Conditional Expectation

$(\Omega, \mathcal{S}, P)$ probability space. $\mathcal{G} \subseteq \mathcal{F}$ a sub-$\sigma$-algebra.

$X$ a r.v. st. $X \in L^1(\Omega, \mathcal{G}, P)$

**Def** $Z = E[X|\mathcal{G}]$ if

i) $Z$ is $\mathcal{G}$-measurable

ii) $\int_A Z \, dP = \int_A X \, dP$ for all $A \in \mathcal{G}$

**Prop 1**

$Z = E[X|\mathcal{G}]$. TFAE: $Z$ is $\mathcal{G}$ measurable

i) $Z = E[X|\mathcal{G}]$

ii) $\int_A Z \, dP = \int_A X \, dP$

iii) $\int_{\Omega} YZ \, dP = \int_{\Omega} YX \, dP$ for all $Y$ odd $\mathcal{G}$-measurable

iv) $Z \in L^1(\Omega, \mathcal{G}, P)$ and $\int_A Z \, dP = \int_A X \, dP$

$\forall A \in \mathcal{G}$. $\mathcal{G}$-system containing $\Omega$ st. $\mathcal{G}(\Omega) = \mathcal{G}$

Markov Processes

**Def** Let $(E, \mathcal{B})$ be a measurable space. Let $(\Omega, \mathcal{F}, P)$ a probability space, and $I \subseteq \mathbb{R}$.

$X = (X_t, t \in I)$ is called a **stochastic process** with state space $(E, \mathcal{B})$ if $X_t: \Omega \to E$ is $\mathcal{F}/\mathcal{B}$-measurable $\forall t \in I$, i.e.

$X_t(\Omega) \in \mathcal{F}$ for all $B \in \mathcal{B}$
Notion: index set can be continuous, discrete, etc.

Can conceive of \( X \) as a very large family of random variables

\[
\mathbb{E} \left[ \mathcal{I} = \{ \omega : \omega = (x_t^i : t \in I), x_t^i \in E \} \right]
\]

We have a projection map \( \pi_t : E^I \rightarrow E \), \( \pi_t(\omega) = x^t \).

Fix \( \omega \), \( t \), \( x_t : E^I \rightarrow E \) \( \omega \), \( x_t(\omega) = x^t \). The coordinate map \( \pi_t(x) = x^t \).

\[
\mathcal{B}^I = \sigma(\left\{ \pi_t : t \in I \right\}) \) the sigma algebra on \( E^I \).
\]

Then \( X : \Omega \rightarrow E^I \) \( \mathcal{F} / \mathcal{B}^I \) measurable. Define \( \mu^I_B = \mathcal{P}(X^{-1}(B)), B \in \mathcal{B}^I \).

Then \( \mu^I \) is a probability measure on \( (E^I, \mathcal{B}^I) \). \( \mu^I \) is called the probability distribution of \( X \).

Def: Two stochastic processes \( X, Y \) have the same distribution if \( \mu^I_X = \mu^I_Y \).

Prop 2

Let \( X, Y \) be stochastic processes. Then \( \mu^I_X = \mu^I_Y \) if and only if \( \mu^I_{X_{t_1}} \cdots X_{t_n} = \mu^I_{Y_{t_1}} \cdots Y_{t_n} \) for all \( t_1, \ldots, t_n \in I \) for all \( \mathcal{B}^I \).

\[
\mu^I_{X_{t_1}} \cdots X_{t_n}(B_1 \times \cdots \times B_n) = \mathcal{P}(X_{t_1}^{-1}(B_1) \cap \cdots \cap X_{t_n}^{-1}(B_n))
\]

\[
= \mu^I_X \left( \left\{ x \in E^I : x_{t_i} \in B_i, \ldots, x_{t_n} \in B_n \right\} \right)
\]

\[
= \mu^I_Y \left( \left\{ x \in E^I : x_{t_i} \in B_i, \ldots, x_{t_n} \in B_n \right\} \right)
\]

\[
= \mu^I_{X_{t_1}} \cdots X_{t_n}(B_1 \times \cdots \times B_n)
\]

(\( \Leftarrow \)) By hypothesis \( \mu^I_X = \mu^I_Y \) on finite dimensional cylinder sets. Finite cylinder sets form a \( \pi \)-system containing \( \mathcal{B}^I \). Thus \( \mu^I_X = \mu^I_Y \) by \( \pi \)-\( \lambda \) theorem.
(Ω, ℱ, ℙ) ℱ₀, ℱ₁, ..., ℱₙ an σ-algebra of Ω
Def: ℱ₁, ..., ℱₙ are conditionally independent given ℱ₀ if
\[ E[Y₁, ..., Yₙ | ℱ₀] = \prod_{i=1}^{n} E[Yᵢ | ℱ₀] \quad \text{for all } Yᵢ ∈ ℱᵢ. \]

Def: Random variables X₁, ..., Xₙ are conditionally independent given ℱ₀ if
\( E(Xₙ₁) \) and \( E(Xₙ₂) \) are conditionally independent given ℱ₀

\( X₁, ..., Xₙ \) conditionally independent given \( ℱ₀ \) if
\( E(Xₙ₁) \) and \( E(Xₙ₂) \) are conditionally independent given \( ℱ₀ \)

Notation: \( ℱ₁ \) and \( ℱ₂ \) are conditionally independent given \( ℱ₀ \)
\[ ℱ₁ ⊥⊥ ℱ₂ \quad \text{for all } X₁, X₂ \]

Thm: \( ℱ₁ ⊥⊥ ℱ₂ \) if and only if
\[ E[Y₁ | σ(F₀, F₂)] = E[Y₁ | F₀] \]

for all bounded \( ℱ₀ \)-measurable \( Y₁ \)

If \( \Rightarrow \) Since \( E[Y₁ | F₀] \) is \( ℱ₀ \)-measurable \( \subseteq σ(F₀, F₂) \)-measurable.

Suffice to show \( ∀ A ∈ σ(F₀, F₂) \):
\[ \int_A E[Y₁ | F₀] dP = \int_A Y₁ dP \]

Let \( C = \{ A ∈ σ(F₀, F₂) : (*) \text{ holds} \} \). Then \( C \) is a σ-algebra of \( Ω \).

Let \( D = \{ A_n ∩ A_l | A_n ∈ F₀, A_l ∈ F₂ \} \). Then \( C = \bigwedge D \) a σ-algebra.

\[ \int A_n l_{A_l} E[Y₁ | F₀] dP = \int A_n l_{A_l} dP \quad (*) \]

Things got somewhat confused here
\[ E[Y, Y_2 \mid \mathcal{F}_0] = E[Y \mid \mathcal{F}_0] E[Y_2 \mid \mathcal{F}_0] \]

\[ \int E[Y \mid \mathcal{A}_2 \mid \mathcal{F}_0] = \int E[E[Y \mid \mathcal{F}_0] \mid \mathcal{A}_2 \mid \mathcal{F}_0] \]

\[ \int_{\mathcal{A}_0} Y_1 \mid \mathcal{A}_2 = \int_{\mathcal{A}_0} E[Y_1 \mid \mathcal{F}_0] \mid \mathcal{A}_2 \quad d\mathcal{P} \]

\[ \int_{\mathcal{A}_0} Y_1 \mid \mathcal{A}_2 = \int_{\mathcal{A}_0} E[Y_1 \mid \mathcal{F}_0] \mid \mathcal{A}_2 \quad d\mathcal{P} \]

\[ \quad (\iff) \quad E[Y, Y_2 \mid \mathcal{F}_0] \]

\[ = E[E[Y, Y_2 \mid \mathcal{F}_0] \mid \mathcal{F}_0] \]

\[ = E[y_2 E[Y \mid \mathcal{F}_0] \mid \mathcal{F}_0] \]

\[ = E[y_2 E[Y \mid \mathcal{F}_0] \mid \mathcal{F}_0] \]

\[ = E[y_2 E[Y \mid \mathcal{F}_0] \mid \mathcal{F}_0] \]

\[ = E[Y \mid \mathcal{F}_0] E[Y_2 \mid \mathcal{F}_0] \]

**Theorem 4**

1) \( \mathcal{F}_1 \perp \perp \mathcal{F}_2 \)

2) \( \mathcal{F}_1 \perp \mathcal{F}_0 \mid \mathcal{F}_0 = \mathcal{G} \) \( \mathcal{F}_0, \mathcal{F}_2 \)

3) \( \mathcal{F}_0 \perp \mathcal{F}_0 \)

Based on 3, \( \mathcal{F}_0 \) is a Hausman
Markov Processes

Def: A stochastic process \( X_t = (X_t, t \geq 0) \) is called a Markov process if
\[
\forall t \geq 0 \quad \sigma(X_t) \subseteq \sigma(X_s, s \leq t) \quad \sigma(X_s, s > t)
\]
\[
F_t \perp \perp F_s
\]

This is the so-called Markov property.

Can be defined for \( t \) in high-dimensional spaces by independence of space divided \( F_s \) on \( F_t \) for \( s \leq t \).

Props:
1. \( F_s \perp \perp F_t \)
2. \( F_s \perp \perp F_s \)
3. \( F_s \perp \perp F_t \)
4. \( F_s \perp \perp F_t \)
5. \( E[Z | F_{s+t}] = E[Z | F_s] + \text{all } Z \text{ bounded, } F_{s+t} \text{ measurable} \)
6. \( E[Z | F_{s+t}] = \chi_{F_s} E[Z | F_s] \quad \forall \chi \in F_{s+t} \quad Z \in F_{s+t} \)

Next the transition function \( \pi \) for Markov process.

Stationary Markov process

Markov chains

Quantum Algorithms

Start reading Markov chain chapter.