

**An investigation into undergraduate students' difficulties in learning the
bivariate normal distribution: A case study of a Kenyan university**

by

NYAMBANE BOSIRE ONYANCHA

submitted in accordance with the requirements

for the degree of

MASTER OF SCIENCE

in the subject

MATHEMATICS EDUCATION

at the

UNIVERSITY OF SOUTH AFRICA

SUPERVISOR: DR U. I. OGBONNAYA

MARCH 2017

Abstract

The low grades that students score in some statistical units in Kenyan universities is of great concern and has evoked research interest in the teaching of some of the units and the students' learning of the statistical content.

The aim of the study was to investigate the difficulties undergraduate students experience in the learning of bivariate normal distribution in a Kenyan university. The research also aimed to answer the following research questions on the difficulties undergraduate students encounter in the learning of bivariate normal distribution.

The first research question was based on the reasons why students find learning of bivariate normal distribution difficult and the second research question was to find the reasons why students experience such difficulties in learning bivariate normal distribution.

The target population for this study included lecturers teaching statistics in the university, and second- and third- year students enrolled or who have previously completed the probability and statistics III unit, where the bivariate normal distribution content is covered. In selecting students for the study, the simple random sampling technique was employed while convenient sampling was used to select lecturers who participated in the study.

A mixed methods design was adopted for this study where both quantitative and qualitative data was collected. A total of 175 students and six lecturers participated in this research study. All students who participated in the study did a bivariate normal distribution test (Appendix 1) designed by the researcher and then filled in a questionnaire (Appendix 2). The lecturers who participated in the study filled in an open-ended questionnaire (Appendix 3).

The results showed that undergraduate students have difficulties in learning bivariate normal distribution. This is because most of them could neither state the bivariate normal distribution nor solve any of the application questions on the content. The students find it difficult to learn and comprehend the bivariate normal distribution equation with its many parameters and constants of the two random independent variables.

The results also showed that students could not state the normal distribution equation nor could they solve questions on the normal distribution, which forms the foundational knowledge required for effective learning of the bivariate normal distribution content.

Based on the results, the study recommended that emphasis should be placed on the basic and foundational knowledge of the normal distribution content and its applications before teaching bivariate normal distribution in probability and statistics III. In addition, it is recommended that all students should be involved in the learning of basic content to enable them to understand all parameters and constants in the equations and their applications. The study also recommends that lecturers revise the foundational knowledge and content related to the bivariate normal distribution before introducing and teaching the bivariate normal distribution content. This study also recommends that the university should consider a change of curriculum by teaching the bivariate normal distribution, as an introductory course to the unit under the multivariate distributions in statistics, in third year of the students' studies.

Key Terms: Bivariate normal distribution; Difficulties in learning; Kenyan university, Normal distribution; Probability; Statistics.

DECLARATION

Name: **Nyambane Bosire Onyancha**

Student number: **48140791**

Degree: **MSc. MATHEMATICS, SCIENCE AND TECHNOLOGY
EDUCATION**

An investigation into undergraduate students' difficulties in learning the bivariate normal distribution: A case study of a Kenyan university

I declare that the above dissertation is my own work and that all the sources that I have used or quoted in the study have been indicated and duly acknowledged by means of complete references.


.....

B N ONYANCHA

03/12/2017
.....

DATE

ACKNOWLEDGEMENTS

In completing this dissertation, I had support from the following individuals and institutions for the research:

- First and foremost, I am sincerely grateful to my supervisor, Dr. Ugorji I. Ogbonnaya for his advice, guidance, inspiration and encouragement during proposal writing. I also thank him for guiding me during my actual research study through all the stages of this work. I am grateful for his kindness and willingness to help in reviewing all chapters of this dissertation. It has been a humbling opportunity for me to work under his guidance and supervision in this research topic.
- I am truly and deeply grateful to the Unisa research directorate for the master's fellowship grant and Unisa student funding for the financial support.
- I thank the vice-chancellor, in charge of academic affairs, in the research university who allowed me to conduct research in the university.
- I am indebted to Mr Geoffrey Barini, a lecturer in one of the Kenyan university, department of mathematics for his utmost support, brotherly love and encouragement during my research and data collection and analysis.
- I am greatly indebted to all lecturers and students of the research university, for without their cooperation and participation in the study, this research would have not been possible or successful.
- I wish also to express my heartfelt and sincere appreciation to my family for their patience and the utmost belief they have in me. Their generous financial support, understanding and encouragement during research for this study are greatly valued since I spent many hours away from home.
- To my wife, Annie, for her physical and mental support. You are a blessing and a noble gift.
- Above all, I am humbled and overwhelmed with God's love, blessings, protection and for giving me strength and courage to start and complete this research study.

DEDICATION

This dissertation is dedicated to my family especially my brothers Robert and Omwoyo, my wife Annie, and daughter, Darlene Nyabiage for their love, support and understanding during my research. I also dedicate this to my parents, Mary Kemunto and Jeremiah Ondeng'i, for the financial and moral support they gave me.

TABLE OF CONTENTS

	Page
Title page	
Abstract	i
Declaration	iii
Acknowledgements	iv
Dedication	v
Table of contents	vi
List of Tables	xi
List of Figures	xiii
List of abbreviations	xv
CHAPTER 1:	1
BACKGROUND OF THE STUDY	1
1.1 THE RESEARCH PROBLEM	3
1.2 THE RESEARCH QUESTIONS	4
1.3 AIM OF THE STUDY	4
1.4 OBJECTIVES OF THE STUDY	5
1.5 RATIONALE FOR THE STUDY	5
1.6 SIGNIFICANCE OF THE STUDY	5
1.7 SCOPE OF THE STUDY	6
1.8 DEFINITION OF TERMS	6
1.9 CONCLUSION	7

CHAPTER 2:	8
THEORETICAL FRAMEWORK AND LITERATURE REVIEW	8
2.1 THEORETICAL FRAMEWORK	8
2.2 LITERATURE REVIEW.....	12
2.2.1 BIVARIATE NORMAL DISTRIBUTION (BND).....	12
2.2.1.2 Properties of the bivariate normal distribution	15
2.2.2 Factors that contribute to students' challenges in learning	17
2.2.2.1 Students background in mathematics	18
2.2.2.2 Lecturers' delivery of content and teaching methods	20
2.2.2.3 Evaluation methods and procedures	24
2.2.2.4 Inadequate teaching and learning resources	25
2.2.2.5 Students' attitudes towards some mathematics content	26
2.2.2.6 Students' lack of interest and inadequate effort in learning mathematics	28
2.3 CONCLUSION	30
CHAPTER 3:	31
RESEARCH METHODOLOGY	31
3.1 RESEARCH DESIGN	31
3.2 SOURCES OF DATA	32
3.3 POPULATION	33
3.4 SAMPLE AND SAMPLING TECHNIQUE	33
3.5 THE RESEARCH INSTRUMENTS	34
3.5.1 Bivariate normal distribution test (BNDT)	35
3.5.2 Bivariate normal distribution student questionnaire	36

3.5.3 Lecturer’s open-ended questionnaire	36
3.6 PROCEDURE FOR DATA COLLECTION	37
3.7 VALIDITY AND RELIABILITY OF RESEARCH INSTRUMENTS	38
3.7.1 Validity	38
3.7.2 Reliability	39
3.8 PILOT STUDY	41
3.9 DATA ANALYSIS	42
3.10 ETHICAL CONSIDERATIONS ON THE STUDY	43
3.11 LIMITATIONS OF THE STUDY	44
3.11 SUMMARY OF THE CHAPTER	44
CHAPTER 4:	46
FINDINGS	46
4.1 THE BIVARIATE NORMAL DISTRIBUTION TEST (BNDT) RESULTS	47
4.1.1 Summary of results of the bivariate normal distribution test	64
4.2 RESULTS FROM THE STUDENT QUESTIONNAIRE	65
4.2.1 Summary of the results from the student questionnaire	80
4.3 THE LECTURER OPEN-ENDED QUESTIONNAIRE	81
4.3.1 Results from the lecturer questionnaire.....	82
4.3.2 Summary of the results from the lecturer open-ended questionnaire.....	86
4.4 SUMMARY OF THE CHAPTER	86

CHAPTER 5:	88
DISCUSSION OF FINDINGS, CONCLUSION AND RECOMMENDATIONS	88
5.1 SUMMARY OF CHAPTERS	89
5.2 DISCUSSION OF THE FINDINGS	90
5.2.1 Can students state the equation of a normal distribution function?	91
5.2.2 Can students obtain the standard deviation given a normal distribution equation?	92
5.2.3 Can students obtain the mean given a normal distribution equation?	92
5.2.4 Can students differentiate a given exponential equation?	93
5.2.5 Can students integrate a given exponential equation?	94
5.2.6 Can students state the bivariate normal distribution equation?	95
5.2.7 Can students calculate the mean, $E(y/X = x)$ of the joint distribution, given the conditional distribution of a variable (e.g. $h(y/X = x)$)?	96
5.2.8 Can students calculate the variance, $Var(y/X = x)$ of the joint distribution, given the conditional distribution of a variable (e.g. $h(y/X = x)$)?	97
5.2.9 Can students calculate the mean of the two independent variables, given a bivariate random density function of a variable?	97
5.2.10 Can students obtain the standard deviations of the two independent variables, given a bivariate random density function of a variable?	98
5.3 REASONS WHY STUDENTS FIND THE LEARNING OF BIVARIATE NORMAL DISTRIBUTION DIFFICULT	100
5.3.1 Summary of the study	102
5.3.2.1 Research Question one (1): What do undergraduate students find difficult in the learning of bivariate normal distribution?	104
5.3.2.2 Research Question two (2): Why do students experience difficulties in learning the Bivariate normal distribution?	104
5.4 IMPLICATIONS OF THE RESEARCH FINDINGS	105

5.5 CONCLUSION	106
5.7 RECOMMENDATIONS	106
5.7 LIMITATIONS OF THE STUDY	108
5.8 SUGGESTIONS FOR FURTHER STUDIES	108
BIBLIOGRAPHY	109
APPENDICES	114
Appendix 1: Bivariate normal distribution test for students	114
Appendix 2: Questionnaire for students	118
Appendix 3: Lecturer open ended-questionnaire	123
Appendix 4: A letter of request to conduct research	125
Appendix 5: A letter of permission to conduct research	127
Appendix 6: Consent form for students	128
Appendix 7: Consent form for lecturers.....	129
Appendix 8: Description of the validity of the test questions	130
Appendix 9: Marking scheme for the student bivariate normal distribution test	132
Appendix 10: Standard normal distribution table:	138

List of Tables

- Table 3.1: The sub – section Reliability – Student questionnaire
- Table 3.2: Overall Reliability – Students’ bivariate normal distribution test
- Table 4.1a: Frequency of the students’ score in the BND test
- Table 4.1b: Students’ achievement as per marks scored on foundational knowledge on the BND
- Table 4.2: Distribution of students’ scores in question 1 of the BND test
- Table 4.3: Distribution of students’ scores in question 2 of the BND test
- Table 4.4: Distribution of students’ scores on question 3 of the BND test
- Table 4.5: Students’ performance on question 4 of the BND test
- Table 4.6: Students’ performance on question 5 of the BND test
- Table 4.7: Students’ performance on question 6 of the BND test
- Table 4.8: Students’ performance on question 7 of the BND test
- Table 4.9: Students’ performance on question 8a of the BND test
- Table 4.10: Students’ performance on question 8b of the BND test
- Table 4.11: Summary of results of statistical analysis of students’ performance in the BND test
- Table 4.12: Results of students’ perception of the effect of their interest in learning BND
- Table 4.13: Results of students’ perception on the effect of lack of background knowledge on learning BND content
- Table 4.14: Results of students’ perception of the effect of large number of students in class on the BND content
- Table 4.15: Results of students’ perception on the effect of inadequate learning resources on learning BND

- Table 4.16: Results of students' perception on the effect of long and complex equations on their learning the BND
- Table 4.17: Results of students' perception on the effect of tests and examinations on their learning of BND
- Table 4.18: Results of students' perception on their mathematical background on their learning of BND
- Table 4.19: Results of students' perception on the effect of the long bivariate normal distribution equation on their learning the BND
- Table 4.20: Results of students' perception on the effect of complex nature of the bivariate normal distribution equation on their learning of the content
- Table 4.21: Results of students' perception on the impact of their negative attitude towards bivariate normal distribution on their learning of the content
- Table 4.22: Results of students' perception on the impact of many parameters and constants in BND equation on their learning of the content
- Table 4.23: Results of students' perception of their background knowledge of the normal distribution equation
- Table 4.24: Results of students' views on having adequate knowledge of differentiation and integration on the learning of BND.
- Table 4.25: Results of students' views on lecturers' testing of difficult content on the BND
- Table 4.26: Results of students' views on their background in mathematical equations and understanding and performance in the bivariate normal distribution
- Table 4.27: Results of students' views on whether they find it difficult to state the bivariate normal distribution equation
- Table 4.28: Results of students' views on whether they understand step-by-step analysis of the bivariate normal distribution equation
- Table 4.29: Students' responses to the difficulties they experience in the learning of bivariate normal distribution.

List of Figures

- Figure 2.1: Schematic diagram showing the interaction of the main factors that impact on students' opportunity to learn
- Figure 2.2: Diagram showing the properties of a bivariate normal distribution.
- Figure 2.3: Diagram showing the normal curve symmetrical about the mean μ with the data values distributed on both sides of the mean.
- Figure 4.1a: Example of a student's solution to question 1 of the BND test
- Figure 4.1b: Example of a student's solution to question 1 of the BND test
- Figure 4.2a: Example of a student's solution to question 2 of the BND test
- Figure 4.2b: Example of a student's solution to question 2 of the BND test
- Figure 4.2c: Example of a student's solution to question 2 of the BND test
- Figure 4.3: Example of a student's solution to question 3 of the BND test
- Figure 4.4: Example of a student's solution to question 4 of the BND test
- Figure 4.5a: Example of a student's solution to question 5 of the BND test
- Figure 4.5b: Example of a student's solution to question 5 of the BND test
- Figure 4.6: Example of a student's solution to question 6 of the BND test
- Figure 4.7a: Example of a student's solution to question 7 of the BND test
- Figure 4.7b: Example of a student's solution to question 7 of the BND test
- Figure 4.8a: Example of a student's solution to question 8 of the BND test
- Figure 4.8b: Example of a student's solution to question 8 of the BND test
- Figure 4.8c: Example of a student's solution to question 8 of the BND test
- Figure 4.8d: Example of a student's solution to question 8 of the BND test
- Figure 4.9: A student stating the difficulties she experiences in learning bivariate normal distribution content

- Figure 4.10: A student stating the difficulties he experiences in learning bivariate normal distribution content
- Figure 4.11: A student stating the difficulties he experiences in learning bivariate normal distribution content
- Figure 4.12: A student stating the difficulties he experiences in the learning the bivariate normal distribution content
- Figure 4.13: A student stating the difficulties she experiences in learning bivariate normal distribution content
- Figure 4.14: A student stating the difficulties he experiences in learning bivariate normal distribution content
- Figure 4.15: A student stating the difficulties she experiences in learning bivariate normal distribution content
- Figure 4.16: A student stating the difficulties she experiences in learning bivariate normal distribution content

List of Abbreviations

BSc:	Bachelor of Science.
BND:	Bivariate Normal Distribution.
BNDT:	Bivariate Normal Distribution Test
SQ:	Student Questionnaire
LOEQ:	Lecturer Open-ended Questionnaire

CHAPTER 1

BACKGROUND OF THE STUDY

The bivariate normal distribution (BND) is a topic forming part of the content in statistics, which some students in Kenyan universities find difficult to learn and understand. This is from my own experience and discussions with students taking the course. From the lecturers point of view, they also find the content complex in terms of its delivery to students.

The BND is a joint density equation involving two normal distributions. From the discussions with students taking the course, it challenges some students in its analysis to obtain the joint mean and joint variance of the conditional distribution and hence understand this distribution.

The BND is introduced in the first year of study in the Kenyan university, as a normal distribution. It is then studied in detail in the second year of study under marginal density function and conditional density distribution of a normal distribution, then as a BND.

From my own experience and discussions with students and lecturers, learning and teaching BND content respectively, learning of BND requires the students to have background knowledge of calculation of the mean, the variance and covariance, the standard deviation, the normal distribution equation, the marginal density function and the conditional probability distribution to understand the content.

The students also need to be equipped with knowledge of both differentiation and integration to enable them to differentiate an exponential equation and to find the integral of an exponential function of a bivariate normal distribution when analyzing the equation. Shojaie, Aminghafari and Mahammadpour (2012) acknowledge that students have difficulties in introductory courses in probability and statistics, and these difficulties include finding the joint distribution of a function and calculation of bivariate expectation and complex integrals.

Notably, for a normal distribution, $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$ where the two parameters μ and σ are the mean and the standard deviation respectively of a population of X – values. π (pi) and e (exponent) are constants. X is a continuous random variable for the distribution.

A graph of $f(x)$ against x is a smooth bell-shaped curve called a normal curve or a normal probability distribution and it is symmetrical about the mean μ .

Given marginal distributions for continuous random variables X and Y, where

$$f(x) = \frac{1}{\sigma_x \sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{x-\mu_x}{\sigma_x} \right)^2}$$

And

$$f(y) = \frac{1}{\sigma_y \sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{y-\mu_y}{\sigma_y} \right)^2}$$

The joint probability density function for the functions $f(x)$ and $f(y)$ is given by;

$$f(x, y) = f(x) \cdot f(y)$$

$$f(x, y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \exp\left(-\left[\frac{1}{2(1-\rho^2)}\right]\left\{\left(\frac{x-\mu_x}{\sigma_x}\right)^2 - 2\rho\left(\frac{x-\mu_x}{\sigma_x}\right)\left(\frac{y-\mu_y}{\sigma_y}\right) + \left(\frac{y-\mu_y}{\sigma_y}\right)^2\right\}\right)$$

is a bivariate normal distribution for two random variables X and Y.

The parameter ρ is the correlation coefficient of x and y.

Shojaie, Aminghafari and Mahammadpour (2012), acknowledge that students have difficulties in introductory courses in probability and statistics which include joint distribution and calculation of bivariate expectation and complex integrals. Similarly, Garfield (1995) acknowledge that, teaching statistical courses is challenging because students have a variety of backgrounds and abilities, and many have had negative experiences with statistics and mathematics. It is possible that the students' attitude towards some statistical courses influences their learning and understanding of the course content and their performance in the courses.

According to Batanero, Burrill, Reading and Rossman (2011), some teachers are willing to learn about some of the content of statistics and spend additional time teaching this subject because they acknowledge its practical importance. However, these teachers feel that their students experience greater difficulties in any other mathematical topic. Students lack motivation and commitment to learning some statistical courses hence they experience difficulties.

Tishkovskaya and Lancaster (2012, p. 4) assert that-

“over the years there has been strong anecdotal evidence that students at university develop antipathy towards statistics and, typically, students at all levels lack interest in learning when taking introductory statistics courses.-”

One possible reason for the students' difficulty in learning the BND might be that they find the equation too long to comprehend. Ben-Zvi and Garfield (2004) asserted that many statistical ideas and rules are complex, difficult, and/or counterintuitive and it is therefore difficult to motivate learners to engage in the hard work of learning statistical content. A BND equation may be considered by some students as difficult and complex. The equation has two constants (π and e) and five parameters, namely; the expected values of two random variables, the standard deviations of the two random variables and the correlation coefficient of the random variables which students need to understand while they learn BND content.

According to Ben-Zvi and Garfield (2004), many learners have difficulty with the underlying mathematics (such as fractions, decimals, proportional reasoning, algebraic formulas and functions, among others) and that could interfere with learning the related statistical concepts. This implies that the students' lack of the pre-requisite statistical knowledge negatively affect their learning of related and advanced statistical content.

Shojaie, Aminghafari and Mahammadpour (2012) asserted that many students find it difficult to calculate the expectation of a function of BND (X, Y), such as $E(|XY|)$, and are not confident of the result. Consequently, this factor discourages them from learning the content.

According to Nardi (2011), the mathematics taught at many universities consists largely of theorems and their proofs, which students tend to memorize with little understanding; when they are required to prove a theorem, students simply regurgitate the proof. To understand statistical concepts well, students should be taught from known content to unknown. They should also be taught from simple to complex content.

The BND is complex and in this study the researcher investigates the difficulties students face in learning the content. This in turn would help lecturers to identify the best approach they should use in teaching the content effectively so that the students learn the topic in a meaningful way.

1.1 THE RESEARCH PROBLEM

The BND content in probability and statistics challenges most students in terms of understanding the equation and its analysis. From experience, interaction and discussions with students and lecturers, most students find the BND equation long to comprehend. The

equation includes two random variables and their respective means, their standard deviations and the correlation coefficient. Hence, the learning and analysis of the BND has proved difficult for some students (Shojaie, Aminghafari and Mahammadpour, 2012). Students' understanding of BND in probability and statistics III on the second- year level in the Kenyan university is very important because it lays a solid foundation for further studies in statistics on the third- and fourth- year levels.

The problem of the study was to investigate the difficulties undergraduate students experience in the learning of BND and hence proffer solutions on what needs to be done to teach BND effectively in the Kenyan university.

1.2 RESEARCH QUESTIONS

This research study attempted to answer the following research questions;

1. What do students find difficult in learning the bivariate normal distribution?

To address this *research question*, the following guiding sub-questions were put forward:

1.1 Can the students state the equation of a normal distribution function?

1.2 Can the students obtain the mean given a normal distribution equation?

1.3 Can the students obtain the the standard deviation given a normal distribution equation?

1.4 Can the students differentiate a given exponential equation?

1.5 Can the students find the integral of a given exponential equation?

1.6 Can the students state the BND equation?

1.7 Given the conditional distribution of a variable (e.g. $h(y/X = x)$), can the students calculate the mean, $E(y/X = x)$ of the joint distribution?

1.8 Given the conditional distribution of a variable (e.g. $h(y/X = x)$), can the students calculate the variance, $Var(y/X = x)$ of the joint distribution?

2. Why do students experience difficulties in learning BND?

1.3 AIM OF THE STUDY

The primary aim of the study was to investigate what undergraduate students in the Kenyan university find difficult in the learning of BND.

1.4 OBJECTIVES OF THE STUDY

The main objective of the study was to investigate what undergraduate students find difficult in the learning of BND in the Kenyan university and why the students experience these difficulties.

The study also aimed to find ways of helping lecturers teach the content of bivariate normal distribution to students more effectively.

1.5 RATIONALE FOR THE STUDY

The rationale for this study emanates from the researcher's experience in studying BND in probability and statistics III, at a Kenyan university. The difficulties the researcher had in learning BND and comprehending the long BND equation evoked his interest to undertake this research study.

In addition, the persistent underperformance of students in the Kenyan university on BND in probability and statistics, and their general lack of interest in statistics and related courses highlight the urgent need for research into the problem and the possibility of finding a solution.

1.6 SIGNIFICANCE OF THE STUDY

The findings of this research will give insight into the Kenyan university students' difficulties in learning BND and statistics in general. This would make a substantial contribution to knowledge which would foster more effective teaching of the content and consequently students' understanding of the BND equation and its analysis.

The study would be important to statistics lecturers in the Kenyan university and perhaps other lecturers in other Kenyan universities in that it could help them to understand the problems students face in learning BND and statistics in general in order to find the correct possible way of teaching the content more effectively. The study would also be important to lecturers teaching statistics in universities outside Kenya.

1.7 SCOPE OF THE STUDY

This study was confined to the investigation into undergraduate students' difficulties in learning of BND in a Kenyan university.

The study also looked into the foundational knowledge related to the BND content acquired by students-; for example, the knowledge of differentiation and integration of exponential equations applicable in the analysis of BND.

1.8 DEFINITION OF TERMS

For the purpose of this research study, the following terms have been defined:

Learning: Learning entails change in a person as regards the individuals' insights, behaviour, perception or motivation and which change leads to the addition of knowledge or the ability to do something that the learner could not do before. This study looks at what the students have gained through their experience in the learning of the bivariate normal distribution and foundational related content.

Teaching: Teaching can be seen as creating opportunities for learning to take place, as well as the process of helping learners to learn.- (Du Plessis et al., 2007).

Bivariate normal distribution: It is a distribution relating to or involving two independent random variables.

Normal probability distribution: This is a distribution in which a random variable, say x , is distributed normally with mean, μ and variance, σ^2 . That is, $X \sim N(\mu, \sigma^2)$. The parameters μ and σ determine the shape of the graph which is symmetrical and bell-shaped.

Normal probability density function: This is an equation used to compute probabilities of continuous random variables which must satisfy the following properties:-

- The graph of the equation must be greater than or equal to zero for all the possible values of the random variable.
- Bell-shaped
- The mean, median, and mode are equal and located at the centre of the distribution

- It is uni-modal (-only 1 mode)
- Symmetrical about the mean
- The curve is continuous – i.e., there are no gaps or holes. For each value of X, there is a corresponding value of Y.
- The curve never touches the x-axis. Theoretically, no matter how far, in either direction, the curve extends, it never meets the x-axis, but it gets increasingly closer to the x-axis.
- The total area under the normal distribution curve is equal to 1.00 or 100%.
- 0.50 or 50% of the area, lies to the left of the mean and 0.50 or 50% lies to the right of the mean.

Standard normal probability distribution: It is a normal distribution with a mean of 0 and a standard deviation of 1. The Z value is the distance between a selected value, designated X, and the population mean, μ , divided by the population standard deviation, σ .

$$z = \frac{(X - \mu)}{\sigma}$$

1.9 CONCLUSION

This chapter dealt with the background of the study, the research problem, the aim of the study, objectives of the study and purpose of the study. Also discussed in this chapter is the rationale for the study, the significance of the study, the scope of the study and the definition of terms used in this study.

The next chapter discusses theoretical framework and literature review of the study.

CHAPTER 2

THEORETICAL FRAMEWORK AND LITERATURE REVIEW

The aim of this study was to investigate students' difficulties in the learning of BND in a Kenyan university. The theoretical framework that guided this study and review of some related literature are presented in this chapter.

2.1 THEORETICAL FRAMEWORK

This study is based on the theoretical perspective of Ingvarson, Beavis, Bishop, Peck and Elsworth (2004), which focuses on the main factors that influence student achievement in mathematics. Although the theory was developed mainly within the general school context, the researcher finds it relevant in learning mathematics, for example, the BND in the university. According to the theory of Ingvarson et al. (2004), the quality of a teacher's instructional practices and activities is fundamental in determining student learning opportunities. The theory advances that the quality of the student opportunity to learn is shaped by teacher knowledge, teacher beliefs, school contextual factors, school system factors and student factors. All these factors can as well affect student learning of a BND in the university.

Teacher knowledge influences the student learning outcomes. Teacher knowledge encompasses knowledge of the content to be taught, knowledge of how students learn and understand that specific content and the knowledge of methods to teach that specific content (Nicette & Marissa, 2014). The knowledge of the mathematics content and how students learn it enables teachers to choose appropriate mathematical tasks and classroom resources that feed into the learning process (Anthony & Walshaw, 2009). Ingvarson et al. (2004) argue that the richer the educational background of the teacher, the more likely the teacher will seek an environment which encourages professional development. A situation where teachers have insufficient knowledge of the content they are teaching leads to one in which students have limited opportunity to learn. The teachers' experience in teaching specific mathematical content positively influences the students' learning and understanding of that content, resulting in better student performance. Consequently, teachers' inadequate knowledge of the content they teach may lead to the difficulties encountered by students in the subject.

Lecturers teaching mathematics, for example, the BND in universities must have a rich content base in the content they are teaching to enable their students grasp the content fully. A sufficient knowhow of what the lecturer is teaching, is paramount to the student understanding of the content and hence the student performance in mathematics. Lecturers' insufficient knowledge of mathematics content they are required to teach, for example the BND, leads to the students' limited opportunity to learn and therefore poor performance.

Teachers' beliefs on teaching and learning shape their approach and influence their teaching and instructional decisions. Teachers' beliefs influence planning, instruction given to the students and classroom management, which consequently influence the students' learning outcomes in mathematics. Poor and/or inadequate planning by teachers, inadequate instruction to students and poor classroom management by teachers may lead to the students' difficulties in learning in mathematics. The teaching of BND requires adequate planning and sufficient instructions by the lecturers to enable students' effective learning of the content.

Ingvarson et al's (2004) theory, also advances that teacher capacity influences student learning outcomes in mathematics. Teacher capacity constitutes teachers' knowledge, beliefs and understandings of teaching in general, and of teaching mathematics in particular. Teacher capacity is shaped by the professional community in the school, their qualifications and initial teacher training, and the involvement of the teacher in professional development, (Ingvarson et al., 2004). Ingvarson et al's (2004) theory, emphasises that the more the teacher has exposure to these environments and opportunities to learn, the greater each teacher's capacity to be an effective teacher. The students' opportunity to learn encompasses classroom practices of teachers, which are partly shaped by the school and the background of the students. At higher learning institutions, the lecturers' understanding of the BND content and statistics in general, and teaching of the content enables effective and meaningful students' learning of the content.

According to Ingvarson et al's (2004) theory, the students' learning outcomes relate to their achievement in mathematics. The outcomes encompass both student knowledge and student attitudes to learning mathematics. Learning outcomes are related to the classroom practices; the more successfully students learn, the more likely it is that the teacher will adopt practices that encourage further successful learning. A situation in which a teacher lacks adequate qualifications in some mathematical concepts, leads to students' difficulties in learning mathematics. Similarly, teachers' inadequate knowledge in mathematics, their lack of

understanding of mathematical content and inadequate exposure to learning opportunities, may all lead to the students' difficulties in learning mathematics. Students' with adequate foundational knowledge in statistics are likely to achieve better grades in BND and subsequent related statistical topics. Students who obtain higher grades in statistics in their earlier years in the university, tend to adopt similar learning practices for further learning in statistics courses.

The student factor encompasses the student background, the student attitude, the student interest and ability level, and their background knowledge in mathematics, all of which influence the teachers' teaching and the students' class participation in mathematics as well as learning outcomes. Students' negative attitudes to mathematical content affect their learning negatively (Ogbonnaya & Mji, 2013). Students who lack interest in mathematics, find it difficult to understand the content taught by the teachers and hence perform poorly in the subject. Also, students with a weak background in mathematics, for example, statistical equations, find learning of advanced content (for example, long statistical equations) difficult; consequently they perform dismally in examinations. The students' weak background in mathematics, negative attitude and lack of interest towards some mathematical and lack of participation in class while teachers are teaching mathematics, may possibly lead to their difficulties in learning mathematics. Students' negative attitude towards the bivariate normal distribution coupled with a lack of adequate background in statistics in general, negatively affect the lecturers' teaching and consequently, the students' learning of the content.

School contextual factors involve the working conditions in the school in which the teacher works. This involves the school teaching and learning conditions, for example, the availability of adequate learning and teaching resources, the time allocated for teaching, the professional work community and other factors. Insufficient teaching and learning resources make learning of mathematics difficult. For effective teaching and meaningful learning to take place, a learning institution should be equipped with adequate learning and teaching resources. Inadequate time allocated to the teaching and learning of any given mathematical content negatively affects students' learning of the content. Similarly, inadequate teaching staff in a learning institution, affects teaching and learning negatively because in that situation, teachers are given a large workload. Poor working conditions in schools, for example, unavailability of adequate teaching and learning resources (Mbugua et al., 2012), inadequate time allocated for teaching mathematics, may also lead to students' difficulties in learning mathematics. The

teaching and learning in the university must be conducive for both the lecturers and students. Effective teaching and learning of BND in statistics, in a university can take place when sufficient teaching and learning resources are adequately available.

The school system factor involves the curriculum and education policies under which the school is run. Each school system has its own policies in education and the government's established curriculum which guides the teaching and learning of a given subject. A change of the school curriculum and policies in education regarding the teaching and learning of mathematics may cause students to experience difficulties while learning mathematics.

Figure 2.1 illustrates the main factors that may lead to students' learning difficulties in mathematics as depicted from Ingvarson et al. (2004) study.

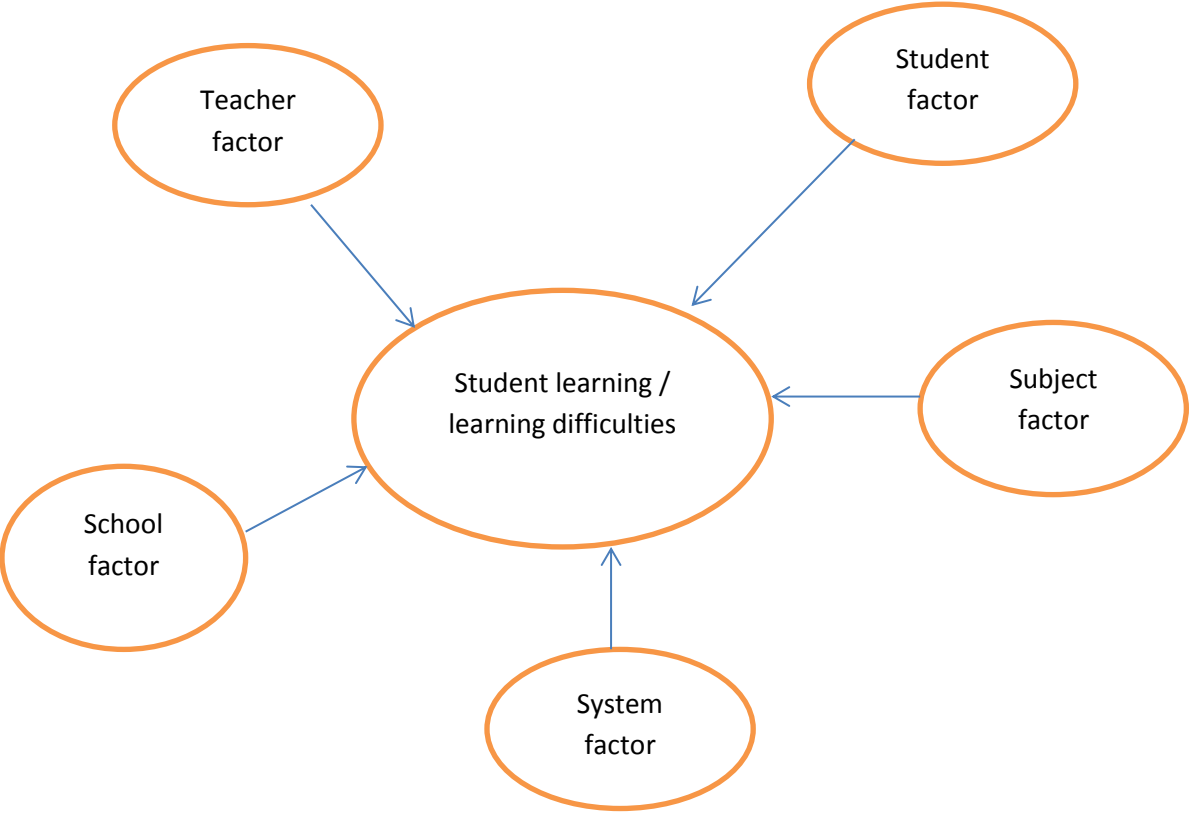


Figure 2.1: Schematic diagram showing the interaction of the main factors that impact on students' opportunity to learn

2.2 LITERATURE REVIEW

The literature review provides a description of the bivariate normal distribution and a critical synthesis of published research studies that are related to difficulties that students encounter in learning statistics and mathematics in general.

2.2.1 Bivariate normal distribution (BND)

The bivariate normal distribution of a continuous random variable X and Y is given by:

$$f(x, y) = \left(\frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \right) \text{Exp}\left(-\left[\frac{Z}{2(1-\rho^2)}\right]\right)$$

$$\text{Where } Z = \left(\frac{(x-\mu_x)^2}{\sigma_x^2} \right) - \left(\frac{2\rho(x-\mu_x)(y-\mu_y)}{\sigma_x\sigma_y} \right) + \left(\frac{(y-\mu_y)^2}{\sigma_y^2} \right)$$

$$\rho = \text{correlation Coefficient, } \rho = \text{Cor}(x, y) = \frac{V_{xy}}{\sigma_x\sigma_y},$$

$$0 \leq x \leq \infty, 0 \leq y \leq \infty, \sigma_x > 0, \sigma_y > 0 \text{ and } -1 \leq \rho \leq 1$$

$\mu_x =$ The expectation of X, $E(x)$ $\sigma_x =$ standard deviatin of X

$\mu_y =$ the expectation of Y, $E(y)$ $\sigma_y =$ standard deviatin of Y

$f(x, y)$ is the joint probability density function.

The marginal density function of X is given by:

$$g(x) = \frac{1}{(2\pi\sigma_x\sigma_y\sqrt{1-\rho^2})} \int_{-\infty}^{+\infty} \exp\left[-\frac{1}{2(1-\rho^2)}\right] \left\{ \left(\frac{x-\mu_x}{\sigma_x}\right)^2 - 2\rho\left(\frac{x-\mu_x}{\sigma_x}\right)\left(\frac{y-\mu_y}{\sigma_y}\right) - \left(\frac{y-\mu_y}{\sigma_y}\right)^2 \right\} dy$$

$$\text{If } u = \frac{x-\mu_x}{\sigma_x} \text{ and } v = \frac{y-\mu_y}{\sigma_y},$$

$$g(x) = \frac{1}{(2\pi\sigma_x\sigma_y\sqrt{1-\rho^2})} \int_{-\infty}^{+\infty} \exp\left[-\frac{1}{2(1-\rho^2)}\right] \{(u)^2 - 2\rho(u)(v) - (v)^2\} dy$$

Then $\frac{dv}{dy} = \frac{1}{\sigma_y}$. It implies that $dy = \sigma_y dv$

From the marginal density function of x , $v^2 - 2\rho uv$ can be written as $(v - \rho u)^2 - \rho^2 v^2$, using quadratic equation knowledge.

$$\begin{aligned} \text{Therefore, } g(x) &= \frac{1}{2\pi\sigma_x\sqrt{1-\rho^2}} \int e^{-\frac{1}{2(1-\rho^2)}[u^2 - \rho^2 v^2 + (v - \rho u)^2]} dv \\ &= \frac{1}{2\pi\sigma_x\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)}u^2(1-\rho^2)} \int e^{-\frac{1}{2(1-\rho^2)}(v-\rho u)^2} dv \end{aligned}$$

But $\frac{1}{\sqrt{2\pi}\sqrt{1-\rho^2}} \int e^{-\frac{1}{2(1-\rho^2)}(v-\rho u)^2} dv$ can be considered as an integral of a normal p.d.f with mean ρu and variance $(1 - \rho^2)$ and hence has an integral equal to 1.

$$\text{Therefore } g(x) = \frac{1}{\sigma_x\sqrt{2\pi}} e^{-\frac{1}{2}u^2}$$

And substituting $u = \frac{x - \mu_x}{\sigma_x}$, back to the equation, it results:

$$g(x) = \frac{1}{\sigma_x\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x - \mu_x}{\sigma_x}\right)^2}$$

$$\text{So that } \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dy dx = 1$$

Hence $f(x, y)$ is a p.d.f

$g(x)$ is normal with mean μ_x and variance δ_x^2 .

It can also be shown that, $h(x)$, the normal p.d.f of Y is also normal with mean μ_y and variance δ_y^2 .

Consider the conditional density, $\phi(Y|X) = \frac{f(x, y)}{g(x)}$ Where $f(x, y)$ is the bivariate normal density function of X and Y and $g(x)$ is the marginal density function of X .

$$\phi(Y|X) = \frac{\frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_x}{\sigma_x}\right)^2 - 2\rho\left(\frac{x-\mu_x}{\sigma_x}\right)\left(\frac{y-\mu_y}{\sigma_y}\right) - \left(\frac{y-\mu_y}{\sigma_y}\right)^2\right]}}{\frac{1}{\sigma_x\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu_x}{\sigma_x}\right)^2}}$$

$$= \frac{\sigma_x \sqrt{2\pi}}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_x}{\sigma_x}\right)^2 - 2\rho\left(\frac{x-\mu_x}{\sigma_x}\right)\left(\frac{y-\mu_y}{\sigma_y}\right) - \left(\frac{y-\mu_y}{\sigma_y}\right)^2\right] + \frac{1}{2}\left(\frac{x-\mu_x}{\sigma_x}\right)^2}$$

(Subtract the powers since base e is the same)

$$\text{Let } u_x = \frac{x-\mu_x}{\delta_x} \text{ and } u_y = \frac{y-\mu_y}{\delta_y},$$

$$\begin{aligned} \text{Hence, } \phi(Y|X) &= \frac{\sigma_x \sqrt{2\pi}}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)}[u_x^2 - 2\rho u_x u_y + u_y^2] + \frac{1}{2}u_x^2} \\ &= \frac{1}{\delta_y \sqrt{2\pi} \sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)}[u_x^2 - 2\rho u_x u_y + u_y^2] + \frac{1}{2}u_x^2} \\ &= \frac{1}{\delta_y \sqrt{2\pi} \sqrt{1-\rho^2}} e^{-\left[\frac{1}{2(1-\rho^2)} - \frac{1}{2}\right]u_x^2 - \frac{1}{2(1-\rho^2)}[u_y^2 - 2\rho u_x u_y]} \\ &= \frac{1}{\delta_y \sqrt{2\pi} \sqrt{1-\rho^2}} e^{-\frac{1}{2}u_x^2 \left[\frac{\rho^2}{1-\rho^2}\right] - \frac{1}{2(1-\rho^2)}[u_y - \rho u_x] + \frac{1}{2(1-\rho^2)}\rho^2 u_x^2} \\ &= \frac{1}{\delta_y \sqrt{2\pi} \sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)}[u_y - \rho u_x]^2} \\ &= \frac{1}{\delta_y \sqrt{2\pi} \sqrt{1-\rho^2}} e^{-\frac{1}{2} \frac{[u_y - \rho u_x]^2}{1-\rho^2}} \end{aligned}$$

$$\text{But } u_x = \frac{x-\mu_x}{\delta_x} \text{ and } u_y = \frac{y-\mu_y}{\delta_y}$$

Therefore,

$$\phi(Y|X) = \frac{1}{\delta_y \sqrt{2\pi} \sqrt{1-\rho^2}} e^{-\frac{1}{2\delta_y^2(1-\rho^2)}[y - \mu_y - \rho \frac{\delta_y}{\delta_x}(x - \mu_x)]^2}$$

$$[y - \mu_y - \rho \frac{\delta_y}{\delta_x}(x - \mu_x)]^2 \text{ can be written as } [y - (u_y + \rho \frac{\delta_y}{\delta_x}(x - \mu_x))]^2$$

$$\phi(Y|X) = \frac{1}{\delta_y \sqrt{2\pi} \sqrt{1-\rho^2}} e^{-\frac{1}{2\delta_y^2(1-\rho^2)}[y - (u_y + \rho \frac{\delta_y}{\delta_x}(x - \mu_x))]^2}$$

The above expression shows that the conditional distribution of Y given $X = x$ is normal with

$$\text{Mean, } E(y|X = x) = u_y + \rho \frac{\delta_y}{\delta_x}(x - \mu_x) \text{ and}$$

$$\text{Variance, } \text{Var}(y/X = x) = \delta_y^2(1 - \rho^2)$$

Similarly, $\phi(X|Y) = \frac{f(x,y)}{h(y)}$ is given by:

$$\phi(X|Y) = \frac{1}{\delta_x \sqrt{2\pi} \sqrt{1 - \rho^2}} e^{-\frac{1}{2\delta_x^2(1-\rho^2)} [x - (u_x + \rho \frac{\delta_x}{\delta_y}(y - \mu_y))]^2}$$

The conditional distribution of X given Y = y is normal with

$$\text{Mean, } E(x/Y = y) = u_x + \rho \frac{\delta_x}{\delta_y}(y - \mu_y) \text{ and}$$

$$\text{Variance, } \text{Var}(x/Y = y) = \delta_x^2(1 - \rho^2)$$

2.2.1.2 Properties of the BND

- The random variables having the bivariate random variables are independent if, and only if, they are uncorrelated.
- For two random variables having the bivariate random variables, the marginal distributions are normal but the converse is not necessarily true.
- The normal regression surface has a maximum at $x = \mu_x$ and at $y = \mu_y$.

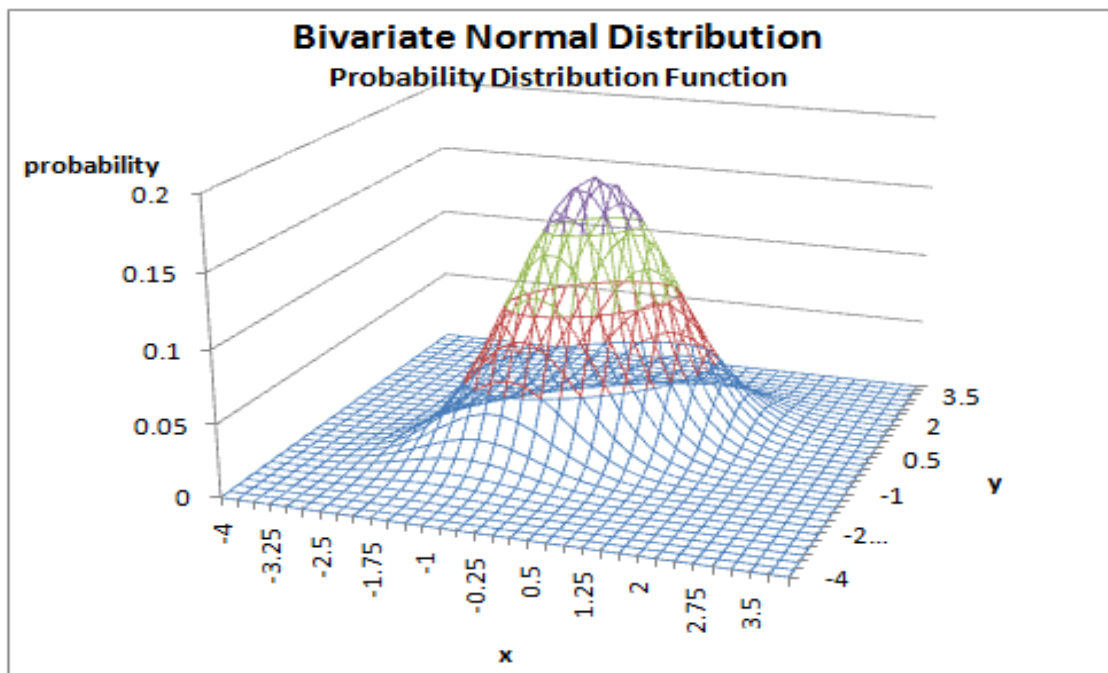


Figure 2.2: Diagram showing the properties of a bivariate normal distribution

Notably, for a normal distribution, $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$ where the two parameters μ and σ are the mean and the standard deviation respectively of a population of X – values. π (pi) and e (exponent) are constants. X is a normal random variable for the distribution.

A graph of $f(x)$ against x is a smooth bell-shaped curve called a normal curve or a normal probability distribution and it is symmetrical about the mean μ .

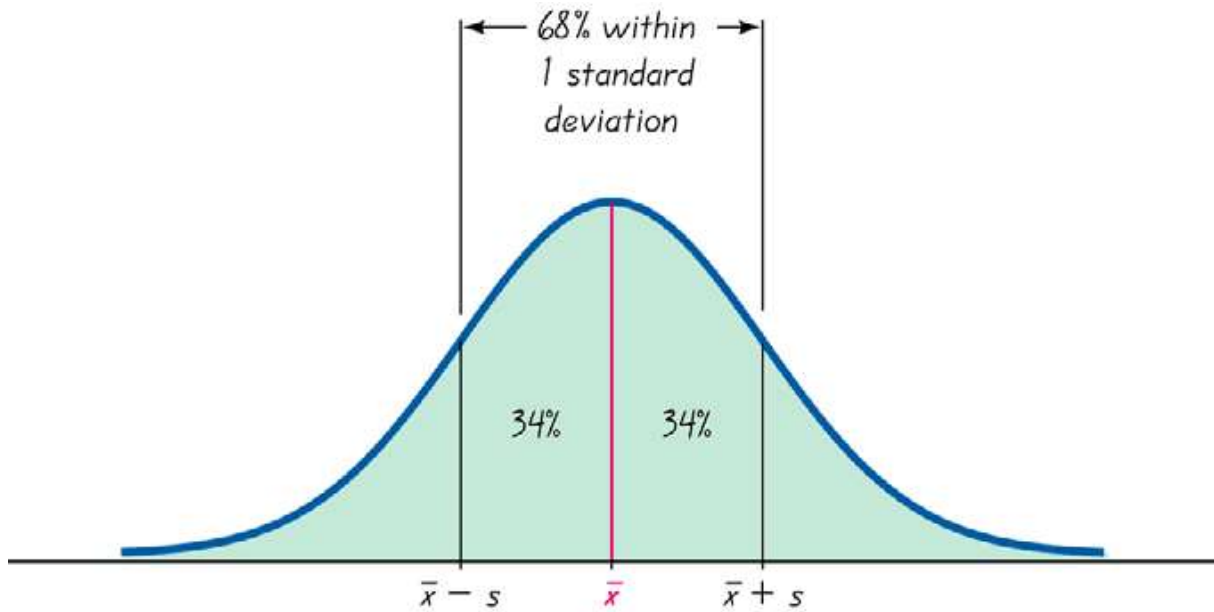


Figure 2.3: Diagram showing the normal curve symmetrical about the mean μ ; data values are distributed on both sides of the mean.

Marginal distributions for random variables X and Y

$$f(x) = \frac{1}{\sigma_x\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu_x}{\sigma_x}\right)^2}$$

And

$$f(y) = \frac{1}{\sigma_y\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{y-\mu_y}{\sigma_y}\right)^2}$$

The joint probability density function for the functions $f(x)$ and $f(y)$ is given by;

$$f(x, y) = f(x).f(y)$$

$$f(x, y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \exp\left(-\left[\frac{1}{2(1-\rho^2)}\right]\left\{\left(\frac{x-\mu_x}{\sigma_x}\right)^2 - 2\rho\left(\frac{x-\mu_x}{\sigma_x}\right)\left(\frac{y-\mu_y}{\sigma_y}\right) - \left(\frac{y-\mu_y}{\sigma_y}\right)^2\right\}\right)$$

a bivariate normal distribution for two random variables X and Y.

The parameter ρ is the correlation coefficient of x and y.

In the Kenyan university, students learn the BND from the first year where it is introduced as the normal distribution and the standard normal distribution. In the second year students learn the marginal density functions of distributions, including the normal marginal density functions. This forms the knowledge on concepts or the foundational knowledge necessary for effective learning the BND. Then still in the second year, students further learn BND under probability and statistics III, where they learn the BND equation, the analysis of the equation stepwise to derive the conditional mean and variance equations for a variable given the other variable.

2.2.2 Factors that contribute to students' challenges in learning

Many studies postulate some of the factors that influence student learning of mathematics in general and learning of statistics in particular. It is worth noting that the researcher acknowledge that there exists limited sources of literature directly connected to the research topic and that few studies have been conducted on the teaching and learning of bivariate normal distribution and precisely the students' difficulties in the learning of the BND. Some of the factors identified in the literature that influence students learning are as follows;- Students' weak background in mathematics (Garfield, 1995; McPhan & Pegg, 2008; Ogbonnaya & Mji, 2013), teachers/lectures teaching methods (Anthony & Walshaw, 2009; Batanero, Burrill, Reading and Rossman, 2011; Nicholson & Darnton, 2003; Tishkovskaya & Lancaster, 2012), evaluation methods and procedures (Ogoli, 1998; Wilkens, 1975);, inadequate teaching and learning resources (Eshiwani, 1983; Mbugua et al, 2012; Ogbonnaya & Mji, 2013);, students' attitudes towards some mathematics content (Anthony & Walshaw, 2009; Greg & John, 2008; Ng'ang'a, 2010) and students' lack of interest or their inadequate effort in learning mathematics (Bell et al, 1983; Davis, 1984; Tishkovskaya & Lancaster, 2012).

2.2.2.1 Students' weak background in mathematics

Background knowledge entails the knowledge students have previously learnt which is applicable in the learning of the current or future study. It is possible that some content students learn was never thoroughly understood and these students find it difficult to learn related content. But students with adequate foundational knowledge find it easy to learn advanced and related mathematical content. Students with a weak foundation in mathematical content, find the current and future learning of mathematics difficult since they lack adequate foundational knowledge which may be related and useful to the current and future learning of mathematics. Teaching statistical courses is challenging because they serve students with varying backgrounds and abilities, many of whom have had negative experiences with statistics and mathematics (Garfield, 1995). It is worth noting that knowledge is better built with a firm foundation which is based on the students' background knowledge. For example, learning of the BND requires a student who is well equipped with adequate pre-requisite knowledge on the basic statistical concepts, differentiation and integration of exponential equations.

Difficulties which students encounter in the learning of BND in the university may be as a result of the difficulties they experienced in earlier study of mathematics, most notably in secondary school since they have inadequate background knowledge. Adequate background in basic statistical courses ensures that the students are well equipped with the knowledge necessary for learning the advanced content in statistics.

Students' poor mathematical background in mathematics affects the learning of mathematics content. Studies (e.g. McPhan & Pegg, 2008) show that it is important for students to see the direct link between what they are studying and what its impact on their life is. If the link is not evident there is very little motivation or reason to study the subject at a higher level. McPhan and Pegg (2008) also show that students seem to find mathematics more difficult and perceive they do not achieve the results they need to go on to higher mathematical study. According to research, for example, Mugisha (2012, p. 18),

“The idea of restricting attention to the firm grounding of concepts and skills as well their application during first year calculus course underlie an important assumption, namely, that the full understanding of calculus from first to fourth year at university

must be calibrated on a gradient so that the next progression is based on a very firm ground of the earlier coverage of the materials.”

Mugisha (2012), goes on to assert that students should also effectively learn the rules for algebraic manipulations and they must understand the mathematics behind the rules. A firm foundation is required in arithmetic and algebra in order to perform well in differentiation and integration (calculus) in higher learning institutions. Inadequate knowledge in the basic arithmetic and algebraic equations, leads to student difficulties in learning and understanding of calculus content.

According to Ogbonnaya and Mji (2013), students’ weak background in mathematics contributes to their poor performance and their success in higher grade levels in school depends to a large extent on how well the student was prepared in the lower grade levels. This is because when students have not mastered and understood the content, for example, the parameters and constants in basic equations, they find it difficult to learn and understand further knowledge which requires the use of the knowledge of concepts foundational to the content under study. Students with inadequate knowledge in statistical equations may find it difficult to learn advanced statistical equations, for example, the BND equation which requires the use of basic foundational knowledge for effective learning.

Michaels and Smit (2008) asserted that, many students admitted to higher education are considerably ill-prepared for tertiary study, particularly learners from poverty-stricken and disadvantaged communities. It is worth noting that students with an excellent and adequate background in mathematics on lower grade levels, perform well in mathematics in higher grade levels. Hence, a good foundation in basic mathematical content is important for students aiming to further their studies in mathematics.

Ogoli (1998) maintains that a poor background in mathematics coupled with long formulae, hinders students’ learning in mathematics. Rosenstein (2005) warns that students who finally take several semesters of calculus in college, are disadvantaged if they have difficulties with arithmetic and algebra. Therefore, a weak background in mathematics, negatively affects student learning of advanced mathematical knowledge as the students further their education.

2.2.2.2 Lecturers' delivery of content and teaching methods

The delivery of content by teachers plays a key role in the learner's performance or difficulties they face in mathematics. Tishkovskaya and Lancaster (2012, p. 2) asserts that, not all statisticians are aware of the full learning potential of their discipline. Some teachers find statistical concepts very complex and too complicated to explain to the students while others lack adequate language to explain the content. In most universities, lecturers employ a lecture method as a means of teaching which is not appropriate when teaching some statistical units, for example, the BND. Other methods like chalkboard illustrations and a learner-centred approach would be appropriate in teaching statistical units, for example, the BND. Use of technology can also be an appropriate method of teaching the BND in which online resources such as the you-tube videos can be accessed to enhance students' understanding of the content.

Nicholson and Darnton (2003) warns that statistics may pose particular problems if the teacher has not studied it earlier or if the teacher is not a statistics subject specialist. This is because the style of teaching mathematics and statistics, generally differs, mainly with regard to the emphasis on statistical concepts and ways of knowing.

Anthony and Walshaw (2009) assert that the knowledge of the mathematics content and how students learn enables teachers to choose appropriate mathematical tasks and classroom resources that feed into the learning process. The mathematical tasks involve group work discussions, a teacher to cater for students' individual differences and students' individual solution of mathematics questions and so on enables effective learning of mathematics.

Lecturer knowledge influences the learning outcome of a student. Fennema and Franke (1992) argue that a teacher with conceptual understanding of mathematics influences classroom instruction in a positive way and it is therefore important for teachers to have mathematics knowledge. Studies, for example, Ganal & Guiab (2014), have shown that the teachers' knowledge of the subject matter and their ability to communicate it are very important factors in the teaching and learning process. Teacher knowledge entails knowledge of the content to be taught, knowledge of how students learn and understand that specific content and the knowledge of the methods to teach the content.

Cobb and Moore (1997) emphasise that,

“To teach statistics well, it is not enough to understand the mathematical theory; it is not even enough to understand also the additional, non-mathematical theory of statistics”.

They believe that to teach statistical units well, one must have a ready supply of real illustrations and know how to use those illustrations to involve students in the development of their critical judgement (Cobb & Moore, 1997).

Teacher behaviour is influenced by the teacher’s understanding of the particular content and knowledge of how students learn it (Koehler & Grouws, 1992). According to Anthony and Walshaw (2009), the importance of the knowledge of the mathematical content and how students learn it enables teachers to choose appropriate mathematical tasks and classroom resources that feed into the learning process.

A number of students at some universities who are taking mathematics courses, especially statistical units, find it difficult to understand some derivations of certain equations, theorems and proofs. According to Nardi (2011), the dependency on unprecedented rigour and proofs in mathematics, particularly at university level, appears to be the reason why mathematics students tend to lose the ability to think for themselves. This is because the mathematics taught consists largely of theorems and their proofs, which students tend to memorize with little understanding.

According to Alexander and Fuller (2005), when all other things are equal, highly qualified teachers produce greater student achievement than comparatively less qualified teachers. Alexander and Fuller (2005) argue that subject matter is necessary for achievement but not sufficient in developing effective teachers because in addition to subject matter, teachers need to know how to teach and how to explain the content.

Nardi (2011) argues that the system of lecturing mathematics, prevalent at most universities encourages students to cope by memorizing what they perceive as a fixed body of knowledge rather than learning to think for themselves. Students normally memorize the proof without necessarily understanding it because all that most of them are focused on, is to pass examinations regardless of whether they have understood the content taught or not. They simply memorize the proofs to pass examinations.

Mugisha (2012) asserts that lecturers ought to help students make connections between different topics to enable them see the bigger picture of the content taught and learnt. Multiple perspectives are important. Fennema and Franke (1992) acknowledge that an awareness of how students acquire the knowledge of mathematics content being addressed, as well as understanding the processes the students use and the difficulties and successes likely to occur, forms part of the teacher's knowledge of student learning. Teachers play an integral part in students' learning and understanding of any given mathematical content.

Teaching methods employed in teaching mathematics affect students' learning of content. Ng'ang'a (2010, p. 32) asserts that,

“Some of the main causes of poor performance in mathematics can be the use of inappropriate teaching methods.”

In universities, lecturers use a lecturer method as the only method of content delivery to students. Very few lecturers embrace the ICT integration in teaching mathematical units.

According to Batanero et al., (2011), some teachers are willing to learn about the content of statistics and spend more time teaching statistics by acknowledging the practical importance of statistics but these teachers feel that their students experience greater difficulties in statistics than in any other mathematical topic.

According to Bunker (1969), university students are taught to consider the context of problems in order to see relationships between elements; to comprehend; “what is given” and “what is wanted” in order to move on to the next step of solving a mathematical problem correctly. Bunker (1969) argue that being able to see and understand relationships is a key to successful problem-solving in mathematics. To solve a mathematical problem correctly, the student needs to understand which topic or content the question has been set from, what has been given in the question and what is required for its solution. This will enable the student to tackle the question correctly from one step to the next with ease while solving the question.

According to Mugisha (2012), lecturers should encourage their students to move back and forth between algebra and geometry. For instance, in solving a system of linear equations, students should realize that they are finding a point where two lines cross. When given a quadratic function, they should recognize and visualize a parabola which it defines and so be

able to draw or sketch it; in other words, the equation and the graph should be seen as different representations of the same mathematical object.

Learning of mathematics may also be difficult due to the approach given its teaching by teachers. Ng'ang'a (2010, p. 31) maintains that,

“most mathematics teachers use the conventional teacher-centred approach such as lecture method” in teaching.

The approach is inappropriate in teaching mathematics since the students do not write all the lesson notes correctly and therefore revision becomes very difficult. Also, mathematical equations (e.g. statistical equations) are long to lecture and also expect students to write the correct equation and take notes. Furthermore, solving a mathematical question requires a step-by-step solution and therefore a student-centred approach should be employed in solving of those questions.

The teachers' delivery of the content and the student understanding of the content taught influence student performance in mathematics. Most of the mathematics taught seems complex to most students at first sight but the content becomes easy when students revise worked examples on the content and also perform more calculations on other practice questions related to the examples done.

According to Yusof and Tall (1999), the traditional methods of teaching mathematics at university, which are intended to inculcate rigorous standards of mathematical proof, often seem to lead students to a 'deficit model' of rote-learning material only to excel in their examinations. Some mathematical proofs, especially statistical proofs, are long for students. When the same proof is examined in an examination, students normally memorize the derivation without necessarily understanding the proof since what the students require is to pass examinations regardless of whether they understand the concepts or not. This is a major hindrance to learning because the students do not foresee a situation where the proof will have to be applied after completion of their studies.

At most universities, examinations are set from the units and content covered in that semester. Once the semester units are complete, there are no more tests on the content in subsequent semesters, and students do not see any reason to continue revising or reading the content.

They only use the learnt knowledge as a foundation for advanced learning in similar units while at university.

According to Tishkovskaya and Lancaster (2012), the traditional lecture-based teaching methods need to be replaced or supplemented by new approaches which reframe the roles and identities of lecturers and students. Students can also improve in learning statistics when lecturers improve the instructional techniques they use in statistics courses.

2.2.2.3 Evaluation methods and procedures

Examinations and continuous assessment tests and the procedures of evaluating students can negatively affect the students' performance in mathematics and therefore might make the learning of some mathematical content difficult. Teachers, lecturers and tutors need to know the stages at which students must be evaluated apart from the normal class lessons. Evaluation can take place at the end of each unit or topic, mid-way through the semester or at the end of the semester on the work already covered.

Ogoli (1998) argues that it is necessary to assess students at the end of each topic in order to evaluate their achievement and understanding. Ogoli (1998) also goes on to assert that, mathematics teachers should mark students' work and offer guidance to ensure students understand the content covered. For basic education, teachers are often required to assess, check and mark students' work as well as guide and help them to thoroughly understand the content. This process would strengthen the students' ability to solve mathematics questions and form a firm foundational knowledge for further learning of mathematics in higher learning institutions.

According to the findings of Wilkens (1975), for the effective instructional evaluation of students' learning, the teacher should mark students' work by being attentive to the procedures followed by individual students in problem-solving.

2.2.2.4 Inadequate teaching and learning resources

Inadequate resources, such as textbooks, may also lead to students experiencing difficulties in learning mathematics. Students require adequate reference materials so that they can use them for practice purposes. Students need not depend only on teachers' notes or examples to learn.

Inadequate teaching and learning resources, for example, textbooks and other revision materials in a learning institution, according to Mbugua et al., (2012), Ogbonnaya and Mji, (2013), can impact negatively on teaching and students' learning.

Mbugua et al., (2012) show that the availability of adequate teaching and learning resources contributes positively to students academic performance. Studies (e.g. Ogbonnaya & Mji, (2013)), have shown that non-availability of adequate teaching and learning resources can hamper teachers' motivation to teach and students' interest to learn. It can make an educational institution become non-conducive to learning.

Most schools do not have libraries, and if there is one it has very few study or reference books. Studies (e.g. Eshiwani (1983)) have noted that the availability of adequate physical facilities such as libraries, essential equipment and materials is important for the effective learning of mathematics.

Eshiwani (1983, p. 16) also noted that-

“The factors that affect performance in Mathematics among secondary school students fall in to class size, streaming effect and school facilities. The class size contributes significantly to poor management and control of the class.”

According to Ogoli (1998, p. 33),

“Most students share mathematics textbooks in groups and they have limited practice, making them less competent and lacking self-confidence in problem solving.” This makes future learning of mathematics difficult.

For effective and meaningful learning to take place, it is paramount that an institution is equipped with adequate teaching and learning resources, for example, a library with sufficient reference books and classrooms and adequate teaching staff among other factors.

Ogoli (1998), maintains that, a poor background in mathematics coupled with a string of formulae, complex procedures and difficult topics, hinder students' learning in mathematics thus making mathematics difficult. Some content in mathematics is complex for students and therefore they need to have adequate foundational knowledge in order to understand the content for effective and meaningful learning to take place.

2.2.2.5 Students' attitudes towards some mathematics content

The negative attitude of some students towards some of the content in mathematics makes learning of the subject difficult for them. The negative attitude towards the study of mathematics may start as early as in the secondary (high) school level of education. According to Tishkovskaya and Lancaster (2012. p. 4), "over the years there has been strong anecdotal evidence that students at university develop antipathy towards statistics and, typically, students at all levels lack interest in learning when taking introductory statistics courses."

Anthony and Walshaw (2009) acknowledge that the positive attitude that develops in students raises the students' comfort level, enlarges their knowledge base, and gives them greater confidence in their capacity to learn and make sense of mathematics. Studies, for example by Ogbonnaya and Mji (2013), show that a negative attitude towards mathematics stems from the students' experiences in mathematics classes over time and that in most classes, mathematics is presented as an abstract subject without connection to the real world and this makes students not see the use of mathematics in real life. Consequently, they become disillusioned with mathematics learning.

The students' attitude, perception and fear regarding some long and complex equations discourage some of them and this makes the content difficult to comprehend. Shojaie, Aminghafari and Mahammadpour (2012) assert that the introductory courses in probability and statistics, which include joint distribution and calculation of bivariate expectation, and the difficulty of calculating complex integrals, are observed for most students. Most students find the calculation of the joint distribution, the expectation of a variable given another variable (the conditional mean and variance) difficult. The students also find the calculation of complex integrals difficult; more so the integrals involved in the analysis of a bivariate normal distribution equation.

Das (1985, p. 4) maintains that;

Some “students with a negative attitude towards mathematics eventually join the teaching profession, carrying with them the same attitude and end up consciously or unconsciously making students believe that mathematics is a difficult subject and it is for the gifted few.”

A negative attitude and wrong perception of some content in mathematics negatively affects a student’s learning and understanding of the content. This therefore implies that there exists no meaningful and effective learning of the content.

A positive attitude towards learning in statistical courses is important after university since it may help students to secure a job in the field of statistics. Therefore, the students should understand the content taught and sufficiently apply the knowledge gained in their future work stations. Tishkovskaya and Lancaster (2012), argue that, the most critical challenges students face in the study of statistics is the fact that the statistical courses affect life-long perceptions of many students and the students’ attitudes towards the value of statistics, and hence of many future employees, employers and citizens. Statistics learnt should be made use of by employees and employers in solving day-to-day problems. But most students do not apply the statistical knowledge acquired in learning institutions to solve day-to-day problems in the work place.

McPhan and Pegg (2008) noted that students normally seem to find mathematics more difficult and they perceive that they are not achieving the results they need to go on to higher mathematical study. Students who score poorly in lower learning grades, find it difficult to pursue mathematics courses in higher learning institutions. These students find it hard to continue learning mathematics beyond a certain level where mathematics is no longer compulsory.

Ogoli (1998, p. 43) indicated that-

“students have negative attitudes towards mathematics and have no aspirations to pursue it beyond secondary school level.”

While Ng’ang’a (2010, p. 31) also noted that-

“students had a negative attitude towards mathematics and there was no link between primary and secondary school mathematics.”

This also means that it is highly probable that the same link between secondary school mathematics and mathematics at higher learning institutions, such as university, may fail to exist because of the students’ negative attitude. Moreover, some students with a negative attitude towards mathematics may apply for courses at the university and be admitted for such courses ignorant of the fact that the courses they have applied for have compulsory mathematics units. The students will then carry the same negative attitude towards mathematics to the higher learning institution.

The negative attitude towards mathematics may make some students not take courses related to mathematics after a certain level of education since they perceive mathematics to be a difficult subject. According to the findings of the United States National Research Council (1989), many people’s school memories were unpleasant because the last mathematics course they did convinced them to take no more.

Some students learn mathematics because it has been made compulsory up to a certain level of education. But once the students have completed that level of education, they do not wish to take any more mathematics courses in advanced and higher level learning institutions because they lack interest in mathematics and perceive it to be a difficult subject.

2.2.2.6 Students’ lack of interest and inadequate effort in learning mathematics

Lack of interest in learning mathematical content, negatively impacts on the students’ learning. Students’ negative interest and wrong perception of the bivariate normal distribution content affect student learning of this content.

Students who are challenged by mathematics generally do not work hard and are not committed; hence they underachieve in their studies. Students with little or no interest in mathematics carry this attitude with them to higher learning institutions (e.g. university) where they may influence other students to develop the same attitude.

Effective and meaningful learning in statistics is essential to enable students to have an interest in learning and understanding the basic introductory courses, parameters and concepts. Learning is effective when students learn content from the known to the unknown.

Rosenstein (2005) noted that most students do not think that it is proper to seek assistance and that the students have not learned that if they are having difficulties in a course, they should seek help as soon as possible.

Bell et al. (1983) note that liking and having an interest in mathematics may lead to higher achievement and in turn that can lead to the willingness to pursue mathematics in subsequent studies. To pursue mathematics in higher and advanced levels of education requires that the student likes the subject and has interest in learning it. The positive interest incorporated with liking mathematics enables the student to learn and understand the content automatically.

For effective learning in mathematics, everyday problem-solving is most essential. Davis (1984) proposes that problem-solving in mathematics gives one the following: the ability; to learn and apply rules; to understand that a problem can be solved in more than one way; to learn to see relationships among different entities; and, to recognize that specific operations are needed in solving a problem.

Ogoli (1998) note that students do not understand most mathematical concepts and that students lack enough practice and revision, hence they are not efficient and confident in computational skills and also that students lack a variety and enough textbooks for practice and revision. Ogoli (1998, p. 47) asserts that-

“students are not exposed to extensive problem solving processes and exercises and that formed the basis for their poor retention, accuracy building and development of confidence in problem solving.”

This study has been conceptualized around the assumption that the difficulties students experience in the learning of a BND emanate from the their inadequate foundational background in mathematics, students’ negative attitude and lack of interest in learning mathematics and the inadequate teaching of some mathematical content. It is worth noting that sufficient content knowledge on the part of teachers leads to students’ increased opportunity to learn. Additionally, students tend to perform better if they are well taught by their teachers. In the current study, the students’ poor performance in statistics, in particular,

the BND and the difficulties they encounter in learning BND need to be addressed. It is against this background that this study was pursued.

2.3 CONCLUSION

This chapter has discussed the theoretical framework and literature related to the research study. The chapter also discussed the teaching and learning of BND content. The researcher acknowledges the limited sources of literature directly connected to the difficulties students experience in the learning of bivariate normal distribution. Therefore the study discussed the difficulties students encounter in learning mathematics and statistics in particular cases because of limited literature on the BND. This chapter discussed the general factors which contribute to students' challenges in learning mathematics, for example, the students' poor background in mathematics, lecturers' methods of content delivery and teaching methods, lecturers' evaluation methods and procedures, availability of the teaching and learning resources, students' attitudes towards some mathematical content, and students' interest, inadequate effort and lack of hard work in mathematics. The literature review focused on the challenges students experience in learning mathematics.

The next chapter discusses the research design, sources of data, population, sample and sampling methods, the research instruments and procedures for data collection and analysis. Also discussed in the chapter 3 is the validity and reliability of the research instruments, the pilot study, data analysis, the limitations of the study and ethical considerations of the study.

CHAPTER 3

RESEARCH METHODOLOGY

This chapter describes the methodology followed in addressing the research questions put forward to seek possible solutions to the difficulties undergraduate students experience in the learning of BND as identified in chapter one. This study employed an explanatory mixed methods design as the research aimed to investigate the difficulties students experience in learning a BND in a Kenyan university.

Therefore, in this chapter, a discussion on the research design of the study, population, sample and sampling techniques, is presented. This chapter also discusses the research instruments, the pilot study, and the procedures for data collection and analysis. This chapter also discusses, data analysis, validity and reliability of the instruments, limitations of the study and ethical considerations.

This study adopted a mixed methods design to collect quantitative and qualitative data.

3.1 RESEARCH DESIGN

The research design of a study is basically the overall approach used to investigate the problem of interest, that is, what is important to shed light on, or answer the question of interest (Gay, 1996).

The research design is a plan which guides the researcher in collecting, analyzing and interpreting the participants' views about the topic under study. And according to Gay (1996), the research design includes the methods of data collection and related specific strategies.

A mixed methods design was adopted for this study where both the quantitative and qualitative data was collected. An explanatory sequential mixed methods design (Creswell & Plano, 2011) consists of first collecting quantitative data and then collecting qualitative data to help elaborate and explain the quantitative results. The quantitative research methodologies focus on data that are mainly numerical and the qualitative research methodologies deal with data that are principally verbal and textual (White, 2005). The researcher's aim in this study

was to investigate the difficulties students encounter in the learning of bivariate normal distribution. Therefore the research investigated the students' perception on how lecturers teach the content. Also, the study looked into how the students learn the bivariate normal distribution content in the Kenyan university. According to White (2005), qualitative research is largely concerned with understanding social phenomena from the perspectives of the participants.

The use of both quantitative and qualitative methods, in combination, provides a better understanding of the research problem and question than either method by itself (Creswell, 2012).

The rationale for this design is that the quantitative data and the results provide a general picture of the research problem; more analysis, specifically through qualitative data collection, which is needed to refine, extend or explain the general picture (Creswell, 2012). The quantitative data, for example, scores, yields specific numbers that can be statistically analyzed to produce results to assess the frequency and magnitude of trends, and can provide useful information, if needed, to describe trends about a large number of people. The qualitative data, for example, an open-ended questionnaire, provides the actual words of respondents in the study, offers many different perspectives on the study topic and provides a complex picture of the situation. For that reason, both quantitative and qualitative research methodologies were appropriate for this study.

To answer the research questions for this study, the research required data that allowed assessment of the extent to which the bivariate normal distribution test (BNDT) for students influences the learning of a BND. This study mainly relied on scores of students on the BNDDT examination for quantitative data collection. The student questionnaire and the lecturer questionnaire also provided qualitative data for this study.

3.2 SOURCES OF DATA

There are two main categories of data sources, namely; primary and secondary sources (Tailor-Powell & Steele, 1996). Primary sources of data involve data collected by the researcher himself/herself. The common methods of primary data collection include tests or

examinations, questionnaires, observations and interviews. Secondary sources of data collection are data sources which have already been collected and analyzed by someone else.

This study relied on BNDT for students, the student questionnaire and the lecturer questionnaire as the primary sources of data collection to get first-hand information on the difficulties students encounter in the learning of bivariate normal distribution.

3.3 POPULATION

A population can be defined as a group of individuals that is the main focus of a research study and to whom the research results can be generalized (Joan, 2009). Population is categorized into the target population and accessible population (Joan, 2009).

McMillan and Schumacher (2001) describe a population as a group of elements or cases, whether individuals, objects or events, that conform to specific criteria and to which researchers intend to generalize the results of their research. All individuals within a certain population normally have common and binding characteristics or traits. Gay (1990) defines a population as the group of interest to the researcher, the group to which the researcher would like the results of the study to be generalized.

This study was conducted in one Kenyan university and its campuses and the targeted group included lecturers teaching statistics in that university and students enrolled or previously enrolled in statistics courses. The targeted student population was, in particular, those who had previously studied bivariate normal distribution content. Second-year students in the department of mathematics were selected for this study since they had completed the study of BND content in the unit probability and statistics (III).

3.4 SAMPLE AND SAMPLING METHOD

White (2005) defines a sample as a group of subjects or situations selected from a larger population. According to Lunsford and Lunsford (1995), a sample is a small subset of the population that has been chosen to be studied while Bless and Achola (1988) define a sample as the sub-set of the whole population which is investigated by a researcher and whose characteristics will be generalized to the entire population.

In total, 50 students participated in the pilot study and 175 students out of approximately 482 (as per first year enrollment in the courses) participated in the main study. However, the number decreases in second and third years since some change their courses (inter-faculty transfer). The students who participated were in their second year and third year of study in the university. The students were enrolled in the following courses offered in the university: BSc Mathematics, BSc Physical Sciences, BSc Mathematics and Computer Science and BSc Statistics.

According to Gay (1992), sampling is a process of selecting a number of individuals for a study in such a way that the individuals represent the larger group called the population from which it is selected. A sample determines the generalizability of the results for the population.

Simple random sampling techniques were used in selecting the students who participated in the study. In simple random sampling, each member of the population has an equal chance of being selected. Hence there is a high probability that all the population characteristics would be represented in the sample (McMillan & Schumacher, 2001).

A convenience sample is a group of elements that are readily accessible to, and therefore convenient for the researcher. According to McMillan and Schumacher (2001), in convenience sampling, a group of subjects is selected on the basis of being accessible. In this study, because there were relatively few lecturers in the department of statistics at the university, the lecturer open-ended questionnaire was administered to two lecturers during the pilot study and six out of approximately 14 lecturers, participated in the actual study. The total number of lecturers could not be ascertained since some of them were part time lecturers. Therefore, a convenient sample of the lecturers was used in this research study.

3.5 RESEARCH INSTRUMENTS

A data collection instrument is a survey, test, scale, rating or tool which is designed to measure the variables, characteristics or information of interest (Pierce, 2009). The data required for this study were gathered using the test and questionnaires. The instruments were developed by the researcher and moderated by statistics lecturers and language experts.

This section describes the data collection instruments for the study, namely; the BNDT, the student questionnaire (SQ) and the lecturer-open ended questionnaire (LOEQ), and how they were designed.

3.5.1 Bivariate normal distribution test (BNDT)

The BNDT (Appendix 1) was developed by the researcher in the form of a diagnostic test on the bivariate normal distribution content and knowledge of concepts foundational to the BND. The test was designed to elicit the students' knowledge of bivariate normal distribution and in the process gain insight into their difficulties, errors and misconceptions on BND content.

The questions in the test were all set in conformity with the curriculum and the course outline, that is: the content covered on BND and other related content on the normal distribution function, differentiation and integration.

The test consisted of eight questions which were based on BND and foundational and related concepts to the bivariate normal distribution, such as the normal distribution function and its application, differentiation and integration of exponential equations, and the conditional mean and variance of the joint distribution of two random variables. A total of 30 marks was allocated for the test and the time duration for the test was one hour. Each question was allocated marks depending on its nature or complexity and the number of steps the students needed to reach the final solution (see Appendix 9).

The test was organized in conformity with the content already covered by the students in the university. The knowledge of concepts foundational to BND, on the normal distribution and its application formed the first three questions of the test; questions four and five were based on differentiation and integration application, Question six was based on the bivariate normal distribution equation and questions seven was an application question on the BND while question eight was a comparison question on the BND function.

The students' solutions on the BND questions, coupled with questions and solutions on the normal distribution, exponential differentiation and integration, gave insight into the students' understanding of the BND and hence the difficulties students experience in learning the BND content.

The total score for each student reflected the performance of the students and hence the difficulties they encountered in learning the BND.

The scores from the test, for the students in each question, were analyzed in a statistical package for social sciences (SPSS) software to help further explain the difficulties students experience in learning the BND. The percentage scores and frequencies of marks scored by students in each question were presented.

The validity and reliability of the instruments are presented in the sections below, 3.7.1 and 3.7.2 respectively

3.5.2 Bivariate normal distribution student questionnaire

The questionnaire (Appendix 2) was developed to find out students views concerning the difficulties they experience in learning BND.

It consisted of two sections. The first section was anchored on demographic information of the students. The second section consisted of 27 likert type items and an open-ended question which sought the students' perception on the factors that contribute to the problems and difficulties they encounter in learning BND.

The questionnaire also sought to elicit the students' views on the teaching and learning of BND.

The students' responses from the questionnaire were analyzed SPSS software to help further elaborate and give more insight into the difficulties the students experience in the learning of BND.

3.5.3 Lecturer's questionnaire (LOEQ)

The lecturer questionnaire (Appendix 3) was developed to find out the lecturers' views concerning the teaching and learning of the bivariate normal distribution, the difficulties students encounter in learning BND and on the factors which contribute to the difficulties students encounter in learning BND. The questionnaire consisted of six open-ended questions. The six questions in the LOEQ were strictly in line with the teaching and learning of BND.

Lecturers were given ample time to answer the questionnaires since most of them were busy preparing for end of semester examinations. So the lecturers needed enough time to answer the questions in their own spare time.

The purpose of the open-ended questionnaire was to probe deeper into the difficulties students experience in learning of BND. The lecturers' responses gave the researcher further information on the teaching and learning of BND content in probability and statistics and hence the difficulties the students encounter in learning this content.

3.6 PROCEDURE FOR DATA COLLECTION

Data collection is the process of gathering and measuring information on variables of interest, in an established systematic fashion that enables one to answer stated research questions, test hypothesis, and evaluate outcomes (Murali et al., 2004).

The collection of data for this study started in April 2015 when the pilot study was conducted. The main study was conducted in March - April 2016.

The researcher explained the purpose of the study to the students and what was expected of them in terms of the research test and filling in of the questionnaire.

They were instructed to clearly show their working and how they had arrived at the answers in the BNDT. The students were also instructed not to write their names and university admission numbers on the instruments.

The researcher administered both the test and the questionnaires to the students and lecturers himself. Both the test and the student questionnaire were administered on the same day during the last few weeks of the semester and before the start of end-of-semester examinations after completing the syllabus.

On completion of the test and filling in of the questionnaire, the researcher collected the scripts and scored them using a marking guideline (see Appendix 9).

3.7 VALIDITY AND RELIABILITY OF THE RESEARCH INSTRUMENTS

3.7.1 Validity

Validity is defined as the extent to which an instrument measures what it purports to measure. According to Schumacher and McMillan (2006), validity refers to the degree of congruence between the explanations of the phenomena and the realities of the world. White (2005), argued that validity is the researcher's conclusion which corresponds to the actual state in reality. Validity requires that an instrument is reliable, however, an instrument can be reliable without being valid (Kimberlin & Winterstein, 2008). Reliability refers to the extent to which administration of the same instrument consistently yields similar results under comparable conditions (De Vos, 2002).

The test questions were all set from the statistics courses the students had completed, namely;- the foundational units related to the BND content and the content on the BND. For a test to be valid, it must be aligned to the content the students are expected to learn (e.g. the normal distribution and its applications, the exponential differentiation and integration and the BND and its applications). All the questions followed the content (see Appendix 8).

After the test was constructed, it was sent to mathematics lecturers teaching statistics and a mathematics doctoral student, also teaching statistics in the university, in order to determine whether the test items were well constructed and within the course outline.

The content validity of the BND test was ascertained by two statistics lecturers teaching in one of the Kenyan university. These lecturers determined the degree to which the BND test measured the intended content. The lecturers noted some changes which they recommended to be effected, for example, correcting 'random variable x' in question one to read, 'random variable X' and in question eight, the equation should read as follows:

$$'f(x, y) = \frac{1}{2\pi\delta_1\delta_2\sqrt{1-\rho^2}} e^{-\frac{1}{1.02}\{(x+2)^2-2.8(x+2)(y-1)+4(y-1)^2\}}, \text{ but not}$$

$'f(x, y) = \frac{1}{2\pi\delta_1\delta_2\sqrt{1-\rho^2}} e^{-\frac{1}{102}\{(x+2)^2-2.8(x+2)(y-1)+4(y-1)^2\}},$ i.e. in the exponent, the denominator should read 1.02 and not 102.

In addition, the researcher's supervisor noted that the time allocation for the test should not be restricted to one hour as per the test instructions but rather the students should be allowed adequate time to complete the test.

The overall impressions for the validity measure, were that the BNDT covered all aspects necessary for the learning and teaching of BND and the content covered in the test was adequate. Thus the suggestions and corrections were implemented in the final BNDT instrument (Appendix 1).

Content validity of the student questionnaire was verified by the supervisor and a statistics lecturer in one of the Kenyan university. They gave the opinion that the questionnaire measured the intended content area of the study. The likert type items had a four point scale which adopted one of the following: (i) not at all, (ii) to a minor extent, (iii) to some extent and (iv) to a major extent; or (i) strongly agree, (ii) agree, (iii) disagree and (iv) strongly disagree. The student questionnaire was considered adequate to serve the intended purpose.

With regard to the open-ended questionnaire, it was initially referred to as lecturer interview guide questions. But with guidance from the supervisor, the interview guide questions were transformed to a lecturer open-ended questionnaire and the number of statements or questions, which totalled thirty, reduced to six. This is because some of the interview questions were combined and others were answering the same concept. The six questions were considered adequate to address its purpose.

3.7.2 Reliability

Reliability is another factor that affects the usefulness of a research instrument. The reliability of a measuring instrument reflects the consistency with which results can be obtained when it is administered repeatedly. The measure of consistency has a reliability coefficient which lies between 0 and 1.

The student questionnaire and the BNDT were piloted to a sample of 50 students from the research university. The aim of conducting the pilot study was to determine its feasibility and further elaborate the suitability of the techniques and instruments used for collecting data.

The reliability of the student questionnaire was determined by computing Cronbach's alpha as a measure of internal consistency of the students' scores from the questionnaires. The Cronbach's alpha (α) for the student questionnaire was found to be 0.715.

In general, if the reliability of a test is closer to 1.0, the test is said to have a very good reliability; but if the reliability of a test is below 0.50, the test would not be considered a very reliable test. This implies that the instrument was reliable (see Table 3.1). The questionnaire had 34 items, 7 items was on the biographic data of the students while 27 were likert scale questions on the students' views on the challenges they encounter in the learning of BND. In terms of the sub-sections of the students' questionnaire, the Cronbach's alpha (α) reliability coefficients were as in Table 3.1:

Table 3.1: Sub – section Reliability – Student questionnaire

Section	Number of Items	Cronbach's alpha (α)
Questions 1-7	6	0.745
Questions 8-20	13	0.875
Questions 21-34	14	0.768

For the bivariate normal distribution test, the internal consistency and reliability of the test was also computed by Cronbach's alpha (α) and a value of 0.826 was obtained as shown in Table 3.2. The 8 questions constituted the BND test with the questions attracting a score of between 2-6 marks (see Appendix 9). Therefore, the 8 BND questions were considered reliable.

Table 3.2: Overall Reliability – Students' bivariate normal distribution test

	Number of items	Cronbach's alpha (α)
Bivariate normal distribution test (Question 1-8)	10	0.826

The reliability of the research instruments was within the the required range and hence they were regarded as having good reliability for the instruments.

3.8 PILOT STUDY

Prior to the main study, a pilot study was conducted. A pilot study is essential to test the validity and reliability of the chosen instruments to be used in the actual research study and ascertain the feasibility of the study. It may reveal the deficiencies in the proposed study and hence they are addressed before the main study.

A pilot study helps to assess and detect the weaknesses in the study so that the weaknesses can be removed before the actual study (Creswell, 2012). The weaknesses which can emanate from questionnaires include: ambiguities in the phrasing of questions:- complexity in the language used; redundant questions etc. The suitability and relevancy of the test and the questionnaires were checked in the pilot study.

A pilot study was conducted to validate and to find the reliability of the research instruments used for the main study. Piloting was also carried out to test-run the whole study with the aim of identifying problem areas in the design to enable the researcher to make necessary amendments before the commencement of the main study.

Some of the weaknesses which could emanate from the test items include: incorrect or more than one correct answer or way of working (Creswell, 2012). The test items must be understood by the respondents and they should be at the appropriate level of complexity to measure the student ability and hence meaningful in terms of the student ability being examined (Creswell, 2012).

A total of 50 students participated in the pilot study. The students who participated in the study were from the department of mathematics, that is, students who were enrolled for the courses: BSc Statistics and Computer Science, BSc. Physical Science, BSc Statistics and BSc. Mathematics course. Two lecturers participated in the pilot study.

Two questions in the BNDT had some incomplete information due to a prior typing error. The errors were adjusted in the students' test before the actual study was conducted.

In question 8 of the test for students, there was a typing error in the bivariate random density function, as 102 instead of 1.02. But the typing mistake had absolutely no effect in the solution of the question as the values of the mean and standard deviation for the two independent random variables could be obtained. Before the main study was conducted, the error was corrected and the question was as follows:

Given that the bivariate random density function of y is:

$$f(x, y) = \frac{1}{2\pi\delta_1\delta_2\sqrt{1-\rho^2}} e^{-\frac{1}{1.02}\{(x+2)^2-2.8(x+2)(y-1)+4(y-1)^2\}}$$

Find the values of:

- a) μ_1 and μ_2
- b) δ_1 and δ_2

All the other questions were correctly framed and had no typing errors.

3.9 DATA ANALYSIS

Data collected from the test for this study was analyzed according to the learners' performance at various sections in a bivariate normal distribution test. Each question was analyzed independently to easily identify the students' difficulties in bivariate normal distribution. The scores collected from the test were analysed using Statistical Package for Social Sciences (SPSS). From the analysis, descriptive statistics were computed and the results were summarized in tables. The descriptive statistics e.g. the mean, the standard deviation, frequency distribution tables and simple percentage statistics on the student performance on the BNDT and the questionnaires, were used for analysis.

Data from the student and the lecturer questionnaires were analyzed using frequencies, percentages, mean and standard deviation to answer the research questions.

The qualitative data collected from the student questionnaire and lecturer open-ended questionnaire, was carefully studied, identified and collated with common features and themes from the respondents' descriptions of experience. That enabled the identification of common themes that emerged from the responses.

The qualitative analysis enabled the researcher to triangulate and explain the quantitative data collected and analyzed from the sample of the population. The results from the sample of students and lecturers were used to generalize for approximately 482 students and 14 lecturers. Approximately 14.86% of the lecturers and 36.31% of the students participated in the study.

3.10 ETHICAL CONSIDERATIONS ON THE STUDY

Ethical issues are the concerns, dilemmas and conflicts that arise over the proper way to conduct research (Neuman, 2011). McMillan and Schumacher (2001) state that-, research ethics are focused on what is morally proper and improper when engaged with participants or when accessing data. Under ethical issues, participants must voluntarily agree to participate and the researcher needs to safeguard against unwanted exposure and loss of anonymity. The researcher needs to fully reveal the procedures of research to the participants at the onset.

In this study, the researcher sought permission (Appendix 4) to conduct research in the research university from the vice-chancellor, academic affairs, the university lecturers in the department of mathematics, and second-year students in the department of mathematics who are registered for the probability and statistics (III) unit.

After being granted permission (Appendix 5) to conduct research, the researcher organized a date with the lecturers teaching statistics in the unit probability and statistics (III), on the day the students would be free in order to collect data. On the day of data collection, the students were assembled by their lecturer who was teaching them the probability and statistics (III) unit with the help of the class student representatives.

Before data collection commenced, the researcher explained to the students the aims of the study, the importance of the study, the need to abide by the procedures for administering an examination for the student test and the questionnaire, the need to sign the consent forms (Appendix 6) and their right to participate or withdraw from the study even after consenting to participate. The researcher informed the students that participation in the study was voluntary. Consent forms were issued to the lecturers and students to request their participation in the study. All participants were also informed that there would be no consequences if they decided to withdraw even after consenting to participate in the study.

The research was conducted in the last few weeks of the semester just before the start of the examinations at the end of the semester after the students had completed the syllabus. This was done to avoid inconveniencing the students' normal class learning and the lecturers' teaching schedules.

The researcher put in place measures to protect the identity of participants against exposure. The anonymity of the participants must be protected at all times during research (Cresswell,

2007). In this study, only the researcher accessed the individual students' data from the test results and the complete questionnaire thus ensuring confidentiality.

The researcher explained to the students that their identity and the information they provided would be treated with strict confidentiality and no information regarding them and their university would be disclosed to any individual or organization. The researcher achieved this by assigning numbers to all the instruments used for data collection. The students were advised against writing their names or admission numbers on the research instruments thus ensuring confidentiality of the students.

On informed consent, the students and lecturers were fully informed about the procedures involved in research and hence there was a need for students and lecturers to consent (Appendix 6 and Appendix 7) to participate. According to McMillan and Schumacher (2001), informed consent can be achieved by providing participants with an explanation of the research, an opportunity to terminate their participation at any time with no penalty and full disclosure of any risks associated with the study.

In this study no risk was envisaged and the ethical rights of the participants to make their decisions about participation in the study were observed.

3.11 LIMITATIONS OF THE STUDY

The number of lecturers who participated in the study were small and so they should have been interviewed to probe deeper into the difficulties students experience in the learning of BND.

It was difficult to assemble all students during data collection. Some of them did not show up to participate in research as much as they were made aware of the time, day and place to assemble for the study.

3.12 SUMMARY OF THE CHAPTER

This chapter discussed the research design; sources of data; population; sample and sampling methods; the research instruments; and procedure for data collection. Also discussed in this

chapter were the validity and reliability of the research instruments; the pilot study; data analysis; the limitations of the study; and ethical considerations of the study.

The next chapter discusses findings of the study on the difficulties students experience in the learning of bivariate normal distribution. The findings from each of the research instruments and their summary are presented in the chapter 4.

CHAPTER 4

FINDINGS

This chapter outlines the findings of the research conducted in the Kenyan university on the difficulties students experience in the learning of bivariate normal distribution content. A total of 175 students out of approximately 482 participated in the study. Six lecturers out of 14 also participated in the research study. The students were selected from a group of second and third year students who had studied the BND content in the probability and statistics III unit.

The students who participated in this study were enrolled in the following courses offered at the university: BSc Mathematics (71 students), BSc Physical Sciences (51 students), BSc Mathematics and Computer Science (31 students) and BSc Statistics (22 students).

The data collected was intended to answer the following research questions;

1. What do students find difficult in learning bivariate normal distribution?

To address this *research question*, the following guiding sub-questions were put forward:

- 1.1 Can the students state the equation of a normal distribution function?
- 1.2 Can the students obtain the mean given a normal distribution equation?
- 1.3 Can the students obtain the the standard deviation given a normal distribution equation?
- 1.4 Can the students differentiate a given exponential equation?
- 1.5 Can the students find the integral of a given exponential equation?
- 1.6 Can the students state the BND equation?
- 1.7 Given the conditional distribution of a variable (e.g. $h(y/X = x)$), can the students calculate the mean, $E(y/X = x)$ of the joint distribution?
- 1.8 Given the conditional distribution of a variable (e.g. $h(y/X = x)$), can the students calculate the variance, $Var(y/X = x)$ of the joint distribution?

2. Why do students experience difficulties in learning BND?

Three instruments were used for data collection. The instruments were; the bivariate normal distribution test (BNDT) for students, the student questionnaire (SQ) and the lecturers' open-ended questionnaire (LOEQ).

The participants

In this research study, a total of 175 students participated in the study. The students who participated in this study were enrolled in the following programmes offered in the university;

- | | |
|---|---|
| 1) BSc Mathematics | - 71 students participated in the study |
| 2) BSc Physical Sciences | - 51 students participated in the study |
| 3) BSc Mathematics and Computer Science | - 31 students participated in the study |
| 4) BSc Statistics. | - 22 students participated in the study |

The number of male and female students who participated in this study were 113 and 62 respectively.

While marking the students' scripts, the researcher observed the following:

- The performance between male and female students was similar.
- Second year students responded better to bivariate normal distribution questions as compared to third year students. This is mainly because it was the last content they had covered in the semester.
- Third year students performed well in normal distribution questions, differentiation and integration questions compared to second year students.
- BSc Mathematics and Computer Science group responded well to the BND test than other groups since they tried to solve all the questions. BSc Mathematics were expected to perform much better than other groups but they did not.
- Third year students were also expected to do better in the test but they did not. This may be attributed to the fact that they covered BND a year before in second year, so they had difficulties recalling the content.

4.1 THE BIVARIATE NORMAL DISTRIBUTION TEST (BNDT) RESULTS

The BND test consisted of eight questions and each question was allocated marks depending on the level of complexity and the number of steps the student wrote to reach the final step.

The following tables illustrate the distribution of students' scores in each question in the BNDT. The frequency of marks scored by the students in each of the BND test illustrates the level of difficulty they experienced in answering the questions.

Students' achievement in the test

The test had 30 marks. The percentage score of each student was then calculated and the frequency and percentages were tabulated as in Table 4.1a.

Table 4.1a: Frequency of students' scores in the BND test

Score	No. of students (%)
0-39	139 (79.4%)
40-49	22 (12.6%)
50-59	11 (6.3%)
60-69	3 (1.7%)
70-100	0 (0.0%)

The students performed poorly in the BNDT as illustrated by Table 4.1a. They scored a low mean mark of 29.3% out of the possible 100 marks.

The percentage scores in the Table 3.1 above represent the grading system used in the Kenyan university. Using the grading system for the BND test, no student scored grade A (70-100 marks). Only 3 (1.7%) students scored grade B (60-69 marks); 11 (6.3%) scored grade C (50-59 marks), 22 (12.6%) scored grade D (40-49 marks) and 139 (79.4%) scored grade F (fail) (0-39 marks). The grades scored in this test demonstrated the students' weaknesses in the learning of BND. It therefore implies that the students have difficulty in learning and understanding the bivariate normal distribution and they essentially lack adequate foundational knowledge on BND (see Table 4.1b).

Table 4.1b: Students' achievement as per marks scored on foundational knowledge on the BND

Question	Marks scored	Frequency	Percent
1. State an equation for a normal distribution function for a random variable X	0	97	55.4%
	1	0	0%
	2	78	44.6%
2. Given that, $X \sim N(1, 4)$, find $P(1 \leq X \leq 3)$	0	77	44.0%
	1	29	16.6%
	2	18	10.3%
	3	51	29.1%
3. Determine the setting μ so that only 1% of the cans of paint will be unacceptable	0	174	99.4%
	1	1	0.6%
	2	0	(0%)
	3	0	(0%)
	4	0	(0%)
	5	0	(0%)

Table 4.1b shows that 97 students scored 0 marks in question 1, 77 students scored 0 marks in question 2 and 174 students scored 0 marks in question 3. This means that the students had inadequate foundational knowledge required for learning of the BND.

Question 1: State an equation for a normal distribution function for a random variable X
(2 marks)

In this question, students were asked to state the equation of a normal distribution function for a random variable X. Table 4.3 shows the scores of the students in the question. Of the 175 sampled students, 97 (55.4%) scored zero, no student scored one mark and only 78 (44.6%) scored the maximum 2 marks.

Table 4.2: Distribution of students' scores in question 1 of the BND test

Mark	No. of students
0	97 (55.4%)
1	0 (0%)
2	78 (44.6%)

Of the 97 students who scored zero marks in this question, over half of them left the question blank. This suggests that they had no idea what to answer since they were allowed adequate time to solve the questions. Only 78 students out of the 175 were able to state the equation correctly.

One of the sampled students stated the equation for a normal distribution function for a random variable X as follows:

[1] State an equation for a normal distribution function for a random variable X (2mks)

$$f(x) = \frac{e^{-(x-\mu)^2 / 2\pi r^2}}{\sigma \sqrt{2\pi}}$$

$$f(x) = \frac{e^{-(x-\mu)^2} (2\pi r^2)}{\sigma \sqrt{2\pi}}$$

Figure 4.1a: Example of a student’s solution to question 1 of the BND test

The student introduced some parameters, e.g. r^2 which is not part of a normal distribution function equation for a variable X. The student could neither recall the equation nor remember all the parameters involved in the equation.

Another student stated the normal distribution function equation as follows:

[1] State an equation for a normal distribution function for a random variable X (2mks)

$$p = \frac{1}{\sigma} \phi\left(\frac{x-\mu}{\sigma}\right)$$

Figure 4.1b: Example of a student’s solution to question 1 of the BND test

The student also could not state the equation and from what the student stated, it is evident that some of the students could not recall the basic and foundational knowledge on the normal distribution required for the learning of a BND. In the equation the student stated, the exponent, pi etc. are missing and the equation stated does not appear as a normal distribution function for a random variable X since it does not contain all the parameters and constants of the equation.

Question 2: Given that, $X \sim N(1, 4)$, find $P(1 \leq X \leq 3)$ (3 marks)

In this question, the students were asked to find the probability of the normal distribution for a random variable X. Table 4.3 shows the scores of the students in the question. Of the 175 sampled students, 77 (44.0%) scored zero, 29 (16.6%) scored one mark, 18 (10.3%) scored 2 marks while only 51 (29.1%) scored the maximum 2 marks (Appendix 9).

This question was testing the foundational knowledge the students have on the normal distribution and whether the students could use the standard normal tables to correctly solve the question.

Table 4.3: Distribution of students' scores in question 2 of the BND test

Mark	No. of students
0	77 (44.0%)
1	29 (16.6%)
2	18 (10.3%)
3	51 (29.1%)

Of the 175 sampled students, 106 (60.6%) students were unable to solve the question correctly and had inadequate knowledge on the calculation of the standard deviation and the probability of a normal random variable.

In the solution of question 2 of the BND test, one of the students solved the question as follows;

[2] Given that, $X \sim (1, 4)$, find $P(1 \leq x \leq 3)$ (3mks)

$\mu = 1$
 $\text{var} = 4$
 $\text{std} = \sqrt{4} = 2$

$\text{Formula} = \frac{(X - 26)}{\text{std}}$
 $= \frac{(1 - 1)}{2} = 0$

A1
 B0
 A0

Figure 4.3a: Example of a student's solution to question 2 of the BND test

From the solution by the student, it is evident that the student had inadequate knowledge on the normal distribution a more reason why the student could not find the probability of the normal distribution for a random variable X. The student only correctly found the square root of variance to obtain the standard deviation but could not move to the next step, that is,

$P(0 \leq Z \leq 1)$, before finally using the standard normal tables to determine the probability of a random variable X as 0.3413.

Another student solved question 2 as follows;

[2] Given that, $X \sim (1, 4)$, find $P(1 \leq x \leq 3)$

(3mks)

$\mu = 1$
 $\text{var} = 4$
 $\text{s.d.} = \sqrt{4} = 2$
 $\left(\frac{x-1}{\text{s.d.}} \right)$

$\frac{1}{\sqrt{2}} X$
 $-\frac{0}{\sqrt{2}} X^0$

A1, B0

Figure 4.3b: Example of a student's solution to question 2 of the BND test

The student made the same mistake in solving the question. The student correctly calculated the standard deviation = 2, but could not find all the Z values to enable him to use the standard normal table to finalize the solution to the question. The student should have obtained $P(0 \leq Z \leq 1)$, and used the standard normal table provided to show that, $P(0 \leq Z \leq 1) = 0.3413$.

Another student solved the question as follows:

[2] Given that, $X \sim (1, 4)$, find $P(1 \leq x \leq 3)$

(3mks)

$f(x) = \frac{1}{\sqrt{2\pi} \cdot 0.5} \cdot e^{-\frac{1}{2} \cdot 0.5 \cdot (x-1)^2}$

$= \frac{1}{4\sqrt{2\pi}} \int e^{-\frac{1}{2} \cdot 0.5 \cdot (x-1)^2}$

$= \frac{1}{4\sqrt{2\pi}} \left[2(x-1)^2 \cdot 0.5 e^{-\frac{1}{2} \cdot (x-1) \cdot 0.5} \right]^3$

M0, B0, A0

Figure 4.3c: Example of a student's solution to question 2 of the BND test

The student did not find the probability of the random variable. The student substituted the values of the variance and the mean in the normal distribution function equation which is a wrong working of the question.

Question 3:

In regulating the blue dye for mixing paint, a machine can be set to discharge an average of μ cm/can of paint. The amount discharged $N(\mu, 0.4cm)$. If more than 6cm is discharged into the paint can, the shade of blue dye, is unacceptable. Determine the setting μ so that only 1% of the cans of paint will be unacceptable. (5 marks)

This question asked the students to find the mean using the standard normal distribution knowledge and the normal distribution. Of the 175 sampled students, 174 (99.4%) students scored zero, 1 (0.6%) student scored 1 mark and no student scored either 2, 3, 4 or 5 marks. None of the students solved the question correctly. In fact, all the students left the question blank apart from two students who attempted the question but incorrectly. The students found it difficult to correctly interpret the question, therefore it was hard to solve.

The probability of the normal distribution is 1% (0.01), the standard deviation = 0.4cm and the value of the random variable greater than 6.

This means that the students had inadequate basic knowledge on the normal distribution which is necessary to solve the question and which is also required for effective learning of a bivariate normal distribution.

Not one of the 175 students was able to solve the question correctly.

Table 4.4: Distribution of students' scores on question 3 of the BND test

Mark	No. of students
0	174 (99.4%)
1	1 (0.6%)
2	0 (0%)
3	0 (0%)
4	0 (0%)
5	0 (0%)

One student who solved question 3 of the research test in finding the mean, tried to solve the question as follows:

[3] In regulating the blue dye for mixing paint, a machine can be set to discharge an average of μ cm/can of paint. The amount discharged $N(\mu, 0.4\text{cm})$. If more than 6cm is discharged in to paint can, shade of blue is unacceptable. Determine the setting μ so that only 1% of the cans of paint will be unacceptable. (5mks)

$N(\mu, 0.4)$
 $n = 6$
 $P = 0.01$
 $Z = \frac{x - \mu}{\sigma}$
 $P(X \geq 6)$
 $x = 2.5$

$t_{0.01} = \frac{5 - \mu}{0.2}$
 $0.0040 = \frac{5 - \mu}{0.2}$
 $0.2(0.0040) = 5 - \mu$
 $\mu = 4.9992$

Page 2 of 4

Figure 4.3: Example of a student’s solution to question 3 of the BND test

This was a question on the mean and the standard deviation of a normal distribution. The student substituted a wrong value, 5 instead of 6. The student again used a standard deviation value of 0.2 instead of using 0.4 cm. The wrong values substituted by the student resulted in the wrong solution of the question. But the student had inadequate knowledge on the procedure of solving the question.

Question 4: Differentiate the equation $y = \frac{1}{8}e^{-4x}$ (3marks)

This question tested the students’ ability to differentiate an exponential function. Table 4.5 shows the frequency of the student scores on the question. Of the 175 sampled students, 21 (12.0%) scored zero, no student scored 1 mark, 4 (2.3%) scored 2 marks and 150 (85.7%) scored the maximum 3 marks.

The students had adequate knowledge of differentiating an exponential function, which is the foundational knowledge applied in the analysis of a bivariate normal distribution equation to obtain equations for the conditional mean and variance of a variable given another variable.

Table 4.5: Results of students' performance on question 4 of the BND test

Mark	No. of students
0	21 (12.0%)
1	0 (0%)
2	4 (2.3%)
3	150 (85.7%)

Few of the students (12%) were unable to solve the exponential differentiation correctly. Figure 4.4 shows the working of one of the students in an effort to solve the differentiation question.

[4] Differentiate the equation $y = \frac{1}{8}e^{-4x}$ (3mks)

$\frac{1}{8} \frac{d}{dx} e^{-4x}$
 let $-4x = u$
 $-4dx = du$
 $dx = \frac{du}{-4}$
 $\frac{d}{du} e^u = e^u$
 $= \left(\frac{-1}{32}\right) e^{-4x}$

Figure 4.4: Example of a student's solution to question 4 of the BND test

Question 5: Find the integral of $\int 3xe^{2x^2} dx$ (4 marks)

In this question, students were required to solve an exponential integral. Table 4.6, shows the scores of the students on the question. Of the 175 students sampled, 74 (42.3%) scored zero, 15 (8.6%) scored 1 mark, 5 (2.9%) scored 3 marks and 81 (46.3%) scored the maximum 4 marks.

Exponential integration skills and knowledge are necessary for the analysis of the bivariate normal distribution equation for the development of the conditional mean and variance equations.

Table 4.6: Results of students' performance on question 5 of the BND test

Mark	No. of students
0	74 (42.3%)
1	15 (8.6%)
2	0 (0%)
3	5 (2.9%)
4	81 (46.3%)

An example of a wrong solution of the question by one of the students who participated in the study is shown in Figure 4.5a:

[5] Find the integral of $\int 3xe^{2x^2} dx$ (4mks)

$u = 3x$
 $\int u dv = uv - \int v du$
 $u = 3x$
 $du = 3$
 $dv = e^{2x^2}$
 $v = \frac{e^{2x^2}}{4x}$
 $\int u dv = uv - \int v du$
 $= \frac{3x e^{2x^2}}{4x} - \int \frac{e^{2x^2}}{4x} \cdot 3$

Figure 4.5a: Example of a student's solution to question 5 of the BND test

The student used integration by parts to solve of the question which is a wrong approach of solving the question. Another student used a similar approach of integration by parts to solve the question thereby, failing to correctly find the solution.

[5] Find the integral of $\int 3xe^{2x^2} dx$

let $u = 3x$ $dv = e^{2x^2}$
 $\int u dv = uv - \int v du$
 $du = 3x^2$ $v = e^{2x^2}$
 $\int 3xe^{2x^2} = 3xe^{2x^2} - \int e^{2x^2} 3x^2$
 $= 3xe^{2x^2} - \int 3x^2 e^{2x^2}$
 $= 3xe^{2x^2} - 3 \int x^2 e^{2x^2}$
 $= 3xe^{2x^2} - 3 \left(\frac{x^3 e^{2x^2}}{3} + c \right)$
 $= 3xe^{2x^2} - x^3 e^{2x^2} + c$

Figure 4.5b: Example of a student's solution to question 5 of the BND test

The student also used integration by parts to solve the question, which is the wrong approach to solving the question.

Question 6: State a bivariate normal distribution equation for two random variables X and Y. (3 marks)

In this question, students were asked to state the equation of a BND for two random variables X and Y. Table 4.6 shows the scores of the students on the question. Of the 175 sampled students, 102 (58.3%) scored zero, and only 73 (41.7%) scored the maximum 3 marks. The majority of the 102 students who scored zero mark, left the question blank while the others stated the wrong or incomplete BND equation.

In this question, the researcher was testing the students' cognitive level on the BND equation. The equation seems long and complex for some students and memorizing it is very difficult for some of them since it has many parameters and constants.

Table 4.7 represents the percentage representation of the sampled students.

Table 4.7: Results of students' performance on question 6 of the BND test

Mark	No. of students
0	102 (58.3%)
1	0 (0%)
2	0 (0%)
3	73 (41.7%)

Figure 4.6 demonstrates the students' inability in stating the BND equation. The student could not write down the parameters and constants contained in the equation.

[6] State a bivariate normal distribution equation for two random variables X and Y.

if X and Y are random variable then (3mks)

$$P(X, Y) = P\{X=x, Y=y\}$$

7

Figure 4.6: Example of a student's solution to question 6 of the BND test

The BND function has five parameters (the expectation of the two independent random variables; the standard deviations of the two independent random variables; and the correlation coefficient of the random variables) and two constants (ρ and the exponent).

Question 7

In a certain population of married couples, the height x ft of husband and the height y ft of the wife have a bivariate normal distribution with parameters $\mu_1=5.8$ ft, $\mu_2=5.3$ ft, $\delta_1= \delta_2=0.2$ inches, and $\rho = 0.6$. Find the conditional p.d.f of y given that x is 6.3 ft. What is the probability that the height of the wife lies between 5.28 ft and 5.92 ft given that the height of the husband is 6.3 ft? (6marks)

In this question, the students' ability to calculate and use the conditional mean and variance hence the standard deviation of a variable given another variable of a joint distribution was tested. Table 4.7 shows the scores of the students on the question. Of the 175 sampled students, 171 (97.7%) students scored zero, 2 (1.1%) students scored 1 mark, 1 (0.6%) student scored 5 marks and only 1 (0.6%) student scored 6 marks.

This implies that the students neither had adequate knowledge on the solution to the question nor were they able to use the tables to find the conditional mean and the standard deviation.

Table 4.8: Students' performance on question 7 of the BND test

Mark	No. of students
0	171 (97.7%)
1	2 (1.1%)
2	0 (0%)
3	0 (0%)
4	0 (0%)
5	1 (0.6%)
6	1 (0.6%)

The students could not apply the conditional mean and variance equations in the solution of question seven of the research test. It demonstrates the students' failure to use the conditional mean and variance equations to find the mean and standard deviation.

The working below shows one of the students' solution to the question.

$$\rho = \frac{\text{cov}(x,y)}{\sigma_1 \sigma_2}$$

$$E(y) = ax + b$$

$$a = \frac{\text{cov}(x,y)}{\text{var}(x)}$$

Figure 4.7a: Example of a student's solution to question 7 of the BND test

The students' working in solving the question was incorrect. The student used the correlation coefficient formula to solve the question instead of using the conditional mean and variance equations. The student tried to calculate the expectation of a variable which was already given in the question as the mean. If the conditional mean and variance equations were supplied to the students for use, some of them could have solved the question correctly.

Another student incompletely solved the question as follows:

$$E(y|x) = \mu_2 - 2\rho \frac{\sigma_2}{\sigma_1} (x - \mu_1)$$

$$5.3 - 2(0.6) \left(\frac{0.2}{0.1} \right) (6.3 - 5.8)$$

$$5.3 - 1.2(0.5) = 4.1$$

μ_1
 $B_0 \frac{6.3 - 5.8}{5}$
 A_0

Figure 4.7b: An example of a student's solution of question 7 of the BND test

The student used the correct formula to solve the question, but the final result was wrong.

Question 8 (a):

Given that the bivariate random density function of y is:

$$f(x, y) = \frac{1}{2\pi\delta_1\delta_2\sqrt{1-\rho^2}} e^{-\frac{1}{1.02}\{(x+2)^2 - 2.8(x+2)(y-1) + 4(y-1)^2\}}$$

Find the values of μ_1 and μ_2

(2 marks)

In this question, students were required to obtain the values of μ_1 and μ_2 from the bivariate random density function equation given. Table 4.9 below, shows the scores of students on the question. Of the 175 sampled students, 168 (96.0%) students scored zero and only 7 (4.0%) students scored the maximum 2 marks.

The students were required to compare and equate the given bivariate random density function to the general bivariate normal distribution equation so as to obtain the two means from the given equation.

From the results of the findings, the students could not equate the equations or/and were unable to relate between the bivariate normal density function given and the bivariate normal distribution equation, to enable them to obtain the values of μ_1 and μ_2 .

It is evident that the students had difficulties in learning the bivariate normal distribution content. Table 4.9 shows the students' performance on question 8a.

Table 4.9: Students' performance on question 8a of the BND test

Mark	No. of students
0	168 (96.0%)
1	0 (0%)
2	7 (4.0%)

One of the students solved the question as follows:

Find the values of:
a) μ_1 and μ_2 (2mks)

μ_1

Exponent of a bivariate random density function:

$$e^{-\frac{1}{2(1-\rho^2)} \left[\left(\frac{x-\mu_1}{\sigma_1} \right)^2 - \frac{(x-\mu_1)\left(\frac{y-\mu_2}{\sigma_2} \right) + \left(\frac{y-\mu_2}{\sigma_2} \right)^2}{\rho} \right]}$$

$\therefore (x+2)^2 = \mu_1$ $x^2 + 4x + 4 = \mu_1$

$x(x+2) + 2(x+2)$ $\mu_1 = \frac{x^2 + 4x + 4}{\sigma_1}$

$x^2 + 2x + 2x + 4$

$\mu_2 = 4(y-1)^2$

$y(y-1) - 1(y-1)$ M_0

$y^2 - y - y + 1$

$y(y^2 - 2y + 1) = \mu_2$ A_0

$\frac{y^3 - 2y^2 + y}{\sigma_2}$

Figure 4.8a: Example of a student's solution to question 8 of the BND test

The student only expanded the sub-equations of the main bivariate normal distribution equation which was a wrong working of the question. The student wrote the exponent of the general bivariate normal function properly but could not equate the corresponding variables to the function given.

Another student solved the question as follows:

Find the values of:

a) μ_1 and μ_2

(2mks)

$$Q = \left(\frac{x-\mu_1}{2}\right)^2 - 2\rho \left(\frac{x-\mu_1}{2}\right) \left(\frac{y-\mu_2}{2}\right) - \left(\frac{y-\mu_2}{2}\right)^2$$

$\mu_1 = 2$ ✓
 $\mu_2 = -2$ ✓

Figure 4.8b: Example of a student’s solution to question 8 of the BND test

The student wrote the wrong exponential function of the bivariate normal distribution. In the solution of the question, the student also wrote 2 in the equation instead of the standard deviations, δ_1 and δ_2 , of the two random variables. Therefore, equating the corresponding variables could not yield the required values.

Question 8 (b):

Sub-question 7.3:

Given that the bivariate random density function of y is:

$$f(x, y) = \frac{1}{2\pi\delta_1\delta_2\sqrt{1-\rho^2}} e^{-\frac{1}{1.02}\{(x+2)^2 - 2.8(x+2)(y-1) + 4(y-1)^2\}}$$

Find the values of δ_1 and δ_2

(2marks)

In this question, students were required to obtain the values of δ_1 and δ_2 from the bivariate random density function equation given. Table 4.10 shows the scores of students on the question. Of the 175 sampled students, all 175 (100.0%) scored zero.

The students were required to relate and equate the given bivariate random density function to the general bivariate normal distribution equation for the two independent variables, so as to obtain the two standard deviations. But from the results of the research findings, the students could not see the relationship between the equation given and the general bivariate normal distribution equation to enable them obtain the values of δ_1 and δ_2 . All 175 students representing 100.0% could not obtain the values. Table 4.10 shows the students' performance on question 8b.

Table 4.10: Students' performance on question 8b of the BND test

Mark	No. of students
0	175 (100.0%)
1	0 (0%)
2	0 (0%)

One of the few students who attempted to solve the question, only tried to solve section (a) but failed to attempt section (b). The solution by the student is as shown in Figure 4.8.

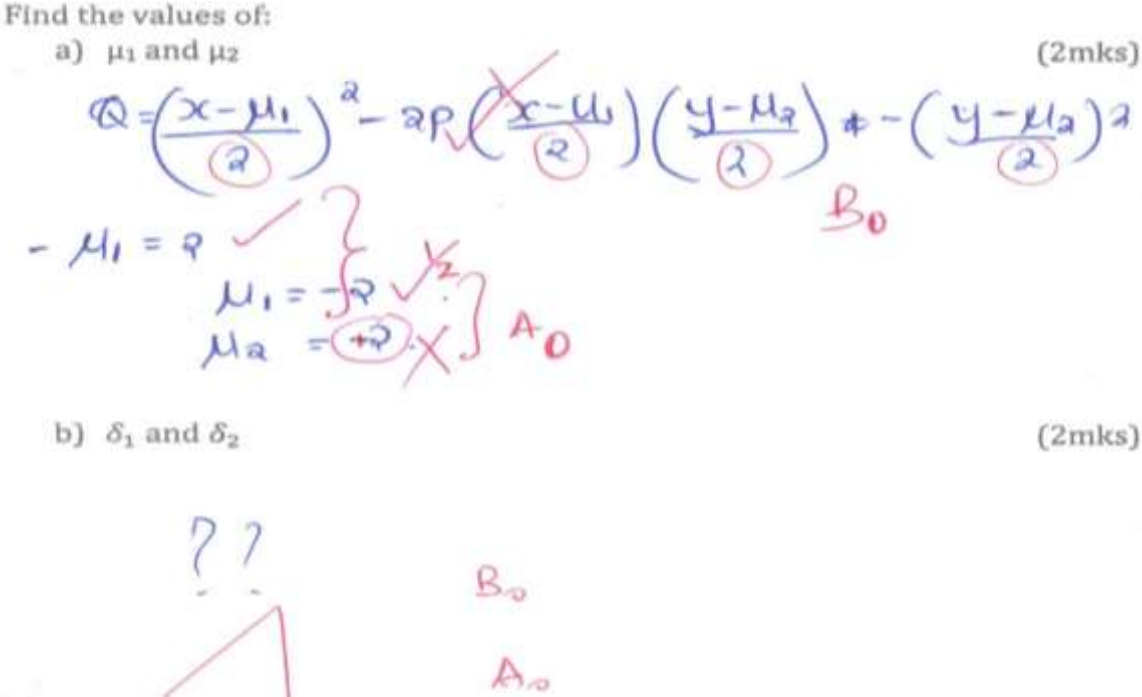


Figure 4.8c: Example of a student's solution to question 8 of the BND test

The student was unable to solve the values of the standard deviations for the two random variables. The student left the question blank.

Another student solved the question as in Figure 4.8d.

b) δ_1 and δ_2 (2mks)

from $u_1 = \frac{x^2 + 4x + 4}{\sigma_1}$

$\sigma_1 = \frac{x^2 + 4x + 4}{u_1}$ M₁

and $u_2 = \frac{(y^2 - 2y + 1) \cdot 4}{\sigma_2}$

$\sigma_2 = \frac{4(y^2 - 2y + 1)}{u_2}$ M₂

$\sigma_2 = \frac{4y^2 - 8y + 4}{u_2}$

Figure 4.8d: Example of a student's solution to question 8 of the BND test

The student expanded the sub-equations of the main equation and could not equate the corresponding variables of the equation.

The pass mark per question was set at half and above of the total marks in that question. Table 4.11 summarises the frequencies and percentages of the students who passed or failed in each of the questions in the test.

Table 4.11: Summary of results of statistical analysis of students' performance in the BND test

Question		Frequency	Percent
State an equation for a normal distribution function for a random variable X	Fail	97	55.4%
	Pass	78	44.6%
Given that, $X \sim N(1, 4)$, find $P(1 \leq X \leq 3)$	Fail	106	60.6%
	Pass	69	39.4%
Determine the setting μ so that only 1% of the cans of paint will be unacceptable	Fail	175	100%
	Pass	0	0%
Differentiate the equation, $y = \frac{1}{8}e^{-4x}$	Fail	21	12.0%
	Pass	154	88.0%
Find the integral, $\int 3xe^{2x^2} dx$	Fail	89	50.9%
	Pass	86	49.1%
State a bivariate normal distribution equation for two random variables X and Y	Fail	103	58.3%
	Pass	73	41.7%
Probability that the height of the wife lies between 5.28 and 5.92 given that the height of the husband is 6	Fail	173	98.9%
	Pass	2	1.1%
Given that the bivariate random density function, find the values of μ_1 and μ_2	Fail	168	96.0%
	Pass	7	4.0%
Given that the bivariate random density function, find the values of δ_1 and δ_2	Fail	175	100.0%
	Pass	0	0%
Total		175	100.0%

4.1.1 Summary of results of the BNDT

The results show that the majority of students did not score up to half of the marks on all the questions except for questions 4 and 5 (differentiation and integration) where 88% and 49.1% of the students were able to solve the questions correctly.

The performance in questions on the bivariate normal distribution (questions 6, 7 and 8, Tables 4.7, 4.8, 4.9 and 4.10) was very poor.

It was also evident that most of the students who correctly solved question 4 on differentiation, also correctly solved question 5 on integration. Yet the students who

performed well in questions 4 and 5 did not perform well in questions on the normal distribution and bivariate normal distributions mainly because they could not recall the long formulas.

More than the expected number of students got very low marks in questions 1 and 6 (stating the normal distribution and the bivariate normal distributions respectively).

The students' knowledge on the exponential differentiation and integration questions seems sufficient for the learning and analysis of the bivariate normal distribution as shown by the research study results (see Tables 4.5 and 4.6).

4.2 RESULTS FROM THE STUDENT QUESTIONNAIRE

The questionnaire consisted of two sections. The first section was anchored on demographic information of the students. The second section consisted of 27 likert type items and an open-ended question which sorts the students' perception on the factors that contribute the difficulties they encounter in learning BND.

Each statement in the research questionnaire was analyzed according to the students' rating of the responses to the statements. In this study, the researcher selected the statements which were regarded as majorly or directly influencing the learning of the bivariate normal distribution content. Each of the statements was analyzed individually.

The themes identified from the questionnaires and the number of students (or lecturers) who responded along the themes are shown in the Table 4.29 for students and section 4.3.2 summary for lecturers.

Student interest in learning the bivariate normal distribution content

Lack of interest in learning a mathematical content, negatively impacts on a student's learning (Mbugua et al., 2012). Students' negative interest and wrong perception of the BND content negatively affects the student learning of this content.

Table 4.12 shows the students' responses on the extent to which they perceive their interest in BND constitutes a problem in learning the content. Table 4.12 shows that 90 students, representing 51.4%, believed that their interest was a problem in learning the content.

Table 4.12: Students' perception of the effect of their interest in learning BND content

	Frequency	Percent
Not at all	25	14.3
To a minor extent	60	34.3
To some extent	73	41.7
To a major extent	17	9.7
Total	175	100.0

Lack of background knowledge on the bivariate normal distribution content

Lack of adequate background knowledge on BND, impacts negatively on the students' learning. The students' inadequate knowledge on BND hampers the lecturers' teaching and hence the students' learning of the content.

Table 4.13 shows the students' responses to their perception of the effect of background knowledge on learning BND. Table 4.13 shows that 152 (88.8%) of the students find learning the BND difficult because of their inadequate knowledge of BND.

Table 4.13: Students' perception on the effect of lack of background knowledge on learning the BND content

	Frequency	Percent
Not at all	5	2.9
To a minor extent	18	13.2
To some extent	62	35.4
To a major extent	90	51.4
Total	175	100.0

Large number of students in a class of bivariate normal distribution

A large population of students in a class negatively affects the students' effective learning of the bivariate normal distribution. The students' individual differences are not catered for with large numbers of students in class, especially on the topic of the BND. In statistics units,

some of the content taught must be written clearly and be seen on the chalkboard by all students to be able to copy the correct notes.

Table 4.14 shows that 135 (77.1%) of the sampled students indicated that the large number of students in the bivariate normal distribution class especially the BSc. mathematics and BSc. physical sciences classes, hinders learning of the BND content. For effective teaching and meaningful learning to take place in statistics, an average class size should not exceed forty students.

Table 4.14: Students' perception of the effect of large number of students in class on the BND content

	Frequency	Percent
Not at all	9	5.1
To a minor extent	31	17.7
To some extent	78	44.6
To a major extent	57	32.6
Total	175	100.0

Inadequate learning resources in the university

Adequate learning resources in a learning institution are very important for any effective learning to take place. Enough textbooks, writing material and reference materials in a learning institution, for example, a university ensures that students learn and research what they need in the shortest possible time thereby enabling effective learning.

Of the 175 sampled students (see Table 4.15), 48.6% of the students indicated that there were inadequate learning resources in the university, which probably affects the student learning of a BND.

Table 4.15 also indicates that 51.4% of the sampled students believe that their university has adequate learning resources and therefore their effective learning of BND is not compromised.

Table 4.15: Results of students' perception on the effect of inadequate learning resources on learning BND

	Frequency	Percent
Not at all	21	12.0
To a minor extent	69	39.4
To some extent	48	27.5
To a major extent	37	21.1
Total	175	100.0

Long and complex equations

Long and complex equations give students problems in learning and understanding and therefore negatively affect their learning.

It was found from this study that most students (96.0%) felt long and complex equations negatively affect their learning of the BND content. Table 4.16 shows that 168 (96.0%) perceive that long and complex equations negatively affect their learning of the BND.

Table 4.16: Students' perception on the effect of long and complex equations on learning of BND

	Frequency	Percent
Not at all	3	1.7
To a minor extent	4	2.3
To some extent	29	16.6
To a major extent	139	79.4
Total	175	100.0

Tests and examinations on the bivariate normal distribution content

The complex nature of tests and examinations given to students on the BND affects their learning. The content examined on the BND in the continuous assesment tests which includes the equation and its applications may lead to some students lose interest learning it. When simple content is examined in the continuous assesment tests, students may perform well.

Table 4.17 illustrates the students' perception of the effects of tests and examinations on the learning of BND. The table illustrates that of the 175 students who were sampled, 159 (90.9%) believe that tests and examinations on the BND content impact negatively on the effective learning of the content.

Table 4.17: Students' perception on the effect of tests and examinations on their learning of BND

	Frequency	Percent
Not at all	2	1.1
To a minor extent	14	8.0
To some extent	60	34.3
To a major extent	99	56.6
Total	175	100.0

Students' mathematical background on the bivariate normal distribution content

The students' inadequate background mathematical knowledge on bivariate normal distribution, impacts negatively on the students' learning. The students' inadequate mathematical knowledge on the BND hampers the lecturers' teaching and hence the students' learning of the content.

Table 4.18 shows the students' responses to their perception of the effect of background knowledge on learning of BND. The table below shows 159 (91.9%) of the students find learning of BND difficult due to inadequate mathematical background knowledge on the BND.

Table 4.18: Students' perception on their mathematical background on their learning of BND

	Frequency	Percent
Not at all	2	1.1
To a minor extent	14	8.0
To some extent	45	25.8
To a major extent	114	65.1
Total	175	100.0

Effect of long bivariate normal distribution equation in learning

Long equations with many parameters and constants become difficult for some students to learn and understand. The long bivariate normal distribution equation, affects students' learning.

Table 4.19 shows the students' responses to their perception on the effect of long bivariate normal distribution. The table below shows 158 (90.3%) of the students find learning BND difficult due to the lengthy nature of the equation.

Table 4.19: Students' perception on the effect of the long bivariate normal distribution equation on their learning of BND

	Frequency	Percent
Strongly agree	78	44.6
Agree	80	45.7
Disagree	16	9.1
Strongly disagree	1	0.6
Total	175	100.0

Complex nature of the bivariate normal distribution equation

Equations which are complex in nature make learning the content difficult. The BND equation is long and has many parameters and constants and therefore seems complex for students.

Table 4.20 shows the students' perception on the effect of complex nature of the BND. The table below shows 160 (91.4%) of the students find learning of BND difficult because of its complex.

Table 4.20: Results of students' perception on the effect of complex nature of the bivariate normal distribution equation on their learning of the content

	Frequency	Percent
Strongly agree	69	39.4
Agree	91	52.0
Disagree	14	8.0
Strongly disagree	1	0.6
Total	175	100.0

Students' negative attitude towards bivariate normal distribution content

The students' negative attitude towards learning of a BND content, negatively impacts on the students' learning. Students' negative attitude and wrong perceptions BND content negatively affect their learning of the content.

Table 4.21 shows the students' responses on the extent to which they perceive their negative attitude in BND constitute a problem in learning the content. The table shows that 101

students, representing 57.7%, perceive that their negative attitude towards learning BND was a problem in learning the content.

Table 4.21: Results of students' perception on the impact of their negative attitude towards bivariate normal distribution on their learning of the content

	Frequency	Percent
Strongly agree	34	19.4
Agree	67	38.3
Disagree	70	40.0
Strongly disagree	4	2.3
Total	175	100.0

Many parameters and constants in the BND equation make its learning, analysis and hence understanding difficult.

The BND equation is long and has many parameters and constants which makes the equation complex and hence makes learning the content difficult. Some equations might be long with many parameters and constants but simple to students because its analysis is easy and hence easily learnt understood.

Table 4.22 shows the students' perception of the impact of many parameters and constants in BND on their learning of the content. Table 4.22 shows that 165 (94.3%) of the students find learning BND difficult due to the many parameters and constants in the BND equation.

Table 4.22: Results of students' perception on the impact of many parameters and constants in BND equation on their learning of the content

	Frequency	Percent
Strongly agree	90	51.4
Agree	75	42.9
Disagree	8	4.6
Strongly disagree	2	1.1
Total	175	100.0

Adequate background knowledge on the normal distribution equation

Lack of adequate background knowledge on the normal distribution, impact negatively on the students' learning of the bivariate normal distribution. The students' inadequate background knowledge on the normal distribution hampers the lecturers' teaching and hence the students' learning of the BND. Table 4.23 shows the students' perception on the effect of inadequate background knowledge of the normal distribution on the learning of BND. The table shows that 74 (41.1%) of the students, find learning of BND difficult due to the their inadequate background knowledge of the normal distribution.

Table 4.23: Results of students' perception of their background knowledge of the normal distribution equation

	Frequency	Percent
Strongly agree	21	12.0
Agree	51	29.1
Disagree	87	49.7
Strongly disagree	16	9.1
Total	175	100.0

Adequate knowledge on differentiation and integration

Students' adequate knowledge on differentiation and integration positively influence their learning of the BND. Table 4.24 shows the students' views on having adequate knowledge of differentiation and integration on the learning of BND. The table shows that 139 (79.4%) of the students have adequate knowledge on differentiation and integration which is necessary for the learning of the BND.

Table 4.24: Results of students' views on having adequate knowledge of differentiation and integration on the learning of BND.

	Frequency	Frequency
Strongly agree	44	25.1
Agree	95	54.3
Disagree	30	17.2
Strongly disagree	6	3.4
Total	175	100.0

Lecturers' testing of difficult content on the bivariate normal distribution

The content the lecturers examine on the BND affects the students' learning of the BND. Testing all the difficult content on the BND leads to student failure in the content.

Table 4.25 shows the students' views and perception of the effect of the lecturers' testing of difficult content on the BND. The table shows that 138 (78.9%) students find the learning of the BND difficult because the lecturers' test difficult content on the BND.

Table 4.25: Results of students' views on lecturers' testing of difficult content on the bivariate normal distribution

	Frequency	Percent
Strongly agree	65	37.1
Agree	73	41.8
Disagree	21	12.0
Strongly disagree	16	9.1
Total	175	100.0

Student background in mathematical equations and their understanding and performance on bivariate normal distribution

Students with a good understanding of basic statistical equations (for example, the normal distribution equation) tend to understand other related but advanced equations (for example, the BND) faster and hence perform well in their studies. Of the 175 sampled students, 132 (75.4%) find the learning of the BND difficult due to their inadequate background in statistical equations that negatively impacts their understanding and performance on the BND content. Table 4.26 shows that 132 (75.4%) of the students, have a weak background in mathematical equations which negatively affects their learning of BND.

Table 4.26: Results of students' views on their background in mathematical equations and understanding and performance on bivariate normal distribution

	Frequency	Percent
Strongly agree	42	24.0
Agree	90	51.4
Disagree	35	20.0
Strongly disagree	8	4.6
Total	175	100.0

Ability of students to state the bivariate normal distribution equation

Long equations with many parameters and constants are difficult to state. The BND equation is long and also has many parameters and constants and students find it difficult to state the equation when called upon to. Table 4.27 shows the students' views on their ability to state the BND. The table shows that 123 (70.3%) of the students find it difficult to state the BND equation, consequently leading to the difficulties they experience in learning the BND.

Table 4.27: Results of students' views on whether they find it difficult to state the bivariate normal distribution equation

	Frequency	Percent
Strongly agree	5	2.9
Agree	47	26.8
Disagree	90	51.4
Strongly disagree	33	18.9
Total	175	100.0

Students' views on their understanding of the step-by-step analysis of the bivariate normal distribution equation

The analysis of equations results in the formation of other equations, for example, the analysis of the BND equation results in the formation of the conditional mean and variance equations of one variable given the other variable. Table 4.28 shows the students' perception of their understanding of the step-by-step analysis of the BND equation. The table shows that 133 (76.0%) of the students do not understand the step-by-step analysis of a BND equation to obtain the conditional mean and variance equations.

Table 4.28: Results of students' views on whether they understand step-by-step analysis of the bivariate normal distribution equation

	Frequency	Percent
Strongly agree	2	1.1
Agree	40	22.9
Disagree	83	47.4
Strongly disagree	50	28.6
Total	175	100.0

Students' responses to the open-ended question on the difficulties they experience in the learning of BND.

At the end of the student questionnaire, the students were probed in their own view, to state any other difficulties they experience in learning of a BND content. The common themes that emerged from the open-ended question are tabulated below.

Table 4.29 shows some of the common difficulties experienced by students according to their statements.

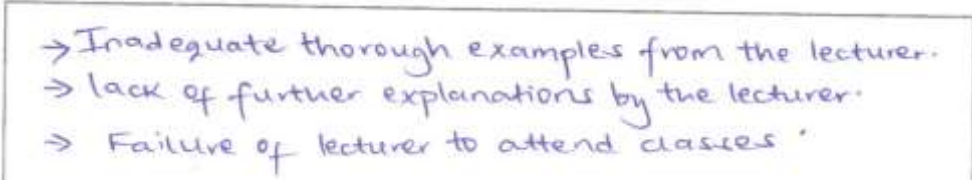
Table 4.29: Students' responses to the difficulties they experience in the learning of bivariate normal distribution.

Response	Frequency	Percent
Lack of interest in BND distribution	13	7.4
BND equation long to memorise	20	11.4
Poor lecturing method	13	7.4
BND equation is boring to learn	11	6.3
Lack of adequate reference materials	22	12.6
Lack of adequate background knowledge on the BND.	17	9.7
Few examples given by lecturers on BND	22	12.6
Inadequate in-depth teaching of the bivariate normal distribution.	13	7.4
Lecturers' tendency of not attending all BND lessons including makeup lessons.	11	6.3
Inexperienced and young lecturers teaching the BND content	8	4.6
Students' inadequate hard work	2	1.1
Insufficient explanation on BND content	8	4.6
Lack of proper guidance on learning BND	2	1.1
Complex nature of the BND questions in examinations	22	12.6
Short teaching and learning time on the BND content	8	4.6
Long and complex BND equations which are hard to memorize and understand	11	6.3

In Table 4.29, 22 students indicated that the lack of adequate examples in class on the BND is a problem in learning its content. A total of 17 students indicated that the lack of further explanation by the lecturers hinders their effective learning of the BND content while 11 students indicated that the failure of lecturers to attend classes regularly negatively affects their learning of the BND.

Figure 4.9 shows an example of a student's response to the difficulties she/he experiences in learning of the BND.

35. In the space provided below, state any other difficulties you face in learning of a bivariate normal distribution.



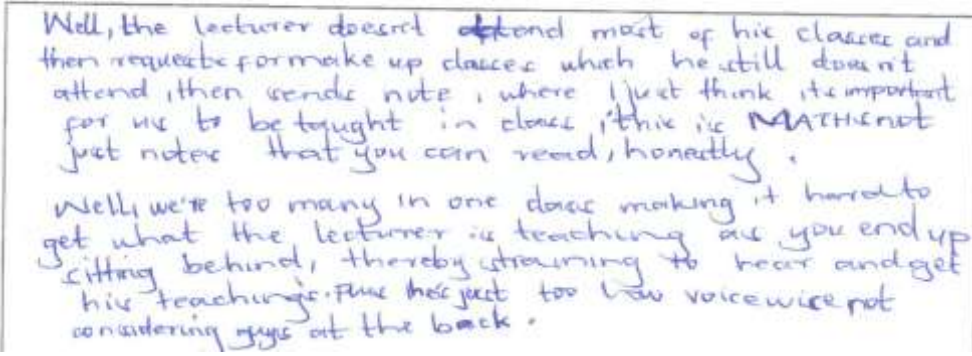
→ Inadequate thorough examples from the lecturer.
→ lack of further explanations by the lecturer.
→ Failure of lecturer to attend classes.

Figure 4.9: A student stating the difficulties she experiences in learning the BND content

It was found from this study (Table 4.29) that 11 students indicated that some lecturers who teach BND do not attend most of the lessons and these lecturers ask for make-up classes which they again do not attend.

Figure 4.10 below shows an example of a student's responses to the difficulties he/she experiences in learning BND.

35. In the space provided below, state any other difficulties you face in learning of a bivariate normal distribution.



Well, the lecturer doesn't attend most of his classes and then requests for make up classes which he still doesn't attend, then sends notes, where I just think it's important for us to be taught in class, this is MATHS not just notes that you can read, honestly.

Well, we're too many in one class making it hard to get what the lecturer is teaching as you end up sitting behind, thereby straining to hear and get his teaching. Plus he's just too low voice wise not considering guys at the back.

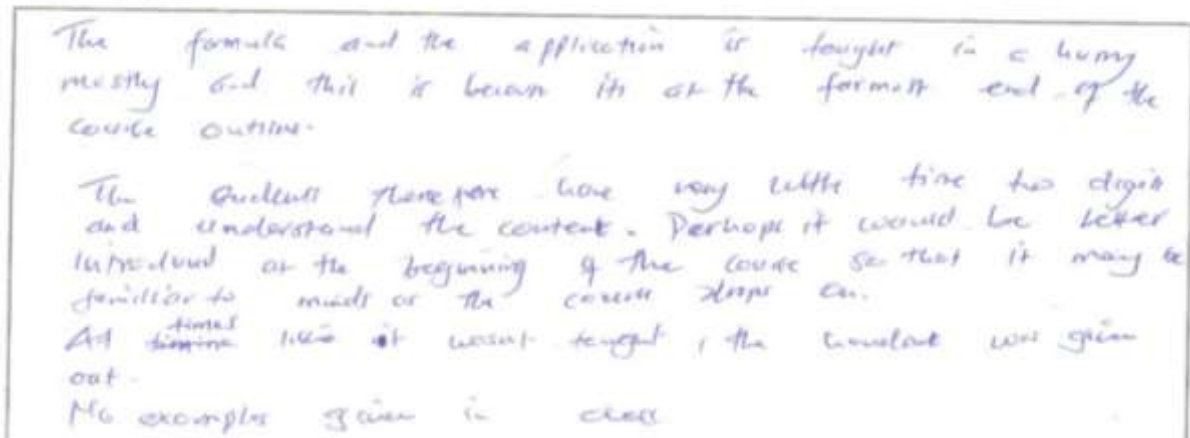
Figure 4.10: A student stating the difficulties he experiences in learning BND content

The study found (Table 4.29) that 8 students felt that the amount of teaching and learning time of the BND is short. They indicated that the BND equation and its application is taught within a limited time and they have too little time to understand the content before examinations.

A total of 22 students (see Table 4.29) indicated that lack of adequate examples in class on the content affects their learning of the content. They indicated that the BND content in probability and statistics III is taught towards the end of the unit, so they do not have enough time to revise the content for the examinations. The course outline is organized in such a way that the content is taught towards the end of the unit.

Figure 4.11 below shows an example of a student's response to the difficulties he/she experiences in learning the BND.

35. In the space provided below, state any other difficulties you face in learning of a bivariate normal distribution.



The formula and the application is taught in a hurry mostly and this is because it is at the former end of the course outline.

The students therefore have very little time to digest and understand the content. Perhaps it would be better introduced at the beginning of the course so that it may be familiar to minds as the course steps on.

At times when it wasn't taught, the student was given out.

No examples given in class.

Figure 4.11: A student stating the difficulties he experiences in learning the bivariate normal distribution content

From Table 4.29, 11 students indicated that the BND is boring to learn and they do not understand it. Table 4.29 also shows that 11 students indicated that the BND equation is long and complex and therefore hard to understand and memorize.

Figure 4.12 shows an example of a student's response to the difficulties he/she experiences in learning BND.

35. In the space provided below, state any other difficulties you face in learning of a bivariate normal distribution.

Actually and sincerely, I don't understand the bivariate normal distribution. It is a tall order for me just from the word go, I couldn't make any meaning to it. I have difficulties in the bivariate normal distribution in all sectors.
Kindly note that I really regret doing bivariate normal distribution.
I HAVE DIFFICULTIES IN BIVARIATE NORMAL DISTRIBUTION AS A WHOLE.

Figure 4.12: A student stating the difficulties he experiences in the learning the bivariate normal distribution content

Table 4.29 also shows that 22 students felt that the BND equation is long and complex, and they were unable to remember it. In addition, 8 students felt that lecturers' insufficient explanation on the BND content is a problem for effective the learning of bivariate normal distribution.

From Table 4.29, 11 students indicated lecturers' tendency of not attending BND lessons regularly, including the make-up lessons, and that this hampers their learning of the bivariate normal distribution.

Figure 4.13 shows an example of a student's response to the difficulties he/she experiences in learning a BND.

35. In the space provided below, state any other difficulties you face in learning of a bivariate normal distribution.

* Calculation of the mgf.
* Use of the distribution tables.
* Long sentenced question
* Requirement of one to cram a long Bivariate equation that one cannot remember half way.
* During examination testing marks are distributed unevenly therefore ~~making a long~~
* which taught in the back years [1st year]; cannot remember but we are tested on it
* NOT ATTENDING CLASSES AND ASKING FOR MAKE-UPS AND NOT ATTENDING TOO!

Figure 4.13: A student stating the difficulties she experiences in learning the bivariate normal distribution content

Figure 4.29 shows, of the 175 sampled students, 2 students indicated that they lack proper guidance on learning BND and 22 students indicated that lecturers work few examples in class but test difficult content in examinations, thereby affecting their learning of BND. Also 8 students felt that the use of inexperienced and young lecturers in teaching the BND hampers their learning. Similarly, 11 students felt that the missing of some lessons by lecturers hampers their learning of the BND content.

Figure 4.14 shows an example of a student's response to the difficulties he/she experiences in learning of a BND.

35. In the space provided below, state any other difficulties you face in learning of a bivariate normal distribution.

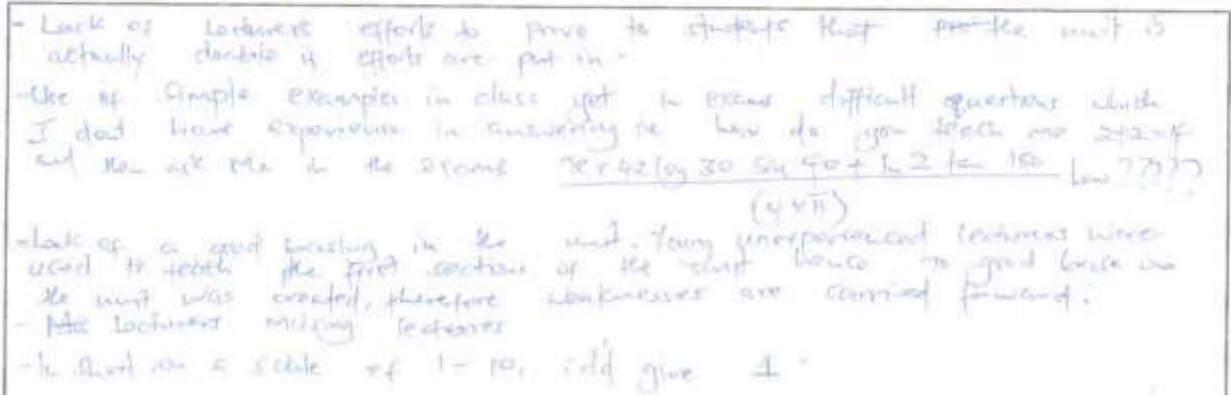


Figure 4.14: A student stating the difficulties he experiences in learning of the bivariate normal distribution content

Figure 4.14 shows an example of a student's response to the difficulties he/she experiences in learning of a BND.

35. In the space provided below, state any other difficulties you face in learning of a bivariate normal distribution.

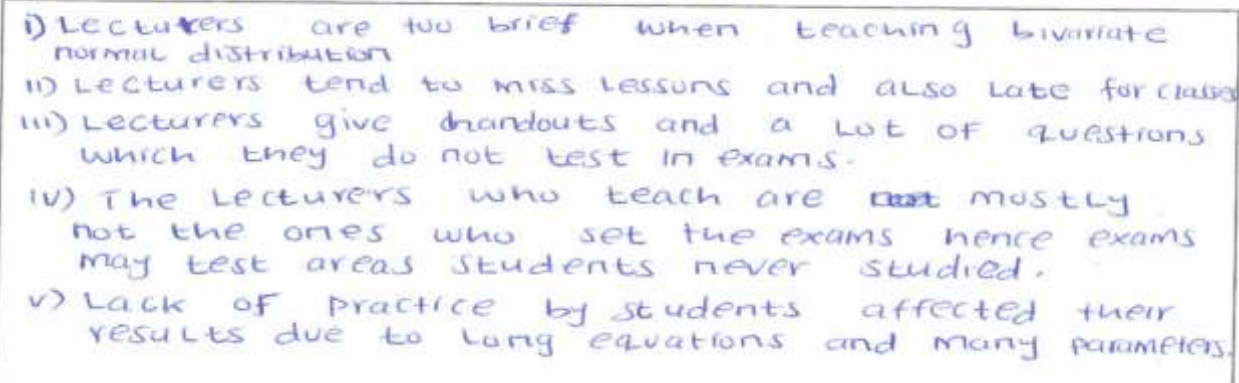


Figure 4.15: A student stating the difficulties she experiences in learning the bivariate normal distribution content

Of the 175 sampled students, 13 students (Table 4.29) indicated that they lack interest in learning the bivariate normal distribution content. Also 11 students felt that long equations negatively affect their learning of the BND.

A total of 13 students also indicated that the lecturers' shallow teaching of the BND affects their learning of the BND content.

Figure 4.16 shows an example of a student's response to the difficulties he/she experiences in learning of a BND.

35. In the space provided below, state any other difficulties you face in learning of a bivariate normal distribution.

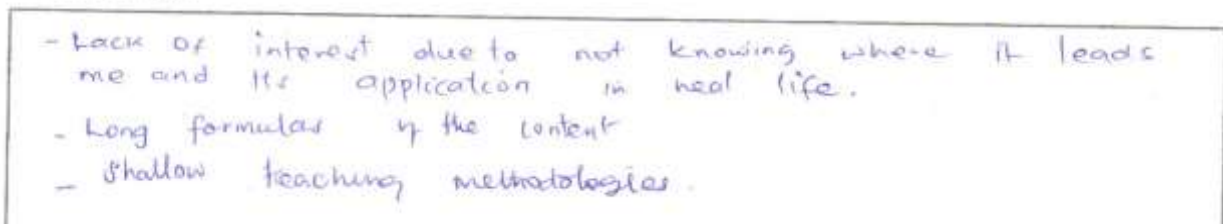


Figure 4.16: A student stating the difficulties she experiences in learning the bivariate normal distribution content

4.2.1 Summary of the results from the student questionnaire

- i. From the student questionnaire, it emerged that the analysis of the BND equation and the derivation of the conditional mean and variance equations is difficult for students.
- ii. The study found that the BND equation is long and complex for some students to comprehend.
- iii. From the research study findings, it was discovered that long and complex equations affect the students' learning of BND content.
- iv. The students' have a negative attitude towards learning of BND content, which negatively hinders their effective learning of this content.
- v. Also discovered was that lack of adequate resources in the higher institution of learning e.g. reference books, hampers the effectiveness of learning of the BND. Though this students can access internet to obtain supplementary resources e.g. you-tube videos and notes.

- vi. It was found from this study, that tests and examinations on the BND content impact negatively on effective learning of the BND since students have the perception that the BND content is very difficult.
- vii. The long bivariate normal distribution function with its five parameters and two constants of the two independent random variables affects the student's learning of the content.
- viii. The study found that students have a perception that the BND equation is long to comprehend and that makes its learning and understanding difficult.
- ix. The students also have a perception that the BND equation is complex with its parameters and constants of the independent variables and hence difficult to understand and remember when examined.
- x. Students have inadequate background knowledge related to the BND which hinders the effective learning of the content.
- xi. The study found that students have a poor background in mathematical equations which hampers the learning of BND.
- xii. The students' have adequate knowledge on differentiation and integration which is applicable in the derivation of the conditional mean and variance equations.
- xiii. The students lack interest in learning BND.
- xiv. The research study findings indicated that students have a perception that lecturers test difficult content in the BND examinations.
- xv. The stepwise analysis of the BND equation to obtain the conditional mean and variance of one variable given another variable is a problem to the students.
- xvi. Students lack adequate foundational knowledge on the normal distribution and other related content useful for the analysis of the BND equation.

4.3 LECTURER OPEN-ENDED QUESTIONNAIRE

The responses from the lecturer questionnaire were used to elaborate on the findings from the test and the students' questionnaire and further understand the difficulties the students experience in the learning of BND.

Six lecturers out of 14 participated in the research study. The lecturers were identified as lecturer 1 to lecturer 6.

4.3.1 Results from the lecturer questionnaire

The following represents the responses of the lecturers:

Lecturer 1:

As, a lecturer, teaching statistics which includes the units on bivariate normal distribution,

Question: How many years have you taught the course?

Response: Five years

Question: How many hours are allocated to the Statistics unit per week in this session?

Response: Three hours split in two hours and one hour lessons in 2 days a week

Question: What is the average size of the class or classes you are teaching in the unit?

Response: Sixty students

Question: Do you think students have difficulties in learning bivariate normal distribution?

Response: Yes

Question: If yes to question 4 above, what do students find difficult in learning the content?

Response: Obtaining of marginal densities

Question: In your opinion, why do the students experience difficulties in learning the bivariate normal distribution content?

Response: The inability of students to integrate the marginal densities

Lecturer 2:

As, a lecturer, teaching statistics which includes the units on bivariate normal distribution,

Question: How many years have you taught the course?

Response: Two years

Question: How many hours are allocated to the Statistics unit per week in this session?

Response: Three hours

Question: What is the average size of the class or classes you are teaching in the unit?

Response: Eighty students

Question: Do you think students have difficulties in learning bivariate normal distribution?

Response: Yes

Question: If yes to question 4 above, what do students find difficult in learning the content?

Response: Derivation and application of the bivariate normal distribution equations

Question: In your opinion, why do the students experience difficulties in learning the bivariate normal distribution content?

Response: Lack of sufficient foundational knowledge on the learning of bivariate normal distribution content

Lecturer 3:

As, a lecturer, teaching statistics which includes the units on bivariate normal distribution,

Question: How many years have you taught the course?

Response: More than nine years.

Question: How many hours are allocated to the Statistics unit per week in this session?

Response: Three hours, mostly split, one hour and two hour lessons.

Question: What is the average size of the class or classes you are teaching in the unit?

Response: Thirty students.

Question: Do you think students have difficulties in learning bivariate normal distribution?

Response: Yes, they do.

Question: If yes to question 4 above, what do students find difficult in learning the content?

Response: The derivation of the conditional mean and variance equations, of one variable given the other, and their applications in solving questions.

Question: In your opinion, why do the students experience difficulties in learning the bivariate normal distribution content?

Response: Students' fear and negative attitude towards long equations and to some extent, inadequate pre-requisite knowledge on bivariate normal distribution.

Lecturer 4:

As, a lecturer, teaching statistics which includes the units on bivariate normal distribution,

Question: How many years have you taught the course?

Response: Two years

Question: How many hours are allocated to the Statistics unit per week in this session?

Response: Three hours

Question: What is the average size of the class or classes you are teaching in the unit?

Response: Eighty five students

Question: Do you think students have difficulties in learning bivariate normal distribution?

Response: Yes

Question: If yes to question 4 above, what do students find difficult in learning the content?

Response: Understanding the bivariate normal distribution function

Question: In your opinion, why do the students experience difficulties in learning the bivariate normal distribution content?

Response: The difficulty comes when trying to remember the bivariate probability density function.

Lecturer 5:

As, a lecturer, teaching statistics which includes the units on bivariate normal distribution,

Question: How many years have you taught the course?

Response: Three years

Question: How many hours are allocated to the Statistics unit per week in this session?

Response: Three hours

Question: What is the average size of the class or classes you are teaching in the unit?

Response: Sixty students.

Question: Do you think students have difficulties in learning bivariate normal distribution?

Response: No, because most of the students have demonstrated good understanding of the course.

Question: If yes to question 4 above, what do students find difficult in learning the content?

Response:

Question: In your opinion, why do the students experience difficulties in learning the bivariate normal distribution content?

Response: It is a bit complex and requires more time to go through the content to understand

Lecturer 6:

As, a lecturer, teaching statistics which includes the units on bivariate normal distribution,

Question: How many years have you taught the course?

Response: Six years

Question: How many hours are allocated to the Statistics unit per week in this session?

Response: Three hours per week

Question: What is the average size of the class or classes you are teaching in the unit?

Response: Approximately sixty students

Question: Do you think students have difficulties in learning bivariate normal distribution?

Response: Yes

Question: If yes to question 4 above, what do students find difficult in learning the content?

Response: Analysis, understanding and using the derived equations on the bivariate normal distribution, e.g. the marginal densities and the conditional mean and variance equations.

Question: In your opinion, why do the students experience difficulties in learning the bivariate normal distribution content?

Response: Lack of adequate knowledge in statistical equations and perception that long equations are difficult to understand.

4.3.2 Summary of lecturers' responses

The results from the lecturers' responses indicate that the students experience difficulties in learning the BND content.

According to the lecturers, students do not have a firm background and pre-requisite knowledge on BND and they find it difficult to do the following:- obtain the marginal density equations of a variable; the probability density function of the BND; analyse the BND equation, and the derivation of the conditional mean, and variance of BND equation and the applications of the equations. These make learning of BND difficult for students.

4.4 SUMMARY OF THE CHAPTER

This chapter discussed the research study findings on the difficulties students experience in the learning of BND content in the Kenyan university.

Description and analysis of the results of the study according to themes and research questions were discussed. Three instruments were used for data collection: the BND test for students, the student questionnaire and the lecturers' open-ended questionnaire.

The results from the test for students were analyzed and presented as descriptive statistics. This chapter also presented the responses from all the student' questionnaires which were also analyzed and presented in frequency tables and percentages to establish the difficulties the students experience in learning BND. The responses from the lecturer questionnaire were also presented in this chapter.

The research study findings revealed that the students had difficulties in learning the BND.

The findings of this study also revealed that the students do not have a firm background on the BND. Moreover, the students do not have adequate pre-requisite knowledge on BND except for the differentiation and integration skills.

The research findings indicated that the students had difficulties in different aspects in statistics including:- marginal densities; probability density functions; analysis of BND equation; and derivations and applications of equations.

The next chapter discusses the research findings, conclusion and recommendations. The chapter also discusses the implications of the research findings, the limitations of the study and suggestions for further studies.

CHAPTER 5

DISCUSSION OF FINDINGS, CONCLUSION AND RECOMMENDATIONS

This chapter provides a summary of the study, discussion of the findings, conclusion and recommendations. Also presented in this chapter are the implications of the research findings, the limitations of the study and suggestions for further research. The conclusion reflects whether the main research question and the other research sub-questions have been answered by the outcomes. In essence, this chapter will indicate how this study could be considered important and provide guidance on future related research studies.

The main aim for this study was to investigate the difficulties undergraduate students experience in the learning of bivariate normal distribution in a Kenyan university.

There were two main research questions. The research questions were as follows:

1. What do students find difficult in learning bivariate normal distribution?

To address this *research question*, the following guiding sub-questions were put forward:

- 1.1 Can the students state the equation of a normal distribution function?
 - 1.2 Can the students obtain the mean given a normal distribution equation?
 - 1.3 Can the students obtain the the standard deviation given a normal distribution equation?
 - 1.4 Can the students differentiate a given exponential equation?
 - 1.5 Can the students find the integral of a given exponential equation?
 - 1.6 Can the students state the BND equation?
 - 1.7 Given the conditional distribution of a variable (e.g. $h(y/X = x)$), can the students calculate the mean, $E(y/X = x)$ of the joint distribution?
 - 1.8 Given the conditional distribution of a variable (e.g. $h(y/X = x)$), can the students calculate the variance, $Var(y/X = x)$ of the joint distribution?
2. Why do students experience difficulties in learning BND?

5.1 SUMMARY OF CHAPTERS

Chapter 1: Introduction of the study

Chapter one dealt with the background of the study, the research problem, the aim of the study, the main objectives of the study and the rationale for the study. Also discussed in this chapter was the significance of the study, the definition of terms used in this study and clarification of concepts.

Chapter 2: Theoretical framework and literature review

In this chapter, the theoretical perspective of Ingvarson et al (2004) was discussed. The factors which influence students' learning of mathematics, which are believed to be the same factors which influence students learning of a bivariate normal distribution were discussed.

This chapter also discussed literature related to the research study. It discussed the difficulties students experience while studying a specific content in mathematics. The chapter dealt with the teaching and learning of bivariate normal distribution, the students' weak background in mathematics, lecturers' methods of content delivery and teaching methods, lecturers' evaluation methods and procedures, availability of the teaching and learning resources, students' attitudes towards some mathematics content and students' interest, inadequate effort and lack of hard work in mathematics. The literature review focused on the difficulties students experience in learning mathematics and especially statistics.

Chapter 3: Research methodology

This chapter discussed the research design, sources of data, population, sample and sampling methods, the research instruments and procedure for data collection. Also discussed in this chapter was the validity and reliability of the research instruments, the pilot study, data analysis, the limitations of the study and ethical consideration of the study.

A mixed methods design was adopted for this study in which both the quantitative and qualitative data were collected. In this study 50 students participated in the pilot study and 175 students out of approximately 482 participated in the main study. The instruments used

for data collection include; the bivariate normal distribution test for students, the student questionnaire and the lecturers' open-ended questionnaire. Permission was sought from the research university to conduct research and all the ethical clearance procedures were fulfilled.

Chapter 4: Findings

This chapter discussed the research study findings. Description and analysis of the results of the study according to themes and research questions was discussed. The results from the test for students were analyzed and presented as descriptive statistics. This chapter also presented the responses from all the students' questionnaire which were also analyzed and presented in frequency tables and percentages to establish the difficulties undergraduate students experience in the learning of bivariate normal distribution in the Kenyan university. The responses from the lecturer open-ended questionnaire were also presented in this chapter.

5.1.1 Summary of the research findings

The research study findings revealed that:-

- i. The students had difficulties in learning the BND. They had difficulties in different aspects in statistics including; marginal densities, probability density functions analysis of bivariate normal distribution equation and derivations and applications of BND equations.
- ii. The students' negative attitude and lack of interest on learning of BND hampered their learning on the topic.
- iii. The students did not have a firm background on the BND or adequate pre-requisite knowledge on the BND.

5.2 DISCUSSION OF FINDINGS

The findings are discussed below, in light of the research questions in connection with the bivariate normal distribution test and the questionnaires.

The test results indicate that the students performed dismally in the BND test questions. The students scored a low mean mark of 29.3%. (see Table 4.1) out of the possible 100 marks. The research study results indicate that questions on differentiation and integration required little effort to find a solution, the more reason why the students (see Tables 4.5, 4.6) solved the problems on differentiation and integration correctly. The students' knowledge on the exponential differentiation and integration was sufficient for some students (see Table 4.5, 4.6), for the effective learning of the BND.

Finally, each of the research questions was addressed with the research findings.

5.2.1 Can students state the equation of a normal distribution function?

A normal distribution function in probability and statistics II, is a basic knowledge acquired by students before they can learn BND in probability and statistics III.

From the research study findings, Table 4.2 shows the students' inability to state the normal distribution equation for a random variable X , which is the foundational knowledge required for effective learning of a BND. It is worth noting that the product of the two marginal density distribution functions of two independent random variables gives rise to the BND function for the two independent variables. The research findings showed that students have difficulties in stating the normal distribution equation for a random variable X , which is basic knowledge in statistics.

Normal distribution content is taught in the first year at the university, in the unit probability and statistics II. This means, therefore that a student must have adequate knowledge in the normal distribution content and the marginal distribution in order to effectively learn and understand BND.

From the research findings, students indicated that the normal distribution content and its application was not well learnt and understood in their first year of study. This is evident from the student test on the BND in which the students (see Table 4.2), were unable to state the equation of the normal distribution function. From the lecturers questionnaire, it emerged that the students lack adequate pre-requisite knowledge (e.g. the normal distribution) for learning BND.

It can therefore be concluded from this study, that the students have inadequate knowledge on the normal distribution.

5.2.2 Can the students obtain the standard deviation given a normal distribution equation?

From the findings of this study, students were unable to find the probability of a normal random variable. The normal distribution equation is pre-requisite knowledge for learning the BND and students with inadequate knowledge on the normal distribution, find learning of the BND a very difficult task. Normal distribution is learnt in the first year in probability and statistics II unit.

From the research study findings, students (see Table 4.2) were unable to solve a normal distribution question to find the probability of the normal distribution.

From the student questionnaire, it emerged that the normal distribution content and its application was not learnt well and understood in the students' previous year of study (Table 4.23). From the student test on the BND (see Table 4.2), students were unable to state the equation of the normal distribution function. They were also unable to calculate the mean, the standard deviation and hence the probability of a normal random variable. The lecturers indicated that the students' inadequate knowledge on equations and lack of a firm foundation (e.g. the normal distribution) affect the students' learning of BND. This study compares to that of Ogbonnaya and Mji (2013) who note that students' weak background in mathematics contributes to their poor performance. This study therefore concludes that the students have inadequate knowledge on the normal distribution.

5.2.3 Can students obtain the mean given a normal distribution equation?

From the test results, none of the students solved an application question on the normal distribution to obtain the mean (Table 4.4).

Students were required to find the mean, μ given that the paint discharge is normally distributed with standard deviation = 0.4cm, where the random variable, say $X > 6cm$ and the probability of the random variable is $1\% = 0.01$.

The value of the mean, μ is determined such that, $P(X > 6) = 0.01$, by letting X be the amount of dye discharged into the can, meaning $X \sim N(\mu, 0.4)$

And therefore it means that,

$$0.01 = P(x > 6) = P\left(\frac{x-\mu}{0.4} > \frac{6-\mu}{0.4}\right) = P\left(z > \frac{6-\mu}{0.4}\right)$$

Therefore, $\frac{6-\mu}{0.4} = 2.33$ (From the Standard Normal Tables)

$$\text{And } \mu = 6 - 0.4(2.33) = 6 - 0.932 = 5.068$$

Hence the mean $\mu = 5.068$

No student interpreted and solved the question correctly (see Table 4.4). The test results showed that questions on the foundational content related to the BND were not answered correctly. The normal distribution and its application questions were poorly done by the students suggesting that they lacked the adequate background knowledge necessary for effective learning of the BND. Question three of the test on the application of the normal distribution knowledge to find the mean, was not attempted by any student.

The student questionnaire results showed that the normal distribution content and its application was a problem to the students and they had not understood it well in their previous year of study. The lecturer questionnaire also revealed that students' lack of adequate knowledge on equations and lack of a firm foundation in statistics affect their learning of BND.

This study therefore reveals that students had inadequate knowledge of how to obtain the mean given a normal distribution equation. This study agrees with McPhan and Pegg (2008) studies who noted that students' inadequate background in mathematics leads to difficulties in learning mathematics in higher grade levels of education. Hence, it can be concluded that the students cannot obtain the mean given a normal distribution equation.

5.2.4 Can students differentiate a given exponential equation?

Many of the students (see Table 4.5) correctly solved the exponential differentiation question. Few students were unable to solve the question. Exponential differentiation skills are necessary for analysis of BND equation. The question on exponential differentiation was done well by the students, (see Table 4.5), which suggests that some of them had adequate knowledge on exponential functions.

On letting $u = -4x$ and differentiating u with respect to x , ($du = -4dx$), the students were unable to substitute the result in the equation, $dy = \frac{1}{8}e^u du$, and also substituting back $u = -4x$. The students should then have obtained the result as, $\frac{dy}{dx} = -\frac{1}{2}e^{-4x}$.

In Figure 4.4, a student wrongly divided by -4 which is not the chain rule method required to correctly solve the question. This further demonstrates that some students had inadequate differentiation skills which hampers their effective learning of BND. In the analysis of the BND equation, to obtain the conditional mean and variance equations, knowledge of both

differentiation and integration of exponential functions is essential. It is worth noting that differentiation and integration of exponential equations is learnt in calculus I and II in the first year at the university. The exponential differentiation is mainly applied in the analysis of a BND equation, to obtain the conditional mean and variance equations. The research results obtained from the test analysis showed that majority of students scored well in the exponential differentiation question (see Table 4.5). It is also evident that students had adequate knowledge on differentiation since the majority attempted the differentiation question, even if some did not solve the question correctly to obtain the correct final result.

The student questionnaire results showed that the students had adequate knowledge of differentiation skills (see Figure 4.24) for learning of BND. This study agrees with Mugisha (2012) who also found that students need to have a firm grounding of concepts and skills in first year calculus courses to enable them fully understand calculus from first year to fourth year at university.

This further explains that the students generally had adequate knowledge on the differentiation of exponential function which is the basic knowledge applied in the analysis of the BND equation.

5.2.5 Can students integrate a given exponential equation?

Exponential integration skills are required in the analysis of a BND equation to obtain the conditional mean and variance equations. Hence, knowledge of integration of exponential functions is foundational to analysis and effective learning of the BND equation.

The students' performance on the question on integration of exponential function (see Table 4.6) indicates that the majority of students solved the exponential integration question correctly. It therefore means that the students had adequate foundational knowledge on integration of the exponential functions which is applied in the analysis of BND equation to obtain the conditional mean and variance equations.

From the student questionnaire, it emerged that the students had adequate knowledge of integration skills (see Figure 4.25) for learning the BND. It can therefore be concluded that students had adequate foundational knowledge on the integration of the exponential functions which is applied in the analysis of bivariate normal distribution equation, and the students can hence integrate a given exponential equation. Findings of this study compares to that of

Mugisha (2012) who found that first year calculus is a pre-requisite course required in understanding further studies in mathematics.

5.2.6 Can students state the bivariate normal distribution equation?

This question tested the students' ability to state the BND equation. From the BND test, 102 (58.3%) students (Table 4.7) were unable to state the BND equation. Some of them stated wrong equations and others stated incomplete BND equation. The BND equation is lengthy for some students. The equation has five parameters and two constants for the two independent random variables.

From the student questionnaire results, students indicated that the BND equation is complex and long to comprehend, hence understanding the equation is difficult. Also, from the student questionnaire, it emerged that the students lack interest in learning BND (see Table 4.12) and they have a negative attitude (see Table 4.21) towards learning the BND. It also emerged that the BND equation is long (see Table 4.19) and the BND equation is complex (see Table 4.20) for students.

The lecturers also indicated that it was difficult for students to understand the probability density function of the BND. The lecturers indicated that the BND is rather complex and that it requires more time to go through the content to understand it well. Also, from the lecturers' questionnaire, it also emerged that students lack adequate pre-requisite background knowledge on learning the BND. Lecturers also indicated that students have difficulties in understanding the probability density function of the BND.

Analysis of the BND equation is long and requires many steps bearing in mind that the equation has many parameters and constants of the two random variables, which makes its learning difficult for some students. Studies, for example, Ben-Zvi and Garfield (2004) indicate that many statistical ideas and rules are complex, difficult, and/or counterintuitive and it is difficult to motivate students to engage in the hard work of learning statistics, which is consistent with the findings of this study.

Most students believe that the longer the equation, the more complex the equation is and thus the more difficult it is to learn and understand the content pertaining to the equation. Some students just look at the length of the equation, and not what the equation contains.

For an equation to be learnt and understood well, it is important that students understand all the parameters and constants in it. The BND equation has five parameters (the expectation of

the two random variables, the standard deviation of the two random variables and the correlation coefficient) and two constants (π and the exponent). All five parameters and the relationship between them must be understood by all the students for effective and meaningful learning of the content to take place. Some students cannot show the relationship between the parameters in the equation and hence it becomes difficult to comprehend the BND equation. It can therefore be concluded from the findings of the test, the student questionnaire and the lecturer questionnaire that the students cannot state the BND equation. Hence, it poses a problem to the students' learning of bivariate normal distribution.

5.2.7 Can the students calculate the mean, $E(y/X = x)$ of the joint distribution, given the conditional distribution of a variable (e.g. $h(y/X = x)$)?

One of the test questions examined the students' ability to calculate the conditional of the joint distribution given the conditional distribution of a variable. 171 students (97.7%), were not able to calculate the mean. Table 4.8 illustrates the difficulties the students encountered in calculating the mean.

One student's solution, Figure 4.7a, illustrates how wrongly the student calculated the expectation of a variable and yet the variable was given as the mean in the question. From the test results, the students (see Table 4.8), left the question blank suggesting that they could not calculate the conditional mean of two random variables. This study findings agrees with Shojaie, Aminghafari and Mahammadpour (2012) studies who also found that students have difficulties in introductory courses in probability and statistics which include joint distribution and calculation of bivariate expectation and complex integrals..

The derivation of the conditional mean and variance equations is long and the steps involved in the analysis of the BND equation are numerous and complex.

From the student questionnaire, it emerged that students do not understand the step-by-step analysis of the BND equation (see Table 4.28), hence its application and use is difficult for them.

From the findings, it also emerged from the lecturers that the derivation of the conditional mean equation in the analysis of the BND equation and the applications of the equation is difficult for students.

The lecturers also indicated that the derivation of the conditional mean equation in the analysis of the BND equation and the applications of the equations is difficult for students.

In conclusion, it was found that the students generally could not calculate the mean of a joint distribution given the conditional distribution of a random variable.

5.2.8 Can students calculate the variance, $Var(y/X = x)$ of the joint distribution given the conditional distribution of a variable (e.g. $h(y/X = x)$)?

This question tested the students' ability to calculate the conditional variance of the joint distribution given the conditional distribution of a variable. From the BNDT, only 2 students calculated the conditional variance of the joint distribution. A total of 97.7% of the students were unable to calculate the variance (see Table 4.28).

Table 4.8 illustrates the difficulties the students encountered in calculating the variance. This suggests that the students could not solve the BND question on the conditional variance of two random variables.

The derivation of the conditional mean and variance equations is long and the steps involved in the analysis of the BND equation are numerous and complex. The student questionnaire findings, indicates that students do not understand a step-by-step analysis of the BND equation (see Table 4.28).

From the findings, it emerged from the lecturers that the derivation of the conditional mean and variance equations in the analysis of the BND equation and the applications and use of the equations is difficult for students.

Therefore, it can be concluded that the students generally could not calculate the variance of a joint distribution given the conditional distribution of a variable.

5.2.9 Can students calculate the mean of the two independent variables, given a bivariate random density function of a variable?

The students in general (96.0%) could not obtain the means of the two independent variables, given a bivariate random density function of a variable; they could not equate and relate the given bivariate random density function to the general BND equation, to obtain the two means

of the two independent variables (Table 4.9). This is in agreement with the study by Bunker (1969), who argued that the ability of a student to see and understand relationships is a key to successful problem solving in mathematics.

Figure 4.8 demonstrates student's weakness in solving the means of a bivariate normal density function. The student questionnaire results indicated that the bivariate normal distribution equation is complex with its parameters and constants of the independent variables and hence difficult to learn and understand.

From the lecturers' questionnaire, it emerged that the bivariate normal distribution is rather complex and that it requires more time to go through the content to understand it well. They indicated that students have difficulties in learning and understanding the BND.

It can therefore be concluded that the students could not obtain the means of the two independent variables by equating and relating the given bivariate random density function to the general BND equation.

5.2.10 Can students obtain the standard deviations of the two independent variables, given a bivariate random density function of a variable?

None of the students solved the test question on obtaining the standard deviations of the two independent variables, given a bivariate random density function of a variable (see Table 4.10). Infact, most of the students left the question blank. The students could not equate and relate the corresponding equations and variables of the given bivariate random density function to the general BND equation, so as to obtain the two standard deviations of the two independent variables.

For example, the student was required to equate the equations and the coefficients to obtain;

$$\frac{(x - \mu_1)^2}{\sigma_1^2} = (x + 2)^2$$

And equating the coefficients of the equations;

$$\frac{1}{\sigma_1^2} = 1;$$

$$\text{Hence, } \sigma_1^2 = 1; \sigma_1 = 1$$

And similarly,
$$\frac{(y - \mu_2)^2}{\sigma_2^2} = 4(y - 1)^2;$$

Therefore;
$$\frac{1}{\sigma_2^2} = 4;$$

It thus implies that,
$$\sigma_2^2 = \frac{1}{4};$$

Hence,
$$\sigma_2 = \frac{1}{2} = 0.5$$

Table 4.10 (chapter 4) illustrates the performance of the students in solving the standard deviations of the two independent variables of a BND.

The BND equation has five parameters and two constants of two independent random variables. The students found it difficult to understand the relationship and association between the parameters in the equation. This possibly explains the students' difficulties in stating the BND equation.

These findings are consistent with those of Shojaie, Aminghafari and Mahammadpour (2012) who acknowledged that students have difficulties in introductory courses in probability and statistics which include joint distribution and calculation of bivariate expectation and complex integrals.

From the student questionnaire, it emerged that the analysis of the BND equation and the derivation of the conditional mean and variance equations are difficult for students.

It also emerged from the lecturers that the students find it difficult in the analysis of the BND equation and the derivation of the conditional mean and variance of the BND equation and the applications of the equations. The above observations are therefore evident reasons why students could not calculate the two standard deviations of the two independent variables. This study compares to Garfield (1995), who found that teaching statistical courses is challenging since they serve students with different backgrounds and abilities, many of whom have had negative experiences with statistics and mathematics.

It can therefore be concluded that the students could not obtain the standard deviations of the two independent variables from the given bivariate random density function.

5.3 REASONS WHY STUDENTS FIND THE LEARNING OF BND DIFFICULT

The results from the students and lecturers indicate that students experience difficulties in learning of a BND. There are many reasons to explain why students encounter such difficulties in the learning of BND.

From the research findings, students indicated that long and complex equations negatively affect their learning of the BND content. The lecturers also indicated that it is difficult for students to understand the probability density function of the BND because it is complex and requires extra time to go through the content to understand it well. These findings are consistent with Ben-Zvi and Garfield (2004) who showed that many statistical ideas and rules are complex, and it is difficult to motivate students to engage in the hard work of learning statistics. This attestation is in line with the BND test taken by the students, where questions seven and eight on the BND were answered very poorly. Not only could students not state the BND equation but they also could not solve the two means and standard deviations from a given bivariate normal density function. It is therefore evident that the students' wrong perception on long equations, which students regard as complex, negatively affects the student learning of that content.

Also the study indicated that inadequate mathematical knowledge on equations related to the BND equation, for example, the normal distribution equation, negatively affects the students' learning of BND content. For effective learning of the BND content, students must be equipped with adequate knowledge on the foundational concepts related to BND before effective learning of the content can take place. A student must have foundational knowledge which includes measures of central tendency, the normal distribution, differentiation and integration of exponential functions, for effective learning of BND to take place. The research study findings also showed that students lacked adequate foundational knowledge and skills necessary for solving application statistical questions on the normal distribution and the BND. It was also found from this study that students lack adequate learning resources (see Table 4.15). Inadequate teaching and learning resources in a learning institution hinders effective learning and teaching of mathematics from taking place. These study findings compare with those of Mbugua et al (2012) and Ogbonnaya and Mji (2013), which showed that a lack of adequate learning and teaching resources impacts negatively on teaching and students' learning. When the learning resources are limited in an institution, students experience difficult times in terms of referencing, research and doing assignments.

Continuous assessment tests and the end of semester examinations, on the BND have a negative influence on the students' learning of the BND. The content on the BND is short but involving in terms of learning and understanding all the concepts within a limited time frame. Students believe that the content examined by the lecturers is difficult. The students indicated that lecturers solve easy examples while teaching the BND content in class (Table 4.29) but later the lecturers examine difficult content in the examinations.

Students weak background in mathematical equations was also implicated in this study as a reason for the difficulties students encounter in learning the BND. From the student questionnaire, the students indicated that they did not understand the normal distribution in the first year. The lecturers also indicated that the students' lack of a firm background and pre-requisite knowledge on BND, negatively affects student learning of the BND. As revealed by Ogbonnaya and Mji (2013), students' weak background in mathematics contributes to their poor performance and while their success in higher grade levels in school depends to a large extent on how well they had been prepared at the lower grade levels. Other studies, for example, Ben-Zvi and Garfield (2004), showed that students' difficulties with the underlying basic mathematics interferes with learning the related statistical concepts. The pre-requisite knowledge on the marginal densities, the normal distribution, differentiation of exponential functions and the application of integration on marginal densities is a basic necessity that students should have for effective and meaningful learning of the BND. The students' learning and understanding in equations, in particular mathematical content, hugely depends on the background knowledge and experience the students have in mathematical equations. A student with adequate background in solving equations finds learning other advanced mathematical equations easy. The performance in BND tests and examinations solely depends on the understanding the student has on the background knowledge in mathematical equations. When a student has insufficient foundational knowledge on simple equations with few parameters and constants, the student finds it difficult to learn other long equations with many parameters and constants. A student must understand all the basic equations and the relationship between the parameters and constants in the equation, so as to understand further equations in advanced related knowledge.

However, many students forget most of the equations once they are finished with the unit containing the equation because the content is no longer tested in future examinations after the semester ends in university. For example, a normal distribution equation has one variable, two parameters and two constants which a student is required to understand before learning

the BND equation. The product of two marginal distributions, say $f(x)$ and $f(y)$, of any two independent variables, leads to the formation of a joint probability density function which is the BND function, $f(x, y)$. This means that a student should be well equipped with adequate background knowledge on equations so that the student can effectively learn the BND.

The students indicated that they do not understand the step-by-step analysis of the BND equation to obtain the conditional mean and variance of a variable given the other variable. This findings can be compared to that of Shojaie, Aminghafari and Mahammadpour (2012) who showed that students have difficulties in introductory courses in probability and statistics which include joint distribution and calculation of bivariate expectation and complex integrals. The stepwise analysis of the BND requires the student to have foundational knowledge on the expectation of a variable (mean), the variance of a variable, the correlation coefficient of two independent random variables, normal distribution, the probability density function, the marginal distribution of a variable and the exponential differentiation and integration of equations. It is important for students to understand the analysis of a BND since it enriches them with all the basic knowledge on the parameters and constants involved in the derivation. It also enables the student to understand the application and use of the derived equations.

It was found from this study that too many students in a class affects the students learning of BND. The students indicated that some of their statistical classes, for example, BSc classes, have many students and that affects their learning of statistics content. For any effective learning of given content to take place, the number of students in a statistics class should not be too high. A class of many learners hinders effective content delivery by the lecturer and student learning of the content. Some statistical content, for example, the BND, requires students to copy all notes meticulously with understanding of all parameters and constants, or copy all notes when the lecturers dictate. Overcrowding of the classroom hinders the lecturers from giving individual attention to the students or even explaining concepts in full. From the research findings, lecturers stated that some BND classes had too many students, for example, the BSc. Physical Sciences whose population in first and second year is averaging more than eighty students. Hence, learning BND is difficult for students since the lecturer cannot give individual attention to all of them.

According to the lecturers, students find it difficult to obtain marginal density equations of a variable and that makes learning of BND difficult. The product of two marginal distributions, say $f(x)$ and $f(y)$, of any two independent variables, leads to the formation of a BND

function, $f(x,y)$. Therefore, if students have difficulties in obtaining the marginal distributions of one variable, it is then very difficult for them to solve a question on a BND.

According to the lecturers, the students' fear, their wrong perception and negative attitude towards long mathematical equations, for example, the BND equation, also negatively affects the student learning of the content. From the student questionnaire results, students indicated that they have a negative attitude towards the BND equation. The students' attitude towards given mathematical content affects their understanding and performance in examinations on the content. With a positive attitude, students can understand and perform well in all examinations testing mathematical content learnt. A student with a negative attitude towards given content will, in all likelihood, not put much effort to understand the content and he/she will not perform well if the content is tested.

It was discovered from both the lecturers' and the students' responses to the questionnaire items that students' lack of interest (see Table 4.11) in BND affect their learning of the content. From the student questionnaire results, students indicated that they lack interest in learning BND. The students have a wrong perception and lack interest in learning equations, for example, the BND equation which they consider long and complex, and therefore they do not take adequate time to understand the content of the equation. The research findings are consistent with those of Tishkovskaya and Lancaster's (2012) study, which found that students at all levels of education, lack interest in learning when taking introductory statistics courses. Some students believe that long equations are generally complex and hence learning and understanding the equations is a toll order for them and so they lose interest in learning that content.

5.3.1 Summary of the study

This study was established to investigate the difficulties students experience in the learning of BND. The study also investigated why students experience such difficulties in learning the BND. Three instruments were used for data collection i.e. the test for students, the student questionnaire and the lecturer open-ended questionnaire. The data collected was analyzed in statistical software to find the difficulties students experience in learning the content.

Results revealed that the students performed poorly in the BND test where they achieved a mean mark of 29.3% out of the possible 100 marks. The research findings showed that students have difficulties in the learning of BND.

The results from the two main research questions and sub-questions were summarised as follows:

5.3.1.1 Research question one (1): What do students find difficult in the learning of bivariate normal distribution?

The results showed that students had difficulty with the following:

- i. Obtaining marginal density equations of a variable of BND;
- ii. Analysing of the BND equation, in the derivation of the conditional mean and variance equations;
- iii. Understanding the probability density function of the BND;
- iv. Stating the BND equation and calculate the conditional mean and variance of a variable given the other variable;
- v. Using the conditional mean and variance equations, derived from the analysis of a BND function, to find the solution to an application question on BND;
- vi. Solving the two means and standard deviations from the bivariate normal density function; and
- vii. Learning the long BND which consists of five parameters and two constants of the two independent random variables.

5.3.1.2 Research question two (2): Why do students experience difficulties in learning the bivariate normal distribution?

From research findings, students experience difficulties in learning the BND. The following are the reasons why students experience difficulties in learning BND;

- i. The long and complex equations negatively affect the student learning of BND. The BND function has five parameters and two constants, for the two independent variables and some students find it hard to comprehend thus making learning of the content difficult.

- ii. Students' lack adequate foundational knowledge on the normal distribution which negatively affects their learning of BND.
- iii. The students' negative attitude and lack of interest on bivariate normal distribution negatively impacts on learning the BND.
- iv. The students' lack of adequate basic knowledge and skills hampers effective learning of a bivariate normal distribution. The background knowledge on the normal distribution is a pre-requisite knowledge necessary for learning of bivariate normal distribution.
- v. The students' weak background in mathematical equations hampers their learning of bivariate normal distribution.
- vi. Lack of adequate learning resources in a learning institution e.g. reference books, hampers the effective learning of bivariate normal distribution.
- vii. Lecturers' testing of difficult content on the bivariate normal distribution impact negatively on learning of the BND.
- viii. Students' believe that the BND equation is complex and long and negatively affects their learning of BND.

5.4 IMPLICATIONS OF THE RESEARCH FINDINGS

It was discovered in this study that students encounter difficulties in learning the BND. The students had major difficulties in stating the normal distribution; finding the probability of a normal distribution variable; calculating the mean for a normal distribution function; stating the bivariate normal distribution; using the conditional mean and variance equations; and, finding the two means and standard deviations of a BND.

The poor performance in the BND test was a clear attestation of the weaknesses and difficulties students have in learning of a BND. The students had inadequate foundational knowledge in the normal distribution and its applications which forms the basis for effective and meaningful learning of the BND. A common area of difficult in this study was that students were unable to use the conditional mean and variance equations to solve the application question and also the inability of students to use the BND equation to determine the two means and the two standard deviations of the two independent random variables.

The implications of the findings of this study to the Kenyan university, lecturers and students at large are as follows:

A firm statistical foundation is required to enable students to effectively learn further concepts in statistics with ease.

The methods of teaching employed in teaching statistics in the university which includes lecture and chalkboard illustrations methods should be incorporated with integration of information communication technology (ICT).

Lecturers' should consider changes in student' assessment format to include continuous assessment assignments which are open book problem-solving to help overcome the burden of memorizing and regurgitating long and complex equations.

Classes for statistical courses in the university should not exceed 30 students. This is to enable lecturers to cater for individual difficulties the undergraduate students experience in learning of statistical courses, for example, the BND.

5.5 CONCLUSION

From the findings of this research, it is evident that students in a BND class and statistics in general, have difficulties in the learning of BND. The students have difficulties in the analysis of the BND equation, obtaining the marginal density equations, the derivation of the conditional mean and variance equations, and understanding the probability density function of the BND. These difficulties are due to: students' inadequate foundational knowledge on BND and other related content, students' lack interest, wrong perceptions and their negative attitude towards long equations.

5.6 RECOMMENDATIONS

Based on the findings of this study, the following recommendations are made.

There is need for students to develop a positive attitude towards learning mathematical content e.g. mathematical equations which seem long and complex when one is introduced to the equation. Students also need to work towards understanding the contents of the equation,

for example, all the parameters and constants, to enable them understand what they learn in class.

The fact that students experience difficulties learning BND, gives credence to the view that; they should research more on what they learn in class. Students should to make maximum use of their institution's library to learn more about what they are taught in class. By referring to several books with similar content, students enrich themselves with the knowledge they acquire from books which enables them to better understand the content. Also, there is need for students and lecturers to utilise free internet resources provided in the university to bridge their statistical content knowledge and pedagogical knowledge gaps respectively. For example, students with difficulties in BND and statistics at large can enhance their understanding and boost their confidence in statistics through the use of online resources such as notes, you-tube videos, mathworld etc.

There is a need for students to be discouraged from memorizing content; but rather, the students should be encouraged to understand the content. For example, students should be encouraged to understand long equations with their parameters and constants, for example, the BND equation rather than cramming the equation. The learning of the equation becomes easier when the equation is understood than when it is memorized.

The fact that students have inadequate foundational knowledge necessary for learning the BND implies that there is need for lecturers to revisit all the foundational topics relating to the BND before teaching its content. This should be done to acquaint students with the background content on the topic, for example, the parameters and constants. There is need for the lecturer to revise content on the normal distribution, its variables and constants, exponential differentiation and integration before the actual teaching of the BND begins.

There is need for lecturers to exhaustively explain to the students the different parts of the BND equation to aid the students' learning. All the parameters and constants in the equation need to be explained to students to enable them to thoroughly learn and understand the equation. Also the lecturer should explain to the students how to obtain the joint BND equation from two normal distributions of two independent variables.

There is need for university to consider changing the curriculum so that students are examined in all the statistics units that they learn from first year to fourth year, in an examination referred to as an end course examination. This would ensure that students consistently and

adequately learn the statistical units they have been taught earlier, in order to always have the adequate foundational knowledge necessary for learning other advanced statistical units while at university.

The researcher also recommends that the university should consider teaching the BND content under the multivariate distributions unit in statistics, in the third year of the students' study as an introductory course to the unit.

5.7 LIMITATIONS OF THE STUDY

Few students are admitted and register for statistics courses with content that includes BND. Hence the sample size was not as large as one would have desired.

Some of the students did not show up for the research study although they had been made aware by their respective class representatives and the mathematics department, of the date and time, and the room where they could assemble to participate in the research.

Only a few lecturers in the departments of mathematics participated in the study. The lecturers who were present on the day of data collection completed the questionnaires. Hence sampling of the lecturers who participated in the study was difficult.

5.8 SUGGESTIONS FOR FURTHER STUDIES

The following suggestions have been put forward to enhance further studies in this field:-

The present research study did not consider in a major way the effects of a number of variables. These variables include- the gender of students, their reading ability, the past student performance in statistical units and their university entry grade, namely the grade in mathematics obtained by candidates who sat for the Kenya Certificate of Secondary Education. Inclusion of any or all of these variables could have influenced the findings of this study. Future studies should consider all these variables with a larger sample size.

Finally, future research should consider involving many Kenyan universities and campuses in order to have a larger population from which a greater sample size of both students and lecturers could be selected.

BIBLIOGRAPHY

- Anthony, G. & Walshaw, M. (2009). Characteristics of Effective teaching of mathematics. A view from the West. *Journal of Mathematics Education*, 2(2) 2009 pp. 147-164.
- Batanero, C., Burrill, G., & Reading, C. (2011). Teaching statistics in school mathematics. Challenges for teaching and teacher education: A Joint ICMI/IASE study. Springer [in press].
- Bell, A. N., Costello, J. & Kiichemann, D. (1983). A review of research in mathematics education. Kneel: Nelson.
- Ben-Zvi, D. & Garfield, J. (2004): The challenges of developing statistical literacy, reasoning and thinking. The Netherlands: Kluwer Academic.
- Bless, C. & Achola, P. (1988). Fundamentals of social research methods: An African perspective, Lusaka: Government Printer.
- Bunker, A. R. (1969). Understanding and teaching mathematics in the primary school (2nd Edn). Sydney: Halstead Press.
- Cobb, G. W. & Moore, D. S. (1997): Mathematics, Statistics, and Teaching. *JSTOR*, 104 (9), 801-823.
- Cresswell, J. W. (2007). Qualitative inquiry and research design, choosing among five approaches, 2nd edition, California: Sage.
- Creswell J., W. (2012). Educational research. Planning, conducting and evaluating quantitative and qualitative research, (4th Edn). – Pearson Education.
- Creswell, J. W. & Plano C. V. L. (2011). Designing and conducting mixed methods research (2nd ed.). Thousand Oaks, CA: Sage.
- Das, R. C. (1985). Science teaching in schools. India: Sterling
- Davis R. B. (1984). Learning mathematics. Ablex, Norwood, New Jersey.
- De Vos, A. S. (2002). Research at Grass root: For the social science and human service professions Pretoria: Van Schaik.

Du Plessis P., Colney, L. & Du Plessis, E. (2007). Teaching and learning in South African schools. Pretoria, South Africa: Van Schaik.

Eshiwani G. S. (1983). Research in education, Nairobi: Kenyatta University

Fennema, E. & Franke, M. L. (1992). Teachers' knowledge and its impact. In D. A. Grouws (Ed), Handbook of research on mathematics teaching and learning: A project of the National Council of teachers of mathematics, 147-164, – New York Macmillan.

Fuller E. & Alexander C. (2005). Effects of teacher qualifications on student achievement in middle school mathematics in Texas. Paper presented at the American Educational Research Association, annual meeting, April 2005, Montreal. Canada

Garfield, J. (1995): How students learn statistics; International Statistical Review, 63(1), 25 - 34

Gay, L. R. (1990). Educational Research, Competencies for Analysis and Application. 3rd Edition. New York: Macmillan Publishing Co.

Gay, L. R. (1992). Educational research: Competencies for analysis and application. 4th Edition. Florida, FL: Macmillan.

Gay, L. R. (1996). Educational research. Competencies for analysis and application. (5th Edn). New Jersey: Prentice-Hall.

https://www.google.com/search?q=properties+of+a+bivariate+normal+distribution+in+three+dimensional&source=lnms&tbn=isch&sa=X&ved=0ahUKEwioieOpvYHWAhWmHsAKHRM1BgIQ_AUICygC&biw=1366&bih=648#imgrc=ANYt39g6aHecRM:

Accessed on 6/02/2017.

Joan, J. C. (2009). Reseach population. Retrieved from <http://www.explorable.com/research-population>

Kimberlin, C. & Winterstein, A. G. (2008). Validity and reliability of measurement instruments used in research. *American Journal of Health System Phamacy*, 65, 2276-2284.

- Koehler, M. S. & Grouws, D. A (1992). Mathematics teaching practices and their effects, In D. A. Grouws (Ed.), Handbook of research on mathematics teaching and learning, 115-125, New York, Macmillan
- Ingvarson L., Beavis A., Bishop A., Peck R. & Elsworth G. (2004). Investigation into Effective Mathematics Teaching and Learning in Australian Secondary Schools, Research report. Australian Council for Educational Research.
- Lunsford, B. R. & Lunsford, T. R. (1995). Research forum: The research sample, part 1: Sampling. JPO, vol. 7, No. 3.
- Mbugua, Z. K., Kibet, K., Muthaa, M. G. & Nkonke R. G. (2012). Factors contributing to students' poor performance in mathematics at Kenya Certificate of Secondary Education in Kenya: A case study of Baringo County, Kenya. American International Journal of Contemporary Research, 2 (6).
- McMillan, J. H. & Schumacher, S. (2001a). Research in education: A conceptual introduction. (5th Edn). New York: Longman.
- McMillan, J. H. & Schumacher, S. (2001b). Research in education: A Conceptual Introduction. New York: Harper Collins.
- McMillan, J. H. & Schumacher, S. (2006). Research in education: Evidence-based Inquiry, London: Pearson Education.
- McPhan G. & Pegg J. (2008). Research paper: The option of selecting higher-level mathematics courses. Transitional Tensions
- Michaels, W. C. & Smit, J. H. (2008). Problem-based learning as a tool for the integration of mathematics and science curriculum. Source: www.pb/2008.com/PDF/00002.pdf accessed 22/04/2015
- Mothata, S. (2000). A dictionary of South African education and training. South Africa. Hodder & Stoughton, Educational Southern Africa.
- Mugisha S. (2012). An investigation in to problem-solving skills in Calculus: The case of Unisa first-year students.

- Murali, K., Dan, C., & Dawn, K. (2004). Responsible conduct in data management. Faculty development and instructional design centre. Northern Illinois University, DeKalb, IL 60115, USA.
- Nardi, E. (2011). Teaching and Learning Mathematics at University level, Springer Science + Business Media.
- Nicette, N. G. & Marrison, R. G. (2014). Problems and difficulties encountered by students towards mastering learning competencies in mathematics.
- Nicholson, J. & Darnton, C. (2003): Mathematics teachers teaching statistics: What are the challenges for the classroom teacher? International Association for Statistical Education.
- Neuman, W. L. (2011). Social research methods. Qualitative and quantitative approaches. 5th Edition. Boston: Allyn and Bacon.
- Ng'ang'a, E. W. F. (2010). Factors leading to students' poor performance in KCSE Mathematics Examination. The case study of public secondary school in Kiambu district–Kenya
- Ogbonnaya, U. I. & Mji, A. (2013). Factors that contribute to students' poor achievement in mathematics in Limpopo South Africa. Proceedings of the London International Conference on Education, London, 4th–6th October 2013. 281-286
- Ogoli, J. (1998). A study of factors contributing to poor performance in mathematics: A case study of Kaveye Girls High School
- Pierce, L. L. (2009). Twelve steps for success in the nursing research journey. *Journal of Continuing Education in Nursing*. 40 (4), 154-162
- Rosenstein, J. G. (2005). What I have learned from many years of teaching Calculus to first–year college students, AMTNJ Conference: Keeping Math on Track: Bridging the gap between high school and College mathematics. Brookdale Community College.
- Shojaie, S. R. H., Aminghafari M. & Mahammadpour, A. (2012): On the expected absolute value of a Bivariate normal distribution. *Journal of Statistics Theory and Applications*, 11: 4, 371-377.

- Taylor-Powell, E. & Steele, S. (1996): Collecting evaluation data: An overview of sources and methods. *Program development and evaluation*, University of Wisconsin-Extension.
- Tishkovskaya, S., & Lancaster, G., A., (2012): Statistical education in the 21st century. A Review of Challenges, Teaching Innovations and Strategies for Reform. *Journal for Statistics Education*, (12)2.
- United States National Research Council (1989). "Everybody counts". A report of the future of mathematics. Washington Academic Press.
- Warkerly, D. D., Mendenhall, W. & Scheaffer, R. L. (2007): *Mathematical Statistics with Applications*. International student edition: Cengage Learning.
- White, C. J. (2005): *Research: A practical guide*. Pretoria: Ithuthuko Investments.
- Weisstein, E. W. "Bivariate Normal Distribution." From *MathWorld*--A Wolfram Web Resource. <http://mathworld.wolfram.com/BivariateNormalDistribution.html> Accessed on 10/08/2016
- Yusof, Y.M. & Tall, D. (1999). Changing attitudes to university mathematics through problem-solving. *Educational Studies in Mathematics*, 37(1): 67-82.

APPENDICES:

Appendix 1: Bivariate normal distribution test for students

INSTITUTE FOR SCIENCE AND TECHNOLOGY EDUCATION

UNIVERSITY OF SOUTH AFRICA, UNISA

TEST ON BIVARIATE NORMAL DISTRIBUTION

Instructions:

This test has eight questions

Answer all questions in the spaces provided

Use the standard normal distribution table attached

Time: 1 hour

[1] State an equation for a normal distribution function for a random variable X
(2mks)

[2] Given that, $X \sim N(1, 4)$, find $P(1 \leq X \leq 3)$ (3mks)

[3] In regulating the blue dye for mixing paint, a machine can be set to discharge an average of μ cm/can of paint. The amount discharged $N(\mu, 0.4\text{cm})$. If more than 6cm is discharged into the paint can, the shade of blue, is unacceptable. Determine the setting μ so that only 1% of the cans of paint will be unacceptable. (5mks)

[4] Differentiate the equation $y = \frac{1}{8}e^{-4x}$ (3mks)

[5] Find the integral of $\int 3xe^{2x^2} dx$ (4mks)

[6] State a bivariate normal distribution equation for two random variables X and Y.
(3mks)

[7] In a certain population of married couples, the height X ft of husband and the height y of the wife have a bivariate normal distribution with parameters $\mu_1=5.8$ ft, $\mu_2=5.3$ ft, $\delta_1=\delta_2=0.2$ inches and $\rho = 0.6$. Find the conditional p.d.f of y given that x is 6.3 ft. What is the probability that the height of the wife lies between 5.28 ft and 5.92 ft given that the height of the husband is 6.3 ft?
(6mks)

[8] Given that the bivariate random density function of y is:

$$f(x, y) = \frac{1}{2\pi\delta_1\delta_2\sqrt{1-\rho^2}} e^{-\frac{1}{1.02}\{(x+2)^2 - 2.8(x+2)(y-1) + 4(y-1)^2\}}$$

Find the values of:

a) μ_1 and μ_2 (2mks)

b) δ_1 and δ_2 (2mks)

End

Thank you for participating in the test

Appendix 2: Student Questionnaire

INSTITUTE FOR SCIENCE AND TECHNOLOGY EDUCATION

UNIVERSITY OF SOUTH AFRICA, UNISA

AN INVESTIGATION INTO UNDERGRADUATE STUDENTS' DIFFICULTIES IN LEARNING THE BIVARIATE NORMAL DISTRIBUTION: A CASE STUDY OF A KENYAN UNIVERSITY

STUDENT QUESTIONNAIRE

The main purpose of this research is to investigate the difficulties students experience in learning bivariate normal distribution. The research also aims to find ways of improving students' understanding of the bivariate normal distribution content.

This questionnaire is one of the methods used to gather information for this research study.

I would be grateful if you could spare some of your time (approximately 30 minutes) to fill in the questionnaire below.

Thanks

NYAMBANE BOSIRE ONYANCHA



P.O. BOX 4142-40200, KISII

TEL: +254722998285

Email: nyambaneo@gmail.com

48140791@mylife.unisa.co.za

This questionnaire has two sections – Sections A & B. Fill Section A with the required details and tick only the appropriate box in section B.

Section A: Personal Information

1. Gender: Male Female

2. What was your mean grade in your Kenya certificate of Secondary Education (KCSE)?

A

A-

B+

B (plain)

B-(minus)

Others (Specify)

3. What is your undergraduate major field of study?

4. Is this your first degree course you are studying?

5. What is your current year of study?

6. When were you admitted in first year?

7. What is the approximate number of students in your course?

Section B: Problem in learning the bivariate normal distribution.

To what extent is each of the following a problem in learning the bivariate normal distribution content in your university?

		Not at all	To a minor extent	To some extent	To a major extent
8.	Student interest in the content				
9.	Lack of background knowledge on the content taught				
10.	Large classes on Bivariate normal distribution				
11.	Inadequate learning resources e.g. textbooks				
12.	Lecturers' teaching methods on the content				
13.	Long and complex equations				
14.	Lecturer interest in teaching the content				
15.	Time allocated to learn the content				
16.	Lecturer preparation to teach the content				

To what extent do you think each of the following influence what you learn on the bivariate normal distribution in the university?

	Not at all	To a minor extent	To some extent	To a major extent
17. Tests and examination on the content				
18. Mathematical background on the content				
19. Lack of reference books				
20. The pass mark set in the mathematics department				

To what extent do you agree or disagree with each of the following statements?

	Strongly Agree	Agree	Disagree	Strongly Disagree
21. Long bivariate normal distribution equation affect your learning of the content				
22. The complex nature of the bivariate normal distribution equation affects your learning of the content.				
23. Students' negative attitude towards bivariate normal distribution content affects students' learning of the content.				
24. Bivariate normal distribution equation has many parameters and constants which makes its learning, analysis and hence understanding difficult.				
25. Lack of reference materials affects learning of the content on bivariate normal distribution.				
26. In first year, normal distribution equation was understood well.				
27. I have adequate differentiation and integration skills.				
28. Lecturers test difficult content on the bivariate normal distribution.				
29. Student background in mathematical equations affects the student understanding and performance on bivariate normal distribution.				
30. Lecturers and tutors use technology e.g. projectors to teach.				
31. Lecturers copy notes on the Board while teaching.				
32. Lecturers and tutors dictate notes including				

long bivariate normal distribution equations.				
33. I am able to state the bivariate normal distribution equation.				
34. I understand the step by step analysis of the normal distribution equation.				

35. In the space provided below, state any other difficulties you face in learning of a bivariate normal distribution.

THANK YOU FOR YOUR PARTICIPATION IN THIS STUDY

Appendix 3: Lecturer open-ended questionnaire

INSTITUTE FOR SCIENCE AND TECHNOLOGY EDUCATION

UNIVERSITY OF SOUTH AFRICA, UNISA

Research Study to investigate the undergraduate students' difficulties in learning the bivariate normal distribution: A case study of a Kenyan university

Lecturer Questionnaire

The main purpose of this research is to investigate students' difficulties in learning bivariate normal distribution. The research also aims to find ways of improving students' understanding of the bivariate normal distribution content. I hope this research will be helpful to the university, lecturers and mathematics students in the teaching of the content.

This questionnaire is one of the methods used to gather information from lecturers for this research study.

I humbly request your support and participation in the study. I would appreciate if you could spare some of your valuable time to answer the following questions which shall guide the research study.

Thanks

Nyambane Bosire Onyancha



P.O. BOX 4142-40200, KISII

TELEPHONE: +254722998285

Email: nyambaneo@gmail.com

48140791@mylife.unisa.co.za

LECTURER QUESTIONNAIRE

As, a lecturer, teaching statistics which includes the units on bivariate normal distribution,

1. How many years have you taught the course?

2. How many hours are allocated to the Statistics unit per week in this session?

3. What is the average size of the class or classes you are teaching in the unit?

4. Do you think students have difficulties in learning bivariate normal distribution?

5. If yes to question 4 above, what do students find difficult in learning the content?

6. In your opinion, why do the students experience difficulties in learning the bivariate normal distribution content?

Appendix 4: A Letter of request to conduct Research

Bosire Nyambane Onyancha

P.O. Box 4142-40200

Kisii

04-03-2015

To The DVC, Academic Affairs,

----- University

P.O Box -----,

-----.

Dear Sir / Madam,

RE: Permission to conduct a Masters of Science in Mathematics, Science and Technology Education Research

The undersigned is a Kenyan citizen who is a distance learning Msc student of Unisa. I am studying Msc. Mathematics Education and I am required to conduct research in Kenyan universities to fulfil the requirements for the award of degree.

The title of the research is “**An investigation into undergraduate students’ difficulties in learning the bivariate normal distribution: A case study of a Kenyan university**”. My target population is students who have done the content and lecturers teaching the content. I am kindly requesting for permission to conduct my research in your university.

I would wish to collect data as from 1st April to 15th April 2015 if I am permitted. I will be collecting both quantitative and qualitative data from the students and lecturers. The information collected shall be treated with utmost confidentiality.

You can kindly contact my supervisor should you need any information regarding my research study.

Dr. Ugorji Ogbonnaya,

University of South Africa,

Institute of Science and Technology Education (ISTE)

P.O. Box 392,

Unisa, 0003

Mobile +27737208026

I am looking forward to hear from you and hope I will be granted permission to conduct my research.

Thanks and kind regards

Yours faithfully,



Onyancha Nyambane Bosire.

(Researcher)

STUDENT NUMBER: **48140791**,



Mobile: 0722998285

Email: nyambaneo@gmail.com

48140791@mylife.unisa.co.za

Appendix 5: A letter of permission to conduct Research

File



**JOMO KENYATTA UNIVERSITY
OF
AGRICULTURE AND TECHNOLOGY**
P. O. Box 62000 - 00200, Nairobi, Kenya, Tel: +254-67 - 53-52181 - 4, 52711, Fax: +254-67 -53- 52017
Email: dvc@aa.jkuat.ac

OFFICE OF THE DEPUTY VICE CHANCELLOR (ACADEMIC)

Ref: JKU/2/003/072

30th March 2015

Bosire Nyambane Onyancha
P.O. Box 4142 - 40200
KISII

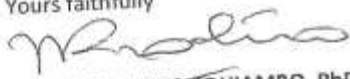
Dear Sir,

RE: REQUEST TO CONDUCT RESEARCH


Your letter on the above subject refers.

On behalf of Jomo Kenyatta University of Agriculture and Technology, I wish to inform you that the request has been granted on condition that the research findings shall be used solely for academic purposes. Please note that the area of your research is and should remain "difficulties students face in the study of Bivariate normal distribution in Statistics." Kindly avail a copy of the research finding to the University once you complete your research.

I wish you all the best as you embark on your research.

Yours faithfully

PROF. ROMANUS ODHIAMBO, PhD.
DEPUTY VICE CHANCELLOR (ACADEMIC AFFAIRS)

cc: COD - STACS



ISO 9001:2008 Certified
Setting trends in Higher Education, Research and Innovation

Appendix 6: Consent Form for Students

INSTITUTE FOR SCIENCE AND TECHNOLOGY EDUCATION

UNIVERSITY OF SOUTH AFRICA, UNISA

Research study to investigate the undergraduate students' difficulties in learning the bivariate normal distribution: A case study of a Kenyan university

Student Consent Form

Dear student,

I am a Masters degree student in Mathematics Education at the University of South Africa. I am carrying out a research on undergraduate students' difficulties in the learning of bivariate normal distribution in Kenya. The findings of the study will be used to offer suggestions on how to address the difficulties students face in learning of bivariate normal distribution.

I have obtained permission of the university to conduct the study. I hereby request you to participate in the study. You will be required to write a test examination and fill in a questionnaire if you choose to participate in the study. A test examination on the bivariate normal distribution shall take approximately 1 hour and filling of the questionnaire shall take approximately 30 minutes.

Your participation in this study is entirely voluntary. You may choose to withdraw from the study at any time. The information you provide shall be treated as confidential and your identity will not be revealed in any publication. The use of the information will be subject to standard research policies that protect the confidentiality of individuals participating in research.

For any further queries concerning this study, please contact Mr. Onyancha Bosire Nyambane (Phone number: +254 722 998 285).

Kindly sign the form to indicate that you consent to participate in the study.

Signature

Date

(Participant)

Appendix 7: Consent Form for Lecturers

INSTITUTE FOR SCIENCE AND TECHNOLOGY EDUCATION

UNIVERSITY OF SOUTH AFRICA, UNISA

Research study to investigate the undergraduate students' difficulties in learning the bivariate normal distribution: A case study of a Kenyan university

Lecturer Consent Form

Dear Prof. / Dr.,

I am a Masters degree student in mathematics education at the University of South Africa. I am carrying out a research on undergraduate students' difficulties in learning of Bivariate normal distribution in Kenya. The findings of the study will be used to offer suggestions on how to address the difficulties students face in learning of bivariate normal distribution.

I have obtained permission of the university to conduct the study. I hereby request you to participate in the study. You are required to fill in the open ended questionnaire. The filling of the questionnaire shall take approximately 20 minutes.

Your participation in this study is entirely voluntary. You may choose to withdraw from the study at any time. The information you provide shall be treated as confidential and your identity will not be revealed in any publication. The use of the information will be subject to standard research policies that protect the confidentiality of individuals participating in this research.

For any further queries concerning this study, please contact Mr. Onyancha Bosire Nyambane (Phone number: +254 722 998 285).

Kindly sign the form to indicate that you consent to participate in the study.

Signature

Date

(Participant)

Appendix 8: Description of the validity of the test questions

The validity of the test questions were as follows:

Question one of the research test required the students to state the equation of a normal distribution for a random variable X . This question was set from the probability and statistics II, the students were taught. It forms part of the foundational knowledge required by the students to effectively learn BND. The question was testing the students' ability to recall the normal distribution equation for a variable with its parameters and constants.

Question two of the test was valid since it tested the students' ability to solve $P(1 \leq X \leq 3)$ given that $X \sim N(1, 4)$ which forms the basic knowledge necessary for the learning of the bivariate normal distribution. The solve the probability of a random variable X for a normal distribution was also testing the students foundational knowledge on the mean, the standard deviation (variance) and the students' ability to use tables to solve the equation.

Question three of the BND test was testing the application of the normal distribution for a random variable in our day to day life situations. The question was testing the students' knowledge on obtaining the mean, standard deviation (variance) of a random variable from a given situation to solve a given question. This question also formed the foundational knowledge required for learning of the bivariate normal distribution.

Question four of the test required the students to solve an exponential differentiation equation. The differentiation skills of solving an exponential equation is essential in the analysis of a bivariate normal distribution equation to obtain the conditional mean and variance equations. Therefore the question was valid.

Question five of the research test tested the students' ability to solve an exponential integration equation. This question was valid since it tested the integration skills of solving an exponential equation is required by the students in the analysis of a bivariate normal distribution equation which gives rise to the conditional mean and variance equations.

Question six of the test required the student to state the bivariate normal distribution equation of two random variables X and Y . This question formed part and parcel of the researchers' area of study and it formed part of the content the students had just studied in probability and statistics III. The BND equation is regarded by many students as long, complex and difficult to

comprehend. This research question therefore tested the students' ability to recall the BND equation, its parameters and constants.

Question seven of the test required the student to use the conditional mean and variance equations, of one variable given the other variable, derived from the analysis of the bivariate normal distribution equation to solve an application question on BND. The question formed part of BND in probability and statistics III and within the course outline hence valid in this test.

Question eight tested the students ability to compare the general BND equation to the bivariate random density function of a variable given the other to obtain the two means and the two standard deviations of the two random variables. This question was valid since it was testing the comparison part of the BND and the students' ability to recall the BND equation with its parameters and constants.

Appendix 9: Marking scheme for the bivariate normal distribution test for students

University of South Africa, Unisa

Bivariate normal Distribution Test

Marking Scheme

Question 1:

State an equation for a normal distribution function for a random variable X (2mks)

Solution

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \left. \begin{array}{l} \text{1 mark} \\ -\infty \leq x \leq \infty, -\infty \leq \mu \leq \infty, \text{ and } \sigma > 0 \end{array} \right\} \text{1 mark}$$

Where $\mu = E(X), \sigma^2 = Var(X), \pi = 3.142$

μ and σ are the parameters of the distribution;

μ = the mean of the random variable X (or of the probability distribution);

σ = the standard deviation of X

Question 2:

Given that, $X \sim N(1, 4)$, find $P(1 \leq X \leq 3)$ (3mks)

Solution

$$\mu = 1 \text{ and } \sigma = 2; \text{ Since } \sigma^2 = 4$$

Then,

$$\begin{aligned} P(1 \leq X \leq 3) &= P\left(\frac{1-1}{2} \leq \frac{X-1}{2} \leq \frac{3-1}{2}\right) \\ &= P\left(\frac{0}{2} \leq \frac{X-1}{2} \leq \frac{2}{2}\right) \\ &= P(0 \leq Z \leq 1) \end{aligned} \left. \begin{array}{l} \\ \\ \end{array} \right\} \text{1 mark}$$

$$= P(0 \leq Z \leq 1) \quad \left. \vphantom{= P(0 \leq Z \leq 1)} \right\} \text{ 1 mark}$$

Therefore, $P(1 \leq X \leq 3) = 0.3413$ (From standard normal tables) $\left. \vphantom{P(1 \leq X \leq 3) = 0.3413} \right\} \text{ 1 mark}$

Question 3:

In regulating the blue dye for mixing paint, a machine can be set to discharge an average of μ cm/can of paint. The amount discharged $N(\mu, 0.4\text{cm})$. If more than 6cm is discharged in to paint can, shade of blue is unacceptable. Determine the setting μ so that only 1% of the cans of paint will be unacceptable. (5mks)

Solution

Let X = Amount of dye discharged into the can

Hence $X \sim N(\mu, 0.4\text{cm})$; Determine μ such that; $\left. \vphantom{\text{Determine } \mu \text{ such that}} \right\} \text{ 1 mark}$

$$P(X > 6) = 0.01 \quad \left. \vphantom{P(X > 6) = 0.01} \right\} \text{ 1 mark}$$

$$0.01 = P(x > 6) = P\left(\frac{x-\mu}{0.4} > \frac{6-\mu}{0.4}\right) = P\left(Z > \frac{6-\mu}{0.4}\right) \quad \left. \vphantom{0.01 = P(x > 6)} \right\} \text{ 1 mark}$$

It therefore implies that, $\frac{6-\mu}{0.4} = 2.33$ (from the standard normal table) $\left. \vphantom{\frac{6-\mu}{0.4} = 2.33} \right\} \text{ 1 mark}$

$$\text{And } \mu = 6 - 0.4(2.33) = 6 - 0.932 = 5.068 \quad \left. \vphantom{\mu = 6 - 0.4(2.33)} \right\} \text{ 1 mark}$$

Hence the mean $\mu = 5.068$, so that only 1% of the cans of paint will be unacceptable

Question 4:

Differentiate the equation $y = \frac{1}{8}e^{-4x}$ (3mks)

Solution

$$\text{let } u = 4x \quad \text{Hence } du = 4dx \quad \left. \vphantom{\text{let } u = 4x} \right\} \text{ 1 mark}$$

$$\text{hence } y = \frac{1}{8}e^{-u}$$

$$\left. \begin{aligned} \text{Therefore, } dy &= -\frac{1}{8}e^{-u} du \\ \frac{dy}{dx} &= -\frac{1}{8}e^{-u} \times 4 \end{aligned} \right\} \mathbf{1 \text{ mark}}$$

$$\frac{dy}{dx} = -\frac{4}{8}e^{-4x}$$

$$\left. \frac{dy}{dx} = -\frac{1}{2}e^{-4x} \right\} \mathbf{1 \text{ mark}}$$

Question 5:

Find the integral of $\int 3xe^{2x^2} dx$ (4mks)

Solution

$$\left. \begin{aligned} \text{let } u &= 2x^2 & \text{Hence } du &= 4xdx \end{aligned} \right\} \mathbf{1 \text{ mark}}$$

$$\frac{1}{4} du = xdx$$

$$\left. \text{Therefore } \int 3xe^{2x^2} dx = \int 3e^{2x^2} (xdx) \right\} \mathbf{1 \text{ mark}}$$

But $xdx = \frac{1}{4} du$, and on substitution, results to;

$$= \int \frac{3}{4} e^u du \quad \left. \right\} \mathbf{1 \text{ mark}}$$

$$= \frac{3}{4} \int e^u du = \frac{3}{4} e^u + C$$

$$= \frac{3}{4} e^{2x^2} + C, \text{ where } C \text{ is a constant} \quad \left. \right\} \mathbf{1 \text{ mark}}$$

Question 6:

State a bivariate normal distribution equation for two random variables X and Y. (3mks)

Solution

$$f(x, y) = \left(\frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \right) \text{Exp} \left(- \left[\frac{\left(\frac{(x-\mu_x)^2}{\sigma_x^2} \right) - \left(\frac{2\rho(x-\mu_x)(y-\mu_y)}{\sigma_x\sigma_y} \right) + \left(\frac{(y-\mu_y)^2}{\sigma_y^2} \right)}{2(1-\rho^2)} \right] \right) \quad \left. \right\} \mathbf{2 \text{ marks}}$$

Or $f(x, y) = \left(\frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}}\right)Exp\left(-\left[\frac{Z}{2(1-\rho^2)}\right]\right)$

Where $Z = \left(\frac{(x-\mu_x)^2}{\sigma_x^2}\right) - \left(\frac{2\rho(x-\mu_x)(y-\mu_y)}{\sigma_x\sigma_y}\right) + \left(\frac{(y-\mu_y)^2}{\sigma_y^2}\right)$

Exp is the exponential function

Correlation Coefficient, $\rho = Cor(x, y) = \frac{V_{xy}}{\sigma_x\sigma_y}$, $0 \leq x \leq \infty$, $0 \leq y \leq \infty$, $\sigma_x > 0$, $\sigma_y > 0$ and $-1 \leq \rho \leq$

1

Question 7:

In a certain population of married couples, the height X of husband and the height y of the wife have a bivariate normal distribution with parameters $\mu_1=5.8$ ft, $\mu_2=5.3$ ft, $\delta_1= \delta_2=0.2$ and $\rho = 0.6$. Find the conditional p.d.f y given that x is 6.3 ft. What is the probability that the height of the wife lies between 5.28 ft and 5.92 ft given that the height of the husband is 6.3?

(6mks)

Solution

Mean = $E(x/Y = y) = \mu_x + \rho \frac{\sigma_x}{\sigma_y}(y - \mu_y)$ and } **1 mark**

Variance, $Var(x/Y = y) = \sigma_x^2(1 - \rho^2)$

Standard deviation = $\sqrt{\delta_2^2(1 - \rho^2)}$ } **1 mark**

Therefore, $\mu_2 + \rho \frac{\delta_2}{\delta_1}(x - \mu_1) = 5.3 + 0.6(6.3 - 5.8)$

= 5.6 ft } **1 mark**

Standard deviation = $\sqrt{\delta_2^2(1 - \rho^2)} = \sqrt{0.2^2(1 - 0.6^2)}$

= $0.2\sqrt{(1 - 0.6^2)}$

= $0.2\sqrt{1 - 0.36}$

$$= 0.16 \quad \left. \vphantom{= 0.16} \right\} \text{ 1 mark}$$

$$P(5.28 \leq Y \leq 5.92 / x = 6.3)$$

$$= \Phi\left(\frac{5.92-5.6}{0.16}\right) - \Phi\left(\frac{5.28-5.6}{0.16}\right) \quad \left. \vphantom{= \Phi\left(\frac{5.92-5.6}{0.16}\right)} \right\} \text{ 1 mark}$$

$$= \Phi(2) - \Phi(-2)$$

$$= 1.954 - 1$$

$$= 0.954 \quad \left. \vphantom{= 0.954} \right\} \text{ 1 mark}$$

Question 8:

Given that the bivariate random density function of y is?

$$f(x, y) = \frac{1}{2\pi\delta_1\delta_2\sqrt{1-\rho^2}} e^{-\frac{1}{1.02}\{(x+2)^2-2.8(x+2)(y-1)+4(y-1)^2\}}$$

Find the values of:

c) μ_1 and μ_2 (2mks)

d) δ_1 and δ_2 (2mks)

Solution

Comparing and equating the general bivariate normal density function and the given function;

$$f(x, y) = \left(\frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}}\right) \text{Exp}\left(-\left[\frac{\left(\frac{(x-\mu_x)^2}{\sigma_x^2}\right) - \left(\frac{2\rho(x-\mu_x)(y-\mu_y)}{\sigma_x\sigma_y}\right) + \left(\frac{(y-\mu_y)^2}{\sigma_y^2}\right)}{2(1-\rho^2)}\right]\right)$$

$$f(x, y)$$

$$= \left(\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\right) \text{Exp}\left(-\left[\frac{\left(\frac{(x-\mu_1)^2}{\sigma_1^2}\right) - \left(\frac{2\rho(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}\right) + \left(\frac{(y-\mu_2)^2}{\sigma_2^2}\right)}{2(1-\rho^2)}\right]\right)$$

$$f(x, y) = \frac{1}{2\pi\delta_1\delta_2\sqrt{1-\rho^2}} e^{-\frac{1}{1.02}\{(x+2)^2-2.8(x+2)(y-1)+4(y-1)^2\}}$$

a) Find the values of μ_1 and μ_2

$$\text{Equating the related equations; } (x + 2)^2 = \frac{(x - \mu_1)^2}{\sigma_1^2}$$

$$x + 2 = x - \mu_1$$

$$\left. \begin{array}{l} \text{And hence } \mu_1 = -2 \\ \text{Similarly } \mu_2 = 1 \end{array} \right\} \mathbf{2 \text{ marks}}$$

b) Find the values of δ_1 and δ_2

$$\text{Equating equations; } \frac{(x - \mu_1)^2}{\sigma_1^2} = (x + 2)^2$$

$$\text{And } \frac{1}{\sigma_1^2} = 1; \text{ hence, } \sigma_1^2 = 1; \sigma_1 = 1 \quad \left. \right\} \mathbf{1 \text{ mark}}$$

$$\text{Also, } \frac{(y - \mu_2)^2}{\sigma_2^2} = 4(y - 1)^2$$

$$\frac{1}{\sigma_2^2} = 4; \text{ It thus implies that, } \sigma_2^2 = \frac{1}{4};$$

$$\text{Hence, } \sigma_2 = \frac{1}{2} = 0.5 \quad \left. \right\} \mathbf{1 \text{ mark}}$$

Appendix 10: The standard Normal Distribution Table

Z Table for Standard Normal Distribution										
z	0	1	2	3	4	5	6	7	8	9
0.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
0.1	.0398	.0439	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0754
0.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
0.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
0.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
0.5	.1915	.1950	.1995	.2019	.2054	.2089	.2123	.2157	.2190	.2224
0.6	.2258	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2518	.2549
0.7	.2580	.2612	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
0.8	.2881	.2910	.2939	.2967	.2996	.3023	.3051	.3079	.3106	.3133
0.9	.3159	.3196	.3212	.3239	.3264	.3299	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3969	.3888	.3907	.3925	.3944	.3962	.3990	.3997	.4015
1.3	.4032	.4049	.4066	.4092	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	.4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
1.7	.4554	.4564	.4573	.4592	.4591	.4599	.4609	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4696	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767
2.0	.4772	.4779	.4783	.4799	.4793	.4799	.4803	.4909	.4812	.4917
2.1	.4821	.4826	.4930	.4834	.4939	.4942	.4946	.4950	.4954	.4857
2.2	.4861	.4964	.4968	.4971	.4875	.4979	.4981	.4994	.4897	.4890
2.3	.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
2.4	.4919	.4920	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
2.5	.4938	.4940	.4941	.4943	.4945	.4946	.4948	.4949	.4951	.4952
2.6	.4953	.4955	.4956	.4957	.4959	.4960	.4961	.4962	.4963	.4964
2.7	.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
2.8	.4974	.4975	.4976	.4977	.4977	.4978	.4979	.4979	.4990	.4991
2.9	.4981	.4982	.4982	.4983	.4984	.4984	.4985	.4985	.4986	.4986
3.0	.4987	.4987	.4987	.4988	.4998	.4989	.4989	.4999	.4990	.4990
3.1	.4990	.4991	.4991	.4991	.4992	.4992	.4992	.4992	.4993	.4993
3.2	.4993	.4993	.4994	.4994	.4994	.4994	.4994	.4995	.4995	.4995
3.3	.4995	.4995	.4995	.4996	.4996	.4996	.4996	.4996	.4996	.4997
3.4	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4998
3.5	.4998	.4998	.4998	.4998	.4998	.4999	.4999	.4998	.4999	.4998
3.6	.4998	.4998	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999
3.7	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999
3.8	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999
3.9	.5000	.5000	.5000	.5000	.5000	.5000	.5000	.5000	.5000	.5000