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

# Application of Markov Chain Model in Studying progression Of Secondary School Students by Sex During The Free Secondary Education: A Case Study of Kisii Central District

Mose Job Nyandwaki[1, 4] Odhiambo E Akelo[2] Ojunga O Samson[3] Prof. Onyango Fredrick[4]
1. Kenya National Bureau of Statistics, Machakos County, PO box 380-90100, Machakos GPO, Kenya
2. Kenya National Bureau of Statistics, PO box 30266-00100, Nairobi, Kenya
3. Ministry of Planning, PO box 115-40111, Pap Onditi, Kisumu, Kenya
4. School of Mathematics, Statistics and Actuarial Science, Maseno University, Private Bag, Maseno, Kenya.
* jobmose2013@gmail.com

**Abstract**

Enrollment forecasting is an essential element in budgeting, resource allocation, and the overall planning for the growth of education sector. This paper demonstrates the use of Markov chain techniques in studying progression of secondary school students from the time of entry/enrollment in form one to graduation after the expected four years in Kenya's secondary school level of education. The target population included all the secondary school students in Kisii Central District. The model was used to determine the district's secondary school completion/dropout rate, retention rate and the expected duration of schooling by sex. It was established that completion rates for male students was higher than that of female students and dropout rates for female students was higher than that of male students. In the long run, it was established that the completion and dropout rates were the absorbing rates. Female students had lower expectation of schooling compared to male students in Kisii Central District. The model is only appropriate in making short period projections.

**Keywords:** Absorbing States, Absorbing Markov Chain, Transition Rates, Dropout Rates, Completion Rates, Fundamental Matrix

## 1. Introduction

Education is widely valued as a central factor in economic, social, and political development of any country. Secondary school education provides a vital link between basic education and further training in tertiary level of education on one hand, and the future world of work, on the other. It is therefore an important sub-sector of education in the preparation of human capital for development and provision of life opportunities. For quality and equality in education to be realized, an understanding of the secondary school student enrollment trend, completion rates, dropout rates, retention rates per class and the expected duration of schooling by sex is needed for planning and proper decision making.

The Markov chain model has been widely used in different fields including education. Mostly it has been applied in a single school, a university or a college but little has been done in applying it to a region or a group of schools. Also the model has not been applied in secondary school level of education. This study therefore, sought to apply the Markov Chain Model to study progress of secondary school students by sex in Kisii Central District, Kisii County, Kenya.

Education system is comparable to a hierarchical organization in which after an academic year, three possibilities arise in the new status of the students; The student may move to the next higher class, May repeat the same class, or may leave the system successfully as graduate or dropout of the system before attaining the maximum qualification (Musiga, Owino and Weke, 2010 and 2011).

In this study, a secondary school system is mo deled using the Markov chain approach in which proportions of students who leave the system either successfully or dropout of the system re separately grouped into double absorbing states. States of the education system were partitioned into two categories; non-absorbing (transient) states which corresponded to the various classes within the education system and absorbing states which also corresponded to the group of all successful graduates, repeaters and dropouts. Thus students enter permanent states; absorbing states, either as graduates or dropouts. This research sought to establish secondary school retention rates, completion rates, dropout rates, expected duration of schooling by sex in Kisii Central District for the purpose of gaining greater insight into areas where the same might be improved. On the basis of the data that was obtained from Kisii Central District, the progress of students from class to class and total enrollment at a given time in secondary school education system was modeled.

The model established will help the stakeholders in the distribution of resources to secondary schools when year of study is put into consideration and, planning and budgeting for the expected future enrollment for quality and equality in education. Equally the model will be used to predict number of dropouts by sex at each class of study which will assist interested parties to make informed decision about secondary education management in view to curb dropouts and reducing the gender disparity in secondary schools. Also, since the initiation of FSE, there is need too to analyze what is currently being financed and what the future needs will be in the country using the case of Kisii Central District secondary school enrollment data.

### 1.1 Objectives of the Study

This study was guided by the following objectives;

1. To establish the secondary school completion rate by sex in Kisii Central District by use of Markov chain Model.
2. To determine the expected duration of schooling by sex in Kisii Central District by use of Markov chain Model.
3. To determine the dropout rate by sex in Kisii Central District by use of Markov chain Model.
4. To determine secondary school retention rates by sex in Kisii Central District by use of Markov Chain Model.
5. To demonstrate the use of Markov Chain model in enrollment projection in Kisii Central district.

### 1.2 Assumptions of the Study

The following assumptions are taken into consideration for the model to be appropriate;

- The study population (Kisii Central District, student's enrollment) is assumed closed i.e. there is no immigration and out migration of students with the neighboring districts.
- Admissions takes place only in form one.
- Dropouts are assumed to be uniformly distributed in the period $(x; x + 1)$, $x$ being the year of study.
- No class repetition.
- Transition from primary to secondary level of education in Kisii Central district is assumed constant.

## 2. Literature Review

Over the years, many techniques have been suggested for forecasting enrollment and students flows at any level in education systems. Wing, 1974 classified them into curve fitting, causal models, attitude surveys and judgmental techniques. Healey et al, 1978 classified them into the judgmental, Markov process, trend analysis, regression, simulation and the ratio techniques.

A Markov model is a stochastic model (one which models random events) which has been used in diverse fields such as computer science, engineering, mathematics, genetics, agriculture economics, education, biology, etc. (Hillis, Maguire, Hawkins and Newhouse, 1986 , Jain, 1986, Stewart, 1994).

Markov chains have been widely used to model stochastic processes and to evaluate time to event data. Musiga et al., 2010 modeled a hierarchical system with a single absorbing state for an education system where dropouts and graduates were grouped together. Later, Musiga et al., 2011, modeled a hierarchical system with double absorbing states for an education system, where graduates were separated from dropouts. These two papers form the basis for this study. In both, the education system under study is narrowed to one institution but this study generalizes to the group of institutions in a district to be specific and takes into consideration sex disaggregation of the data.

Generally, most of these studies are carried out in primary level of study, universities and single institutions of learning. Little has been done in modeling progress of a group of students in secondary schools in Kenya using Markov chains.

## 3 Model Development

The Markov model is based on an underlying stochastic process in which a system in one state, say, $s_i$ moves to a subsequent state, say, $s_j$. The states are sometimes referred to as the Current state and the Next state. The act of moving from one state to the next is referred to as a step or transition.

In this study, a Markov chain model with t non-absorbing states; 1, 2…., t corresponding to the classes of the education system and r absorbing states corresponding to the various final qualifications was considered. This implies that N = t + r, where N is the total number of possible states of the education system. Transition probabilities between absorbing states should be represented by one, hence the use of identity matrix. Transition from an absorbing state to a non-absorbing state which is impossible, is represented by zero, hence the matrix of zeroes. Transition from non-absorbing states to absorbing states are possible, likewise transitions between non-absorbing states.

### 3.1 Absorbing Markov Chain

A Markov Chain is absorbing if it has at least one absorbing state and if from every state it is possible to go to an absorbing state (not necessarily in one step).In an absorbing Markov Chain, a state which is not absorbing is called transient. If we have an absorbing Markov chain with t transient states and r absorbing states, the transition probability matrix **P**, will take the following canonical form;

$$\mathbf{P}=\begin{bmatrix} \boldsymbol{Q} & \boldsymbol{R} \\ \boldsymbol{O} & \boldsymbol{I} \end{bmatrix}$$

Where;

**Q** is a $t \times t$ matrix, $q_{ij}$ being the probability that a student who is in class i at time *(t-1)* will be in class j at time t; $i, j = 1, \quad 2, \quad 3\ldots, t,$

**R** is a non-zero $t \times r$ matrix, $r_{ik}$ being the probability that a student in class i at time *(t-1)* will graduate with final education k at time $t;\ i = 1, 2, 3\ldots, t\ \&\ k = 1, 2, \quad 3\ldots, r;$

**O** is an $r \times t$ zero matrix and,

**I** is an $r\times r$ identity matrix.

The first *t* states are transient states and the last *r* states are absorbing states (Beck and Pauker, 1983).

The $ij^{th}$ entry, $\boldsymbol{P}_{ij}^{\boldsymbol{n}}$ of the matrix $\boldsymbol{P}^n$ gives the probability that the Markov chain, starting in state $s_i$ will be in state $s_j$ after *n* steps by Chapman-Kolmogorov theorem. The canonical form of the matrix $\boldsymbol{P}^n$ is given as;

$$\boldsymbol{P}^{\boldsymbol{n}}=\begin{bmatrix} \boldsymbol{Q}^{\boldsymbol{n}} & \boldsymbol{R}^{\boldsymbol{n}} \\ \boldsymbol{0} & \boldsymbol{I} \end{bmatrix}$$

Where;

$\boldsymbol{Q}^{\boldsymbol{n}}$ is a $t\times t$ matrix which gives the probability that a student who is in class *i* will be in class *j, n* years later; $i; j = 1, 2, 3\ldots, t,$

$\boldsymbol{R}^{\boldsymbol{n}}$ = $(\mathbf{I}+\mathbf{Q}+\mathbf{Q}^2+\ldots\ldots+\mathbf{Q}^{n-1})\mathbf{R}$. is a $t\times r$ matrix which gives the probability that a student who is in class *i* will graduate with final education *k* within n years, $i = 1, 2\ldots, t;\ k = 1, \quad 2\ldots, r.$ It is also called the completion rate,

**O** is an $r\times t$ matrix of zeros which gives transition probabilities from absorbing states to non absorbing states in n steps and,

**I** is a $r \times r$ identity matrix which gives transition probabilities between absorbing states in *n* steps.

Urakabe et al, 1986 and Silverstein et al, 1988 concluded that while the probability matrix summarizes transition probability of the cohort, the transient states analysis allows prediction or prognosis for an individual subject, given their starting state, current state and cycle.

### 3.2 The Fundamental matrix

For an absorbing Markov chain, the matrix **N** is called the fundamental matrix where;

$\mathbf{N}= (\mathbf{I}-\mathbf{Q})^{-1}= \mathbf{I}+\mathbf{Q}+\mathbf{Q}^2+\ldots\ldots$ (3.1)

Where the $ij^{th}$ entry $n_{ij}$ of the matrix **N** is the expected number of times the process is in the transient state $s_j$ given that it started in the transient state $s_i$.

Hence **N** gives the average number of cycles that a subject resides in transient states before absorption, given a specified starting state.

The states of the education system were denoted by integers *1, 2, 3,. . . , n* at time *t= 0, 1, 2,. . .* while $p_{ij}$ denoted the probability that a student in class *i* at time *t-1* will be in class *j* at time *t,* then the transition matrix,

$\boldsymbol{P}=((P_{ij}));\ i,j= 1, 2, \ldots, N.$

The non-absorbing states(transient states) were four and they were represented by values *1, 2, 3,* and *4*. This implies that the **Q** component of the transition matrix **P** is a $4 \times 4$ matrix.

The number of absorbing states were two and they were represented by values *5* and *6*.

The absorbing state *5* represents graduation from the system after attaining the maximum qualification and state *6* represents dropping out of the system before attaining the maximum qualification. Hence the **R** component of matrix **P** was a $4 \times 2$ matrix.

According to (**?**), the purpose of the transition matrix is to represent the probability of movement between states in a single time period. In this case, it was the probability that a student will reach a particular state by the end of the year of study.

### 3.3 Initial Transition Matrix

By letting $n_{ij}(t)$ to represent the number of students in class *i* at time *(t - 1)* who will be in class *j* at time *t* and $n_i(t-1)$, the number of students in class *i* at time *t- 1*, and by assuming the multinomial distribution, the initial transition probabilities were estimated by;

$P_{ij}= nij(t)/\ n_i(t-1)$, where; $i; j = 1, 2\ldots, t:$

This was the proportion of students who were in class *i* at time *(t- 1)* who ended up being in class *j* at time *t*.

### 3.4 The n-step Transition Matrix

The *n*-step transition probability matrix takes the canonical form below as per Beck and Pauker, 1983 and Musiga et al., 2011 and the Kolmogorov theorem;

$$\boldsymbol{P^n}=\begin{bmatrix} \boldsymbol{Q^n} & \boldsymbol{R^n} \\ \boldsymbol{0} & \boldsymbol{I} \end{bmatrix}$$

The solution to this *n*-step transition matrix gives the state of a student *n*-steps(years later). The elements of the *n*-step transition probability matrix represents the probabilities that an object in a given state will be in the next state *n*-steps later.

### 3.5 Completion rates

Musiga and Owino in (Musiga et al., 2011) defined the dropout rate from class *i, n* years later by;

$$r_{ik}^{(n)} = \sum_{j=1}^{s} q_{ij}^{(n-1)} r_{ik}, \quad i; j = 1; 2\ldots, s$$

Where

$p_{ij}^{(n-1)}$ is the probability that a student in class *i* will be in class *j, n- 1* years later and $r_{ik}$ is the probability that a student in class *i* at time *t - 1* graduates with final education *k* at time *t*. It is the $(i, k)^{th}$ element of the product $\boldsymbol{Q^{n-1}}\mathbf{R}$.

Therefore the cumulative dropout rate within y years from class i will be given by;

$$r_{ik}^{(y)} = \sum_{n=1}^{y} r_{ik}^{(n)} \qquad i=1,2,\ldots,t \text{ and } k=1, 2, \ldots, r.$$

Where, $r_{ik}^{(y)}$ is the $(i,k)^{th}$ element of $(\boldsymbol{I}+\boldsymbol{Q}+\boldsymbol{Q}^2+\ldots\ldots+\boldsymbol{Q}^{y-1})\boldsymbol{R}$.

### 3.6 Absorbing rates

Assuming that students will remain in the system indefinitely, then the absorbing rate is given by;

$r_{i1}^{(\infty)} = \sum_{n=1}^{\infty} r_{i1}^{(n)}$

$= (\mathbf{I}+\mathbf{Q}+\mathbf{Q}^2+\ldots\ldots..)\mathbf{R}.$

$= (\mathbf{I}-\mathbf{Q})^{-1}\mathbf{R}$

The solution to this gives the absorbing rate under double absorbing states system. (Uche, 1980 and Musiga et al., 2011)

### 3.7 Retention rates

According to (Dworkin, 2005), retention rates are of interest not only to learning institutions hopeful of maintaining or increasing enrollment, but act as a social economic indicator of well being for the community as a whole. It is therefore, considered to be in the best public interest to maintain high retention rates.

## 4 Model Fitting

### 4.1 Initial transition Probabilities

By letting $n_{ij}(t)$ represent the number of students in class *i* at time *(t - 1)* who will be in class *j* at time *t*, and $n_i(t-1)$ to represent the number of students in class *i* at time *(t - 1),* and by assuming the multinomial distribution, the transition probabilities can be estimated by;

$P_{ij}= n_{ij}(t)/ n_i(t-1), \; i,j= 1, 2, .., N$

$P_{ij}$ is the proportion of students who are in class *i* at time *(t- 1)* who ends up being in class *j* at time *t.*

#### 4.1.1 Important Notations

Some of the notations used in the subsequent sections are defined below;

- subscript d represents district
- subscript m represents male students
- subscript f represents female students

Such that $\mathbf{P}_d$ =District transition probability matrix, $\mathbf{P}_m$ =male students transition probability matrix, and $\mathbf{P}_f$ =female students transition probability matrix, etc

### 4.2 The Initial Transition Probability matrix for Kisii Central District

From the data obtained from Kisii Central District, students enrollment in Forms 1, 2, 3 and 4 for the year 2011 and enrollment for the same students in forms 2, 3, and Form 4s who graduated the following year, 2012 were as shown in Table I. The dropout proportions before attaining maximum qualification for students who were in Forms *1, 2, 3* and *4* were *(65/5329)=0.012197, (345=5531) = 0:062376, (199/6180)=0.032201 and (265/4550)=0.058242,* respectively. The proportion of students who graduated successfully after reaching Form 4 is given by *(4285/4550) = 0.941758*. This gives rise to the **R** component of the matrix **P** for the district. The **Q** component of the matrix **P**, whose states are transient states, its elements represent the proportions of students who proceeded to Forms 2, 3 and 4 the following year, 2012. The proportions of students who proceeded to Forms 2, 3 and 4 in the year 2012 are *(5264/5329) = 0.987803, (5186/5531) = 0.937624* and *(5981/6180)= 0.941758* respectively.

Thus, assuming time homogeneity, the district transition probability matrix **P**, is;

$$\mathbf{P}_d = \begin{bmatrix} 0.000000 & 0.987803 & 0.000000 & 0.000000 & 0.000000 & 0.012197 \\ 0.000000 & 0.000000 & 0.937624 & 0.000000 & 0.000000 & 0.062376 \\ 0.000000 & 0.000000 & 0.000000 & 0.967799 & 0.000000 & 0.032201 \\ 0.000000 & 0.000000 & 0.000000 & 0.000000 & 0.941758 & 0.058242 \\ 0.000000 & 0.000000 & 0.000000 & 0.000000 & 1.000000 & 0.000000 \\ 0.000000 & 0.000000 & 0.000000 & 0.000000 & 0.000000 & 1.000000 \end{bmatrix}$$

The **Q** component is;

$$\mathbf{Q}_d = \begin{bmatrix} 0.000000 & 0.987803 & 0.000000 & 0.000000 \\ 0.000000 & 0.000000 & 0.937624 & 0.000000 \\ 0.000000 & 0.000000 & 0.000000 & 0.967799 \\ 0.000000 & 0.000000 & 0.000000 & 0.000000 \end{bmatrix}$$

and the **R** component is;

$$\mathbf{R}_d = \begin{bmatrix} 0.000000 & 0.012197 \\ 0.000000 & 0.062376 \\ 0.000000 & 0.032201 \\ 0.941758 & 0.058242 \end{bmatrix}$$

### 4.3 The Initial Transition Probability matrix by gender

Considering the same data by sex, for the male students, the proportions who dropout out before attaining maximum qualification in Forms 1, 2, 3 and 4 were *(40/2771)=0.014435, (145/2718)=0.053348, (88/3018)=0.029158* and *(89/1977)=0.045018,* respectively. The proportion of male students who graduated successfully after reaching Form 4 was *(1888/1977) = 0.954982*. This gives rise to the **R** component of the matrix **P** for the male students. The **Q** component for the transition matrix **P**, for male students was given by the proportions of the male students who proceeded to Forms 2, 3 and 4 in the year 2012. They were given as *(2731/2771)=0.985565, (2573/2718)=0.946652* and *(2930/3018)=0.970842* respectively.

Thus, the transition probability matrix **P** for male students, with the double absorbing states, assuming time homogeneity is;

$$\mathbf{P}_m = \begin{bmatrix} 0.000000 & 0.985565 & 0.000000 & 0.000000 & 0.000000 & 0.014435 \\ 0.000000 & 0.000000 & 0.946652 & 0.000000 & 0.000000 & 0.053348 \\ 0.000000 & 0.000000 & 0.000000 & 0.970842 & 0.000000 & 0.029158 \\ 0.000000 & 0.000000 & 0.000000 & 0.000000 & 0.954982 & 0.045018 \\ 0.000000 & 0.000000 & 0.000000 & 0.000000 & 1.000000 & 0.000000 \\ 0.000000 & 0.000000 & 0.000000 & 0.000000 & 0.000000 & 1.000000 \end{bmatrix}$$

The **Q** component for the male students is;

$$\mathbf{Q}_m = \begin{bmatrix} 0.000000 & 0.985565 & 0.000000 & 0.000000 \\ 0.000000 & 0.000000 & 0.946652 & 0.000000 \\ 0.000000 & 0.000000 & 0.000000 & 0.970842 \\ 0.000000 & 0.000000 & 0.000000 & 0.000000 \end{bmatrix}$$

and the **R** component is;

$$\mathbf{R}_m = \begin{bmatrix} 0.000000 & 0.014435 \\ 0.000000 & 0.053348 \\ 0.000000 & 0.029158 \\ 0.954982 & 0.045018 \end{bmatrix}$$

Similarly, the transition probability matrix P for female students, with the double absorbing states, assuming time homogeneity is;

$$\mathbf{P}_f = \begin{bmatrix} 0.000000 & 0.990227 & 0.000000 & 0.000000 & 0.000000 & 0.009773 \\ 0.000000 & 0.000000 & 0.928902 & 0.000000 & 0.000000 & 0.071098 \\ 0.000000 & 0.000000 & 0.000000 & 0.964896 & 0.000000 & 0.035104 \\ 0.000000 & 0.000000 & 0.000000 & 0.000000 & 0.931597 & 0.068403 \\ 0.000000 & 0.000000 & 0.000000 & 0.000000 & 1.000000 & 0.000000 \\ 0.000000 & 0.000000 & 0.000000 & 0.000000 & 0.000000 & 1.000000 \end{bmatrix}$$

Its **Q** component is;

$$\mathbf{Q}_f = \begin{bmatrix} 0.000000 & 0.990227 & 0.000000 & 0.000000 \\ 0.000000 & 0.000000 & 0.928902 & 0.000000 \\ 0.000000 & 0.000000 & 0.000000 & 0.964896 \\ 0.000000 & 0.000000 & 0.000000 & 0.000000 \end{bmatrix}$$

and the **R** component is;

$$\mathbf{R}_f = \begin{bmatrix} 0.000000 & 0.009773 \\ 0.000000 & 0.071098 \\ 0.000000 & 0.035104 \\ 0.931597 & 0.068403 \end{bmatrix}$$

**4.4 Completion rates**

Students in Form 4 were grouped into those who dropped out of the system before attaining the maximum qualification and those who actually graduated from the system. The completion rate is the $(i, k)^{th}$ element of $(\boldsymbol{I} + \boldsymbol{Q} + \boldsymbol{Q}^2 + \ldots + \boldsymbol{Q}^{(x-1)})\boldsymbol{R}$:

Within one year i.e. in the year 2013, the district completion rate will be given by;

$$(\boldsymbol{I}+\boldsymbol{Q}_d)\boldsymbol{R}_d = \begin{bmatrix} 0.000000 & 0.073812 \\ 0.000000 & 0.092568 \\ 0.911432 & 0.088568 \\ 0.941758 & 0.058242 \end{bmatrix}$$

Within two years it will be;

$$(\boldsymbol{I}+\boldsymbol{Q}_d+\boldsymbol{Q}_d^2)\boldsymbol{R}_d = \begin{bmatrix} 0.000000 & 0.103636 \\ 0.854581 & 0.145419 \\ 0.911432 & 0.088568 \\ 0.941758 & 0.058242 \end{bmatrix}$$

In summary the completion rates within x years using the absorbing states model is as in Table II in the Appendices. From Table II, it is clear that by 2013, 7.3812% of the students who were in Form 1 in the year 2011 had dropped out of the system before attaining the maximum qualification. By the year 2015, 84.4158% of the students who were in Form 1 in the year 2011 are expected to graduate from the system. At the same time a maximum of 15.5842% of the same students will drop out of the system without attaining the maximum qualification. Also for the students in Form 2, 85.4581% of the students are expected to graduate from the system after attaining maximum qualification by the year 2014. For those in Form 3, 91.1432% of the students are expected to successfully graduate from the system by the year 2013.

**4.5 Completion rates by sex**

In addition to the above analysis, the results were further disaggregated by sex. The students were grouped into those who dropped out of the system before attaining the maximum qualification and those who actually graduated from the system by gender. The completion rate is the $(i, k)^{th}$ element of $(\boldsymbol{I} + \boldsymbol{Q} + \boldsymbol{Q}^2 + \ldots + \boldsymbol{Q}^{(x-1)})\boldsymbol{R}$: in both cases. Within one year i.e. in the year 2013, the completion rate for female will be given by;

$$(\boldsymbol{I} + \boldsymbol{Q}_f)\boldsymbol{R}_f = \begin{bmatrix} 0.000000 & 0.080176 \\ 0.000000 & 0.103706 \\ 0.898894 & 0.101106 \\ 0.931597 & 0.068403 \end{bmatrix}$$

While that for male students will be;

$$(\boldsymbol{I} + \boldsymbol{Q}_m)\boldsymbol{R}_m = \begin{bmatrix} 0.000000 & 0.067013 \\ 0.000000 & 0.080950 \\ 0.927137 & 0.072863 \\ 0.954982 & 0.045018 \end{bmatrix}$$

In summary, the completion rates for female and male students within x years using a double absorbing states model is as in Tables III and IV in the Appendices respectively; From Table III we can clearly see that by the year 2013, 8.0176% of female students who were in Form 1 in the year 2011 had dropped out of the system before attaining the maximum qualification and 82.6824% of the same students are expected to graduate from the system successfully by the year 2015. For the female students who were in Form 2 in the year 2011, a maximum of 83.4985% are expected to graduate successfully from the system and a maximum of 16.5015% are expected to drop out before completing their secondary education successfully by the year 2014. Also, for those who were in Form 3, a maximum of 89.8894% are expected to graduate while 10.1106% are expected to drop. Finally for those who were in Form 4, 93.1597% graduated successfully while 6.8403% dropped.

Table IV in the Appendices, indicates that 6.7013% of the male students who were in Form 1 in the year 2011 dropped out of the system by 2013 compared to 8.0176% of their female classmates in the same period. By the year 2015, 86.5007% of the male students who were in Form 1 the year 2011 are expected to graduate from the system. At the same time a maximum of 13.4993% of the same male students will drop out of the system without attaining the maximum qualification compared to 17.3176% of the female students in the same class. Also for the students in Form 2, 87.7676% of male students are expected to graduate from the system after attaining maximum qualification by the year 2014. For those in Form 3, 92.7137% of male students are expected to successfully graduate from the system by the year 2013.

**4.6 The expected duration of study**

Here we consider the fundamental matrix **N**. The matrix gives the number of cycles that a subject resides in transient states before absorption, given a specified starting state. The fundamental matrix **N** is given as;

$$N = I + Q + Q^2 + \ldots = (I-Q)^{-1}$$

To compute $N$ for the district, we first get the matrix;

$$I - Q_d = \begin{bmatrix} 1.000000 & -0.987803 & 0.000000 & 0.000000 \\ 0.000000 & 1.000000 & -0.937624 & 0.000000 \\ 0.000000 & 0.000000 & 1.000000 & -0.967799 \\ 0.000000 & 0.000000 & 0.000000 & 1.000000 \end{bmatrix}$$

and its inverse is;

$$(I - Q_d)^{-1} = \begin{bmatrix} 1.000000 & 0.987803 & 0.926188 & 0.896364 \\ 0.000000 & 1.000000 & 0.937624 & 0.907432 \\ 0.000000 & 0.000000 & 1.000000 & 0.967799 \\ 0.000000 & 0.000000 & 0.000000 & 1.000000 \end{bmatrix}$$

The expected duration of study according to Beck and Pauker, 1983 is given by;

$$(\mathbf{I} - \mathbf{Q}_d)^{-1}(1111)^T = \begin{bmatrix} 3.810354 \\ 2.845056 \\ 1.967799 \\ 1.000000 \end{bmatrix}$$

This result gives the total expected duration in school till completion. From the result, the expected duration of study for a student in Forms 1, 2, 3 and 4 is 3.810354, 2.845056, 1.967799 and 1.00000 years respectively.

**4.7 The expected duration of study by sex**

Considering the fundamental matrix **N** for both cases, we can deduce the expected duration of schooling for either sex. The matrix as stated earlier gives the number of cycles that a subject resides in transient states before absorption, given a specified starting state. From the fundamental matrix **N**;

$$N = I + Q + Q^2 + \ldots = (I-Q)^{-1};$$

we get that **N** for male students;

$$(I - Q_m)^{-1} = \begin{bmatrix} 1.000000 & 0.985565 & 0.932987 & 0.905783 \\ 0.00000 & 1.000000 & 0.946652 & 0.919050 \\ 0.000000 & 0.000000 & 1.000000 & 0.970842 \\ 0.000000 & 0.000000 & 0.000000 & 1.000000 \end{bmatrix}$$

while that for female students is;

$$(I - Q_f)^{-1} = \begin{bmatrix} 1.000000 & 0.990227 & 0.919824 & 0.887534 \\ 0.00000 & 1.000000 & 0.928902 & 0.896294 \\ 0.000000 & 0.000000 & 1.000000 & 0.964896 \\ 0.000000 & 0.000000 & 0.000000 & 1.000000 \end{bmatrix}$$

The expected duration of schooling for male students is given by;

 

$$(\mathbf{I} - \mathbf{Q}_m)^{-1}(1111)^T = \begin{bmatrix} 3.824335 \\ 2.865702 \\ 1.970842 \\ 1.000000 \end{bmatrix}$$

Similarly that for female students will be;

$$(\mathbf{I} - \mathbf{Q}_f)^{-1}(1111)^T = \begin{bmatrix} 3.797585 \\ 2.825196 \\ 1.964896 \\ 1.000000 \end{bmatrix}$$

This result gives the total expected duration in school till completion for either sex. From the result, the expected duration of study for a male student in Forms 1, 2, 3 and 4 is 3.824335, 2.865702, 1.970842 and 1.00000 years respectively while that for the female students in the same order is 3.797585, 2.825196, 1.964896 and 1.00000 years.

**4.8 Absorbing rates**

In the long run, the absorbing rate under double absorbing states is given by;

$(I - Q)^{-1}R$

and in this study the District absorbing rates were established to be;

$$(I - Q_d)^{-1}R_d == \begin{bmatrix} 0.844158 & 0.155842 \\ 0.854581 & 0.145419 \\ 0.911432 & 0.088568 \\ 0.941758 & 0.058242 \end{bmatrix}$$

From the result, it was established that in the long run, students who were in Forms 1, 2, 3 and 4, a maximum of 15.5842%, 14.5419%, 8.8568% and 5.8242% respectively dropped out of the system without attaining maximum qualification. Also considering the same students in that order, a maximum of 84.4158%, 85.4581%, 91.1432% and 94.1758% respectively successfully graduated from the system. Incidentally, in the long run, the rates of completion of each form are the same as the absorbing rates because there was no repetition.

**4.9 Absorbing rates by sex**

The absorbing rate was also computed with respect to gender. In the long run, the absorbing rate under double absorbing states is as discussed earlier. The absorbing rates for the male students are given as;

$$(I - Q_m)^{-1}R_m = \begin{bmatrix} 0.865007 & 0.134993 \\ 0.877676 & 0.122324 \\ 0.927137 & 0.072863 \\ 0.954982 & 0.045018 \end{bmatrix}$$

while that for female students was;

$$(\mathbf{I} - \mathbf{Q}_f)^{-1}\mathbf{R}_f = \begin{bmatrix} 0.826824 & 0.173176 \\ 0.834985 & 0.165015 \\ 0.898894 & 0.101106 \\ 0.931597 & 0.068403 \end{bmatrix}$$

From the result, it was established that in the long run, male students who were in Forms 1, 2, 3 and 4 in the year 2011, a maximum of 13.4993%, 12.2324%, 7.2863% and 4.5018% respectively dropped out of the system without attaining maximum qualification while for the female students in the same order, 17.3176%, 16.5015%, 10.1106% and 6.8403% respectively dropped out. Also considering the same students in that order, 86.5007%, 87.7676%, 92.7137% and 95.4982% of male students successfully graduated from the system while 82.6824%, 83.4985%, 89.8894% and 93.1597% of female students respectively graduated successfully from the system. These rates are the same as the proportions in the long run in Table III and IV in the Appendices below.

**4.10 Retention rates**

Retention rates can be determined from the transition probabilities. In this study, the retention rates for Forms 1, 2, 3 and 4 was established to be 0.987803, 0.937624, 0.967799 and 0.941758 respectively. It can be seen that Form 1 had the highest retention rate while Form 2 had the lowest. The same was also determined by gender. Male student's retention rates in the district were 0.985565, 0.946652, 0.970842 and 0.954982 for Forms 1, 2, 3 and 4 respectively while that of female students were 0.990227, 0.928902, 0.964896 and 0.931597 in the same order. It can be deduced that the retention rate for female students in Form 1 is higher than that of male students. However, in Forms 2, 3 and 4 the rates were higher for male students.

### 4.11 Enrollment Projections

Using the transition matrix and the completion rate model, it is possible to project the number of students in future though in a short period. For example the number of students in the year 2012, 2013, 2014,.. can be estimated using the transition matrix. Considering that for those in Form 1, the year 2011, only 0.9878030 proceeded to Form 2 in the year 2012, and 0.9261880 are expected to proceed to Form 3 in the year 2013, 0.8963640 are expected to proceed to form 4 in the year 2014 and finally only 0.8441580 are expected to graduate from the system successfully in the year 2015. This approach can be used to project enrollment in the district though in a short duration. To get the total number of students in the district in subsequent years, we need to know those who transit from primary to secondary in the district. This will be treated as inflows into the system but cannot be estimated by this model. Hence the short fall of the model.

Therefore, the total District enrollment in the year 2013 will be given by the sum of those who join Form 1 the same year, 98.78030% of those in Form 1 the year 2012, 92.6188% of those in Form 2 the year 2012, 89.63640% of those in Form 3 the year 2012.

## 5 Conclusion

The following conclusions were drawn from the above analysis and discussions;

- The completion rates for male students is higher than that of female students in Kisii Central District.
- The expected duration of schooling is higher for male students compared to the female students.
- Female students were found to have lower expectation of schooling than their male counterparts.
- The dropout rates for female students is higher than that of male students in Kisii Central District.
- The transition rate from Form 1 to Form 2 is higher for female students than that of male students. In all other Forms, transition rate for the male students is higher.
- Form two has the least retention rate in the District.
- The model can be used to project enrollment in the district though in a short duration.
- In the long run, completion and dropout rates were found to be the same as the absorbing rates when repetition is not allowed.

### 5.1 Recommendations

From the above conclusions, the following recommendations were made;

- Studies and sensitization should be done in Kisii Central District to increase the completion rates for female students.
- Research should be done to determine the causes of the gender disparity in expectation of schooling in Kisii Central District
- Research should be done to get into the root cause of female dropouts in Kisii Central District and recommend the way forward in order to reduce the gender parity in the district.
- In a closed hierarchical system, Markov chain model can be used to study how the system progresses with time.
- Considering the retention rates, research should be done to get into the insight of the reasons as to why there is a low retention rate in Form 2 in the District and the disparity in the same between the genders.
- This model can only be used in a short period of time which is one of the weaknesses of the model.

## References

Beck, J. R. & Pauker, S. G. (1983), "The Markov Process in Medical Prognosis", *Med Decis Making* **3**, 419–458.

Boltyanskiˇı, V. G., Gamkrelidze, R. V. & Pontryagin, L. S. (1956), "On the Theory of Optimal Processes", Dokl. Akad. Nauk SSSR (N.S.) **110**, 7–10.

Dworkin, J. B. (2005), "Retention: How important is it really? and graduation rates", http://www.pnc.edu/co/Retention.pdf .

Healey, M. T. & Brown, D. J. (1978), "Forecasting University Enrollment by Ratio Smoothing", *Higher*

*Education* **7**, 417–430.

Hillis, A., Maguire, M., Hawkins, B. & Newhouse, M. M. (1986), "The Markov Process as a General Method for Non-parametric Analysis of Right-Censored Medical Data", *J Chron Diseases* **39**, 595–604.

Jain, S. (1986), "Markov chain Models and its Applications", *Comp Biomed Res* **19**, 374–378.

Musiga, L. A., Owino, J. O. & Weke, P. G. (2010), "Modeling a Hierarchical System with a Single Absorbing State, University of Nairobi .

Musiga, L. A., Owino, J. O. & Weke, P. G. (2011), "Modeling a hierarchical system with double absorbing states", *International Journal of Business and Public Management,* **1**, Mount Kenya University, 158–167.

Silverstein, M. D., Albert, D., Hadler, N., & Ropes., M. (1988), "Prognosis in SLE: Comparison of Markov Model to life table analysis", *J clinic Epidemiologic* **41**, 623–633.

Stewart, W. J. (1994), "Introduction to the Numerical Solutions of Markov chains", Princeton University Press, Princeton.

Uche, P. I. (1980), "A transition model of academic survival in a single channel system", *International Journal of Education, Science and Technology* **17**, 177–187.

Urakabe, S., Orita, Y., Fukuhara, Y., Ando, A., Ueda, N., Furkawa, T. & Abe., H. (1986), "Prognosis of Chronic Glomerlonephritis in Adult Patients Estimated on the basis of the Markov Process", *Clin Nephrol* **3,** 48–53.

Wing, P. (1974), "Higher education enrollment forecasting", NCHEMS .

**Appendix: Tables**

**Table I: Kisii Central District Secondary School Enrollment 2011/2012**

| Class | Enrollment 2011 | | | Proceeded to 2012 | | | Dropped | | |
|---|---|---|---|---|---|---|---|---|---|
| | Male | Female | Total | Male | Female | Total | Male | Female | Total |
| **Form 1** | 2771 | 2558 | 5329 | 2731 | 2533 | 5264 | 40 | 25 | 65 |
| **Form 2** | 2718 | 2813 | 5531 | 2573 | 2613 | 5186 | 145 | 200 | 345 |
| **Form 3** | 3018 | 3162 | 6180 | 2930 | 3051 | 5981 | 88 | 111 | 199 |
| **Form 4** | 1977 | 2573 | 4550 | 1888 | 2397 | 4285 | 89 | 176 | 265 |

***Source: Extracted from Kisii Central District Education Statistics Database***

**Table II: Kisii Central District Secondary school completion rates within x years.**

| Yrs(x) | Form 1 | | Form 2 | | Form 3 | | Form 4 | |
|---|---|---|---|---|---|---|---|---|
| | Completion | Drop | Completion | Drop | Completion | Drop | Completion | drop |
| **1** | 0 | 0.012197 | 0 | 0.062376 | 0 | 0.032201 | 0.941758 | 0.058242 |
| **2** | 0 | 0.073812 | 0 | 0.092568 | 0.911432 | 0.088568 | 0.941758 | 0.058242 |
| **3** | 0 | 0.103636 | 0.854581 | 0.145419 | 0.911432 | 0.088568 | 0.941758 | 0.058242 |
| **4** | 0.844158 | 0.155842 | 0.854581 | 0.145419 | 0.911432 | 0.088568 | 0.941758 | 0.058242 |
| **5** | 0.844158 | 0.155842 | 0.854581 | 0.145419 | 0.911432 | 0.088568 | 0.941758 | 0.058242 |

**Table III: Kisii Central District Female Students completion rates within x years.**

| Year(x) | Form 1 | | Form 2 | | Form 3 | | Form 4 | |
|---|---|---|---|---|---|---|---|---|
| | Completion | Dropouts | Completion | Dropouts | Completion | Dropouts | Completion | Dropouts |
| **1** | 0 | 0.009773 | 0 | 0.071098 | 0 | 0.035104 | 0.931597 | 0.068403 |
| **2** | 0 | 0.080176 | 0 | 0.103706 | 0.898894 | 0.101106 | 0.931597 | 0.068403 |
| **3** | 0 | 0.112466 | 0.834985 | 0.165015 | 0.898894 | 0.101106 | 0.931597 | 0.068403 |
| **4** | 0.826824 | 0.173176 | 0.834985 | 0.165015 | 0.898894 | 0.101106 | 0.931597 | 0.068403 |
| **5** | 0.826824 | 0.173176 | 0.834985 | 0.165015 | 0.898894 | 0.101106 | 0.931597 | 0.068403 |

**Table IV: Kisii Central District Male students Completion Rates within x years.**

| Year(x) | Form 1 | | Form 2 | | Form 3 | | Form 4 | |
|---|---|---|---|---|---|---|---|---|
| | Completion | Dropouts | Completion | Dropouts | Completion | Dropouts | Completion | Dropouts |
| **1** | 0 | 0.014435 | 0 | 0.053348 | 0 | 0.029158 | 0.954982 | 0.045018 |
| **2** | 0 | 0.067013 | 0 | 0.080950 | 0.927137 | 0.072863 | 0.954982 | 0.045018 |
| **3** | 0 | 0.094217 | 0.877676 | 0.122324 | 0.927137 | 0.072863 | 0.954982 | 0.045018 |
| **4** | 0.865007 | 0.134993 | 0.877676 | 0.122324 | 0.927137 | 0.072863 | 0.954982 | 0.045018 |
| **5** | 0.865007 | 0.134993 | 0.877676 | 0.122324 | 0.927137 | 0.072863 | 0.954982 | 0.045018 |