Supported by tetfund TERTIARY EDUCATION TRUST FUND

A MULTI-STATE MARKOV APPROACH TO MONITORING STUDENTS' ACADEMIC PERFORMANCES

B. M. Oseni* and L. A. Oderinde
Department of Statistics, Federal University of Technology, Akure, Ondo State, Nigeria
*Corresponding author: bmoseni@futa.edu.ng

Received: August 23, 2021 **Accepted:** October 07, 2021

Abstract: Monitoring students' academic performances are becoming increasingly important in the face of different technological advancement which may serve as a distraction to weak students. Performance is often monitored on a student-to-student bases which is highly time-consuming. It is also done through some common statistical analysis which has been grossly inefficient as progressions of students through the grades are not often considered. An all-encompassing way of evaluating students' academic performance and progression using multi-state model is presented and demonstrated using cumulative grade point averages (CGPA) of a set of students who are expected to graduate at the same time. The results obtained based on the data show that students are about 12 times more likely to move from probation to pass than move in the opposite direction. Also, students are about 80% likely to move from pass grade to third class than move from third class to pass. Based on these results and other results obtained, we conclude that movement from probation to second class lower is more likely than movement from second class lower to first class. Hence with some learning support being offered to students in the bottom half of the grade ladder, performances can be improved.
Keywords: Multi-state, class of graduation, transition probability, transition intensity, prevalence

Introduction

All universities monitor student's academic progress with the intention of not only meeting a high standard of learning and teaching but also to identify students whose academic progress is less than satisfactory so as to offer them appropriate learning support, resources and assistance. In the event that such students continue to make unsatisfactory academic progress, despite being offered learning support, they may be required to withdraw from the programs.

An indicator of academic performance employed by universities and employers is the cumulative grade point average (CGPA). Most universities use 5 point-scale while others use 4-point and 7-point scales. In a 5 point scale, the minimum CGPA requirement before students are advised to withdraw is mostly 1.0, while a CGPA of 2.4 or higher is considered an indicator of good academic performance. Though CGPA may not be the only factor associated with subsequent career success; it still remains the most common factor used by administrators to evaluate progression in an academic environment. Qualities such as empathy and social skills, communication skills, conflict management, leadership, collaboration, cooperation and team capabilities are also important in the environment.

A number of statistical approach are being employed in monitoring students' performance progression in terms of GPA/CGPA, most popular of which includes the ordered probit model approach as used by Erdem *et al.* (2007), regression and descriptive statistics as employed by Dayıoğlu and Türüt-Aşık (2007), t-test, Analysis of Variance and correlation among others. The multi-state approach may be considered superior owing to its ability to estimate the length of stay in a particular GPA/CGPA and to calculate the probabilities of transition either to lower or upper GPA/CGPA. Though this approach has been widely employed in medicine to examine various progressions through stages, it is relatively new or non-existing in education.

The evolution over time of longitudinal failure time data is most often described using multi-state models (Hougaard, 1999). "A multi-state model is defined as a model for a stochastic process which at any time occupies one of a set of discrete states" (Hougaard, 2000). In medicine, the states can describe conditions like healthy, diseased, diseased with complication and dead. A change of state is called a transition or an event. The state structure specifies the states in which transitions from one state to another is possible. In spite of the many advantages of applying multi-state models, these models are not often applied. Meira-Machado *et al.* (2009) states the reasons for this as daunting mathematical theory and lack of available software. As a multi-state process evolves over time, history is naturally generated. This history contains information on previous states visited, times of entry into previous states and length of stay in states. The simplest and most applied multi-state model assumes that the transition to a future state is only dependent on the present state (Markov property) and that the transition intensities are constant over time (i.e. Time homogenous transition rates).

Table 1: The States of the model based on CGPA

State	1	2	3	4	5	6	7
Class	Probation	Pass	Third Class	Second Class lower	Second Class Upper	First Class	Graduated/ Withdrawal
Range	0 – 0.99	1.00 – 1.49	1.50 – 2.39	2.40 – 3.49	3.50 – 4.49	4.50 – 5.00	Absorbing State

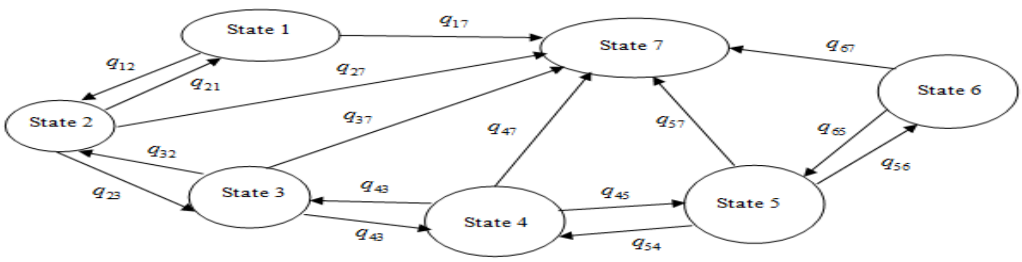

Fig. 1: Transition diagram for the model

In applying a multi-state model to academic performance monitoring, the states are partitioned according to the classes of CGPA given in Table 1. When a student graduates or withdraws from the system, it is considered that the student has been absorbed. All other states are considered non-absorbing. Visitation among the non-absorbing states is expected to be to the immediately preceding state or the succeeding state. The transition diagram depicting the expected possible movement between the states is shown in Fig. 1. State 7 is the absorbing state while the rest of the states (1-6) are transient states. A student may visit state 7 either as a result of graduation, voluntary withdrawal or non-voluntary withdrawal. The non-voluntary withdrawal may take the form of expulsion, advice to withdraw due to poor performance or even change of course of study.

Materials and Methods

Multi-state model

A discrete state space stochastic process $(X(t), t \in T)$ **where** $T = [0, \tau]$ is the period of observation and $X(t)$ is finite and consist of several states is called a multi-state process (Meira-Machado *et al*., 2009). The states in the model can be visited by individuals under observation at some periods. History F_{t^-} is created as the process evolves over time. This history consists of previous states visited, length of stay in the states, transition time into previous states and other related information (Oseni & Anjorin, 2016). The model is completely characterized by the transition probabilities and transition intensities. The transition probability is generally defined as;

$$p_{ij}(s,t) = P\,[X(t) = j \mid X(s) = i, F_{t^-}] \qquad (1)$$

Where $p_{ij}(s,t)$ is the probability that the process will be in state *j* at time *t* given that it was in state *i* at time *s*. The transition probabilities can also be represented as the $(i, j)^{th}$ entries of a matrix called stochastic matrix given by Oseni & Anjorin (2016);

$$P(t) = \begin{pmatrix} p_{11}(t) & p_{12}(t) & \cdots & p_{1n}(t) \\ p_{21}(t) & p_{22}(t) & \cdots & p_{2n}(t) \\ \vdots & \vdots & & \vdots \\ p_{n1}(t) & p_{n2}(t) & \cdots & p_{nn}(t) \end{pmatrix}. \qquad (2)$$

Each row of the stochastic matrix (2) summed up to unity and represents the probability of assessing every other state from state *i*. The transition intensities, which is the instantaneous rate of moving from state *i* to state *j*, is given by;

$$q_{ij}(t, F_{t^-}) = \lim_{\delta t \to 0} \frac{p_{ij}(t, t+\delta t)}{\delta t}. \qquad (3)$$

The transition intensity can also be put in a matrix similar to (2) to form the intensity matrix. The matrix is constructed in such a way that the sum of each row is zero and the off-diagonal elements are non-negative. Thus, the diagonal element is given by;

$$q_{ii}(t, F_{t^-}) = -\sum_{i \neq j} q_{ij}(t, F_{t^-}) \qquad (4)$$

The implication is that at every instant where the off-diagonal element, say q_{ij}, of the transition intensity matrix is positive, change of state occurs from state *i* to state *j*. This movement reduces the probability of subjects remaining at the current state *i*, thus the transition intensity of the diagonal element is represented by (4).

It may also be of interest to examine how some characteristics of the students may have affected their performances. For example, the mode of admission of the student could have an impact on their performances. While many of the students may have been admitted through UTME, some of the students could have been admitted by DE. Students who came in by DE begin their studies at 200 Level since they are expected to have covered 100 Level syllabus from their previous institution. This variable is added through a generalized linear regression to each level of the model as implemented by Jackson *et al.* (2003). A proportion hazards model which relates the covariates, say z, to the transition intensity q_{ij} is given by Marshall and Jones (1995) as;

$$q_{ij}(z) = q_{ij} \exp(\beta_{ij}^{\,T} z) \qquad (5)$$

Where β_{ij} is the vector of regression coefficients associated with the covariates *z* for the transition between states *i* and *j*. The transition intensity matrix constructed from (5) can be used to compute the transition probability matrix $P(s,t \mid z)$ with elements $p_{ij}(s,t \mid z)$ denoting the contribution of each observation to the likelihood function. Though, this may not be investigated in this work due to an insufficient number of DE students in the data used.

Maximum likelihood

The estimates of the unknown parameters in the model are determined by the method of maximum likelihood. These estimates can be computed from the transition probability matrix which depends on the unknown parameters through the Kolmogorov differential equations (Cox & Miller, 1965). When the transition intensity matrix is constant, the equations are solved by;

$$P(t) = Exp(tQ) \qquad (6)$$

Where *Q* is the intensity matrix scaled by *t*. This matrix exponential is quite difficult to evaluate as noted by Moler and van Loan (2003). Thus, the likelihood is at first computed through the product of probabilities of transition between observed states over all individual *i* and observation times *j* (Jackson *et al*., 2003).

$$L(Q) = \prod_{ij} p_{s(t_{ij})s(t_{i,j+1})}(t_{i,j+1} - t_{ij}) \qquad (7)$$

Where each component L_{ij} is the $S(t_{ij})$ th row and $S(t_{i,j+1})$ th column of the transition probability matrix *P*(*t*) evaluated at $t = t_{i,j+1} - t_{ij}$. The maximum likelihood of q_{rs} is obtained by maximizing *L*(*Q*) in terms of $\log(q_{rs})$, using the optimization algorithms as implemented in R development software function "optim".

Data

The data used for the study were obtained from students' enrolment record of the Department of Mathematical Sciences, the Federal University of Technology, Akure for 2008/2009 and 2009/2010 academic sessions and performance records of the students from the period when they were given admission to the period when they are expected to have graduated. The data cover the period 2008/2009 -2012/2013 sessions, since the students are expected to undergo their study for a period of five academic sessions for students admitted through UTME and four academic sessions for students admitted through DE. Students who were unable to complete their study in the period 2012/2013 are considered lost hence were not captured in the study. Only DE students were considered from the enrolment records of 2009/2010 session since they are the only category of students admitted in that session who are able to complete the study in 2012/2013 sessions.

The record contains the performance indicators of the students which include semester grade point averages, the cumulative grade point averages, the total load units and the total credit points. However, only cumulative grade point averages were used in the study. This is because the cumulative grade point averages is the only indicator uses both the total load units and total credit points in its computation and it is also generally accepted as an indicator of performance progression of students undergoing courses of study in Universities.

Result and Discussion

A model with state structure based on the six grades which CGPA can be classified and graduation as the absorbing sixth state, as shown in Fig. 1, is examined. The model is fitted to the data described earlier. The decision to designate graduation as the end point was based on the primary objective of this study, which is to determine the rate at which one is able to maintain a particular CGPA till the end (that is, graduation period) and for how long a student is going to maintain a particular class before she drops. A frequency table of pair of consecutive state used in summarizing the data is given in Table 2. This counts the number of times an individual in state *i* moves to state *j*.

Table 2 shows that none of the students visited state 6 and all the students eventually visited state 7 (graduating state). This implies that at no point in time do any of the students had a CGPA which placed them in the first class bracket. Only 40 out of the 55 students eventually graduated with 7 graduating in state 3 (third class), 22 graduating in state 4 (second class lower), 11 graduated in state 5 (second class upper).

Table 2: Frequency of the observed states transitions

	To State						
From State	**1**	**2**	**3**	**4**	**5**	**6**	**7**
1	5	6	0	0	0	0	0
2	5	92	9	0	0	0	0
3	0	13	140	19	0	0	7
4	0	0	13	132	8	0	22
5	0	0	0	3	59	0	11
6	0	0	0	0	0	0	0

In other to increase the rate at which the estimates computed by using "optim" converges to the maximum likelihood estimate, a crude initial estimate of the transition intensity matrix is computed at first, using some assumed intensity values given by;

$$Q = \begin{pmatrix} 0.003 & 0.002 & 0.001 & 0.004 & 0.000 & 0.000 & 0.001 \\ 0.100 & 0.014 & 0.100 & 0.101 & 0.000 & 0.000 & 0.201 \\ 0.210 & 0.102 & 0.011 & 0.011 & 0.012 & 0.000 & 0.200 \\ 0.110 & 0.001 & 0.201 & 0.011 & 0.203 & 0.001 & 0.023 \\ 0.114 & 0.189 & 0.201 & 0.013 & 0.122 & 0.001 & 0.001 \\ 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 \\ 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 \end{pmatrix} \quad (8)$$

At first, the computation of the maximum likelihood estimates of the transition intensity matrix is done without using covariates. These estimates are given in Table 3.

From the estimates in Table 3, it is clear that it is impossible for students to jump states. This implies that before progression from any state *i* to any state *k* ($k = i \pm 2$), a student must pass through a state *j* ($j = i \pm 1$). Therefore students with third class CGPA will be unable to attain second class upper CGPA unless they pass through a second class lower CGPA. The estimates also revealed that none of the students considered eventually graduated with a pass or in a state of probation/fail.

Table 3: Parameter estimates of the intensity matrix without Covariates

q_{ij}	Estimates	Confidence Interval
q_{11}	-0.54545	-1.21412, -0.24505
q_{12}	0.54545	0.24505, 1.21412
q_{21}	0.04717	0.01963, 0.11333
q_{22}	-0.13208	-0.22301, -0.07822
q_{23}	0.08491	0.04418, 0.16318
q_{32}	0.07263	0.04217, 0.12508
q_{33}	-0.21788	-0.29820, -0.15919
q_{34}	0.10615	0.06771, 0.16641
q_{37}	0.03911	0.01864, 0.08203
q_{43}	0.07429	0.04313, 0.12793
q_{44}	-0.24571	-0.33131, -0.18223
q_{45}	0.04571	0.02286, 0.09141
q_{47}	0.12571	0.08278, 0.19092
q_{54}	0.04110	0.01325, 0.12742
q_{55}	-0.19178	-0.32382, -0.11358
q_{57}	0.15068	0.08345, 0.27209

There are also some statistics being revealed when the ratios of the transition intensities are computed. For example, the ratio of q_{12} and q_{21} show that students are about 12 times more likely to move from state 1 to 2 than from state 2 to 1. This implies that students in probation are about 12 times more likely to move to a pass grade than students in pass grade moving to probation. The ratio of q_{23} and q_{21} also shows that students in pass grade are about 80% more likely to move into third class grade than moving into probation. Some other ratios also revealed that students in third class grade are about 50% more likely to move into second class lower than move into pass. Students in second class lower grade are about 60% more likely to slip into third class grade than climb into second class upper grade. The students with second class lower grade are about three times more likely to graduate with their grade than students with third class. Also, students in second class upper grade are about 20% more likely to graduate with the grade than students with second class lower grade.

Estimation of how often students migrate between the grades can be done using the transition probabilities. With these probabilities, it is possible to predict the likelihood of a student with a certain class of CGPA improving and moving into another class in some specified number of semesters. These probabilities are given in Table 4. From the Table, students in probation will move into pass grade with probability 0.39360, 0.58242, 0.66234, 0.68532 and 0.6793 after one, two, three, four and five semesters, respectively. Clearly, this probability increases as the number of semesters become bigger. This is an indication that students in probation have good chances of migrating from the grade to pass if they are allowed to continue their studies for longer semesters. Examination of other transition probabilities reveals that if the number of semesters projected for a student to move into the next grade is increased, the chances of moving into that grade equally increase. Though, it is easier to move from a lower grade to the next grade than to move from a higher grade to the next.

Table 4: Estimates of the transition probabilities after various semesters of studies

State	1	2	3	4	5	6	7
Transition probabilities after one semester (1-step)							
1	0.58820	0.39360	0.01731	0.00062	0.00001	0.00000	0.00026
2	0.03404	0.88880	0.07175	0.00371	0.00006	0.00000	0.00165
3	0.00128	0.06137	0.80993	0.08441	0.00195	0.00000	0.04105
4	0.00003	0.00222	0.05908	0.78601	0.03680	0.00000	0.11586
5	0.00000	0.00003	0.00123	0.03308	0.82625	0.00000	0.13941
6	0.00000	0.00000	0.00000	0.00000	0.00000	1.00000	0.00000
7	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	1.00000
Transition probabilities after two semester (2-step)							
1	0.35940	0.58242	0.05247	0.00377	0.00009	0.00000	0.00185
2	0.05037	0.80778	0.12268	0.01229	0.00037	0.00000	0.00651
3	0.00388	0.10494	0.66540	0.13501	0.00630	0.00000	0.08446
4	0.00020	0.00736	0.09449	0.62403	0.05944	0.00000	0.21449
5	0.00000	0.00020	0.00397	0.05344	0.68391	0.00000	0.25848
6	0.00000	0.00000	0.00000	0.00000	0.00000	1.00000	0.00000
7	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	1.00000
Transition probabilities after three semester							
1	0.23129	0.66234	0.09073	0.00978	0.00035	0.00000	0.00551
2	0.05728	0.74534	0.15892	0.02306	0.00105	0.00000	0.01436
3	0.00671	0.13593	0.55451	0.16289	0.01148	0.00000	0.12847
4	0.00051	0.01380	0.11400	0.50046	0.07226	0.00000	0.29897
5	0.00002	0.00056	0.00722	0.06496	0.56705	0.00000	0.36018
6	0.00000	0.00000	0.00000	0.00000	0.00000	1.00000	0.00000
7	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	1.00000
Transition probabilities after four semester (4-step)							
1	0.15871	0.68532	0.12558	0.01796	0.00087	0.00000	0.01157
2	0.05927	0.69481	0.18454	0.03437	0.00207	0.00000	0.02495
3	0.00929	0.15785	0.46862	0.17573	0.01657	0.00000	0.17194
4	0.00093	0.02058	0.12298	0.40543	0.07834	0.00000	0.37173
5	0.00004	0.00111	0.01043	0.07043	0.47093	0.00000	0.44706
6	0.00000	0.00000	0.00000	0.00000	0.00000	1.00000	0.00000
7	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	1.00000
Transition probabilities after five semester (5-step)							
1	0.1168	0.6793	0.1547	0.0274	0.0017	0.0000	0.0201
2	0.0587	0.6523	0.2024	0.0453	0.0034	0.0000	0.0380
3	0.0114	0.1731	0.4014	0.1788	0.0211	0.0000	0.2141
4	0.0014	0.0271	0.1251	0.3317	0.0799	0.0000	0.4347
5	0.0001	0.0018	0.0133	0.0718	0.3917	0.0000	0.5213
6	0.0000	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000
7	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	1.0000

Another important estimate which may be obtained from the intensity matrix in Table 3 is the mean sojourn time $1/q_{ii}$. The estimate is the average semesters in a single stay in a state (Jackson, 2014). The mean sojourn time for the states is given in Table 5. From the table, it is clear that students in probation have mean sojourn time of 1.83333 which implies that on the average, students are expected to be out of probation after two semesters. Out of the five states that were visited by students, the sojourn time in state 2 is the biggest. This implies that students stay in third class grade for an average of 8 semesters. The mean sojourn time of states 3, 4 and 5 are respectively 4.58974, 4.06977 and 5.21429, respectively.

Table 5: Estimates of the mean sojourn time in each state

States	estimates	SE	L	U
1	1.83333	0.74846	0.82365	4.08078
2	7.57143	2.02355	4.48420	12.78412
3	4.58974	0.73495	3.35341	6.28188
4	4.06977	0.62063	3.01830	5.48753
5	5.21429	1.39358	3.08817	8.80416

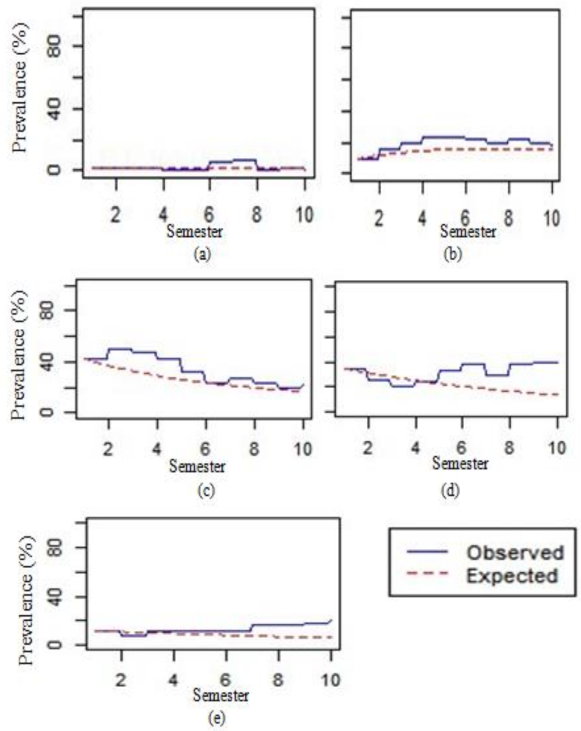

Fig. 2: The prevalence plot for observed and expected number of students in; (a) State 1, (b) State 2, (c) State 3, (d) State 4 and (e) State 5

Prevalence

A rough indication of the goodness of fit of the multi-state model is assessed through the prevalence plot given in Fig. 2. The observed and the expected frequencies of the students at each state are examined. Fig. 2(a) shows that the model properly estimated the number of students in State 1, since the observed and expected frequencies almost coincide at every semester. The model underestimated the number of students in States 2 and 3, though the trend is well estimated. Over-estimation also occurred in States 4 and 5, especially when the number of semesters spent is greater than 5. Generally, the model is very good for monitoring the performances of the students who are in States 3 and below. These are the categories of students who might need some learning support as their CGPA is in third class and below.

Conclusion

A method of monitoring students' academic progression based on a multi-state approach has been presented. This method is used in monitoring the academic progress of students of some undergraduate students who were admitted in session beginning in 2008 and 2009. The class of student considered in the session beginning in 2009 are the students who were expected to graduate in 2013. The states were classified based on class of CGPA. The absorbing state is the graduation or withdrawal state. The student is expected to be monitored for a maximum of 11 periods. In the last period, absorption is expected to have taken place. The first 10 periods are the semesters under which the students undergo their studies and cumulate the grade point average. The results show that it is easier to climb from the bottom part of the class ladder to the middle, while it is quite difficult to climb from the middle to the top. This is an indication that learning support offered to the students in the bottom part of the grade ladder will significantly impact on the progress made during their studies.

References

Cox DR & Miller HD 1965. *The Theory of Stochastic Processes*. London: Chapman and Hall.

Dayıoğlu M & Türüt-Aşık S 2007. Gender differences in academic performance in a Large Public University in Turkey. *Higher Education*, 53: 255-277.

Erdem C, Şentürk İ & Arslan CK 2007. Factors affecting grade point average of University students. *The Empirical Economics Letters* 6(5): 359-368.

Hougaard P 1999. Multi-state models: A review. *Lifetime Data Analysis*, 5: 219-264.

Hougaard P (Ed.) 2000. Probability aspects of multi-state models. *In: Analysis of Multivariate Survival Data. Statistics for Biology and Health*. New York: Springer.

Jackson CH 2014. *Multi-state Modelling with R: The MSM Package*. Cran.r-project.org, Available on: http://cran.r-project.org/web/packages/msm/vignettes/msm-manual.pdf.

Jackson CH, Sharples LD, Thompson SG, Duffy SW & Couto E 2003. Multistate Markov models for disease progression with classification error. *The Statistician,* 52(2): 193-209. doi: 10.1111/1467-9884.00351

Marshall G & Jones RH 1995. Multistate models and diabetic retinopathy. *Statistics in Medicine,* 14: 1975-1983.

Meira-Machado L, deUna-Alvarez J, Cadarso-Suarez C & Andersen PK 2009. Multi-state models for the analysis of time-to-event data. *Statistical Methods in Medical Research,* 18(2): 195-222. doi: 10.1177/0962280208092301

Moler C & van Loan C 2003. Nineteen dubious ways to compute the exponential of a matrix, twenty-five years later. *SIAM Review,* 45(1): 3-49.

Oseni BM & Anjorin IH 2016. Modelling of road traffic accidents: A multi-state Markov Approach. *Sri Lankan J. Appl. Stat.,* 17(2): 135-147.