

POWER AND SIZE OF NORMAL DISTRIBUTION AND ITS APPLICATIONS

Budi Pratikno

Department of Mathematics, Faculty Mathematics and Natural Sciences
Jenderal Soedirman University
bpratikto@gmail.com

Jajang

Department of Mathematics, Faculty Mathematics and Natural Sciences
Jenderal Soedirman University

Setianingsih

Department of Mathematics, Faculty Mathematics and Natural Sciences
Jenderal Soedirman University

R. Sudarwo

Open University, UPBJJ-UT, Purwokerto, Indonesia

ABSTRACT. *The research studied power and size of normal distribution and its applications on linear regression model. The power and size formulas are derived, and the unrestricted test (UT), restricted test (RT) and pre-test test (PTT) are used. The recommendation of the test is given by choosing maximum power and minimum size, and also graphical analysis. The result showed that the power and size for large standard deviation (σ) tend to be identical and flat. In simulation study, the graphs of the UT, RT, and PTT are still similar to the previous research (Pratikno, 2012), where the PTT tend to lie between UT and RT.*

Keyword: *power, linear regression model, and size.*

ABSTRAK. Riset ini mengkaji *power* dan *size* distribusi normal dan aplikasinya pada model regresi linier. Formula *power* dan *size* diturunkan, dan *unrestricted test (UT)*, *restricted test (RT)* dan *pre-test test (PTT)* digunakan. Rekomendasi uji tersebut diberikan dengan memilih *power* maksimum dan *size* minimum, dan juga analisis grafik. Hasilnya menunjukkan bahwa *power and size* untuk simpangan baku besar (σ) cenderung identik dan mendatar. Dalam kajian simulasi, grafik UT, RT, dan PTT masih mirip dengan riset sebelumnya (Pratikno, 2012), di mana PTT cenderung berada di antara UT dan PTT.

Kata kunci: *power, model regresi linier, dan size*

1. Introduction

Normal distribution is often called as Gaussian distribution. The probability density function (pdf) of variabel random X with parameter mean (μ) and variance (σ^2) is defined as

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, -\infty < x < \infty, -\infty < \mu < \infty \text{ and } 0 < \sigma < \infty. \quad (1)$$

To compute the pdf and cumulativie distribution function (cdf) in the equation (1), *R-code* is used. Due to this difficulties, here, the values of the power for rejecting null hypothesis (H_0) under alternatif hypothesis (H_1) is also computed using *R-code*.

The power and size in term of univariate normal distribution havel already studied by many authors, such as Pratikno (2012), Khan (2005) and Yunus (2010). Moreover, Pratikno (2012) used the power to compute the values of the power of the tests: unrestricted test (UT), restricted test (RT) and pre-test test (PTT) in testing intercept using non-sample prior information (NSPI) on model regression linear model. Many authors have already contributed in developing this research area especially in estimation area, such as Khan (2005, 2008), Khan and Saleh (1997, 2005, 2008), Khan and Hoque (2003), and Saleh (2006).

The values of the power and size are very significance criteria in testing UT, RT and pre-test test (PTT) on regression model (RM). Therefore, we used power and size to test UT, RT dan PTT by choosing maximum power and minimum size of them. Furthermore, this research is focused to derive fromula of the power and size of the normal distribution and compute them using *R-code*. To more clear, we also did graphical analysis of the power and size the UT, RT and PTT.

The research presented the introduction in Section 1. The reconstruction of the power and size formula of the normal distribution are given in Section 2. A simulation study and their (power and size) applications on the multivariate simple regression model (MSRM) model are presented in Section 3, and Section 4 described conclusion

2. Power and Size Normal Distribution

Cummulative distribution function (cdf) of the normal distribution with parameter μ and σ for $x = x_1$ and $x = x_2$ random variable X is given as

$$P(x_1 < X < x_2) = \int_{x_1}^{x_2} f(x) dx = \int_{x_1}^{x_2} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx \quad (2)$$

To compute the equation (2), *R-code* is used. This is due to the probability integral of the cdf of the equation (2) is difficult. Here, we need numerical analysis and theory of the Maclourin series. To compute the values of the power and size for two-side hypothesis $H_0: \mu = \mu_0$ versus $H_a: \mu \neq \mu_0$, the cdf is used. The power ($\pi(\mu)$) and size (α) formulas are then given as, respectively,

$$\begin{aligned} \pi(u) &= P(\text{reject } H_0 \mid \text{under } H_1 : \mu = \mu_1) = \int_{x_1}^{x_2} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx \\ &= \int_{x_1}^{x_2} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu_1}{\sigma}\right)^2} dx \\ &= \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\mu_1}{\sigma}\right)^2} \int_{x_1}^{x_2} \sum_{i=0}^{\infty} \frac{1}{i!} \left(-\frac{x^2}{2\sigma^2}\right)^i \sum_{i=0}^{\infty} \frac{1}{i!} \left(-\frac{x\mu_1}{2\sigma^2}\right)^i dx \quad (3) \end{aligned}$$

$$\begin{aligned} \alpha &= P(\text{reject } H_0 \mid \text{under } H_0 : \mu = \mu_0) = \int_{x_1}^{x_2} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx \\ &= \int_{x_1}^{x_2} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu_0}{\sigma}\right)^2} dx \\ &= \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\mu_0}{\sigma}\right)^2} \int_{x_1}^{x_2} \sum_{i=0}^{\infty} \frac{1}{i!} \left(-\frac{x^2}{2\sigma^2}\right)^i \sum_{i=0}^{\infty} \frac{1}{i!} \left(-\frac{x\mu_0}{2\sigma^2}\right)^i dx \quad (4) \end{aligned}$$

For hypothesis $H_0: \mu = 6$ versus $H_a: \mu \neq 6$, $\sigma = 4$, range 0 to 4, and $\mu_1 = 4$, we then got the power is 0,813701, and for $\mu_1 = 8$ the power is 0, 223284. Similarly, the value of the size under H_0 for $\mu_1 = 4$ is 0,1609764. Furthermore, we presented graphs of the power and size for several σ as follows

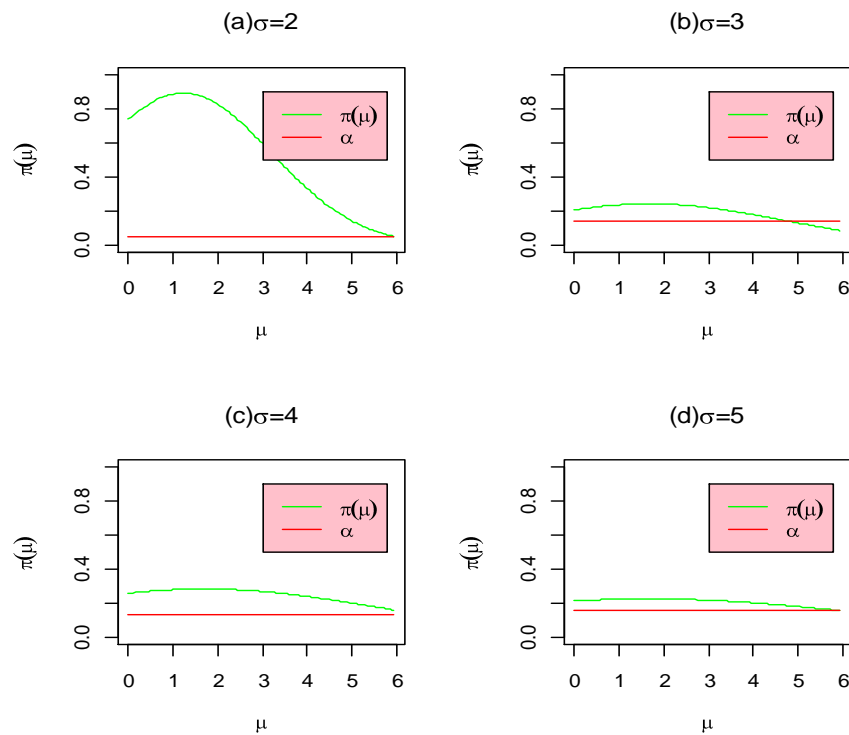


Figure 1. The graphs of power and size of the normal distribution for some selected σ

Figure 1. showed that the *power* tend to be close to *size* for $\mu = 0$ to 6, on $x_1 = 0$ and $x_2 = 3$ when the values of the σ increases. The graphs are similar to the curve of the normal distribution, i.e. the curve tend to be leptocurtic for large the σ , and they will be flat for small the σ .

3. A Simulation Study of the Power dan Size on Generate Data

Following Pratikno (2012), then the *power* and *size* of the UT, RT and PTT on multivariate simple regression model (MSRM) for 4 dependent variables, namely $y_1 = 5 + 2x$, $y_2 = 2 + 4x$, $y_3 = 3 - 5x$ and $y_4 = 6 + 8x$, with *level of significance* 0,05 in *two-side* hypothesis $H_0 : \beta_{0i} = 0$ versus $H_0 : \beta_{0i} \neq 0$, the graphs of the UT, RT and PTT is then obtained as follows.

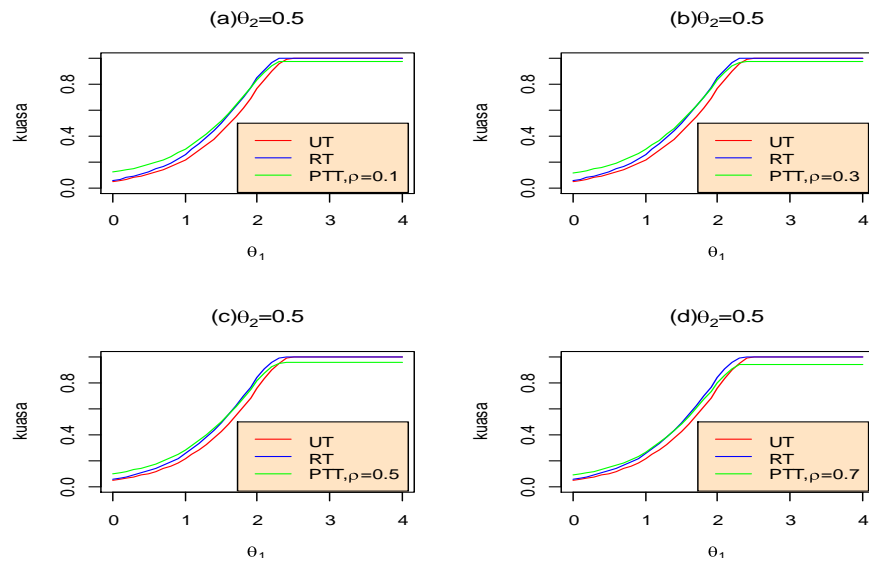


Figure 2. Power of the UT, RT, and PTT for $\rho = 0,1, 0,3, 0,5, 0,7, \theta_2 = 0,5$

From Figure 2., we see that PTT lies between UT and RT, it means the PTT can be an alternative choice of the test. In the context of maximum power, the PTT still follows the previous research, Pratikno (2012). Similarly, we also got that the alternative choice of the minimum size of under H_0 .

4. Conclusion

The research studied power and size of normal distribution and its applications on regression linear model. The power and size formulas of the normal distribution are then reconstructed, and they are used to compute some tests, namely test (UT), restricted test (RT) and pre-test test (PTT). The result showed that the power and size for large standard deviation (σ) tend to be identical and flat. In simulation study, the graphs of the UT, RT, and PTT are still similar to the previous research, i.e. the PTT tend to lie between UT and RT.

References

Khan, S., *Estimation of Parameters of The Multivariate Regression Model with Uncertain Prior Information and Student-t Errors*, Journal of Statistical Research, **39**(2) (2005), 79-94

- Khan, S., *Shrinkage Estimators of Intercept Parameters of Two Simple Regression Models with Suspected Equal Slopes*, Communications in Statistics-Theory and Methods, **37** (2008), 247-260.
- Khan, S. and Saleh, A. K. Md. E., *Shrinkage Pre-test Estimator of The Intercept Parameter for A Regression Model with Multivariate Student-t Errors*, Biometrical Journal, **39** (1997), 1-17.
- 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, **37** (2003), 43-55.
- Khan, S. and Saleh, A. K. Md. E., *Estimation of Intercept Parameter for Linear Regression with Uncertain Non-Sample Prior Information*, Statistical Papers, **46** (2005), 379-394.
- Khan, S. and Saleh, A. K. Md. E., *Estimation of Slope for Linear Regression Model with Uncertain Prior Information and Student-t Error*, Communications in Statistics-Theory and Methods, **37**(16) (2008), 2564-2581.
- Pratikno, B., *Tests of Hypothesis for Linear Regression Models with Non Sample Prior Information*, Dissertation, 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.
- Yunus., R. M., *Increasing Power of M-test through Pre-testing*, Unpublished Ph.D. Thesis, University of Southern Queensland, Australia, 2010.