PRESENT STATE AND FUTURE PROSPECTS OF
STOCHASTIC PROCESS THEORY

J. L. Doob

The theory of stochastic processes has developed sufficiently in the past
two decades so that one can now properly give a survey of its present state
and hazard some guesses as to its future development. It is clear that no
mathematician can reliably predict what the mathematics of the next twenty
years will be, even in his own field. In fact, if he could know what it would
be, he would negate his own foreknowledge by developing the new theories
himself, long before the twenty years were up! Thus the guesses to be given
below are simply the guesses of someone who is now interested in stochastic
processes, and is considering the areas in which he would like to obtain new
results.

Probability theory advances in two ways. On the one hand, probability
problems lead to problems in other fields, in differential and integral equations,
for example, which can be formulated and solved with little or no knowledge
of the probability background. On the other hand, there are the peculiarly
probabilistic problems, say on the convergence of mutually independent
random variables, or on the continuity of sample functions of stochastic
processes, which must be attacked by the methods peculiar to probability.
The measure theoretic formalization of probability concepts twenty-one years
ago by Kolmogorov made possible further progress in the latter direction,
progress which had been woefully retarded up to that time. Using this for­
malization, it is now generally accepted that the most useful mathematical
definition of a stochastic process is simply that a stochastic process is a family
of random variables. Although this definition is so general as to seem pointless,
anything less general appears to be insufficient. In most studies, the random
variables of a stochastic process have been numerical-valued, and have been
indexed by a real-valued parameter. This case will be called the standard case
below. Our survey will first cover the standard case, going into its general
definition, and describing the types of standard stochastic processes that have
been studied intensively. An indication of the present state of the theory,
and the possibilities for future development, will also be given. Finally, a few
remarks will be made on new possibilities in the standard case, and on the non­
standard case.
Before proceeding to the body of the discussion, we remark that, in view of the wide definition of a stochastic process, anyone can define a type of stochastic process. In fact, anyone with a rudimentary knowledge of probability can hope in five minutes or so to define a type of stochastic process that has not appeared in the literature. The real problem is to devise some relationship between random variables leading to a property of a family of random variables that will prove interesting and important, and the only properties discovered up to this time are the property of mutual independence, the Markov property, the stationarity property, the martingale property. We have mentioned here only properties specific enough to lead to interesting results, excluding, for example, the property that the random variables of the process have finite second moments. This meager list of four properties indicates that probabilists are sadly in need of some new ideas!

1. Basic concepts involved in discussions of standard processes.

Let \( \{x(t), t \in T\} \) be a standard stochastic process, that is, a family of measurable functions, defined on a (probability) measure space, with linear parameter set \( T \). The random variable \( x(t) \) is a function of the point \( w \) of the specified measure space, with value \( x(t, w) \) at the point \( w \). Fixing \( w \), \( x(t, w) \) defines a function of \( t \), and the functions of \( t \) obtained in this way are called the \textit{sample functions} of the process. The given measure of \( w \) sets determines a measure of sets of sample functions. By definition of random variable, the joint distributions of the finite sets of the random variables of the process are known. These distributions are called the \textit{finite dimensional distributions} of the process. In the past, stochastic processes have been characterized by the properties of these finite-dimensional distributions. For example, if \( T \) is the whole line, the process is called \textit{stationary} if, for every finite parameter set \( t_1, \ldots, t_n \), the joint distribution of the random variables \( x(t_1 + h), \ldots, x(t_n + h) \) does not depend on the number \( h \). It remains to be seen whether other methods of classification will ever prove fruitful. According to a classical measure theorem of Kolmogorov, corresponding to any specification of mutually consistent finite dimensional distributions, there is a stochastic process with the specified distributions. Moreover, associated with this existence theorem, and with the method of classification by finite-dimensional distributions is the fact that the finite dimensional distributions have been the \textit{only} specifications of the process. Thus, any properties of the process have necessarily been obtainable from these joint distributions, either directly using the properties of measure functions, or by the intermediary of some additional principle which need only be consistent with the properties of measure functions. Such additional principles were introduced by the speaker in 1937 and later, in order to make

349
possible the discussion of such properties as boundedness, continuity, and measurability, of the sample functions of a stochastic process with a non-denumerable parameter set.

Thus, with the help of certain additions to the basic measure-theoretic formalization of probability, there is now an adequate formalization of standard stochastic processes, and it does not appear that much more work will be done here except in the direction of alternative approaches. For example, a more algebraic approach is now gaining in popularity, in which events are interpreted not as sets of a measure space but as elements of a Boolean algebra. The sample functions must be defined indirectly, in such an approach, and this is a disadvantage. However such an approach has the advantage that it automatically avoids some of the difficulties involved in studying sample functions, and this is an advantage if one is not interested in the sample functions for their own sake.

2. **Standard stochastic processes with mutually independent random variables.**

Stochastic process theory has been limited for much of its life, in fact during the centuries before this one when the term *stochastic process* was invented, to the study of finite or denumerably infinite sequences of mutually independent random variables. The study of the distributions of sums of such random variables appears inexhaustible, not only from the point of view of general theory but of refinement of known results. One prediction it is safe to make is that this study will continue, and will remain fruitful.

The continuous parameter analogue of sums of mutually independent random variables is a process \( \{x(t), t \in T\} \) of independent increments, with parameter set \( T \) an interval. That is, if \( t_1 < \ldots < t_n \) are parameter points, it is supposed that the random variables

\[
x(t_2) - x(t_1), \ldots, x(t_n) - x(t_{n-1})
\]

are mutually independent. The general study of these processes, initiated by de Finetti in 1929, has now finished its first task. That is, the general character of the sample functions is known, the (infinitely divisible) class of distributions involved has been fully described, and many special properties have been discovered. The Brownian motion process, in which the above differences have normal distributions with zero means and variances a constant multiple of the parameter differences, has been the most studied special case, and the fact that new and important properties of this process are still being discovered suggests that there will be much further general work on these processes, besides the obvious refinement in detail. In fact, as is remarked below, there is even
now no satisfactory treatment of diffusion processes, which depend essentially on Brownian motion processes.

3. **Standard Markov processes.**

The study of Markov processes was initiated by Markov (in a special case) at the beginning of this century. The stochastic process \( \{x(t), t \in T\} \) is a Markov process if, when \( t_1 < \ldots < t_n \) are parameter values, and \( A \) is a linear Borel set, the equality

\[
P\{x(t_n, w) \in A \mid x(t_1), \ldots, x(t_{n-1})\} = P\{x(t_n, w) \in A \mid x(t_{n-1})\}
\]

is true with probability 1. That is, somewhat roughly, the \( x(t) \) process is a Markov process if the conditional probability of a future event, given present and past states, only depends on the present state.

Standard Markov processes have been studied most intensively in two cases, the stationary chain case, in which \( x(t) \) only has integer values, so that the probability of a transition from state \( i \) to state \( j \) in time \( s \) is given by a number which we can write in the form \( p_{ij}(s) \), and the diffusion case, which is defined below. When the parameter set \( T \) is the set of positive integers, the basic properties of Markov chains, essentially the asymptotic properties of \( p_{ij}(t) \) for large \( t \), have been known for years, but, when \( T \) is the interval \((0, \infty)\), new problems arise, for example the continuity properties of the sample functions and regularity properties of the transition probability functions must now be investigated. These problems have been solved in the non-pathological cases, but not even all the obvious fundamental problems in this area have been solved in the general case, in spite of recent important work by Lévy. There remains the systematic study of the non-stationary case, which has been barely begun.

The standard Markov processes of diffusion type can be roughly described as follows. The parameter set is an interval, and there are functions \( m, \sigma \), called *diffusion coefficients*, for which

\[
x(t + h) - x(t) = m[t, x(t)]h + \sigma(t, x(t))\xi_{t, h} + \ldots
\]

Here the remainder is negligible in comparison with the terms displayed, as \( h \to 0 \), and \( \xi_{t, h} \) is a random variable which is normally distributed, independent of the class of random variables \( \{x(s), s \leq t\} \), and has expectation 0, dispersion \( h \). In the simplest non-trivial special case, \( m \) vanished identically, \( \sigma \) is identically a positive constant, and the corresponding process is the Brownian motion process. This process was first studied (unrigorously) by Bachelier at the turn of the century, before either Markov processes or processes with independent increments, to both of which classes this process belongs, had received general definitions.
The class of diffusion processes is not rigorously defined; that is one of the difficulties of this subject. There are almost as many different approaches to these processes as there are investigators, and in fact it is not yet known under what conditions the various approaches lead to the same processes. There appears to be general agreement that the definitions should be phrased in such a way that almost all sample functions of a diffusion process are continuous (under the appropriate conventions necessary for such a statement). Under various regularity conditions, the transition probability functions of a diffusion process satisfy parabolic differential equations involving the diffusion coefficients, first discussed systematically by Kolmogorov in 1931, but the exact role of these equations is not yet clear. Conversely, given a pair of diffusion coefficients, there are various ways of defining the corresponding diffusion process: by solving the corresponding parabolic differential equations to find the transition probabilities (Feller); by semi-group methods, at least in the stationary case (Yosida); by rewriting the above difference equation in the form of an integral equation, replacing $\xi_{t+h}$ by the differential element of a Brownian motion process, and solving for $x(t)$ in terms of this Brownian motion process (Ito). As already remarked, it is not clear when these methods, for a specified pair of diffusion coefficients, yield the same stochastic process, or even when the transition probability functions of the processes obtained in any way other than the first satisfy the parabolic partial differential equations characteristic of diffusion processes. In addition, there are various operations commonly performed on diffusion processes leading to new ones, such as stopping the trajectories when they reach a specified closed set, and it is not yet clear when the transition probability functions of the new processes satisfy these same parabolic equations, with the same diffusion coefficients but different initial conditions. In short, there is obviously much work to be done in this area, and much is being done. If $x(t)$ is an $n$-dimensional vector-valued random variable, with $n \geq 1$, it is to be expected that the corresponding problems may become considerably more difficult. The corresponding diffusion processes are non-standard in that case, of course.


These stochastic processes have already been defined. Since the definition is so general, it is not to be expected that many specific results can be obtained for these processes, and in fact the general theorems applicable to these processes are centered around one theorem, the ergodic theorem (law of large numbers) in its various forms. Further developments must be based on more specialized hypothesis.

The processes called stationary in the wide sense, and characterized by the
condition that $E\{x(t)x(s)\}$ depend only on $t - s$ (in the real case, the conjugate sign is of course unnecessary) have received considerable attention in the past and under various special hypotheses, can be expected to receive considerable further study, but, just as for stationary processes, further specialization is undoubtedly necessary to obtain new results. The study of these processes has been and will continue to be intensified by the interest they have in statistics.

5. **Standard martingales.**

Martingales and semimartingales are relative newcomers. The first general martingale theorems were proved by Lévy about twenty years ago, but these processes were first studied systematically by the speaker, in 1940. A stochastic process \( \{x(t), \: t \in T\} \) is a martingale if, when \( t_1 < \ldots < t_n \) are parameter values, the equality

\[
E\{x(t_n) \mid x(t_1), \ldots, x(t_{n-1})\} = x(t_{n-1})
\]

is true with probability 1. That is, roughly, the conditional expectation of future values, given present and past values, is the present value. For a semimartingale, the equality is replaced by inequality (\( \geq \)). It is likely that the most important general theorems on these processes are now known, that is, the convergence properties of martingale and semimartingale sequences, the continuity properties of sample functions of continuous parameter processes of these types, and the invariance properties of these types under various frequently used transformations. However there remain many fruitful possibilities for research in the applications of martingale theory. For example, the martingale equation above states that a certain value is an integral average. This suggests the possibility of applying martingale theory to the study of the first boundary value (Dirichlet) problem for elliptic, or more generally, for parabolic partial differential equations. The most general such application would be the following. Suppose that \( u \) is a function, defined on some topological space, with the property that the value of \( u \) at a point \( z \) of an open set \( D \) is a weighted average of the values of \( u \) on the boundary of \( D \). Then it is natural to try to find a system of continuous probability trajectories (sample trajectories of a stochastic process whose random variables have values in the topological space on which \( u \) is defined) such that \( u(z) \) is the average value of \( u \) over the boundary of \( D \), average here meaning with the weighting defined by the distribution of the first point in which the probability trajectories from \( z \) meet the boundary of \( D \). This program has been carried out in detail by the speaker for \( u \) the solution of a second order linear differential equation in one dimension, for \( u \) harmonic in \( n \) dimensions, and more generally for \( u \) the solution of the heat equation in \( n + 1 \) dimensions. In each case, if \( z(t) \) is the position
of the trajectory issuing from the point \( z \), at time \( t \), where the trajectory is stopped when it reaches the boundary of \( D \), the \( u[z(t)] \) stochastic process is a martingale. The properties of martingales can then be used to show that functions in the class of \( u \), under mild restrictions, have boundary limits along probability trajectories, these limits being the assigned values if \( u \) is a generalized first boundary value problem solution. The general abstract problem has not yet been studied, however, although a positive solution would have the greatest interest. This problem illustrates the importance, for a class of functions \( u \) defined on a topological space, of finding the stochastic processes \( \{z(t), t \in T\} \), whose random variables have values in the space, for which the stochastic process \( \{u[z(t)], t \in T\} \) is a martingale. The properties of martingales can then be used to derive new properties of the functions in the given class. Somewhat more generally, the familiar subfunctions of differential equation theory, such as the subharmonic functions, lead to semimartingales.


In the previous sections, we have discussed the most important types of standard stochastic processes. An obvious question is: Can one expect that important new types will be discovered? No scientist with any historical sense would give a negative answer to such a question, but progress has certainly been slow in the development of new types. What has been sought is a simple property with interesting implications, and we have already noted that only the independence, Markov, stationarity, and martingale properties (or closely related ones) have satisfied this criterion.

It is possible that the successful study of non-linear relations in other parts of mathematics will suggest new types of stochastic processes. One example of such a possibility is the following. Let \( L \) be a differential operator, and consider the differential equation \( L(f) = A \). If \( f \) is the \( x \) coordinate of a particle subject to a stochastic driving force, this equation can be interpreted as the equation of motion of a particle, when, if the particle is a molecule, the equation is called Langevin’s equation. This differential equation has been studied in detail if \( L \) is a linear differential operator, with constant coefficients, and if \( A \) is the sample function of a stationary process, \( f(t) \) can be found explicitly in terms of \( A \), and the \( f(t) \) process, with appropriate initial conditions, is stationary. There is no corresponding theory if \( L \) is a non-linear differential operator, although such a theory might be very illuminating in many applications, for example in the analysis of the rocking of a ship exposed to wave action. Here again one would expect, under reasonable conditions, to find a stationary solution, obtaining in this way a new class of stationary processes.

An obvious generalization of the standard processes is to suppose that the random variables of the process have values in an \( n \)-dimensional Euclidean space, for \( n > 1 \). Some work has been done in generalizing the results for \( n = 1 \) to this case, but much remains to be done even at this elementary level of generalization. Recently there has been work on processes whose random variables take on values in a Riemannian manifold, and this is obviously just the beginning of such work. There has also been some work on random variables with values in a Banach space, but no significantly new types of stochastic process with such random variables have been discovered as yet.

Finally, one can generalize the standard processes by allowing other than linear parameter sets. An example is the Poisson process in which events are distributed in \( n \)-space instead of in time, the expected number of events in an open set of the space being proportional to the volume of the open set. Here the most natural parameter set is the class of all Lebesgue measurable sets in \( n \) dimensions, and the random variable corresponding to such a Lebesgue measurable set is the number of events in the set. Only the barest beginnings have been made in discussing types of processes with non-linear parameter sets, although it would seem at first that here is the easiest context in which new and interesting types of stochastic processes could be defined.