AMS 241: Bayesian Nonparametric Methods – Spring 2018
Instructor: Athanasios Kottas

Stochastic processes: basic concepts and definitions

Consider a probability space $(\Omega, \mathcal{F}, P)$, where Ω is the **sample space** of the experiment, an **index set** T, and a **state space** S. A **stochastic process** is a collection

$$\mathcal{X} = \{X(\omega, t) : \omega \in \Omega, t \in T\}$$

such that:
(1) For any n and any set of index points $t_i \in T$, $i = 1,...,n$, $(X_{t_1}, ..., X_{t_n})$ is an n-dimensional random variable (random vector) defined on the probability space $(\Omega, \mathcal{F}, P)$ and taking values in $S^n \equiv S \times ... \times S$. (Hence, for each fixed $t_i \in T$, $X_{t_i}(\cdot) \equiv X(\cdot, t_i) : (\Omega, \mathcal{F}, P) \rightarrow S$ is a random variable.)
(2) For any fixed $\omega \in \Omega$, $X_\omega(\cdot) \equiv X(\omega, \cdot) : T \rightarrow S$ is a function defined on T and taking values in S, referred to as a **sample** (or **sample path** or **realization**) of the stochastic process $\mathcal{X}$.

Conditions (1) and (2) indicate that a stochastic process $\mathcal{X}$ can be viewed either as a collection of random variables $\{X_t : t \in T\}$ or as a collection of random functions $\{X_\omega : \omega \in \Omega\}$.

Depending on the nature of T and S, we can have discrete-time or continuous-time stochastic processes (countable or uncountable T, respectively) and discrete-state or continuous-state stochastic processes (countable or uncountable S, respectively). For the details below, assume that S is a (countable or uncountable) subset of $\mathbb{R}^d$, $d \geq 1$.

Conditions (1) and (2) also indicate that for the study of a stochastic process both distributional properties and properties of sample paths are important. With regard to the former, the distribution function of the random vector $(X_{t_1}, ..., X_{t_n})$,

$$F_{\boldsymbol{t}}(x_1, ..., x_n) = \Pr(X_{t_1} \leq x_1, ..., X_{t_n} \leq x_n),$$

contains all the information for the specific index points $\boldsymbol{t} = (t_1, ..., t_n)$. The collection of all these distribution functions $F_{\boldsymbol{t}}$, as $\boldsymbol{t}$ ranges over all possible vectors of index points of any (finite) length n, is the set of **finite-dimensional distributions** (*f.d.d.s*) of the stochastic process $\mathcal{X}$.

The **Kolmogorov consistency conditions** ensure existence of a stochastic process associated with a set of f.d.d.s. Formally, assume that for each (finite) n and for each set of index points $\boldsymbol{t} = (t_1, ..., t_n)$ (in some index set T), we define a distribution function $F_{\boldsymbol{t}}$. If the collection of all such distribution functions satisfies the Kolmogorov consistency conditions:
(a) $F_{(t_1,...,t_n,t_{n+1})}(x_1, ..., x_n, x_{n+1}) \rightarrow F_{(t_1,...,t_n)}(x_1, ..., x_n)$ as $x_{n+1} \rightarrow \infty$, and
(b) For all n, $\boldsymbol{x} = (x_1, ..., x_n)$, $\boldsymbol{t} = (t_1, ..., t_n)$, and any permutation $\pi = (\pi(1), ..., \pi(n))$ of $\{1, 2, ..., n\}$, $F_{\pi \boldsymbol{t}}(\pi \boldsymbol{x}) = F_{\boldsymbol{t}}(\boldsymbol{x})$, where $\pi \boldsymbol{x} = (x_{\pi(1)}, ..., x_{\pi(n)})$ and $\pi \boldsymbol{t} = (t_{\pi(1)}, ..., t_{\pi(n)})$,
then there exists a probability space $(\Omega, \mathcal{F}, P)$ and a collection $\mathcal{X} = \{X_t : t \in T\}$ of random variables, defined on $(\Omega, \mathcal{F}, P)$, such that the set of $F_{\boldsymbol{t}}$ is the set of f.d.d.s of $\mathcal{X}$.

It is important to note that f.d.d.s do not characterize a stochastic process, that is, they do not always yield complete information about properties of sample paths. It is possible to have two (or more) stochastic processes with the same set of f.d.d.s but with different sample paths. Such processes are called *versions* of one another. (Under conditions on the stochastic process $\mathcal{X}$, it can be shown that there exists a version $\mathcal{Y}$ of $\mathcal{X}$ with some specific property satisfied by its sample paths, e.g., right-continuity or differentiability.)

Using the information provided by the set of f.d.d.s, we can define several useful functions for a stochastic process $\mathcal{X}$. (For all the definitions below, we assume that the required expectations exist.) For any $t \in T$, the **mean function** of $\mathcal{X}$ is

$$\mu(t) \equiv \mathrm{E}(X_t) = \int x \, \mathrm{d}F_t(x).$$

For any $t_i, t_j \in T$, the **covariance function** is given by

$$c(t_i, t_j) \equiv \mathrm{Cov}(X_{t_i}, X_{t_j}) = \mathrm{E}(X_{t_i} X_{t_j}) - \mu(t_i)\mu(t_j)$$

and the **correlation function** by

$$r(t_i, t_j) \equiv \mathrm{Corr}(X_{t_i}, X_{t_j}) = \frac{\mathrm{Cov}(X_{t_i}, X_{t_j})}{\sqrt{\mathrm{Var}(X_{t_i})\mathrm{Var}(X_{t_j})}},$$

provided $\mathrm{Var}(X_{t_i}) > 0$ and $\mathrm{Var}(X_{t_j}) > 0$. An important property of the covariance function is that it is a non-negative definite function, that is, $\sum_{i=1}^{k}\sum_{j=1}^{k} z_i z_j c(t_i, t_j) \geq 0$, for all (finite) k and for any t_1,...,$t_k \in T$ and real constants z_1,...,z_k.

If $c(t_i, t_j) = 0$, for all t_i, t_j with $t_i \neq t_j$, then the stochastic process $\mathcal{X}$ is typically called a **white noise** process. (If X_{t_i} and X_{t_j} are independent for all t_i, t_j with $t_i \neq t_j$, $\mathcal{X}$ is sometimes called a strictly white noise process.) We say that $\mathcal{X}$ is a stochastic process with **uncorrelated (orthogonal) increments** if for any $t_i < t_j < t_k < t_l \in T$, $\mathrm{Cov}(X_{t_j} - X_{t_i}, X_{t_l} - X_{t_k}) = 0$ ($\mathrm{E}((X_{t_j} - X_{t_i})(X_{t_l} - X_{t_k})) = 0$). The process $\mathcal{X}$ has **independent increments** if for any $t_i < t_j < t_k < t_l \in T$, $X_{t_j} - X_{t_i}$ and $X_{t_l} - X_{t_k}$ are independent.

Theory and methods for stochastic processes are considerably simplified under the assumption of (strong or weak) **stationarity**, that imposes certain structure on the set of f.d.d.s (strong stationarity) or the mean function and the covariance function (weak stationarity). Stationarity also has deeper consequences, including spectral representations and ergodic theory.

A stochastic process $\mathcal{X}$ is **strongly (or strictly) stationary** if its f.d.d.s are invariant under *time* shifts, that is, for any (finite) n, for any t_0 and for all $t_1, ..., t_n \in T$, $(X_{t_1}, ..., X_{t_n})$ and $(X_{t_1+t_0}, ..., X_{t_n+t_0})$ have the same distribution (and, as a result, $F_{\boldsymbol{t}} = F_{\boldsymbol{t}+t_0}$, where $\boldsymbol{t} = (t_1, ..., t_n)$ and $\boldsymbol{t} + t_0 = (t_1 + t_0, ..., t_n + t_0)$). A stochastic process $\mathcal{X}$ is **weakly stationary** if its mean function is constant and its covariance function is invariant under *time* shifts. That is, for all $t \in T$, $\mathrm{E}(X_t) = \mu$, and for all $t_i, t_j \in T$, $\mathrm{Cov}(X_{t_i}, X_{t_j}) = c(t_i - t_j)$, a function of $t_i - t_j$ only.