Lecture 29: **Brownian Motion, Brownian Bridge, Application of Brownian Bridge, Kolmogorov-Smirnov Test**

Definition 1. X_t for $t \in [0, \infty)$ is a Brownian motion if X_t is sample continuous
$\mathbb{E}X_t = 0, \operatorname{cov}(X_t, X_s) = \min(t, s)$

Existence

From finite-dim distribution, Gaussian.
Prove sample continuous.

About sample continuity:

$X_t, \mathbb{E}X_t = 0, \mathbb{E}X_t X_s = \min(t, s)$

$\mathcal{N}(0,1)(c, \infty) \le e^{-\frac{c^2}{2}}, c > 0$

$$\mathcal{N}(0,1)(c, \infty) = \frac{1}{\sqrt{2\pi}} \int_c^\infty e^{-\frac{x^2}{2}} dx \le \frac{1}{\sqrt{2\pi}} \int_c^\infty \frac{x}{c} e^{-\frac{x^2}{2}} dx = \frac{1}{\sqrt{2\pi}} \frac{1}{c} e^{-\frac{c^2}{2}}$$

$$c > \frac{1}{\sqrt{2\pi}} \Rightarrow (*) \le e^{-\frac{c^2}{2}}$$

$c < \frac{1}{\sqrt{2\pi}} \Rightarrow \mathcal{N}(0,1)(c, \infty) \le \mathcal{N}(0,1)(0, \infty) = \frac{1}{2} \le e^{-\frac{1}{2}(\frac{1}{\sqrt{2\pi}})^2} = e^{-\frac{c^2}{2}}$

$s < t, X_t - X_s \sim \mathcal{N}(0, t - s)$

$\mathbb{E}[X_t - X_s]^2 = t + s - 2s = t - s$

$$X_t, t \in [0,1], n \ge 1, V_k = X_{\frac{k+1}{2^n}} - X_{\frac{k}{2^n}}, k = 0, ..., 2^n - 1, \operatorname{var}(V_k) = \frac{1}{2^n}$$

$$\mathbb{P}\left(\sup_k |V_k| \ge \frac{1}{n^2}\right) \le 2^n \mathbb{P}(|V_t| \ge \frac{1}{n^2}) = 2^n \mathbb{P}\left(\left|\frac{1}{\frac{1}{2^{\frac{n}{2}}}}\right| \ge 2^{\frac{n}{2}} \frac{1}{n^2}\right) \le 2^n e^{-\frac{1}{2} 2^n \frac{1}{n^2}} = 2^n e^{-\frac{2^{n-1}}{n^2}}; \sum_n < \infty$$

$$\Rightarrow \left(\sup_k |V_k| \ge \frac{1}{n^2} \text{ i.o}\right) = 0$$

$$t = \sum_{j=1}^\infty \frac{t_j}{2^j}, t_j = 0, 1, 2, ... \qquad t(n) = \sum_{j=1}^n \frac{t_j}{2^j}$$

$X_{t(n)} - X_{t(n-1)} \in \{0, V_k, k = 0, ...\}$

$$X_{t(n)} = 0 + X_{t(1)} + \sum_{2 \le j \le n} (X_{t(j)} - X_{t(j-1)}) \xrightarrow{\text{a.s.}} Z_t$$

$Z_t = X_t$ for dyadic t

Since Z_t agrees with X_t on a dense set of t, sufficient to show Z_t continuous:

take $n_0(\omega)$ s.t. for $n \ge n_0(\omega), \sup_k |V_k| \le \frac{1}{n^2}$

take $t, s, |t - s| \le 2^{-n} \Rightarrow t(n) = \frac{k}{2^n}, s(n) = \frac{\varphi}{2^n}, |k - \varphi| \in \{0, 1\}$

$|X_{t(n)} - X_{s(n)}| \in \{0, |V_k|\} \Rightarrow |X_{t(n)} - X_{s(n)}| \le \frac{1}{n^2}$

$|Z_t - Z_s| \le |Z_t - X_{t(n)}| + |X_{t(n)} - X_{s(n)}| + |X_{s(n)} - Z_s|$ (Z_t defined as limit of $X_{t(n)}$)

$$\le \sum_{m \ge n} \frac{1}{m^2} + \frac{1}{n^2} + \sum_{m \ge n} \frac{1}{m^2} \le \frac{c}{n} \to 0 \text{ as } n \to \infty$$

Definition 2. $B_t, t \in [0,1]$, is a Brownian bridge if B_t is Gaussian, $\mathbb{E}B_t = 0$, $\mathbb{E}B_t B_s = s(1-t), s < t$ and B_t continuous.

Example.

$B_t = X_t - tX_T$ is a Brownian bridge, X_t is Brownian motion.

$s < t \mathbb{E}B_sB_t = \mathbb{E}(X_t - tX_1)(X_s - sX_1) = s - st - ts + st = s(1-t)$

Notice $B_0 = B_1 = 0 \Leftarrow$ motivation for the name "bridge"

Applicatin of Brownian bridge

$u_1, ..., u_n$ iid, uniform on $[0, 1]$

Empirical cdf $\Rightarrow$ True cdf, $\frac{1}{n}\sum_{i=1}^{n}\mathbf{I}(u_i \le t) \xrightarrow{\text{LLN}} \mathbb{P}(u_i \le t) = t$

Convergence rate:

For each t, $X_{n,t} = \sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n}(\mathbf{I}(u_i \le t) - t\right) \to \mathcal{N}(0, t(1-t))$

$X_{n,t}$ random process, as $n \to \infty$, $X_t \to$?

For finite t's - F

$(X_{n,t})_{t\in F} \xrightarrow{\text{in law}} (B_t)_{t\in F}$

Check: $\mathbb{E}X_{n,t}X_{n,s} = \mathbb{E}(\mathbf{I}(u_i \le t) - t)(\mathbf{I}(u_k \le s) - s) = s - ts - ts + ts = s(1-t)$

So Brownian bridge is the limit of the empirical distribution.

Kolmogorov-Smirnov test

$F(t) = \mathbb{P}(x \le t), F_n(t) = \frac{1}{n}\sum_{i=1}^{n}\mathbf{I}(X_i \le t)$

$\sup_t \sqrt{n}|F_n(t) - F(t)| \overset{\text{in dist}}{=\!=\!=\!=\!=} \sup_t \sqrt{n}\left|\frac{1}{n}\sum \mathbf{I}(u_i \le t) - t\right|$

If F - continuous, $\mathbf{I}(X_i \le F^{-1}(y)) = \mathbf{I}(\underbrace{F(X_i)}_{=u_i} \le y)$ br

$\to \sup_{t\in[0,1]} |B_t|$ test if data from F or not

$(C[0,1], ||\cdot||_\infty)$ - complete separable space

Consider $(C[0,\infty), ?)$

$C([0,\infty)), d_n(f,g) = \sup_{0\le x\le n} |f(x) - g(x)|, d(f,g) = \sum_{n\ge 1}\frac{1}{2^n}\frac{d_n(f,g)}{1 + d_n(f,g)}$

meaning: $d(f_j, f) \to 0$ iff $d_n(f_j, f) \xrightarrow{j\to\infty} 0$ $\forall n$

so d metrize the space.

$(C([0,\infty)), d)$ - complete separable space

e.g. completness, Cauchy sequence $\Rightarrow$ Cauchy sequence in each dim
same for separable

$\{\mathbb{P}_n\}_{n\ge1}$ - laws on $(C([0,\infty)), d)$

$\mathbb{P}_n \to \mathbb{P} \Rightarrow \{\mathbb{P}_n\}$ - uniformly tight, i.e. $\forall\epsilon\exists K$ - compact in $C([0,\infty))$, $\forall n$ $\mathbb{P}_n(K) > 1 - \epsilon$

$\{\mathbb{P}_n\}$ - uniformly tight on complete separable space
$\Leftrightarrow$ totally bounded in ρ, β
$\Leftrightarrow$ for any subsequence of $\mathbb{P}_n$, there exists a converging subsequence

To prove convergence, only need to prove convergence of finite-dim distributions, the rest is to show uniform tightness.

Theorem (Arzela-Ascoli). K - compact in $(C([0,\infty)), \delta) \Leftrightarrow K$ - closed and K uniformly bounded and equicontinuous on each $[0, n]$