STAT 620: Asymptotic Statistics **Spring 2022**

Lecture: Mar 1

Lecturer: Xianyang Zhang

1 Projection

1.1 Hilbert space

A vector space $\mathcal{H}$ is a **Hilbert space** if it is a complete normed vector space and have the inner product. $\Big\langle \cdot, \cdot \Big\rangle : \mathcal{H} \times \mathcal{H} \to \mathbb{R}$ which is linear in both arguments and $\Big\langle u, u \Big\rangle = \|u\|^2$

Example 1 $\mathbb{R}^n$ is a Hilbert space with $\Big\langle X, Y \Big\rangle = X^\top Y$

Example 2 Suppose P is a measure of $\mathcal{X}$. Then

$$L^2(P) = \left\{ f : \mathcal{X} \to \mathbb{R}, \int f^2(x)\, dP(x) < \infty \right\}$$

is a Hilbert space with

$$\Big\langle f, g \Big\rangle = \int f(x)\, g(x)\, dP(x) \quad \text{and} \quad \|f\| = \left(\int f^2(x)\, dP(x) \right)^{1/2}.$$

1.2 Projection

Let $\mathcal{S} \subseteq \mathcal{H}$ be a closed linear subspace of $\mathcal{H}$, i.e. $\mathcal{S}$ contains 0 and all linear combinations of elements in itself. For any $v \in \mathcal{H}$, we define the projection of v onto $\mathcal{S}$ as

$$\pi_{\mathcal{S}}(v) = \operatorname*{argmin}_{s \in \mathcal{S}} \|s - v\|^2.$$

1.3 Theorem

The projection $\pi_{\mathcal{S}}(v)$ exists, is unique and is uniquely determined by the equality:

$$\Big\langle v - \pi_{\mathcal{S}}(v), a \Big\rangle = 0$$

for all $a \in \mathcal{S}$.

Proof: (1) We first show that if the orthogonality condition is satisfied, then $\pi_{\mathcal{S}}(v)$ must be the projection of v onto $\mathcal{S}$. Consider any $s \in \mathcal{S}$. We have

$$\|s - v\|^2 = \|s - \pi_{\mathcal{S}}(v) + \pi_{\mathcal{S}}(v) - v\|^2 = \|s - \pi_{\mathcal{S}}(v)\|^2 + \|\pi_{\mathcal{S}}(v) - v\|^2 + 2\Big\langle s - \pi_{\mathcal{S}}(v), \pi_{\mathcal{S}}(v) - v \Big\rangle \geq \|\pi_{\mathcal{S}}(v) - v\|^2,$$

where we have used the fact that $\Big\langle s - \pi_{\mathcal{S}}(v), \pi_{\mathcal{S}}(v) - v \Big\rangle = 0$. The equality holds when $\|s - \pi_{\mathcal{S}}(v)\|^2 = 0$, that is $s = \pi_{\mathcal{S}}(v)$.

(2) Next we show the projection $\pi_{\mathcal{S}}(v)$ satisfies the orthogonality condition. For any $a \in \mathcal{S}$ and $c \in \mathbb{R}$, we have

$$\|v - \pi_{\mathcal{S}}(v) - ca\|^2 - \|v - \pi_{\mathcal{S}}(v)\|^2 = c^2\|a\|^2 - \ 2c \left\langle a, v - \pi_{\mathcal{S}}(v) \right\rangle.$$

In order to have the above $\geq\ 0\ \forall c \in \mathbb{R}$, we must have

$$\left\langle a, v - \pi_{\mathcal{S}}(v) \right\rangle \ = \ 0.$$

2 Conditional Expectation

Define

$\mathcal{S} = \Big\{$linear span of $g(Y)$ for all measurable functions g and some random variable Y with $g(Y) \in L^2(P)\Big\}$,

where $Y \sim P$.

2.1 Definition

Suppose X and Y are random variables. We define the conditional expectation of X given Y, $\mathbb{E}[X \mid Y]$ as the projection of X onto $\mathcal{S}$, that is $\mathbb{E}[X \mid Y]$ is the unique function of Y such that

$$\mathbb{E}[\{X - E(X \mid Y)\}\, g(Y)] \ = \ 0,\ \ \forall\, g \in \mathcal{S}.$$

2.2 Properties

(1) $\mathbb{E}[X] = \mathbb{E}[\mathbb{E}(X \mid Y)]$. It follows directly by choosing $g(Y) = 1$.

(2) $\mathbb{E}[X\, g(Y) \mid Y] = g(Y)\, \mathbb{E}[X \mid Y]$.

2.3 Proposition

Let T_n be a sequence of random variables and S_n be sequence of subspaces of $L^2(P_n)$. Let $\hat{T}_n = \pi_{S_n}(T_n)$. Also let, $\sigma^2(X) = \mathrm{var}(X)$. If $\frac{\sigma^2(T_n)}{\sigma^2(\hat{T}_n)} \to 1$ as $n \to \infty$, then

$$\frac{T_n - \mathbb{E}(T_n)}{\sigma(T_n)} - \frac{\hat{T}_n - \mathbb{E}(\hat{T}_n)}{\sigma(\hat{T}_n)} \xrightarrow{P} 0.$$

Proof: Note that $\mathbb{E}[\{T_n - \pi_{S_n}(T_n)\}\, S] \ = \ 0\,, \forall\, S \in S_n$. Let

$$A_n = \frac{T_n - \mathbb{E}(T_n)}{\sigma(T_n)} - \frac{\hat{T}_n - \mathbb{E}(\hat{T}_n)}{\sigma(\hat{T}_n)}.$$

Then $\mathbb{E}(A_n) = 0$. Thus we just need to show that

$$\mathrm{var}(A_n) \to 0.$$

Toward this end, we note that

$$\begin{aligned}
\mathrm{var}(A_n) =& \mathrm{var}\left(\frac{T_n - \mathbb{E}(T_n)}{\sigma(T_n)}\right) + \mathrm{var}\left(\frac{\hat{T}_n - \mathbb{E}(\hat{T}_n)}{\sigma(\hat{T}_n)}\right) - 2\,\mathrm{cov}\left(\frac{T_n - \mathbb{E}(T_n)}{\sigma(T_n)}, \frac{\hat{T}_n - \mathbb{E}(\hat{T}_n)}{\sigma(\hat{T}_n)}\right) \\
=& 1 + 1 - \frac{2}{\sigma(T_n)\,\sigma(\hat{T}_n)}\,\mathrm{cov}(T_n, \hat{T}_n).
\end{aligned}$$

Noting that $\mathbb{E}(T_n) = \mathbb{E}(\hat{T}_n)$, we have

$$\operatorname{cov}(T_n, \hat{T}_n) = \mathbb{E}[T_n \hat{T}_n] - \mathbb{E}[T_n]\,\mathbb{E}[\hat{T}_n] = \mathbb{E}[(T_n - \hat{T}_n + \hat{T}_n)\,\hat{T}_n] - \mathbb{E}^2(\hat{T}_n) = \mathbb{E}(\hat{T}_n^2) - \mathbb{E}^2(\hat{T}_n) = \sigma^2(\hat{T}_n).$$

Then

$$\operatorname{var}(A_n) = 2 - 2\,\frac{\sigma(\hat{T}_n)}{\sigma(T_n)} \to 0 \text{ as } n \to \infty.$$

2.4 Hajek Projection

Let $X_1, X_2, \ldots, X_n$ be independent. Let

$$S = \left\{ \sum_{i=1}^{n} g_i(X_i) : g_i(X_i) \in L^2(P) \right\}.$$

If $\mathbb{E}(T^2) < \infty$, then the projection $\hat{T}$ of T onto S is given by

$$\hat{T} = \sum_{i=1}^{n} \mathbb{E}[T \mid X_i] - (n-1)\mathbb{E}[T].$$

Proof: Note that

$$\mathbb{E}[\mathbb{E}(T \mid X_i)|X_j] = \begin{cases} \mathbb{E}(T \mid X_i) & \text{if } i = j, \\ \mathbb{E}(T) & \text{if } i \neq j. \end{cases}$$

Now,

$$\mathbb{E}[\hat{T} \mid X_j] = \sum_{i=1}^{n} \mathbb{E}[\mathbb{E}(T \mid X_i) \mid X_j] - (n-1)\,\mathbb{E}[T] = (n-1)\,\mathbb{E}[T] + \mathbb{E}[T \mid X_j] - (n-1)\,\mathbb{E}[T] = \mathbb{E}[T \mid X_j].$$

Then $\mathbb{E}[(T - \hat{T}) \mid X_j] = 0$. So

$$\mathbb{E}[(T - \hat{T})\, g_j(X_j)] = \mathbb{E}[\mathbb{E}\{(T - \hat{T})\, g_j(X_j) \mid X_j\}] = \mathbb{E}[g_j(X_j)\,\mathbb{E}\{(T - \hat{T}) \mid X_j\}] = 0.$$

Hence

$$\mathbb{E}\Big[(T - \hat{T}) \sum_{j=1}^{n} g_j(X_j)\Big] = 0,$$

which completes the proof.