Random Elements in a Hilbert Space

Tan-Jianbin

School of Mathematics
Sun Yat-sen University

Seminar on Statistics 105c

Table of Contents

Table of Contents

Definition

H is a separable Hilbert space. If $\mathcal{X}$ is a random element $(\Omega, \mathcal{F}, P) \to (H, \mathcal{B}(H))$ and $\int ||\mathcal{X}||dP < \infty$. The mean of $\mathcal{X}$, $E\mathcal{X} = \int \mathcal{X}dP$.

$E\mathcal{X} = m \Leftrightarrow E\langle \mathcal{X}, x\rangle = \langle m, x\rangle$, $\forall x \in H$.

$E||\mathcal{X}||^2 < \infty$, we can similarly define $Var(\mathcal{X}) = E||\mathcal{X} - E\mathcal{X}||^2 = E||\mathcal{X}||^2 - 2E\langle \mathcal{X}, E\mathcal{X}\rangle + ||E\mathcal{X}||^2 = E||\mathcal{X}||^2 - ||E\mathcal{X}||^2$.

Definition

If $E||\mathcal{X}||^2 < \infty$, covariance operator for $\mathcal{X}$: $\mathcal{K} = E(\mathcal{X} - E\mathcal{X}) \otimes (\mathcal{X} - E\mathcal{X})$.

$E||\mathcal{X}||^2 < \infty \Rightarrow Var(\mathcal{X})$ exists $\Rightarrow \mathcal{K}$ exists. $(\mathcal{X} - E\mathcal{X}) \otimes (\mathcal{X} - E\mathcal{X}) \in B_{HS}(H)$, then $\mathcal{K} \in B_{HS}(H)$.

$E(\mathcal{X} - E\mathcal{X}) \otimes (\mathcal{X} - E\mathcal{X}) = E(\mathcal{X} \otimes \mathcal{X}) - E\mathcal{X} \otimes E\mathcal{X}$.

Property

$\mathcal{X}$ is a random element s.t. $E\mathcal{X} = 0$, $E||\mathcal{X}||^2 < \infty$, $\forall\ x, y \in H$
$\langle \mathcal{K}x, y\rangle = E\langle \mathcal{X}, x\rangle\langle \mathcal{X}, y\rangle$ *and* $\mathcal{K} \gg 0$.

Proof.

$\langle \mathcal{K}x, y\rangle = \langle E(\mathcal{X} \otimes \mathcal{X})x, y\rangle = E\langle(\mathcal{X} \otimes \mathcal{X})x, y\rangle = E\langle \mathcal{X}, x\rangle\langle \mathcal{X}, y\rangle$. □

$\langle \mathcal{K}x, y\rangle = Cov(\langle \mathcal{X}, x\rangle, \langle \mathcal{X}, y\rangle)$.

Definition

$(\Omega, \mathcal{F}, P)$ probability space, H_i separable Hilbert spaces, $\mathcal{X}_i$ random elements in H_i s.t. $E||\mathcal{X}_i||_i^2 < \infty$ and $E\mathcal{X}_i = 0$, cross-covariance operator: $\mathcal{K}_{12} = E(\mathcal{X}_2 \otimes \mathcal{X}_1) \in B_{HS}(H_2, H_1)$.

$\langle \mathcal{K}_{12}x, y\rangle_1 = E\langle \mathcal{X}_1, y\rangle_1 \langle \mathcal{X}_2, x\rangle_2 \Rightarrow |\langle \mathcal{K}_{12}x, y\rangle_1| \leq (E\langle \mathcal{X}_1, y\rangle_1^2)^{1/2}(E\langle \mathcal{X}_2, x\rangle_2^2)^{1/2} = \langle \mathcal{K}_1 y, y\rangle_1^{1/2} \langle \mathcal{K}_2 x, x\rangle_2^{1/2}$ and $\mathcal{K}_{12}^* = \mathcal{K}_{21}$.

The generalized correlation measure $\mathcal{R}_{12} = \mathcal{K}_1^{-1/2}\mathcal{K}_{12}\mathcal{K}_2^{-1/2}$ can be define in multivariate analysis case.

But this is not right when we focus on infinite-dim case since that compact operator is not invertible in infinite-dim.

Theorem

$\exists\, \mathcal{R}_{12} \in B(H_2, H_1)$, $||\mathcal{R}_{12}|| \leq 1$ *s.t.* $\mathcal{K}_{12} = \mathcal{K}_1^{1/2}\mathcal{R}_{12}\mathcal{K}_2^{1/2}$.

Proof.

$\forall y \in Im(\mathcal{K}_2^{1/2})$, $\mathcal{K}_2^{-1/2} = (\mathcal{K}_2^{1/2})^\dagger$, $x = \mathcal{K}_2^{-1/2}y$, then $||\mathcal{K}_1^{-1/2}\mathcal{K}_{12}\mathcal{K}_2^{-1/2}y||_1$
$= ||\mathcal{K}_1^{-1/2}\mathcal{K}_{12}x||_1 = (\langle \mathcal{K}_{12}x, \mathcal{K}_1^\dagger\mathcal{K}_{12}x\rangle_1)^{1/2} \leq \langle \mathcal{K}_1\mathcal{K}_1^\dagger\mathcal{K}_{12}x, \mathcal{K}_1^\dagger\mathcal{K}_{12}x\rangle_1^{1/4}\langle \mathcal{K}_2x, x\rangle_2^{1/4} \Rightarrow$
$||\mathcal{K}_1^{-1/2}\mathcal{K}_{12}x||_1 \leq ||\mathcal{K}_1^{-1/2}\mathcal{K}_{12}x||_1^{1/2}||\mathcal{K}_2^{1/2}x||_2^{1/2} \Rightarrow ||\mathcal{K}_1^{-1/2}\mathcal{K}_{12}\mathcal{K}_2^{-1/2}y||_1 \leq ||y||_2$.

Extend $\mathcal{K}_1^{-1/2}\mathcal{K}_{12}\mathcal{K}_2^{-1/2}$ to H_2 and get $\mathcal{R}_{12} \in B(H_2, H_1)$ s.t. $||\mathcal{R}_{12}|| \leq 1$. □

Lemma

H separable Hilbert space, define $\mathcal{G} = \cup_{C\in\mathcal{C}} \cup_{T\in H^*} T^{-1}(C)$, $\mathcal{C}$ is the collection of all the open subset of R, then $\sigma(\mathcal{G}) = \mathcal{B}(H)$.

Proof.

It's easy to show that $\sigma(\mathcal{G}) \subset \mathcal{B}(H)$. It's suffices to show that $\sigma(\mathcal{G})$ contains all the open balls since H is separable. Take COB $\{e_n\}$.

$\forall\ r > 0,\ x \in H,\ B(x;r) = \{y \in H; ||x-y|| < r\} = \{||y||^2 < 2\langle x,y\rangle + ||x||^2 + r\}$
$= \cup_{q\in Q}[\{||y||^2 < q\} \cap \{q < 2\langle x,y\rangle + ||x||^2 + r\}]$ and $\{q < 2\langle x,y\rangle + ||x||^2 + r\}) \in \mathcal{G}$,
$\{||y||^2 < q\} = \{\sum_n \langle y,e_n\rangle^2 < q\} = \cup_{q_1\in Q}\{\langle y,e_1\rangle^2 < q_1\} \cap \{\sum_{k\geq 1}\langle y,e_n\rangle^2 < q - q_1\}$
$\in \sigma(\mathcal{G}) \Rightarrow B(x;r) \in \sigma(\mathcal{G})$. $\square$

Theorem

$(\Omega, \mathcal{F}, P)$ probability space, $\mathcal{X} : (\Omega, \mathcal{F}) \to (H, \mathcal{B}(H))$, $\mathcal{X}$ is measurable if and only if $\langle \mathcal{X}(\cdot), x \rangle$ is measurable for all $x \in H$, and F_X is uniquely determined by $F_{\langle \mathcal{X}(\cdot), x \rangle}$, $\forall\ x \in H$.

Proof.

Let $\mathcal{G} = \cup_{C \in \mathcal{C}} \cup_{T \in H^*} T^{-1}(C)$, then $\mathcal{B}(H) = \sigma(\mathcal{G})$ and we just need to check that $\forall A \in \mathcal{G}$, $\mathcal{X}^{-1}(A) \in \mathcal{F}$. $\forall A \in \mathcal{G}$, $\mathcal{X}^{-1}(A) = \mathcal{X}^{-1}(T^{-1}(C)) = \{\omega \in \Omega; T(\mathcal{X}(\omega)) \in C\}$ $= \{\omega \in \Omega; \langle \mathcal{X}(\omega), x_T \rangle \in C\} \in \mathcal{F} \Rightarrow \mathcal{X}$ is measurable.

Let $\mathcal{G}' = \{\cap_n A_n;$ some finite $A_n \in \mathcal{G}\ \}$, then $\mathcal{G}'$ is π system and $\sigma(\mathcal{G}') = \sigma(\mathcal{G})$, $P(\mathcal{X} \in \cap_n A_n) = P(\cap_n(\mathcal{X} \in A_n)) = P(\cap_n\{\mathcal{X} \in T_n^{-1}(C_n)\}) = P(\cap_n\{\langle \mathcal{X}(\omega), x_n \rangle \in C_n\})$. $\square$

Definition

$(\Omega, \mathcal{F}, P)$ *is a probability space,* $\mathcal{X}_1$ *and* $\mathcal{X}_2$ *are independence if* $\forall A, B \in \mathcal{F}$, $P(\mathcal{X}_1 \in A, \mathcal{X}_2 \in B) = P(\mathcal{X}_1 \in A)\ P(\mathcal{X}_2 \in B)$.

We can similarly define the distribution $F_{\mathcal{X}} := P \circ \mathcal{X}^{-1}$.

There is a question remained: Did the sequence of independence random elements exist? To be precisely, if $\mathcal{X}_n$: $(\Omega_n, \mathcal{F}_n, P_n) \to (H_n, \mathcal{B}(H_n))$, do we have a probability space $(\Omega, \mathcal{F}, P)$ s.t. $\{\mathcal{X}_n\}$ is sequence of independence random elements. This is true by Kolmogorov's extension theorem.

Theorem

$\mathcal{X}_i$ *independence* $\Leftrightarrow \langle \mathcal{X}_i, x_i \rangle_i$ *independence*, $\forall\ x_i \in H_i$.

Proof.

$"\Rightarrow":\ P(\cap_i \langle \mathcal{X}_i, x_i \rangle_i \in C_i) = P(\cap_i \mathcal{X}_i \in T_i^{-1}(C_i)) = \prod_i P(\langle \mathcal{X}_i, x_i \rangle_i \in C_i)$.
$"\Leftarrow":\ P(\cap_i \{\mathcal{X}_i \in \cap_{n_i} T_{n_i}^{-1}(C_{n_i})\}) = P(\cap_i \cap_{n_i} \{\langle \mathcal{X}_i, x_{n_i} \rangle \in C_{n_i}\}) =$
$\prod_i P(\cap_{n_i} \{\langle \mathcal{X}_i, x_{n_i} \rangle \in C_{n_i}\}) = \prod_i P(\mathcal{X}_i \in \cap_{n_i} T_{n_i}^{-1}(C_{n_i}))$. $\square$

Let $\mathcal{X}_n \in \mathcal{B}(H)$. $(H, \mathcal{B}(H), (F_{\mathcal{X}_n})_{n\in N})$ is the statistical structure.

Theorem

If $\mathcal{X}_n$ iid with $E||\mathcal{X}_n|| < \infty$, then $\frac{\sum_n \mathcal{X}_n}{n} \to E\mathcal{X}_1$ a.s..

Proof.

Define the truncating element of $\mathcal{X}_n$: $\mathcal{Y}_n = \mathcal{X}_n I_{\{||\mathcal{X}_n|| \leq n\}}$. One can show $\frac{\sum_{j\leq k_n} \mathcal{Y}_j - E\sum_{j\leq k_n} \mathcal{Y}_j}{k_n} \to 0$ a.s., $k_n = [\alpha^n]$, $\alpha > 1$.

$$||\frac{E\sum_{j\leq k_n} \mathcal{Y}_j}{k_n} - m|| = ||\frac{\sum_{j\leq k_n} E(\mathcal{Y}_j - \mathcal{X}_j)}{k_n}|| \leq \frac{\sum_{j\leq k_n} E||\mathcal{X}_j|| I_{\{||\mathcal{X}_j|| > j\}}}{k_n} \to 0 \Rightarrow$$

$\frac{E\sum_{j\leq k_n} \mathcal{Y}_j}{k_n} \to m \Rightarrow \frac{\sum_{j\leq k_n} \mathcal{Y}_j}{k_n} \to m \Rightarrow \frac{\sum_{j\leq k_n} \mathcal{X}_j}{k_n} \to m \Rightarrow \frac{\sum_n \mathcal{X}_n}{n} \to m.$ □

M is a metric space. Let $\mathcal{P}(M)$ denote the collection of all probability measures defined on $(M, \mathcal{B}(M))$.

Definition

$\{\mu\}, \{\mu_t\}_{t\in I} \subset \mathcal{P}(M)$, then $\mu_n \xrightarrow{w} \mu$ if $\forall\ f \in C_b(M)$ s.t. $\int f d\mu_n \to \int f d\mu$. We say $\mathcal{X}_n \xrightarrow{d} \mathcal{X}$ if $F_{\mathcal{X}_n} \xrightarrow{w} F_{\mathcal{X}}$.

The definition of weak convergence can defines a topology structure on $\mathcal{P}(M)$.

Definition

$\{\mu_t\}_{t\in I}$ *is tight if* $\forall\ \varepsilon > 0$, $\exists$ *compact* $W \in \mathcal{B}(M)$ *s.t.* $\inf_{t\in I}\mu_t(W) \geq 1-\varepsilon$.

Prohorovs theorem: M be a complete separable metric space. $K \subset \mathcal{P}(M)$ is tight $\Leftrightarrow K \subset\subset \mathcal{P}(M)$.

$S^\delta = \{x \in H; d(x,S) \leq \delta\}$, $B^r(y) = \{x \in H : |\langle x, y_k\rangle| \leq r\}$

Lemma

Let $M = H$, *then* $\{\mu_t\}_{t\in I}$ *is tight if* $\forall\ \varepsilon, \delta > 0$, $\exists$ *finite* $\{y_k\}$, $S = span\{y_k\}$ *and* $r > 0$ *s.t.* $\inf_{t\in I}\mu_t(S^\delta) \geq 1-\varepsilon$, $\inf_{t\in I}\mu_t(B^r(y)) \geq 1-\varepsilon$.

Theorem

If $\langle \mathcal{X}_n, x\rangle \xrightarrow{d} \langle \mathcal{X}, x\rangle$, $\forall\ x \in H$ *and* $\forall\ \varepsilon, \delta > 0$, $\exists$ *finite-dim* $S \subset H$ *s.t.* $\inf_{n\geq 1} P(\mathcal{X}_n \in S^\delta) \geq 1-\varepsilon$, *then* $\mathcal{X}_n \xrightarrow{d} \mathcal{X}$.

Proof.

$\{F_{\mathcal{X}_n}\}$ is tight $\Rightarrow$ if $\mathcal{X}_{n_k} \xrightarrow{d} \mathcal{Y}$, $\mathcal{X}_{n_g} \xrightarrow{d} \mathcal{Z}$, then $\forall x \in H$, $\langle \mathcal{X}_{n_k}, x\rangle \xrightarrow{d} \langle \mathcal{Y}, x\rangle$ and $\langle \mathcal{X}_{n_g}, x\rangle \xrightarrow{d} \langle \mathcal{Z}, x\rangle \Rightarrow \langle \mathcal{Y}, x\rangle \stackrel{d}{=} \langle \mathcal{Z}, x\rangle \stackrel{d}{=} \langle \mathcal{X}, x\rangle \Rightarrow \mathcal{Y} = \mathcal{Z} = \mathcal{X}$. □

Definition

$\mathcal{X}$ *Gaussian element if* $\langle \mathcal{X}, x\rangle$ *is a Gaussian random variable* $\forall\ x \in H$. *Noticed* $E\langle \mathcal{X}, x\rangle = \langle m, x\rangle$. $Var\langle \mathcal{X}, x\rangle = \langle \mathcal{K}x, x\rangle$. *We mark that* $\mathcal{X} \sim \mathcal{N}(m, \mathcal{K})$.

Theorem

$\mathcal{X}_n$ *iid with* $m = 0$ *and* $E||\mathcal{X}_1||^2 < \infty \Rightarrow \xi_n = n^{-1/2}\sum_n \mathcal{X}_n \xrightarrow{d} \mathcal{X}$, $\mathcal{X} \sim \mathcal{N}(0, \mathcal{K})$, $\mathcal{K} = E(\mathcal{X}_1 \otimes \mathcal{X}_1)$.

Proof.

$\forall\ \varepsilon, \delta > 0$, $\{e_n\}$ COB, $S_j = span\{e_n, n \leq j\}$, $P(\xi_n \in S_j^\delta) = P(||P_{S_j^\perp}\xi_n|| \leq \delta) = 1 - P(||P_{S_j^\perp}\xi_n|| > \delta) \geq 1 - \frac{E||P_{S_j^\perp}\xi_n||^2}{\delta^2} = 1 - \frac{E||P_{S_j^\perp}\mathcal{X}_1||^2}{\delta^2} \geq \varepsilon$, if j is large enough.

$\langle n^{-1/2}\sum_n \mathcal{X}_n, x\rangle = n^{-1/2}\sum_n \langle \mathcal{X}_n, x\rangle \xrightarrow{d} N(0, \langle \mathcal{K}x, x\rangle)$ □

Let $m_n = \frac{\sum_i \mathcal{X}_i}{n}$ and $\mathcal{K}_n = \frac{\sum_i (\mathcal{X}_i - m_n)\otimes(\mathcal{X}_i - m_n)}{n-1}$, if $E||\mathcal{X}_1|| \leq \infty$, $m_n \to m$ a.s..
Moreover, if $E||\mathcal{X}_1||^2 \leq \infty$, $\sqrt{n}(m_n - m) \xrightarrow{d} \mathcal{N}(0, \mathcal{K})$ and $\mathcal{K}_n \to \mathcal{K}$ a.s..

Table of Contents

We focus on random process $\{X_t\}_{t\in E}$, E is compact. We want to make the process $\{X_t\}_{t\in E}$ become a random element.

Let $(E, \mathcal{B}(E), \mu)$ another measure space, $\mu(E) < \infty$. We want to make the map $\mathcal{X} : \Omega \to L^2(E)$, $\mathcal{X}(\omega) = X_t(\omega)$ measurable.

Theorem

$(\Omega \times E, \mathcal{F} \times \mathcal{B}(E), P \times \mu)$ is a product measurable space. If $X : \Omega \times E \to R$ measurable and $X(\omega, \cdot) \in L^2(E)$, define $\mathcal{X} : \Omega \to L^2(E)$, $\mathcal{X}(\omega) = X(\omega, \cdot)$, then $\mathcal{X}$ is measurable.

Proof.

$f \in L^2(E)$, $\langle X(\omega, \cdot), f\rangle_2 = \int X(\omega, t)\ f(t) d\mu(t) \Rightarrow \langle X(\omega, \cdot), f\rangle_2$ measurable. $\square$

Theorem

If $X(*,t) : \Omega \to R$ is measurable $\forall t \in E$, and $X(\omega,\cdot) : E \to R$ is continuous $\forall \omega \in \Omega$, then X is jointly measurable.

Proof.

Let $X_n(\omega,t) = \sum_{n_k} I_{E_{n_k}}(t) X(\omega, t_{n_k})$, $E = \cup_{n_k} E_{n_k}$, $d(E_{n_k}) < 1/n$, then $\sup_{t\in E} ||X(\omega,t) - X_n(\omega,t)|| \to 0$. Claim that X_n is jointly measurable since $\forall B \in \mathcal{B}(R)$, $X_n^{-1}(B) = \cup_{n_k}(X^{-1}(B, t_{n_k}) \times E_{n_K}) \in \mathcal{F} \times \mathcal{B}(E)$. $\square$

One sufficient condition to have continuous modifications is Kolmogorov criterion: $\exists\ \alpha, \beta, C$ s.t. $E|X_{t_1} - X_{t_2}|^\alpha \le C|t_1 - t_2|^{1+\beta}$. It means that all the processes satisfied Kolmogorov criterion and take values in $L^2(E)$ can be viewed as a random element.

Let $m(t) := EX_t$ and $K(s,t) := Cov(X(s), X(t))$. Processes with well-defined m and K are referred to as second-order processes.

Definition

$\{X_t\}_{t\in E}$ is a mean-square continuous processes if it is a second-order processes and $\lim_{t_n \to t} ||X_t - X_{t_n}||_2 = 0$.

Theorem

$\{X_t\}_{t\in E}$ is mean-square $\Leftrightarrow$ m and K are continuous.

Proof.

"$\Leftarrow$": $E(X_t - X_{t_n})^2 = E(X_t - m(t) - (X_{t_n} - m(t_n)) + m(t) - m(t_n))^2 = K(t,t)$ $+K(t_n,t_n) - 2K(t,t_n) + (m(t) - m(t_n))^2$.

"$\Rightarrow$": $|EX_t - X_{t_n}| \leq (E(X_t - X_{t_n})^2)^{1/2} \to 0 \Rightarrow m(t_n) \to m(t)$. Let $m(t) = 0$, $\forall t \in E$, $|K(t,t) - K(t,t_n)| = Cov(X_t, X_t - X_{t_n}) \leq K^{1/2}(t,t)(E(X_t - X_{t_n})^2)^{1/2} \to 0$ $\Rightarrow K(t_n,t) \to K(t,t) \Rightarrow K(t_n,t_n) \to K(t,t) \Rightarrow K$ is jointly continuous. $\square$

$X : \Omega \times E \to R$ is jointly measurable, and $\{X(,t)\}_{t\in E}$ is mean-square continuous process. m and K are mean and covariance function of process.

$E\int X_t^2 d\mu(t) = \int EX_t^2 d\mu(t) < \infty \Rightarrow \int X^2(\omega,t)d\mu(t) < \infty$ a.s.. We can make a modification of X s.t. $X(\omega,\cdot) \in L^2(E)$, $\forall \omega \in \Omega$.

Then define $\mathcal{X} : \Omega \to L^2(E)$, $\mathcal{X}(\omega) = X_t$. Then $\mathcal{X}$ is a random element s.t. $E||\mathcal{X}||_2^2 < \infty$, let $\mathcal{K}$ be the covariance operator.

Property

$E\mathcal{X} = m$, $(\mathcal{K}f)(t) = \int K(t,s)f(s)d\mu(s)$.

Proof.

$E\langle \mathcal{X}, f\rangle_2 = E\int X(\omega,t)f(t)d\mu(t) = \int m(t)f(t)d\mu(t) = \langle m, f\rangle_2$, $\forall\ f \in L^2(E)$.
Let $m = 0$ and define $(Tf)(t) = \int K(t,s)\ f(s)d\mu(s)$, then $\langle Tf, g\rangle_2 = \int g(t)$ $\int K(t,s)f(s)d\mu(s)d\mu(t) = \int\int E(X_tX_s)f(s)g(t)d\mu(s)d\mu(t) = E\langle \mathcal{X}, f\rangle_2\langle \mathcal{X}, g\rangle_2$. □

$\{X_{i,s}\}$ mean-square continuous on compact E_i, $K_{ij}(s,t) = Cov(X_{i,s}, X_{j,t})$.
Then $\forall\ f_j \in L^2(E_j)$, $(\mathcal{K}_{ij}\ f_j)(t) = \int_E K_{ij}(t,s)f_j(s)d\mu_j(s)$.

For a generalized case, we can drop the condition: X is jointly measurable.

Theorem

If $m(t) = 0$, then $\exists$ random element $\mathcal{X}$ s.t. $(\mathcal{K}_{\mathcal{X}} f)(t) = \int K(t,s) f(s) d\mu(s)$.

To do more, we should firstly define the stochastic integration of $f \in L^2(E)$: $I_X(f) = \int_E X_t \ f(t) d\mu(t)$.

Firstly, $\forall n > 0$, $\exists$ some finite balls E_{n_k} s.t. $E = \cup_k E_{n_k}$ and $d(E_{n_k}) \leq 1/n$, take $t_k \in E_{n_k}$, define $I_{X,n}(f) = \sum_k X_{t_k} \int_{E_{n_k}} f(x) d\mu(x)$.

We note that $X_{t_k} \in L^2(\Omega) \Rightarrow I_{X,n}(f) \in L^2(\Omega)$.

Theorem

$\{I_{X,n}(f)\}$ *is Cauchy, we define* $\lim_n I_{X,n}(f) = I_X(f)$.

Proof.

$$
\begin{aligned}
&E\left(\sum_k X_{t_k} \int_{E_{n_k}} f(x)d\mu(x) - \sum_j X_{t_j} \int_{E_{m_j}} f(x)d\mu(x)\right)^2 \\
&= \sum_{k_1,k_2} K(t_{k_1}, t_{k_2}) \int_{E_{n_{k_1}}} f(x)d\mu(x) \int_{E_{n_{k_2}}} f(x)d\mu(x) \\
&\quad + \sum_{k,j} K(t_k, t_j) \int_{E_{n_k}} f(x)d\mu(x) \int_{E_{m_j}} f(x)d\mu(x) \\
&\quad -2 \sum_{j_1,j_2} K(t_{j_1}, t_{j_2}) \int_{E_{m_{j_1}}} f(x)d\mu(x) \int_{E_{m_{j_2}}} f(x)d\mu(x)
\end{aligned}
$$

And they all converge to $\int\int_{E\times E} K(u,v)f(u)f(v)d\mu(u)d\mu(v)$. □

If X is jointly measurable, $\langle X_t, f\rangle_2 = I_{X_t}(f)$, $\forall\, f \in L^2(E)$.

R is symmetric, positive and continuous kernel on $E \times E$, E is compact.

Theorem

$\mathcal{X}: \Omega \to H(R)$ measurable, define $X_t: \Omega \to \mathcal{R}$, $X_t(\omega) = \mathcal{X}(\omega)|_{\cdot=t}$ is a random process.

If $X(,t)$ is a stochastic process on E and $X(\omega,\cdot) \in H(R)$, define $\mathcal{X}: \Omega \to H(R)$, $\mathcal{X}(\omega) = X(\omega,\cdot)$, then $\mathcal{X}$ is measurable.

Proof.

$\mathcal{X}(\omega)|_{\cdot=t} = \langle \mathcal{X}(\omega), R(\cdot,t)\rangle \Rightarrow X_t = \langle \mathcal{X}, R(\cdot,t)\rangle$ is measurable.

$\langle X(\omega,\cdot), R(\cdot,t)\rangle = X(\omega,t) \Rightarrow \langle \mathcal{X}, R(\cdot,t)\rangle = X_t$ is measurable. $\square$

Theorem

$\mathcal{X}$ is random element valued in $H(K)$ s.t. $E||\mathcal{X}||^2 < \infty$, we have:
(a) $EX_t = \langle m, R(\cdot,t)\rangle$, $K(s,t) = Cov(X_s, X_t) = \langle \mathcal{K}R(\cdot,t), R(\cdot,s)\rangle$.
(b) X_t is a mean-square continuous process on E.
(c) $K \in H(R) \otimes H(R)$.

Proof.

$EX_t = E\mathcal{X}(\cdot)_{\cdot=t} = E\langle \mathcal{X}, R(\cdot,t)\rangle = \langle m, R(\cdot,t)\rangle$.
$Cov(X_s, X_t) = Cov(\langle \mathcal{X}, R(\cdot,s)\rangle, \langle \mathcal{X}, R(\cdot,t)\rangle) = \langle \mathcal{K}R(\cdot,t), R(\cdot,s)\rangle$.
Since $\mathcal{K} \in B_{HS}(H(R)) \Rightarrow K \in H(R) \otimes H(R)$. $\square$

One can prove that $H(K)$ must be finite dimensional since compact $\mathcal{K} = I$, which is shown by $\langle \mathcal{K}K(\cdot,t), K(\cdot,s)\rangle_K = \langle K(\cdot,t), K(\cdot,s)\rangle_K$.

Table of Contents

$1 \le i \le n$,$1 \le j \le r$, Let $\mathcal{X}(t)$ be a $L^2(E)$ valued process, which m and K are continuous. And $\mathcal{X}$ be a random element. We consider a model $Y_{ij} = \mathcal{X}_i(T_{ij}) + \varepsilon_{ij}$. $\mathcal{X}_i$, T_{ij} and ε_{ij} are iid.

Now we interesting in estimating $m = E\mathcal{X}$. A local smoothing based estimator of m is obtained by:

$$arg\min_{a_0,a_1} \sum_{i,j} W_h(T_{ij} - t)\ (Y_{ij} - a_0 - a_1(T_{ij} - t))^2$$

$W_h(\cdot) = h^{-1}W(\frac{\cdot}{h})$ and W is a symmetric probability density function on $[-1, 1]$, bounded variation and $\int u^2 W(u)du = C \neq 0$. Let $m_h(t) = \hat{a}_0$.

Under some regularity conditions and $E\sup_t |\mathcal{X}_i(t)|^q < \infty$, $q > 2$. One can show that if $h \to 0$, $n \to \infty$ in such a way that $(h^2 + h/r)^{-1}(\log n/n)^{1-2/q} \to 0$, then

$$\sup_{t\in[0,1]} |m_h(t) - m(t)| = O\left(h^2 + (\frac{(1+(hr)^{-1})\log n}{n})^{1/2}\right) \text{ a.s.}$$

Theorem

If r is bounded, $\sup_{t\in[0,1]} |m_h(t) - m(t)| = O(h^2 + (\frac{\log n}{nh})^{1/2})$ *a.s..*

If $r_n^{-1} \lesssim h \lesssim (\frac{\log n}{n})^{1/4}$, $\sup_{t\in[0,1]} |m_h(t) - m(t)| = O\left((\frac{\log n}{n})^{1/2}\right)$ *a.s..*

Similarly, one can approximate $K_{h_R}(s,t) = \hat{a}_0 - m_{h_m}(s)m_{h_m}(t)$:

$$\arg\min_{a_0,a_1,a_2} \frac{1}{nr(r-1)} \sum_i [\sum_{j \neq k} W_{h_R}(T_{ij}-s)W_{h_R}(T_{ik}-t)$$
$$(y_{ij}y_{ik} - a_0 - a_1(T_{ij}-s) - a_2(T_{ik}-t))^2]$$

Theorem

1. *If r is bounded and $h_R^2 \lesssim h_m \lesssim h_R$:*

$$\sup_{s,t\in[0,1]} |K_{h_R}(s,t) - K(s,t)| = O\left(h_R^2 + \left\{(\log n / (nh_R^2)\right\}^{1/2}\right) \qquad a.s.$$

2. $r_n^{-1} \lesssim h_m, h_R \lesssim (\log n/n)^{1/4}$:

$$\sup_{s,t\in[0,1]} |K_h(s,t) - K(s,t)| = O\left(\{\log n/n\}^{1/2}\right) \qquad a.s.$$

Now we assume that $T_{ij} \sim U(0,1)$ and $\mathcal{X}$ takes values in $W_q[0,1]$, which is equipped with $||f||_W^2 = ||f||_2^2 + ||f^{(q)}||_2^2$.

Recall COB of $W_q[0,1]$: $\{e_n\}$, $||\mathcal{X}||_W^2 = \sum_i (1+\gamma_i)\langle \mathcal{X}, e_i\rangle_2^2$. We will get a approximation of m by:

$$arg\min_{v\in W_q[0,1]} f_{rn,\eta}(v) = arg\min_{v\in W_q[0,1]} \frac{1}{nr}\sum_{i,j}\left(Y_{ij} - v\left(T_{ij}\right)\right)^2 + \eta\left\|v^{(q)}\right\|_2^2$$

Theorem

Let $\mathcal{P}(q, C_1)$ be the collection of probability measures for $\mathcal{X}$ s.t. $E||\mathcal{X}^{(q)}||_2^2 \le C_1$. Then $\exists\ C_2 > 0$ that depends on C_1 only, then

$$\lim\sup_n \ \sup_{P\in\mathcal{P}(q,C_1)} P(||\hat{m} - m||_2^2 > C_2((nr)^{-\frac{2q}{2q+1}} + n^{-1})) > 0$$

Theorem

If $\eta \asymp (nr)^{-\frac{2q}{2q+1}}$, *then* $||\hat{m} - m||_2^2 = O_p((nr)^{-\frac{2q}{2q+1}} + n^{-1})$.

One can approximate $\hat{K}(s,t) = \hat{v}(s,t) - \hat{m}(s)\hat{m}(t)$. Let $H = W_q[0,1] \otimes W_q[0,1]$:

$$\arg\min_{v \in H} \frac{1}{nr(r-1)} \sum_i \sum_{j \neq k} (y_{ij} y_{ik} - v(t_{ij}, t_{ik}))^2 + \eta ||v||_H$$

Theorem

If $\eta \asymp (\frac{\log n}{rn})^{\frac{2q}{2q+1}}$ *and more conditions, then* $||\hat{K} - K||_2^2 = O_p((\frac{\log n}{rn})^{\frac{2q}{2q+1}} + n^{-1})$.