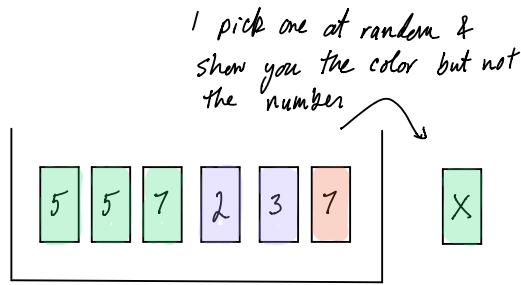


Lecture 20: Conditional expected value with respect to a sub- σ -field

(1)

Let's start with a motivation.

Consider a box with numbered tickets which are colored.



Before you know the color your best guess for X is

$$E(X) = \int_X dP = \left\{ \begin{array}{l} \text{average of the} \\ \text{ticket numbers} \\ \text{in the box.} \end{array} \right.$$

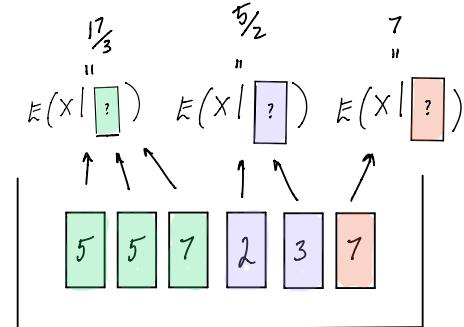
After I tell you the color is green your new best guess for X is

$$\begin{aligned} E(X | ?) &= \left\{ \begin{array}{l} \text{average green} \\ \text{ticket number} \end{array} \right. \\ &= \frac{5+5+7}{3} = \frac{17}{3} \end{aligned}$$

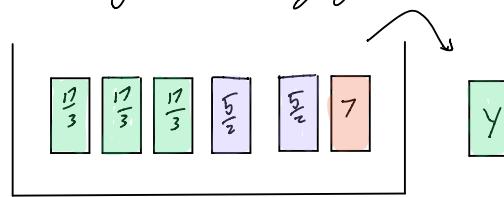
If you wanted to automate this prediction you could pre-compute

$$E(X | ?) \quad E(X | ?) \quad E(X | ?)$$

This can be thought of as a map from ticket to prediction value



or as a granular smoothing of X



where $Y = E(X | \text{color})$.

Notice two key facts about $Y = E(X | \text{color})$.

- (i) The collection of events that we can place bets on for Y is less than for X
- (ii) If A is an event that corresponds to a bet on Y , i.e. $A = \{Y = \frac{22}{3}\} \cup \{Y = 7\}$, then

$$\underbrace{\int_A X dP}_{\frac{5+5+7+7}{6}} = \underbrace{\int_A Y dP}_{\frac{17}{3} \cdot \frac{1}{2} + 7 \cdot \frac{1}{6}}$$

Now make the correspondence

(3)

\mathcal{J} = the collection of tickets

\mathcal{F} = the possible bets on all tickets

\mathcal{Q} = the possible bets on color

For $w \in \mathcal{J}$, $X(w) = \text{ticket \#}$

$E(X|\mathcal{Q}) = Y$ maps $w \in \mathcal{J} \mapsto$ are X & tickets
with the same
color as w

and we have

$E(X|\mathcal{Q})$ is \mathcal{Q} -measurable and

$$\int_A X dP = \int_A E(X|\mathcal{Q}) dP \quad \forall A \in \mathcal{Q}.$$

Theorem: (existence of $E(X|\mathcal{Q})$)

Let $(\mathcal{J}, \mathcal{F}, P)$ be a probability space and $X \in Q(\mathcal{J}, \mathcal{F}, P)$ be a possibly extended r.v.

If $\mathcal{Q} \subset \mathcal{F}$ is a σ -field then \exists a P -unique extended r.v. $E(X|\mathcal{Q}) \in Q(\mathcal{J}, \mathcal{F}, P)$ such that

- (i) $E(X|\mathcal{Q})$ is \mathcal{Q} -measurable \Leftarrow more granular
- (ii) $\int_A X dP = \int_A E(X|\mathcal{Q}) dP \quad \forall A \in \mathcal{Q}$

Proof:

Start by assuming $X \geq 0$.

Let $v(\cdot) = \int \cdot dP$ be a measure on $(\mathcal{J}, \mathcal{Q})$.

and $\bar{P}(\cdot) = P(\cdot)$ but only defined over $(\mathcal{J}, \mathcal{Q})$.

Now to show (ii) we want a $E(X|\mathcal{Q}) @ \mathcal{Q}$

$$v(A) = \int_A E(X|\mathcal{Q}) d\bar{P}$$

↑
replaced P with \bar{P} since
 $E(X|\mathcal{Q})$ is supposed to be $@ \mathcal{Q}$

So look for this as $\frac{dv}{d\bar{P}}$

Notice $v \ll \bar{P}$ since

$$\bar{P}(A) = 0, A \in \mathcal{Q} \Rightarrow \int_A X dP = 0 \quad P\text{-a.e.}$$

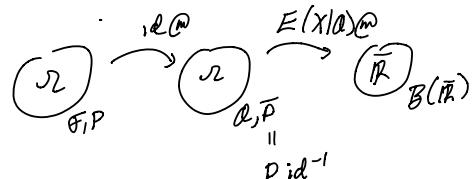
$$\Rightarrow v(A) = \int_A X dP = 0$$

\therefore By The Radon-Nikodym Thm \exists a unique $dV/d\bar{P} \in \mathcal{H}(\mathcal{J}, \mathcal{Q})$ s.t. $\forall A \in \mathcal{Q}$

$$\int_A \frac{dV}{d\bar{P}} d\bar{P} = V(A) = \int_A X dP$$

$$\text{So set } E(X|\mathcal{Q}) := \frac{dV}{d\bar{P}} \in \mathcal{H}(\mathcal{J}, \mathcal{Q})$$

By the change of variables Thm



we have

$$E(X|\mathcal{Q}) \in \mathcal{Q}(\mathcal{J}, \mathcal{Q}, \bar{P}) \Leftrightarrow E(X|\mathcal{Q}) \circ id \in \mathcal{Q}(\mathcal{J}, \mathcal{F}, P)$$

$$\text{and } \int_A E(X|\mathcal{Q}) dP = \int_{id^{-1}(A)} E(X|\mathcal{Q}) \circ id dP$$

$$\stackrel{\text{def}}{=} \int_A E(X|\mathcal{Q}) d\bar{P}$$

$$= \int_A X dP \leftarrow v(A)$$

as was to be shown.

Now just suppose $X \in Q(\mathcal{J}, \mathcal{F}, P)$.

Assume $X \in Q^+(\mathcal{J}, \mathcal{F}, P)$ w.l.o.g.

$\therefore v(\cdot) := \int \cdot dP$ is a finite measure on \mathcal{Q}

$$\therefore E(X^+|\mathcal{Q}) := \frac{dv}{d\bar{P}} \in L_1(\mathcal{J}, \mathcal{Q}, \bar{P})$$

by Thm "props of RND".

$\therefore E(X^+|\mathcal{Q}) \in L_1(\mathcal{J}, \mathcal{F}, P)$, change of variables

$$\therefore E(X|\Omega) := \underbrace{E(X^+|\Omega)}_{\in L_1(\Omega, \mathcal{F}, P)} - \underbrace{E(X^-|\Omega)}_{\in \mathcal{N}(\Omega, \mathcal{F})} \stackrel{P\text{-a.e. defined}}{\in} Q^+(\Omega, \mathcal{F}, P)$$

and $\forall A \in \Omega$

$$\begin{aligned} \int_A E(X|\Omega) dP &= \int_A \underbrace{E(X^+|\Omega)}_{\in L_1} dP - \int_A \underbrace{E(X^-|\Omega)}_{\in \mathcal{N}} dP \\ &= \int_A X^+ dP - \int_A X^- dP \quad \text{by Prop 3} \\ &= \int_A X dP \end{aligned}$$

This establishes (i) & (ii) & $E(X|\Omega) \in Q^+(\Omega, \mathcal{F}, P)$.

For uniqueness suppose $\tilde{E}(X|\Omega)$ is another version.

Now by Thm on indefinite integrals in Lecture 11

$$\begin{aligned} \int \cdot \tilde{E}(X|\Omega) dP &\stackrel{(ii)}{=} \int \cdot E(X|\Omega) dP \text{ on } \Omega \\ \Rightarrow \tilde{E}(X|\Omega) &= E(X|\Omega) \text{ P-a.e.} \\ \Rightarrow \tilde{E}(X|\Omega) &= E(X|\Omega) \text{ P-a.e.} \quad \text{QED} \end{aligned}$$

Remark: By construction we have

$$X \in Q^+(\Omega, \mathcal{F}, P) \Rightarrow E(X|\Omega) \in Q^+(\Omega, \mathcal{F}, P)$$

$$X \in Q^-(\Omega, \mathcal{F}, P) \Rightarrow E(X|\Omega) \in Q^-(\Omega, \mathcal{F}, P).$$

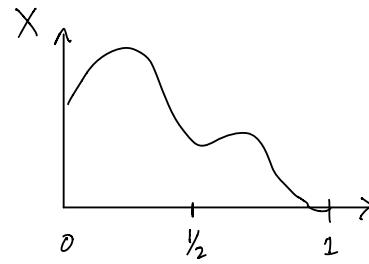
Remark: One way to think about $E(X|\Omega)$ is

$$E(X|\Omega)(w) = E(X \mid \text{the sets } A \in \Omega \text{ s.t. } w \in A)$$

A slightly more rigorous view is to think of $E(X|\Omega)(w)$ as the weighted average of X over the "smallest Ω -set" containing w , i.e. a smoothing or granulation of X , or a projection of X onto the space of Ω -measurable functions.

Example:

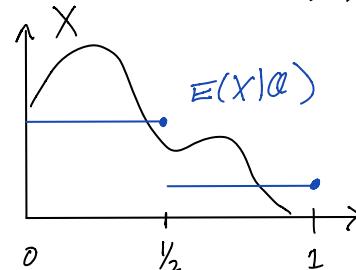
$$(\Omega, \mathcal{F}, P) = ([0,1], \mathcal{B}([0,1]), P)$$



$$\Omega = \{\emptyset, \Omega_1, [0, \frac{1}{2}], (\frac{1}{2}, 1]\}$$

Guess at $E(X|\Omega)$ & show it has the correct properties

$$E(X|\Omega)(w) := \begin{cases} \frac{E(I_{[0, \frac{1}{2}]} X)}{P([0, \frac{1}{2}])} & \text{if } w \in [0, \frac{1}{2}] \\ \frac{E(I_{(\frac{1}{2}, 1]} X)}{P((\frac{1}{2}, 1])} & \text{if } w \in (\frac{1}{2}, 1] \end{cases}$$



$E(X|\Omega)$ is Ω -measurable (it's a simple function w.r.t. Ω)

$$E(X|\Omega) \in Q(\Omega, \mathcal{F}, P)$$

Also if $A = [0, \frac{1}{2}]$

$$\begin{aligned} \int_A E(X|\Omega) dP &= \frac{E(I_{[0, \frac{1}{2}]} X)}{P([0, \frac{1}{2}])} P(A) \\ &= \int_A X dP \end{aligned}$$

and similarly for $A = \emptyset, \Omega_1$ or $(\frac{1}{2}, 1]$.

$\therefore E(X|\Omega)$ has the desired properties and is the P-a.e. unique such choice.

(6)

Example

Let (Ω, \mathcal{F}, P) be a probability space and $X \in \mathcal{L}_2(\Omega, \mathcal{F}, P)$. Suppose

$$\{\emptyset, \Omega\} = \mathcal{F}_0 \subset \mathcal{F}_1 \subset \dots \subset \mathcal{F}$$

is an increasing sequence of sub- σ -fields

Then

$$E(X|\mathcal{F}_0), E(X|\mathcal{F}_1), \dots, E(X|\mathcal{F})$$

$$\begin{array}{ccc} E(X) & \xrightarrow{\text{increasing resolution approx}} & X \\ \parallel & & \parallel \end{array}$$

Example:

This example shows how to understand $E(X|\mathcal{Q})$ as a projection when $X \in L_2(\Omega, \mathcal{F}, P)$.

Define

$$S := \{Y \in L_2(\Omega, \mathcal{F}, P) : Y \text{ is } \mathcal{Q}\text{-measurable}\}$$

Notice that S a closed linear subspace of $L_2(\Omega, \mathcal{F}, P)$ by the Closure thm.

The projection $P_S X$ satisfies

$$X - P_S X \perp w \quad \forall w \in S$$

$$\therefore E((X - P_S X)w) = 0 \quad \forall w \in S$$

$$\therefore E(Xw) = E(P_S Xw) \quad \forall w \in S$$

Given $A \in \mathcal{Q}$, set $w = \mathbf{1}_A \in S$ so that

$$\int_A X dP = \int_A P_S X dP$$

Since $P_S X \in S \subset L_2(\Omega, \mathcal{F}, P)$ we have

$$E(X|\mathcal{Q}) = P_S X. \quad P\text{-a.e.}$$

(7)

Example

$$\Omega = [-1, 1]$$

$$\mathcal{F} = \mathcal{B}([-1, 1])$$

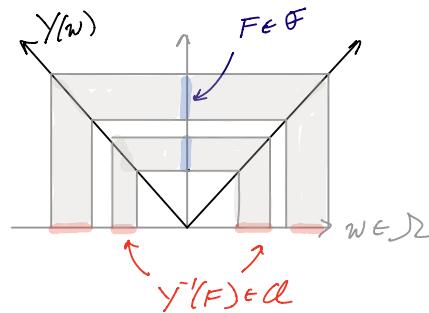
$$dP = f(w)dw$$

Let $X \in \mathcal{L}_2(\Omega, \mathcal{F}, P)$ and define

$$Y(w) = |w| \quad \text{on } w \in \Omega \quad \text{so that}$$

$$E(Y) = Y^{-1}(\mathcal{F}) \subset \mathcal{F}$$

pull back
of a single
map



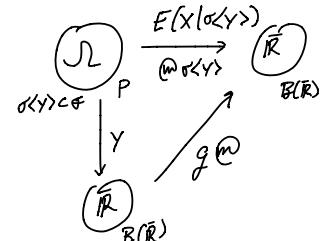
To make a guess at what $E(X|\sigma(Y))$ is notice

$$E(X|\sigma(Y)) \in \sigma(Y)$$

↓ Lecture 8

$$E(X|\sigma(Y))(w) = g(Y(w))$$

$$= g(|w|)$$



So $E(X|\sigma(Y))(w) = g(|w|)$ for some measurable g . Since the "smallest $\sigma(Y)$ sets containing w are of the form $\{w, -w\}$ " set

$$E(X|\sigma(Y))(w) = X(w) \frac{s(w)}{s(w) + s(-w)} + X(-w) \frac{s(-w)}{s(w) + s(-w)}$$

= weighted ave of X over $\{w, -w\}$ with weights given by the density of P w.r.t. Lebesgue measure.

(9) Since this $E(X|\sigma(Y))$ is a \mathcal{Q} function of w it is $\sigma(Y)$ -measurable.

Since $E(X|\sigma(Y)) \geq 0$ it is in $\mathcal{Q}(\Omega, \mathcal{F}, P)$.

To show that it is the conditional expected value of X w.r.t. $\sigma(Y)$ just check it integrates the same as X over sets $A \in \sigma(Y)$.

Indeed if $A \in \sigma(Y)$ then $A = -A$ so

$$\begin{aligned} \int_A E(X|\sigma(Y)) dP &= \int_A X(w) \frac{\delta(w)}{\delta(w) + \delta(-w)} \delta(w) dw \\ &\quad + \underbrace{\int_A X(-w) \frac{\delta(-w)}{\delta(w) + \delta(-w)} \delta(w) dw}_{\text{By change of variables}} \\ &\quad \text{this is } \int_{-A} X(w) \frac{\delta(w)}{\delta(w) + \delta(-w)} \delta(-w) dw \\ &= \int_A X(w) \frac{\delta(w)}{\delta(w) + \delta(-w)} [\delta(w) + \delta(-w)] dw \\ &= \int_A X(w) dP \end{aligned}$$

(10) Properties of $E(X|\mathcal{Q})$

These come in two flavors:

- 1) Properties of $E(X|\mathcal{Q})$ that resemble properties of integration & expected value
- 2) Properties of $E(X|\mathcal{Q})$ which mimic that of a smoothing operator.

Assumptions: To state the results more

clearly assume throughout this section that X, Y, X_1, X_2, \dots are quasi-integrable extended r.v.s on a prob space (Ω, \mathcal{F}, P) & $\mathcal{Q}, \mathcal{Q}_1, \mathcal{Q}_2, \dots$ are sub- σ -fields of \mathcal{F} .

Theorem: (Smoothing properties of $E(X|\mathcal{Q})$)

$$(s1) E(E(X|\mathcal{Q})) = E(X)$$

$$(s2) \text{ If } \mathcal{Q}_1 \subset \mathcal{Q}_2 \text{ (i.e. } \mathcal{Q}_1 \text{ smooths more) Then} \\ E(E(X|\mathcal{Q}_2)|\mathcal{Q}_1) \stackrel{P-a.e.}{=} E(X|\mathcal{Q}_1)$$

$$= E(E(X|\mathcal{Q}_1)|\mathcal{Q}_2)$$

$$(s3) X \in \mathcal{Q}^\pm(\Omega, \mathcal{F}, P) \iff E(X|\mathcal{Q}) \in \mathcal{Q}^\pm(\Omega, \mathcal{F}, P)$$

$$X \in L_1(\Omega, \mathcal{F}, P) \iff E(X|\mathcal{Q}) \in L_1(\Omega, \mathcal{F}, P)$$

$$(s4) \text{ If } X \text{ is } \mathcal{Q}\text{-measurable then}$$

$$E(XY|\mathcal{Q}) \stackrel{P-a.e.}{=} X E(Y|\mathcal{Q})$$

and in particular

$$E(X|\mathcal{Q}) \stackrel{P-a.e.}{=} X$$

(Note: (s4) implicitly assumes XY is
quasi-integrable ... along with X & Y)

(11)

we will also give the result for the expected value properties of $E(X|\Omega)$ &

then prove both together (since proving (s4), for example, uses the following linearity results).

Theorem: (expected value properties of $E(X|\Omega)$)

$$(e_1) \quad E(cX|\Omega) \stackrel{P.a.e.}{=} cE(X|\Omega) \quad \forall c \in \mathbb{R}.$$

$$(e_2) \quad E(X+Y|\Omega) = E(X|\Omega) + E(Y|\Omega)$$

P-a.e. on the set where

$E(X|\Omega) + E(Y|\Omega)$ is defined (i.e. not $\pm\infty$).

Note: There is an implicit assumption here that $X, Y, X+Y \in \mathcal{Q}(\Omega, \mathcal{F}, P)$ but not necessarily of the same class.

$$(e_3) \quad X \stackrel{P.a.e.}{\leq} Y \Rightarrow E(X|\Omega) \stackrel{P.a.e.}{\leq} E(Y|\Omega)$$

$$(e_4) \quad |E(X|\Omega)| \leq E(|X||\Omega)$$

(e5) If $0 \leq X_n \uparrow X$ P-a.e. then

$$\xrightarrow{\text{MCT}} 0 \leq E(X_n|\Omega) \uparrow E(X|\Omega) \quad \text{P-a.e.}$$

(e6) If $0 \leq X_n$ P-a.e. $\forall n \in \mathbb{N}$ then

$$\xrightarrow{\text{Fatou}} E(\liminf_n X_n|\Omega) \leq \liminf_n E(X_n|\Omega)$$

(e7) If $X_n \xrightarrow{a.e.} X$ then

$$\xrightarrow{\text{Def}} E(X_n|\Omega) \xrightarrow{a.e.} E(X|\Omega)$$

on the set $\{E(\sup_n |X_n| |\Omega) < \infty\}$

(12)

Before we prove the above two Theorems we need The following lemma.

Lemma: If $X \in \mathcal{Q}(\Omega, \mathcal{F}, P)$ then $\forall A \in \Omega$

$$E(I_A X|\Omega) \stackrel{P.a.e.}{=} I_A E(X|\Omega)$$

Proof:

Clearly $I_A E(X|\Omega)$ is $\cap \Omega$ & in $\mathcal{Q}(\Omega, \mathcal{F}, P)$.

Also $\forall \tilde{A} \in \Omega$

$$\int_{\tilde{A}} I_A E(X|\Omega) dP = \int_{A \cap \tilde{A}} E(X|\Omega) dP$$

$$\stackrel{\text{def}}{=} \int_{A \cap \tilde{A}} X dP \quad \text{since } A \cap \tilde{A} \in \Omega$$

$$= \int_{\tilde{A}} I_{\tilde{A}} X dP$$

$$\stackrel{\text{def}}{=} \int_{\tilde{A}} E(I_{\tilde{A}} X|\Omega) dP$$

$$\therefore E(I_A X|\Omega) \stackrel{P.a.e.}{=} I_A E(X|\Omega) \quad \text{by uniqueness}$$

QED

Proof of (s1)-(s4) & (e1)-(e7):

(s1) follows simply by the requirement:

$$\int_A X dP = \int_A E(X|\Omega) dP \quad \text{when } A = \cup \tilde{A} \in \Omega.$$

(s2) we show $E(X|\Omega_1)$ is a version of $E(E(X|\Omega_1)|\Omega_2)$ and $E(E(X|\Omega_2)|\Omega_1)$.

Note $E(X|\Omega_1)$ is both Ω_1 & Ω_2 measurable and in $\mathcal{Q}(\Omega, \mathcal{F}, P)$.

Now $\forall A \in \mathcal{Q}_1$, $E(E(X|\mathcal{Q}_2)|\mathcal{Q}_1)$ satisfies (13)

$$\int_A E(E(X|\mathcal{Q}_2)|\mathcal{Q}_1) dP \stackrel{\text{def}}{=} \int_A E(X|\mathcal{Q}_2) dP$$

... but so does $E(X|\mathcal{Q}_1)$ since

$$\begin{aligned} \int_A E(X|\mathcal{Q}_1) dP &\stackrel{\text{def}}{=} \int_A X dP \\ &\stackrel{\text{def}}{=} \int_A E(X|\mathcal{Q}_2) dP \quad \text{since } A \in \mathcal{Q}_1 \subset \mathcal{Q}_2 \end{aligned}$$

$$\therefore E(X|\mathcal{Q}_1) \stackrel{\text{P-a.e.}}{=} E(E(X|\mathcal{Q}_2)|\mathcal{Q}_1).$$

$$\text{& similarly } E(X|\mathcal{Q}_1) \stackrel{\text{P-a.e.}}{=} E(E(X|\mathcal{Q}_1)|\mathcal{Q}_2).$$

(53) This is just what the remark was following
The Theorem establishing the existence of
 $E(X|\mathcal{Q})$.

(e1) This clearly follows from Big 3.

(e2) : We show if $X, Y, X+Y \in \mathcal{Q}(\mathcal{D}, \mathcal{F}, P)$

then $\mathbb{I}_A [E(X+Y|\mathcal{Q})] \stackrel{\text{P-a.e.}}{=} \mathbb{I}_A E(X|\mathcal{Q}) + \mathbb{I}_A E(Y|\mathcal{Q})$
where $A := \{E(X|\mathcal{Q}) + E(Y|\mathcal{Q}) \text{ is defined}\}$. ↑
equals
if the sum is defined.

case 1: $X \& Y$ are in $\mathcal{Q}^-(\mathcal{D}, \mathcal{F}, P)$.

Now by (53) $E(X|\mathcal{Q}) \in \mathcal{Q}^-$ & $E(Y|\mathcal{Q}) \in \mathcal{Q}^-$

so we may change them on P -null sets so
that $E(X|\mathcal{Q}) > -\infty$ & $E(Y|\mathcal{Q}) > -\infty$
everywhere on \mathcal{D} .

Now $E(X|\mathcal{Q}) + E(Y|\mathcal{Q}) > -\infty$ & is
therefore in \mathcal{Q}^- & is \mathcal{Q} -measurable
by the closure theorem.

we also have $\forall A \in \mathcal{Q}$

$$\int_A E(X|\mathcal{Q}) + E(Y|\mathcal{Q}) dP$$

$$\stackrel{\text{Big 3}}{=} \int_A E(X|\mathcal{Q}) dP + \int_A E(Y|\mathcal{Q}) dP$$

$$\stackrel{\text{def}}{=} \int_A X dP + \int_A Y dP$$

$$\stackrel{\text{Big 3}}{=} \int_A X+Y dP$$

$$\therefore E(X+Y|\mathcal{Q}) = E(X|\mathcal{Q}) + E(Y|\mathcal{Q}) \text{ P-a.e. on } \mathcal{D}.$$

case 2: $X, Y, X+Y \in \mathcal{Q}^+(\mathcal{D}, \mathcal{F}, P)$.

For $-\infty < c \leq 0$ define

$$A_c := \{E(X|\mathcal{Q}) \geq c\} \cap \{E(Y|\mathcal{Q}) \geq c\} \in \mathcal{Q}$$

Now the previous lemma shows

$$E(\mathbb{I}_{A_c} X|\mathcal{Q}) \stackrel{\text{P-a.e.}}{=} \mathbb{I}_{A_c} E(X|\mathcal{Q}) \geq c > -\infty$$

$$E(\mathbb{I}_{A_c} Y|\mathcal{Q}) \stackrel{\text{P-a.e.}}{=} \mathbb{I}_{A_c} E(Y|\mathcal{Q}) \geq c > -\infty$$

The previous case now applies by (53) & gives

$$E(\mathbb{I}_{A_c} X + \mathbb{I}_{A_c} Y|\mathcal{Q}) \stackrel{\text{P-a.e.}}{=} E(\mathbb{I}_{A_c} X|\mathcal{Q}) + E(\mathbb{I}_{A_c} Y|\mathcal{Q})$$

ii Lemma

ii Lemma

$$\mathbb{I}_{A_c} E(X+Y|\mathcal{Q}) = \mathbb{I}_{A_c} E(X|\mathcal{Q}) + \mathbb{I}_{A_c} E(Y|\mathcal{Q})$$

$$\therefore E(X+Y|\mathcal{Q}) = E(X|\mathcal{Q}) + E(Y|\mathcal{Q}) \text{ P-a.e. on }$$

$$A_\infty := \bigcup_{c \in \mathcal{Q}^+} A_c = \{E(X|\mathcal{Q}) > -\infty \text{ and } E(Y|\mathcal{Q}) > -\infty\}$$

A similar argument shows

$$E(X+Y|\mathcal{Q}) = E(X|\mathcal{Q}) + E(Y|\mathcal{Q}) \text{ P-a.e. on }$$

$$A_\infty := \{E(X|\mathcal{Q}) < \infty \text{ and } E(Y|\mathcal{Q}) < \infty\}$$

To finish simply notice $A \cup A_{-\infty}$ is
exactly the set $\{E(X|\Omega) + E(Y|\Omega)\}$ is defined. (15)

(e3): show $X \stackrel{\text{P.a.e.}}{\leq} Y \Rightarrow E(X|\Omega) \stackrel{\text{P.a.e.}}{\leq} E(Y|\Omega)$.

This follows from the indefinite integral results in Lecture II. Indeed

$$\begin{aligned} \int_A X dP &\leq \int_A Y dP \quad \forall A \in \mathcal{Q} \\ \Rightarrow \int_A E(X|\Omega) dP &\leq \int_A E(Y|\Omega) dP, \quad \forall A \in \mathcal{Q} \\ \xrightarrow{\text{Lecture II}} E(X|\Omega) &\stackrel{\text{P.a.e.}}{\leq} E(Y|\Omega) \end{aligned}$$

(e4): show $|E(X|\Omega)| \leq E(|X||\Omega)$

This follows from (e3) since it implies
 $-E(|X||\Omega) \leq E(X|\Omega) \leq E(|X||\Omega)$ P-a.e.

(e5): show if $0 \leq X_n \uparrow X$ P-a.e. then
 $\overbrace{\lim^{\text{MC}}}^{\text{MC}} 0 \leq E(X_n|\Omega) \uparrow E(X|\Omega)$ P-a.e.

Note that (e3) establishes that

$$0 \leq E(X_n|\Omega) \uparrow \quad \text{P-a.e.}$$

$\therefore \lim_n E(X_n|\Omega)$ exists P-a.e. and is
in $\mathbb{Q}^-(\Omega, \mathcal{F}, P)$ by non-negativity & closure.

Similarly $X \notin \mathbb{Q}^-(\Omega, \mathcal{F}, P)$.

We show $\lim_n E(X_n|\Omega) \stackrel{\text{P.a.e.}}{=} E(X|\Omega)$.

Indeed $\forall A \in \mathcal{Q}$

$$\begin{aligned} \int_A \lim_n E(X_n|\Omega) dP &\stackrel{\text{MC}}{=} \lim_n \int_A E(X_n|\Omega) dP \\ &\stackrel{\text{def}}{=} \lim_n \int_A X_n dP \\ &\stackrel{\text{MC}}{=} \int_A \lim_n X_n dP \\ &\quad X \end{aligned}$$

\therefore by the characterizing properties of $E(X|\Omega)$ we have

$$\lim_n E(X_n|\Omega) \stackrel{\text{P.a.e.}}{=} E(X|\Omega)$$

(e6) Factor. This is similar to (e5).

(e7) show that if $X_n \xrightarrow{\text{a.e.}} X$ &
 $X, X_n \in \mathbb{Q}(\Omega, \mathcal{F}, P)$ then

$$I_A E(X_n|\Omega) \xrightarrow{\text{a.e.}} I_A E(X|\Omega)$$

where $A := \{E(\sup_n |X_n| |\Omega) < \infty\}$.

Let $0 \leq c < \infty$ & set

$$A_c := \left\{ E\left(\sup_n |X_n| |\Omega\right) \leq c \right\} \in \mathcal{Q}$$

$$\therefore \underbrace{E\left(\sup_n |I_{A_c} X_n| |\Omega\right)}_{\substack{\text{Lemma} \\ =}} \leq c$$

$$\begin{aligned} &= I_{A_c} E\left(\sup_n |X_n| |\Omega\right) \\ &\leq c \end{aligned}$$

$$\therefore E\left(\sup_n |I_{A_c} X_n| |\Omega\right) \in L_1(\Omega, \mathcal{F}, P)$$

$$\therefore \sup_n |I_{A_c} X_n| \in L_1(\Omega, \mathcal{F}, P) \text{ by (S3)}$$

Now setting $Y = \sup_{n \in \mathbb{N}} |I_{A_n} X_n|$ (17)

$$E\left(\liminf_n \underbrace{(Y + I_{A_n} X_n)}_{\geq 0} | \mathcal{Q}\right) \stackrel{\text{P-a.e.}}{\leq} \liminf_n E(Y + I_{A_n} X_n | \mathcal{Q})$$

by Fatou. Then by canceling $E(Y | \mathcal{Q})$ from both sides (possible since $Y \in L_1 \Rightarrow E(Y | \mathcal{Q}) < \infty$ P-a.e.) we get

$$E\left(\liminf_n I_{A_n} X_n | \mathcal{Q}\right) \stackrel{\text{P-a.e.}}{\leq} \liminf_n E(I_{A_n} X_n | \mathcal{Q})$$

Therefore

$$I_{A_c} E\left(\liminf_n X_n | \mathcal{Q}\right) \stackrel{\text{P-a.e.}}{=} I_{A_c} \liminf_n E(X_n | \mathcal{Q})$$

Similarly

$$I_{A_c} \limsup_n E(X_n | \mathcal{Q}) \stackrel{\text{P-a.e.}}{\leq} I_{A_c} E\left(\limsup_n X_n | \mathcal{Q}\right)$$

so that

$$I_{A_c} \lim_n E(X_n | \mathcal{Q}) \stackrel{\text{P-a.e.}}{=} I_{A_c} E(X | \mathcal{Q})$$

$$\therefore I_A \lim_n E(X_n | \mathcal{Q}) \stackrel{\text{P-a.e.}}{=} I_A E(X | \mathcal{Q})$$

$$\text{for } A := \bigcup_{c \in \mathbb{Q}^+} A_c = \left\{ E\left(\sup_n |X_n|\right) < c \right\} \in \mathcal{Q}.$$

(18)

(54) Show that if $X, Y, XY \in \mathcal{Q}(\Omega, \mathcal{F}, P)$ and $X \in \mathcal{Q}$ then $E(XY | \mathcal{Q}) \stackrel{\text{P-a.e.}}{=} X E(Y | \mathcal{Q})$.

case 1: $X, Y \in \mathcal{G}(\Omega, \mathcal{F}, P)$ & $X \in \mathcal{Q}$

By the lemma $\exists \tilde{A} \in \mathcal{Q}$

$$\int_{\tilde{A}} I_{\tilde{A}} E(Y | \mathcal{Q}) dP = \int_{\tilde{A}} I_{\tilde{A}} Y dP.$$

By the structure thm \exists non-negative simple functions ($\in \mathcal{Q}$) X_n s.t.

$$0 \leq X_n \uparrow X.$$

By linearity (e2) & MCT (e5) we have

$$\int_{\tilde{A}} X_n E(Y | \mathcal{Q}) dP \stackrel{\text{e2}}{=} \int_{\tilde{A}} X_n Y dP.$$

$$\stackrel{\text{e5}}{\downarrow} \int_{\tilde{A}} n \rightarrow \infty \quad \stackrel{\text{e5}}{\downarrow} \int_{\tilde{A}} n \rightarrow \infty$$

$$\int_{\tilde{A}} X E(Y | \mathcal{Q}) dP = \int_{\tilde{A}} X Y dP$$

since $XE(Y | \mathcal{Q})$ is $\in \mathcal{Q}$ by closure & in $\mathcal{G}(\Omega, \mathcal{Q})$

$$\text{we have } X E(Y | \mathcal{Q}) \stackrel{\text{P-a.e.}}{=} E(XY | \mathcal{Q})$$

Case 2: $X, Y, XY \in \mathcal{Q}(\Omega, \mathcal{F}, P)$ & $X \in \mathcal{Q}$

$$XY = (X^+ - X^-)(Y^+ - Y^-)$$

$$\begin{aligned} &= (X^+ Y^+ + Y^- Y^-) - (X^- Y^+ + X^+ Y^-) \\ &\quad \underbrace{(X^+ Y^+ + Y^- Y^-)}_{=(XY)^+} \quad \underbrace{(X^- Y^+ + X^+ Y^-)}_{=(XY)^-} \end{aligned}$$

$$\therefore E(XY|\alpha) \stackrel{def}{=} E((XY)^+|\alpha) - E((XY)^-|\alpha) \quad (19)$$

$$\begin{aligned} &= E(X^+Y^+|\alpha) + E(X^-Y^-|\alpha) \\ &\quad - E(X^-Y^+|\alpha) - E(X^+Y^-|\alpha) \\ &\stackrel{\text{P-a.e.}}{=} \underset{\text{case 1}}{X^+} E(Y^+|\alpha) + \underset{\text{case 2}}{X^-} E(Y^-|\alpha) \\ &\quad - X^- E(Y^+|\alpha) - X^+ E(Y^-|\alpha) \\ &= (X^+ - X^-)(E(Y^+|\alpha) - E(Y^-|\alpha)) \\ &\stackrel{\text{P-a.e.}}{=} X E(Y|\alpha). \end{aligned}$$

QED

Remark: Notice the technique for showing the linearity (e2) & the DCT (e7)... regularize with \mathcal{I}_{A_ϵ} , use the lemma to show $\mathcal{I}_{A_\epsilon} E(X|\alpha) \stackrel{\text{a.e.}}{=} \mathcal{I}_{A_\epsilon} E(Y|\alpha)$ then take a countable union over A_ϵ sets.

Let's do a quick example that shows the generality of (e2). (20)

Example:

$$(\Omega, \mathcal{F}, P) = ([0, 1], \mathcal{B}([0, 1]), \text{Lebesgue measure})$$

$$\mathcal{Q} = \{\emptyset, \Omega, [0, \frac{1}{2}], [\frac{1}{2}, 1]\}$$

$$X(w) = \frac{1}{w}$$

Now P-a.e. we have

$$E(X|\alpha) = \begin{cases} \infty & \text{on } [0, \frac{1}{2}] \\ 2\log 2 & \text{on } [\frac{1}{2}, 1] \end{cases} \quad \& \quad E(-X|\alpha) = -E(X|\alpha)$$

and

$$\mathcal{I}_{[\frac{1}{2}, 1]} E(X - X|\alpha) = \mathcal{I}_{[\frac{1}{2}, 1]} E(X|\alpha) + \mathcal{I}_{[\frac{1}{2}, 1]} E(-X|\alpha)$$

but you can't remove the indicator since the RHS is not defined.

The next result is a useful tool for showing $E(X|\sigma(\rho))$ has a particular form.

Theorem: (π -generators are enough)

Let $\mathcal{P} \subset \mathcal{F}$ be a π -system
and $X \in L(\Omega, \mathcal{F}, P)$. If $Y \in L(\Omega, \sigma(\mathcal{P}), P)$
s.t. $\forall A \in \mathcal{P}$

$$\int_A X dP = \int_A Y dP \quad (*)$$

then $Y \stackrel{\text{P.a.e.}}{=} E(X|\sigma(\mathcal{P}))$.

Proof:

If (*) holds then

$$\int \cdot X^\pm dP = \int \cdot Y^\pm dP \quad \text{on } \Omega$$

Since $\int x^\pm dP$ & $\int y^\pm dP$ are both finite (21) measures on $(\Omega, \sigma(\mathcal{P}))$ which agree on a π -system \mathcal{P} we have

$$\sigma > \int x^\pm dP = \int y^\pm dP \text{ on } \sigma(\mathcal{P})$$

$$\therefore \int x dP = \int y dP \text{ on } \sigma(\mathcal{P}).$$

$$\therefore Y = E(X | \sigma(\mathcal{P})) \quad \text{QED}$$

Remark: The above claim can be extended if it is a-priori known that

$\int x^\pm dP$ & $\int y^\pm dP$ are σ -finite measures over \mathcal{P} .

Conditional expected value of X given Y

If $X \in \mathcal{Q}(\Omega, \mathcal{F}, P)$ then we can define $E(X | Y_1, Y_2, \dots)$ as follows

$$E(X | Y_1, Y_2, \dots) := E(X | \sigma(Y_1, Y_2, \dots))$$

Notice that by definition we have

$$E(X | \sigma(Y_1, Y_2, \dots)) \subset \sigma(Y_1, Y_2, \dots)$$

there exist a Borel measurable $g: \mathbb{R} \rightarrow \mathbb{R}$ s.t.

$$E(X | Y_1, Y_2, \dots) = g(Y_1, Y_2, \dots)$$

by "two fans of two fans them" in Lecture 9.

In general, the notation

$$E(X | Y_1=y_1, Y_2=y_2, \dots)$$

but **warning** this is only meaningful for a fixed g since we are free to change $E(X | Y_1, \dots)$ on a P -null set.

Now notice that this g must satisfy (22)

$$E(I_{(Y_1, Y_2, \dots) \in B} g(Y_1, Y_2, \dots))$$

$$= \int g(\vec{Y}) dP \quad \leftarrow E(X | \sigma(Y_1, Y_2, \dots))$$

$$\stackrel{\text{def}}{=} \int_X dP \quad \leftarrow \sigma(Y_1, Y_2, \dots) \text{ sets}$$

$$= E(I_{(Y_1, Y_2, \dots) \in B} X) \quad (\star)$$

$$\forall B \in \mathcal{B}(\mathbb{R}^\infty).$$

Also recall that

$\mathcal{B}(\mathbb{R}^\infty)$ coordinate projection

mappings

$$\begin{aligned} \mathcal{B}(\mathbb{R}^\infty) &\xrightarrow{\text{coordinate projection}} \mathcal{B}(\mathbb{R}) \\ &= \sigma(\pi_1, \pi_2, \dots) \quad \leftarrow \mathcal{B}((-\infty, x]: x \in \mathbb{R}) \\ &= \sigma(\pi_i^{-1}((-\infty, x]): i \in \mathbb{N}, x \in \mathbb{R}) \end{aligned}$$

by generators are enough

$$= \sigma(\underbrace{\pi_{i_1}^{-1}((-\infty, x_1]) \cap \dots \cap \pi_{i_n}^{-1}((-\infty, x_n])}_{\text{finite coordinate proj}})$$

generators

Again by generators are enough we have

$$\sigma(Y_1, Y_2, \dots) = \sigma(\vec{Y}^{-1}(B): B \in \mathcal{B}(\mathbb{R}^\infty))$$

$$= \sigma(\vec{Y}^{-1}(B): B \in \underset{\text{finite coord}}{\text{proj}})$$

$$= \sigma(\underbrace{\{Y_{i_1} \leq x_1, \dots, Y_{i_n} \leq x_n\}}_{\text{Pi-system}}: x_i \in \mathbb{R})$$

new

Combining with our π -system Thm above gives this corollary which works for finite or infinite sequences.

Corollary: If $X \in L_1(\Omega, \mathcal{F}, P)$ & Y_1, Y_2, \dots are r.v.s on (Ω, \mathcal{F}, P) then

$E(X|Y_1, Y_2, \dots)$ is the P -a.e. unique r.v. on (Ω, \mathcal{F}, P) that satisfies

$$E(X|Y_1, Y_2, \dots) = g(Y_1, Y_2, \dots)$$

where $g: \mathbb{R}^\infty \rightarrow \mathbb{R}$ is Borel measurable, $g(Y_1, Y_2, \dots) \in L_1(\Omega, \mathcal{F}, P)$ and

$$\begin{aligned} E(I_{\{Y_1 \leq x_1\}} \cdots I_{\{Y_n \leq x_n\}} g(Y_1, Y_2, \dots)) \\ \stackrel{\text{(*)}}{=} E(I_{\{Y_1 \leq x_1\}} \cdots I_{\{Y_n \leq x_n\}} X) \end{aligned}$$

for all $x_1, \dots, x_n \in \mathbb{R}$ & $n \geq 1$

(23)

Example:

Suppose X & Y_1, Y_2, \dots, Y_n are r.v.s taking values in a discrete set $Q \subset \mathbb{R}$. Then

$$g(y_1, \dots, y_n) := \begin{cases} \sum_{x \in Q} x \frac{P(X=x, Y_1=y_1, \dots, Y_n=y_n)}{P(Y_1=y_1, \dots, Y_n=y_n)} & \text{if } > 0 \\ 0 & \text{o.w.} \end{cases}$$

(24)

has the property that

$$E(X|Y_1, \dots, Y_n) \stackrel{P\text{-a.e.}}{=} g(Y_1, \dots, Y_n).$$

This follows since $\forall y_1, \dots, y_n \in Q$

$$\begin{aligned} E(I_{\{Y_1=y_1, \dots, Y_n=y_n\}} X) \\ \stackrel{\substack{\text{Pi-sgs} \\ \text{generators}}}{=} \sum_{x \in Q} x I_{X=x} \\ = \sum_{x \in Q} x P(X=x, Y_1=y_1, \dots, Y_n=y_n) \\ = \begin{cases} \sum_{x \in Q} x \frac{P(X=x, Y_1=y_1, \dots, Y_n=y_n)}{P(Y_1=y_1, \dots, Y_n=y_n)} P(Y_1=y_1, \dots, Y_n=y_n) & \text{when this is} \\ 0 & \text{o.w.} \end{cases} \\ = g(y_1, \dots, y_n) P(Y_1=y_1, \dots, Y_n=y_n) \\ = E(I_{\{Y_1=y_1, \dots, Y_n=y_n\}} g(Y_1, \dots, Y_n)) \end{aligned}$$

and g is trivially \mathcal{G} and integrable as it's a discrete map.