Mathematical Theory and Modeling
ISSN 2224-5804 (Paper) ISSN 2225-0522 (Online)
Vol.4, No.5, 2014

Modeling e-learning implementation level for Primary Schools in Nairobi County through Stratified Random Sampling

Benard Kipyegon Bett [1], Romanus O. Odhiambo[2], George O. Orwa[3]

1. Department of Statistics and Actuarial Sciences, Jomo Kenyatta University of Agriculture and Technology, P.O Box 62000– 00200, Nairobi, Kenya

2. Department of Statistics and Actuarial Sciences, Jomo Kenyatta University of Agriculture and Technology, P.O Box 62000– 00200, Nairobi, Kenya

3. Department of Statistics and Actuarial Sciences, Jomo Kenyatta University of Agriculture and Technology, P.O Box 62000– 00200, Nairobi, Kenya

* E-mail of the corresponding author: b.kipyegon@yahoo.com

Abstract

Recent policy initiatives on education have focused on improving access to education and retaining pupils in schools through equity and enhancing quality education. However, the Government of Kenya has not unveiled detailed programme that will be a roadmap in implementation of e-learning policies in Kenyan primary and secondary schools set in the year 2006 under the Ministry of Information and Communications. The Kenya Institute of Curriculum Development (KICD), formerly KIE, has worked towards production of e-learning content and materials but the consumption of digital print materials to be produced by publishers have been given little attention in terms of provision of ICT infrastructures at school level. Thus there has been a need to carry out a research to evaluate the e-learning implementation level in Kenyan primary schools.

The findings of the research are of great benefit to scholars, government and education private developers. The mathematical equation on finite union of sets is based on the concept of set theory and probability theory models. The scholars will appreciate the utility of mathematical concepts in solving real problem wherefore the link may not be easily observed in usual instructions. On the other hand, the government may find this applied equation, a useful tool to be used in provision of ICT infrastructures at various education centres while emphasis is on equity of resources as exposed by the model. Both the government and private developers can use it to implement e-learning policies to a particular proportion or percentage.

The research was carried out in Nairobi County by drawing three independent samples using stratified random sampling strategy from private and public primary schools. Neyman allocation scheme was applied in determination of strata sample sizes per region whereas purposive sampling was applied to obtained regions of the county. Similar questionnaires were administered to selected fifty-one schools and observation used concurrently.

The research results are numeric vectors for mean of e-learning implementation level (Y) per stratum per region. The mean e-learning implementation level is greater than 40% in each of the stratum, though some regions have higher and others have lower values. From the Shapiro-Wilk test, all samples have the variable Y from non-Gaussian family except one case, prompting the non-parametric test. At 5% level of significance, Kruskal-Wallis test carried out shows that the sample per region per stratum forms a stratified population.

Keywords: Modeling E-learning, E-learning implementation level, Set and probability theory, Stratified random sampling

1. Introduction

Information and Communication and Technology (ICT) and ICT related Innovations have been rooted in every sector in Kenya. To mention a few, in banking, for example, ATMs, e-banking like electronic transfer and mobile banking all utilise the internet facility (communication) with relevant technology to offer services (required information).

The education sector has broadly utilised the innovations in ICT. This is quite evident in curriculum development bodies such as at publishers' domain where editors do contact investigations on currency of the information at hand using Internet facility, while some are already working on production of materials for electronic learning (e-learning. Many publishers also published textbooks for Computer Studies for secondary school level and relevant practical tips are included.

Most secondary and some primary teachers nationwide use computers in computing pupils' marks. Numerous secondary schools offer the course of Computer Studies to the students as an optional subject. The school must have Computer laboratory in order to offer the subject effectively since the subject in itself is practical oriented and Kenya National Examination Council (KNEC) assesses practicals other than theory part in their evaluation of curriculum exams. However, ICT and Computer usage has not been extended largely to Kenyan Primary and Secondary schools as it ought to be.

Kenyan Schools are yet to explore and appreciate the value of electronic learning (e-learning) since the government of Kenya had promised to implement it by the year 2012. The Constitution of Kenya talks of free and compulsory basic education. The mechanism to achieve it lies in the hands of policy makers; putting e-learning in place depends largely on government funding.

Private schools will be obliged to fully implement e-learning once the government of Kenya will have implemented e-learning policies in public primary schools. This is due to the contribution e-learning will impact by raising quality of education in the public schools and specifically on the improved passing in the national examinations by pupils in public schools that will force the private schools to cope up with competition that will follow. The government, can through e-learning policy regulations that will be instituted demand all primary schools have accessibility to e-learning resources. The approved law will then demand all private and public primary schools implement e-learning policies in pursuance of education policies in line to Vision 2030.

The aims of this Paper are as follows:

- Review ICT Infrastructures and stages necessary for effective implementation of e-learning.
- Assess if the education stakeholders in Kenya are prepared to implement e-learning policies.
- Evaluate e-learning implementation level (Y) from primary data and respective mean vectors in R.

2 Review of Literature

2.1 E-learning

The Organisation for Economic Co-operation and Development, OECD (December, 2005), defined e-learning as the use of information and communication technology (ICT) to enhance and support learning.

According to Kaplan-Leiserson(2003:1), e-learning covers a wide set of applications and processes such as Web-based learning, computer-based learning, virtual classrooms and digital collaboration. It includes the delivery of content via internet facility, audio and videotape, satellite broadcasting, interactive television (TV), CD-ROM, and many more.

Further, Rovai(2003), stated e-learning to consist of multiple components, e.g. learning material, learning software, academic and technical support, presentation of content and interaction. All components must work together in an efficient manner.

On the other hand, Ghirardini(2011) defined e-learning as the use of computer and internet technologies to deliver a broad array of solutions to enable learning and improve performance. Two approaches of e-learning based on the definition thus follows — Self-paced e-learning and facilitated/instructor-led. Self-paced learners are alone and completely independent while facilitated and instructor-led e-learning provide different levels of support from tutors/instructors and collaboration amongst learners. Further, blended learning combines different training media like technologies, activities and events to create optimum training programme for specific audience. Blended incorporates traditional instructor-led training with electronic formats of e-learning, Bersin (2004).

Different authors refer to e-learning based on internet using different names. To name a few, American Federation of Teachers (2000) referred to it as online learning or technology-enhanced learning or internet-based distance learning, while Fresen(2005) termed it as web-based learning or web-supported learning.

2.2 E-learning policies, implementation and research

Below is a review of how some selected countries in different continents adopted e-learning policies and executed it to achieve quality education.

The Republic of Korea (ROK) is a successful e-learning platform that can be desired by other nations. Hwang *et al* (2010) carried out a survey of ICT innovations in education in the ROK; ICT policies and initiatives including legal framework, organizational structure, budget and policy implementation process with focus on infrastructure, curriculum, teacher training, quality assurance system, monitoring and evaluation systems with global contribution at the centre of interest of the research. Globally, ICT policy in education within the ROK has been recognized as the best practice. Many organisations were set up through policies implemented that had various roles to carry out e-learning policies in ROK through their strong cooperation. Some of such organisations include Ministry of Education and Technology (MEST), Korea Education and Information Service (KERIS), and 16 Metropolitan Provincial Offices of Education (MPOEs). Since 1996, the development of ICTs within education system in ROK has been implemented under three national master plans. The first Master Plan (1996 – 2000) was focused on establishment of a world-class ICT infrastructure in elementary and secondary schools, the second Master Plan (2001 – 2005) aimed at enhancing quality of education by allowing open access to educational content and provision of teacher training for integration of ICT into classroom teaching practices, while the third Master Plan (2006 – 2010) was focused on the creation of sustainable learning environments with u-Learning and future education through more flexible and secure educational services such as the development of digital textbooks. Their overall scheme of monitoring and evaluation of ICT policy in education was on measuring ICT in education for all schools from primary to upper secondary level, ICT literacy tests for students as well as performing on external evaluation of major national ICT projects. Measuring ICT in education determined the status of ICT infrastructure, integration of ICT in classroom activities and other school-level outcomes using ICT in education. 43 performance indicators were used to measure and analyse the progress of ICT use in school education according to region and school level. The indicators were grouped into three classes, namely: input, utilization and outcome. Input covered aspect of support (human resource) and hardware/software. Hardware/software was on provision of computer equipment for teachers and student, relevant software and network environment. Utilization indicators dealt with issues of internet usage in teaching and learning activities, school homepage operation and community of practice. The outcome indicator dealt with issues of certificates of ICT, educational software contests, among others. KERIS was mandated to analyse main aspects of ICT education in ROK through use of

- Descriptive statistics to reveal the characteristics of target group.
- Analysis of what and how much the school characteristics are responsible for the overall indicators.
- Time series analysis of common indicators to compare the progress of ICT levels.

E-learning in Africa as a continent has been lagging behind for many years compared to other continents. Some African countries have worked towards implementing e-learning policies. Diallo (2012) reported that many African universities have shifted from correspondence to e-learning modes of delivery of distance education due to recent developments in technology, namely, increases in internet connectivity and access to high speed broadband connections. For example, at African Virtual University (AVU), open educational resources are being accessed in 193 countries worldwide. The 2000 Dakar Framework for Action recognizes use of ICT as a key strategy for achieving “Education for All” goals. In 2003, the New Partnership for Africa’s Development (NEPAD) prioritized efforts towards bridging the digital divide between Africa and the developed world by fast-tracking support to ICT initiatives in fostering access to education. Different African countries faced different constraints towards success of e-learning. Bappa(2012) survey findings affirmed those variations in constraints. For example, ‘Bandwidth is limited’ is major in Zambia but least likely identified constraint in Kenya, ‘Electricity is limited’ is a major constraint in Nigeria but least likely identified constraint in South Africa, among many others.

In sub-Saharan Africa region, Mauritius and South Africa have wide access to telecommunications infrastructure like fixed telephone lines, mobile cellular and international internet bandwidth. International Telecommunication Union, ITU, (2007) rated Mauritius and South Africa to have network index of 141.6 and 104.7 respectively. In South Africa, the national Department of Education, DOE, (2004) published its White Paper on E-Education focusing on policies and implementation. Useful terms related to e-learning are detailed in DOE(2004) like ICT represent the convergence of information and communication technology and are combinations of networks, hardware and software, means of communication, etc., E-learning using ICT resources, and online learning. Provincial education departments analysed delivery of e-learning at school level (e.g. KwaZulu-Natal (2006) and

Guateng(2007) were interested in Western Cape's leadership). Regulatory authorities in education like the South African Qualifications Authority and the Higher Education Quality Committee vested interest in understanding the national ICT landscape to inform more broadly on work they were to do in various institutional contexts. Many universities have their own policies on e-learning and are operational (ICTS and Higher Education in South Africa, www.cet.uct.ac.za).

In Kenya, nationally on e-learning, the Ministry of Information and Communications, MIC, (2006) set policies, regulations and strategies to govern formulation and implementation of e-learning in all levels of education system including training programmes. For example, the Government of Kenya recognizes the role played by various institutions in provisions of ICT education and training. The government is required to promote ICT school education at primary, secondary, tertiary and community levels by developing curricula and ensuring teachers/trainers possess requisite skills. Section 2.5 of MIC (2006) comprehensively address on policy framework of electronic learning. It should provide affordable access to ICT infrastructures and regulate content for different levels of education, among others. Notably, Kenyan e-learning programmes are to be for export just as developed nations and superpowers are already on it e.g. UK university distance learning/online learning programmes. Section 3.3.4 of MIC (2006) covers various strategies that will be applied to implement e-learning policies. KICD (4th July 2013) had invited a tender from publishing firms to supply digital publication (e-learning content) materials for evaluation and approval.

Some private sponsors involve in implementing e-learning in Kenya include Samsung Electronics. According to report in www.humanipo.com, Samsung Electronics has launched a pilot program named Samsung Smart School aimed at imparting skills in classroom settings for Kenya's primary school pupils. The program has been launched at Nairobi's Lavington Primary School. It uses computers and mobile phones as the ICT infrastructures. It has also rolled on e-learning implementation of policies at secondary and tertiary levels not only in Kenya but also in Nigeria and South Africa under different dubbing names.

Mwazemba (2011), in his published article, address various forms of digital revolution on publishers' side. In the article, he posted various forms of digital publishing to include multimedia affair which encompasses all formats such as: pdfs, videos, animations, audios, etc. He cited exclusive advantage of multimedia formats that they can be decoded using DVDs and TV screens (just like access to movies), or using computers. He concluded conditionally that digital publishing will only take off if publishers will be ready to invest and either train or employ editors who are abreast with digital publishing.

Otieno(2008) carried out a research on the framework for evaluating ICT use in teacher education in a case study on primary teacher training colleges (PTTC) in Kenya. The research findings were that the quantity of computer use in PTTC was 14 percent and 75 percent of time used on computers was on learning of ICT skills. This indeed is useful in assessing availability of trained personnel in regard to e-learning and ICT education at primary schools.

Odhiambo(2009) carried out a research on comparative study of e-learning platforms used in Kenyan universities and discovered that e-learning platforms in Kenyan universities were not different from one university to another. The findings revealed that downloading and uploading of e-learning lecture notes and assignments were aspects used and little use on aspects of interactivity and other modes of usability of e-learning platforms perceived to be secondary to users.

Kenyan primary schools are at initial stages of implementation of e-learning policies and a gap will be left if framework for measuring ICT infrastructures at initial stages is not developed. Robust models ought to be adopted to assess the available ICT infrastructures at initial development level and not just concentrating on productive stages for e-learning as many researchers have aimed in the past.

3 Research Methodology

3.1 Research variables

This research presents a unique strategy that links the contribution of various variables in the research. It considers technological factors, power sources, security, accommodation and electronic and storage devices that accounts for sound implementation of e-learning. Mwazemba(2011) regarded those variables as multimedia format decoders of digital publishing which directly impact the e-learning i.e. they form basis on analysis of successful and effectively implemented e-learning. Absence of such multimedia format decoders at learning environment simply implies traditional modes of learning where technological advances are mysteries to embrace.

3.2 Research Design

The research utilises descriptive research design. There is one dependent variable, namely, percentage implementation level of e-learning. The independent variables of the research constitute: Technological factors, Electronic and storage devices, Trained personnel/experts, Support services.

The research measures the numbers on all independent variables except for Support services that is recorded as available or not available. The dependent variable is then computed from the numerical data of independent variables objectively..

3.3 Population

The population of the research was all the public and private primary schools in Nairobi County, Kenya. The working population was 500 primary schools from the county of Nairobi comprising 200 public and 300 private primary schools.

3.4 Sampling Design

The working population was subdivided into three regions based on geographical locations via purposive sampling technique. Stratified random sampling (Cochran(1977)) technique was applied on the three regions created in Nairobi County by collecting data from selected private and public primary schools.

The sample size was purposively chosen to be 51 for ease of management. The strata sample sizes were calculated in accordance to Neyman allocation scheme(Nassiuma(2000)), since it is more efficient than proportional allocation scheme. As a result, 21 public primary schools and 30 private primary schools were respective strata sample sizes by assigning fixed sample standard deviation, s_i be 10 for each stratum. A simple random sampling strategy was applied in each stratum to form strata samples by selecting units (schools) from the population. Thus, the sampler assigned numbers 1-300 for private schools stratum and 301-500 for public schools. Partitioning of the population into three regions per stratum was executed and using lottery method, sampling elements were selected with sample strata sizes being major pivot factors.

3.5 Data Collection

The researcher gathered primary type of data. It was done by use of questionnaires administered to sampling elements selected in the areas of survey as the major research tools. The researcher also utilised the observation and interview guides while administering questionnaires in order to verify the primary data collected. The research assistants visited all selected schools and administer questionnaire to a teaching staff member or administrator while observing. The types of questions adopted were open, closed ended questions and likert scale questions. The structure of the questionnaire was as follows: It has three sections. The first section covered respondent's data, second section covered school information and third section covered on e-learning resources available with utility frequencies. Filled questionnaires supplied required data, which were recorded in MS Excel ready for synthesis and analysis.

3.6 Research Procedures

The research had a pre-test carried out using the prepared collection tools to check for clarity of items by administering to three different schools in the county. During the pre-test period the researcher trained hand-on personnel who were assigned to cover the stated survey regions to collect data. Necessary adjustments after pre-test were effected and the final tool was administered by the trained field personnel to the schools via hand-delivery to give room for interview and observation guides. The researcher served as supervisor over a team of trained field personnel. The response rate was high.

4 Data Analysis and Data Presentation

4.1 Probability Concepts

- Empirical or Statistical Probability – The probability of an event, P(E), is based on actual experience over a period of time by computing proportion of items that each event occurred, Kothari(2009), Meyer(1970), Murray *et al*(2009) and Parzen(1992).
- Simple event, compound event and Sample space – Simple event is an event that cannot be decomposed while compound event can be decomposed into simple events, Kothari(2009) and Feller(1968). In die-tossing experiment, six simple events form the sample space (set consisting of all possible sample points), S, e.g. face may show 1 or 2 or 3 or 4 or 5 or 6 = {E_1, E_2, E_3, E_4, E_5, E_6}= S. An event say odd number shows up is a compound event constituting of {E_1, E_3, E_5} = N_o.
- Statistical dependence, independence and mutually exclusive events – Two events are said to be ***statistically independent*** if occurrence of an event A will not affect the probability of occurrence of another event B, Kothari(2009), Wackerly *et al* (2008) and Murray *et al*(2009). A case of tossing a coin and a die are independent events. Two events are said to be ***statistically dependent*** if occurrence of an event A will affect the probability of occurrence of another event B, Kothari(2009),Wackerly *et al* (2008) and Murray *et al*(2009). A case of selecting second sampling unit where units are drawn without replacement is an example of dependent events. A set of events such that only one event can occur on one trial are said to be ***mutually exclusive events***, Kothari(2009). Mutually exclusive events cannot occur simultaneously, i.e. given event A has occurred then probability of occurrence of event B is zero. Thus mutually exclusive events form examples of dependent events and not vice versa.
- Probability axioms – Let P(A) be probability of event A in sample space S associated with an experiment(A is subset of S), Kothari(2009) and Murray *et al*(2009). The following holds:

Axiom 1: $P(A) \geq 0$.

Axiom 2: $P(S) = 1$

Axiom 3: For a sequence of pairwise mutually exclusive events A_1,A_2, A_3,… in S, that is $(A_i \cap A_j = \phi, \forall i \neq j)$.

$$P(A_1 \cup A_2 \cup A_3 \cup ...) = P(A_1) + P(A_2) + P(A_3) + ... \qquad (1)$$

By Axiom 3, finite cases of independent and dependent events are defined as follows.

$P(A_1 \cap A_2)$ $P(A_1).P(A_2)$For independent events,

$$P(A_1 \cap A_2) = P(A_1).P(A_2) \qquad (2)$$

Conditional probability for general case (dependent or independent events) is given as

$$P({}^{A_1}\!/\!_{A_2}) = \left(\frac{P(A_1 \cap A_2)}{P(A_2)} \right) \qquad (3)$$

It is read as probability of an event $P(A_2) > 0$ A_1 given that event A_2 has occurred.

The finite union for two statistically independent events A_1 and A_2 have their probability according to Murray *et al*(2009) as follows.

$$P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2) \qquad (4)$$

Similarly, for events A, B and C, Murray *et al*(2009) gave probability of union set as below.

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C) \qquad (5)$$

4.2 Concepts Data Synthesis

The primary (raw) data constituted for each school the following: Average class population, Number of classes, Number of computer and availability of internet facility, Number of televisions and DVDs, and school performance (KCPE 2011 mean score).The primary data was prepared to have three different types of sample data from the three identified regions from Nairobi County. The computer and internet facilities data was synthesized as the ratio of respective resource to the particular school mean class pupils' size (**B** = $B_i = x_i \div m_i$), the data for the number of television and DVDs was synthesized as available resource per class(**C =** C_i=**{**0 or 1}) while the power supply simply recorded available or not available assumed binary digits (**A =** A_i= {0, 1}). Hereby m_i = mean number of pupils per class in i^{th} school, i =1, 2, …, n and n = sample size for specific region of interest. Considering the variables A = power supply, B = technological factors, C = electronic and storage factors as of interest in e-learning. The variables B and C form non-adjoint sets i.e. independent non-mutually exclusive events ($B \cap C \neq \phi$).Both B and C each depend on power supply A for its operation. The random variable of interest is modeled as follows based on probability space formed in regard to elements constituting e-learning facilities.

$$Y_i = A_i \cap (B_i \cup C_i) = A \cap (B \cup C) \quad (6)$$

The Venn diagram representation of $B \cap C$ on main e-learning variables is given below.

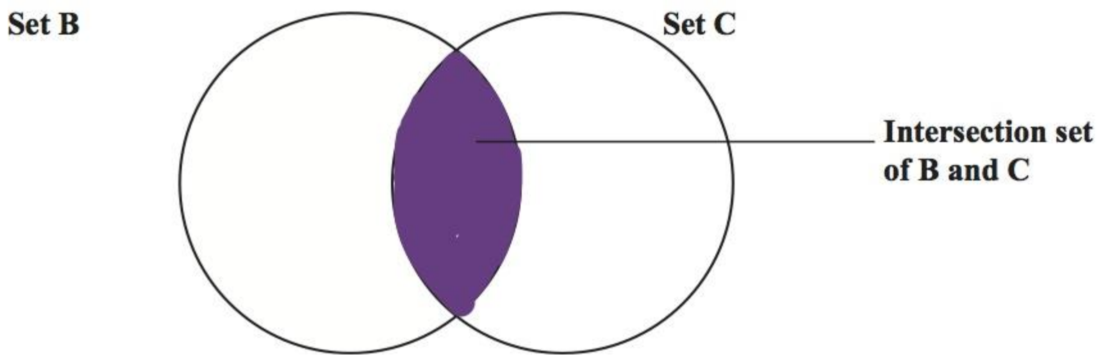

Figure 1: Union and intersection sets ($B \cap C$)

Consequently, the probability of Y_i is given as

$$P(Y_i) = P[A_i \cap (B_i \cup C_i)] = P[A \cap (B \cup C)]$$
$$= P(A \cap B) + P(A \cap C) - P[A \cap (B \cap C)] \quad (7)$$
$$= P(A).P(B) + P(A).P(C) - P(A).[P(B).P(C)]$$
$$= P(A).P(B/A) + P(A).P(C/A) - P(A).[(B \cap C)/A] \quad (8)$$

Conditional probabilities pivoted on conditional distributions of the data collected assists in evaluation of the implemented percentage level of e-learning and the percentage not implemented aids in budget projection estimate. In real sense, power supply(A) must exist for workability of e-learning infrastructures as per this model.

The consequent implemented percentage level is purely conditional as proved in equation (8) above. The techniques in analysis that the research utilised are descriptive, Bayesian and inferential statistics.

The construction of equation (7) constitutes denominator of the definition of Bayes Theorem in three variables, according to Kothari(2009) and Gelman*et al*(2008), say, A, B and C as follows:

$$P(B/A) = P(A \cap B) \div [P(A \cap B) + P(A \cap C)$$
$$= P(A \cap B) \div [P(Y_i) + P(A \cap \{B \cap C\})] \quad (9)$$

Accordingly, by Bayes Theorem, the modeled variable of interest can be computed using formulation of equation (9), whereby events B and C need not be mutually exclusive i.e. variables B and C can be present simultaneously.

For each school that form a sample, $P(Y_i)$ is generated and for n schools for a particular sub-region, mean of $P(Y_i)$) is computed using R-program. Q-Q normal plots and Shapiro-Wilk test of normality are used to test normality of $P(Y_i)$, and if the distribution of $P(Y_i)$ is not known, then a non-parametric test is carried out (Kruskal Wallis test) to test hypotheses in regard to their sample means if they were significally different from each other at 5% level of significance. The critical region is based from Chi-square distribution and hypothesis testing followed.

Thereafter, population mean is computed which allows for inferences made regarding Nairobi County. The population parameter inferred after computations, *mean of* $P(Y_i)$ is the overall implemented percentage level desired and its consequently *100% – Mean[*$P(Y_i)$*]* is the percentage level not implemented, which is useful in estimation of budget projection. The analysed data is presented in a vector form, pie charts and bar charts.

4.3 Data analysis and Stakeholder's Preparedness

Synthesis of raw data is done by using given mathematical definitions and coding so as to run stated programs in R to obtain the modeled data for P(Y) as follows. Relevant statistical analysis on P(Y) are carried out in R and all tests executed at 5% significance level.

Let x_i = Number of computers (with internet resource availability), z_i = Availability of TVs (or DVDs) and average number of pupils per class be m_i. Define TVs /DVDs and power present be 1 for Yes and 0 for No, Internet be 1 for Yes and 0 for No. Then, synthesized data are computed as follows:

A=A_i=[0,1],**C**=C_i=[0,1],**B**=B*i*={ x_i divided by m_i, $\forall m_i \geq x_i, B=1$ }and

$$P(Y_i) = P(A).P(B) + P(A).P(C) - P(A).[P(B).P(C)]$$

The row matrices for private and public schools per region in Nairobi County are presented, with useful analysis executed as follows.

- *Public Schools in Region 1*

B= (0.018867925, 0.037735849, 0.095541401, 0.087837838, 0.093023256, 0.108108108, 0.007874016), **C**= (0,0,1,1,1,1,1) and **A**=(1,1,1,1,1,1,1).

By assessment of pupil-computer ratio, B, it is evident that the highest percentage observed is 10.8108108% which indicates 10 pupils can be assigned to use a single computer in learning. In this sample, 0% have B=50%. Thus there is evidence of low level of preparedness to utilise computer in e-learning in this region from selected samples. The power supply is available in all the schools selected whereas two out of seven schools did not have television.

The e-learning implementation level, $P(Y_{11})$ denoted as Y_{11}, is obtained using R and its output is

$P(Y_{11})$= (0.01886792, 0.03773585, 1.000000, 1.000000, 1.000000, 1.000000, 1.000000).

The highest implemented e-learning level is 100% and the smallest in the sample is 1.886792%. Its sample mean, P_{11}= 0.722372 or 72.2372% and sample variance is V_{11}=0.2248385. This can easily demonstrate high level of stakeholder's preparedness, but still B is significant in assessment and should be invested in. The Shapiro-Wilk Normality test on null hypothesis that $P(Y_{11})$is normally distributed is rejected since its p-value shown below is less than 5%. It thus confirms that Y_{11}does not belong to Gaussian family.

Its Q-Q plot (Figure 2) verifies that, it does not originate from normal distribution. Notably, Shapiro-Wilk test is more affirmative than usage of Q-Q plots.

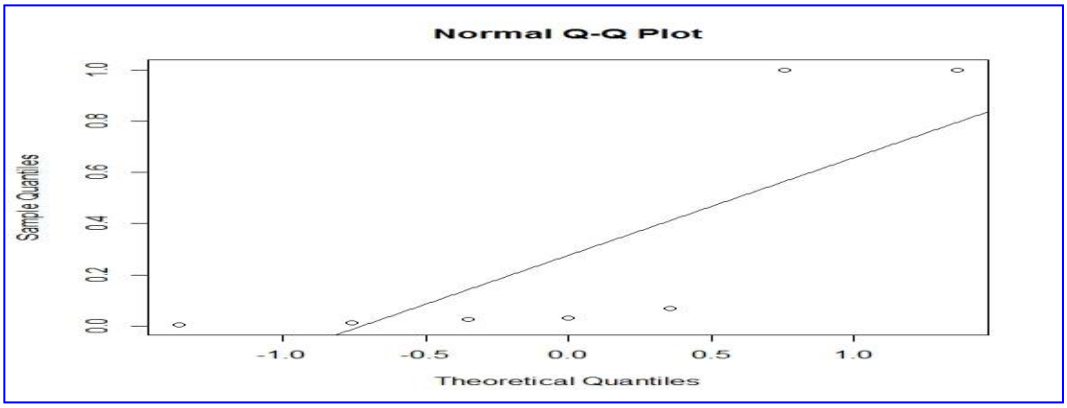

Figure 2: Q-Q Plot for Public Schools in Region 1

- *Private Schools in Region 1*

B= (0.66666667, 0.11111111, 0.06250000, 0.00000000, 0.10000000, 0.00000000, 0.07500000, 0.00000000, 0.07142857, 0.55555556), **C**= (1,0,0,0,0,0,0,0,0,1) and **A**=(1,1,1,1,1,1,1,1,1,1).

By assessment of pupil-computer ratio, B, it is evident that the highest percentage observed is 66.666667% which indicates 2 pupils can be assigned to use a single computer in learning. In this sample, two out of ten schools (20% of the sample) have B>50%. Thus there is evidence that 80% of the sample have low level of preparedness to utilise computer in e-learning in this region. The power supply is available in all the schools selected whereas two out of ten schools had television/DVDs.

The e-learning implementation level, $P(Y_{12})$, denoted as Y_{12}, is obtained using R and its output is

$P(Y_{12})$= (1.00000000, 0.11111111, 0.06250000, 0.00000000, 0.10000000, 0.00000000, 0.07500000, 0.00000000, 0.07142857, 1.00000000).

The highest implemented e-learning level is 100% and the smallest in the sample is 0%. It is evident that 20% of the sample had Y=100% which had C=1 and B>0.5 are prepared to effectively implement e-learning using B and C resources. However, 80% of the sample has low level of preparedness to implement e-learning. Its sample mean is P_{12}=0.242004 or 24.2004% and sample varianceV_{12}=V_{12} 0.1612578. The Shapiro-Wilk Normality test confirms that $P(Y_{12})$does not belong to Gaussian family at 5% level of significance. Its Q-Q plot in Figure 3 verifies that, it does not originate from normal distribution.

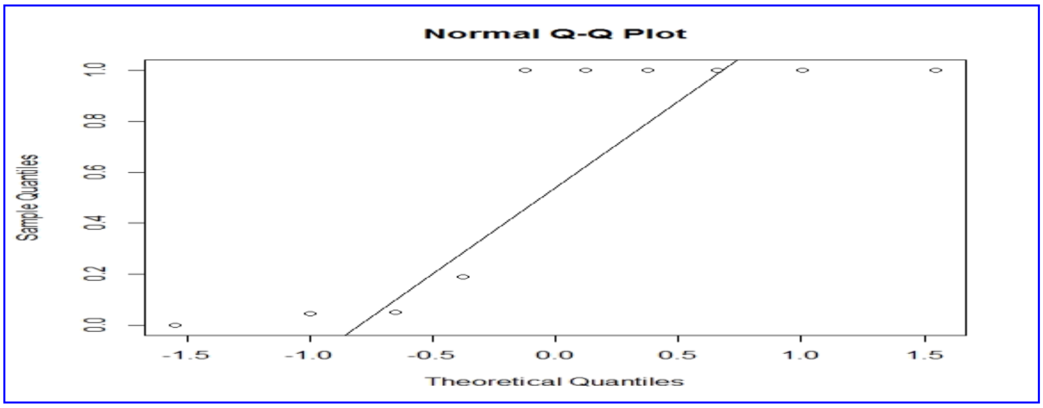

Figure 3: Q-Q Plot for Private Schools in Region 1

- *Public Schools in Region 2*

B = (0.0000000,0.0000000,0.1267606,0.1773050,0.0250000,0.0000000,0.0000000), **C**=(0,0,1,0,1,0,0) and **A**=(1,1,1,1,1,1,1).

By assessment of pupil-computer ratio, B, it is evident that the highest percentage observed is 17.7305% which indicates at least 5 pupils can be assigned to use a single computer in learning. In this sample, four schools did not have computers, whereas the other three had B<20%. Thus there is evidence that the sample had very low level of preparedness to utilise computer in e-learning in this region. The power supply is available in all the schools selected whereas two out of seven schools had television/DVDs.

The e-learning implementation level, $P(Y_{21})$, denoted asY_{21}, is obtained using R and its output is

$P(Y_{21})$ =(0.000000, 0.000000, 1.000000, 0.177305, 1.000000, 0.000000, 0.000000).

The highest implemented e-learning level is 100% and the smallest in the sample is 0%. Only two schools that had televisions/DVDs is reflected to have y=100%, whereas four schools have y=0% and one school had y<20%. In this sample, the level of preparedness is poor and a lot of investment on infrastructures is a necessity if e-learning is to be a reality. Its sample mean is P_{21}=0.3110436 or 31.10436% and sample variance isV_{21}=0.2257001.The Shapiro-Wilk Normality test confirms that Y_{21}does not belong to Gaussian family at 5% level of significance. Its Q-Q plot in Figure 4 verifies that, it does not originate from normal distribution.

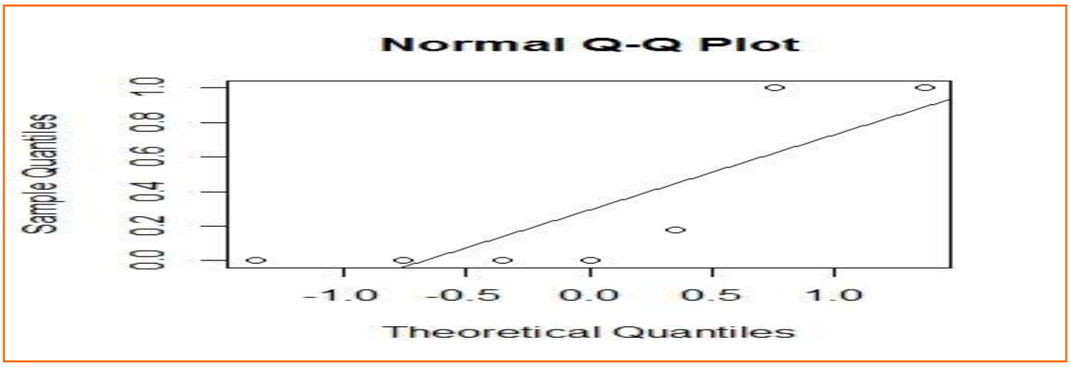

Figure 4: Q-Q plot for public schools in Region 2

- *Private Schools in Region 2*

B= (0.18779343, 0.43478261, 0.32000000, 0.17241379, 0.09523810, 0.50000000,0.03571429, 0.00000000, 0.28571429, 0.20000000), **C**= (0,0,0,0,0,0,0,0,1,1) and **A**=(1,1,1,1,1,1,0,0,1,1).

By assessment of pupil-computer ratio, B, it is evident that only one school had B= 50% , whereas one had B=0% and eight schools had 0%<B<50%. Thus, there is evidence that 90% of the sample have low level of preparedness to utilise computer in e-learning in this region. The power supply is available in 8 schools whereas two out of ten schools had television/DVDs.

The e-learning implementation level, $P(Y_{22})$, denoted as Y_{22}, is obtained using R and its output is

$P(Y_{22})$ = (0.1877934, 0.4347826, 0.3200000, 0.1724138, 0.0952381, 0.5000000, 0.0000000, 0.0000000, 1.0000000, 1.0000000).

The highest implemented e-learning level is 100% (occurring twice) and the smallest in the sample is 0% (occurring twice). Only three schools had y greater than or equal to 50%, while seven had y<50%. Thus, there is evidence that 70% of the sample had low level of preparedness to implement e-learning. Its sample mean is P_{22} = 0.3710228 or 37.10228% and sample variance is V_{22}=0.1376578.The Shapiro-Wilk Normality test gives evidence that Y_{22} belong to Gaussian family at 5% level of significance. Its Q-Q plot in Figure 5 shows that Y_{22} originates from normal distribution.

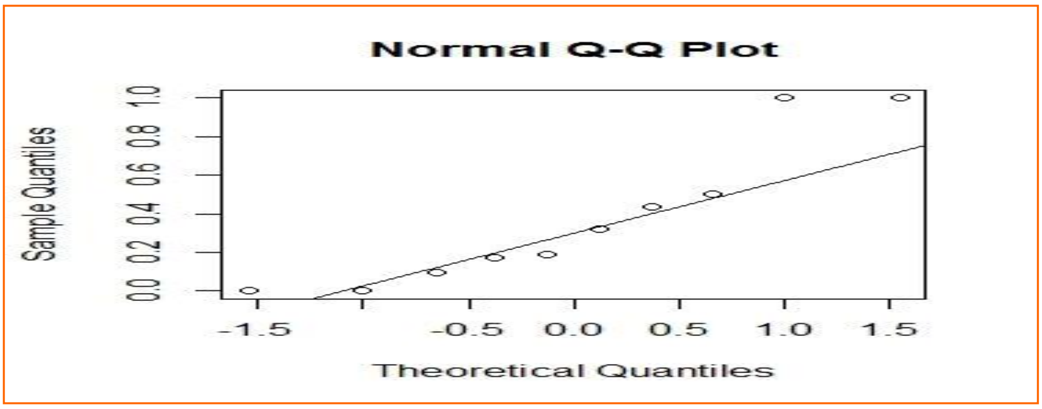

Figure 5: Q-Q Plot for Private Schools in Region 2

- *Public Schools in Region 3*

B = (0.014634146, 0.068027211, 0.005464481, 0.030000000, 0.026666667, 0.011111111, 0.011235955), **C**=(1,0,0,0,0,1,0) and **A**=(1,1,1,1,1,1,1).

By assessment of pupil-computer ratio, B, it is evident that only one school had B= 6.8% , whereas six schools had 0%<B<4%. Thus, there is evidence that 85.71% of the sample have very low level of preparedness to utilise computer in e-learning in this region. The power supply is available in all the sampled schools whereas two out of seven schools had television/DVDs.

The e-learning implementation level, $P(Y_{31})$, denoted as Y_{31}, is obtained using R and its output is

$P(Y_{31})$ = (1.000000000, 0.068027211, 0.005464481, 0.030000000, 0.026666667, 1.000000000, 0.011235955).

The highest implemented e-learning level is 100% (occurring twice as C replica) and the smallest in the sample is 0.5464481%. Thus, the level of preparedness is very low, especially in regard to computer usage which is sensitive to class size. Its sample mean is P_{31} = 0.3059135or30.59135% and sample variance is V_{31} = 0.2252189.The Shapiro-Wilk Normality test confirms that Y_{31} does not belong to Gaussian family at 5% level of significance. Its Q-Q plot in Figure 6 verifies that, it does not originate from normal distribution.

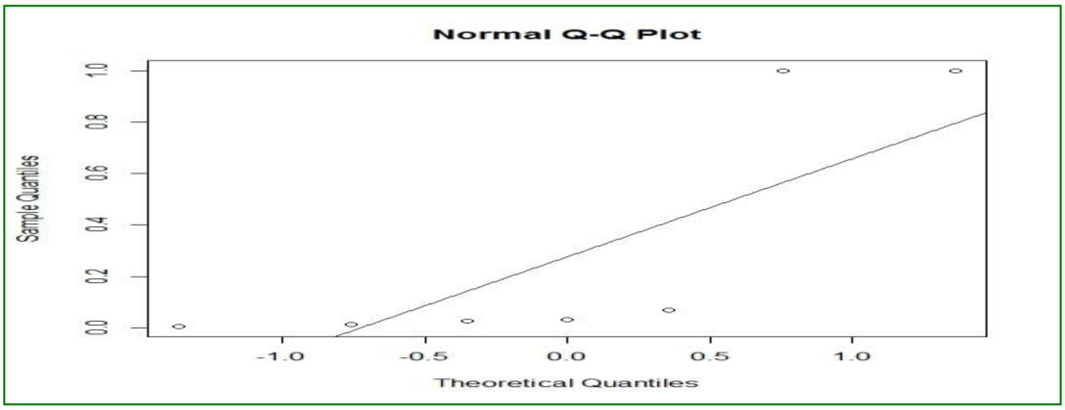

Figure 6:Q-Q Plot for Public Schools in Region 3

- *Private Schools in Region 3*

B = (0.10416667, 0.05000000, 0.00000000, 0.04545455, 0.15873016, 0.20652174, 0.19047619, 1.00000000, 0.18823529, 0.23255814), **C**= (1,0,0,0,1,1,0,0,1,1) and **A**=(1,1,1,1,1,1,1,1,1,1).

By assessment of pupil-computer ratio, B, it is evident that only one school had B= 100% (every pupil in a class can access to a computer), whereas one had B=0% and eight schools had 0%<B<25%. Thus, there is evidence that 90% of the sample have low level of preparedness to utilise computer in e-learning in this region. The power supply is available in all the selected schools whereas 50% of the schools had television/DVDs.

The e-learning implementation level, $P(Y_{32})$, denoted as Y_{32}, is obtained using R and its output is

$P(Y_{32})$ = (1.00000000, 0.05000000, 0.00000000, 0.04545455, 1.00000000, 1.00000000, 0.19047619, 1.00000000, 1.00000000, 1.00000000).

The highest implemented e-learning level is 100% (occurring in six schools due to C=1 in five schools and B=1 in one school) and the smallest in the sample is 0%. This appears high level of preparedness, but by assessing B the stakeholders have low level of preparedness towards e-learning. Its sample mean is P_{32} = 0.6285931or62.85931% and sample variance is V_{32} = 0.2321728.The Shapiro-Wilk Normality test confirms that Y_{32} does not belong to Gaussian family at 5% level of significance. Its Q-Q plot in Figure 7 verifies that, it does not originate from normal distribution.

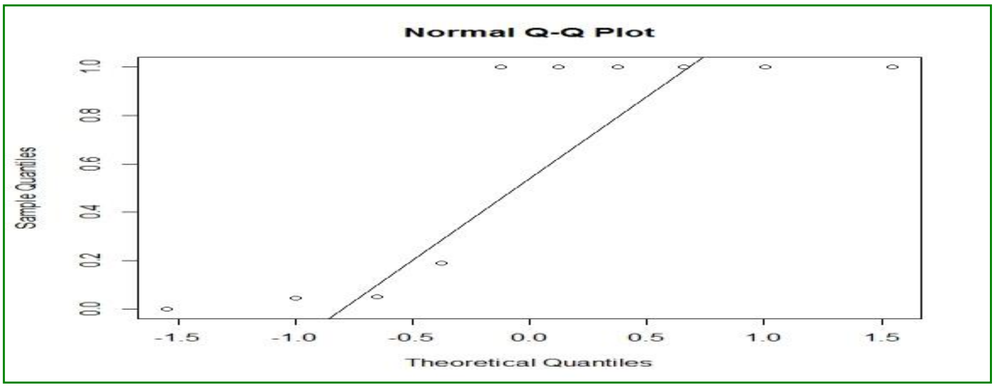

Figure 7:Q-Q Plot for Private Schools in Region 3

4.4 Hypothesis testing

The null hypothesis that all means of independent samples or regions partitioned is usually tested using One Way ANOVA test in parametric analysis. First, the test is still robust with violation of normality assumption on the dependent variable(y is skewed or follows kurtotic distributions). However, platy-kurtosis can have profound effect when the group sizes are small. Secondly, when the assumption of homogeneity of variance is violated Welch test, and Brown & Forsythe tests can be done. In case where the two key assumptions are violated, the Kruskal-Wallis H Test is appropriate (https://statistics.laerd.com/annova).Through the previous analysis, normality test for y and sample sizes for each region is small. Therefore, the non-parametric test is suitable and reliable to test the null hypothesis.

- *Kruskal-Wallis test in Public Schools*

The synthesized data for all the Public Primary Schools in the three independent regions of the county were tested for location parameters of the distribution y_1 in null hypothesis, if they are the same in each group at 5% level of significance. The R output for kruskal.test is presented below.

Table 1: Kruskal-Wallis R Output for public schools

Kruskal-Wallis rank sum test
data: Scorey1 by Method
Kruskal-Wallis chi-squared = 4.385, df = 2, p-value = 0.1116

Since the p-value = 0.1116 is greater than 5%, the null hypothesis is not rejected. This leads to conclusion that the three independent samples in Public Primary Schools stratum, all have their location parameters from the same population and population parameters can thus be deduced using sample statistics.

- *Kruskal-Wallis test in Private Schools*

The synthesized data for all the Private Primary Schools in the three independent regions of the county were tested for location parameters of the distribution y_2 in null hypothesis, if they are the same in each group at 5% level of significance. The R output for kruskal.test is presented below.

Table 2: Kruskal-Wallis R Output for private schools

Kruskal-Wallis rank sum test
data: Scorey2 by Method
Kruskal-Wallis chi-squared = 3.3267, df = 2, p-value = 0.1895

Since the p-value = 0.1895 is greater than 5%, the null hypothesis is not rejected. This leads to conclusion that the three independent samples in Private Primary Schools stratum, all have their location parameters from the same population and population parameters can thus be deduced using sample statistics.

4.5 Data Presentation

From foregoing section, the summary statistics are given as follows. The mean vector, **P**, is presented in vector form as

P= (**P_{11},P_{21},P_{31},P_{12},P_{22},P_{32}**) = (72.2372%,31.10436%,30.59135%,24.2004%,37.10228%,62.85931%).

The stratum mean for e-learning implementation levels are 0.446443 and 0.4138733 for public and private schools respectively, with variances of 0.2427037 and 0.1915377 respectively. In relation to Mukhopadhyay(1998) and Nassiuma(2000), in the context of stratified random sampling;

the an unbiased mean estimator $\overline{y}_w = \sum_i w_i y_i, \forall w_i = N_i \div N$ for overall population mean $\overline{Y}$.

The unbiased population estimator for overall population mean $\overline{Y}_i$ is 0.4269012. Using the mean vector **P**, pie charts representing the two strata mean vectors are presented below.

Figure 8: Pie Chart representation of mean vectors

The comparison of e-learning implementation levels per strata per region is presented in Figure 9 in the bar graph below, using the foregoing summary statistics.

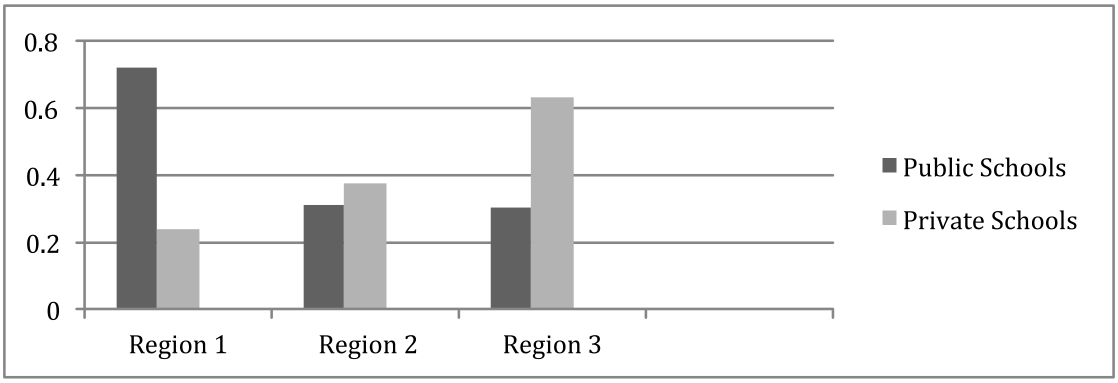

Figure 9: Comparison of strata mean vectors per Region

It is evident that the highest e-learning implementation level is realized in Region 1 Public Schools followed by Private Schools in Region 3. The lowest e-learning implementation level, according to the empirical model fitted is seen in Region 1 Private Schools.

5 Conclusions and Recommendations

The study reveals that the implementation e-learning level per strata per region varies. Using the equation on finite union of sets, the mean implementation e-learning level stands at 44.64430% in public primary schools and 41.38733% in private primary schools (this is only valid after establishing that the three regions per strata form a population as outline under Kruskal-Wallis test). Thus, the overall population estimator for e-learning implementation in primary schools is 42.69012%. It is notably vital to state that from the mean values presented, all schools do not have the same mean value, but some have higher while others lower, and budgetary allocations and funding should be school-based and not regional-based.

From the study, the relevant stakeholders have low-level of preparedness to implement e-learning. They have minimum requirements to implement e-learning policies in Nairobi County, owing to availability of e-learning infrastructures necessary for implementing e-learning. However, the variables like e-learning infrastructures physical security, trained personnel availability and sustenance of e-learning through relevant digital content once implemented were not tackled by the research. It is thus necessary for the government and concern education stakeholders to prepare budget that can look into holistic implementation and sustenance of e-learning for its role in improving quality of education.

The mathematical equation applied to establish implementation e-learning level in R is rigorous and its workability is acceptable on numeric vectors of e-learning variables of significance since it returns interval data in line to probability axioms i.e. at all times it restricts the output to be within the range of zero and one, extreme ends inclusive. This gives an objective way of deciding on implementation e-learning level by logical argument of education variables concern. However, the major weakness of the applied equation is observed whenever either B or C is 0 or 1 so that B or C that is non-zero or unity dominates. For example, if B= 0, A = 1 and C = 0.9, then implementation level Y = 0.9 (C dominates the implementation level). But if B = 1, A = 1 and C = 0.9, the Y = 1 (B dominates). The equation applied does not display such weakness as long both B and C lies strictly between 0 and 1, excluding 1 and 0. The applied equation is thus assuming that B and C equally contributes to implementation e-learning level, though it may not be true from one school to next based on teacher's and learner's knowledge of utility of such e-learning infrastructures. It may thus, be necessary for further research be done and adjust the mathematical equation applied by introduction of weights on B and C so that this exhibited weakness can be overcome. Again, the Set C can be designed to return a continuous primary data through further decision on its usability in e-learning to class size or streams per class.

References

Andrew Gelman, John B. Carlin, Hal S. Stern and Donald B. Rubin(2008), *Bayesian Data Analysis 2nd edition(Texts in Statistical Science),* Chapman & Hall/ CRC, A CRC Press Company, London and New York

Beatrice Ghirardini(2011), *E-learning methodologies: A Guide for designing and developing e-learning courses*, Food and Agricultural Organization(FAO), Rome

C. R. Kothari(2009), *Quantitative Techniques*, Vikas Publishing House PVT Ltd., New Delhi

Cramer H.(1973), *The Elements of Probability and some of its Applications 2nd edition,* Huntington, New York Krieger

Dae Joon Hwang, Hye-Kyung Yang and Hyeonjin Kim(2010), *E-learning in the Republic of Korea*, UNESCO Institute for Information Technologies in Education, Moscow

Dankit K. Nassiuma(2000), *Survey Sampling: Theory and Methods*, Nairobi University Press, Nairobi

Dennis D. Wackerly, William Mendenhall and Richard L. Scheaffer(2009), *International*

Edwin J. Purcell and Dale Varberg(1987), *Calculus with Analytic Geometry*, Prentce-Hall International(UK) Limited, London

Feller, W. (1968), *An Introduction to Probability Theory and its Applications 3rd edition,* New York: Wiley

Fresen, J. (2005), *Quality assurance practice in online (web-supported)learning in higher education: An exploratory study,* Doctoral thesis, University of Pretoria

humanipo Home to African Tech, Samsung brings e-learning in Kenyan Primary Schools. Website: www.humanipo.com, 25-01-2013

ITU(International Telecommunication Union)(2007), *Measuring the information society 2007:ICT Opportunity Index and World telecommunication/ICT.[Online]* Available at Website: http://www.itu.int/ITU-D/ict/publications/ict-oi/2007/material/table1.html#low, (on 10-01-2013).

John Mwazemba(September 2011), Digital revolution: Are Kenyan publishers prepared?, *14th Nairobi International Book Fair– 2011 Exhibitors KPA Magazine*, Page 19, [Print]

Kaplan-Leiserson, E. (2003), *Glossary [Online]*, Available at Website:www.learningcircuits.org/glossary.html, (on 12-01-2013).

Lund Research Ltd, *One Way ANNOVA Test*, Website: https://statistics.laerd.com/annova, 15-04-2014.

Meyer, P.L. (1970), *Introductory Probability and Statistical Applications, 2nd edition Reading*, Mass; Addison-Wesley

Ministry of Information and Communications(January 2006), *National Information &Communications Technology (ICT) Policy*, Republic of Kenya

Murray R. Spiegel, John Schiller and R. Alu Srinivasan(2009), *Schaum's Outlines of Probability and Statistics 3rd edition*, McGraw-Hill Companies Inc., New York

Omondi Oscar Odhiambo(2009),*Comparative Study of E-learning Platforms used in Kenyan Universities: Case study of Jomo Kenyatta of Agriculture and Technology and United States International University*, Masters, United States International University, Nairobi, Kenya

Oredo John Otieno(2008), *A framework for Evaluating ICT use in Teacher Education: A case study of the Primary Teacher Training Colleges in Kenya*, Masters, United States International University, Nairobi, Kenya

Organisation for Economic Cooperation and Development(December 2005), *Policy Brief: E-learning in Tertiary Education*, [Online], Available at www.oecdbookshop.org , 14/1/2013

Parimal Mukhopadhyay(1998), *Theory and Methods of Survey Sampling*, Asoke K. Ghosh, Prentice-Hall of India Private Limited, M-97, New Delhi

Parzen, E. (1992), *Modern Probability Theory and its Applications*, Wiley-Interscience, New York

Willium G. Cochran(1977), *Sampling Techniques*, Wiley Eastern Limited, New Delhi

Acknowledgement

I will like to acknowledge the enormous contribution of my supervisors Dr. George O. Orwa and Prof. Romanus O. Odhiambo who took much of their time assessing and correcting my work. Many thanks are also in order to Dr. G. A. Waititu for advice in data analysis specially in usage of R.

Gratitude goes to my colleagues: Victor, Grace, Ismail and Hezron.

Finally, I thank my family (Jackline Bett- wife and sons- Mark & Mike) for all their time tolerance during my dedication to this research.