An Application of Markov Modeling to the Student Flow in Higher Education in Sudan

Dr. Rahmtalla Yousif Adam

Department of Statistics, University of Tabuk, Kingdom of Saudi Arabia

Abstract: ***This paper focuses on modeling students flow in educational system with a stochastic process depending mainly on Markov Chains to predict the number of graduate students for the coming years. This paper curried out in the high education system in Sudan under the enlargement of intakes during (2001-2011). The model has been applied to estimate the predicting number and ratios of students in different levels for graduate students after studying period without delaying, the ratio of whole graduations, the delaying ratio and dismissed students over the batches. The results indicated that The ratio of delaying seems to be decreasing all over the batches (61%,45%,42%,33%). From the mean absolute error (MAE) test, the model shows that there is homogeneous and no significant difference in comparing the actual numbers with predicted numbers.***

Keywords: Markov Chains; Transition matrix; batches; stochastic process, Kordofan

1. Introduction

The period from (1990 – 2011) witnessed an enlargement in high education by establishment of new universities and high institutes, both in the capital and the states. Therefore, the number of intake of students was increased. This leads to expansion of foundation of the new universities in the states. The paper problem, there is no strategic plan for graduations. Also there are many students delaying and dismissed according to academic regulations. In addition, they may face difficulties adjusting to the university environment and regulations. The main objectives of this paper are to investigate the flow of students in the Faculty of Education in Kurdofan University via Markov Chains analysis. A typical B.Sc. program requires 132 credit hours and can be completed in 8 semesters of studies i.e. 4 years, assuming that the student enrolls for every semester. The student must complete the registration procedure personally according to the time given by the college. The registration should not be later than the end of the 1st week from the commencing of the semester. The student should satisfy all the needs of registration for each semester. Many applications of Markov chains technique occur in educational system such that the paper of F.W.O Saporu and W.P.T Chinamo (1992) addressed by Statistical analysis of data from University of Zimbabwe Educational System and they described the educational advancement of student through the undergraduate degree programme. The paper has reported valuable insights as a result of using Markov Analysis.

The rest of the paper is structured as follows: In Section two we outline the Markov modeling estimation of transition probabilities. In section three we present the source of data. Section four presents and discusses the results. In section five summarizes the paper and concludes.

2. Markov Chains Modeling

The theory of stochastic process deals with system which develops in time or space in accordance with probabilistic laws and its concept is based on expanding the random variable concept to include time. The function $X_{(t,s)}$ is called a stochastic process, when X random variable a function of S possible outcomes of an experiment (state space), t is the parameter set of process (time), so that the set of possible values of an individual random variables $X_{n(xt)}$ of a stochastic process $X_{n,} n \geq 1, X_{(t),} t \in T$ is known as it's state space. Markov Chains are the simplest mathematical models for the random phenomena evolving in time.

The stochastic process with discrete parameter space denoted by $X_{n,}$ (n = 0,1,2,….) or the stochastic process with continuous parameter $X_{n,} n \geq 0$ is called the Markov chain . If $X_n = i,$ this means that process in state i at time n supposing that, the process in the state (i) then there's a constant probability (P_{ij}) the process can be in the state j as in the coming step:

$$P[X_{n+1,} = j / X_n = i, X_{n-1} = i, n-1 \ldots\ldots. X_1 = i, X_0] = P_{ij} \quad \ldots\ldots..(1)$$

For all states $i_1, \ldots.., i_{n-1,} i, j$ and $n \geq 1$. The process (1) represents Markov chain and illustrate that the conditional distribution for any future state is n+1; with repeated to previous state $X_0; X_1; \ldots\ldots, X_{n-1}$ and the present state X_n is independent from previous state only depends on present state. This property called Markov property.
Since the probability Pij is non-negative and the process must transfer to some state then

$$\sum_{j=0}^{k} Pij = 1 \quad 0 \le Pij \le 1 \forall i, j \ge 0 \quad \text{........ (2)}$$

2.1 Probability Transition Matrix

If there is a limit state n elements then transition probability for i and j values can organize P matrix called probability transition matrix. The transition matrix

$$P_{ij} = \Pr(X_{n+1} = j / X_n = i$$

satisfy $P_{ij} > 0$, $\sum P_{ij} = 1$ for all j

These probabilities can be written in the matrix form

$$P = \begin{pmatrix} p_{11} & p_{12} & \cdots\cdots & p_{1n} \\ p_{21} & p_{22} & \cdots\cdots & p_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ p_{n1} & p_{n2} & \cdots\cdot & p_{nn} \end{pmatrix}$$

This is called the probability transition matrix of Markov chain

2.2 The n- step transition probability

The probability transition of random processes $\{X_n : n = 0,1,2,\cdots\}$ from state i to state j after n step called the n-step transition probability and denoted by $p_{ij}^{(n)}$ defined as:

$$p_{ij}^{(n)} = P(X_{n+m} = j \mid X_m = i), \quad n \ge 0, \ i, j = 0,1,2,\ldots$$

These indicate to the probability transition of random processes from state i to state j after n step, and we can write in term of matrix $P^{(n)}$ as follow:

$$P^{(n)} = \begin{pmatrix} p_{00}^{(n)} & p_{01}^{(n)} & p_{02}^{(n)} & \cdots \\ p_{10}^{(n)} & p_{11}^{(n)} & p_{12}^{(n)} & \cdots \\ p_{20}^{(n)} & p_{21}^{(n)} & p_{22}^{(n)} & \cdots \\ \vdots & \vdots & \vdots & \vdots \end{pmatrix}$$

The interpretation of matrix is:

1- If n =1 then $p_{ij}^{(n)}$ became probability transition of random processes from state i to state j with one step denoted by p_{ij} .

2- If n = 0 then

$$p_{ij}^{(0)} = P(X_0 = j \mid X_0 = i) = \begin{cases} 1, & i = j, \\ 0, & i \ne j. \end{cases}$$

3- For all n = 0,1,.. The matrix $P^{(n)}$ became a random process convince the two above properties.

The definition of n – step transition probability leads us to Chapman – Kolmogorov equations which take the frequency relations for computing the probabilities transition after n step .This property my help us to predict and forecast for several steps or several years in the future all that depend on kind of probability transition of random processes from state i to state j after n + m step , this last step set to be in k state after n step.

2.3 Chapman – Kolmogorov equations

If $\{X_n : n = 0,1,2,\cdots\}$ Markov chain with limit m states and transition probability matrix (TPM) $P = (p_{ij})$ then :

$$p_{ij}^{(n)} = \sum_{k=1}^{m} p_{ik}^{(r)} \, p_{kj}^{(n-r)}, \quad \forall r = 1,2,\ldots,n-1.$$

2.4 Maximum Likelihood Function

We have a Markov Chain with state 0, 1 , 2 , 3 , ….. With unknown transition matrix P, the likelihood function is:

$$L = \prod_{i,j \in s} .P_{ij}^{nij}$$

n_{ij} : the number of times has state j following state i, to maximize the function:

$\sum_{i \in s} p_{ij} = 1$ Indicate that each row of transition matrix equal 1 and then:

$$\sum_{i \in s} p_{ij} = \sum \frac{n_{ij}}{\sum n_{i.}}$$

Where n_{ij} the transition count for $(i, j)^{th}$ cell and $n_{i.}$ is the i^{th} row total transition count.

So that random variable n_{ij} depends on the parameter P_{ij}

$$\ln L = \sum n_{ij} L P_{ij}$$

3. Data Sources

Data was obtained from student's intake records of Kordofan University, Faculty of Education, of batches, (2005-2006, 2006-2007, 2007-2008, and 2008-2009). The Faculty of Education applies semester system, it has three departments, but they have a unique record. historical data of student concern, the pass all throw students in the first and second term, the repeaters students in two semesters, the students who either graduated from or withdrawn academically. Simplifying assumptions are made about the process in order to make it mathematically tractable, these are enumerated below. Assume that there are four years in university = n. The student can only proceed from one part to the next. That is no double promotion mean after the end of class i in the end of the year t; the student moved to class (i+1) (success situation).The entry into the system is only through part i_0 entry is not permitted once the academic session starts. Student can remain in the same class i in the beginning of year t+1 (final situation).The student when he ended (finished) the class n in the end of year t he get out of system, (he is graduated).

For the purpose of this paper, some notions explains as follow

B(t): Number of students who came at the beginning of the year (t).

$X_i(t)$: Number of students at the level (i) at the beginning of the year (t).

$X_{ii}(t)$: Number of students who repeated at the level (i) at the beginning of the year (t+1) (repeat case).

$X_{(i+1)}(t)$: Number of students who passed from level i at the end of year t to the beginning of year (t+1) (pass at first time).

$W_i(t)$: Number of students withdrawn in level i.

$G_n(t)$: Number of students who graduated from the faculty (who came at year (t) after n years).

$\hat{G}(t)$: Number of students who graduated from faculty (Who came at year (t) after n years without delay).

Table 1: The student Advancement through Batches [(2005-2006, (2006-2007, (2007-2008, and (2008-2009)]

years / Batch (1) (2005-2006)	1st	2nd	3rd	4th	years / Batch (3) (2007-2008)	1st	2nd	3rd	4th
Pass all through $X_{(i+1)}(t)$	139	152	139	132	Pass all through $X_{(i+1)}(t)$	170	174	145	134
Repeaters $X_{ii}(t)$	130	10	10	7	Repeaters $X_{ii}(t)$	33	7	25	11
Withdrawn $W_i(t)$	13	-	3	-	Withdrawn $W_i(t)$	14	-	4	-
Total	282	162	152	139	Total	217	181	184	145
years / Batch (2) (2006-2007)	1st	2nd	3rd	4th	years / Batch (4) 20078-2009)	1st	2nd	3rd	4th
Pass all through $X_{(i+1)}(t)$	172	157	140	136	Pass all through $X_{(i+1)}(t)$	210	212	198	182
Repeaters $X_{ii}(t)$	58	16	15	4	Repeaters $X_{ii}(t)$	54	9	5	4
Withdrawn $W_i(t)$	10	2	2	-	Withdrawn $W_i(t)$	16	2	3	-
Total	240	195	157	140	Total	280	223	206	186

4. Results and Discussion

Simplifying assumptions are made about the process in order to make it mathematically tractable, these are enumerated below. Assume that there are four years in university = n.

1) The student can only proceed from one part to the next. That is no double promotion mean after the end of class i in the end of the year t; the student moved to class (i+1) (success situation).
2) The entry into the system is only through part i_0 entry is not permitted once the academic session starts.
3) Student can remain in the same class i in the beginning of year t+1 (final situation).
4) 4-The student when he ended (finished) the class n in the end of year t he get out of system, (he is graduated).
5) Following equations represent the way to calculate Probability Transition Matrix.

1- The probability of student who transfer from the class (i) to $(i+1)$ successively is:

$$p_{i,i+1(t)} = \frac{X_{i,i+1(t)}}{X_{i(t)}} \quad \text{……….. (4-1)}$$

2- The probability of students who repeated at the level (i) at the beginning of the year (i+1) (repeat case).

$$p_{i,i(t)} = \frac{X_{i,i(t)}}{X_{i(t)}} \quad \text{………… (4-2)}$$

3- The probability of students who withdrawal at the level (i).

$$p_{i,0(t)} = \frac{X_{i,0(t)}}{X_{i(t)}} \quad \text{……………… (4-3)}$$

Table 2: The corresponding Ratios of students are calculated according to equations (4-1), (4-2), (4-3)

Batch 1 (2005-2006)					*Batch 3 (2007-2008)*				
$p_{i,i+1(t)}$	0.493	0.938	0.914	0.95	$p_{i,i+1(t)}$	0.783	0.961	0.833	0.924
$p_{i,i(t)}$	0.461	0.062	0.066	0.05	$p_{i,i(t)}$	0.152	0.039	0.144	0.076
$p_{i,0(t)}$	0.046	0	0.02	0	$p_{i,0(t)}$	0.065	0	0.023	0
$\sum_{i \in s} p_{ij}$	1	1	1	1	$\sum_{i \in s} p_{ij}$	1	1	1	1

Batch 2 (2006-2007)					Batch 4 (2008-2009)				
$p_{i,i+1(t)}$	0.717	0.897	0.892	0.971	$p_{i,i+1(t)}$	0.75	0.951	0.961	0.978
$p_{i,i(t)}$	0.242	0.091	0.096	0.029	$p_{i,i(t)}$	0.193	0.04	0.024	0.022
$p_{i,o(t)}$	0.042	0.011	0.013	0	$p_{i,o(t)}$	0.057	0.009	0.015	0
$\sum_{i\in s} p_{ij}$	1	1	1	1	$\sum_{i\in s} p_{ij}$	1	1	1	1

4.2 The calculation of the rows of matrices will be as follow

1- Batch1) matrix

$$P=\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0.046 & 0.46 & 0.493 & 0 & 0 & 0 \\ 0 & 0 & 0.062 & 0.914 & 0 & 0 \\ 0.02 & 0 & 0 & 0.066 & 0.914 & 0 \\ 0 & 0 & 0 & 0 & 0.05 & 0.95 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

2- Batch2 matrix

$$P=\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0.042 & 0.242 & 0.717 & 0 & 0 & 0 \\ 0.011 & 0 & 0.091 & 0.897 & 0 & 0 \\ 0.013 & 0 & 0 & 0.096 & 0.892 & 0 \\ 0 & 0 & 0 & 0 & 0.029 & 0.971 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

3-Batch3 matrix

$$P=\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0.065 & 0.152 & 0.783 & 0 & 0 & 0 \\ 0 & 0 & 0.039 & 0.961 & 0 & 0 \\ 0.023 & 0 & 0 & 0.144 & 0.833 & 0 \\ 0 & 0 & 0 & 0 & 0.076 & 0.924 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

4-Batch4 matrix

$$P=\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0.057 & 0.193 & 0.750 & 0 & 0 & 0 \\ 0.009 & 0. & 0.040 & 0.951 & 0 & 0 \\ 0.015 & 0 & 0 & 0.024 & 0.961 & 0 \\ 0 & 0 & 0 & 0 & 0.022 & 0.978 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

4.3 Prediction of Gradations

In order to estimate the situation of student throughout the year of study and to predict the number of students after four years should follow 3 steps:

1- multiply the probability of student at the beginning of the year (t).those Xi(t) with p_{15}^{4} from matrix but there are four state for each student that (gradation ,withdrawn , third level ,or fourth level) then the result being as:

$$p_{15}^{4} * \frac{X_1(t)}{4} \quad \text{……………………………….(4-4)}$$

Secondly: in the second level after four years the student has three chances (graduation ,withdrawal ,or fourth level) then the result :

$$p_{25}^{4} * \frac{X2(t)}{3} \quad \text{……………………….. (4-5)}$$

Thirdly: in the third level after four years the student has two chances (graduation, withdrawal) then the result:

$$p_{35}^{4} * \frac{X2(t)}{2} \quad \text{…………………………….(4-6)}$$

Fourthly: the total of above results is Number of students who graduated from the faculty after 4 yeas). $G_n(t)$.

Table 3.The table represents the expected graduate student after four years

Batches	1	2	3	4
$G_4(batch)$	152	172	176	223
Graduates%	54%	76%	81%	80%

From the table (3) we notice that the percentage of graduation after four years was very low in the first Batch 54% but there was gradual increasing in percentage through batches reach up to 80%

Table 4: Students who promoted without delay, the withdrawn, and remain students in the (1st ,2nd , 3rd&4th class) after four years

	Withdrawn	1st level	2nd level	3rd level	4th level	B*	Total
Second raw of P^4	0.095	0.07	0.055	0.128	0.262	0.391	1
No. of students in Batch1	27	20	15	36	74	110	282
Second raw of P^4	0.078	0.003	0.017	0.085	0.265	0.557	1
Students of Batch2	18	1	4	19	60	126	228
Second raw of P^4	0.1	0.001	0.004	0.059	0.258	0.579	1
Students of Batch3	22	0	1	13	56	126	217
Second raw of P^4	0.093	0.002	0.007	0.037	0.191	0.67	1
Students of Batch4	26	1	2	10	53	188	280

From the table (4)the column B* represents the prediction of students who graduate without delay after they spending four years among the batches so the follow-up the fluctuation of student becomes clear. for example in the ratio of batch1 (0.095, 0.07, 0.055, 0.128 and 0.262) and only 110 students have been graduated. For further details of student ratio of delaying we can construct a table from equations:

$$\frac{B^*}{G_4(bach)}$$ = the ratio of graduate student without delay ….(4-7)

$$\frac{B^*}{B_4(t)}$$ = the ratio of graduate students of those from beginning of the year (t).(4-8)

Table 5: Students graduate and ratio of delaying

	B^*	$B_{(t)}$	$G_4(batch)$	$\frac{B^*}{G_4(bach)}$	Delay Ratio B*	$\frac{B^*}{B_4(t)}$	Delay Ratio
Batch1	110	282	152	72%	28%	39%	61%
Batch2	126	228	172	73%	27%	55%	45%
Batch3	126	217	176	71%	29%	58%	42%
Batch4	188	280	223	84%	16%	67%	33%

From the table 5.the ratio of graduated students (without delaying) of the predicted number is high (72, 73, 71, 84,) compared with the ratio of graduated students (without delaying) of the total student number at the beginning of year.

The system of Faculty of Education besides the delaying students and some students who have dismissed in batches form the table 6.below we find high in batch (1and3) 10%,10% respectively.

Table 6: The dismissed student after four years

	Total st. in beginning of year	*The ratio*	*No. of lost st.*
Batch1	282	10%	27
Batch2	228	8%	18
Batch3	217	10%	22
Batch4	280	11%	31

4.3 Evaluations of Data

In order to estimate whether the estimate for the transition probability matrix over period of four years it remains constant. This was investigated by setting up the hypothesis.

$H_\circ : P_{ij}(t) = P_{ij}$ t = 0.1.2.3

$H_1 : P_{ij}(t)$ is different for each t = 0.1.2.3

Then we can test the significances by the mean absolute error (MAE) . by using the variation between different groups the actual and predicted, so that we use the F. test such that:

$$F = \frac{\max S_i^2}{\min S_i^2}$$

Table 7: The predicted & actual number of graduate's students over batches

Batches	*1*	*2*	*3*	*4*	S_i^2
The predicted graduate	152	172	176	223	903.6
The actual graduate	132	136	134	184	627.7

Then the calculated value of test is: S^2_{actual} = 627.7 and the $S^2_{predicted}$ = 903.6

Hence the data is compatible with the null hypothesis at (α = 0.5%) .

5. Conclusions

The model of Markov Chain has homogeneous probability transition matrix according to fitted test. Then the model has been compatible to estimation of graduate students compared with the intake students in duration of university study throughout the batches. So that from the study there is very low ratio of graduation student 54% of batch1 because there is high delaying in the first level the probability of repeated student is 0.46. But at other levels (2, 3, 4) low delaying probability of repeating (0.062,0.066,0.05) and withdrawals probability (0, .002, 0) respectively, so that 39 % of total registered students this ratio is very low to satisfy the needs of schools in fact the majority of the graduates are teachers brought from schools to be qualified. In the batches (2,3,4) the predicted graduates after four years have been increased gradually (76%,81%,80%).we conclude that, the reasons are , the students were taken from General Intake Office ,in the first year the college was not completely separated of the university, the academic atmosphere was not good enough and libraries were not habilitated besides the shortage of lectures.

The ratio of delaying seems to be decreasing all over the batches (61%, 45%, 42%, 33%). Referring to the certain procedures applied in the faculty, the ratio of missed students (10%,8%,10% and 11%) in spite of decreasing of delaying students, but the missed students increasing all over the batches , because of severe academic rules of cheating cases.

References

[1] Awda (1987): “Planning of Technical Education with Application on Teaching Data of Kingdom of Arabia Saudi”, *Journal of Tanmiayt Al-Rafidian*, Vol. (22), University of Mosul, IRAQ, PP.291-306.
[2] Bhat, U. N. (1972): “*Elements of Applied Stochastic Processes*”, John-Wiley & Sons, New York.
[3] Chase, C. W. (1995): “Measuring Forecast Accuracy”, *Journal of Business Forecasting*, (Fall), PP.23-25.
[4] Cinlar, E. (1975): “*Introduction To Stochastic Processes*”, Prentice-Hall Company, New Jersey.
[5] Gani, J. (1963): “Formula of Projecting Enrolments and Degrees Awarded in Universities”, *Journal of Royal Statistical Society*, Vol.(126), PP.400-409.
[6] Ibrahim, B. Y. (1995): “Planning of Higher Education using Markov Chain”, *Journal of Tanmiayt Al-Rafidian*, Vol. (48), University of Mosul, IRAQ, PP.225-236
[7] Rodovilsky, Z. & Eyck, J. T. (2000): “Forecasting with Excel”, *Journal of Business Forecasting*, (Fall), PP.22-27.
[8] Ross, S. (1981): “*Stochastic Processes*”, John-Wiley &
[9] Saporu, F. W. & Chinamo, W. P. (1992): “A Statistical Analysis of Data From University of Zimbabwe”, E*ducational System Using Markov Model*.