



JOURNAL OF AFRICAN RESEARCH AND DEVELOPMENT (JAFRED)



J.Afr.Res.Dev: www.jafred.co.ke

ISSN 2517-9365: ISBN: 978-9966-098-38-2

7(2017):1-029, 2017, Article No. JAFRED-01-029-17

Accepted on 20th July, 2017; Reviewed on 25th August, 2017 and Published on 27th October, 2017

AN ANALYSIS OF THE PERFORMANCE OF STUDENTS IN SEVEN STATISTICS UNITS AT UNIVERSITY OF ELDORET

By JOURNAL OF

Ayubu Anapapa Okango

Department of Mathematics and Computer Science, University of Eldoret, P.O. Box 1125-30100
Eldoret, Kenya.

Author Contributions

This work has been fully done and written by Ayubu Anapapa Okango.

Editors

1. James Dan, Department of Mathematics and Statistics, University of Birmingham.
2. Ken Lamadhi, Lecturer in Statistics, Dar Es Salaam University.

Reviewers

1. Prof Mary James, Lecturer in Statistics; University of London.
2. Dr. Henry Yap, Lecturer in Statistics, Ohio University.

ORIGINAL RESEARCH ARTICLE

1. **Corresponding Author: Ayubu Anapapa Okango, Email: anapapaa@yahoo.com**
2. **Citation: Ayubu A, O (2017): An Analysis of The Performance of Students in Seven Statistics Units at University of Eldoret, Journal of African Research and Development, (J. Afr. Res. Dev. (Kericho)) Vol. 2: Issue 2 (2017) pp: 19 - 29.**

Copy right: © Ayubu, A.O (2017): This is an open access article distributed for unrestricted use, distribution, and reproduction in any medium, provided the original author and source is credited.

ABSTRACT

This study analyzed the performance in seven statistics units of students who were admitted for a Bachelor of Science (Applied Statistics with Computing) course at University of Eldoret, Kenya in 2012. The units were taught by the same lecturer over a period of three years. The data comprised Continuous Assessment Test (CAT) and exam marks obtained by all the students in this class. The overall mean mark for the seven units was 56.7% ($n = 614$, $SD = 8.27$). There was no significant difference ($p = 0.1601$) in the mean score of male students value of 56.6% ($n = 386$, $SD = 8.72$) and that of female students value of 57.3% ($n = 209$, $SD = 7.50$). Results showed that government students performed better than the self-sponsored ones ($p = 0.0002$). There was a significant difference in mean performance by year ($p = 0.0006$). The mean score was highest in second year (57.8%) followed by fourth year (56.5%) and lastly third year (53.9%). Simple linear regression models were developed for the seven units with the final exam mark being the dependent variable and the CAT mark being the independent variable. The study established in general there was a positive correlation between CAT marks and final exam marks ($p < 0.001$, $R^2 = 0.0441$). However only a small proportion of the variation in the exam mark can be explained by the CAT mark. There was significant difference in mean performance from one unit to another ($p = 0.0001$).

Key Words: *Performance, analysis, statistics, sponsorship, gender*

1. Introduction

Analyzing the performance of students in a course is a vital activity since it helps to assess the understanding of the course's contents among different demographics, check the trend in the results or predict the outcome in other courses among other uses. In statistics, such analyses can be used to compare the performance between theoretical and practical courses, evaluate new teaching methods or assess the level at which students have grasped specific statistical methodologies.

A number of studies have been done on the analysis of performance of courses at university level. (Smith *et al*, 1994) investigated students' difficulties in first-year statistics examinations at university. They hypothesized that difficulty with language was an important factor in student performance in statistics examination. They

found that there was no correlation between student performance and the linguistic complexity of the questions as measured by formal measures of lexical density. (Kakish *et al*, 2012) researched on the performance of students enrolled in traditional versus hybrid elementary statistics courses at Georgia Gwinnett College (GGC). They discovered that there is no significant difference between the performances of the two groups. (Dayton, 1990) evaluated student performance in applied statistics courses which emphasize data analysis. He observed that applied statistics courses for students in professional fields other than the field of statistics itself pose a unique set of instructional problems. He recommended that the design of such courses must recognize the roles which these professionals will assume and should attempt to provide training that emphasizes the professional development of the students. (Anja *et al*, 2008) demonstrated the analytical use of a statistical tool ANOM-Analysis of

Means (more commonly used by quality control engineers), mixed model analysis and other statistical tools (used in social sciences) to investigate differences found in Croatian high school student performance in mathematics examinations. Analysis of variance results showed that mathematics scores were on average significantly different across regions, programs and school sizes. Analysis of means (ANOM) was used to identify and visualize the differences among those individual programs/counties in which students achieved results that were significantly better/worse than the average. (Merkling *et al*, 2013) examined the effect of taking pre-calculus in high school on final letter grades in Calculus I, II, and III at Valparaiso University. The results showed that students who took pre-calculus in high school on average received higher final letter grades in Calculus I, II, and III. (Kakish *et al.*, 2012) studied the predictors of success in a business statistics course. They found that GPA was the strongest predictor of performance in the business course. (Rochelle and Dotterweich, 2007) found that three independent variables were significantly related to the final grade earned by students in a business statistics course at a medium-size, regional state university. They were the number of student absences, grade

earned in a previous introductory quantitative methods course (algebra and differential calculus), and GPA.

This study analyzed the performance in seven statistics units of students who were admitted for a Bachelor of Science (Applied Statistics with Computing) course at University of Eldoret, Kenya in 2012. The units were taught by the same lecturer over a period of three years

2. Method

Data was collected from all students who enrolled for Bachelor of Science (B.Sc.) in Applied Statistics with Computing at University of Eldoret in 2012. The selected seven statistics units were taught by the same lecturer from second year to fourth year. This eliminated the inter lecturer bias. There were three second year courses, one third year course and three fourth year courses. For each student their gender and sponsorship status were recorded. Students at University of Eldoret can either be government sponsored or self-sponsored. For each unit the continuous assessment test (CAT) marks and final exam marks for the students were recorded. The units and their names are shown in table 1

Table 1: Unit Description

Unit	Unit Name	Year
I	Introduction to Applied Statistics	2
II	Principles of Statistical Inference	2
III	Categorical Data Analysis	2
IV	Computing Methods and Data Analysis	3
V	Applied Time Series	4
VI	Scientific Computing	4
VII	Biometry Methods	4

Microsoft Excel 2010 was used enter and manage the data. Data analysis was performed in Stata 12 and R 3.1.1 (R Core Team, 2014).

3.1 Descriptive Statistics

Descriptive analysis was performed on the final exam mark for the seven units. The results are shown in table 2.

Table 2: Descriptive Statistics

Unit	N	Mean	Standard Deviation	Minimum	Maximum
I	95	59.2	8.403	44	84
II	92	57.1	9.900	40	86
III	92	57.0	7.944	40	73
IV	86	53.9	6.254	43	72
V	82	56.9	10.069	40	91
VI	83	54.6	6.852	40	84
VII	84	58.1	6.408	40	71
Overall	614	56.7	8.269	40	91

Table 2 shows that mean final exam mark ranged between 53.9 for unit IV to 59.2 for unit I. The overall mean for the seven units was found to be 56.7 implying that the mean grade is C. This shows that the performance of students is average. The performance is reflected in the final graduation statistics for this cohort that graduated in November, 2016. Out of 65 graduates in the class, 3(4.6%) had first class, 30(46.2%) had second upper class while 32(49.2%) had second lower class.

The distribution of the final exam mark is depicted in figure 1 and figure 2.

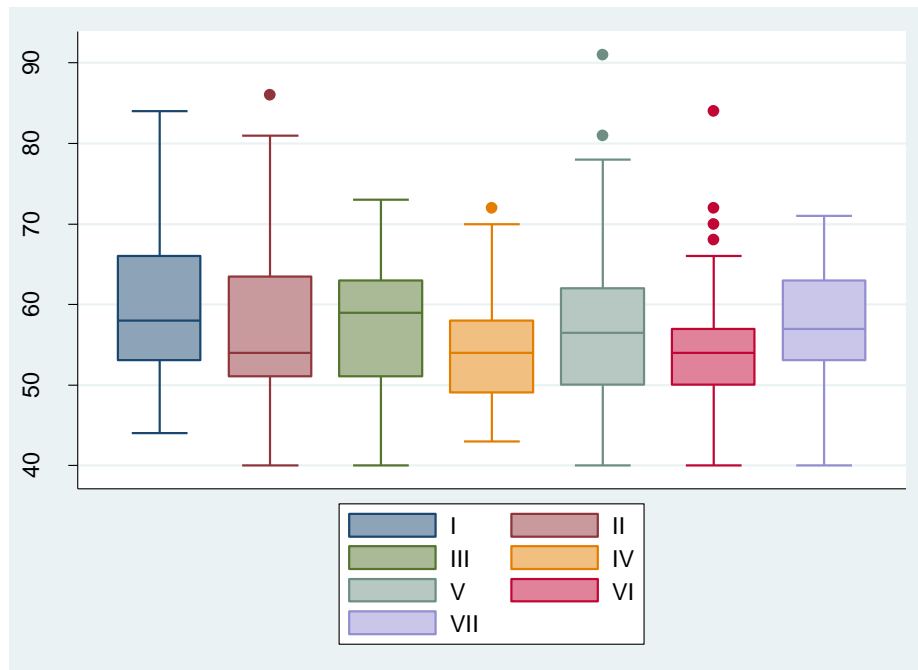


Figure 1: Boxplots for Final Exam Marks

Figure 1 shows that the median marks for all the units are between 55 and 60. Units II, IV, V and VI had some outliers on the right end indicating a few students performed exceptionally well in those units away from the general pattern of the other students.

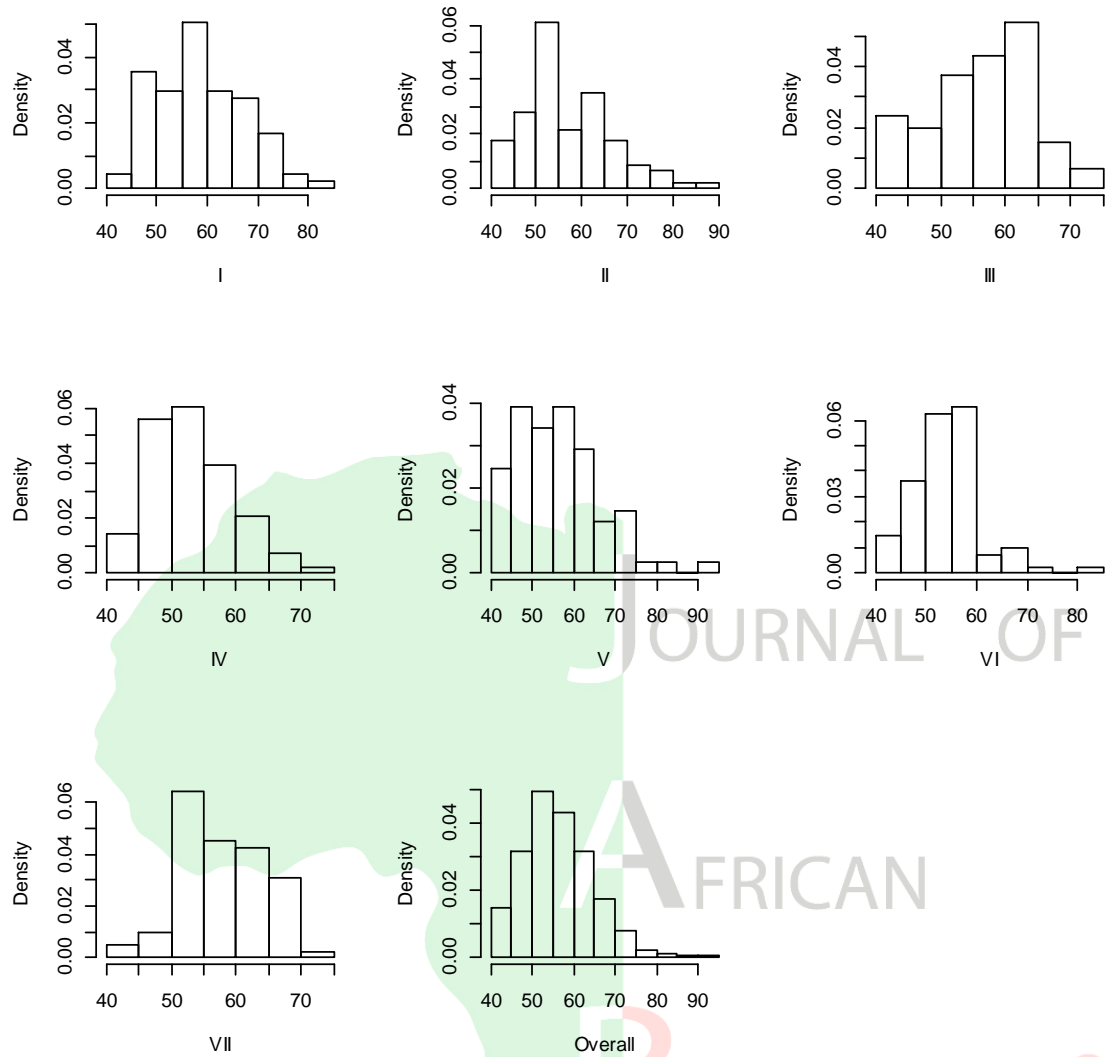


Figure 2: Histograms for Final Exam Marks

Figure 2 shows that the distribution for units I and IV is almost normal whereas the others are slightly skewed.

3.2 Performance by Gender

Independent samples *t*-test was carried out to determine if there was any significant difference in the performance of the seven statistics units between males and females. The results are shown in table 3. For each unit, the table shows the number of student by gender (N), the mean score (Mean), the was 57.3 for females and 56.6 for males.

standard deviation, the *t* statistic and the *p* value of the *t* statistic. A difference in mean performance was deemed significant if the associated *p* value was less than 0.05. The table shows that females performed better than males in four units. However the difference was significant for only unit II (*p* = 0.0174). On the overall the mean mark

Table 3: Performance by Gender

Unit	Gender	N	Mean	Standard Deviation	<i>t</i> statistic	<i>p</i> value
I	Female	32	59.3	6.945	0.1095	0.4565
	Male	63	59.1	9.107		
II	Female	30	60.7	9.678	2.1453	0.0174
	Male	57	56.0	9.676		
III	Female	30	56.8	7.290	0.0492	0.4804
	Male	56	56.9	8.529		
IV	Female	30	53.2	5.453	0.9359	0.1760
	Male	54	54.5	6.655		
V	Female	29	57.1	8.387	0.1647	0.4348
	Male	51	56.7	11.164		
VI	Female	29	55.0	5.754	0.3247	0.3731
	Male	52	54.4	7.529		
VII	Female	29	58.7	5.929	0.5850	0.2801
	Male	53	57.8	6.788		
Overall	Female	209	57.3	7.504	0.9951	0.1601
	Male	386	56.6	8.715		

3.3 Performance by Sponsorship

Students at University of Eldoret can either be government sponsored or self-sponsored. It was interesting to compare performance in these two groups since self-sponsored students having made a deliberate decision to apply specifically for the course would be more motivated in the course as opposed to some government sponsored students for whom the course might not be their first choice. On the other hand the average KCSE entry points for self-sponsored students are usually lower than that for government students. Oluoch and Ronoh, (2017) compared the academic performance between government and self-sponsored students at Rongo University and found that there was significant difference in performance on second class (both upper and lower divisions), pass and supplementary.

Independent samples *t*-test was carried out to determine if there was any significant difference in the performance of the seven

statistics units between government and self-sponsored students. The results are shown in table 4. The table shows that government sponsored students performed better their self-sponsored counterparts in all units except one. The difference in the mean performance was significant for unit I ($p = 0.0098$), unit II ($p = 0.0229$), unit IV ($p = 0.0167$) and unit V ($p = 0.0241$). The overall mean for government sponsored students was found to be 57.7 whereas that for the self-sponsored students was 55.1. The difference was significant ($p = 0.0002$).

3.4 Performance by Year

There were three units in second year, one unit in third year and three units in fourth year. A one way analysis of variance was done to find if there was significant difference in mean performance across the years. The results are shown in table 5 and table 6.

Table 4: Performance by Sponsorship

Unit	Sponsorship	N	Mean	Standard Deviation	t statistic	p value
I	Government	63	60.6	8.249	2.3747	0.0098
	Self	32	56.4	8.112		
II	Government	59	59.1	10.498	2.0268	0.0229
	Self	28	54.6	7.724		
III	Government	57	56.7	7.891	0.1421	0.4437
	Self	29	57.0	8.565		
IV	Government	57	55.0	6.432	2.1631	0.0167
	Self	27	51.9	5.377		
V	Government	55	58.4	10.436	2.0064	0.0241
	Self	25	53.6	8.945		
VI	Government	55	54.9	7.669	0.5612	0.2881
	Self	26	54.0	5.028		
VII	Government	57	58.4	6.518	0.4812	0.3158
	Self	25	57.6	6.474		
Overall	Government	403	57.7	8.558	3.6006	0.0002
	Self	192	55.1	7.485		

Table 5: Performance by Year

Year	N	Mean	Standard Deviation	Minimum	Maximum
Two	279	57.8	8.813	40	86
Three	86	53.9	6.254	43	72
Four	249	56.5	8.031	40	91

Table 5 shows that the mean mark was 57.8 in second year 53.9 in third year and 56.5 in third year.

Table 6: Analysis of Variance for Performance by Year

Source	Sum of squares	Degrees of freedom	Mean square	F	P value
Between groups	1002.694	2	501.347	7.49	0.0006
Within groups	40913.034	611	66.961		
Total	41915.728	613	68.378		

Table 6 shows that there was a significant difference in mean performance by year. The

F statistic value was found to be 7.49 with a p value of 0.0006.

3.5 Relationship between CAT Marks and Final Exam Mark

The study investigated the relationship between CAT marks and final exam marks. The final CAT marks being an aggregate of sit – in CATs, take way CATs and assignments were used in the analysis. Simple linear regression models were developed for the seven units with the final exam mark being the dependent variable and the CAT marks being the independent variable. The results are shown in table 7. The fitted models were of the form $y = \beta_0 + \beta_1x$ where y was the final exam

mark and x the CAT marks. All the models showed that there was a positive relationship between CAT marks and final exam marks. A model was deemed significant if the associated p value was less than 0.05. The only non-significant models were for unit I ($p = 0.8683$) and unit II ($p = 0.1554$). The R^2 statistic is used to give the proportion of variation in the dependent variable that can be explained by the independent variable. This ranged between 0.0003 for unit I to 0.2739 for unit V. This shows that final exam mark cannot be well predicted from the CAT mark.

Table 7: Regression Analysis between CAT Marks and Final Exam Mark

Unit	Coefficient	<i>t</i>	<i>P</i> value	<i>R</i> ²	Adjusted <i>R</i> ²	<i>F</i>	<i>P</i> value	
I	β_1	0.052	0.170	0.868	0.0003	-0.0105	0.03	0.8683
	β_0	36.748	5.490	0.000				
II	β_1	0.485	1.430	0.155	0.0223	0.0114	2.05	0.1554
	β_0	27.960	4.160	0.000				
III	β_1	0.339	2.190	0.031	0.0506	0.0401	4.80	0.0310
	β_0	31.885	10.770	0.000				
IV	β_1	0.486	3.600	0.001	0.1334	0.1231	12.93	0.0005
	β_0	24.117	8.790	0.000				
V	β_1	0.965	5.490	0.000	0.2739	0.2649	30.18	0.0000
	β_0	15.027	3.950	0.000				
VI	β_1	0.489	2.220	0.029	0.0575	0.0459	4.95	0.0289
	β_0	24.825	5.590	0.000				
VII	β_1	0.856	4.100	0.000	0.1703	0.1602	16.83	0.0001
	β_0	22.530	5.590	0.000				
Overall	β_1	0.441	5.310	0.000	0.0441	0.0425	28.23	0.0000
	β_0	27.857	16.520	0.000				

3.6 Performance between Units

A one way analysis of variance was done to establish whether there was significant difference in mean performance for the seven units. The results are shown in table 8.

Table 8: Analysis of Variance for Performance by between Units

Source	Sum of squares	Degrees of freedom	Mean sum of square	<i>F</i>	<i>P</i> value
Between groups	1821.284	6	303.547	4.6	0.0001
Within groups	40094.444	607	66.053		
Total	41915.728	613	68.378		

Table 8 shows that there was a significant difference in mean performance among the seven units. The *F* statistic value was found to be 4.6 with a *p* value of 0.0001.

4. Summary and Conclusions

The study found that the performance of students in the statistics units was fair with an overall mean score of 56.7. There was no significant difference in the mean score between male and female students. The research hypothesized that self-sponsored students perform better than government

sponsored students. This is because self-sponsored students purposely selected and applied to be admitted to the course and were thought to have a keen interest in the Applied Statistics course. For some government sponsored students however, the course was not their first choice. On the contrary results showed that government students performed better than the self-sponsored ones.

It was expected that as years progressed students would be more familiar with statistical concepts and therefore perform better. However the mean performance was highest in second year at 57.8 followed by fourth year at 56.5 and lastly in third year at 53.9. There was significant difference in mean performance from one unit to another. The study established in general there was a positive correlation between CAT marks and final exam marks.

References

- Dayton, C. M. (1990). Evaluating student performance in applied statistics courses which emphasize data analysis. *Proceedings of International Conference on Teaching Statistics*, 3, 475 - 478.
- Habus-Korbar, A. Luzar-Stiffler, H and Bazdan, V. (2008). Investigation of Preliminary Student Test Performance Data Using the ANOM Statistical Tool and Mixed Model Analysis. *Journal of Computing and Information Technology*, 4, 287 - 294.
- Kakish, K. M., Pollacia, L. and Heinz, A. (2012). Analysis of the Effectiveness of Traditional Versus Hybrid Student Performance for an Elementary Statistics Course. *International Journal for the Scholarship of Teaching and Learning*, 6(2), 46 – 54.
- Merkling, K., Dorman, H. and Kinzel, N. (2013). Statistical Analysis of Student Performance in Mathematics Courses. *Celebration of Undergraduate Scholarship*, 285.
- Monroe, S. R., Moreno, A. and Segall, M. (2016). Student Performance Determinants in a Business Statistics Course at a Large Urban Institution. *Academic and Business Research Institute* [online], Available at <http://www.aabri.com/LV11Manuscripts/LV11082.pdf>.
- Oluoch, J. and Ronoh, L. K. (2017). Comparative Analysis of Academic Performance between PSSP and GSSP Students in Kenyan Universities: Case of Rongo University. *Journal of Education and Practice*, 8(18), 155-161
- R Core Team (2014). R: A language and environment for statistical computing. *R Foundation for Statistical Computing, Vienna, Austria*. URL <http://www.R-project.org/>
- Rochelle, C. F. and Dotterweich, D. (2007). Student Success in Business Statistics. *Journal of Economics and Finance Education*, 6 (1), 19 - 24.
- Smith, N. F., Wood, L. N., Gillies, R. K. and Perrett, G. (1994). Analysis of Student Performance in Statistics. *Proceedings of Mathematics Education Research Group of Australasia*, 17, 539-545.