

Chapter 4

Further Statistical Analysis

In Chapter Three we used some basic statistical methods for exploring and analysis the data. In this chapter we used multiple regression to fit a model with several possible predictors of the outcome.

4.1 Multiple Regression Analysis

Linear regression analysis is used to fit a model predicting the mean of each outcome, given the determinants. There were nine determinants in this study, namely, (1) gender, (2) religion, (3) degree duration, (4) high school type, (5) home province, (6) seniority level, (7) faculty of study, (8) major field of study, and (9) the time taken to complete the test. Eight were categorical and one (the time taken to complete the test) was continuous. Table 4.1 shows the reference categories used for the categorical determinants.

Variable	Characteristic variable	Reference group
Gender	Binary	Male
Religion	Binary	Islamic
Degree duration	Binary	2 or 4 years
High school type	Binary	Religious school
Home province	Binary	5 Southern border provinces
Seniority level	Categorical with 4 groups	1 st year
Faculty of study	Categorical with 6 groups	Education
Major field of study	Categorical with 7 groups	Science

Table 4.1: Categorical determinants and reference groups for regression analysis

Multiple linear regression analysis was used to test the association between the English Vocabulary Skill Test score and all determinants. A reduced model, produced by omitting determinants with p-values exceeding 0.05 using back ward elimination, gave an r-squared value of 0.222 and contained three predictors – students seniority level, faculty and major field of study.

Since faculty and major field are strongly associated, we recoded major field to include the faculty, given nine categories instead of seven as shown in the Table 4.2.

Variable	Characteristic variable	Reference group
Gender	Binary	Male
Religion	Binary	Islamic
Degree duration	Binary	2 or 4 years
High school type	Binary	Religious school
Home province	Binary	5 Southern border provinces
Seniority level	Categorical with 4 groups	1 st year
Recoded major field of study included faculty	Categorical with 9 groups	Science major in Faculty of Education

Table 4.2: Categorical determinants and reference groups for regression analysis when recoded major field of study

Table 4.3 shows the reduced multiple regression model for English Vocabulary Skill. The result gave an r-squared value of 0.218 and showed that the students' seniority level and recoded major field of study were both related to their English Vocabulary Skill Test score. In addition, students majoring in languages in the Faculty of Humanities and Social Sciences did substantially better than other students in the test, and the first-year and third-year students also performed better than the others.

Determinant	Coefficient	St Error	z-value	95% CI	p-value
Constant	6.6782	0.3233	20.66	6.0446, 7.3119	0.0000
studentYr	baseline:	1			0.0000
2	-1.1452	0.2413	-4.75	-1.6182, -0.6722	0.0000
3	0.1394	0.2698	0.52	-0.3895, 0.6683	0.6057
4	-0.6049	0.2598	-2.33	-1.1142, -0.0957	0.0202
majorGrp2	baseline:	Sci..Educ			0.0000
Sci..SciTech	-0.5937	0.2786	-2.13	-1.1398, -0.0476	0.0334
CommSci	-2.6275	0.4180	-6.29	-3.4468, -1.8082	0.0000
SocSci..Hum	-1.7013	0.5038	-3.38	-2.6887, -0.7138	0.0008
Educ..Educ	-0.8192	0.5770	-1.42	-1.9502, 0.3117	0.1561
IslamicStud	-1.9278	0.3289	-5.86	-2.5725, -1.2832	0.0000
Arts	-0.8857	0.4138	-2.14	-1.6967, -0.0746	0.0326
Lang..Educ	-1.2994	0.3520	-3.69	-1.9894, -0.6093	0.0002
Lang..Hum	0.9117	0.3462	2.63	0.2331, 1.5902	0.0086
r-sq: 0.2175 df: 773 RSS: 3810.2453 s: 2.2202					

Table 4.3: Reduced multiple regression model for English Vocabulary Skill

Prince of Songkla University
Pattani Campus