IOP Conf. Series: Materials Science and Engineering **352** (2018) 012054 doi:10.1088/1757-899X/352/1/012054

Description of the supporting factors of final project in Mathematics and Natural Sciences Faculty of Syiah Kuala University with multiple correspondence analysis

Asep Rusyana, Nurhasanah and Maulizasari
Study Program of Statistics, Syiah Kuala University, Banda Aceh, 23111, Indonesia

Email: asep.rusyana@unsyiah.ac.id

Abstract. Syiah Kuala University (Unsyiah) is hoped to have graduates who are qualified for working or creating a field of work. A final project course implementation process must be effective. This research uses data from the evaluation conducted by Mathematics and Natural Sciences Faculty (FMIPA) of Unsyiah. Some of the factors that support the completion of the final project are duration, guidance, the final project seminars, facility, public impact, and quality. This research aims to know the factors that have a relationship with the completion of the final project and identify similarities among variables. The factors that support the completion of the final project at every study program in FMIPA are (1) duration, (2) guidance and (3) facilities. These factors are examined for the correlations by chi-square test. After that, the variables are analyzed with multiple correspondence analysis. Based on the plot of correspondence, the activities of the guidance and facilities in Informatics Study Program are included in the fair category, while the guidance and facilities in the Chemistry are included in the best category. Besides that, students in Physics can finish the final project with the fastest completion duration, while students in Pharmacy finish for the longest time.

1. Introduction

The process of education in universities begins with the introduction of basic knowledge in the field of interest and ends with research in the form of a final project. The final project is a combination of cognitive, psychomotor and affective competencies in students by the field of science that is learned. The final project is a compulsory course that must be taken by the student as the main requirement to obtain a Bachelor of Science degree. National Accreditation Board for Higher Education (2008) stated that the best time to the completion of the task is less or equal to six months. Since 2013, the number of undergraduate graduates in FMIPA of Syiah Kuala University has continued to increase. FMIPA has a Faculty Quality Assurance Unit (SJMF) team that acts as a quality assurance institution. In 2016 the SJMF MIPA team conducted an academic evaluation of the completion of the students' final assignment. The goal was to know student constraints in doing the final project. The scope of research conducted by SJMF includes characteristics of students, duration of final project completion, guidance process, facilities, final project seminar, and general and quality effect. This research provides qualitative data information consisting of several variables with various categories. Each variable will be analyzed firstly through the chi-square test, then the related variables will be analyzed further using correspondence analysis.

IOP Conf. Series: Materials Science and Engineering **352** (2018) 012054 doi:10.1088/1757-899X/352/1/012054

2. Literature review

2.1. Chi-square test

In the correspondence analysis, the strength of relationships among categories is determined based on the value of chi-square. The chi-square value indicates the relationship between variables [2].

The hypothesis as follows:

H_o: There is no relationship between rows and columns

H_1: There is relationship between rows and columns

The formula for calculating the chi-square test:

$$X^2 = \Sigma_{ij} \frac{(x_{ij} - E_{ij})^2}{E_{ij}} \quad (1)$$

Where :

x_{ij} = The observed frequencies on the i-th row and the jth columns

E_{ij} = The expected frequencies on the i-th row and the jth columns

Where :

$$E_{ij} = \frac{x_{i.} x_{.j}}{x_{..}} \quad (2)$$

[1].

With the following criteria:

H_o is rejected, if X^2 value > X^2 table with [α, df = (a-1) (b-1)], where α is an error of type I, a and b are the number of rows and columns, respectively.

H_o can not be rejected, if X^2 value ≤ X^2 table with [α, df = (a-1) (b-1)].

Or

P-value < α then H_o is rejected, and *P-value* > α then H_o cannot be rejected

2.2. Multiple Correspondence Analysis

The calculation of multiple correspondence analysis is a looping application of a simple correspondence analysis algorithm using the Burt matrix [5]. The Burt matrix is a combination of the following matrices:

$$\mathbf{Z}'\mathbf{Z} = n \begin{bmatrix} \mathbf{P}_r & \mathbf{P} \\ \mathbf{P} & \mathbf{P}_c \end{bmatrix} \quad (3)$$

where **P** is the correspondence matrix.

The matrix $\mathbf{P}_r$ and matrix $\mathbf{P}_c$ are diagonal matrices with diagonal elements marginal columns or marginal row matrix P, respectively.

2.2.1. Correspondence matrix

The relative frequency of matrix is:

$$\mathbf{P}_{axb} = (p_{ij}) = (\frac{n_{ij}}{n}) \quad (4)$$

With $n = \sum_i \sum_j n_{ij}$, $\mathbf{P}_{axb}$ is called the correspondence matrix, n_{ij} is the number of the object on row i and column j, and p_{ij} is the proportion of object on row i and column j.

The diagonal matrix of row number element elements is the matrix $\mathbf{D_r}$ of size (a x a), as follows:

$$\mathbf{D}_r = diag(r) = \begin{bmatrix} p_{1.} & 0 & \cdots & 0 \\ 0 & p_{2.} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & p_{a.} \end{bmatrix} \quad (5)$$

The diagonal matrix of column number element elements is the matrix $\mathbf{D_c}$ of size (b x b), as follows:

$$\mathbf{D}_c = diag(c) = \begin{bmatrix} p_{.1} & 0 & \cdots & 0 \\ 0 & p_{.2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & p_{.b} \end{bmatrix} \quad (6)$$

[4].

2.2.2. Decomposition of singular value

Singular decomposition is done by reducing the dimensions of data based on the greatest diversity of data (eigenvalues /inertia) while maintaining optimum information. The matrix of decomposition value yields matrix **A** with size *a x m* and matrix **B** with size *b x m* with m is the best low rank for **Λ**. The coordinate matrix ($\mathbf{P} - \mathbf{rc}'$) is **AΛB'** where **r = P 1** and **c = P` 1**. Where **Λ** diagonal matrices have the diagonal elements of the singular value of the matrix ($\mathbf{P} - \mathbf{rc}'$)**.**

2.2.3. Inertia

The value of inertia shows the contribution of the i-th row in total inertia. The column profiles are represented by the dimensional spaces d ≤ m, the coordinates of rows of the matrices formed by taking the first d columns of $\mathbf{F} = \mathbf{D_r^{-1}A\Lambda}$ for d = 2 or 3, and the coordinates of the column profiles are *j* rows of the matrices formed by taking *k* first columns of $\mathbf{G} = \mathbf{D_c^{-1}B\Lambda}$. Total inertia will explain all the information in all spaces $\boldsymbol{tr(E) = n\sum_{i=1}^{m} \lambda_i}$ where $\boldsymbol{E = D_r^{-\frac{1}{2}}(P - rc`)D_c^{-\frac{1}{2}}\left(D_r^{-\frac{1}{2}}(P - rc`)D_c^{-\frac{1}{2}}\right)`}$] and $\lambda_i$ are the singular values obtained from $\boldsymbol{D_r^{-\frac{1}{2}}(P - rc`)D_c^{-1/2}}$ [4].

3. Data and Methods

This research uses data from the research conducted by Quality Assurance Unit Team of Faculty of Mathematics and Natural Sciences of Syiah Kuala University in 2016. The respondents are 62 students consisting of students of Mathematics, Physics, Biology, Chemistry, Pharmacy and Informatics Study Program who have completed the final project in the academic year 2015/2016. The variables are defined as in Table 1.

Stages performed in the study are as follows:

1. Describe the characteristics of the respondent and the general description of the completion of the final project of FMIPA students of Syiah Kuala University.
2. Perform a chi-square test to know the relationship between the duration of the completion of the final task, the assessment of the students in completing the final task based on the guidance, facilities, the final seminar and the general impact and the quality of the study program at FMIPA.
3. Perform data processing using multiple correspondence analysis. The importance of each variable must be found in correspondence plot [7].

Table 1. Variables and Categories in this research.

No	Variable	Category
1	Study Program (Y)	1 = Biology, 2 = Pharmacy 3 = Physics 4 = Informatics 5 = Chemistry 6 = Mathematics
2	Completion duration of final project (X_1)	1 = ≤ 6 month 2 = 7-8 month 3 = 9-10 month 4 = 11-12 month 5 = >12 month (BAN-PT, 2008)
3	Guidance (X_2)	1 = Poor; 2 = Fair; 3 = Good; 4 = Excellent
4	Facilities (X_3)	1 = Poor; 2 = Fair; 3 = Good; 4 = Excellent
5	Final project seminar (X_4)	1 = Poor; 2 = Fair; 3 = Good; 4 = Excellent
6	Quality and general impact (X_5)	1 = Poor; 2 = Fair; 3 = Good; 4 = Excellent

4. Results and Discussion

4.1. Descriptive Analysis

Respondents in this research are students from five study programs at Faculty of Mathematics and Natural Sciences Syiah Kuala University. Some information about characteristics of respondents is in Table 2.

Based on Table 2 most of the respondents are women that is equal to 76%. Almost all students in good health condition when completing the final task that is equal to 97%. Only 11% of the student's finance is from their own while preparing the final task, while the rest is financed by others. The final assignment theme is determined by the students themselves 58% and 40% are determined by their supervisor. Furthermore, the determination of supervisor by the study program is as much as 77%.

Table 2. Percentage of respondent characteristics.

No.	Variabel	Category	n	Percentage
1	Gender	Male	15	24
		Female	47	76
2	Health condition	Healthy	60	97
		Not Healthy	2	3
3	Cost of education	Themselves	7	11
		Others (parent, scholarship, etc.)	55	89
4	Determination of topic of final project	Themselves	36	58
		Supervisor	25	40
		Others	1	2
5	The supervisor is determined by the Study Program	Yes	48	77
		No	14	23

4.2. Chi-square test

Table 3. The result of the chi-squared test between the variable of the study program and supporting variable of final project completion.

No.	Variable	Chi-Square	df	*p-value*
1	Guidance (X_1)	21.069	10	0.021*
2	Facility (X_2)	20.512	10	0.025*
3	Final project seminar (X_3)	16.310	10	0.091
4	Quality and general impact (X_4)	10.725	5	0.057
5	Duration (X_5)	36.853	20	0.012*

Note : * = Significant with α 0.05

Chi-square test is performed to find out the variables that are related to the study programs. More can be seen in the following Table 3.

By using the α value of 0.05, then the rejection criterion is used if p-value < α then reject H_0. Based on these criteria, guidance activities, facility and duration have a relationship with study program, the other words each of study programs have characteristics.

4.3. Multiple correspondence analysis

Based on the Chi-Square test, the duration of completion of the final project, guidance activities and facilities are available variables. Correspondence analysis results obtain the singular value, the value of inertia, the inertia proportion, and the cumulative proportion presented in the following Table 4.

Table 4. The singular value, the value of inertia, the proportions, and the cumulative proportions.

Dimension	Singular value	The value of Inersia	The Proportion of Inertia	Cumulative proportion
1	0.7109	0.5054	0.1555	0.1555
2	0.6745	0.4550	0.1400	0.2955
3	0.6080	0.3697	0.1138	0.4093
...	...	...	...	...
13	0.2827	0.0799	0.0246	1.0000
Total	6.263	3.25	1.00	-

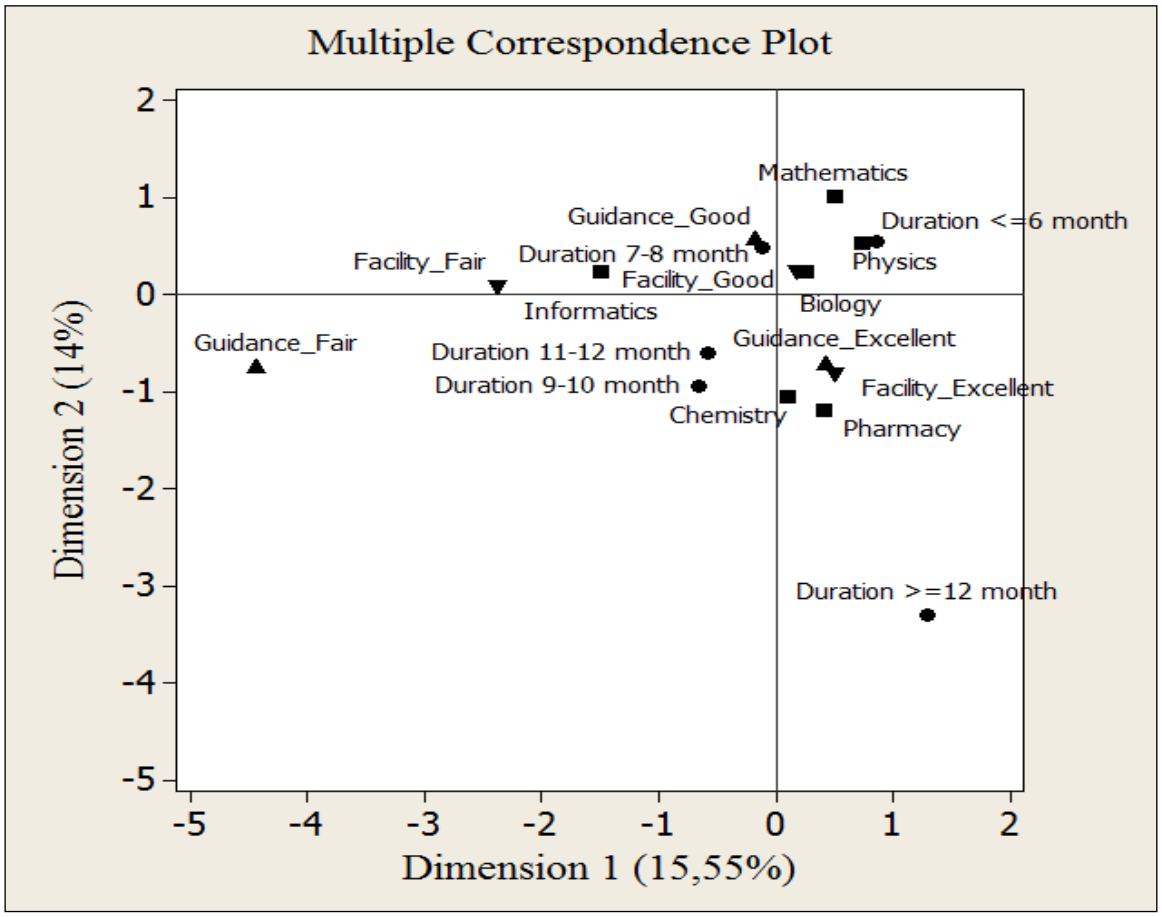

Figure 1. The plot of multiple correspondence analysis.

Table 4 shows that there are 13 dimensions of the results of multiple correspondence analysis. The data consist of 4 variables with 17 categories, so that this research can produce 17 – 4 = 13 dimensions. There are some criteria used to determine the representative model [6]. In the study, multiple correspondence analysis is only capable of displaying 2-dimensional graphics so that the first two root characteristics describe 29.55% from all information. They are 15.55% from the first dimension and 14% from the second dimension.

The result of multiple correspondence analysis in the form of correspondence plot can be seen in figure 1.It can explain about the scoring of students as respondents that:

1. Study Program of Mathematics has good facility and guidance, and ideal final project completing duration. These can be caused by computer laboratories and many Ph.D. lecturers.
2. Study Program of Physics has a good facility, good guidance, and the fastest duration. These can be caused because lecturers provide much time to guide student research.
3. Study Program of Biology has the good facility, good guidance, and duration for completing the final task between 7 and 8 months. They can be influenced by complete biology laboratories and tools.
4. Study Programs of Chemistry and Pharmacy consist of an excellent facility, excellent guidance, and greater than 7-month duration. These can be caused by advanced laboratories and also professor and Ph.D. lecturers.
5.Study Program of Informatics has a fair facility, good guidance activities and 7 to 8-month duration. These can be caused because this study program has not had a permanent building yet.

5. Conclusions

The conclusions of this study are as follows:

1. The factors that have a relationship with the completion of the final task in each study program at FMIPA Syiah Kuala University are facilities available, guidance activities, and the duration of completion of the final task.
2. Guidance activities and facilities available in Chemistry and Pharmacy are included into the excellent category. Whereas the fastest duration is owned by Physics Study Program. Guidance activities and available facilities in Informatics are included in the category fair while Pharmacy completed the final task for the longest time.
3. Chemistry and Pharmacy Study Programs have similar characteristics.

References

[1] Abdi H and Williams L J 2010 *Correspondence Analysis Encyclopedia of Resign Desain* (ed: Neil Salkind) (United States of America: Thousand Oaks: Sage Publications) 11

[2] Hair J F, Black W C, Babin B J and Anderson R E 2010 *Multivariate Data Analysis A Global Perspective 7th Edition* (United States of America: Pearson Education International)

[3] Johnson A R and Wichern D W 2007 *Applied Multivariate Statistical Analysis 6th Edition* (United States of America. Pearson Education International)

[4] Mattjik A A and Sumertajaya I M 2011 *Sidik Peubah Ganda dengan Menggunakan SAS* (Bogor: Statistics Departement of IPB) (in Indonesian)

[5] Nenadic O and Greenarce M 2005 *Computation of Multiple Correspondence Analysis with code in R*

[6] Rusyana A, Nurhasanah, Marzuki and Flancia M 2016 SARIMA Model for Forecasting Foreign Tourists at the Kualanamu International Airport. *Proc. Int. Conf. Mathematics, Statistics, and their Applications (ICMSA)* Banda Aceh IEEE p 155

[7] Rusyana A, Salwa N and Muzamil 2009 Analisis konjoin untuk menilai pembukaan Program Studi Statistika di Universitas Syiah Kuala *J. Pythagoras* **5** 14-23 (in Indonesian)