

Statistical Analysis on Students' Performance on ST2063 Statistical Programming Package for Semester 2 2006/2007 Session: A Comparative Analysis on Gender, Programme, Status and Year of Study

SURIANI HASSAN¹, FAUZIAH SULAIMAN², KHADIZAH GHAZALI³,
SITI RAHAYU MOHD HASHIM⁴ AND ZAINODIN HAJI JUBOK⁵
^{1, 3, 4, 5} School of Science and Technology,
Universiti Malaysia Sabah,
Locked Bag No. 2073, 88999 Kota Kinabalu, Sabah,
MALAYSIA

Abstract: - ST2063 Statistical Programming Package is a core subject for most students in School of Science and Technology, Universiti Malaysia Sabah. This study discussed the students' performance on ST2063 Statistical Programming Package for Semester 2 2006/2007 Session on gender, programme, students' status and year of study. The data consisted of ST2063 Statistical Programming Package results of 269 students in Semester 2 2006/2007 Session. This study utilized Independent-Samples t-Tests and Analysis of Variance for the analysis methods. From this study, it was statistically significant that female students performed better than male students on ST2063 Statistical Programming Package total marks for Semester 2 2006/07 Session. There were significant mean differences for ST2063 Statistical Programming Package total marks between first and third year students, between first and second year students and between second and third year students. Biotechnology students performed better than Plant Technology students, Conservation Biology students, Physics with Electronics students, Marine Science students and Environmental Science students on ST2063 Statistical Programming Package total marks for Semester 2 2006/07 Session. There was no significant mean difference between the status of the students on students' performance on ST2063 Statistical Programming Package.

Key-Words: Marks, Gender, Status, Programme, Year of Study

1 Introduction

Universiti Malaysia Sabah is located in East Malaysia. School of Science and Technology is one of the faculties in Universiti Malaysia Sabah and ST2063 Statistical Programming Package is a core subject for most students.

Some students register ST2063 Statistical Programming Package during their first year of study; others register in second, third or fourth year of study, depending on their degree programs requirement. Sometimes there are also students repeat the subject in order to improve their grades, in addition their PNGKs. Student is only allowed to repeat a subject if the student's result was C- and below [1].

Schuyten, Dekeyser and Goeminne from Department of Data-Analysis, University of Gent, Belgium suggested towards an electronic

independent learning environment for statistics in higher education. The results indicated that the cognitive outcomes of students' learning in a computer-based environment were neither fostered nor hindered by students' attitudes towards computers. Students with better prior knowledge and having fewer problems with statistics had a better perception of the learning access to the contents [2].

A past study discussed the students' performance on ST2063 Statistical Programming Package results for 2004/2005 Session. The study compared the results of two semesters for 2004/2005 Session. From the study, significant different of students' performance on ST2063 Statistical Programming Package was identified based on student's school and programme [3].

This paper discussed the students' performance on ST2063 Statistical Programming Package for

Semester 2 2006/2007 Session. The paper focused on the mean differences for gender of the students, status of the students for the subject, the year of study and the programme of the students when registered the subject.

2 Methodology

The data comprised from ST2063 Statistical Programming Package results of 269 students in Semester 2 2006/2007 Session. The total marks that were used to analyze the students' performance consisting of the coursework marks and the final examination marks for each student. This study utilized Independent-Samples t-Tests and Analysis of Variance (ANOVA) for the analysis methods.

During the lectures and laboratory sessions, students are exposed using the Statistical Package for Social Sciences (SPSS) software, starting from the introduction of using SPSS, the SPSS menus especially the Analyze Menu and also the Graph Menu.

In the Analyze Menu, students are exposed with the explore dialogue box to check for the normality of the data distribution, the descriptive dialogue box for descriptive statistics like mean and variance, the correlation dialogue box and the compare means dialogue box for t-Test (One-Sample t-Test, Paired-Samples t-Test and Independent-Samples t-Test) and Analysis of Variance (ANOVA).

3 Results

The numbers of students registered for ST2063 Statistical Programming Package for Semester 2 2006/07 Session were 269 students. The crosstabulation table of Table 1 shows the number of students registered this subject for the semester respective to students' gender and programmes. There were 99 (36.8%) male students and 170 (63.2%) female students registered this subject for the semester.

Table 1 shows the highest number of students registered the subject for this semester were Conservation Biology students', which were 99 (36.8%) students. The second highest number of students registered the subject for this semester was Biotechnology students', which were 63 (23.4%) students. The second smallest number of students registered the subject for this semester were Industrial Chemistry students', which were 6 (2.23%) students and the smallest number of students registered the subject for this semester were

Mathematics with Economy students', which were 4 (1.49%) students'.

Table 1. Crosstabulation table between Gender and Programme of the students

Programme	Male		Female		Total	
	N	%	N	%	N	%
Conservation Biology	20	7.43	79	29.4	99	36.8
Biotechnology	23	8.55	40	14.9	63	23.4
Physics with Electronics	28	10.4	13	4.83	41	15.2
Plant Technology	6	2.23	12	4.46	18	6.69
Marine Science	8	2.97	6	2.23	14	5.2
Aquaculture	3	1.12	9	3.35	12	4.46
Environmental Science	7	2.6	5	1.86	12	4.46
Industrial Chemistry	3	1.12	3	1.12	6	2.23
Mathematics with Economy	1	0.37	3	1.12	4	1.49
Total	99	36.8	170	63.2	269	100

The crosstabulation table of Table 2 shows the number of students registered this subject for the semester respective to status of the students and year of study. The status of the students was divided into two categories; first time or repeat subject. First time means it is the first time a student takes the subject, whereas repeat subject means the student repeats the subject to improve the result. There were 216 (80.3%) students registered this subject for first time and 53 (19.7%) students repeated the subject to improve their results.

All 137 (63.43%) first year students registered this subject for their first time. All 5 (1.86%) fourth year students repeated the subject to improve their results.

Table 2. Crosstabulation table between Year of Study and Status of the students

Year Of Study	Status				Total	
	First Time		Repeat Subject			
	N	%	N	%	N	%
1	137	50.9	0	0	137	50.9
2	75	27.9	2	0.7	77	28.6
3	4	1.5	46	17.1	50	18.6
4	0	0	5	1.9	5	1.9

Total	216	80.3	53	19.7	269	100
-------	-----	------	----	------	-----	-----

Normality assumption must be fulfilled for testing Independent Samples t-Test and Analysis of Variance (ANOVA) for ST2063 Statistical Programming Package total marks [4]. Figure 1 shows that the Normality assumption for ST2063 Statistical Programming Package total marks for Semester 2 2006/2007 Session was fulfilled using the Normal Q-Q Plot since the points fall more or less in a straight line [5].

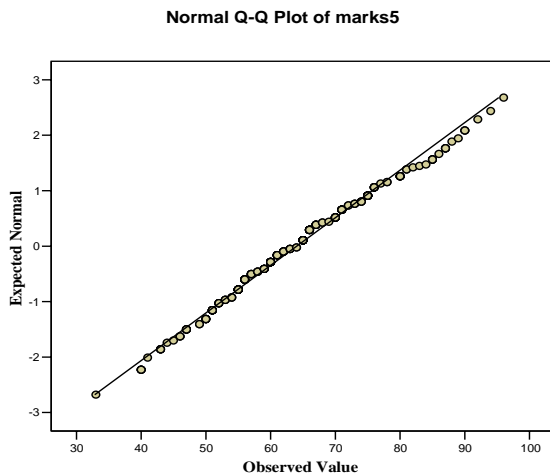


Fig. 1. Normal Q-Q Plot of total marks

Table 3 shows the mean and standard deviation for the total marks respective to the genders and status of the students. Female students had higher mean total marks compared to male students. Students who repeated the subject had higher mean total marks compared to first time students. Further analysis using Independent t-test were done to test whether there were significant means differences total marks between female and male students and between first time and repeat students.

Table 3. Mean and standard deviation for the total marks respective to Genders and Status of the students

Variable		Mean	Standard Deviation
Gender	Male	62.1616	11.76878
	Female	65.1353	11.47879
Status	First Time	63.8519	12.37613
	Repeat Subject	64.8113	8.11470

Table 4 shows the mean and standard deviation for the total marks respective to year of study the students registered for the subject. Second year students had the highest mean total marks and first year students had the lowest mean total marks. Further analysis using Analysis of Variance (ANOVA) was done to test whether there were significant means differences total marks among year of study.

Table 4. Mean and standard deviation for the total marks respective to Year of Study

Year Of Study	Mean	Standard Deviation
1	58.9781	9.69268
2	72.8961	11.68173
3	64.5600	7.77794
4	61.2000	13.02690

Table 5 shows the Independent Samples t-Test for Gender and Status of the students. Taking alpha at 5% significance level, t-Test for Equality of Means for Gender showed that the average of total marks was statistically difference between the male and female students ($t = -2.03$, $\text{Sig-p} = 0.043$). The mean difference was 2.97. Therefore, it was statistically significant that female students performed better than male students.

T-Test for Equality of Means for status of the students showed that the average of total marks was statistically no difference between the first time students registered for this subject and the students repeated the subject ($t = -0.687$, $\text{Sig-p} = 0.494$).

Table 5. Independent Samples Test

Variable	Means		
	t	Sig.	Mean difference
Gender	-2.03	0.043	-2.97
Status	-.687	0.494	0.494

Table 6 shows the analysis of variance for the year of study and programme of the students. The analysis of variance showed that at least one of the average of total marks were statistically difference among the year of study ($F = 31.663$, $\text{Sig-p} = 0.000$). The analysis of variance showed that at least one of

the average of the total marks were statistically difference among the programmes ($F = 15.548$, $\text{Sig-p} = 0.000$).

Table 6. Analysis of Variance

Variable	F	Sig.
Year of Study	31.663	0.000
Programme	15.548	0.000

Table 7 shows the Tukey Multiple Post-Hoc Comparisons for total marks for the year of study. The results showed that there were significant mean differences for total marks between first and third year students ($\text{Sig-p} = 0.000$), between first and second year students ($\text{Sig-p} = 0.005$) and between second and third year students ($\text{Sig-p} = 0.000$).

Table 7. Tukey Multiple Post-Hoc Comparisons for Year of Study

(I) Year Of Study	(J) Year Of Study	Mean Difference (I-J)	Standard Error	Sig.
1	2	-13.918	1.432	.000
1	3	-5.582	1.661	.005
2	3	8.336	1.826	.000

Table 8 shows the Tukey Multiple Post-Hoc Comparisons for total marks for the students' programmes. The results showed that there were significant mean differences for total marks between Biotechnology (HG07) students with students from several programmes.

There were significant mean differences for total marks between Biotechnology students and Plant Technology (HS23) students ($\text{Sig-p} = 0.000$), between Biotechnology students and Conservation Biology (HS03) students ($\text{Sig-p} = 0.000$), between Biotechnology students and Physics with Electronics (HS22) students ($\text{Sig-p} = 0.000$), between Biotechnology students and Marine Science (HS40) students ($\text{Sig-p} = 0.000$), and between Biotechnology students and Environmental Science (HS11) students.

Table 8. Tukey Multiple Post-Hoc Comparisons for Students' Programmes

(I) Programme	(J) Programme	Mean Difference (I-J)	Standard Error	Sig.
HG07	HS23	11.70	2.60	0.00

HG07	HS03	15.42	1.57	0.00
HG07	HS22	17.96	1.95	0.00
HG07	HS40	12.54	2.88	0.00
HG07	HS11	13.42	3.06	0.00

Two-way ANOVA involves the operation of analyzing the variances for one dependent variable based on two independent variables [6]. Table 9 showed the interaction effects between two independent variables. The Univariate Analysis of Variance results showed that there were no significant interaction effect between gender and status of the students ($F=0.002$, $\text{Sig-p}=0.967$), no significant interaction effect between gender and programme of the students ($F=0.594$, $\text{Sig-p}=0.782$), no significant interaction effect between gender of the students and year of the study ($F=0.256$, $\text{Sig-p}=0.857$), no significant interaction effect between status of the students and year of the study ($F=2.652$, $\text{Sig-p} =0.105$), no significant interaction effect between status of the students and students' programmes ($F=1.585$, $\text{Sig-p} =0.165$) and no significant interaction effect between year of the study and students' programmes ($F=1.281$, $\text{Sig-p} =0.260$).

From Table 9, the Univariate Analysis of Variance results showed that there were no significant interaction effect between the first independent variable and the second independent variable. Thus the influence of the first independent variable on total marks did not depend on the second independent variable.

Table 8. The Interaction Effects between Two Independent Variables

INTERACTION EFFECTS	F	Sig.
Gender * Status	0.002	0.967
Gender * Programme	0.594	0.782
Gender * Year of study	0.256	0.857
Status * Year of study	2.652	0.105
Status * Programme	1.585	0.165
Year of study * Programme	1.281	0.260

4 Discussions and Conclusion

This study showed the average of ST2063 Statistical Programming Package total marks for Semester 2 2006/07 Session was statistically difference between the male and female students. Female students performed better than male students on ST2063 Statistical Programming Package total marks for Semester 2 2006/07 Session. The mean difference was 2.97. This confirms to existing literature that the learning approach by gender was different. In Australia, young boys were lagging behind their female peer in terms of performance in mathematics [7].

The results showed that there was no significant mean difference for ST2063 Statistical Programming Package total marks for status of the students; between the first time students registered for this subject and the students repeated the subject. A study showed there was significant mean difference on the students' performance on ST1033 General Mathematics and Statistics total marks between the first time students registered for this subject and the students repeated the subject [8].

The Tukey Multiple Post-Hoc Comparisons results showed that there were significant mean differences for ST2063 Statistical Programming Package total marks between first and third year students, between first and second year students and between second and third year students. Second year students performed better for than first year students with mean difference 13.92. Third year students performed better than first year students with mean difference 5.582. Second year students performed better than third year students with mean difference 8.336.

The results showed that there were significant mean differences for ST2063 Statistical Programming Package total marks between Biotechnology students with students from several programmes. Biotechnology students performed better than Plant Technology students, Conservation Biology students, Physics with Electronics students, Marine Science students and Environmental Science students on ST2063 Statistical Programming Package total marks for Semester 2 2006/07 Session. The outstanding performance by this Biotechnology students will provide excellent human capitals as Malaysia is aspires to be a biotechnology hub as clearly spelled out in the Malaysia National Biotechnology Policy that was launched on the 28th April 2005 [9].

The interaction results showed that the gender's influence of the students on total marks did not depend on the status of the students, the influence of

the gender total marks did not depend on the students' programmes, the influence of the gender total marks did not depend on the year of study, the influence of the status total marks did not depend on the year of study, the influence of the students' total marks did not depend on the students' programmes and the influence of the students' year of study total marks did not depend on the students' programmes.

5 Future Work

Further study should be done on difference of practicalities and technicalities in using computer applications and theoretical parts in Mathematics / Statistics on gender and status of the students (first time students and repeat students).

References:

- [1] Universiti Malaysia Sabah, *Prospektus Prasiswazah Universiti Malaysia Sabah*, Kota Kinabalu, Universiti Malaysia Sabah, 2007
- [2] Schuyten, G., Dekeyser, H. and Goeminne, K., Towards an electronic independent learning environment for statistics in higher education, *Education and Information Technologies Springer Science + Business Media B.V.*: Vol.4, No.4, 1999, pp 409-424
- [3] Suriani, H., Siti Rahayu, M.H., and Khadizah, G. Students' Performance on Statistical Programming Package ST2063, *Borneo Science: A Journal of Science and Technology*, Universiti Malaysia Sabah, Kota Kinabalu, Malaysia, Vol 19, 2006, pp409 - 413, ISSN 1394-4339
- [4] Carver, R.H., Nash, J.G., *Doing Data Analysis With SPSS*. Thomson Brooks/Cole, Duxbury, 2006
- [5] Coakes, S. J. and Steed, L. G., *SPSS: Analysis Without Anguish: Version 14.0 for Windows*. John Wiley and Sons Australia, Ltd., Sydney, 2007
- [6] Mohd Salleh Abu and Zaidatun Tasir, *Pengenalan Kepada Analisis Data Berkomputer SPSS 10.0 for Windows*, Kuala Lumpur, Venton Publishing, 2001
- [7] Richardson, J. T. E and King, E., Gender Differences in the Experience of Higher Education: Quantitative and Qualitative Approaches, *Educational Psychology*, Vol 11, 1991, pp363-382
- [8] Suriani H, Rajasegeran R and Noraini A, Students' Performance on ST1033 General Mathematics and Statistics, *Proceedings of the*

*Third International Conference on Research
And Education In Mathematics 2007 ICREM3,*
Universiti Putra Malaysia, Kuala Lumpur,
Malaysia, 2007, pp409 – 413, ISBN 978-983-
3455-89-8

- [9] Malaysian Biotechnology Information Centre,
Biotechnology in Malaysia,
<http://www.biotek.gov.my/>