Application of Finite Markov Chain to a Model of Schooling

Abba Auwalu (Corresponding author)
College of Remedial and Advanced Studies, P.M.B. 048, Kafin Hausa, Jigawa State, Nigeria
E-mail: abbaauwalu@yahoo.com
Lawan Bulama Mohammed
Bauchi State University, Gadau, Nigeria
E-mail: lawanbulama@gmail.com
Afis Saliu
Gombe State University, Gombe, Nigeria
E-mail: saliu.afis@yahoo.com

Abstract
In this paper, we focused on the application of finite Markov chain to a model of Schooling. The sample for the study was selected from one Secondary School in Nigeria. Four sets of graduates' data were used. Thus, 474 students were used as the study sample. Primary data were used as the methodology, while the knowledge of finite Markov chain was used to analyse the data collected for the study. Base on the result obtained in this study when compared with the actual values from the actual data, indicated that the model formulated, proves very efficient and can be accepted as suitable for predicting the Academic progress of the sampled school.
Keywords: Matrix; Markov chain; Probability vector; Model of schooling

1. Introduction

This case study focused on the application of finite Markov chains to a model of schooling, which gives an essential idea in operational research, particularly, the theory of Markov chains.

A Markov chain have special property that probabilities involving how a stochastic process will evolve in the future depend only on the present state of the process and so are independent of events in the past. That is; a Markov chain is generally a Mathematical model that describes, in probabilistic terms, the dynamic behaviour of certain types of processes over time. In other words, it assumes that given a transition matrix, knowledge of the current state occupied by the process is sufficient to completely describe the future probabilistic behaviour of the process. The key features of Markov chain are as follows:

Having a process that occur over fixed time interval such that;

(a) At each stage (integral number of time period), there is a finite set of possible outcomes called states.

(b) At every stage, there are fixed probabilities of going from one state to another state called transition probability.

A Markov process is a process consisting of a set of objects and a set of states such that;

(i) At any given time each object must be in a state (distinct objects need not be in distinct states),

(ii) The probability that an object moves from one state to another state (which may be the same as the first state) in one time period depends only on those two states.

The stages of the process may be finite or infinite. If the number of states is finite or countably infinite, the Markov process is a Markov chain. A finite Markov chain is one having a finite number of states.

1.1 Model of schooling

Suppose that a course consist **n** stages, each having duration of one time period (say one year). At the end of each stage, the promotion of a student to the next stage (or completion of the course) is decided by an examination. Suppose that a student may leave the course, but that once drops out, does not join the course again. At the end of a stage, a student's progress can be described by one of the following three possibilities:

(1) The student passes the examination and goes to the next stage.

(2) The student fails the examination and repeats the same stage.

(3) The student leaves the course at all.

Assuming that the probabilities of proceeding to next stage, repeating the stage, completing and leaving the course does not depend on the previous performance of the student. Under these assumptions, the progress of a student under going such a course, until the student either drops out or complete all the **n** stages, can be describe by a Markov chain with states: 0,1,2,….,**n**,**n**+1. Where; state 0, corresponds to a student entering the course in it is first stage, state **n**+1, to a student leaving the course and the intermediate state **i**, (1≤**i**≤**n**), to a student successfully completing the first **i** stages. The transition matrix of the above assumptions has the form;

Table 1.

$$\begin{pmatrix} r_1 & p_1 & 0 & . & . & . & 0 & 0 & q_1 \\ 0 & r_2 & p_2 & . & . & . & 0 & 0 & q_2 \\ 0 & 0 & r_3 & . & . & . & 0 & 0 & q_3 \\ . & . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . & . \\ 0 & 0 & 0 & . & . & . & r_n & p_n & q_n \\ 0 & 0 & 0 & . & . & . & 0 & 1 & 0 \\ 0 & 0 & 0 & . & . & . & 0 & 0 & 1 \end{pmatrix}$$

Where; p_i, q_i and r_i, are the transition probabilities associated with the Possibilities; (a), (b) and (c) above, at stage (i-1).

1.1.1 **Model of schooling in secondary school level**

In secondary school program, it is considered one time period should be one year and define states 0 through 8 as the conditions of being; entering the program in it is first stage, JS 1, JS 2, JS 3, SS 1, SS 2, SS 3, graduated as SSCE holder, and dropped, respectively. With the assumption that discharged (dropped) students never re-enter the program and graduates remain graduates. The transition matrix will also have the same form as that in table 1. The aim and objectives of this research work is to formulate a Mathematical model of schooling by using the knowledge of Markov chain, which could be use in predicting future academic performance and progress of students.

1.1.2 **Definition of Terms**

Transition matrix: This is a square matrix with non-negative entries and i^{th} row sums to one.

Stochastic matrix: This is a matrix with the property that each row is a probability vector. While, if each row and each column of a matrix is a probability vector, the matrix is said to be double stochastic. Stochastic matrix are also called Transition matrix.

Eigenvalues and Eigenvectors: Let $\mathbf{A} = (a_{ij})$ be an n x n square matrix. Then λ (complex number) is said to be an eigenvalue of the matrix **A** if only if there exist a non null vector **V** of order n, such that $\mathbf{V'A} = \lambda\mathbf{V'}$. Thus, the eigenvalues of a matrix A are nothing but the roots of the nth degree algebraic equation $|\mathbf{A}-\lambda\mathbf{I}| = 0$, known as **characteristic equation.** With each eigenvalue, we can distinguish two sets of eigenvectors. The first set is the set of vectors **V** that satisfies the equation $\mathbf{V'A} = \lambda\mathbf{V'}$. This set of vectors is called **left-eigenvectors** of **A** associated with eigenvalue λ. The other set, is the set of vectors **W**, such that the equation $\mathbf{AW'} = \lambda\mathbf{W'}$ is satisfied. This set of vectors is called **right-eigenvectors** of **A** associated with eigenvalue λ. **Multiplicity of eigenvalues:** An eigenvalue of a matrix has two types of multiplicities; algebraic multiplicity and geometric multiplicity. An algebraic multiplicity of the eigenvalue λ is the number of roots of the characteristic equation, $|\mathbf{A}-\lambda\mathbf{I}| = 0$. While a geometric multiplicity of the eigenvalue λ is the dimension of the eigenvector $\mathbf{U}_\lambda$ (left) corresponding to it. In fact, the geometric multiplicity of an eigenvalue never exceeds it is algebraic multiplicity. Also, if the algebraic multiplicity is just one, then it is geometric multiplicity is also one. However, the reverse does not hold.

Diagonal matrix: This is a square matrix $\mathbf{A} = (a_{ij})$ whose entries above and below the main diagonal are all zero, that is $a_{ij} = 0$ for all $i \neq j$. Also, an n x n square matrix **A** is said to be diagonalizable if there exist another matrix **U** such that $\mathbf{UAU^{-1}} = \mathbf{D}$, where **D** is a diagonal matrix. In other word, the rows of **U** are the left eigenvectors of **A** and the columns of $\mathbf{U}^{-1}$ are the right eigenvectors of **A**, corresponding to the respective eigenvalues. Also, the diagonal elements of **D** are the respective eigenvalue of the matrix **A.** Now, let $\mathbf{U}^{(i)}$ be ith row vector of the matrix **U** and $\mathbf{V}^{(i)}$ the ith column vector of the **matrix** $\mathbf{U}^{-1}$. Hence,

$$\mathbf{A} = \mathbf{U}^{-1}\mathbf{D}\mathbf{U}, \quad \therefore \quad \mathbf{A} = \sum_{i=1}^{n} \lambda_i \mathbf{V}^{(i)}\mathbf{U}^{(i)}$$

This expression is called **SPECTRAL** representation of the matrix **A,** it is important in the fact that it implies

$$\mathbf{A}^m = \sum_{i=1}^{n} \lambda_i^{m}\mathbf{V}^{(i)}\mathbf{U}^{(i)}$$

Inverse matrix: The matrix **B** is said to be the inverse of the matrix **A** if $\mathbf{AB} = \mathbf{BA} = \mathbf{I}$, where **I** is the identity matrix of appropriate order.

Singular and Non-singular matrix: The square matrix **A** is called is non–singular if and only if it possesses an inverse. Otherwise, it is singular matrix. In other words, a matrix whose determinant is zero is called a singular. Also, a matrix can be inverted if and only if it is non-singular matrix.

Powers of stochastic matrices: Denote the nth power of a matrix **P** by $\mathbf{P}^n = (P_{ij}^{(n)})$. If **P** is stochastic, then $P_{ij}^{(n)}$ represents the probability that an object moves from state i to state j in n time periods. It follows that $\mathbf{P}^n$ is also a stochastic matrix.

Denote the proportion of objects on state i at the end of the nth time period by $x_i^{(n)}$, and designate $\mathbf{x}^{(n)} = (x_i^{(n)}, x_2^{(n)},..., x_N^{(n)})$ the fixed (distribution) vector for the end of the nth time period. Accordingly, $\mathbf{x}^{(0)} = (x_i^{(0)}, x_2^{(0)},..., x_N^{(0)})$ represents the proportion of objects in each state at the beginning of the process. $\mathbf{x}^{(n)}$ is related to $\mathbf{x}^{(0)}$ by the equation $\mathbf{x}^{(n)} = \mathbf{x}^{(0)}\mathbf{P}^n$.

Regular matrix: Suppose **A** is an m x m matrix, then **A** is regular matrix if we can find an integral value n, such that $\mathbf{A}^n$ is a positive matrix. In other words, matrix **A** is regular if one of its powers contains only positive element.

Markov chain: This is a sequence of discrete random variables; x_0, x_1, x_2,...with the property that, the conditional probability distribution of x_{n+1} given $x_0, x_1, x_2, ..., x_n$ depends only on the value of x_n , but not further on $x_0, x_1, x_2, \ldots, x_{n-1}$. That is, for any given set of values h, i,….j in discrete space,

$\Pr(x_{n+1} = j / x_0 = h,...,x_n = i) = \Pr(x_{n+1} = j / x_n = i) = P_{ij}$, for n = 0,1,2,..., and i, j = 1,2,3,… The conditional probability P_{ij} is called **Transition probability**.

Finite Markov chain: This is a discrete time parameter stochastic process in which the future state of the system is dependent only on the present state and is independent of the past history and the number of states are finite or countably infinite.

Absorbing Markov chain: This is a Markov chain that has one or more states called **absorbing states** for which transition from such a state to any other state is not possible. If a state i is an absorbing state, then $P_{ii} = 1$ and $P_{ij} = 0$ for all $j \neq i$. In other words, a state a_i is called absorbing if the system remains in the state a_i once it enters there. A state which is not absorbing is called **transient.**

Regular Markov chain: A Markov chain is called regular if for some integer power, n, of the transition matrix, the elements of the matrix are positive. That is, a Markov chain is regular if there is sometimes at which it is possible to be in any states regardless of the starting state.

Ergodic Markov chain: A Markov chain is said to be Ergodic if it is possible to go from every state to the other. Obviously, a regular Markov chain is Ergodic, because if the nth power of the transition matrix is positive, it is possible to move from every state to every state in n steps. On the other hand, an Ergodic chain is not necessarily regular.

2. Preliminaries

The general concept of set theory throughout the twentieth century have invaded an ever widening portion of mathematics, and few branches has been as thoroughly influenced by the trend as has the theory of probability. The theory of probability began in seventeenth century in France, when two great mathematicians; Blaise pascal and Pierre de Fermat corresponds over two problems from games of chance. Also many scientists like Borel (1909), Gibbs (1901), Francis Galton (1822-1911) and Karl Pearson (1857-1936) to mention but a few have greatly contributed to the concept of probability. Most study of probability deals with independent trials processes. Modern theory of probability studies chance processes for which the knowledge of previous outcome influences predictions for future experiments.

In 1907, Andrew Andreyevich Markov (1856-1922), began the study of chance processes for which the knowledge of each outcome is relevant only in predicting the next outcome. These processes were called **Markov chains**, in respect of the Russian mathematician, A.A Markov, who helped to develop the theory of stochastic processes. He also, developed Markov chains as a natural extension of sequences of independent random variables.

In his paper, A.A Markov (1908) proved that for a Markov chain with positive transition probabilities and numerical states, the average of the outcomes converges to the expected value of the limiting distribution (fixed vector). In a later paper, (1913), Andrew Markov proved the central limit theorem for such chains. He also choose a sequence of 20,000 letters from pushkin's Eugene onegin (a novel) to see if this sequence can be approximately considered as a simple chain, and he obtained the Markov chain with transition matrix below

$$\begin{pmatrix} 0.128 & 0.872 \\ 0.663 & 0.337 \end{pmatrix}$$

With fixed vector (0.432, 0.568) indicating that we should expect 43.2 percent vowels and 56.8 percent consonants in the novel. This was borne out by the actual count.

According to Moore yakel (1938), the probability model for a sequence of random observations entails a different random variable for each observation. To be more explicit, let us consider a random variable x_1 which records the outcome of the first observation, x_2 which records the second outcome, and so on. Thus $x_1, x_2, x_3, ..., x_{10}$ records the first ten outcomes of a developing process. This set of values for the random variables is called the state space, and each individual value is called a state of the process.

In the same way, Laurie Snell (1939), added that; the study of process for which the knowledge of each outcome is relevant only in predicting the next outcome is called Markov chains. He described Markov chain as follows:

Given a set of states; $S = \{s_1, s_2,..., s_n\}$. The process starts in one state of these states at time zero and move successfully from state to another at unit time intervals. Each move is called a step. The probability p_{ij} means the process moves state s_i to state s_j which depend only on the state s_i occupy before the step.

Moore Yakel further added that, the random variables $x_1, x_2, x_3, ..., x_n, ...$ are called Markov chain if they satisfy Markov principle which is as follows; $Pr(x_n = j / x_{n-1} = k, x_{n-2} = l,..., x_2 = t, x_1 = u) = Pr(x_n = j / x_{n-1} = k)$, for whatever the states $\{j, k, l, ..., t, u\}$ and time $n = 0, 1, 2,...$

Howard (1923), provides a square description of a Markov chain as a frog jumping around on a set of lily pads. The frog starts on one of the pads and then jumps around from lily pad to another with appropriate transition probabilities.

According to Kemeny, Snell and Thompson, the land of OZ is blessed by many things but not by good weather. Thus, they never have two nice days in a row, if they have a nice day, they are just as likely to have snow or rain, and only half of the time is a change to a nice day. Assuming that, it is a nice day in the land of OZ, with this information, we form Markov chain as follows: Taking the states as the kinds of weather; rain (R), nice day (N), and snow (S). From the above information, they determined the transition probabilities and represented it in a transition matrix as;

$$\begin{pmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \end{pmatrix}$$

In fact, there are many areas where the knowledge of Markov chain is applicable. These areas includes: Biology (genetic), Sociology (social mobility), Business (production, replacement, quality control, personnel screening test) as well as the physical sciences, among which is it is application to a model of schooling.

In the sequel, we will make use of the following Theorems:

Theorem 2.1 Let **P** be transition matrix at a Markov chain. The ijth entry $\mathbf{P}_{ij}^{(n)}$ of the nth power $\mathbf{P}^n$ of **P** gives the probability that the Markov chain started in state s_i will be in state s_j.

Theorem 2.2 A stochastic matrix is Ergodic if $\lim_{n\to\infty} \mathbf{P}^n$ exists: that is, if each $\mathbf{P}_{ij}^{(n)}$ has a limit as $n \to \infty$.

We denote the limit matrix, (necessary a stochastic matrix), by **L**. The components of $\mathbf{x}^{(\infty)}$, defined by the equation $\mathbf{x}^{(\infty)} = \mathbf{x}^{(0)}\mathbf{L}$, are the limiting state distributions and represent the approximate proportions of objects in the various states of a Markov chain after large number of time periods.

Theorem 2.3 A stochastic matrix is Ergodic if and only if the only eigenvalue λ of magnitude 1 is 1 itself and, if $\lambda = 1$ has a multiplicity k, there exist k linearly independent (left) eigenvectors associated with this eigenvalue.

Theorem 2.4 If every eigenvalue of a matrix **P** yields linearly (left) eigenvectors in number equal to it is multiplicity, then there exists a non-singular matrix **M,** whose rows are left eigenvectors of **P**, such that $\mathbf{D} = \mathbf{MPM}^{-1}$ is a diagonal matrix. The diagonal elements of D are the eigenvalues of **P**, repeated according to multiplicity. For a diagonalizable, Ergodic, n x n matrix **P**, the limit matrix **L** is calculated as;

$$\mathbf{L} = \mathbf{M}^{-1}(\lim_{n\to\infty} \mathbf{D}^n)\mathbf{M}.$$

Theorem 2.5 If a stochastic matrix is regular, then 1 is an eigenvalue of multiplicity one, and all other eigenvalues λ_i satisfy $|\lambda_i| < 1$.

Theorem 2.6 A regular matrix is Ergodic.

3. Results and Discussion

The population of the study was made up of four sets of graduates in one Government College. The sample was made up of four hundred and seventy four (474) students of the College.

The Data for the study was obtained from the school records for some years back and are given in tabular form below:

1(i): **1997-2003 SET**

	JS1	JS2	JS3	SS1	SS2	SS3	G
NEC	100	96	92	90	89	87	67
NPR	90	91	85	85	86	67	67
NRP	05	02	04	03	02	15	00
NDO	05	03	03	02	01	05	00
NNA	100	06	01	05	04	01	00

1(ii): **1998-2004 SET**

	JS1	JS2	JS3	SS1	SS2	SS3	G
NEC	67	64	62	59	60	66	60
NPR	58	62	55	54	55	60	60
NRP	03	01	05	03	02	02	00
NDO	06	01	02	02	03	04	00
NNA	67	06	00	04	06	11	00

1(iii): **1999-2005 SET**

	JS1	JS2	JS3	SS1	SS2	SS3	G
NEC	120	98	90	80	88	80	65
NPR	90	88	80	75	78	65	65
NRP	05	06	06	01	02	10	00
NDO	25	04	04	04	08	05	00
NNA	120	08	02	00	13	02	00

1(iv): **2000-2006 SET**

	JS1	JS2	JS3	SS1	SS2	SS3	G
NEC	98	100	95	90	85	78	68
NPR	90	95	85	83	75	68	68
NRP	05	04	06	06	05	06	00
NDO	03	01	04	01	05	04	00
NNA	98	10	00	05	02	03	00

Where; NEC = Number of students in each class, NPR = Number of students progressing, NRP = Number of students repeating, NDO = Number of students dropping out, NNA = Number of students newly admitted, G = Number of students graduating.

The first row of the tables represent the classes for each column, the second row represent the number of the students in each class, the third row represent the number of students proceeding to the next class, the fourth row represent the number of students repeating a class, the fifth row represent the number of students dropping out from a class, while the sixth row represent the number of students that join the school newly, and eighth column, G, denotes the number of students that graduated or complete successfully at the end of the programme.

The secondary school course of education, in Nigeria, consists of six (6) styles, each having duration of one year (one time period). At the end of every year, the promotion of a student to the next stage or completing the programme is decided by an examination. It is also seen that some student leaves the course, but once discharged (dropped out), the student would not join again. Thus, at the end of a stage, students' academic progress is given as in tables above, by the following possibilities; the students passes the examination and proceed to the next stage, students that fails the examination repeat the same stage and students that leave the course altogether before the examination is considered to have dropped out during the next stage. It is assumed that the probabilities of promotion to the next stage, repetition of the current stage and leaving the course do not depend on the previous performances of the students. Under these assumptions, the progress of a student undergoing secondary school course of education until a student either drops out or completes all the six (6) stages can be described by a Markov chain with states; 0, 1, 2, 3, 4, 5, 6, 7, and 8. where state 0 represent the arrival time for the new stage, states 1 through 8 represent the conditions of being; JS 1, JS 2, JS 3, SS 1, SS 2, SS 3, graduated as SSCE holder, dropped, respectively. Applying the above principle to 1(i), (ii), (iii) and (iv) lead to the matrices below which is called transition matrix:

(i)

$$\begin{pmatrix} 0 & 100 & 6 & 1 & 5 & 4 & 1 & 0 & 0 \\ 0 & 5 & 90 & 0 & 0 & 0 & 0 & 0 & 5 \\ 0 & 0 & 2 & 91 & 0 & 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 4 & 85 & 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 & 3 & 85 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 & 2 & 86 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 15 & 67 & 5 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

(ii)

$$\begin{pmatrix} 0 & 67 & 6 & 0 & 4 & 6 & 11 & 0 & 0 \\ 0 & 3 & 58 & 0 & 0 & 0 & 0 & 0 & 6 \\ 0 & 0 & 1 & 62 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 5 & 55 & 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 & 3 & 54 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 & 2 & 55 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 & 60 & 4 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 0 & 120 & 8 & 2 & 0 & 13 & 2 & 0 & 0 \\ 0 & 5 & 90 & 0 & 0 & 0 & 0 & 0 & 25 \\ 0 & 0 & 6 & 88 & 0 & 0 & 0 & 0 & 4 \\ 0 & 0 & 0 & 6 & 80 & 0 & 0 & 0 & 4 \\ 0 & 0 & 0 & 0 & 1 & 75 & 0 & 0 & 4 \\ 0 & 0 & 0 & 0 & 0 & 2 & 78 & 0 & 8 \\ 0 & 0 & 0 & 0 & 0 & 0 & 10 & 65 & 5 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

(iv)

$$\begin{pmatrix} 0 & 98 & 10 & 0 & 5 & 2 & 30 & 0 & 0 \\ 0 & 5 & 90 & 0 & 0 & 0 & 0 & 0 & 3 \\ 0 & 0 & 4 & 95 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 6 & 85 & 0 & 0 & 0 & 4 \\ 0 & 0 & 0 & 0 & 0 & 83 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 5 & 75 & 0 & 5 \\ 0 & 0 & 0 & 0 & 0 & 0 & 6 & 68 & 4 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

In order to bring the result closer to reality, we form table 3 by adding (i), (ii), (iii) and (iv), respectively. We get:

$$\begin{pmatrix} 0 & 385 & 30 & 3 & 14 & 25 & 17 & 0 & 0 \\ 0 & 18 & 328 & 0 & 0 & 0 & 0 & 0 & 39 \\ 0 & 0 & 18 & 336 & 0 & 0 & 0 & 0 & 9 \\ 0 & 0 & 0 & 21 & 305 & 0 & 0 & 0 & 13 \\ 0 & 0 & 0 & 0 & 13 & 297 & 0 & 0 & 9 \\ 0 & 0 & 0 & 0 & 0 & 11 & 294 & 0 & 17 \\ 0 & 0 & 0 & 0 & 0 & 0 & 33 & 260 & 18 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 4 \end{pmatrix}.$$

Above, the non-zero vector for each row is not a probability vector (since the sum of it is components does not sum to one). However, since the components of each row are non-negative, each vector has a unique scalar

multiple which will make it a probability vector; this can be done by multiplying each component of the vector by the reciprocal of the sum of the components of that vector. Thus, for the case of first row; let vector **v** = (0, 385, 30, 3, 14, 25, 17, 0, 0), sum of the components is 0 + 385 + 30 + 3 + 14 + 25 + 17 + 0 + 0 = 474.
This implies that; **v**/474 = (0, 0.8, 0.1, 0, 0, 0.1, 0, 0, 0), which is a probability vector for **v**. Hence, others follow in the same way and we obtained:

$$\begin{pmatrix} 0 & 0.8 & 0.1 & 0 & 0 & 0.1 & 0 & 0 & 0 \\ 0 & 0 & 0.9 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.1 & 0.9 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.1 & 0.9 & 0.9 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.9 & 0 & 0.1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.1 & 0.8 & 0.1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

This is clearly a stochastic matrix, since has each of it is row as a probability vector. Now, denote above matrix by **P**. So that the stochastic matrix **P**, depict the model upon which the research is centred.
We know check whether **P** is regular, Ergodic and also calculate limiting matrix **L** = $\lim_{n \to \infty} \mathbf{P}^n$, if it exists. Rather than raise **P** to successive higher powers to ascertain whether it is regular, let us instead determine it is eigenvalues by solving the characteristic equation; $|\mathbf{P}-\lambda\mathbf{I}| = 0$, where **P** is the above matrix, **I** is 9 x 9 identity matrix, λ represent the eigenvalues of **P** and $|\mathbf{P}-\lambda\mathbf{I}|$ is the determinant of $\mathbf{P}-\lambda\mathbf{I}$. Hence,

$$|\mathbf{P}-\lambda\mathbf{I}| = \begin{pmatrix} -\lambda & 0.8 & 0.1 & 0 & 0 & 0.1 & 0 & 0 & 0 \\ 0 & -\lambda & 0.9 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -\lambda+0.1 & 0.9 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -\lambda+0.1 & 0.9 & 0.9 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -\lambda+0.1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -\lambda & 0.9 & 0 & 0.1 \\ 0 & 0 & 0 & 0 & 0 & 0 & -\lambda+0.1 & 0.8 & 0.1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\lambda+1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\lambda+1 \end{pmatrix}.$$

This implies
$(1-\lambda)^2 (0.1-\lambda)^4 (0-\lambda)^3 = 0$. Thus; $\lambda_1 = 1$ (double root), $\lambda_2 = 0.1$ (three times), $\lambda_3 = 0$ (four times). Hence, by theorem 2.5, **P** is not regular.
Now, to find the corresponding left eigenvectors **U** for the eigenvalues; λ_1, λ_2 and λ_3, we solve the system of equation $\mathbf{U'P} = \lambda\mathbf{U'}$, for each λ, where $\mathbf{U} = (u_1, u_2, u_3, u_4, u_5, u_6, u_7, u_8, u_9)$.
For $\lambda_1 = 1$: We obtain the left eigenvectors by solving the equations; $\mathbf{UP} = \lambda_1\mathbf{U}$.
We obtained two linearly independent left eigenvectors **(**0, 0, 0, 0, 0, 0, 0, 1, 0) and **(**0, 0, 0, 0, 0, 0, 0, 0, 1) corresponding to $\lambda_1 = 1$. Hence, by theorem 2.3, **P** is Ergodic. We also obtained the other left eigenvectors corresponding to $\lambda_2 = 0.1$ and $\lambda_3 = 0$ respectively as: (-36, 4, -5, 45, 0, 0, 0, 0, 0) and (0, 0, 0, 0, 0, 0, 9, -8, -1).
But, $\lambda_2 = 0.1$ possesses the four linearly independent eigenvectors according to it is algebraic multiplicity as: (-36, 4, -5, 45, 0, 0, 0, 0, 0), **(**0, -36, 4, -5, 45, 0, 0, 0, 0), **(**0, 0, -36, 4, -5, 45, 0, 0, 0), and **(**0, 0, 0, -36, 4, -5, 45, 0, 0).
Similarly, $\lambda_3 = 0$ possesses the three linearly independent eigenvectors according to it is algebraic multiplicity as: (0, 0, 0, 0, 0, 0, 9, -8, -1), (0, 0, 0, 0, 0, 9, -8, -1, 0) and (0, 0, 0, 0, 9, -8, -1, 0, 0).
By theorem 2.4, **P** is diagonalizable with non-singular matrix **M** whose rows are the left eigenvectors of **P**, and Matrix **D** having the eigenvalues of **P** as the diagonal elements repeated according to it is multiplicity. Thus;

$$\mathbf{M} = \begin{pmatrix} -36 & 4 & -5 & 45 & 0 & 0 & 0 & 0 & 0 \\ 0 & -36 & 4 & -5 & 45 & 0 & 0 & 0 & 0 \\ 0 & 0 & -36 & 4 & -5 & 45 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 9 & -8 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 9 & -8 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 9 & -8 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$ and

$$\mathbf{D} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Since, matrix **P** is diagonalizable, Ergodic, 9 x 9, hence the limiting matrix **L** exists, where
$\mathbf{L} = \lim_{n\to\infty} \mathbf{P}^n = \mathbf{M}^{-1}(\lim_{n\to\infty} \mathbf{D}^n)\,\mathbf{M}$
Using spectral representation on the matrix **D**, we have;

$$\lim_{n\to\infty} \mathbf{D}^n = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

By Calculation, we obtained,

$$\mathbf{M}^{-1} = \begin{pmatrix} -0.028 & -0.003 & 0.004 & 0.136 & 0.138 & 0.12 & 0.122 & 0.763 & 0.232 \\ 0 & -0.028 & -0.003 & -0.014 & 0.125 & 0.125 & 0.125 & 0.872 & 0.126 \\ 0 & 0 & -0.028 & 0.012 & -0.005 & 0.136 & 0.121 & 0.216 & 0.121 \\ 0 & 0 & 0 & 0.111 & 0.099 & 0.1 & 0.1 & 0.9 & 0.1 \\ 0 & 0 & 0 & 0 & 0.111 & 0.099 & 0.101 & 0.9 & 0.101 \\ 0 & 0 & 0 & 0 & 0 & 0.111 & 0.099 & 0.901 & 0.099 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.111 & 0.889 & 0.111 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$ Hence,

$$\mathbf{L} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.77 & 0.23 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.87 & 0.13 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.88 & 0.12 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.9 & 0.1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.9 & 0.1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.9 & 0.1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.89 & 0.11 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

The above matrix **L** is basis upon which the research is centred.
In order to determine the academic progress of the sampled school, we use the following equations:
(a) $\mathbf{X}^{(n)} = X^{(0)}\mathbf{P}^n$
(b) $\mathbf{X}^{(\infty)} = X^{(0)}\mathbf{L}$
Where: $\mathbf{X}^{(n)}$ is the distribution (fixed) vector for the end of the nth time period, $\mathbf{X}^{(0)}$ is the initial probability vector, P^n is the nth power of the stochastic matrix **P**, $\mathbf{X}^{(\infty)}$ is the probability on the long run and **L** is the limiting matrix.
The academic progress could be determine over a short and long periods. For instance, suppose there are 90 students in year one. How many students could the school expect to proceed to year two?
The answer to this is that; since there are 90 students in year one, the probability vector is $\mathbf{X}^{(0)} = (0, 1, 0, 0, 0, 0, 0, 0, 0)$, using $X^{(n)} = X^{(0)}\mathbf{P}^n$, but n = 1 (one year), then $\mathbf{X}^{(1)} = X^{(0)}\mathbf{P}$, where; $\mathbf{X}^{(1)}$ is the fixed vector for the end of year one, and **P** is the stochastic matrix. Therefore, $\mathbf{X}^{(1)} = (0, 0.1, 0.9, 0, 0, 0, 0, 0, 0)$. Thus; 10 percent of the students (about 9 students) will repeat year one, while the remaining 81 students will proceed to year two.

Now, suppose there are: 110, 105, 97, 85, 83, and 74 students in JS 1, JS 2, JS 3, SS 1, SS 2, and SS 3, respectively. How many students could the school expect to successfully complete the program on the long run? On the other hand, if the school has a population of 356 students in SS 3, how many students will be expected to be drop (discharge) from the school?

The answer to this is that; there is 110 + 105 + 97 + 85 + 83 + 74 = 554 students currently, and $\mathbf{X}^{(0)}$ = (0, 0.20, 0.19, 0.18, 0.15, 0.15, 0.13, 0, 0).

Using $\mathbf{X}^{(\infty)} = X^{(0)}\mathbf{L}$, we have $\mathbf{X}^{(\infty)}$ = (0, 0, 0, 0, 0, 0, 0, 0.889, 0.111). Thus; 88.9 percent of the students (or about 493 students) will successfully graduate or complete their schooling and 11.1 percent of the students (or about 61 students) will drop out from the programme on the long run.

If on the other hand, there are 356 students in SS 3, then $X^{(0)}$ = (0, 0, 0, 0, 0, 0, 1, 0, 0), and $\mathbf{X}^{(\infty)} = X^{(0)}\mathbf{L}$ = (0, 0, 0, 0, 0, 0, 0, 0.89, 0.11). Thus; 11 percent of the 356 students (or about 39 students) will be dropped from the school with remaining 317 eventually becoming SSCE holders.

4. Conclusion

From the results obtained in this work, it is evident that, the study of chance processes for which the knowledge of each outcome is relevant only in predicting the next outcome. Also, the result revealed that the academic performance and progress of the school, upon which the research is centred, is above average. All these are achieved through the model formulated by applying the idea of finite Markov chain.

Based on the result obtained, when compared with the actual values from the actual data, indicated that the model formulated can be accepted as suitable for predicting the academic progress of the sampled school.

References

Abubakar, U.Y. (2005), *Homogeneous and non-homogeneous Markov chain models for Asthma,* ABACUS: The journal of Mathematical Association of Nigeria (MAN), vol. 32(2B), 262-270.

Bronson, R. *Schaum's Outline of Theory and Problems of Operations Research,* 2nd Edition, McGraw-Hill Inc., New York, 1982.

David, S. M. and James, W. Y. *Applicable Finite Mathematics,* 3rd Edition, Houghton Mifflin Company, London, 1974.

Gillespie, C. C. *Dictionary of Scientific Biography,* vol. II, Scribner's Sons, Canada, 1970.

Iosifescu, M. *Finite Markov Process and their Applications,* 6th Edition, John Wiley and Sons, New York, 1980.

LiberMann, G. J., et al., *Introduction to Operations Research,* 3rd Edition, Holden Day Inc., Oakland, 1982.

Miller, H. D. and Cox, D. R. The *Theory of Stochastic Processes,* Chapman and Hall, Boston, 1992.

Novel, B. and Daniel, J. W. *Applied Linear Algebra,* 3rd Edition, Prentice Hall, New Jersey, 1988.

Osuala, E. C. *Introduction to Research Methodology,* Revised Edition, Ilupeju Press Ltd, Benin City, 1987.

Snell, G. T., et al., Introduction to Finite Mathematics, 2nd Edition, Prentice-Hall, Eagle-wood Cliffs, New Jersey, 1974.

Thomas, H. A. *Dynamic Probabilistic System,* vol. I, John Wiley and Sons, New York, 1971.

Yakasai, B. M. and Hussain, N. *MATRICES*; A Handbook for Applied Sciences Students, First Edition, Gidan Dabino Publishers, Jakara, Kano, 2004.