le \theta \le 0.2$, and they begin decrease on $0.2 \le \theta \le 0.8$ (sigmoid), and then they tend to be flat (asymptote to zero) on $\theta \ge 0.8$. Here, we note that the period of the probability one of the curves on the graph (b) is more length, that is $0 \le \theta \le 0.8$. This is due to the different value of the null hypothesis. Generally, we note that they (both of the graphs) will be maximum on one, and the minimum is zero.

Following the previous research of Pratikno (2012), Khan and Pratikno (2013), Khan (2003), Khan and Saleh (1995), Khan and Hoque (2003), Saleh (2006), Yunus (2010), and Yunus and Khan (2007), we noted that the maximum power (one) and minimum size are chosen for the better indicator testing. I therefore conclude that the graph of the power will be not so similar (or the same as) when *n*, θ, and rejection area are different. When they increase (or decrease), the graph will be different in achieving of the maximum power going to be one.

3.2 The Power on the Gamma Distribution

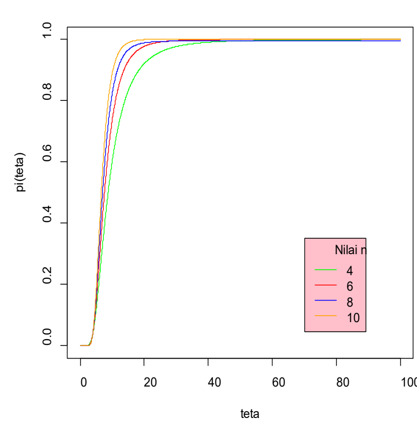
Following the previous research of the Pratikno (2012) and Pratkno et al. (2019, 2020) and the probability distribution function (pdf) of Gamma distribution of the variable random *X* with parameter $\kappa > 0$ and $\theta > 0$ $f(x,\kappa,\theta) = \frac{1}{\theta^{\kappa}\Gamma(\kappa)} x^{\kappa-1} e^{-x/\theta}, x > 0$, the formula of the power of the Gamma distribution in testing $H_0 : \theta = 0.1$ versus $H_1 : \theta > 0.1$, is then given as

$$\pi(\theta) = 1 - P\left(\frac{2}{\theta}\sum_{i=1}^{n} x_i \leq \frac{\theta_0}{\theta}\chi^2_{(2n\kappa_0;\alpha)}\right) = 1 - \int_0^{\frac{\theta_0}{\theta}\chi^2_{(2n\kappa_0;\alpha)}} \frac{1}{2^n\Gamma(n)} x^{n-1} e^{-\frac{x}{2}} dx. \tag{4}$$

Similarly, the formula of the size of the Gamma distribution is then written, for $\theta = 0.1$, as

$$\alpha(\theta) = 1 - P\left(\frac{2}{\theta}\sum_{i=1}^{n} x_i \leq \frac{\theta_0}{\theta}\chi^2_{(2n\kappa_0;\alpha)}\right) = 1 - \int_0^{\frac{\theta_0}{\theta}\chi^2_{(2n\kappa_0;\alpha)}} \frac{1}{2^n\Gamma(n)} x^{n-1} e^{-\frac{x}{2}} dx \tag{5}$$

Here, the value of the θ on the equation (5), is exactly 0.1. So the curve is not sigmoid, but it will be straight line and constant. Using the equation (4), we then simulate on several *n* and κ_0 for $\theta_0 = 2$ and $\alpha = 0,01$ to get the graphs of the power. The graph of the power of the Gamma distribution is then given in Figure 2.



(a) $\kappa_0 = 2$

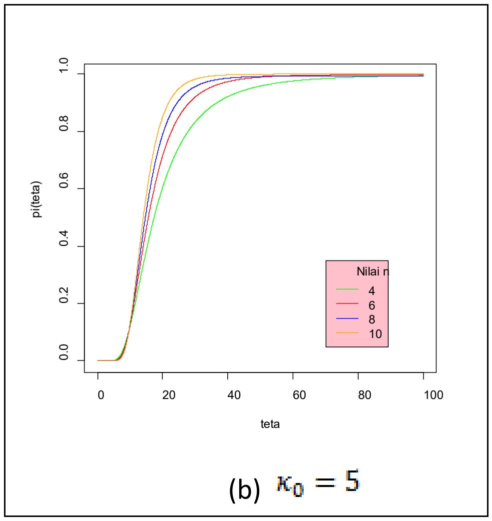


Figure 2. The Power of the Gamma Distribution

From Figure 2 (a) and (b), we see that all the curves are influenced by *n* and κ_0. The highest curve is occurred on large *n* and small κ_0 (see both Figure 2). we also noted that the curves will be move to the right (more skew to the right) for large κ_0, and the lower curve is occurred when they have small *n*. Generally, all the curve tends to be one faster when the parameter shape is small, and they will go to be one on short period of the parameter shape. Here, we note that the 1st graph has short period of the parameter shape between 0 to 0.2, but the 2nd graph is around 0 to 0.4. It is clear that they tend to be smooth (skew to the right) when n and κ_0 are increase. We also noted that they are faster to one for small *n* and κ_0. Here, we also note that both of the graphs (a) and (b) will be maximum on one, and the minimum is zero in small parameter shape. Following the previous research of such as Pratikno (2012), Khan and Pratikno (2013), Khan (2003), Khan and Saleh (1995), we must to choose the maximum power (one) as an indicator testing.

4 Conclusion

The research studied power and size on Binomial (discrete) and Gamma (continuous) distribution. The rejection area is used to find the formula of the power, and the *R-code* is then used to figure and graphically analysis. The result showed that the influencing factor of the curves of the power are parameters shape and sample size (*n*) , and we also note that all the curve tend to be more faster to one for small *n* and κ_0.

References

Casella, G. dan Berger, R. L. (2002). *Statistical Inference, 2nd* ed. USA: Thomson Learning Inc.

Khan, S. and Pratikno, B. Testing Base Load with Non-Sample Prior Information on Process Load. *Statistical Papers*, 54 (3), 605-617, (2013).

Khan, S. Estimation of the Parameters of two Parallel Regression Lines Under Uncertain Prior Information. *Biometrical Journal, **44**, 73-90*, (2003).

Khan, *S.* and Saleh, A. K. Md. E. Preliminary test estimators of the mean based on p-samples from multivariate Student-t populations. *Bulletin of the International Statistical Institute. 50th Session of ISI, Beijing, 599-600*, (1995).

Khan, S. and Hoque, Z. Preliminary test estimators for the multivariate normal mean based on the modified W, LR and LM tests. *Journal of Statistical Research, Vol 37, 43-55, (*2003).

Khan, S., Pratikno, B. Ibrahim, A.I.N., and Yunus, R.M. The correlated bivariate noncentral *F* distribution and Its application. *Communications in Statistics—Simulation and Computation*, 45 3491–3507 (2016).

Pratikno, B, Parwati, M.S.R and Bon, A.T. (2020). The Decision Case on Gaussian Binary Data*Proceedings of the 5th NA International Conference on* Industrial Engineering and Operations Management Detroit, Michigan, USA, August 10-14, 2020, IEOM Detroit Conference, Link proceedings papers or abstracts here : http://ieomsociety.org/detroit2020/proceedings/

Pratikno, B ., Layyinah, S. Y., Pratidina, G. M., dan Suryaningtiyas, Y. D. (2019). *The Power of Weibull and Exponential Distribution On Testing Parameters Shape*. Journal of IOP Conference Series Vol 255(1), pp 012029.

Pratikno, B.. Test of Hypothesis for Linear Models with Non-Sample Prior Information. Unpublished PhD Thesis, University of Southern Queensland, Australia, (2012).

Saleh, A. K. Md. E. Theory of preliminary test and Stein-type estimation with applications. *John Wiley and Sons, Inc., New Jersey,* (2006).

Wackerly, D. D., Mendenhall, W., dan Scheaffer, R. L. (2008). *Mathematical Statistics with Applications*, *7th ed.* USA: Thomson Learning, Inc.

Yunus, R. M. Increasing power of M-test through pre-testing. *Unpublished PhD Thesis, University of Southern Queensland, Australia*, (2010).

Yunus ,R. M. and Khan, S. Test for intercept after pre-testing on slope a robust method. *In: 9th Islamic Countries Conference on Statistical Sciences (ICCS-IX): Statistics in the Contemporary World - Theories, Methods and Applications,* (2007).

Acknowledgement
I thankfully to the colleague for providing me the ideas and data.

Biographies

Budi Pratikno, currently works as a lecturer in Department of Mathematics, Universitas Jenderal Soedirman, author of several international journals related to testing intercepts with non-sample prior information (NSPI), which are widely published in international journals such as Statistical Papers, Far East Journal of Mathematics Science, Springer, JSTA, ISSOS, IJET, IJAST, IOP series, IEOM and others. In the study of NSPI as the main research the author, we try to improve population inference using non-sample. The author completed his undergraduate education at the Mathematics Department (UGM), then continued his Masters in Statistical Science at La Trobe University, Melbourne, Australia, and his S3 in Statistical Science (Ph.D) at the University of Southern Quensland (USQ), Toowoomba, Brisbane, Australia.

Mashuri, currently works as a lecturer in Department of Mathematics, Universitas Jenderal Soedirman, author of several international journals related to mathematical dynamic, which are widely published in international journals such as Far East Journal of Mathematics Science, Advances and Applications in Fluid Mechanics, Far East Journal of Applied Mathematics and others. The author completed his undergraduate education (S.Si) at the Mathematics Department (UNPAD), then continued his Masters (M.Si.) and doctorate program (Dr.) in Mathematics at ITB Bandung, Indonesia.

Kurniasih, A. and Azizah, N. are students in Department of Mathematics, Universitas Jenderal Soedirman.

Abdul Talib Bon is a professor of Production and Operations Management in the Faculty of Technology Management and Business at the Universiti Tun Hussein Onn Malaysia since 1999. He has a PhD in Computer Science, which he obtained from the Universite de La Rochelle, France in the year 2008. His doctoral thesis was on topic Process Quality Improvement on Beltline Moulding Manufacturing. He studied Business Administration in the Universiti Kebangsaan Malaysia for which he was awarded the MBA in the year 1998. He's bachelor degree and diploma in Mechanical Engineering which his obtained from the Universiti Teknologi Malaysia. He received his postgraduate certificate in Mechatronics and Robotics from Carlisle, United Kingdom in 1997. He had published more 150 International Proceedings and International Journals and 8 books. He is a member of MSORSM, IIF, IEOM, IIE, INFORMS, TAM and MIM.