36-752: Advanced Probability Theory **Spring 2018**

Lecture 8: February 22

Lecturer: Alessandro Rinaldo *Scribes: Heejong Bong*

Note: *LaTeX template courtesy of UC Berkeley EECS dept.*

Last Time:

1. Stochastic Process:
   Let T be a index set. $\forall t \in T$, X_t: a random variable on $(\Omega, \mathcal{F}, P)$ taking values in $(\mathcal{X}_t, \mathcal{F}_t)$.
   i.e., X_t: $\omega \in \Omega \mapsto X_t(w)$.
2. Also, this can be represented as a set of random functions from T into $\mathcal{X} = \cup_{t\in T}\mathcal{X}_t$.
   i.e., $\omega \mapsto f(w,t)$: a random function at t.

New Material:

8.1 Stochastic Process

8.1.1 Projections and Cylinder sets

How do we represent probability distributions on $\cap_{t\in T}\mathcal{X}_t$? We first need the notion of measurability on the space of functions from T into $\mathcal{X}$.

Definition 8.1

1) *For $t \in T$, $\pi_t : \mathcal{X} \to \mathcal{X}_t$, given by $\pi_t(f) = f(t)$, is called a* ***projection*** *or an* ***evaluation functional****.*
2) *A* ***one-dimensional cylinder set*** *is a set of the form $\Pi_{t\in T} B_t$ where $B_{t_1} \in \mathcal{F}_{t_1}$ for one $t_1 \in T$, and $B_t = \mathcal{X}_t$, $\forall t \neq t_1$.*
3) *A* k***-dimensional cylinder set*** *is a set of the form $\Pi_{t\in T} B_t$ where $B_{t_i} \in \mathcal{F}_{t_i}$, $i = 1, \ldots, k$ for some $\{t_1, \ldots, t_k\} \subset T$, and $B_t = \mathcal{X}_t$, $\forall t \notin \{t_1, \ldots, t_k\}$.*

Definition 8.2 *The* ***product σ-field****, $\Pi_{t\in T}\mathcal{F}_t$, is the smallest σ-field on Ω containing all one-dimensional cylinder sets.*

Note that, if T is finite, this definition is equivalent to the previous definition on product σ-fields.

Lemma 8.3 *The product σ-field is the smallest σ-field such that all π_t, $t \in T$ are measurable.*

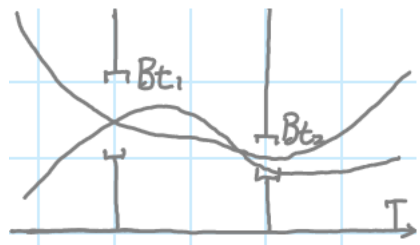

Figure 8.1: two examples of realizations of a two-dimensional cylinder

Proof For a fixed $t_1 \in T$ and for each $B_{t_1} \in \mathcal{F}_t$, $\pi_{t_1}^{-1}(B_{t_1}) = \Pi_{t \in T} B_t$ where $B_t = \mathcal{X}_t$, $\forall t \neq t_1$. ■

Theorem 8.4 *Define $X : \Omega \to \mathcal{X}$ by setting $X(\omega)$ to be a function f such that*

$$f(t) = X_t(w), \forall t$$

Then, X is $\mathcal{F}/\Pi_t \mathcal{F}_t$-measurable if and only if X_t is $\mathcal{F}/\mathcal{F}_t$-measurable.

8.1.2 Kolmogorov's extension theorem

Now, how do we specify probability on stochastic processes as a set of realizations? We cannot describe the exact probability distribution for the realizations on every t. However, with the Kolmogorov's extension theorem, we can handle with the probability of realizations on the finite number of t's.

Definition 8.5 *For $v \subset T$, let $(\mathcal{X}_v, \mathcal{F}_v)$ be the product space and the product σ-field restricted to v. Also, suppose that P_v is a measure on $(\mathcal{X}_v, \mathcal{F}_v)$. For $u \subset v$, the* ***projection*** *on P_v, $\pi_u(P_v)$, is defined to be*

$$[\pi_u(P_v)](B) := P_v(\{x \in \mathcal{X}_v : x_u \in B\}) \text{ for } B \in \mathcal{F}_u$$

, which is also said to be the ***marginal distribution*** *of P_v over $\mathcal{X}_u$.*

Theorem 8.6 *(Kolmogorov's extension theorem) Let $\mathcal{X}_t = \mathbb{R}$, $\forall t \in T$. Then, $\mathcal{X} = \mathbb{R}^T$. Assume that for every finite subset $v \subset T$, P_v is well-defined over $(\mathbb{R}^v, \mathcal{B}^{|v|})$ and that P_v is consistent. (i.e., $\pi_u(P_v) = P_u$.) Then, $\exists! P$ on $(\mathbb{R}^T, \mathcal{B}^T)$ such that $\pi_v(P) = P_v$, $\forall v \subset T$.*

Note again that, if T is finite, then it is equivalent to the case of product measures of random vectors.

8.1.3 Example: Brownian motion

Suppose that $W_t : t \geq 0$ satisfies

1. $W_0 = 0$ with probability 1.
2. If $0 \leq t_0 < t_1 < \ldots < t_k$, $W_{t_i} - W_{t_{i-1}}$ are mutually independent.
   i.e., W_t's have independent increments.
3. $\forall 0 \leq s < t$, $W_t - W_s \sim \mathcal{N}(0, t)$.

Some useful facts of $\mathcal{W}_t$ are as follows:

1. $\mathbb{E}[W_t] = 0$
2. $\mathbb{E}[W_t^2] = t$, $\forall t \geq 0$
   $\because W_t = W_t = W_t - W_0 \sim \mathcal{N}(0, t)$.
3. $$\begin{aligned} Cov[W_s, W_t] &= \mathbb{E}[W_s(W_t - W_s)] + \mathbb{E}[W_s^2] \\ &= \mathbb{E}[(W_s - W_0)(W_t - W_s)] + \mathbb{E}[W_s^2] \\ &= \mathbb{E}[W_s^2] \text{since } W_s - W_0 \perp W_t - W_s \text{ with } 0 \leq s < t \\ &= s = \min\{s, t\}. \end{aligned}$$

To apply the Kolmogorov's extension theorem, we need a distribution of $(W_{t_1}, \ldots, W_{t_k})$, $0 \leq t_1 < \ldots < t_k$. Notice that $(W_{t_1}, W_{t_2} - W_{t_1}, \ldots, W_{t_k} - W - t_k)$ has a joint multivariate Gaussian distribution with the mean 0 and the diagonal covariance matrix. Thus, $(W_{t_1}, \ldots, W_{t_k})$ has a *pdf*,

$$f(x_1, \ldots, x_k) = \Pi_{i=1}^k \frac{1}{\sqrt{2\pi(t_i - t_{i-1})}} \exp\left(-\frac{(x_i = x_{i-1})^2}{2(t_i - t_{i-1})}\right)$$

where $t_0 = x_0 = 0$. It is the same as the distribution of $(S_1, \ldots, S_k)$ where $X_i \sim \mathcal{N}(0, t_i - t_{i-1})$ independently and $S_i = X_1 + \ldots + X_i$.

Now, $\forall i = 1, \ldots, k$, let $g_i(x_1, \ldots, x_k) = (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_k)$. If $\mu_{t_1, \ldots, t_k}$ is the distribution of $(W_{t_1}, \ldots, W_{t_k})$, then $\mu_{t_1, \ldots, t_{i-1}, t_{i+1}, \ldots, t_k} = \mu_{t_1, \ldots, t_k} g^{-1}$. Hence, the probabilities defined on $(W_{t_1}, \ldots, W_{t_k})$ are consistent and then the Kolmogorov's extension theorem shows the existence of probability distribution for the stochastic process, the Brownian motion.

8.2 L^p-spaces

Let $(\Omega, \mathcal{F}, \mu)$ be a measure space. Specifically, suppose that $\Omega = [0, 1]$ OR $\mathbb{R}$, $\mathcal{F} = \mathbb{B}^1$, and μ: the Lebesgue measure. Let $\mathcal{L}^p$, for $p > 0$, be the set of real-valued functions such that $\int |f|^p d\mu < \infty$. For $f \in \mathcal{L}^p$, let $\|f\|_p = [\int |f|^p d\mu]^{\frac{1}{p}}$.

Then,

1. $\|f\|_p \geq 0$
2. If $f, g \in \mathcal{L}^p$, then $f + g \in \mathcal{L}^p$, and
   $$\|f + g\|_p \leq \|f\|_p + \|g\|_p (\because \text{ triangle inequality}).$$
3. If $f \in \mathcal{L}^p$, then $\forall a \in \mathbb{R}$, $af \in \mathcal{L}^p$, and $\|af\|_p = |a|\|f\|_p$.

i.e., $\|\cdot\|_p$ behaves like a norm on $\mathcal{L}^p$.

However,

4. $\|f\|_p = 0 \nrightarrow f = 0$. i.e., f may be nonzero on any measure zero set.

Hence, let L^p to be a set of equivalence classes in $\mathcal{L}^p$ where $f \sim g$ when $f = g$ a.e. μ. Then, L^p is a normed vector space with norm $\|f\|_p$ for $[f] \in L^p$, where $[f]$ is the set of all functions g such that $g = f$ a.e. μ.

If $p = \infty$, let

$$\begin{aligned}\|f\|_\infty =& \sup\{t \geq 0 : \mu(\{\omega : |f(\omega)| \geq t\})\} \\ =& \inf\{\alpha \geq 0 : \mu(\{\omega : |f(\omega)| > \alpha\})\}\},\end{aligned}$$

which is said to be an **essential supremum** of f. Note that it is different from $\sup_x |f(x)|$.

Proposition 8.7

1) $\lim_{p\to\infty} \|f\|_p = \|f\|_\infty$.

2) If $\mu(\Omega) = 1$ *and* $p \leq q$, *then* $\|f\|_p \leq \|f\|_q$.

Before proving the second proposition, I would introduce a useful inequality: the Hölder's inequality.

Lemma 8.8 *(Hölder's inequality) If* $f \in \mathcal{L}^p$ *and* $g \in \mathcal{L}^q$ *where* $\frac{1}{p} + \frac{1}{q} = 1$, *then*

1) $fg \in \mathcal{L}^1$, *and*

2) $\|f\|_p \|q\|_q \geq \|fg\|_1$.

In case of $p = q = 2$, *we get the Cauchy-Schwartz inequality.*

Proof of Proposition 2)

$$\begin{aligned}\int |f|^p d\mu =& \int |f|^p \cdot 1 d\mu \\ \leq& \left(\int |f|^{p\frac{q}{p}} d\mu\right)^{\frac{p}{q}} (\int 1^{\frac{q}{q-p}} d\mu)^{1-\frac{p}{q}} \text{ by the Hölder's inequality} \\ & \text{where } \frac{q}{p} \text{ and } \frac{q}{q-p} \text{ are conjugate} \\ =& (\|f\|_q)^p [\mu(\Omega)]^{1-\frac{p}{q}}.\end{aligned}$$

Hence, $\|f\|_p \leq \|f\|_q$. ■

In sum, if $\mu(\Omega) = 1$ and $1 \leq p < q$, then $\|f\|_1 \leq \|f\|_p \leq \|f\|_q \leq \|f\|_\infty$.