Batanero, C. (Ed.), Training Researchers in the Use of Statistics, 73-85.

Printed in Granada, Spain.

ANN-LEE WANG

HOW MUCH CAN BE TAUGHT ABOUT STOCHASTIC PROCESSES AND TO WHOM?

Researchers quite often need to model and analyse real-world random phenomena using stochastic processes. Learning stochastic processes requires a good knowledge of the probability theory, calculus, matrix algebra and a general level of mathematical maturity. However, not all researchers have a good foundation in probability and mathematics. In this paper, we discuss the different approaches to the teaching of a first course in stochastic processes to researchers. Difficulties in the understanding of stochastic processes and the various mathematical techniques used in stochastic processes are discussed. Proposal for the core topics of such a course and ways of teaching them are put forward.

1. INTRODUCTION

The abundance of books on stochastic processes shows the perceived usefulness and applicability of stochastic processes in modelling random phenomena. These books range from the more elementary books such as Bhat (1984) and Taylor and Karlin (1994), intermediate books such as Cox and Miller (1965), Karlin and Taylor (1975) to more specialised books such as Gardiner (1997) and Tijms (1994).

Researchers in science, engineering, computing, business studies and economics quite often need to model real-world situations using stochastic models in order to understand, analyse, and make inferences about real-world random phenomena. Finding a model usually begins with fitting some existing simple stochastic process to the observed data to see if this process is an adequate approximation to the real-world situation. If these simple models are found to be inadequate, ways of extending these models may then be developed.

Learning stochastic processes requires a good knowledge of the probability theory, advanced calculus, matrix algebra and a general level of mathematical maturity. Nowadays, however, less probability theory, calculus, matrix algebra and differential equations are taught in the undergraduate courses. This makes it a little bit difficult to teach stochastic processes to researchers.

Faced with a group of researchers, one has to decide on the amount of mathematical techniques that the students are able to understand in a reasonable time span. The researchers' prior mathematical knowledge puts a limit on the amount and the level of stochastic processes that one can teach. It is necessary to think about what and how to teach a first course in stochastic processes to researchers.

In this paper, we shall first consider the different types of stochastic models and some reasons why researchers use stochastic models. A short account of the different approaches to the teaching of stochastic processes is then given, and the core topics for a first course is put forward. The obstacles and the difficulties inherent in the understanding of stochastic processes are discussed, and some suggestions for the teaching of these topics are proposed. Some comments about the misuse of stochastic models are stated.

2. STOCHASTIC MODELS

A first approximation for a model is usually a deterministic one. Sometimes adding some variation or some stochastic elements to a deterministic model may give a better representation of the random phenomenon under study, so that there is a need to use stochastic models. Stochastic models are constructed with certain assumptions in mind. If the model proved to be badly inadequate, the assumptions are most probably wrong. In case the model seems to fit the data collected, it may provide a better understanding of the random phenomenon.

Stochastic models may be broadly classified into two groups: the purely empirical models which would only give a good representation of the data and models that attempt to incorporate some links between the estimated parameters and the underlying physical process. Examples of a purely empirical model are the models of daily rainfall observations of Stern and Coe (1984), who used non-stationary Markov chains to fit the occurrence of rain and gamma distributions with varying parameters to fit the rainfall amounts. Examples of the second group are the stochastic models for rainfall considered by Rodriguez-Iturbe, Cox and Isham (1987) where the parameters of the models have a physical interpretation. A more detail grouping of stochastic models into four different types is given by Cox (1997).

Stochastic models are used in several fields of research. Some models used in the engineering sciences are models of traffic flow, queueing models, reliability models, spatial and spatial-temporal models. In the computer sciences, the queueing theory is used in performance models to compare the performance of different computer systems. Applications of random walk models are to be found in electric networks (see Doyle and Snell, 1984) and in physics, chemistry and biology (see Weiss, 1983 for a non-technical description). Various diffusion processes models are used as stochastic financial models. Stochastic models are also used to construct epidemic models. More examples of the use of stochastic processes can be found in Gardiner's book (1997). For a short and succinct introduction to stochastic modelling, please refer to Isham (1991). The mathematical techniques and the numerical computation used in stochastic models are not very simple.

In an introductory course, the hope is to teach researchers a small number of stochastic models effectively to enable them to start thinking about the applications of stochastic processes in their area of research. A researcher could use one of these small numbers of stochastic models as a starting point for modelling the random phenomenon that he/she is investigating. This simple model might be adequate for the researcher's interest, and if it is not, later modification could be made.

These small numbers of stochastic models are the core topics to be taught in an introductory course on stochastic processes directed to researchers in the physical sciences, engineering, operational research and computing science. These researchers have a stronger background in mathematics and probability than researchers in the biological sciences. There also seems to be a low usage of stochastic processes in recent biological and health sciences publications as mentioned by Harraway et al. (2001).

Table 3 of their paper indicates that only about 1.71% of the papers surveyed used some stochastic processes, and the topics from stochastic processes that appear in these papers are transition matrices, random walk, Markov processes, time series, periodogram, cross correlation and random intervention analysis. Spatial analysis merits a classification by itself and about 4.47% of the papers surveyed used this method of analysis. Notwithstanding this evidence, some very highly complex stochastic models could be used in the biological sciences. Such models are outside the scope of this paper and will not be discussed here.

I consider the core topics to be discrete time Markov chains with stationary transition probabilities, the Poisson process and birth and death processes. The type of researchers attending the course will determine what other specialised topics such as renewal theory and queueing theory need to be taught.

The needs of different research areas are different, since some stochastic models are only used in a particular research area. It is not reasonable to require all researchers to learn models that they will not need in their research. In some area, such as physics, different notations and nomenclature are used. These have to be explained to the researchers and perhaps some other adjustments to the lecture notes need to be carried out for certain applications of stochastic processes. Researchers, while learning the three core topics, would have picked up a fair amount of probability reasoning and mathematical techniques that are very useful in the learning and understanding of more specialised topics.

Some textbook writers have different views on what should be taught first. For example, Tijms (1984) and Kao (1997) start off with renewal processes, the Poisson process, proceed to Markov chains and then other more specialised topics. However, understanding renewal theory may require more mathematical skills than the average researcher has. This approach may be suitable for the more mathematical minded researchers. In general, it is better to begin with Markov chains.

Even if one chooses to start off with Markov chains, there are two ways in which the topic is laid out in textbooks. Books such as Cox and Miller (1965) and Bhattacharya and Waymie (1990) begin with random walk. Other books such as Karlin and Taylor (1975), Taylor and Karlin (1994) and Kulkarni (1995) begin with Markov chains and random walks are introduced as examples of Markov chains.

Starting a stochastic process course with random walk has its advantages. Feller (1957) remarked in Chapter III that “exceedingly simple methods may lead to far reaching and important results” in random walk. It is possible to discuss many results that are sometimes counter-intuitive using simple methods. Unfortunately, such elementary methods may not be so easily understood by some students. My experience is that not all students are comfortable with combinatorial methods, and I think that it is better to teach random walks as examples of Markov chains.

3. OBSTACLES IN THE LEARNING OF STOCHASTIC PROCESSES

There appears to be a lack of papers on the teaching of stochastic processes. Writers of textbooks sometimes will suggest what selection of the chapters in their books can be used to make up a course in stochastic processes in the prefaces to their books. A few writers do comment on how students/researchers should learn stochastic processes. Some comments are:

> *"Although the basic concepts of useful stochastic models are at the core simple and intuitive, in fact many students find it difficult to translate a specific applied probability problem into an appropriate stochastic model. The student can only acquire the skills of modelling new situations by a considerable amount of practice in solving problems on his*

own " (Tijms, 1986, p. XI).

"The best way to learn the material in this book is by solving problems given in exercises " (Kulkarni, 1995, p. X).

Therefore, the advice is practice, practice and more practice. The only paper commenting on the teaching of stochastic processes that I have come across is the one by Cox and Davison (1994). In it, they remarked that:

"A genuine understanding of the key concepts and principles of the probability theory is essential for any work in stochastic processes. These concepts are simple, but treacherously so. There is probably no branch of mathematics in which it is so easy to advance plausible sounding arguments that are in fact wrong! " (Cox & Davison, 1994, p. 25).

Not many researchers have "a genuine understanding of the key concepts and principles of probability". All of them have some knowledge of the probability theory, and there is not enough time to teach both probability and stochastic processes in an introductory stochastic processes course to researchers. What can be done is to selectively teach the most relevant topics and their applications in selected stochastic processes. It is necessary to teach with as little mathematics as possible. Each new concept has to be taught in small doses with graphical displays and diagrams where possible. Successive new concepts are built on previously taught concepts. Teaching using such small building blocks progressively may make it easier for the researchers to understand and to learn the required concepts. Researchers need to be reminded every now and then that they have to put in a fair amount of effort into learning too. Some reinforcement in the form of requiring researchers to attempt tutorial problems frequently has to be carried out. In this section, some obstacles in the learning of stochastic processes and some ways of overcoming them will be discussed.

3.1. INTUITION

Researchers' intuitions get in the way of their understanding of stochastic processes. Since some results in stochastic processes are counter-intuitive, these would add to the frustration of researchers in their attempts to understand stochastic processes.

A demonstration, via simulation, of a Bernoulli process at the beginning of the course might serve as a warning to researchers about intuition. The simulation of ten thousand tosses of an unbiased coin can be used as an example. The result of the simulation can be used in various ways. The proportion of heads is most probably very close to 0.5. Moore (1990) on pages 97-98 recounted that G. L. Buffon, Karl Pearson and John Kerrich had actually tossed coins for respectively 4040, 24,000 and 10,000 times. The observed proportions of heads were respectively 0.5069, 0.5005 and 0.5067.

Given these facts, if one then displays the sequence of heads and tails, the researchers may or may not be surprised to see the runs of heads and tails. Another graphical display showing the lead of head over tail (represented on the vertical axis) as the number of tosses increases (represented on the horizontal axis) will perhaps lead the researchers to think about counter-intuitive happenings in stochastic processes. Most researchers would expect to see the points oscillating around the horizontal axis. What they actually see is a graph where for a large proportion of the time, the points are above (or below) the horizontal axis. This seems to run counter to their intuitive feelings of how a random process should behave. The graphical display might just keep them in the straight and narrow path of thinking in terms of probability and stochastic processes.

Readers who may want to read more about intuition and probability and ways to resolve the conflicts and to build on it for undergraduates and school children may read Borovcnik and Peard (1996) and the extensive (ninety-one references) bibliography therein.

3.2. TRANSLATING A WORD PROBLEM INTO PROBABILITY STATEMENTS

Anyone who has any experience in teaching mathematics would realise that students have difficulties in translating a word problem into mathematical statements. The same thing holds true in stochastic processes as has been remarked by Tijms (1986). One way to overcome this is to use examples that are as close to the subject matter of the researchers' interest as possible while teaching. These examples have to be preceded by simple unambiguous examples that will help researchers to understand the various concepts in stochastic processes.

As an illustration of translating a word problem into probability statements in stochastic processes, consider the case of a Markov chain. The definition of the various states of a Markov chain in an example has to be taught carefully and thoroughly. States have to be defined such that they conform to the Markov chain assumptions. Emphasis has to be placed on what the answers that are being sought are. For the same random phenomenon, different state spaces may be defined for different objectives. We can use the coin tossing example to illustrate this.

In the coin tossing example, we can consider at least three different stochastic processes as shown below:

1. The lead of head (H) over tail (T): in other words a simple random walk. The states may be defined as $X_0=0$ and for $n \geq 1, X_n = Z_1 + ... + Z_n$ where the Z_i are independently and identically distributed random variables with equal probability of taking on the values of 1 or –1 depending on whether H or T occurs at the n-th toss.
2. Success run: Suppose we consider the occurrence of H as a success. The states of a Markov chain, W_n, may be defined as the length of a success run at the n-th toss of a coin.
3. Result of two consecutive tosses of a coin: Define a Markov chain where the states X_n, for $n > 1$ are defined according to the (n-1)-th and the n-th tosses.

These examples may be followed by a few simple examples of coloured balls in boxes, i.e. examples involving the withdrawing, replacing and inter-changing of the balls in various manners. An illustration of this type of example is the following:

Example 1 (Karlin & Taylor, 1975). N black balls and N white balls are placed in two urns so that each urn contains N balls. At each step, one ball is selected at random from each urn and the two balls interchange. The state of the system is the number of white balls in the first urn. Determine the transition probability matrix of the Markov chain.

These coin tossing and coloured balls in boxes examples may serve as an introduction to the state space and the time parameter space of a Markov chain. These examples are usually devoid of ambiguity and are on the whole non-emotional, non-political and non-cultural. Researchers would be able to understand the various concepts in stochastic processes with less interference. Examples from the subject matter of the researchers are then used to show how the concept of Markov chain can be used in the area of interest to the researchers.

In contrast to state space, researchers usually do not have much difficulty in defining the time parameter space of a Markov chain.

3.3. TRANSITION PROBABILITY

Most undergraduates find great difficulties in determining the transition probability matrix, ***P***, of a Markov chain. This is in part due to the fact that they have great difficulties in defining the state space and the time parameter space of a Markov chain. This usually arises because of students' failure to fully understand the nature of the random process of a given problem. A Markov chain is a random process that evolves with time. It moves from state to state in accordance with the transition probabilities of the Markov chain. A future state is dependent only on the present state and not on the past history of the Markov chain. Given the state of the process at time n, X_n, students have to be able to write down all the possible states of the process at time (n+1), X_{n+1}. They would then be in a position to evaluate the transition probability $P[X_{n+1}|X_n]$. Without a full understanding of the random process under study, students may not be able to write down all the possible mutually exclusive and exhaustive events for X_{n+1} and so be unable to find the transition probability.

Falk (1986) discussed three issues concerning the learning and understanding of conditional probability. Let us consider these issues in the context of the transition probability of a Markov chain.

Consider two events A and B. The first issue raised by Falk (1986) is that if students perceived that it is reasonable to ask about $P[A|B]$, some of them would hesitate at evaluating $P[B|A]$ as this seems "unnatural" to them. In a Markov chain, however, other than the transition probability $P[X_{n+1}|X_n]$, it is often quite meaningful to consider $P[X_n|X_{n+1}]$.

The second issue raised by Falk (1986) is that in evaluating $P[A|B]$, has the sample space changed with the extra information provided by the event B? For transition probability, a tree diagram or similar diagrams showing the possible progression with time of the Markov chain is helpful in concentrating the minds of researchers. Hopefully they will be able to evaluate $P[X_{n+1}|X_n]$.

The third issue raised by Falk (1986) may be termed as translating a word problem into probability statements. Students have to decide whether the word problem is in the form of $P[A|B]$ or $P[B|A]$. In a Markov chain, this issue may not arise if the students have understood the random process fully. Researchers may also be reminded to concentrate on $P[X_{n+1}|X_n]$.

In a course for researchers, demonstrations of how to obtain transition probabilities can be carried out initially with some coin tossing examples and simple examples of coloured balls in boxes. A review of conditional probability will be helpful. These examples are then to be followed by examples from the researchers' subject area. These examples have to be carefully explained in detail to show how probability and conditional probability are used in practice. Researchers with a better grasp of probability theory can be required to work out the transition probabilities themselves.

3.4. MATHEMATICAL TECHNIQUES

In order to understand and use stochastic processes, researchers need to know a fair amount of mathematical techniques. The mathematical techniques used range from the sum of geometric series, difference equations, recursive relations, generating functions, convergence, matrix operation, and conditional expectation to differential equation.

These techniques need to be either reviewed or taught at length in context as the needs arise. Textbooks usually compiled the mathematical techniques in an appendix or as a group within each chapter. This is good for a smooth uninterrupted flow to the discussion of the various topics. In contrast, to be effective, the mathematical techniques have to be developed in front of the researchers as the teaching of each topic progresses. The effectiveness will be enhanced when the examples used to illustrate the techniques are in the subject matter of the researchers.

Among the various techniques, the technique of first-step (or last-step) analysis is very useful in stochastic processes. In this method the probability of an event is obtained by considering what would happen after the first step has been taken (or the last step has been arrived at). The theorem of total probability is then applied to obtain the desired probability. As an example, consider the derivation of the Chapman-Kolmogorov's equation for $P_{ij}^{(n+1)}$, i.e., the probability of arriving at state *j* in the *(n+1)*th step given that the process was initially at state *i*. Let the state at time *n* be X_n.

$$P_{ij}^{(n+1)} = P[X_{n+1} = j | X_0 = i] = \sum_{k \varepsilon S} P[X_{n+1} = j, X_1 = k | X_0 = i]$$
$$= \sum_{k \varepsilon S} P[X_{n+1} = j | X_0 = i, X_1 = k] \, P[X_1 = k | X_0 = i] = \sum_{k \varepsilon S} P_{ik} P_{kj}^{(n)}$$

In the above, in the second equation we consider all possible transitions that can occur in the first step and then we use the law of total probability. In the third equation, we use the formula for $P[A \cap B | C]$ for any three events *A, B* and *C*. In the fourth equation, we use the definition of transition probability and apply the Markov property to obtain the Chapman-Kolmogorov's equation for $P_{ij}^{(n+1)}$.

Using this first step analysis technique, the probability of ruin in the gambler's ruin problem can be shown to satisfy a set of difference equations. The use of the first-step analysis technique is not as straightforward here as in the case of $P_{ij}^{(n+1)}$.

Consider a particle moving along the line interval [*0, c*]. It moves one step to the right (one positive unit) with probability p and one step to the left with probability q where $p > 0, q > 0$ and $p + q = 1$. Suppose the particle is initially at point a or $X_0 = a$. Let u_i be the probability that the particle reaches *0* before it reaches *c*, with $a \geq 1$ and $c > a$. Let E_i denote the event that the particle reaches *0* before it reaches *c* when it is initially at *i*. Let F_i denote the event that the particle take one negative step when it is at *i.* Then using the first-step analysis and the theorem of total probability we have,

$$P[E_i] = P[E_i \cap F_i] + P[E_i \cap F_i^c]$$

This can then be written as

$$P[E_i] = P[F_i]P[E_i|F_i] + P[F_i^c]P[E_i|F_i^c] = qu_{i-1} + pu_{i+1}$$

Students usually have difficulties in ascertaining $P[E_i|F_i]$ and $P[E_i|F_i^c]$. Let us consider $P[E_i|F_i]$ first. An extra effort seems to be needed by the students to understand that $P[E_i|F_i] = P[E_{i-1}]$. When the particle reaches $(i-1)$, the movement of the particle will start afresh from a new starting point. In this case the new starting point or the new initial point is $(i-1)$. And so we have $P[E_i|F_i] = P[E_{i-1}]$. Similarly, one can obtain $P[E_i|F_i^c] = P[E_{i+1}]$.

The first-step analysis technique may be generalised and used to prove the Chapman-Kolmogorov equation, $P_{ij}^{(m+n)}$, and then to prove the first entrance decomposition theorem, to establish the mean first passage time and the mean time to absorption. As an illustration, we can prove the first entrance decomposition theorem as follows. Define $T_j = \min[n \geq 1 | X_n = j]$. Then

$$\begin{aligned} P_{ij}^{(n)} &= P[X_n = j|X_0 = i] = \sum_{\nu=1}^{n} P[X_n = j, T_j = \nu|X_0 = i] \\ &= \sum_{\nu=1}^{n} P[T_j = \nu|X_0 = i]P[X_n = j|X_0 = i, T_j = \nu] \\ &= \sum_{\nu=1}^{n} f_{ij}^{(\nu)} P[X_n = j|T_j = \nu] = \sum_{\nu=1}^{n} f_{ij}^{(\nu)} P_{jj}^{(n-\nu)} \end{aligned}$$

where $f_{ij}^{(\nu)} = P[T_j = \nu|X_0 = i]$, and is called the first passage time.

Before considering the first passage time, the time to absorption, limiting distribution and stationary distribution, evaluation of the powers of $\boldsymbol{P}$ and the classification of states have to be taught. Powers of $\boldsymbol{P}$ can be obtained by several methods. A review of these methods may be considered. Some $P_{ij}^{(n)}$ may be obtained by straightforward probability argument. For example, $P_{00}^{(n)}$ of an unrestricted simple random walk may be obtained in this manner. Tree diagrams are also very useful in the determination of certain $P_{ij}^{(n)}$.

The partitioning of the state space can be carried out by using a transition graph. Drawing a transition graph poses no difficulty to the students. The classification of the state space into equivalent classes and the identification of transient and recurrent states also do not seem too difficult. Of course it is still possible to set problems that will tie some students into knots while they try to draw a transition graph and partition the state space.

What appear to be difficult for the students are the understanding of first passage times and the evaluation of first passage probabilities. The stationary distribution of a finite Markov chain may be obtained by solving a set of equations. Researchers most probably know enough matrix calculus to be able to find the stationary distribution. Sometimes the set of equations can be readily solved as the solution involved only a geometric sum. Matrix algebra is also needed to obtain the probability of absorption and the mean time to absorption.

Computer packages for matrix operation may be used if they are available and researchers are familiar with them or are willing to learn to use them. Some textbooks provide computer program codes (for example Kao, 1997, who uses MATLAB) or offer for sale software packages (for example Tijms, 1994) for analysing stochastic models. One software package for generating large Markov chain models and analysing the models is MARCA obtainable from North Carolina State University.

3.5. BIRTH AND DEATH PROCESSES

The Poisson process is a good and simple introduction to continuous time discrete state space Markov chains in terms of concepts and the mathematical techniques used. In deriving the probability that n events occur in the time interval (0, t), the researchers are introduced to the use of the differential equation in stochastic processes. The technique of generating function may also be introduced here to obtain the required probabilities. As exponential distribution features prominently in continuous time discrete state space Markov chain, a thorough review of its properties and the relationship with the Poisson distribution has to be carried out. Generalisations of the Poisson process can also be taught.

In a demonstration of the derivation of the relevant probabilities in the birth and death process, it may be better to start off with just one individual in the process. The relevant differential equation is then written out. This approach gives the researchers an intuitive feel for the mathematical techniques used. As most researchers are familiar with differential equations, it will not be too difficult for them to understand the derivation. A generating function approach is then used to solve the differential equation. A variation of the technique of first-step (last-step) analysis used here will yield the backward (forward) equations.

One would begin the discussion using a pure birth process with initially only one organism in the colony. A graphical display like the one in Cox and Miller (1965, p. 156) is a good visual starting point. This can then be followed by the demonstration of the derivation of $P_{1n}(t)$, i.e., the probability that there are $n \geq 1$ organisms present at time t, when initially one organism was present. The derivation of the relevant probabilities for a birth and death process will then be easier for the researchers to accept and understand. For example, the starting point of the derivation of the relevant probabilities of a birth and death process may look like this:

$$\begin{aligned} P_{1n}(t) &= [\lambda\delta t + o(\delta t)]P_{2n}(t-\delta t) \\ &+ [1-\lambda\delta t - \mu\delta t + o(\delta t)]P_{1n}(t-\delta t) + o(\delta t) \end{aligned}$$

which gives

$$\frac{dP_{1n}(t)}{dt} = -(\lambda+\mu)P_{1n}(t) + \lambda P_{2n}(t), \qquad n \geq 1.$$

Using generating function one might arrive at

$$\frac{\delta G}{\delta t} = (\lambda G - \mu)(G-1)$$

$$\text{where } G(\theta,t) = \sum_{n=0}^{\infty} \theta^n P_{1n}(t), \quad |\theta| < 1.$$

And the solution for $\mu > \lambda$ is

$$G(\theta,t) = \frac{\mu(1-\theta) - (\mu - \lambda\theta)\exp[(\mu-\lambda)t]}{\lambda(1-\theta) - (\mu - \lambda\theta)\exp[(\mu-\lambda)t]}$$

These equations look intimidating, but would not be quite so if one has understood the derivation of a Poisson process and a pure birth process. In the derivation of the Poisson process, it would have been shown that for $n \geq 1$,

$$P_n(t+h) = P_n(t)P_0(h) + P_{n-1}(t)P_1(h) + \sum_{i=2}^{\infty} P_{n-i}(t)P_i(h)$$

where $P_n(t) = P[X(t) = n]$. From the above equation, one would obtain the following:

$$\frac{dP_n(t)}{dt} = -\lambda P_n(t) + \lambda P_{n-1}(t) .$$

The solution is

$$P_n(t) = \frac{e^{-\lambda t}(\lambda t)^n}{n!}, \quad n = 0,1,2,...$$

As for the pure birth process, one may start with

$$P_{1n}(t) = P[X(t) = n | X(0) = 1] = [1 - \lambda\delta t + o(\delta t)]P_{1n}(t - \delta t) + \lambda\delta t P_{2n}(t - \delta t) + o(\delta t) .$$

For $n \geq 2$, $\frac{dP_{1n}(t)}{dt} = -\lambda P_{1n}(t) + \lambda P_{2n}(t)$. Let $G_1(z,t) = \sum_{n=0}^{\infty} z^n P_{1n}(t)$, $|z| < 1$.

The differential equation can then be written as

$$\frac{\delta G_1}{\delta t}\left\{\frac{1}{G_1(G_1 - 1)}\right\} = \lambda$$

The solution is $G_1(z,t) = \frac{ze^{-\lambda t}}{1 - z(1 - e^{-\lambda t})}$, which gives $P_{1n}(t) = e^{-\lambda t}(1 - e^{-\lambda t})^{n-1}$, n=1,2,3,…

Here the introduction of differential equations and their solutions are carried out in stages and in increasing complexity. This slow introduction will help to ease the way for the researchers to use these techniques in stochastic processes. It is more effective to start teaching a topic from the beginning.

Various other topics in birth and death processes may be explored depending on the interest of the researchers. A discussion of an embedded process in the birth and death process may be carried out. This embedded process turns out to be a simple random walk with one absorbing state. Other results may then be introduced without proof. The researchers will just have to take many things on faith.

4. SOME WORDS OF CAUTION

Some researchers may become over zealous in the application of stochastic processes. They may assume that the model they have set up explained all there is to be explained about the random phenomenon under study. This is usually not the case. Klemes (1994) had written forcefully, in colourful language and with examples on this point for the case of models in hydrology. Klemes warned against attempts "to create knowledge about nature by misguided manipulation of mathematical formulae". His other comments and the comments he quoted in his paper (forty-three references) were for the main part concerned with the building of hydrologic models. These comments are, however, also relevant when one tries to construct stochastic models.

If a deterministic model is adequate, there is no necessity to consider a stochastic model.

5. IMPLICATIONS FOR RESEARCHERS IN STATISTICAL EDUCATION

More research needs to be carried out on how to teach stochastic processes to researchers. One area is the writing up of suitable illustrative examples in the various subject areas. Cox and Davison (1994) gave some suggestions as to what types of examples are suitable for some fields of engineering sciences. The actual examples still need to be written up. One way to obtain these examples is to look at information in as many textbooks as possible. Another way is to seek help from experienced researchers. It will be nice if it is possible to come up with materials that are like Saville's (2001) workshops for agricultural researchers. In the mean time, we have to make do with textbooks and the researchers who want to learn stochastics processes as our guides.

Another area is the research into the teaching of mathematical techniques in stochastic processes. Some researchers in mathematical education might have done some work on some of these techniques. Perhaps there are things that we can learn from them. At the present time, more graphical displays, line graphs, diagrams and simpler examples, such as those covered in this paper, can be used to illustrate the mathematical techniques.

6. IMPLICATIONS FOR THE TRAINING OF RESEARCHERS

Researchers have to be taught with as little mathematics as possible. Where possible some computer simulation of the stochastic processes under study may be demonstrated to the researchers. Basic concepts and mathematical techniques have to be carefully explained with, where possible, examples drawn from the researchers' field of interest. Software packages, if available, may be used.

7. CONCLUSION

Stochastic processes are new topics for most researchers. Most researchers begin the study of stochastic processes with a vague understanding of the essential concepts and principles of probability and usually with a lack of mathematical skills. The task of teaching them stochastic processes is not an easy one. The material for the course has to be taught in small quantities and at the same time alternate with tutorial problems.

Graphs and diagrams need to be used to help researchers to understand and learn the concepts. Using these graphs and diagrams in tutorial problems will help them to better understand the tutorial problems and to solve them too. Researchers need to be reminded to think and to practise, practise and do more practice in solving problems on his/her own.

Mathematics features prominently in stochastic processes. One cannot avoid using mathematics. What can be done is to use as little mathematics in teaching as possible.

New concepts may be introduced by using coin tossing and coloured balls in boxes examples. These examples are mostly devoid of ambiguity. When the researchers are comfortable with the new concepts then examples from the researchers' area of interest can be used to illustrate the applications of the new concepts.

The concepts and techniques learned in Markov chain, Poisson process and birth and death processes will enable the researchers to understand other specialised stochastic models with more ease.

APPENDIX

Websites and a paper related to software for stochastic processes:

1. Tijms, H. C. (1994) at Vrije Universiteit, Department of Econometrics, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands. Package name: STOCHAST.
2. Website: http://www.econ.vu.nl/medewerkers/tijms/default.htm. Reviewed in OR/MS Today, February 1993, 60-62.
3. MARCA: http://www.csc.ncsu.edu/faculty/WStewart/MARCA/marca.html
4. Goel, P. K., Peruggia, M, & An, B. (1997) Computer-aided teaching of probabilistic modeling for biological phenomena, *The American Statisticians*, 51, 164-169.

REFERENCES

Bhat, U. N. (1984). *Elements of applied stochastic processes*, Second Edition. New York: Wiley.

Bhattacharya, R. N., & Waymire, E. C. (1990*). Stochastic processes with applications.* New York: Wiley.

Borovcnik, M., & Peard, R. (1996). Probability. In A. J. Bishop, K. Clements, C. Keitel, J. Kilpatrick, & C. Laborde (Eds.), *International handbook of mathematics education* (Vol. 1, pp. 239-287). Dordrecht: Kluwer.

Cox, D. R. (1997). The current position of statistics: A personal view (with discussion). *International Statistical Review*, 65, 261-290.

Cox, D. R., & Davison, A. C. (1994). Some comments on the teaching of stochastic processes to engineers. *International Journal of Continuing Engineering Education*, 4, 24-30.

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

Doyle, P. G., & Snell, J. L. (1984) *Random walks and electric networks*, Mathematical Association of America. http://front.math.ucdavis.edu/math.PR/0001057

Falk, R. (1986) Conditional probabilities: Insights and difficulties. In R. Davidson, & J. Swift (Eds), *Proceedings of the Second International Conference on Teaching Statistics* (pp. 292-297). Victoria, Australia: International Association for Statistical Education.

Feller, W. (1957) *An introduction to probability theory and its applications*. Volume I. New York: Wiley.

Gardiner, C. W. (1997). *Handbook of stochastic methods for physics, chemistry and the natural science.* Berlín: Springer-Verlag.

Harraway, J., Manly, B., Sutherland, H., & McRae, A. (2001) Meeting the statistical needs of researchers in the biological and health sciences. In C. Batanero (Ed.), *Training researchers in the use of statistics* (pp. 177-195). Granada: International Association for Statistical Education and International Statistical Institute.

Isham, V. (1991). Modelling stochastic phenomena. In D. V. Hinkley, N. Reid, & E. J. Snell (Eds.), *Statistical theory and modelling* (pp. 177-203). London: Chapman.

Kao, E. P. C. (1997). *An introduction to stochastic processes*. Chicago: Duxbury Press.

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

Klemes, V. (1994) Statistics and probability: Wrong remedies for a confused hydrologic modeller, in V. Barnett, & K. F. Turkman (Eds.) *Statistics for the Environment 2: Water Related Issues* (pp. 345-366). Chichester: Wiley.

Kulkarni, V. G. (1995) *Modeling and analysis of stochastic systems*. London: Chapman.

Moore, D. (1990) Uncertainty. In L. Steen (Ed.), *On the shoulders of giants: New approach to numeracy*, (pp. 95-137). Washington: National Research Council,

Rodriguez-Iturbe, I., Cox, D. R., & Isham, V. (1987) Some models for rainfall based on stochastic point processes. *Proceedings of the Royal Society of London*, Series A, 410, 269-288.

Saville, D. J. (2001) A hands-on, interactive method of teaching statistics to agricultural researchers. In C. Batanero (Ed.), *Training researchers in the use of statistics* (pp. 197-213). Granada: International Association for Statistical Education and International Statistical Institute.

Stern, R. D., & Coe, R. (1984) A model fitting analysis of daily rainfall data. *Journal of the Royal Statistical Society,* Series A, 147, 1-34.

Taylor, H. M., & Karlin, S. (1994) *An introduction to stochastic modeling*. Boston: Academic Press.

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

Tijms, H. C. (1994) *Stochastic models: An algorithmic approach*. Chichester: Wiley.

Weiss, G. H. (1983) Random walks and their applications, *American Scientist*, 71, 65-71.

Ann-Lee Wang
University of Malaya,
Institute of Mathematical Sciences,
50603 Kuala Lumpur Malaysia.
E-mail: j2wang@ns1.umcsd.um.edu.my