STOCHASTIC PROCESSES AND RANDOM FIELDS

K. Grill
Institut für Statistik und Wahrscheinlichkeitstheorie, TU Wien, Austria

Keywords: Stochastic process, mean value function, covariance function, correlation theory, Gaussian process, strictly stationary process, wide-sense stationary process, Markov process, independent increment process, martingale, stochastic differential equation, self-similar process, fractal, point process, renewal process, empirical process, invariance principle, random field, Gaussian random field, Gibbs random field, phase transition

Contents

Summary

A stochastic process can be considered either as a family of random variables, indexed by a subset T of the real numbers, the so-called parameter space, or as a random function, that is, a random variable taking values in some function space. Stochastic processes have important applications in many fields of science, including biology, physics, chemistry, and even finance. Basically, they can be viewed as models for the evolution of a (biological, physical, etc.) system in time, where the observed variable undergoes random changes. One can also consider more general parameter spaces, so that the stochastic process becomes a random function of more than one variable. This type of stochastic processes is usually called a random field. Sometimes one of the

coordinates of the parameter vector is interpreted as time, whereas the others are spatial variables; in this case, one speaks of a stochastic space-time model. The reason why random fields get special treatment is that many of the methods of the theory of stochastic processes rely heavily on the natural order of the (one-dimensional) parameter space, for which there is no easy replacement in higher dimensions.

In order to understand a stochastic process, it is important to have a clear description of its dependence structure. There are many sensible ways to do this, and so in the study of stochastic processes, a number of special classes of processes evolved, each with its own particular methods.

1. Introduction

1.1. Basic Notions and Definitions

In 1827, the British botanist Robert Brown observed that pollen grains suspended in water perform some form of rapid zigzag movement. He first attributed this movement to the fact that there was some kind of "live power" in those grains, which, after all, are a live organic substance. Yet, soon it turned out that a suspension of (lifeless) dye particles exhibited the same behavior, so it was obvious that this movement, which became popular under the name "Brownian motion", was caused by some other phenomenon.

In 1905, Einstein developed the first mathematical theory of Brownian motion, which was later awarded the Nobel Prize. His explanation was that the reason for the irregular movement of small particles is the thermal oscillation of the water molecules. As a Brownian particle gets hit by a molecule, it is displaced by a small distance, and the combined effects of a large number of such collisions result in the visible movement of the particle. So, as Brownian motion is caused by the random thermal oscillation of the molecules, it is obvious that the spatial position of the Brownian particle at a given time instance is a random variable, and the path of the particle should be considered as the graph of a random function of time.

Thus, Brownian motion becomes one of the first examples of a random function, or, as the usual name is today, a stochastic process. Actually, one cannot really speak about "the" Brownian motion; in fact, there are a number of different mathematical models, and each of them has its particular merits and flaws. The most prominent among them is of course the one that goes back to Einstein's work, and this is called the Wiener process, owing to the fact that Einstein's model was later refined and extended by Wiener. Very often, the name "Brownian motion" is used as a synonym for the Wiener process, but it is better to make a clear distinction between the physical phenomenon and the particular mathematical model. The formal definition of the Wiener process is given in *Construction of Random Functions and Path Properties.* It has many desirable properties, and it will turn up as an example for many important types of processes.

Since its beginnings in the first half of the last century, the theory of stochastic processes has developed into a large and important field of mathematical research. Its applications range from the study of physical phenomena like Brownian motion and

ferromagnetism over biological and chemical processes to the modeling of the evolution of stock exchange rates.

Now, it is time to give a formal definition of a stochastic process. First, there are two sets: the parameter space T and the state space X. A stochastic process on T with values in X is simply a collection of random variables $(\xi(t), t \in T)$ with $\xi(t) \in X$ for all $t \in T$. Usually, both X and T are subsets of the real line $\mathbf{R}$, the most common choices for T being (maybe infinite) intervals. If the set T is at most countable, then ξ is called a discrete parameter process, otherwise one speaks of a continuous parameter process. A similar distinction can be made for the state space. It is clear that a discrete parameter process reduces to a (finite or infinite) sequence of random variable.

A different way of looking at a stochastic process is to consider it as a randomly chosen function (i.e. instead of picking an individual random $\xi(t)$ value for each value of t, one randomly chooses the whole function ξ, and then gets the individual random variables by evaluating this particular choice of ξ at t). This makes ξ a random element of some function space, and the particular instance of the process ξ that is observed is called a trajectory (or path) of ξ.

For both the parameter space T and the state space X, more general choices are possible. If X is a d-dimensional set, then ξ is called a multivariate process; this situation is quite common, and usually this doesn't create any real difficulties compared to the one-dimensional case. For the parameter space, on the other hand, using higher dimensions results in a significant increase of technical and theoretical problems. Thus the theory of processes with a parameter space that is not one-dimensional has forked off as the theory of "random fields". This will be discussed in a later section, and until then, all processes will be assumed to have a one-dimensional parameter space, in most cases one has $T = [0, \infty)$. In this one-dimensional setting, the parameter is often interpreted as a time variable, and in the sequel, the word "time" is used synonymously with the parameter t.

2. Important Concepts and Methods

2.1. Dependency

In probability theory, one of the most basic concepts is that of a sequence of independent random variables. This would suggest that a simple example of a stochastic process would be one that has all the random variables $\xi(t)$ independent. It turns out, however, that such a process doesn't have very nice properties. In particular, its trajectories are not continuous, not even measurable. Still, such processes are sometimes considered as so called white noise processes, but in order to really make use of those, it is necessary to define them as some kind of generalized functions or distributions, in a fashion similar to the one that is used to define the Dirac delta-"function".

So, a decent stochastic process has values that are not independent, and it is of prime importance to properly define their dependence structure. To this end, one usually uses

the concept of filtration. This is simply an increasing family of sigma-algebras, i.e., a family of sigma-algebras $(F_t, t \in T)$ with $F_s \subset F_t$ for $s \le t$. The sigma-algebra F_t can be interpreted as the set of all event observable up to time t, i.e., the set of all events whose occurrence can be decided based on one's knowledge of the past up to time t. Thus, F_t comprises the knowledge one has gathered about the process up to time t. In many cases, F_t is the sigma-algebra generated by the random variables $(\xi(s), s \le t)$ which means that F_t is the set of all events that can be expressed in terms of these random variables. This need not always be the case, reflecting the fact that, besides monitoring the process ξ up to time t, one might be able to get some additional information (e.g., by observing a second process that runs in parallel).

It should be clear that the value of $\xi(t)$ itself is known at time t; this amounts to saying that $\xi(t)$ is measurable with respect to F_t, and gives rise to the following definition:

A stochastic process ξ is said to be adapted to the filtration F_t if for any $t, \xi(t)$ is F_t measurable.

In many cases, one needs a slightly stronger condition:

A stochastic process ξ is said to be progressively measurable with respect to the filtration F_t if for every t, the mapping $(\omega, s) \mapsto \xi(s, \omega)$ with $s \le t$ and $\omega \in \Omega$ is measurable with respect to the product of F_t and the Borel sets. This is mainly a technical condition, implying in particular that the individual trajectories are Borel-measurable (remember that there is always an underlying probability space $(\Omega, F, \mathbf{P})$, and that a random variable is just an F-measurable function on Ω).

Once the dependency structure of a process is fixed, it is possible (though not always easy) to calculate the finite-dimensional marginal distributions of the process (i.e., the joint distributions of $\xi(t_1), ..., \xi(t_n)$), and by Kolmogorov's existence theorem (see *Construction of Random Functions and Path Properties* for details), the process itself is determined. This way of constructing the process leads one to consider two stochastic processes equivalent if they have the same finite-dimensional marginals, and if this is the case, each of these processes is called a "version" of the other. In the case of an uncountable parameter space, this concept leads to some peculiarities, as the following example shows:

Assume that τ is uniformly distributed on [0,1]. Let $\xi(t) = 0$ for all $t \ge 0$, and $\eta(t) = n$ if $t = \tau / n \, (n = 1, 2, ...)$, and $\eta(t) = 0$ for all other values of t. It is immediate that η and ξ are equivalent processes, but in spite of that, the trajectories of η are unbounded and discontinuous with probability one, whereas those of ξ are even constant. This example serves to show that many important properties of trajectories are not shared by all versions of a process (the mathematical reason behind this is that, e.g., the set of all bounded or continuous functions does not constitute an event, i.e., these sets are not measurable).Thus, saying, for example, that a given process has continuous trajectories,

is not strictly correct, and in the following discussion it will always be meant as a short formulation for the more accurate but awkward statement that there is a version of the process that has the desired property.

2.2. Correlation Theory

For random variables, the expectation and variance are important parameters: the expectation, in some sense, gives the typical size of a realization of the random variable, whereas the variance indicates how much different realizations deviate from this typical value. When one has more than one random variable, then the covariance

$$\mathbf{Cov}(\xi,\eta) = \mathbf{E}(\xi\eta) - \mathbf{E}(\xi)\mathbf{E}(\eta) \tag{1}$$

enters the picture. For example, the variance of the sum $\xi+\eta$ can be expressed by the formula

$$\mathbf{Var}(\xi+\eta) = \mathbf{Var}(\xi) + \mathbf{Var}(\eta) + 2\mathbf{Cov}(\xi,\eta). \tag{2}$$

One can base a fairly large part of the theory of stochastic processes on the study of these three quantities. Namely, one defines:

The mean value function of the process ξ is given by

$$m(t) = \mathbf{E}(\xi(t)), \tag{3}$$

and its covariance function is defined as

$$R(s,t) = \mathbf{Cov}(\xi(s),\xi(t)) \tag{4}$$

(There is no need to separately define a variance function as the variance of $\xi(t)$ is simply given by $R(t,t)$).

For the mean value function, there are no obvious restrictions, but it is interesting to take a short look at the properties of the covariance function. This function is obviously symmetric, i.e.,

$$R(s,t) = R(t,s), \tag{5}$$

and, which is less obvious, it is positive definite, which means that for any real $x_1,...,x_n$ and for any $t_1,...,t_n \in T$, one has

$$\sum_{i=1}^{n}\sum_{j=1}^{n} x_i x_j R(t_i,t_j) \geq 0. \tag{6}$$

On the other hand, for any symmetric, positive definite function R, there exists a stochastic process ξ that has R as its covariance function (one possible choice is a Gaussian process - see later in this chapter for a definition).

It is by no means true that the mean and covariance functions uniquely determine the process ξ, except for very rare special cases (even for a single random variable, there are infinitely many distributions that give the same mean and variance), so the conclusions made by correlation theory are valid for a large number of processes. This generality, of course, has to be paid for with a restriction on the type of statements that can be made in this context. In particular, one can only consider linear functionals of the process. For these, the mean and variance are readily calculated if one only knows the mean and correlation function of the process; for a nonlinear functional like $\sin(\xi(t))$, on the other hand, it is not even possible to calculate the expectation.

2.3. Convergence

One often wants to study properties of stochastic processes like continuity or one-sided continuity, or the limiting behavior for $t \to \infty$ that involve some kind of limiting operation. In probability theory, there are a few different concepts of convergence that can be applied in such an investigation.

The weakest notion of convergence is convergence in probability. This is defined in the following way:

A sequence ξ_n converges in probability to the random variable ξ if, for any $\epsilon > 0$,

$$\lim \mathbf{P}(|\xi_n - \xi| > \epsilon) = 0. \qquad (7)$$

This concept is so weak that it can hardly be used to arrive at very meaningful conclusions. A process has to satisfy only very mild regularity conditions in order to be, for example, continuous in probability. Thus, for example, the Poisson process (see *Construction of Random Functions an*

A little more meaningful is convergence in square mean. A sequence ξ_n of random variables converges in square mean to the random variable ξ if

$$\mathbf{E}((\xi_n - \xi)^2) \to 0,\ n \to \infty. \qquad (8)$$

Still, this is a somewhat weak concept (the Poisson process, for instance, is still continuous in square mean), but it has the advantage that in deciding whether convergence in square mean holds, one only needs to consider the mean and covariance function of a process, and this makes convergence in square mean the natural concept of convergence in the setting of correlation theory.

The most interesting type of convergence is convergence with probability one. In this case, one demands that convergence (in the usual sense) takes place for all trajectories

of the process except those in a set of probability zero. Unfortunately, here the problems mentioned at the end of Section 2.1 arise, namely, one can only find a version of the process that is, for example, continuous with probability one. So, one has to concentrate on conditions that ensure that a given process has aversion that is continuous (or right continuous, or has a limit for $t \to \infty$, or something similar). Here is one such result, stated without proof:

Theorem 1 *(Kolmogorov) If the process ξ satisfies*

$$\mathbf{E}(|\xi(t+h)-\xi(t)|^{\alpha}) \le Ch^{\beta} \tag{9}$$

with $0<\alpha<\beta$ and $0<C<\infty$, then it has a version with continuous trajectories.

One important concept in this connection is that of a separable process. A process ξ is called separable if there is a countable subset S of T such that for any $t \in T$, $\xi(t)$ is contained in the set of all limit points of sequences $\xi(s_n)$ with $s_n \in \mathbf{S}$ and $s_n \to t$. The set S is called a separable set of the process ξ, and one has the following result, which is again stated without proof:

Theorem 2 *For any stochastic process ξ, there exists a separable process equivalent to it (in other words, every stochastic process has a separable version).*

3. Types of Stochastic Processes

3.1. Gaussian Process

A Gaussian process is a stochastic process with the property that all its finite-dimensional distributions are multivariate normal distributions. These processes are completely characterized by their mean value and covariance functions. One example of a Gaussian process is the Wiener process. One of the many ways to define it is as a mean zero Gaussian process with covariance function

$$R(s,t) = \min(s,t), (s,t \ge 0). \tag{10}$$

Another example is the so-called Brownian bridge. Its parameter space is the interval [0,1], and its covariance function is given by

$$R(s,t) = \min(s,t) - st \tag{11}$$

(Again, the mean value of the function is 0). The Wiener and Brownian bridge process are closely related. The process

$$B(t) = W(t) - tW(1), \tag{12}$$

for instance, is a Brownian bridge, and another way to obtain the Brownian bridge is to consider the conditional distribution of $W(t)$ under $W(1)=0$.

Both the Wiener and the Brownian bridge process are Markov processes, as is the process obtained by letting

$$\xi(t) = e^{-\alpha t/2} W(\alpha e^{t}), \tag{13}$$

where $\alpha > 0$ and W is a Wiener process. This is the so-called Ornstein-Uhlenbeck process, and it is not only Gaussian and Markov, but also stationary (definitions of the Markov property and stationarity are given later in this section; see also *Markov Processes* and *Stationary Processes).* Integrating this process over item, i.e, taking

$$\eta(t) = \int_0^t \xi(t)\, dt, \tag{14}$$

one arrives at another process that is often used as a model for Brownian motion; some times the process ξ is called the Ornstein –Uhlenbeck velocity process, and the process η is called the Ornstein-Uhlenbeck position process. One (among many) differences between the process η and the Wiener process is that the process η has differentiable trajectories, whereas the Wiener process doesn't. Yet, for $a \to \infty$, the process η approaches the wiener process.

It is interesting to note that a Gaussian process admits a representation in terms of a sequence of independent normal random variables. One way to get such a representation is the so-called Karhunen-Loéve expansion which states:

Theorem 3 *Let ξ be a mean zero Gaussian process with state space $[a,b]$ and covariance function R. Then one has the representation*

$$\xi(t) = \sum_{n=1}^{\infty} \lambda_n^{-1/2} \phi_n(t) \eta_n, \tag{15}$$

where λ_n and ϕ_n are the eigenvalues and (orthonormal) eigenfunction of the Fredholm-type equation

$$\phi(t) = \lambda \int_a^b R(s,t)\phi(s) ds, \tag{16}$$

and η_n is a sequence of independent standard normal random variables.

Thus, for example, the Wiener process on the interval [0,1] can be represented as

$$W(t) = \sum_{n=1}^{\infty} \frac{2\sqrt{2}}{(2n+1)\pi} \sin \frac{(2n+1)\pi t}{2} \eta_n. \tag{17}$$

-
-
-

Bibliography

Cramér, H., and Leadbetter, M. (1966) *Stationary and Related Stochastic Processes.* New York: Wiley [This is one of the most important references on stationary processes]

Doob, J. (1945) *Stochastic Processes.* New York: Wiley [This is *the* fundamental reference work on stochastic processes]

Karlin, S., and Taylor, H.M. (1974) *A First Course in Stochastic Processes.* New York: Academic Press [This is a very readable introduction to the theory of stochastic processes]

Karlin, S., and Taylor, H.M. (1981) *A Second Course in Stochastic Processes.* New York: Academic Press [Not only the sequel to the first part, but also interesting in its own right, covers slightly advanced topics like diffusions, branching processes, etc.]

Lawler, G.F., (1995) *Introduction to Stochastic processes.* New York: Chapman and Hall CRC Press [A nice introduction: not too hard to read, makes you solve problems using computer]

Lifshits, M.A. (1995) *Gaussian Random Functions.* Dordrecht: Kluwer [An up-to-date view of the subject of Gaussian processes]

Liptset, R. Sh., and Shiryayev, A. N. (1989) *Theory of Martingales* Dordrecht: Kluwer [An introduction to the theory of martingales and its applications]

Malyshev, V. A., and Minlos, R. A. (1991) *Gibbs Random Fields.* Dordrecht: Kluwer [A discussion of Gibbs random fields and phase transitions]

Rozanov, Yu. A. (1995) *Probability Theory, Random Processes and Mathematical Statistics* Dordrecht: Kluwer [This book is another fundamental reference for many aspects of stochastic processes]

Yaglom, A.M. (1962) *An Introduction to the Theory of Stationary Random Functions.* Englewood Cliffs, NJ, USA: Prentice–Hall [This is a basic reference on stationary processes, giving in particular a good introduction to spectral theory]

Biographical Sketch

Karl Grill received the Ph.D. degree from TU Wien in 1983. Since 1982 he is with TU Wien where he became an Associate Professor in 1988. He was a visiting Professor in the Department of Statistics, University of Arizona during 1991-92. From February to August 1994, he held NSERC Foreign Researcher Award, Carleton University, Ottawa, Canada