SigmaMu 1(2) (2022)

Journal of Mathematics, Statictics and Data Science

https://journals.balaipublikasi.id

Path Analysis of the Decomposition Method to Know the Factors Influence on the Completion of the Thesis

Gilang Primajati[1*], Fatmawati1, Wirajaya Kusuma [2]

[1]Mataram University, Indonesia
[2]University of Bumigora, Indonesia

Received: 23 Juli
Revised: 7 August
Accepted: 27 August
Published: 31 Sept

Corresponding Author Email:
Gilang Primajati
gilangeuler@gmail.com

Abstract: This study aims to perform modeling and determine the factors that significantly and not significantly influence the completion of the thesis. The analysis used is the path analysis of the decomposition method. The data used is primary data by distributing questionnaires. Based on the results of the analysis that has been done, the path analysis model for the factors that influence the completion of the thesis are:

$$Y_3 = -0.0312x_1 + 0.1159x_2 + 0.3575x_3 - 0.2233y_1 + 0.0905y_2 + 0.897$$

The factors that significantly effect are facilities (x_3), selection of specialization courses (y_1), and selection of courses (y_2). While lecturer factor (x_1) and motivation factor (x_2) have no effect. Supporting facilities factor and selection of supporting courses factor only directly affect the completion of minithesis. While the selection of specialization courses also has indirect effect to the completionof the thesis.

Key Word : Path Analysis, Decomposition Method

INTRODUCTION

Thesis is a scientific essay resulting from student research which is part of the academic education requirements and is also the final gate that must be passed by all undergraduate students to earn their bachelor's degree. In the process of working on the thesis does not always run smoothly. Students often find difficulties in writing their thesis. This will cause the length of time needed for a student to complete the thesis. According to [1] in his research, it states that there are several supporting factors and obstacles that influence the process of writing a thesis including: internet facilities, motivation, references in the form of previous thesis, parental support, difficulty finding problems, supporting courses, and supporting facilities.

Based on the results of observations made on Information Technology Program Students at Bumigora University, several factors also influenced the completion of the thesis, namely the selection of areas of interest and the selection of supporting courses. By choosing an area of interest, students can determine which supporting courses to take next. In choosing a field of interest, students consider several factors, namely lecturers, motivation, and supporting facilities. So, we need an analysis that can determine whether these factors have a direct effect on the completion of

the thesis or have an indirect effect (through intermediary variables). The analysis that can be used is path analysis. Path analysis is one of the multivariate analyzes used to determine the relationship pattern between the independent (exogenous) and dependent (endogenous) variables.

In this study, a model of the relationship between exogenous and endogenous variables was formed, namely the completion of the thesis and it was determined how much direct and indirect influence the exogenous variables had. The method used is the decomposition method. This method is a model building method by including all exogenous variables into the model, without eliminating variables that have no significant effect.

LITERATURE REVIEW

Path Analysis

Path analysis which is known as path analysis was developed in 1920 by an expert named Sewall Wright [2]. Path analysis and multiple regression analysis are related. Path analysis is a further development or expansion of multiple regression analysis. So that regression analysis can be said as a special form of path analysis.

In path analysis, it is known that there are terms of direct influence and indirect influence. Direct influence can be interpreted as the influence of an exogenous variable on endogenous variables that occurs without going through other endogenous variables. Meanwhile, the indirect effect is the effect of an exogenous variable on endogenous variables through other endogenous variables contained in the causal model being analyzed [3].

Before using path analysis, you must first model the relationship between variables, which is called a path diagram. The relationship between variables in the path diagram is expressed by two types of arrows, namely straight arrows with one head, arrows indicating a causal relationship between two variables and curved lines with two arrows indicating a correlational relationship between two variables [2].

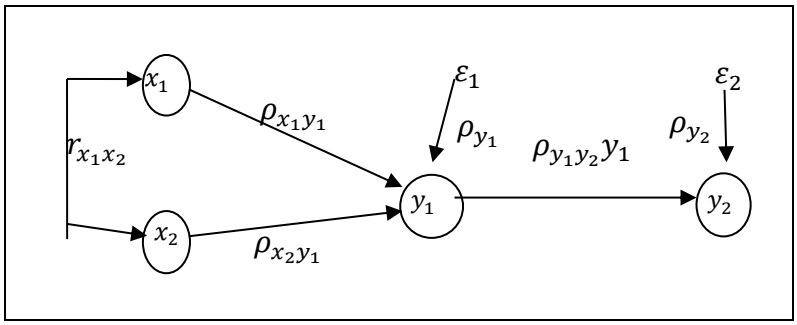

Figure 1. Model the relationship between variabels

In conducting path analysis there are several assumptions that must be met, namely the relationship between variables must be linear, additive, and causal and have a recursive relationship pattern. Residual variables should not be correlated with each other. The data used is an interval scale. Must meet classical assumptions as well as multiple regression analysis [4]. The method that can be used in path analysis is the decomposition method. The decomposition method is a model building method by including all exogenous variables into the model, without eliminating insignificant variables [3].

Exogenous and Endogenous Variables

The strength of the relationship between variables in a correlation is expressed by the correlation coefficient. Meanwhile, in path analysis there is a path coefficient. The path coefficient shows the strong influence of exogenous variables on endogenous variables [4]. In research [5], the calculation of the path coefficient is not much different from the method of estimating parameters in ordinary linear regression. The steps for calculating the path coefficient are as follows:

- Draw a path diagram along with its structural equation
- Calculating the correlation coefficient between variables using the *Product Moment Coefficient (r)* equation 1:
- $$\boldsymbol{r_{x_i x_j}} = \frac{\boldsymbol{n(\sum_{h=1}^{n} x_{ih}x_{jh}) - (\sum_{h=1}^{n} x_{ih})(\sum_{h=1}^{n} x_{jh})}}{\sqrt{\{\boldsymbol{n\sum_{h=1}^{n} x_{ih}^2 - (\sum_{h=1}^{n} x_{ih})^2}\}\{\boldsymbol{n\sum_{h=1}^{n} x_{jh}^2 - (\sum_{h=1}^{n} x_{jh})^2}\}}} \quad (1)$$

  Next, a correlation matrix between variables is formed, namely $\boldsymbol{R_{x_i x_j}}$ and $\boldsymbol{R_{x_i y}}$ for each of the substructures formed
- Calculating the inverse of the correlation matrix between exogenous variables $\boldsymbol{R_{x_i x_j}}^{-1}$ of each substructure.
- Calculating all path coefficients (ρ_{yx_i}) where $i = 1,2,3, \ldots k$ with the equation 2:

  $$\boldsymbol{[\rho_{yx_i}] = \left[R_{x_i x_j}\right]^{-1}[R_{x_i y}]} \quad (2)$$

In path analysis, the coefficient of determination is used to calculate the value of the residual coefficient (ρ_y). The residual coefficient is the magnitude of the influence of other variables outside the model that are not observed. To calculate the value of the residual coefficient, the following equation 3:

$$\boldsymbol{\rho_{y_k} = \sqrt{1 - R^2}} \quad (3)$$

In research [5], after the path coefficient value is obtained, simultaneous and partial testing is carried out like ordinary regression analysis. Testing simultaneously (simultaneously) using the F test. The hypothesis used in the simultaneous test is:

$H_0 : \rho_{yx_1} = \rho_{yx_2} = \cdots = \rho_{yx_k} = 0$

$H_1 : \rho_{yx_1} = \rho_{yx_2} = \cdots = \rho_{yx_k} \neq 0$

The F test statistics used are as equation 4:

$$\boldsymbol{F = \frac{(n-k-1)R^2}{k(1-R^2)}} \quad (4)$$

where n is the number of samples, k is the number of exogenous variables, and R^2 is the coefficient of determination. The decision making criterion is that if the value of $F_{hitung} > F_{\alpha\,;(k,n-k-1)}$ then H_0 is rejected. Meanwhile, the partial test is used using the t test, where the hypothesis used is as follows:

$H_0 : \rho_{yx_k} = 0$

$H_1 : \rho_{yx_k} \neq 0$

The t test statistic for the partial test is, as equation 5.

$$\boldsymbol{t = \frac{\rho_{yx_i}}{\sqrt{\frac{(1-R^2)C_{ii}}{n-k-1}}}} \quad (5)$$

where n is the number of samples, k is the number of exogenous variables, ρ_{x_i} is the path coefficient, C_{ii} is the inverse value of the correlation matrix, and R^2 is the coefficient of determination. The decision making criterion is if the value of $t_{hitung} > t_{\frac{\alpha}{2};(n-k-1)}$ then H_0 is rejected.

Furthermore, according to the assumptions in the path analysis, it is necessary to test the residual assumptions, namely whether the residuals meet the assumptions of identical, independent, and normality assumptions or meet the IIDN assumptions with a mean of 0 and a variance of σ^2 as in the case of multiple regression analysis.

METHOD

The population in this study were students of the Information Technology Study Program, Faculty of Engineering, Bumigora University, from class of 2019 to class of 2022. The data used is primary data obtained by distributing questionnaires directly to 100 people.

In this study, there are two types of variables, namely exogenous and endogenous variables. Exogenous variables include lecturers (x_1), motivation (x_2), and supporting facilities (x_3Meanwhile, the endogenous variables are interest field selection (y_1), course selection (y_2), and thesis completion (y_3).

The first step in this study was to develop a Likert scale questionnaire. Furthermore, the questionnaire was tested on 30 respondents and a validity and reliability test was carried out. If the data is valid and reliable, then the required sample size is determined and the questionnaire is redistributed. Furthermore, the data obtained were analyzed using the path analysis decomposition method. The first step is to create an overall path diagram. Then formed each substructure and substructural equations. For each substructure, the value of the path coefficient is sought using formula (2), after previously looking for the correlation between variables using formula (1). Calculation of the residual coefficient is also carried out using formula (3). The path coefficient that has been obtained is then tested simultaneously with formula (4) and partially using formula (5). After testing the path coefficient, proceed with testing the residual assumptions. The final step is to determine which factors have a direct and indirect effect on the endogenous variables.

RESULTS AND DISCUSSION

Based on the relationship between each exogenous variable and endogenous variable, the overall path diagram can be formed as follows:

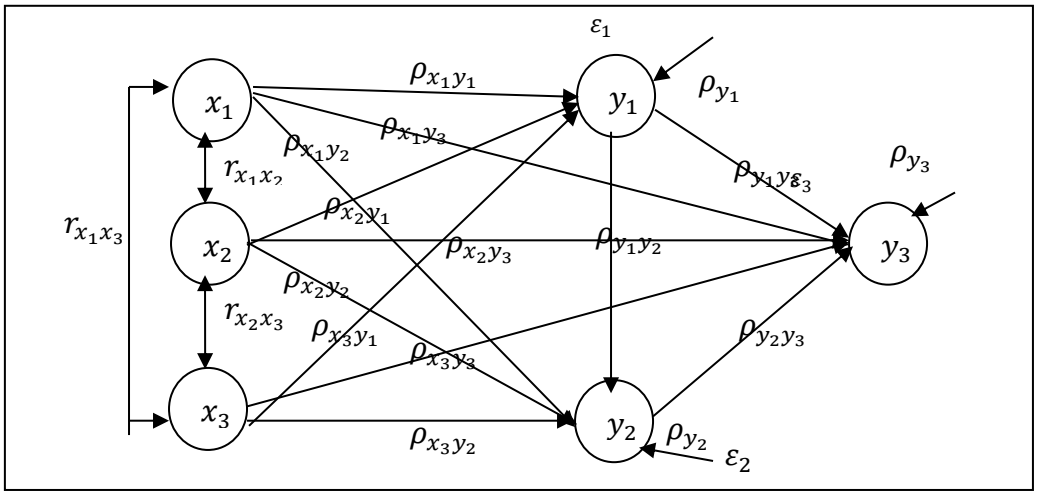

Figure 2. Overall Path Diagram

Based on the path diagram above, the substructural equation can be formed as follows:

$$Y_1 = \rho_{x_1y_1}x_1 + \rho_{x_2y_1}x_2 + \rho_{x_3y_1}x_3 + \rho_{y_1}\varepsilon$$

$$Y_2 = \rho_{x_1y_2}x_1 + \rho_{x_2y_2}x_2 + \rho_{x_3y_2}x_3 + \rho_{y_1y_2}y_1 + \rho_{y_2}\varepsilon$$

$$Y_3 = \rho_{x_1y_3}x_1 + \rho_{x_2y_3}x_2 + \rho_{x_3y_3}x_3 + \rho_{y_1y_3}y_1 + \rho_{y_2y_3}y_2 + \rho_{y_3}\varepsilon$$

For the three structural equations, then analyzed using path analysis decomposition method.

Substructural equation 1 describes the effect of x_1, x_2, and x_3 on y_1. The following is the path coefficient value obtained along with the results of the substructure path coefficient significance test 1 :

Table 1. Path coefficient for substructure equation 1

From	to	Koef. path	t_{hit}	Sig.	F_{hit}	Sig.	R^2
x_1		−0,119	−1,133	No			
x_2	y_1	0,0013	0,0126	No	1,354	No	0,0406
x_3		0,203	1,8781	No			

Based on table 1 it can be seen that the value of $F_{hit} < F_{tabel} = 2{,}699$ and all values of $t_{hit} < t_{tabel} = 1{,}985$ which means that both simultaneously and partially x_1, x_2, dan $x_3$have no effect on y_1.

Substructural equation 2 describes the effect of x_1, x_2, x_3 and y_1 on y_2. The following is the path coefficient value obtained along with the results of the substructure path coefficient significance test 2 :

Table 2. Path coefficient for substructure equation 2

From	to	Koef. pat	t_{hit}	Sig.	F_{hit}	Sig.	R^2
x_1		0,0694	0,0631	No			
x_2	y_2	0,0752	0,7377	No	1,8979	No	0,962
x_3		−0,0704	−0,6460	No			
y_1		0,2640	2,6188	Yes			

Based on table 2 it can be seen that the value of $F_{hit} < F_{tabel} = 2.4675$ which means that simultaneously x_1, x_2, x_3, dan $y_1$have no effect on y_2. Meanwhile, partially only y_1 has an effect when compared to the value of $t_{tabel} = 1.9853$.

Substructural equation 3 describes the effect of x_1, x_2, x_3, y_1 and y_2 on y_3 . The following is the path coefficient value obtained along with the results of the substructure path coefficient significance test 3.

Table 3. Path coefficient for substructure equation 3

From	to	Koef. path	t_{hit}	Sig.	F_{hit}	Sig.	R^2
x_1		−0,0312	−0,0312	No			
x_2		0,1159	1,2088	No			
x_3	y_3	0,3575	3,4899	Yes	4,5251	Yes	0,194
y_1		−0,2233	−2,2808	Yes			
y_2		0,0905	3,466	Yes			

Based on table 3 it can be seen that the value of $F_{hit} > F_{tabel} = 2{,}3112$ which means that simultaneously x_1, x_2, x_3, y_1 dan $y_2$have an influence on y_3 While partially x_3, y_1, dan y_2 have an effect on $y_3$when compared to the value of $t_{tabel} = 1{,}9855$. Nilai $\rho_{y_3}$is 0.897.

Furthermore, residual testing is carried out for substructure 3. Based on the results obtained for the substructural equation 3, the residuals fulfill the identical, independent, and normal assumptions so that direct and indirect effects on variable y_3 can be calculated. The results obtained can be seen as follows:

Table 4. Direct Influence, Indirect Influence, and Total

Variabel	Influence						Total
	Direct	Indirect					
		x_1	x_2	x_3	y_1	y_2	
$\boldsymbol{x_1}$	-	-	-	-	-	-	-
$\boldsymbol{x_2}$	-	-	-	-	-	-	-
$\boldsymbol{x_3}$	0,3575	-	-	-	-	-	0,3575
$\boldsymbol{y_1}$	$-0,2233$	-	-	-	-	0,0239	$-0,1994$
$\boldsymbol{y_2}$	0,0905	-	-	-	-	-	0,0905

Based on table 4 above, it can be seen that the factors that have a direct effect on y_3 are the variable x_3 of 0.3573 and y_2 of 0.0905. While the factor $y_1$has a direct and indirect effect on y_3, namely through y_2 of -0.1994.

4. CONCLUSION

Based on the results and discussion that has been done, the following conclusions can be drawn for this study:

- The path analysis model obtained for the factors that influence the completion of the thesis are

$$Y_3 = -0,0312x_1 + 0,1159x_2 + 0,3575x_3 - 0,2233y_1 + 0,0905y_2 + 0,897$$

- The factors that have a direct influence on the completion of the thesis are the supporting facilities and the selection of courses. Meanwhile, the selection factor for the field of interest has an indirect effect, namely through the selection of courses. Meanwhile, for lecturers and motivation factors have no influence on the completion of the thesis.

REFERENSI

Wiyatmo, Y., Mundilarto, Suharyanto, dan Widodo E, 2010, Efektivitas Bimbingan Tugas Akhir Skripsi (TAS) Mahasiswa Jurusan Pendidikan Fisika FMIPA UNY, Prosiding Seminar Nasional Peneltian Pendidikan dan Penerapan MIPA.

Usman, M., 2009, Model Linear Terapan, Sinar Baru Algensindo, Bandung.

Roflin, E., 2009, Penggunaan metode Trimming pada Analisis Jalur dalam Menentukan Model Kausal Dana Alokasi Umum Kabupaten/Kota di Provinsi Sumatera Selatan, Jurnal Penelitian Sains.

Sugiyono, 2012, Statistika untuk Penelitian, Alfabeta, Bandung.

Lumenta, C.Y., Kekenusa J.S. dan Hatidja D., 2012, Analisis Jalur Faktor-Faktor Penyebab Kriminalitas di Kota Manado, Jurnal Ilmah Sains, Vol. 12 No. 2 : 78-82.

Hadi, W. 2009. Hubungan Antara Self-Efficacy Dengan Penyesuaian Akademik Dan Prestasi Akademik (Studi Pada Mahasiswa FIP Universitas Negeri Surabaya). Jurnal Ilmiah Ilmu PendidikanVol. 9 No. 1.Kartono. 1999. Patologi Sosial. Jakarta: Rajagrafindo Persada.

Riduwan dan Sunarto H. 2007. Pengantar statistik untuk penelitian (pendidikan, sosial, ekonomi, komunikasi, dan bisnis). Alfabeta. Bandung.

Sarwono, J. 2007. Analisis Jalur untuk Riser Bisnis dengan SPSS. Andi. Yogyakarta.

Somantri, A dan Sambas. 2006. Aplikasi Statistika Dalam Penelitian. Pustaka Setia. Bandung

Suliyanto. 2005. Analisis Data Dalam Aplikasi Pemasaran. Ghalia Indonesia. Bogor.

Supranto. 2004. Analisis Multivariant Arti dan Interpretasi. Rineka Cipta. Jakarta.

Riduwan & E.A. Kuncoro, 2007, Cara Menggunakan dan Memaknai Analisis Jalur (Path Analysis), Alfabeta, Bandung.

Sarwono, J., 2007, Analisis Jalur untuk Riset Bisnis, Andi, Yogyakarta.

Supranto, J., 2004, Analisis Mutivariat Arti dan Interpretasi, Rineka Cipta, Jakarta.

Somantri, A. & S.A. Muhidin, 2006, Aplikasi Statistika dalam Penelitian, Pustaka Setia, Bandung.

Denis, Daniel J. and Joanna Legerski. (2006). Causal Modeling and the Origins of Analisis Jalur . University of Montana

Duncan, O. D., & Hodge, R. W. (1963). Education and occupational mobility: A regression analysis. The American Journal of Sociology , 629-644.

Garson, David (2011) Analisis Jalur. Diunduh dari http://faculty.chass.ncsu.edu/garson Hair, Joseph F. et al. (2010). Multivariate Data Analysis: A Global Perspective. New Jersey: Pearson Prentice Hall Johnson, Richard A. and Wichern, Dean W.(2002). Applied Multivariate Statistical Analysis. New Jersey: Prentice Hall Lleras ,

Christy (2011) Analisis Jalur Pennsylvania State University University Park Pennsylvania USA Olabutiyi, Moses. E.(2006). A User's Guide to Analisis Jalur Maryland: University Press of America Perdana, M.S.P. (2011). Uji Utilitas Multidimensional Self Concept Scale. Skripsi. Jakarta: Universitas Kristen Krida Wacana

Sarwono, Jonathan.(2010). Analisis Jalur Untuk Riset Bisnis dengan SPSS. Edisi 5. Yogyakarta: Penerbit Andi. Streiner, David L. Finding Our Way: An Introduction to Analisis Jalur . Can J Psychiatry, Vol 50, No.2 February 2005

Schumacker, Randall E. and Richard G. Lomax .(1996) A beginner's guide to structural equation modeling . New Jersey: Lawrence Erlbaum Associates Inc Wright,

Sewal. (1920). The relative importance of heredity and environment in determining the piebald pattern of guinea-pigs. Proceedings of the National Academy of Sciences , 320-332.

Wolfle, L. M. (2003). The introduction of Analisis Jalur to the social sciences, and some emergent themes: An annotated bibliography. Structural Equation Modeling , 1- 34.