

## **AN OVERVIEW OF HOMOGENEITY OF VARIANCE TESTS ON VARIOUS CONDITIONS BASED ON TYPE 1 ERROR RATE AND POWER OF A TEST** (Sorotan Ujian Kehomogenan Varians pada Pelbagai Keadaan Berdasarkan Kadar Ralat Jenis 1 dan Kuasa Ujian)

NUR FAZLIN ABDULLAH & NORA MUDA\*

### *ABSTRACT*

In most statistical analyses, the data variance used is assumed to be homogeneous, but not all cases follow the assumption. Therefore, the homogeneity of variance assumption testing should be carried out prior to performing the main analysis. There are various statistical tests of variance homogeneity that exist and to obtain the best statistical test in the testing of variance equality, this study makes a comparison of statistical tests against assumptions met, assumptions violated and the existence of outlier. The comparison is based on the Type 1 Error rate and the power of the statistical test. For normal distribution, the comparison is between parametric statistical tests such as the Fisher test, the Bartlett test, the Levene test, the Brown-Forsythe test, and the Cochran C test. While for chi-squared distribution and outlier data, the comparison is between parametric and nonparametric statistical tests. The nonparametric statistical tests used are the Mood test, the Ansari-Bradley test, and the Fligner-Killeen test. The data used is the result of a normal and Chi-squared Monte Carlo simulation. The results showed that almost all the parametric statistical tests can control Type 1 Errors well in almost all situations. For the Chi-squared distribution only the Brown-Forsythe parametric statistical test was found to be robust. But most of the robust tests on non-normal data are nonparametric statistical tests. While for normal data with heterogeneous variance, the power of the test of all parametric statistical tests is seen to increase and exceed 0.80 as the size effect increases. On non-normal distributions, the power of the test is smaller than normal, but the value will increase as the size effect increases. The case was different for the Fisher test, the Bartlett test, and the Cochran C test, which was tested against data with 10% outlier in one group. The power of the test for the 1:2 variance ratio is seen as large, but the value is decreasing as the size effect increases. Thus, it can be concluded that none of the statistical tests were found to be robust and suitable for use in all the conditions set.

*Keywords:* Monte Carlo simulation; nonparametric test; outlier; parametric test; robust test

### *ABSTRAK*

Dalam kebanyakan analisis statistik, varians bagi data yang digunakan adalah bersifat homogen. Tetapi bukan semua kes yang andaian kehomogenan varians dicapai. Oleh sebab itu, pengujian andaian kehomogenan varians perlu dilakukan sebelum analisis utama dijalankan. Terdapat pelbagai ujian statistik kehomogenan varians yang wujud dan untuk mendapatkan ujian statistik yang terbaik dalam pengujian kesamaan varians, kajian ini membuat perbandingan ujian-ujian statistik terhadap andaian dipenuhi, andaian tidak dipenuhi dan kewujudan data terpencil. Perbandingan yang dilakukan adalah berdasarkan kadar Ralat Jenis 1 dan kuasa ujian statistik. Bagi taburan normal, perbandingan adalah antara ujian-ujian statistik berparameter seperti ujian Fisher, ujian Bartlett, ujian Levene, Ujian Brown-Forsythe dan ujian Cochran C. Manakala bagi taburan khi-kuasa dua dan data terpencil, perbandingan yang dibuat adalah antara ujian statistik berparameter dan tak berparameter. Ujian statistik tak berparameter yang digunakan adalah ujian Mood, ujian Ansari-Bradley dan ujian Fligner-Killeen. Manakala data yang digunakan adalah hasil simulasi Monte Carlo yang tertabur secara normal dan Khi-kuasa dua. Hasil kajian mendapati hampir kesemua ujian statistik

berparameter dapat mengawal Ralat Jenis 1 dengan baik pada hampir kesemua keadaan. Bagi taburan Khi-kuasa dua hanya ujian statistik berparameter Brown-Forsythe yang didapati teguh. Tetapi kebanyakan ujian yang teguh pada keadaan data bukan normal adalah ujian statistik tak berparameter. Manakala bagi data normal dengan varians heterogen, kuasa ujian bagi kesemua ujian statistik berparameter dilihat meningkat dan melebihi 0.80 apabila kesan saiz semakin meningkat. Pada taburan bukan normal, kuasa ujian adalah lebih kecil berbanding normal tetapi nilai tersebut akan meningkat seiring dengan peningkatan kesan saiz. Kes ini berlainan pula bagi ujian Fisher, ujian Bartlett dan ujian Cochran C yang diuji terhadap data dengan kewujudan 10% nilai terpencil pada satu kumpulan. Kuasa ujian tersebut bagi nisbah varians 1:2 dilihat besar tetapi nilai tersebut semakin menurun apabila kesan saiz semakin meningkat. Maka, boleh disimpulkan bahawa tiada satu ujian statistik pun yang didapati teguh dan sesuai digunakan pada kesemua keadaan yang ditetapkan.

*Kata kunci:* simulasi Monte Carlo; ujian tak berparameter; terpencil; ujian berparameter; ujian teguh

## References

- Agunbiade D.A. & Iyaniwura J.O. 2010. Estimation under multicollinearity: A comparative approach using Monte Carlo methods. *Journal of Mathematics and Statistics* **6**(2): 183-192.
- Ahad N.A., Yin T.S., Othman A.R. & Yaacob C.R. 2011. Sensitivity of normality tests to non-normal data. *Sains Malaysiana* **40**(6): 637-641.
- Alabi O.O., Ayinde K. & Olatayo T.O. 2008. Effect of multicollinearity on power rates of the ordinary least squares estimators. *Journal of Mathematics and Statistics* **4**(2): 75-80.
- Anderson T.W. 1962. On the distribution of the two-sample Cramer-von Mises criterion. *Annals of Mathematical Statistics* **33**(3): 1148-1159.
- Anderson T.W. & Darling D.A. 1952. Asymptotic theory of certain 'Goodness of Fit' criteria based on stochastic processes. *Annals of Mathematical Statistics* **23**(2): 193-212.
- Bartlett M.S. 1937. The statistical conception of mental factors. *British Journal of Psychology* **28**(1): 97-104.
- Bradley J.V. 1978. Robustness?. *British Journal of Psychology* **31**(2): 144-152.
- Brown M.B. & Forsythe A.B. 1974. Robust tests for the equality of variances. *Journal of the American Statistical Association* **69**(346): 364-367.
- Carroll R.J. & Schneider H. 1985. A note on Levene's tests for equality of variances. *Statistics & Probability Letters* **3**(4): 191-194.
- Chambers J.M., Cleveland W.S., Kleiner B. & Tukey P.A. 1983. *Graphical Methods for Data Analysis*. Boston: Duxbury Press.
- Cochran W.G. 1941. The distribution of the largest of a set of estimated variances as a fraction of their total. *Annals of Eugenics* **11**(1): 47-52.
- Cochran W.G. & Cox G.M. 1957. *Experimental Design*. NY: John Wiley and Sons.
- Cohen J. 1988. *Statistical Power Analysis for the Behavioral Sciences*. 2nd Ed. Hillsdale, NJ: Lawrence Erlbaum Associates Publishers.
- Conover W.J. & Iman R.L. 1981. Rank transformations as a bridge between parametric and nonparametric statistics. *The American Statistician* **35**(3): 124-129.
- Conover W.J., Johnson M.E. & Johnson M.M. 1981. A comparative study for homogeneity of variances, with applications to the outer continental shelf bidding data. *Technometrics* **23**(4): 351-361.
- Conover W.J., Guerrero-Serrano A.J. & Tercero-Gómez V.G. 2018. An update on 'A comparative study of tests for homogeneity of variance'. *Journal of Statistical Computation and Simulation* **88**(8): 1454-1469.
- Ercheg-Hurn D.M. & Mirosevich V.M. 2008. Modern robust statistical methods: An easy way to maximize the accuracy and power of your research. *American Psychologist* **63**(7): 591-601.
- Fligner M.A. & Killeen T.J. 1976. Distribution-free two-sample tests for scale. *Journal of the American Statistical Association* **71**(353): 210-213.
- Gastwirth J.L., Gel Y.R. & Miao W. 2009. The impact of Levene's test of equality of variances on statistical theory and practice. *Statistical Science* **24**(3): 343-360.
- Ghasemi A. & Zahediasl S. 2012. Normality tests for statistical analysis: A guide for non-statisticians. *International Journal of Endocrinology Metabolism* **10**(2): 486-489.
- Gorbunova A.A. & Lemeshko B.Y. 2012. Application of parametric homogeneity of variances tests under violation of classical assumption. *Proceedings of the 2nd Stochastic Modeling Techniques and Data Analysis International Conference*, pp. 253-260.
- Gravetter F.J. & Wallnau L.B. 2000. *Statistics for the Behavioral Sciences*. 5th Ed. Belmont: Wadsworth -

Thomson Learning.

- Hartley H.O. 1950. The use of range in analysis of variance. *Biometrika* **37**(3/4): 271-280.
- Kim Y.J. & Cribbie R.A. 2018. ANOVA and the variance homogeneity assumption: Exploring a better gatekeeper. *British Journal of Mathematical and Statistical Psychology* **71**(1): 1-12.
- Klotz S. & Johnson N.L. 1993. *Breakthroughs in Statistics: Volume 1: Foundations and Basic Theory*. NY: Springer.
- Kolmogorov A.N. 1956. *Foundations of the Theory of Probability*. 2nd English Eds. NY: Chelsea Publishing Company.
- Lee H.B., Katz G.S. & Restori A.F. 2010. A Monte Carlo study of seven homogeneity of variance tests. *Journal of Mathematics and Statistics* **6**(3): 359-366.
- Legendre P. & Borcard D. 2000. Statistical comparison of univariate tests of homogeneity of variances. *Unpublished*.
- Lemeshko B.Y., Lemeshko S.B. & Gorbunova A.A. 2010a. Application and power of criteria for testing the homogeneity of variances. Part I. Parametric criteria. *Measurement Techniques* **53**(3): 237-246.
- Lemeshko B.Y., Lemeshko S.B. & Gorbunova A.A. 2010b. Application and power of criteria for testing the homogeneity of variances. Part II. Nonparametric criteria. *Measurement Techniques* **53**(5): 476-486.
- Levene H. 1960. Robust tests for equality of variances. In Olkin I. (ed.). *Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling*. CA: Stanford University Press.
- Mazahreh A.S., Hammad H. & Abu-Jaber H. 2009. The attitudes of instructors and faculty members about the quality of technical education programs in community colleges in Jordan. *Journal of Social Sciences* **5**(4): 401-407.
- Mendes M., Ozcaya Turhan N. & Gurbuz F. 2006. A new alternative in testing for homogeneity of variances. *Journal of Statistical Research* **40**(2): 65-83.
- Nordstokke D.W. & Zumbo B.D. 2010. A new nonparametric Levene test for equal variances. *Psicológica* **31**(2): 401-430.
- O'Brien R.G. 1978. Robust techniques for testing heterogeneity of variance effects in factorial designs. *Psychometrika* **43**(3): 327-342.
- O'Brien R.G. 1981. A simple test for variance effects in experimental designs. *Psychological Bulletin* **89**(3): 570-574.
- Oladejo N. & Adetunde I. 2009. A numerical test on the Riemann hypothesis with applications. *Journal of Mathematics and Statistics* **5**(1): 47-53.
- Overall J.E. & Woodward J.A. 1976. A robust and powerful test for heterogeneity of variance. University of Texas Medical Branch Psychometric Laboratory.
- Patil K.P. & Kulkarni H.V. 2022. An uniformly superior exact multi-sample test procedure for homogeneity of variances under location-scale family of distributions. *Journal of Statistical Computation and Simulation* **92**(18): 3931-3957.
- Peterman R.M. 1990. The importance of reporting statistical power: The forest decline and acidic deposition example. *Ecology* **71**(5): 2024-2027.
- Rana M.S., Midi H. & Imon A.R. 2008. A robust modification of the Goldfeld-Quandt test for the detection of heteroscedasticity in the presence of outliers. *Journal of Mathematics and Statistics* **4**(4): 277-283.
- Samiuddin M. & Atiqullah M. 1976. A test for equality of variances. *Biometrika* **63**(1): 206-208.
- Shapiro S. & Wilk M. 1965. An analysis of variance test for normality (complete samples). *Biometrika* **52**(3/4): 591-611.
- Sharma D. & Kibria B.G. 2013. On some test statistics for testing homogeneity of variances: A comparative study. *Journal of Statistical Computation and Simulation* **83**(10): 1944-1963.
- Smirnov S.S. 1935. Zur Copepodenfauna des Amur-Limans. K faune Copepoda Amurskogo limana. *Issledovaniya Morei SSSR, Explorations des Mers de l'URSS* **22**: 41-53.
- Thomas L. & Juanes F. 1996. The importance of statistical power analysis. *Animal Behaviour* **52**(4): 856-859.
- Tomarken A.J. & Serlin R.C. 1986. Comparison of ANOVA alternatives under variance heterogeneity and specific noncentrality structures. *Psychological Bulletin* **99**(1): 90-99.
- Underwood A.J. 1997. *Experiments in Ecology: Their Logical Design and Interpretation using Analysis of Variance*. Cambridge: Cambridge University Press.
- Vorapongsathorn T., Taejaroenkul S. & Chukiat V. 2004. A comparison of type I error and power of Bartlett's test, Levene's test and Cochran's test under violation of assumptions. *Songklanakarin Journal of Science & Technology* **26**(4): 537-547.
- Wang Y., Rodríguez de Gil P., Chen Y.H., Kromrey J.D., Kim E.S., Pham T., Nguyen D. & Romano J.L. 2017. Comparing the performance of approaches for testing the homogeneity of variance assumption in one-factor ANOVA models. *Educational and Psychological Measurement* **77**(2): 305-329.
- Weerahandi S. 1995. ANOVA under unequal error variances. *Biometrics* **51**(2): 589-599.
- Yi Z., Chen Y.H., Yin Y., Cheng K., Wang Y., Nguyen D., Pham T. & Kim E.S. 2022. Brief research report: A comparison of robust tests for homogeneity of variance in factorial ANOVA. *The Journal of Experimental*

*Nur Fazlin Abdullah & Nora Muda*

*Education* **90**(2): 505-520.  
Zar J.H. 1999. *Biostatistical Analysis*. 4th Ed. NJ: Prentice Hall.

*Department of Mathematical Sciences*  
*Faculty of Science and Technology*  
*Universiti Kebangsaan Malaysia*  
*43600 UKM Bangi*  
*Selangor DE, MALAYSIA*  
*E-mail: fazlinabdullah03@yahoo.com, noramuda@ukm.edu.my\**

Received: 17 October 2022  
Accepted: 2 November 2022

---

\*Corresponding author