IOSR Journal of Research & Method in Education (IOSR-JRME)
e- ISSN: 2320–7388,p-ISSN: 2320–737X Volume 8, Issue 5 Ver. II. (Sep. – Oct. 2018), PP 06-10
www.iosrjournals.org

Testing of Normal Multivariate and Homoscedasticity of Variance-Covariance on Learning Achievement and Credits That Have Been Granted

Rusli[1], Syamsul Bakhri Gaffar[2], Jasruddin[3], M. Arifin Ahmad[4]
[1,2,3,4] *Education Department, Universitas Negeri Makassar, Indonesia*
Corresponding Author: Rusli1

Abstract : *A good future is the dream of every human being. In order to achieve a good future, education is the answer. A place where the process of formation of quality human resources is through a higher education institution, especially State Universities (PTN). Therefore, this study focuses on Student Achievement of Universitas Negeri Makassar class of 2014. The purpose of this research is to know the difference of learning achievement and credits that have been granted of class 2014 based on entry point, gender, school accreditation and source of cost. This comparison used a two population multivariate test as it consists of two dependent variables; GPA and Credits that have been granted. While for each population consists of four variables, namely gender, entrance selection, school accreditation and source of cost. However, in doing the comparison, the assumption of multivariate test must also be fulfilled first. In this case, the study was conducted only for testing the multivariate and multivariate normal assumptions of homoscedasticity of two populations. After that, the result of tests performed was good and fulfilled for the assumption of normal multivariate and multivariate homoscedasticity.*

Keywords: *Education, GPA, Graduate Credit, Normal Multivariate, Multivariate Homoscedasticity*

Date of Submission: 03-09-2018 Date of acceptance: 18-09-2018

I. Introduction

A good future is the dream of every human being. One way to achieve a good future is through education because education is the most important element in human life. Education is not easy to achieve because it takes seriousness and sincerity in achieving it. Therefore, education is very important for human life. A place where the process of formation of quality human resources is through higher education institutions, especially State Universities (PTN).

Makassar State University is one of the state universities has succeeded in creating quality human resources. This is due to the selection of students, upbringing during the lecture, and the completion of the course follows the rules that have been set by the government.

According to the Law Number 2 Year 2015 on the Admission of New Students of Undergraduate Program at State Universities, the pattern of admission of new undergraduate program is conducted through: (1) National Selection of State University Entrance (SNMPTN), (2) Joint Selection of State University Entrance (SBMPTN), and (3) Independent Selection.

Based on the above pattern of admission, it is important to conduct research on the potential of input quality (SNMPTN, SBMPTN, and Mandiri), as a reference of UNM to determine the percentage of quotas for each of the existing channels.

In order to know the characteristics of the students, it is necessary to conduct a comprehensive study on the achievement of student achievement based on the selection of the entry by gender, school accreditation, and financing (bidik misi and regular).

The purpose of this study is to determine the differences in student achievement and graduate credit force of 2014 based on entrance, gender, school accreditation and source of cost. This comparison uses a two population multivariate test as it consists of two dependent variables ie GPA and Graduate Credit. While for each population consists of four variables, namely gender, entrance selection, school accreditation and source of cost. But in doing the comparison, the assumption of multivariate test must also be fulfilled first. So in this case, the study was conducted only to test the multivariate and multivariate normal assumptions of homoscedasticity of two populations.

II. Literature Review

2.1 Learning Achievement

Learning achievement is one of the main goals that students want to achieve as participants of learning activities. Achievement is the result of an activity that has been done, created both individually and in groups [2]. While Chomadi states that "learning achievement is a result achieved during a lesson at a certain period in an educational institution where results are expressed by judgment, which can be realized in numbers or symbols" [4], achievement is what has been created, , rewarding results obtained by way of work perseverance. Darmadi [6] states that "learning achievement is a skill or success obtained by a person after performing an activity and learning process so that in a person experiencing behavioral changes in accordance with the competence of learning.

2.2 Multivariate Analysis

Multivariate data is the measurement of data on some variables that are generally denoted as xij. The notation shows the particular price of the i-th observation of the jth variable if there are p variables and there are as many as n observations, which can be shown in the matrix form as below, where xnp is the nth observational object for the pth variable [3].

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ x_{n1} & x_{n2} & x_{n3} & x_{np} \end{bmatrix} \qquad (1)$$

2.3 Variant-Covarian Matrix

In order to analyze data easier, sample variance-covariance can be arranged in a *p* x *p*-sized matrix as follows.

$$S = \begin{bmatrix} S_{11} & S_{12} & \cdots & S_{1p} \\ S_{21} & S_{22} & \cdots & S_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ S_{p1} & S_{p2} & \cdots & S_{pp} \end{bmatrix} \qquad (2)$$

The main diagonal of the matrix expresses variance, and besides the main diagonal represents covariance [3].

2.4 Assumption Test for Multivariate Data

Basically, the main distribution and problems that arise in the multivariate analysis is the multivariate normal distribution. The multivariate normal is an extension of the normal univariate. The assumptions that must be met include the data on the free variable should be multivariate normal distribution and the similarity of the variance matrix of covariance between groups / populations. Therefore a multivariate normality test is needed to determine whether the data follows a multivariate normal distribution [5].

2.4.1 Multivariate Normality Test

To check data on whether multivariate normal distribution, can be seen from scatter plot or Q-Q plot between square distance $\left(d_j^2\right)$ with the quantile value of Chi-Square distribution $\left(\frac{j-0,5}{n}\right)$. If the result plot describes a straight line then the data can be expressed as normal multivariate [3].

Hypothesis testing:

H0: multivariate normal distribution data.

H1 : the data is not multivariate normal distribution.

A normal multivariate examination is performed with the following steps [3].

1. Calculate the square distance value *(d^2)* for each observation.

$d_j^2=(X_j - \overline{X})^t S^{-1} X_j - \overline{X})$, *j* = 1, 2, 3, …, *n*.

2. Sets the value of d_i^2 all observations obtained from the above calculations such that $d_1^2, \leq d_2^2 \leq d_3^2 \leq \ldots \leq d_n^2$

3. Create a Q-Q plot or Chi-Square plot with the value d_j^2 as the X axis and the upper quantile value

$$q_{i,p}(p_i) = \chi_p^2\left(\frac{n-i+\frac{1}{2}}{n}\right)$$

as the Y axis.

A criterion fails to reject H0, which means multivariate normal distribution data, visually visible from scatter plot or Q-Q plot formed. If the plot forms a straight line then the data follows the normal distribution. Furthermore, the criteria for the fulfillment of normal multivariate assumptions can be determined by the test statistic formulated as follows [3].

$$r_q = \frac{\sum_{j=1}^{n}(x_j - x)(q_j - \overline{q})}{\sqrt{\sum_{j=1}^{n}(x_j - \bar{x})^2}\sqrt{\sum_{j=1}^{n}(q_j - \bar{q})^2}} \qquad (4)$$

Rejection region: Reject $r_Q < r_{n,\alpha}$

Where r_Q is the correlation coefficient between $q_j = X^2_{p;\left(\frac{j-0.5}{n}\right)}$ and $X_j = d_j^2$ and $r_{n,\alpha}$ is the critical point of the scatter plot (Q-Q plot) in the correlation coefficient test table for normality.

2.4.2 Matrix Equations Test of Variance-Covariance

Examination of the equality matrix of covariance variance between two populations or more is done with Box's M test formulated as follows [4]..

Hypothesis testing:

H0 : $\sum 1 = \sum 2 = \ldots = \sum k = \sum$ (covariance matrix is multivariate homoscedasticity)

H1 : at least one $\sum i \neq \sum j$ (covariance matrices not multivariate homoscedasticity)

Test Statistics: $C = (1-u)M$, where

$$u = \left[\sum_l \frac{1}{(n_l - 1)} - \frac{1}{\sum_l (n_l - 1)}\right]\left[\frac{2p^2 - 2p - 1}{6(p+1)(q-1)}\right]$$

$$M = \left[\sum_l (n_l - 1)\right] ln|S_{pooled}| - \sum_l [(n_i - 1) ln|S_l|]$$

$$S_{pooled} = \frac{1}{\sum_l (n_l - 1)}\{(n_1 - 1)S_1 + \cdots + (n_g - 1)S_g\}$$

The critical area rejects H_o

if $C > X^2_{\theta(\alpha)}$ with $\gamma = \frac{1}{2}p(p+1)(q-1)$

2.5 Multivariate Normality Test

Multiple comparison statistic method requires the assumption of normality distribution with hypothesis:

H_0: Normally distributed multivariate data

H_1: Data is not multivariate normal distribution.

If $X_1, X_2, \ldots, X_n$ are normally multivariate distributed then $(X - \mu)^2\ \Sigma^{-1}((X - \mu)$ is distributed γ_p^2 based on these properties then multivariate normal distribution can be done in each population by making scatter plots from the value $d_i^2 = (X_i - \overline{X})^t S^{-1} X_i - \overline{X})$, $i = 1, 2, 3, \ldots, n$. and the value of *Chi-Square* (*qi*).

The stages of making this scatter plot are as follows:

1. Determine the average vector value $\overline{X}$
2. Determine the value of the variance matrix - covariance *S*.
3. Determining the magnitude of a common carpark or square distance of each observation point by its average vector $d_i^2 = (X_i - \overline{X})^t S^{-1} X_i - \overline{X})$, $i = 1, 2, 3, \ldots, n$.
4. Sorting values d_i^2 from small to large $d_1^2, \leq d_2^2 \leq d_3^2 \leq \ldots \leq d_n^2$
5. Determine the value of $p_i = \frac{i - 1/2}{n}$, $n = 1, 2, 3, \ldots, n$.
6. Determine the value of q_i such that $\int_{-\infty}^{q_i} f(\chi^2) d\chi^2 = p_i$ atau $q_{i,p}(p_i) = \chi_p^2(\frac{n - i + \frac{1}{2}}{n})$
7. Make scatter-plot 0r Q-Q Plot d_i^2 with q_i
8. If the plot scatter (Q-Q) this tends to form a straight line, then the received H_0 means the data is normally distributed.

III. Research Methods

3.1 Data Source

Data collection in this research was done by documentation technique. Documentation data in this research were sourced from (1) database of Academic Information System (SIA) of UNM and (2) database of SNMPTN - SBMPTN central committee, (3) database of school accreditation system. Data obtained from SIA UNM in the form of student achievement data (GPA), gender, graduation credit, and source of cost. Data obtained from the central committee data base is the data of entrance selection, while the data obtained data base accreditation system accreditation school accreditation school grades school students from the origin. The number of sample members in this study is 3436 Students.

3.2 Research Variables

The variables used in this study are Student Achievement (GPA), Graduate Credit, Entrance Path Selection (SNMPTN, SBMPTN, Mandiri), Gender, Source of Cost (Bidikmisi and Nonbidikmisi), and Accreditation of origin school.

3.3 Data Analysis Method

The method of analysis used in the study was as follows:

1. Identify the data obtained into two dependent variables namely GPA and Credits that have been granted of UNM Students class of 2014.
2. . Identify the variables studied such as Entrance Path, Gender, Source of Cost, and School Accreditation.
3. Perform normal multivariate assumption test for both dependent variable that is GPA and Credit Graduated.
4. Next test the homoscedasticity assumption of Variance-Covariance Matrix using Box's value M.

IV. Results And Discussion

Parametric Multivariate analysis is one of the analytical tools in statistics that have several assumptions that must be met. This assumption is the normality assumption of covariance variables-data and homoscedasticity matrix. This study used two bound variables namely GPA and credits that have been granted of UNM Students class of 2014. Here is a result of normal multivariate and homoscedasticity variance-covariance matrix:

4.1 Normal Multivariate Assumption Test

Based on the calculation results according to steps Mahalonobis distance and Chi-Square value (q_i) from the data in this study obtained scatter-plot between the distance hormones with qi as follows

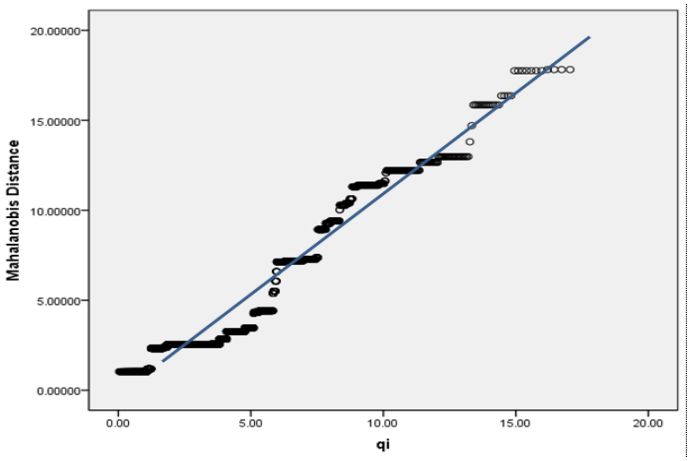

Figure 1. Chi-square plot with Mahalonobis Distance

Based on the result of scatter plot in figure 1, it can be seen that the plot of the hormone long distance and the chi-square tends to form a straight line, and because of the large number of samples using the central limit theorem, it can be concluded that the research data was multivariate normal distribution. Furthermore, we will examine the similarity of variance / covariance matrix test through Box's M.

4.2 Assay Test of the Variant-Covariance Matrix Diversity

This homoscedasticity test can use the value of Box's M. The hypothesis of this test is:

Ho : $\sum 1 = \sum 2 = \ldots = \sum n$

H_1 : at least one delivered a pair Σ which is not the same.

The test criterion for the significant value (α) is Ho is rejected if $C > \chi p(p+1)(g-1)/2(\alpha)$, otherwise H_o is accepted.

Based on table 1, the value of *P*_value = 0.706> α = 0.05, means the matrix of variance and homogeneous covariance. So the two dependent variables ie GPA and Credit Graduated meet the assumption of homogeneity based on related independent variables.

Table 1. Value of test statistics Homogeneity matrix variant-covariance

variant-covariance

Statistics	**Statistics value**
Box'M	1072.351
F	8.536
df1	117
df2	13499.027
Siq.	0.706

Based on the research, the dependent variable that is GPA and Credit pass fulfil normal test of multivariate and homoscedasticity of variance-covariance matrix so that can be used for further research like MANOVA manipulation test.

V. Conclusion

In assumption testing of multivariate normal distribution of dependent variable of GPA and credits that have been passed showed statistically significant with 95% confidence level. This indicates that both the variables and their associated free variables can be used in the parametric statistical methods. In addition, the assumption of the variant-covariant variant of these two variables is also met which is indicated by the Box'M statistical value.

For further research it is better to do formal assumption test analysis related to multivariate normality test. If this test is not met by formal test it can be used nonparametric comparison method.

References

[1] Anonim, 2012. Undang-Undang Nomor 12 tahun 2012, tentang Pendidikan Tinggi.
[2] Syaiful Bahri Djamara. 1994. Prestasi Belajar dan Kompetensi Guru. Surabaya: Usaha Nasional
[3] Maulidya. 2007. Perbandingan Analisis Diskriminan dan Regresi Logistik. Tugas Akhir S1: Jurusan Matematika, UNS.
[4] Herman Hidayat & Kartono Sukarto. 2010. Pengaruh Minat Belajar Siswa terjadap Prestasi Belajar Mata Diklat PDTM. Jurnal
[5] Pendidikan Teknik Mesin. Vol. 10 No. 1, Juni 2010. Versi Elektronik. Diakses dari http: www. journal.unness.ac.id Diakse 18 Oktober 2017
[6] Sharma, Subhash. 1996. Applied Multivariate Techniques. Canada: John Wiley & Sons.
[7] Nico. 2012. Prestasi Belajar. (http://elnicovengeance.wordpress.com /2012/09/30/ prestasi-belajar/ diakses 16 agustus 2017)

Rusli1. "Testing of Normal Multivariate and Homoscedasticity of Variance-Covariance on Learning Achievement and Credits That Have Been Granted" IOSR Journal of Research & Method in Education (IOSR-JRME) , vol. 8, no. 5, 2018, pp. 06-10.