Lecture 13

Conditional Expectations

Let $A \in \mathcal{A}$ with $P(A) > 0$.

Define: $P_A(B) = \frac{P(B \cap A)}{P(A)}$

P_A is a probability measure on $\mathcal{A}$. P_A is called the conditional probability relative to A. Let Y be a random variable.

$$E_A(Y) = \int_\Omega Y(\omega) P_A(d\omega) = \frac{1}{P(A)} \int_\Omega Y(\omega) P(d\omega)$$

is called the conditional expectation. If $P(A) = 0$ we define $P_A(B) = 0 \ \forall B \in \mathcal{A}$.

Next $\{A_n : n \geq 1\}$ be a countable measurable partition of Ω. i.e.

$$\Omega = \bigcup_{n=1}^{\infty} A_n, \ A_n \in \mathcal{A} \ A_n \cap A_m = \phi. \ m \neq n.$$

$$P(B) = \sum_{n=1}^{\infty} P(A_n \cap B) = \sum_{n=1}^{\infty} P(A_n) P_{A_n}(B)$$

$$E(Y) = \sum_{n=1}^{\infty} \int_{A_n} Y(\omega) P(d\omega) = \sum_{n=1}^{\infty} P(A_n) E_{A_n}(Y)$$

Let $\mathcal{G}$ be a Borel field generated by $\{A_n\}$.
If Y is an integrable random variable.

$$E_{\mathcal{G}}(Y) = \sum_{n=1}^{\infty} E_{A_n}(Y) I_{A_n}$$

I_{A_n} is the indicator of A_n.

$E_{\mathcal{G}}(Y)$ is a discrete random variable that assumes the value $E_{A_n}(Y)$ on the set A_n.

$$\begin{aligned} E(Y) &= \sum_{n=1}^{\infty} \int_{A_n} Y(\omega) P(d\omega) \\ &= \sum_{n=1}^{\infty} \int_{A_n} E_{\mathcal{G}}(Y) P(d\omega) \\ &= \int_\Omega E_{\mathcal{G}}(Y) P(d\omega) \end{aligned}$$

Also if $A \in \mathcal{G}$.

$$\int_A Y dP = \int_A E_{\mathcal{G}}(Y) dP$$

Notice that $E_{\mathcal{G}}(Y)$ is a random variable over $\mathcal{G}$.
Suppose $\exists$ two functions φ_1 & φ_2 over $\mathcal{G}$ $\ni$

$$\int_A Y dP = \int_A \varphi_i dP \ i = 1, 2. \ \forall A \in \mathcal{G}.$$

Let $A = \{\omega : \varphi_1(\omega) > \varphi_2(\omega)\} \Rightarrow A \in \mathcal{G}$.

Also

$$\int_A \varphi_1 dP = \int_A Y dP = \int_A \varphi_2 dP$$
$$\Rightarrow \int_A (\varphi_1 - \varphi_2) dP = 0$$
$$\Rightarrow P(A) = 0$$

Similarly if $A = \{\omega : \varphi_2(\omega) > \varphi_1(\omega)\}$.

$\Rightarrow P(A) = 0$
$\Rightarrow \varphi_1 = \varphi_2$ almost surely.
$\Rightarrow E_{\mathcal{G}}(Y)$ is unique up to a set of measure 0.
$\Rightarrow E_{\mathcal{G}}(Y)$ is really an equivalence class of random variables.

We write $E(Y|\mathcal{G}) = E_{\mathcal{G}}(Y)$ and use either notation.

Theorem: If $E[|Y|] < \infty$ and $\mathcal{G}$ is a Borel subfield of $\mathcal{A}$, then $\exists$ an equivalence class of integrable random variables $E(Y|\mathcal{G})$ over $\mathcal{G}$ such that

$$\int_A Y dP = \int_A E(Y|\mathcal{G}) dP, \ \forall A \in \mathcal{G}.$$

Note that this is true for every Borel subfield of $\mathcal{A}$ not just one generated by a partition.

Definition: Given an integrable random variable Y and a Borel subfield $\mathcal{G}$, the conditional expectation $E(Y|\mathcal{G})$ of Y relative to $\mathcal{G}$ is any one of the equivalence classes of random variables on Ω satisfying

i) it belongs to $\mathcal{G}$ (i.e. defined over $\mathcal{G}$)
ii) it has the same integral as Y over any set in $\mathcal{G}$.

If $Y = I_B$, $B \in \mathcal{A}$

Then define $P(B|\mathcal{G}) = E[I_B|\mathcal{G}]$.
is the conditional probability of B relative to $\mathcal{G}$.

Note that $P(B|\mathcal{G})$ is a random variable.

$$P(B \cap A) = \int_A P(B|\mathcal{G})dP, \forall A \in \mathcal{G}.$$

Note that

$$\int_A [Y - E[Y|\mathcal{G}]]dP = 0, \forall A \in \mathcal{G}.$$
$$\Rightarrow E\{[Y - E[Y|\mathcal{G}]]Z\} = 0, \forall \text{ bounded } Z \in \mathcal{G}.$$

$Z \in \mathcal{G}$ means Z is a random variable over $\mathcal{G}$, i.e. Z is a measurable function defined over $\mathcal{G}$; measurable w.r.t. $\mathcal{G}$.

$\Rightarrow Y = Y' + Y''$
where $Y' = E(Y|\mathcal{G})$ and $Y'' \perp \mathcal{G}$.

where $Y'' \perp \mathcal{G}$ means $E(Y''Z) = 0$ $\forall$ bounded $Z \in \mathcal{G}$.

Y' is the projection of Y on $\mathcal{G}$.
and Y'' is its orthogonal complement.

For the Borel field $\mathcal{A}\{X\}$ generated by the random variable X,

we write

$$E(Y|X) = E(Y|\mathcal{A}(X))$$

Similarly we can define $E(Y|X_1, \cdots, X_n)$

Theorem: One version of the conditional expectation $E(Y|X) = \varphi(X)$, where φ is a Borel measurable function on $\mathbb{R}$. Moreover if

$$\lambda(B) = \int_{X^{-1}(B)} Y dP, \forall B \in \mathcal{B}(\mathbb{R})$$

and let ν be the probability measure on $\mathbb{R}$ induced by X, then $\varphi = \frac{d\lambda}{d\nu}$, the so-called Radon-Nikodym derivative.

Theorem: Let Y and YZ be integrable random variables and $Z \in \mathcal{G}$, then

$$E[YZ|\mathcal{G}] = ZE[Y|\mathcal{G}] \text{ a.s.}$$

Properties:

i. If $X \in \mathcal{G}$, then $E[X|\mathcal{G}] = X$ a.s.

ii. $E[X_1 + X_2|\mathcal{G}] = E[X_1|\mathcal{G}] + E[X_2|\mathcal{G}]$

iii. If $X_1 \leq X_2$, then $E[X_1|\mathcal{G}] \leq E[X_2|\mathcal{G}]$

iv. $|E(X|\mathcal{G})| \leq E(|X||\mathcal{G})$

v. If $X_n \uparrow X \Rightarrow E[X_n|\mathcal{G}] \uparrow E[X|\mathcal{G}]$

vi. If $X_n \downarrow X \Rightarrow E[X_n|\mathcal{G}] \downarrow E[X|\mathcal{G}]$

vii. If $|X_n| \leq Y$ and $E[Y] < \infty$ and $X_n \to X$, then

$$E[X_n|\mathcal{G}] \to E[X|\mathcal{G}]$$

viii. **Cauchy-Schwartz**

$$[E(|XY|\,|\mathcal{G})]^2 \leq E(X^2|\mathcal{G})E(Y^2|\mathcal{G}).$$

ix. **Jensen's Inequality**
If φ is a convex function on $\mathbb{R}$ and X and $\varphi(X)$ are integrable random variables, then

$$\varphi[E(X|\mathcal{G})] \leq E(\varphi(X)|\mathcal{G})$$

Theorem: If Y is integrable and $\mathcal{G}_1 \subset \mathcal{G}_2$ are fields, then

$$E(Y|\mathcal{G}_1) = E(Y|\mathcal{G}_2) \text{ iff } E(Y|\mathcal{G}_2) \in \mathcal{G}_1$$

and

$$E_{\mathcal{G}_1}(E_{\mathcal{G}_2}(Y)) = E_{\mathcal{G}_1}(Y) = E_{\mathcal{G}_2}(E_{\mathcal{G}_1}(Y))$$

In particular

$$\begin{gathered} E\{E(Y|X_1, X_2)|X_1\} \\ = E(Y|X_1) = E\{E(Y|X_1)|X_1, X_2\} \end{gathered}$$

Martingales
Let $\{x_n\}$ denote independent random variables with zero mean and let

$$X_n = \sum_{j=1}^{n} x_j$$

$$E(X_{n+1}|x_1, \cdots, x_n) = E(X_n + x_{n+1}|x_1, \cdots, x_n)$$

$$
\begin{aligned}
&= E(X_n|x_1, \cdots, x_n) + E(x_{n+1}|x_1, \cdots, x_n) \\
&= E(X_n|X_1, \cdots, X_n) + E(x_{n+1}) \\
&= X_n + 0 = X_n. \\
&E(X_{n+1}|x_1, \cdots, x_n) = X_n
\end{aligned}
$$

Definition: The sequence of random variables and Borel fields $\{X_n, \mathcal{A}_n\}$ is called a <u>martingale</u> iff for each n

$$(a)\ \mathcal{A}_n \subset \mathcal{A}_{n+1} \text{ and } X_n \in \mathcal{A}_n$$

$$(b)\ E[|X_n|] < \infty$$

$$(c)\ X_n = E[X_{n+1}|\mathcal{A}_n] \text{ a.s.}$$

It is called a supermartingale iff

$$(c)\ X_n \geq E[X_{n+1}|\mathcal{A}_n] \text{ a.s.}$$

and a submartingale iff

$$(c)\ X_n \leq E[X_{n+1}|\mathcal{A}_n] \text{ a.s.}$$

<u>smartingale</u> refers to all three varieties

Condition (a) is referred to as $\{X_n\}$ is <u>adapted</u> to $\{\mathcal{A}_n\}$

Condition (b) says that all the random variables are integrable.

Condition (c) implies that if $n < m$, then $X_n = E[X_m|\mathcal{A}_n]$

This follows because

$$
\begin{aligned}
E[X_m|\mathcal{A}_n] &= E[E[X_m|\mathcal{A}_{m-1}]|\mathcal{A}_n] \\
&= E[X_{m-1}|\mathcal{A}_n]
\end{aligned}
$$

and by induction

An equivalent statement is for $n \leq m$ and each $A \in \mathcal{A}$

$$\int_A X_n dP = \int_A X_m dP.$$

Theorem: Let $\{X_n, \mathcal{A}_n\}$ be a submartingale and φ be an increasing convex function defined on $\mathbb{R}$. If $\varphi(X_n)$ is integrable for every n, then

$$\{\varphi(X_n), \mathcal{A}_n\} \text{ is also a submartingale.}$$

Proof: Use Jensen's Inequality.

Corollary: If $\{X_n, \mathcal{A}_n\}$ is a submartingale, then so is $\{X_n^+, \mathcal{A}_n^+\}$

Corollary: If $\{X_n, \mathcal{A}_n\}$ is a martingale, then $\{|X_n|, \mathcal{A}_n\}$ is a submartingale.

Corollary: If $\{X_n, \mathcal{A}_n\}$ is a supermartingale, then so is $\{X_n \wedge A, \mathcal{A}\}$ where A is any constant.

recall $a \wedge b = min\{a, b\}$.

Let $\{X_n, \mathcal{A}_n\}$ be a martingale, suppose $\varphi(x)$ is any of the functions

$$\varphi(x) = |x|,\ x^2,\ x^+,\ x^-,\ x \vee A$$

Then

$$\{|X_n|\},\ \{X_n^2\},\ \{X_n^+\},\ \{X_n^-\},\ \{X_n \vee A\}$$

are all submartingales.

Definition: $\{Z_n\}$ is an <u>increasing process</u> iff it satisfies

$$i)\ Z_1 = 0, Z_n \le Z_{n+1},\ n \ge 1$$

$$ii)\ E[Z_n] < \infty \text{ for each } n.$$

Theorem: Any submartingale $\{X_n, \mathcal{A}_n\}$ can be written as

$$X_n = Y_n + Z_n$$

where $\{Y_n, \mathcal{A}_n\}$ is a martingale and Z_n is an increasing process.

Corollary: Any supermartingale $\{X_n, \mathcal{A}_n\}$ can be written as

$$X_n = Y_n - Z_n$$

where $\{Y_n, \mathcal{A}_n\}$ is a martingale and Z_n is an increasing process.