fied the changes on a technical basis, and have, in turn, suggested instructional problems and problems of curriculum.

5. Mr. Smith and Mr. Trussell demonstrated the drawing of harmonic curves by the use of two simple machines which were equipped with cones instead of gears. These allowed the continuous passage from a given curve to others involving different frequencies. A small mirror was so connected with moving parts of the second machine that a spot of light was caused to trace a curve with such great speed that a line of light in the form of the curve appeared on the screen. Stereoscopic effects were produced by use of lights of two different colors from suitably oriented directions. The development of cylindrical curves, whose projections are Lissajou curves, was demonstrated by using transparent celluloid sheets.

6, 7. Professor Smith demonstrated, by means of oscillograms, the Lissajou and epitrochoidal curves which have been found to be of great value in certain ultrafrequency work used by the Signal Corps. All of the curves were constructed by Professor Gaba, using the University of Nebraska curve tracing machine.

LULU L. RUNGE, *Secretary*

WHAT IS A STOCHASTIC PROCESS?

J. L. DOOB, University of Illinois

1. Introduction. A stochastic process is simply a probability process; that is, any process in nature whose evolution we can analyze successfully in terms of probability. We shall not attempt an exhaustive description. On the empirical side, a discussion of the nature of probability would take us too far afield (and might sidetrack us into philosophy), and on the mathematical side the definitions would require too much high-powered mathematics. We shall limit ourselves to a description of a stochastic process in simple terms, followed by a discussion of a few important particular types. A stochastic process was described above as an empirical entity. On the mathematical side, certain concepts have characteristics in close correspondence with those of these processes, and will be called mathematical stochastic processes. The point is that although at an elementary level probability courses frequently deal with urns, dice and events, a strictly mathematical treatment is possible, with no immediate empirical flavor.

There is a popular prejudice that probability is the subject which deals with gambling games, and perhaps with life insurance and statistics, but is otherwise useless. Moreover its principal method is the calculation of permutations and combinations, than which nothing could be more boring. This calculation looms large in elementary courses, however, only because calculations need little mathematical preparation, either for student or teacher. It is considered easier to perform long calculations than to develop a general point of view, bolstered by theory. This may be good pedagogy, up to a certain point, but gives students a false picture of the subject. The basic reason for the applicability of permutation-combination theory to the study of probability will be noted below. It will be seen, however, that considerably more than the counting of favorable and unfavorable events by such methods is necessary to deal with the problems to be discussed.

2. Concept of a chance variable. Perhaps the basic difference between the older mathematical probability and that of the last fifteen years lies in the stress now put on the concept of a chance variable, and in the development of the whole subject in terms of that concept. In this spirit, we shall develop the idea of a stochastic process from that of a chance variable.

A chance variable, sometimes called a stochastic variable, or a random variable, or (especially by statisticians) a variate, is (non-mathematically speaking) the numerical result of an experiment to which probability analysis is to be applied. There are various possible results, with varying probabilities assigned to them. Thus let N_t be the number of telephone subscribers who will initiate a call tomorrow in Chicago between 4 P.M. and t minutes after 4. Then, fixing t, the practical determination of the necessary number of central telephone facilities needed to prevent more than a given percentage of lost calls is based on assumptions about N_t. For each t, certain probabilities suggested by theoretical considerations are assigned to the various possible N_t values $0, 1, \cdots$. The character of the chance variable N_t depends on t. For example, the more probable values of N_t increase as t does. In many cases the chance variable y under consideration can take on any one of a continuous set of values (for example let y be the value of t when the first subscriber to initiate a call after 4 does so). In such cases individual values of y may each have 0 probability, and intervals of y-values are assigned probabilities. The probability characteristics of any chance variable y are determined by the specification of every probability of the type. Prob. $\{y<k\}$, where k is any number. Thence the probability that y satisfy other types of condition, for example Prob. $\{a \leqq y \leqq b\}$, is determined. The function

$$F(k) = \text{Prob. } \{y < k\}$$

is called the distribution function of y.

Now suppose $y_1, \cdots, y_n$ are n chance variables. To discuss them together, not only the probability relations of each y_j must be given (that is, its distribution function), but we need also the combined distribution: If $k_1, \cdots, k_n$ are any numbers, the probability $F(k_1, \cdots, k_n)$ that $y_j < k_j$ simultaneously for all j is supposed known. All the probability relations of $y_1, \cdots, y_n$ are determined by the knowledge of the "distribution function" F.

Any set of values of n chance variables $y_1, \cdots, y_n$ determines a point P in n-dimensional space. A condition imposed on the y_j can be interpreted as a condition that P lie in some portion of this space. If the probability distribution of $y_1, \cdots, y_n$ is determined by setting the probability that P lie in a given region R proportional to the integral of e raised to a second degree polynomial, integrated over R, then $y_1, \cdots, y_n$ are said to be normally distributed, or to have a Gaussian (n-variate) distribution. The normal distributions are the most common distributions of theoretical statistics because they really do arise frequently in practice, and also because their common statistical parameters are easy to compute.

3. Concept of a stochastic process. Having finished all these preliminaries, we finally come to the concept of a stochastic process. As we have said, a stochastic process in practice is any process whose evolution we find it possible to follow and predict in terms of probability. Now the evolution of such a process is measured by a number (or can be expressed in terms of such measurements) registered in some way, perhaps continuously (at each instant of time), perhaps at discrete times (say every hour), depending on the nature of the process and our analysis of it. Each measurement is by hypothesis a chance variable. From this point of view, a stochastic process is,—and this is the usual definition,—merely a family of chance variables $\{y_t\}$ where t represents the time. Here t may run through all numbers, or only through the integers, or more generally through any set of numbers. If the process consists of a recurring experimental procedure, it is frequently convenient to think of the y_t with t (an integer) ranging to ∞ in both directions even though physically we cannot make (and cannot have made) these infinitely many measurements.

Throughout this exposition, we have not introduced any pure mathematical analysis. For example our "definition" of a chance variable was merely a description of a certain type of experimental procedure, and not a very explicit one at that. This is not the place to go into the proper mathematical definitions. We can only say that the mathematical chance variable is a certain type of function (a measurable function), so that a mathematical stochastic process is a family of (measurable) functions. The analytical study of probabilities is the study of measurable functions. Some of the simple examples we shall discuss below are far simpler to describe physically than to analyze mathematically. What makes an elementary treatment of probability possible is that in many problems the chance variables concerned can only assume a finite number of values, and the time parameter t is allowed to run through only finitely many values. Because of these two characteristics, such problems can always be solved by combinatorial methods. Any finite repetition of a gambling game has these simple properties; and this is one reason for the central place of gambling games in elementary treatments of probability. In our discussion here we are considering the stochastic processes of nature, and carefully evading a proper mathematical analysis (which incidentally dates back less than ten years).

4. Examples. We conclude with several examples of stochastic processes. First consider repeated coin tossing. Here the family of chance variables is $y_1, y_2, \cdots$, with $y_j=1$ (heads) or $y_j=0$ (tails), with respective probabilities

p and $1-p$. The chance variable y_j corresponds to the jth toss. If the coin is evenly balanced, $p=1/2$. It is supposed as usual that the probability of a set of n tosses containing m heads and $n-m$ tails in some given order is $p^m(1-p)^{n-m}$. It is convenient in the mathematical analysis, as we have remarked above, to think of the tosses as continuing in both directions in time, so that we have $\cdots, y_0, y_1, \cdots$. Physically this is impossible, conceptually it is easy to swallow (but even if one gags, the results are easy to translate in terms of finitely many y_j), and the mathematical basis is not very difficult. The translation just noted is of course essential if the theorems are to make any empirical sense at all. A more important interpretation of the same mathematical setup would be the succession of male and female births in a community, or simple examples of Mendel's laws of heredity.

Of this stochastic process we note first that there is no change in the probability relations with time; that is, the probability relations of any $y_m, \cdots, y_n$ are the same as those of $y_{m+1}, \cdots, y_{n+1}$. Such a process is called stationary, or temporally homogeneous. Now what are proper questions that can be asked of such a process? There are various well known ones. For example, the number of heads in n throws divided by n approaches p, the probability of heads in a single throw: formally, if according to the usual rules the probability that

$$\lim_{n\to\infty} \frac{y_1 + \cdots + y_n}{n} = p$$

is calculated, the answer is 1. Although the tosses are independent, there is thus still an average character, a fact which has caused a considerable amount of head shaking, and induced peculiar remarks in well known books. Knowing that $s_n/n \to p$ $(s_n = y_1 + \cdots + y_n)$ one might ask how quickly $\sigma_n = s_n/n - p$ goes to 0, or whether the probability distribution of the chance variables $\sigma_1, \sigma_2, \cdots$ has some limiting form, as n increases. These questions have been answered by the law of iterated logarithm, which states that $\lim_{n\to\infty} \sup \sqrt{n}\,\sigma_n/\sqrt{2pq \log\log n} = 1$ with probability 1, and the central limit theorem, which states that $\sqrt{n}\,\sigma_n$ has more and more nearly a normal distribution as $n \to \infty$.

In the above example, the y_j were mutually independent. The next most simple type of connection is that of a Markoff process. Consider the chance variable $s_1, s_2, \cdots$ in the previous example. Evidently if $n > m$, knowing $s_1, \cdots, s_m$; that is, knowing the number of heads obtained in each toss up through the mth, considerably restricts s_n, the number of heads obtained in n tosses. For example, $s_n \geqq s_m$. It is easy to compute the probability of the various values of s_n if $s_1, \cdots, s_m$ have preassigned values. The answers will, of course, depend on these preassigned values. Now actually it is evident, since $s_n = s_m + y_{m+1} + \cdots + y_n$, where $y_{m+1}, \cdots, y_n$ are independent of $s_1, \cdots, s_m$, that the assigned value of s_m is all that is relevant to the computation. The conditional probabilities depend only on the value assigned to s_m, not on those assigned to $s_1, \cdots, s_{m-1}$. Such a process is called a Markoff process; formally the hypothesis is that whenever $t_1 < t_2$, the probability distribution of y_{t_2} calculated under the assumption that y_t has been assigned values for all $t \leqq t_1$ depends only on the value assigned to y_{t_1}.

Let y_t be the size of a population under statistical analysis. Whenever the population's change depends on the state of the population at a given moment, but, this being known, is independent of how the present state was attained, the y_t determine a Markoff process. It is then evident that Markoff processes will have general application in statistical studies of the population trends of a species and in epidemiology (if the number of infected individuals depends only on those already infected). Important questions to be asked here are: what is the asymptotic character of y_t for large t, is there a limiting value, any kind of periodicity, etc.?

Let y_t be a chance variable depending on the time t, which we assume runs through all values. Suppose that the increments of y_t: $(y_{t+h} - y_t)$ over non-overlapping time intervals are independent, that is if $t_1 < \cdots < t_n$ we suppose that $y_{t_2} - y_{t_1}, \cdots, y_{t_n} - y_{t_{n-1}}$ are mutually independent. The stochastic process is then called a process with independent increments, or a differential process. These arise in many connections. For example let $y_t = N_t$ where N_t was defined above in the telephone example. Then the y_t determine a differential process: the number of subscribers initating a phone call in any time interval can under ordinary circumstances be supposed independent of the calls offered before this time interval (or later). Or, to take an example from physics, let y_t be the number of radioactive disintegrations of a given substance by time t. Again we have a differential process. Or let y_t be the amount of money to be paid out by an insurance company to its claimants between time 0 and time t. Again the process can frequently be assumed a differential process.

In the examples of differential processes considered above, it is clear that y_t, considered as a function of t has as graph a set of horizontal lines, no matter how the process turns out. The proof that in general the y_t of a differential stochastic process are continuous in t except for non-oscillatory discontinuities (jumps)* is quite complicated, and is less than ten years old. Aside from simple changes of scale, there is only one differential process in which the functions y_t are actually continuous in t (with probability 1)—that in which the increments $y_{t+h} - y_t$ have normal distributions. The best known application of this process is to the haphazard Brownian movements of small particles immersed in a liquid. About 35 years ago it was shown that if y_t is a coördinate value of such a particle at time t, the movement can be analyzed fairly accurately by considering the y_t as the determining family of chance variables of a differential stochastic process whose increments have normal distributions. Thus the probability analysis predicts continuous curves as the particle trajectories. That is certainly desirable, but it was noted at once that this same probability analysis predicts infinite

* The meaning of this statement is the following. The probability can be computed that y_t as a function of t will be continuous at a point, everywhere continuous, etc. The probability that the y_t of a differential process will have any non-oscillatory discontinuity is 0 (if very minor restrictions are imposed on the process to eliminate degenerate cases).

velocities for the particles, and that the length of a trajectory between any two of its points is infinite. This may seem remarkable physically, but man can predict anything after it happens, and in fact physicists have asserted that the particle movements as observed seem to have these peculiar properties.

We go finally to a simple but very important type of stationary stochastic process. Suppose that the distribution of y_t is normal, and even that for any $t_1, \cdots, t_n$ $y_{t_1}, \cdots, y_{t_n}$ have an n-variate normal distribution (for all n). We shall call such a process a stationary normal process. It is easy to see that apart from a scaling factor, and a centering constant, such a process is completely determined by a single function $\rho(h)$, the correlation coefficient of the two chance variables y_t, y_{t+h}, which measures the connection between these chance variables. (This is independent of t since the process is stationary.) Necessary and sufficient conditions are known that a function be such a correlation function, but many questions are still unanswered about these processes, and many more will remain unformulated until the need arises. The simplest case is the degenerate case $\rho(h) \equiv 1$, when y_t is independent of t: the process never changes from its initial state, and the theory of probability is quite superfluous. At the other extreme is the case $\rho(h) = 0$ for all $t \neq 0$, when y_s, y_t are independent of each other if $s \neq t$. The first non-trivial class of stationary normal stochastic processes to investigate, is certainly the subclass of Markoff processes, of which the above two cases are degenerate special cases. For a Markoff process, aside from these two cases, the connection between y_t and y_{t+h} goes down exponentially, as $h \to \infty : \rho(h) = e^{-c|h|}$ (c a positive constant). In work on the Brownian movement dating back about nine years, and giving a better approximation than the earlier work, it was found that the particles had finite velocities after all (but infinite accelerations) and in fact that each component of the velocity of a particle at time t can be considered as a chance variable y_t, where the y_t determine one of the Markoff processes just described.

In this case of a Markoff process, with $\rho(h) = e^{-c|h|}$, it is known that the y_t are continuous functions of t, with probability 1, but it is not known generally which correlation functions $\rho(h)$ determine processes with continuous y_t.

Another example of a stationary normal process arises in electricity. The spontaneous thermal movements of the electrons in any wire cause current fluctuations in all electric circuits. If y_t is the current in a given wire at time t, y_t determines a stationary normal stochastic process. The correlation function depends on the particular circuit, and on the resistances, capacities, and inductances. This phenomenon is important in radio since the current fluctuations cause receiver noises which cannot be entirely eliminated. These current fluctuations are a disturbing influence in making any kind of delicate electrical measurements. A somewhat similar phenomenon occurs in radio tubes, known as the shot effect. The study (not yet completed) of these electrical disturbances is thus the study of certain types of stochastic processes.