Should simple Markov processes be taught by mathematics teachers?

Ann-Lee Wang
Institute of Mathematical Sciences
University of Malaya
50603 Kuala Lumpur
Malaysia
wangal@um.edu.my

Song-How Kon
Institute of Mathematical Sciences
University of Malaya
50603 Kuala Lumpur
Malaysia
shkon@um.edu.my

1. Introduction

Simple Markov processes with discrete states in continuous time are taught in a variety of undergraduate courses in different disciplines such as the sciences, engineering, economics, computing, social sciences and medical sciences. As a consequence, circumstances arise where these simple Markov processes are sometimes taught by biologists, engineers or economists. Perhaps in these situations these simple Markov processes are taught with much less mathematics than they would be in a course meant for statistics and mathematics majors. In this paper, some aspects of the teaching of simple Markov processes to statistics and mathematics majors are considered. In these days of universal university education for all, a fair number of students enter university with minimum scholastic attainment. Teaching techniques and content of courses have to be redesigned to cater for students with varying background and level of mathematical abilities.

Although simple Markov processes are usually taught in a second course of lectures on stochastic processes, students still have difficulties in learning these topics. Some students would be taking these lectures with minimum passes in the prerequisite courses. In Section 2, the scope and content to be taught are described. Learning simple Markov processes requires the use of more mathematics and mathematical techniques than is customary in a first course in statistics. The possibilities of mathematics teachers teaching these simple Markov processes and some of the problems involved are considered in Section 3. Section 4 lists some concluding remarks.

2. Scope and content

The simple Markov processes to be taught are Poisson process and birth-death processes. This series of lectures is aimed at statistics and mathematics majors who have taken at least first courses in probability theory, calculus, linear algebra and discrete-state space Markov chain. These lectures will most probably form part of a second course in stochastic processes. These simple Markov processes are to be taught at a much lower level than some standard textbooks, such as Cox and Miller (1965), Karlin and Taylor (1975) and Ross (1996). These lectures are also to be taught at a slightly lower level than textbooks such as Taylor and Karlin (1994) and Ross (2000).

For each process, probabilistic argument is used to derive sets of differential-difference equations. Useful properties are then obtained from these equations by using various mathematical techniques. In the following, differential-difference equation will be refered to as differential equation.

Students are introduced to the use of differential equation in stochastic processes for the first time in Poisson process. For the targeted students, only one definition is used for deriving the differential equation which yield the distribution of the Poisson process. The probabilistic argument leading to the setting up of the set of differential equations is not that difficult for students to understand. This differential equation can be solved recursively. This is a technique which is easy for students to understand. The generating function method of solution for this differential equation can be set as a tutorial question later on once the method has been introduced.

In Cox and Miller (1965), the backward equation of a linear birth process is derived by considering the probability of the transition of the population size from i individuals to j individuals in time t, $P_{ij}(t)$ say. Most standard textbooks do it in this manner too. This method may be a bit difficult for the targeted students. At the first instance, it may be easier for students to understand if it is assumed that initially there is only one individual in the population or colony as is done in Beaumont (1983). Once students are familiar with the technique used, the derivation of the more general equation may be taught. In this way, students learn the techniques in small doses and in increasing complexity.

The derivation of the backward equation is less complicated than the derivation of the forward equation and will be taught first. Assuming there is one individual initially, the backward equation derived can be solved by using the technique of generating function. Partial fraction method is then applied to the equation involving generating function to yield a solution. The expected value and the variance of the number of individuals at time t as well as $P_{1n}(t)$, $n = 1,2,...$, may be obtained from the generating function. Let the time to the *n*-th birth be T_n. The probability distribution of T_n may also be evaluated.

The derivation of the forward equation is next taught. Using generating function one would arrive at a partial differential equation that students are unfamiliar with. The form of the general solution is stated and the solution is obtained by using the boundary conditions. Students may be asked to verify that the general solution as stated satisfies the partial differential equation. When students have understood the derivation of the backward and forward equations for the case when initially there is only one individual, it is then time to show students the derivation of the general backward and forward equations.

Similar techniques are used to derive and then to solve the backward and forward differential equations in linear birth-death processes. Some simple and direct applications are interspersed throughout the course.

3. Mathematics teachers as instructors

An assortment of different mathematical concepts, techniques and probabilistic argument are used in these lectures on simple Markov processes. Although individually these concepts, techniques and argument are not too unfamiliar to students, having to deal with all of them in such a short space of time may be a bit too difficult for some students. Students well-versed in basic manipulative skills may still find the probabilistic argument a bit difficult to understand. An average student finds the combination of probabilistic argument and mathematical techniques difficult to handle.

Since so much mathematics are used in these lectures, the question of whether someone who specialises in functional analysis or some other field in pure mathematics may do a better job at teaching these simple Markov processes naturally arise. What are the impediments and advantages? Certainly such a person being not too familiar with Markov processes would spend more time and more thoughts on how to teach these processes. While trying to write the lecture notes, one would have realised that the application of mathematical techniques in this setting cannot be that straight forward. Quite often, a direct approach to obtain a quantity may be too difficult for students at this level. Direct computation may sometimes be circumvented by looking at the probabilistic nature of the situation. This may not be apparent to a mathematics teacher. On the other hand, someone with a pure mathematics background may think that certain steps or subskills that a statistician may just gross over, are non-trivial and should be emphasised for the benefit of the students.

A mathematics teacher may not be too familiar with different possible applications of Markov processes and so may not be able to make use of these applications in teaching. However, this may not be a drawback. The examples used in the lectures are limited to simple applications of Markov processes. In order to apply Markov processes in some field, one would have to have a good grasp of the subject matter. An ideal teaching and learning situation would be to have the subject matter explained by an expert and then to have a statistician explain the mathematics involved. Unfortunately, there usually is not enough time in an undergraduate course to do this. The person teaching Markov processes have to take on both roles, i.e. being an expert in the subject matter and teaching the mathematics involved. Hence only simple application of Markov processes will be taught because it is not possible to be an expert in a variety of unrelated fields.

Furthermore, knowing a theoretical process and being able to apply them to problems are two very different skills. From the cognitive aspect, Lovett and Greenhouse(2000) stated that:

> *"Knowledge tends to be specific to the context in which it is learned."*

From an instructor's viewpoint, Wang (2001) highlighted students' difficulty in translating word problems into probability statements. From a textbook writer's point of view, Tijms (1986) in the preface to his book remarked that

> *".... many students find it difficult to translate a specific applied probability problem into an appropriate stochastic model. "*

In addition to these perceived difficulties there is usually insufficient time in class to explain real-world random phenomena in probabilistic terms. A real-world random phenomenon that cannot be explained in a few sentences may not be effectively used in these lectures. A long description would mean students would have to take a longer time just to grasp and understand the phenomenon being described. The extra time taken would have been better used in understanding the techniques used. Short qualitative descriptions of possible applications of simple Markov processes would suffice. For example, the death and immigration process may be used to model the number of pedestrians in a given section of foot path (Cox and Miller (1965)). Simple examples, with the derivation of the relevant differential equations, such as general birth process, death and immigraion process, birth and immigration process, linear growth model with immigration, finite birth process and simple M/M/1 queueing system may be taught. Simple birth-death processes with general birth rates, l_i and general death rates, m_i can be given as examples without going through the derivation of the differential equations involved.

When the instructor of the course is a mathematics teacher, there is an inherent danger of over-emphasising mathematical rigour at the expense of the probabilistic nature of the materials. Moreover, there may be an inclination to develop the general theory instead of presenting each of the simple Markov processes in succession to reinforce learning. Perhaps one can consider Moore's (2001) and Cox's (1998) comments. Moore (2001) stated that:

> *"Mathematically-trained teachers often imagine that they can gain efficiency by presenting general principles or structures first, followed by concrete "special cases." This doesn't work. Few people learn from basic principles down to special cases."*

Cox (1998) remarked that:

> *"Few students can appreciate general principles without repeated illustrations."*

Consistent with Moore's and Cox's comments, one could consider the approach proposed by Wang (2001). Wang suggested that the materials should be ″taught in small quantities and at the same time alternating with tutorial problems ″. Poisson process and linear birth-death processes lend themselves easily to this approach. One begins with Poisson process, introducing probabilistic argument and mathematical techniques that will be used repeatedly and in increasingly more complicated ways in linear birth process, linear birth-death process and linear death process.

4. Concluding remarks

This set of lectures will always be difficult for weak students. A fair amount of mathematical skills is required to learn and apply Poisson process and birth-death processes. To facilitate learning, the necessary mathematical techniques should be taught in small doses repeatedly and in increasing complexity. Mathematics teachers willing to put in the necessary effort should be able to teach these simple Markov processes. However, they should adopt an appropriate approach in teaching the topics in line with the probabilistic nature of the course. Where undergraduates are concerned, realistically, one could only teach some simple linear birth-death processes. A thorough knowledge of these processes will be a good foundation to build on.

REFERENCES

Beaumont, G. P. (1983) *Introductory applied probability.* Chichester: Ellis Horwood Limited.

Cox, D. R. (1998) Some remarks on statistical education. *The Statistician*, 47, 211-213.

Cox, D. R. and Miller, H. D. (1965) *The theory of stochastic processes.* London: Chapman and Hall.

Karlin, S. and Taylor, H. M. (1975) *A first course in stochastic processes*. Second Edition. San Diego: Academic Press.

Lovett, M. C. and Greenhouse, J. B. (2000) Applying cognitive theory to statistics. *The American Statistician*, 54, 196-206.

Moore, D. S. (2001) Undergraduate programs and the future of academic statistics. *The American Statistician*, 55, 1-5.

Ross, S. (1996) *Stochastic processes*. Second Edition. New York: John Wiley.

Ross, S. (2000) *Introduction to probability models*. Seventh Edition. San Diego: Harcourt Science and Technology.

Taylor, H. M. and Karlin, S. (1995) *An introduction to stochastic modeling*. Boston: Academic Press.

Tijms, H. C. (1986) Stochastic modeling and analysis: A computational approach. Chichester: John Wiley.

Wang, A. L. (2001) How much can be taught about stochastic processes and to whom? In C Batanero (Ed.), *Training researchers in the use of statistics*, 2001, 73-85. Granada: International Associations for Statistical Education and International Statistical Institute.