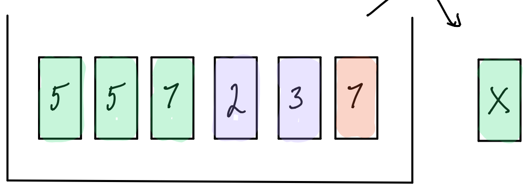
Lecture 20: Conditional expected value with respect to a sub-σ-field

Lets start with a modivation.

Consider a box with numbered tickets which are colored.

I pick one at random & show you the color but not the number



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

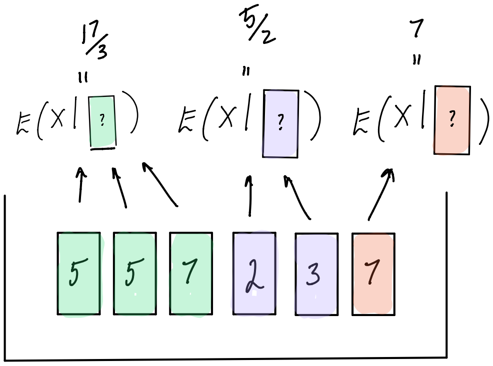
$$E(X \mid \boxed{?}) = \left\{ \begin{array}{l} \text{average green} \\ \text{ticket number} \end{array} \right.$$

$$= \frac{5+5+7}{3} = \frac{17}{3}$$

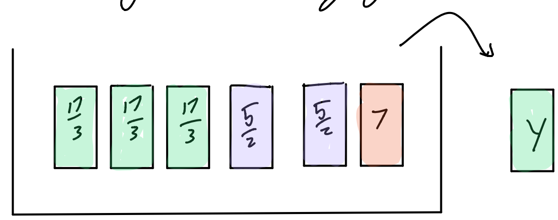
If you wanted to automate this prediction you could pre-compute

$$E(X \mid \boxed{?}) \quad E(X \mid \boxed{?}) \quad E(X \mid \boxed{?})$$

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



or as a granular smoothing of X



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

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

(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

Ω = the collection of tickets

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

$\mathcal{A}$ = the possible bets on color

For $\omega \in \Omega$, $X(\omega) =$ ticket #

$E(X|\mathcal{A}) = Y$ maps $\omega \in \Omega \longmapsto$ ave X $\forall$ tickets with the same color as ω

and we have

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

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

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

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

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

(i) $E(X|\mathcal{A})$ is $\mathcal{A}$-measurable $\leftarrow$ more granular than X

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

Proof:

Start by assuming $X \geq 0$.

Let $\nu(\cdot) = \int_\cdot X\,dP$ be a measure on $(\Omega, \mathcal{A})$.

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

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

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

So look for this as $d\nu/d\bar{P}$

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

Notice $\nu <\!\!<\!\!< \bar{P}$ since

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

$$\Rightarrow \nu(A) = \int_\Omega I_A X\,dP = 0$$

$\therefore$ By the Radon-Nikodym Thm $\exists$ a $\bar{P}$-unique $d\nu/d\bar{P} \in \mathfrak{M}(\Omega, \mathcal{A})$ s.t. $\forall A \in \mathcal{A}$

$$\int_A \frac{d\nu}{d\bar{P}}\,d\bar{P} = \nu(A) = \int_A X\,dP$$

So set $E(X|\mathcal{A}) := \frac{d\nu}{d\bar{P}} \in \mathfrak{M}(\Omega, \mathcal{A})$

By the change of variables Thm

$$\begin{tikzcd} \underset{\mathcal{F}, P}{\Omega} \arrow[r, "id @"] & \underset{\mathcal{A}, \bar{P} \,=\, P\,id^{-1}}{\Omega} \arrow[r, "E(X|\mathcal{A}) @"] & \underset{\mathcal{B}(\bar{\mathbb{R}})}{\bar{\mathbb{R}}} \end{tikzcd}$$

we have

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

and

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

$$\overset{c.v.}{=} \int_A \underbrace{E(X|\mathcal{A})}_{\frac{d\nu}{d\bar{P}}}\,d\bar{P}$$

$$= \int_A X\,dP \leftarrow \nu(A)$$

as was to be shown.

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

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

$\therefore$ $\nu(\cdot) := \int_\cdot X^+\,dP$ is a finite measure on $\mathcal{A}$

$\therefore$ $E(X^+|\mathcal{A}) := \frac{d\nu}{d\bar{P}} \in L_1(\Omega, \mathcal{A}, \bar{P})$ by Thm "props of RND"

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

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

and $\forall A \in \mathcal{A}$

$$\int_A E(X|\mathcal{A})\,dP = \int_A \underbrace{E(X^+|\mathcal{A})}_{\in L_1}\,dP - \int_A E(X^-|\mathcal{A})\,dP$$

by Big 3

$$= \int_A X^+ dP - \int_A X^- dP$$

$$= \int_A X\,dP$$

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

For uniqueness suppose $\widetilde{E(X|\mathcal{A})}$ is another version.

Now by Thm on indefinite integrals in Lecture 11

$$\int_\cdot \widetilde{E(X|\mathcal{A})}\,dP \overset{(ii)}{=} \int_\cdot E(X|\mathcal{A})\,dP \quad \text{on } \mathcal{A}$$

$$\Rightarrow \widetilde{E(X|\mathcal{A})} = E(X|\mathcal{A}) \quad \bar{P}\text{-a.e}$$

$$\Rightarrow \widetilde{E(X|\mathcal{A})} = E(X|\mathcal{A}) \quad P\text{-a.e}$$

QED

<u>Remark</u>: By construction we have

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

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

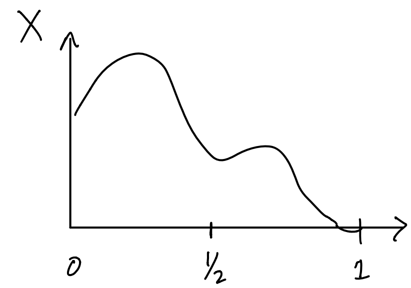
<u>Remark</u>: One way to think about $E(X|\mathcal{A})$ is

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

A slightly more vigorous view is to think of $E(X|\mathcal{A})(\omega)$ as the weighted average of X over the "smallest $\mathcal{A}$-set" containing ω, i.e. a smoothing or granulation of X, or a projection of X onto the space of $\mathcal{A}$-measurable functions.

<u>Example</u>:

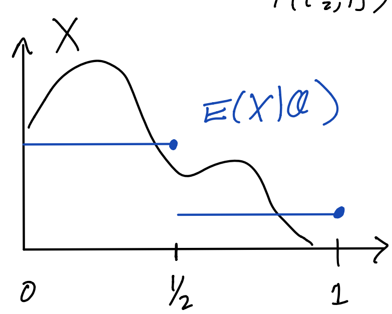
$$(\Omega,\mathcal{F},P) = ((0,1], \mathcal{B}((0,1]), P)$$



$$\mathcal{A} = \{\phi, \Omega, (0,\tfrac{1}{2}], (\tfrac{1}{2},1]\}$$

Guess at $E(X|\mathcal{A})$ & show it has the correct properties

$$E(X|\mathcal{A})(\omega) := \begin{cases} \dfrac{E(I_{(0,\frac{1}{2}]}X)}{P((0,\frac{1}{2}])} & \text{if } \omega \in (0,\tfrac{1}{2}] \\[2ex] \dfrac{E(I_{(\frac{1}{2},1]}X)}{P((\frac{1}{2},1])} & \text{if } \omega \in (\tfrac{1}{2},1] \end{cases}$$



$E(X|\mathcal{A})$ is $\mathcal{A}$-measurable (it's a simple function w.r.t $\mathcal{A}$)

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

Also if $A = (0,\tfrac{1}{2}]$

$$\int_A E(X|\mathcal{A})\,dP = \frac{E(I_{(0,\frac{1}{2}]}X)}{\cancel{P((0,\frac{1}{2}])}}\,\cancel{P(A)}$$

$$= \int_A X\,dP$$

and similarly for $A = \phi, \Omega$ or $(\tfrac{1}{2},1]$.

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

Example

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

$$\{\phi,\Omega\}=\mathcal{F}_0 \subset \mathcal{F}_1 \subset \cdots \subset \mathcal{F}$$

is an increasing sequence of sub-σ-fields. Then

$$\underset{E(X)}{\underset{\parallel}{E(X|\mathcal{F}_0)}},\ E(X|\mathcal{F}_1), \ldots, \underset{X}{\underset{\parallel}{E(X|\mathcal{F})}}$$

$\longrightarrow$ increasing resolution approx to X

Example:

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

Define

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

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

The projection $\mathcal{P}_S X$ satisfies

$$X - \mathcal{P}_S X \perp W \quad \forall W \in S$$

$$\therefore \quad E\big((X-\mathcal{P}_S X)W\big)=0 \quad \forall W \in S$$

$$\therefore \quad E(XW) = E(\mathcal{P}_S X\, W) \quad \forall W \in S$$

Given $A \in \mathcal{A}$, set $W = I_A \in S$ so that

$$\int_A X\,dP = \int_A \mathcal{P}_S X\,dP$$

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

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

Example

$\Omega = [-1,1]$

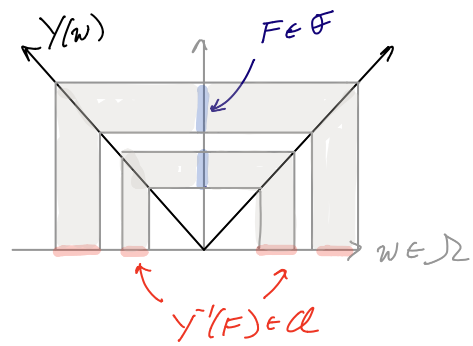
$\mathcal{F} = \mathcal{B}([-1,1])$

$dP = \delta(\omega)\,d\omega$

Let $X \in \mathcal{N}(\Omega,\mathcal{F},P)$ and define $Y(\omega)=|\omega|$ on $\omega \in \Omega$ so that

$$\sigma\langle Y\rangle = Y^{-1}(\mathcal{F}) \subset \mathcal{F}$$

Pull back of a single ⓜ map



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

$E(X|\sigma\langle Y\rangle)$ ⓜ $\sigma\langle Y\rangle$

$\Updownarrow$ Lecture 8

$E(X|\sigma\langle Y\rangle)(\omega) = g(Y(\omega))$

$= g(|\omega|)$

$$\begin{tikzcd}
\underset{\sigma\langle Y\rangle\subset\mathcal{F}}{\Omega}\ P \arrow[r, "E(X|\sigma\langle Y\rangle)", "ⓜ\ \sigma\langle Y\rangle"'] \arrow[d, "Y"] & \overline{\mathbb{R}}\ \mathcal{B}(\overline{\mathbb{R}}) \\
\overline{\mathbb{R}}\ \mathcal{B}(\overline{\mathbb{R}}) \arrow[ur, "g\ ⓜ"'] &
\end{tikzcd}$$

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

$$E(X|\sigma\langle Y\rangle)(\omega) = X(\omega)\frac{\delta(\omega)}{\delta(\omega)+\delta(-\omega)} + X(-\omega)\frac{\delta(-\omega)}{\delta(\omega)+\delta(-\omega)}$$

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

Since this $E(X|\sigma\langle Y\rangle)$ is a @ function of $|w|$ it is $\sigma\langle Y\rangle$-measurable.

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

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

Indeed if $A \in \sigma\langle Y\rangle$ then $A = -A$ so

$$\int_A E(X|\sigma\langle Y\rangle)dP$$

$$= \int_A X(w)\frac{f(w)}{f(w)+f(-w)} f(w)dw$$

$$+ \underbrace{\int_A X(-w)\frac{f(-w)}{f(w)+f(-w)} f(w)dw}_{\text{By change of variables this is } \int_{-A} X(w)\frac{f(w)}{f(w)+f(-w)} f(-w)dw}$$

$$= \int_A X(w)\frac{f(w)}{\cancel{f(w)+f(-w)}}[\cancel{f(w)+f(-w)}]dw$$

$$= \int_A X dP$$

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 $X, Y, X_1, X_2, \ldots$ are quasi-integrable extended r.v.s on a prob space $(\Omega,\mathcal{F},P)$ & $\mathcal{Q}, \mathcal{Q}_1, \mathcal{Q}_2, \ldots$ are sub-σ-fields of $\mathcal{F}$.

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

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

(s2) if $\mathcal{Q}_1 \subset \mathcal{Q}_2$ (i.e. $\mathcal{Q}_1$ smooths more) then

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

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

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

and in particular

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

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

we will also give the result for the expected value properties of $E(X|\mathcal{Q})$ & then prove both together (since proving (s4), for example, uses the following linearity results).

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

(e1) $E(cX|\mathcal{Q}) \overset{P\text{-a.e.}}{=} cE(X|\mathcal{Q}) \quad \forall c \in \mathbb{R}$.

(e2) $E(X+Y|\mathcal{Q}) = E(X|\mathcal{Q}) + E(Y|\mathcal{Q})$ P-a.e. on the set where $E(X|\mathcal{Q})+E(Y|\mathcal{Q})$ is defined (i.e. not $\pm\infty \mp\infty$).

> 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.

(e3) $X \overset{P\text{-a.e.}}{\leq} Y \Rightarrow E(X|\mathcal{Q}) \overset{P\text{-a.e.}}{\leq} E(Y|\mathcal{Q})$

(e4) $|E(X|\mathcal{Q})| \leq E(|X|\,|\mathcal{Q})$

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

MCT

$0 \leq E(X_n|\mathcal{Q}) \uparrow E(X|\mathcal{Q})$ P-a.e.

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

Fatou

$E(\liminf_n X_n|\mathcal{Q}) \leq \liminf_n E(X_n|\mathcal{Q})$

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

DCT

$E(X_n|\mathcal{Q}) \xrightarrow{a.e.} E(X|\mathcal{Q})$

on the set $\{E(\sup_n |X_n|\,|\mathcal{Q}) < \infty\}$

Before we prove the above two theorems we need the following lemma.

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

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

Proof:

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

Also $\forall \tilde{A} \in \mathcal{Q}$

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

$$\overset{\text{def}}{=} \int_{A\cap\tilde{A}} X dP \quad \text{since } A\cap\tilde{A} \in \mathcal{Q}$$

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

$$\overset{\text{def}}{=} \int_{\tilde{A}} E(I_A X|\mathcal{Q}) dP$$

$\therefore E(I_A X|\mathcal{Q}) \overset{P\text{-a.e.}}{=} I_A E(X|\mathcal{Q})$ by uniqueness.

QED

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

(s1) follows simply by the requirement:

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

(s2) we show $E(X|\mathcal{Q}_1)$ is a version of $E(E(X|\mathcal{Q}_1)|\mathcal{Q}_2)$ and $E(E(X|\mathcal{Q}_2)|\mathcal{Q}_1)$.

Note $E(X|\mathcal{Q}_1)$ is both $\mathcal{Q}_1$ & $\mathcal{Q}_2$ 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

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

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

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

$$\overset{\text{def}}{=} \int_A E(X|\mathcal{Q}_2)\,dP \quad \text{since } A\in\mathcal{Q}_1\subset\mathcal{Q}_2$$

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

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

(s3) This is just what the remark was following the Theorem establishing the existance of $E(X|\mathcal{Q})$.

(e1) This clearly follows from Big 3.

(e2): We show if $X, Y, X+Y \in \mathcal{Q}(\Omega,\mathcal{F},P)$ then $I_A\left[E(X+Y|\mathcal{Q})\right] \overset{P\text{-a.e.}}{=} I_A E(X|\mathcal{Q}) + I_A E(Y|\mathcal{Q})$ where $A := \{E(X|\mathcal{Q})+E(Y|\mathcal{Q}) \text{ is defined}\}$.

equals $I_A[E(X|\mathcal{Q})+E(Y|\mathcal{Q})]$ if the sum is defined.

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

Now by (s3) $E(X|\mathcal{Q})\in\mathcal{Q}^-$ & $E(Y|\mathcal{Q})\in\mathcal{Q}^-$ so we may change them on P-null sets so that $E(X|\mathcal{Q})>-\infty$ & $E(Y|\mathcal{Q})>-\infty$ everywhere on Ω.

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

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

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

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

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

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

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

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

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

Now the previous lemma shows

$$E(I_{A_c}X|\mathcal{Q}) \overset{P\text{-a.e.}}{=} I_{A_c}E(X|\mathcal{Q}) \ge c > -\infty$$

$$E(I_{A_c}Y|\mathcal{Q}) \overset{P\text{-a.e.}}{=} I_{A_c}E(Y|\mathcal{Q}) \ge c > -\infty$$

The previous case now applies by (s3) & gives

$$\underbrace{E(I_{A_c}X + I_{A_c}Y|\mathcal{Q})}_{\overset{\text{Lemma}}{\|}\ I_{A_c}E(X+Y|\mathcal{Q})} \overset{P\text{-a.e.}}{=} \underbrace{E(I_{A_c}X|\mathcal{Q}) + E(I_{A_c}Y|\mathcal{Q})}_{\overset{\text{Lemma}}{\|}\ I_{A_c}E(X|\mathcal{Q}) + I_{A_c}E(Y|\mathcal{Q})}$$

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

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

A similar argument shows

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

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

15

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

(e_3): show $X \overset{P\text{-a.e.}}{\leq} Y \Rightarrow E(X|\mathcal{Q}) \overset{P\text{-a.e.}}{\leq} E(Y|\mathcal{Q})$.

This follows from the indefinite integral results in Lecture 11. Indeed

$$\int_A X\,dP \leq \int_A X\,dP \quad \forall A \in \mathcal{Q}$$

$$\Longrightarrow \int_A E(X|\mathcal{Q})\,dP \leq \int_A E(Y|\mathcal{Q})\,dP, \quad \forall A \in \mathcal{Q}$$

$$\overset{\text{Lecture 11}}{\Longrightarrow} E(X|\mathcal{Q}) \overset{P\text{-a.e.}}{\leq} E(Y|\mathcal{Q})$$

(e_4): show $|E(X|\mathcal{Q})| \leq E(|X|\,|\mathcal{Q})$

This follows from (e_3) since it implies

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

(e_5): show if $0 \leq X_n \uparrow X$ P-a.e. then

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

Note that (e_3) establishes that

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

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

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

We show $\lim_n E(X_n|\mathcal{Q}) \overset{P\text{-a.e.}}{=} E(X|\mathcal{Q})$.

16

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

$$\int_A \lim_n E(X_n|\mathcal{Q})\,dP \overset{MCT}{=} \lim_n \int_A E(X_n|\mathcal{Q})\,dP$$

$$\overset{\text{def}}{=} \lim_n \int_A X_n\,dP$$

$$\overset{MCT}{=} \int_A \underbrace{\lim_n X_n}_{X}\,dP$$

$\therefore$ by the characterizing properties of $E(X|\mathcal{Q})$ we have

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

(e_6) Fatou. This is similar to (e_5).

(e_7) show that if $X_n \xrightarrow{a.e.} X$ & $X, X_n \in \mathcal{Q}(\Omega,\mathcal{F},P)$ then

$$I_A E(X_n|\mathcal{Q}) \xrightarrow{a.e.} I_A E(X|\mathcal{Q})$$

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

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

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

Lemma

$$\therefore E(\sup_n |I_{A_c} X_n|\,|\mathcal{Q})$$

$$\overset{P\text{.a.e}}{=} I_{A_c} E(\sup_n |X_n|\,|\mathcal{Q})$$

$$\leq c$$

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

$\therefore \sup_n |I_{A_c} X_n| \in L_1(\Omega,\mathcal{F},P)$ by (s_3)

(17)

Now setting $Y = \sup_n |I_{A_c} X_n|$

$$E\Big(\liminf_n \big(\underbrace{Y + I_{A_c} X_n}_{\geq 0}\big)\Big|\mathcal{Q}\Big) \overset{P\text{-a.e.}}{\leq} \liminf_n E\big(Y + I_{A_c} X_n \big| \mathcal{Q}\big)$$

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

$$E\big(\liminf_n I_{A_c} X_n \big| \mathcal{Q}\big) \overset{P\text{-a.e.}}{\leq} \liminf_n E\big(I_{A_c} X_n \big| \mathcal{Q}\big)$$

Therefore

$$I_{A_c} E\big(\liminf_n X_n \big| \mathcal{Q}\big) \overset{P\text{-a.e.}}{\leq} I_{A_c} \liminf_n E(X_n|\mathcal{Q})$$

similarly

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

so that

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

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

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

(18)

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

Case 1: $X, Y \in \mathcal{M}(\Omega, \mathcal{F}, P)$ & $X \;ⓜ\, \mathcal{Q}$

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

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

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

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

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

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

$$e5 \Big\downarrow n \to \infty \qquad\qquad e5 \Big\downarrow n \to \infty$$

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

Since $X E(Y|\mathcal{Q})$ is $ⓜ\,\mathcal{Q}$ by closure & in $\mathcal{M}(\Omega, \mathcal{Q})$ we have $X E(Y|\mathcal{Q}) \overset{P\text{-a.e.}}{=} E(XY|\mathcal{Q})$

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

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

$$= \underbrace{(X^+Y^+ + X^-Y^-)}_{=(XY)^+} - \underbrace{(X^-Y^+ + X^+Y^-)}_{=(XY)^-}$$

$$\therefore \; E(XY|\mathcal{Q}) \overset{\text{def}}{=} E((XY)^+|\mathcal{Q}) - E((XY)^-|\mathcal{Q})$$

$$\underset{(e_2)}{\overset{P\text{-a.e.}}{=}} E(X^+Y^+|\mathcal{Q}) + E(X^-Y^-|\mathcal{Q})$$

$$- E(X^-Y^+|\mathcal{Q}) - E(X^+Y^-|\mathcal{Q})$$

$$\underset{\text{case } 1}{\overset{P\text{-a.e.}}{=}} X^+E(Y^+|\mathcal{Q}) + X^-E(Y^-|\mathcal{Q})$$

$$- X^-E(Y^+|\mathcal{Q}) - X^+E(Y^-|\mathcal{Q})$$

$$= (X^+ - X^-)(E(Y^+|\mathcal{Q}) - E(Y^-|\mathcal{Q}))$$

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

QED

Remark: Notice the technique for showing the linearity (e_2) & the DCT (e_7)... regularize with I_{A_c}, use the Lemma to show $I_{A_c}E(X|\mathcal{Q}) \overset{a.e.}{=} I_{A_c}E(Y|\mathcal{Q})$ then take a countable union over A_c sets.

Lets do a quick example that shows the generality of (e_2).

Example:

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

$\mathcal{Q} = \{\phi, \Omega, (0,\tfrac{1}{2}], (\tfrac{1}{2},1]\}$

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

Now P-a.e. we have

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

and

$$I_{(\frac{1}{2},1]}E(X-X|\mathcal{Q}) \overset{a.e.}{=} I_{(\frac{1}{2},1]}E(X|\mathcal{Q}) + I_{(\frac{1}{2},1]}E(-X|\mathcal{Q})$$

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

The next result is a useful tool for showing $E(X|\sigma\langle\mathcal{P}\rangle)$ has a particular form.

Theorem: (π-generators are enough)

Let $\mathcal{P} \subset \mathcal{F}$ be a π-system and $X \in L_1(\Omega, \mathcal{F}, P)$. If $Y \in L_1(\Omega, \sigma\langle\mathcal{P}\rangle, P)$ s.t. $\forall A \in \mathcal{P}$

$$\int_A X\,dP = \int_A Y\,dP \qquad (*)$$

then $Y \overset{P\text{-a.e.}}{=} E(X|\sigma\langle\mathcal{P}\rangle)$.

Proof:

If $(*)$ holds then

$$\int_\cdot X^{\pm}\,dP = \int_\cdot Y^{\pm}\,dP \quad \text{on } \mathcal{P}$$

Since $\int_\cdot X^\pm dP$ & $\int_\cdot Y^\pm dP$ are both finite measures on $(\Omega, \sigma\langle \mathcal{P}\rangle)$ which agree on a π-system $\mathcal{P}$ we have

$$\infty > \int_\cdot X^\pm dP = \int_\cdot Y^\pm dP \quad \text{on } \sigma\langle \mathcal{P}\rangle$$

$\therefore$ $\int_\cdot X dP = \int_\cdot Y dP$ on $\sigma\langle \mathcal{P}\rangle$.

$\therefore$ $Y \overset{P\text{-a.e.}}{=} E(X|\sigma\langle\mathcal{P}\rangle)$ QED

Remark: The above claim can be extended if it is a-priori known that $\int_\cdot X^\pm dP$ & $\int_\cdot Y^\pm dP$ 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

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

Notice that by definition we have

$$E(X|\sigma\langle Y_1,Y_2\dots\rangle) \text{ ⓜ } \sigma\langle Y_1,Y_2,\dots\rangle$$

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

$$E(X|Y_1,Y_2,\dots) \overset{P\text{-a.e.}}{=} g(Y_1,Y_2,\dots)$$

by "ⓜ fns of ⓜ fns Thm" in Lecture 9.

In general, the notation $E(X|Y_1=y_1,Y_2=y_2,\dots)$ means $g(y_1,y_2,\dots)$ 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

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

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

$$= \int_{\vec{Y}^{-1}(B)} g(\vec{Y})\,dP$$

$\sigma\langle Y_1,Y_2,\dots\rangle$ sets

$$\overset{def}{=} \int_{\vec{Y}^{-1}(B)} X\,dP$$

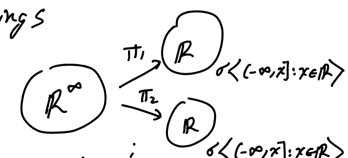
$$= E(I_{(Y_1,Y_2,\dots)\in B}\, X) \qquad (*)$$

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

Also recall that

coordinate projection mappings

$$\mathcal{B}(\mathbb{R}^\infty) = \sigma\langle \pi_1,\pi_2,\dots\rangle$$



$$= \sigma\langle \pi_i^{-1}((-\infty,x]) : i\in\mathbb{N},\ x\in\mathbb{R}\rangle$$

by generators are enough

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

Again by generators are enough we have

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

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

$$= \sigma\langle \underbrace{\{Y_{i_1}\le x_1,\dots,Y_{i_n}\le x_n\}}_{\pi\text{-system}} : \begin{matrix} n\in\mathbb{N} \\ x_i\in\mathbb{R}\end{matrix}\rangle$$

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

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

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

$$E\left(I_{\{Y_{i_1} \le x_1\}} \cdots I_{\{Y_{i_n} \le x_n\}} \, g(Y_1, Y_2, \dots)\right) \overset{(*)}{=} E\left(I_{\{Y_{i_1} \le x_1\}} \cdots I_{\{Y_{i_n} \le x_n\}} X\right)$$

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

Example:

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

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

has the property that

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

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

$$\underbrace{E\Big(I_{\{Y_1=y_1, \dots, Y_n=y_n\}}}_{\pi\text{-sys generators}} \underbrace{X}_{=\sum_{x \in R} x I_{X=x}}\Big)$$

$$= \sum_{x \in R} x \, P(X=x, Y_1=y_1, \dots, Y_n=y_n)$$

$$= \begin{cases} \sum_{x \in R} x \dfrac{P(X=x, Y_1=y_1, \dots, Y_n=y_n)}{P(Y_1=y_1, \dots, Y_n=y_n)} \underbrace{P(Y_1=y_1, \dots, Y_n=y_n)}_{\text{when this is } >0} \\ 0 \quad \text{o.w.} \end{cases}$$

$$= g(y_1, \dots, y_n) P(Y_1=y_1, \dots, Y_n=y_n)$$

$$= E\left(I_{\{Y_1=y_1, \dots, Y_n=y_n\}} \, g(Y_1, \dots, Y_n)\right)$$

and g is trivially Ⓜ and integrable as its a discrete map.