Conditioning on a sub-σ-algebra (R § 13.2, B § 34)

- So far, only conditioned on a single RV X.
- Equivalently, this is conditioning on the σ-algebra generated by X: $\sigma(X) = \{\{X \in B\} : B \subseteq \mathbb{R} \text{ Borel}\}$.
- More generally, we may wish to condition on something more than a single RV, such as a sequence of RVs $X_1, X_2, \dots, X_n$ or some other set of events.

Given a Prob. space $(\Omega, \mathcal{F}, P)$.

def: Given any sub-algebra $\mathcal{G} \subseteq \mathcal{F}$, define $P(A \mid \mathcal{G})$ and $E[Y \mid \mathcal{G}]$ to be $\mathcal{G}$-measurable RVs satisfying:

$$E[P(A \mid \mathcal{G}) \mathbb{1}_G] = P(A \cap G) \qquad (13.2.1)$$

$$E[E[Y \mid \mathcal{G}] \mathbb{1}_G] = E[Y \mathbb{1}_G] \quad \text{for any } G \in \mathcal{G}. \qquad (13.2.2)$$

Note: If $\mathcal{G} = \sigma(X)$, this definition is precisely the one we gave before.

$P(A \mid X)$ short hand notation for $P(A \mid \sigma(X))$

$E[Y \mid X]$ " " $E[Y \mid \sigma(X)]$

Similarly, we write $E[Y|X_1, X_2]$ as shorthand for $E[Y|\sigma(X_1,X_2)]$ where

$$\sigma(X_1, X_2) = \sigma(\{\{X_1 \le a\} \cap \{X_2 \le b\} : a, b \in \mathbb{R}\})$$

σ-algebra generated by X_1 & X_2.

come back to this in §14 Martingales

Do other examples 1st

Example 1: Let $\mathcal{G} = \{\phi, \Omega\}$ ← trivial σ-algebra.

Then $P(A|\mathcal{G}) = P(A)$

& $E[Y|\mathcal{G}] = E[Y]$

both constants "no information"

why? We don't know anything about Y so best guess is $E[Y]$

Example 2: Let $\mathcal{G} = \mathcal{F}$ ← full σ-algebra for $(\Omega, \mathcal{F}, P)$.

Equivalently, if $A \in \mathcal{G}$ and X is $\mathcal{G}$-measurable,

then $P(A|\mathcal{G}) = \mathbb{1}_A$

& $E[X|\mathcal{G}] = X$

"Perfect Information"

why? We know whether or not A has occurred.

Prop 13.2.6: Let X & Y be RVs & let $\mathcal{G}$ be a sub-σ-algebra. Suppose that $E[Y] < \infty$ and $E[XY] < \infty$, and that X is $\mathcal{G}$-measurable. Then with probability 1,

$$E[XY|\mathcal{G}] = X \cdot E[Y|\mathcal{G}].$$

<u>Pf</u>: First note that $X \cdot E[Y|\mathcal{G}]$ is $\mathcal{G}$-measurable (since X is by assumption & $E[Y|\mathcal{G}]$ is a $\mathcal{G}$-meas. RV by definition).

Now if $X = \mathbb{1}_{G_0}$ with $G_0 \in \mathcal{G}$, then for $G \in \mathcal{G}$

$$E\left[X \cdot E[Y|\mathcal{G}]\, \mathbb{1}_G\right] = E\left[E[Y|\mathcal{G}]\, \mathbb{1}_{G_0 \cap G}\right]$$

$$= E\left[Y\, \mathbb{1}_{G_0 \cap G}\right] \quad \text{by def'n of } E[Y|\mathcal{G}]$$

$$= E\left[XY\, \mathbb{1}_G\right] \quad \text{so Eqn (13.2.2) holds for } X = \mathbb{1}_{G_0}.$$

Can show that (13.2.2) holds for general X using linearity of expectation & monotone convergence arguments. It follows that $X \cdot E[Y|\mathcal{G}]$ is a version of $E[XY|\mathcal{G}]$.

<u>Prop 13.2.7</u>: Let Y be a RV with finite mean and let $\mathcal{G}_1 \subseteq \mathcal{G}_2$ be two sub-σ-algebras. Then with prob. 1,

$$E\left[E[Y|\mathcal{G}_2] \mid \mathcal{G}_1\right] = E[Y|\mathcal{G}_1].$$

In words, conditioning first on $\mathcal{G}_2$ & then on $\mathcal{G}_1$ is equivalent to conditioning just on the smaller sub-σ-alg. $\mathcal{G}_1$.

(R 13.4.5) Assume $P(\{1\}) = P(\{2\}) = P(\{3\}) = \frac{1}{3}$.

Example 1: Let $\Omega = \{1,2,3\}$ & $\mathcal{F} = 2^{\Omega}$. Define RVs X & Y by

$X(1) = X(2) = 5$
$X(3) = 6$
and $Y(\omega) = \omega$ $\forall \omega \in \Omega$. Let $Z = E[Y|X]$.

a) Describe $\sigma(X)$.

b) Describe $Z(\omega)$ for each $\omega \in \Omega$.

Ans: a) $\sigma(X) = \{\phi, \Omega, \{\omega : X(\omega) = 5\}, \{\omega : X(\omega) = 6\}\}$
$= \{\phi, \Omega, \{1,2\}, \{3\}\}$

b) $Z = E[Y|X] = E[Y|\sigma(X)]$ by def.

$Z(\omega) = E[Y|\sigma(X)](\omega)$

$\omega = 1$: $Z(1) = E[Y|\sigma(X)](1)$ → if $\omega = 1$, then $X = 5$ or $\omega = 2$

$= \dfrac{Y(1) + Y(2)}{2}$ → So we look at average of $Y(\omega)$ for $\omega \in \{1,2\}$

$= \dfrac{1+2}{2} = \dfrac{3}{2}$

$\omega = 2$: $Z(2) = Z(1)$ since $\underline{\{1,2\} \in \sigma(X)}$

$\omega = 3$: $Z(3) = E[Y | \sigma(X)](3)$ $\longrightarrow$ so $X = 6$

$= Y(3) = 3$

Thus, $Z(\omega) = E[Y | \sigma(X)](\omega) = \begin{cases} 3/2 & \text{if } \omega \in \{1,2\} \\ 3 & \text{if } \omega \in \{3\} \end{cases}$

this partition of Ω is given by $\sigma(X)$

* If we condition on some other σ-algebra, we'll (likely) get a different RV $Z(\omega)$.

Example 2: Suppose that $(\Omega, \mathcal{F}, P)$ is a probability space s.t. $\Omega = \{a, b, c, d, e, f\}$, $\mathcal{F} = 2^{\Omega}$, P uniform.

Let X, Y, & Z be RVs s.t.

$$X(\omega) = \begin{cases} 1 & \text{if } \omega = a \\ 3 & \text{if } \omega \in \{b, c\} \\ 5 & \text{if } \omega \in \{d, e\} \\ 7 & \text{if } \omega = f \end{cases}$$

$$Y(\omega) = \begin{cases} 2 & \text{if } \omega \in \{a, b\} \\ 1 & \text{if } \omega \in \{c, d\} \\ 7 & \text{if } \omega \in \{e, f\} \end{cases} \qquad Z(\omega) = \begin{cases} 3 & \text{if } \omega \in \{a, b, c, d\} \\ 2 & \text{if } \omega \in \{e, f\} \end{cases}$$

Q. What is $E[X \mid \mathcal{G}]$ for $\mathcal{G} \subseteq \mathcal{F}$?

The average of $X(\omega)$ for over all ω which are consistent with the current information, i.e. $\mathcal{G}$.

① Let $\mathcal{G} = \sigma(Y)$, i.e. info about exact value of Y

$$= \{\phi, \Omega, \underbrace{\{a, b\}, \{c, d\}, \{e, f\}}_{\text{these are the essential "atoms" of } \sigma(Y)}, \dots\}$$

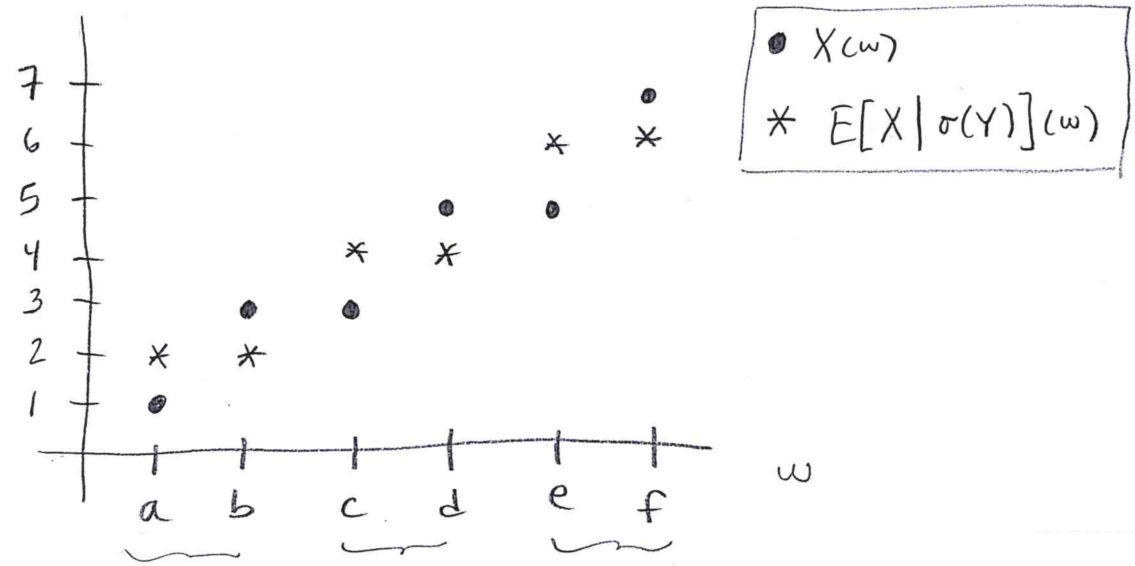
So, $E[X \mid Y = 2] = \dfrac{X(a) + X(b)}{2} = \dfrac{1 + 3}{2} = 2$

$E[X \mid Y = 1] = \dfrac{X(c) + X(d)}{2} = \dfrac{3 + 5}{2} = 4$

$E[X \mid Y = 7] = \dfrac{X(e) + X(f)}{2} = \dfrac{5 + 7}{2} = 6$

Then we write this as

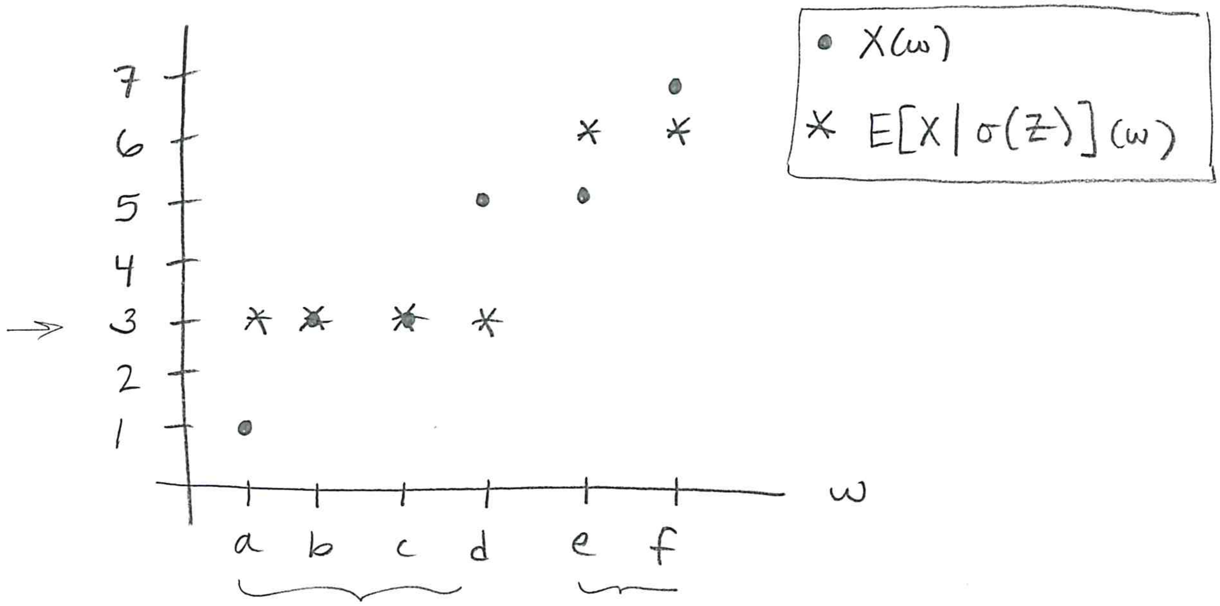
$$E[X \mid \sigma(Y)](\omega) = \begin{cases} 2 & \text{if } \omega \in \{a,b\} \\ 4 & \text{if } \omega \in \{c,d\} \\ 6 & \text{if } \omega \in \{e,f\} \end{cases}$$



(2) Now let $\mathcal{G} = \sigma(Z) = \{\phi, \Omega, \underbrace{\{a,b,c,d\}, \{e,f\}}_{\text{essential "atoms"}}\}$

Then
$$E[X \mid \sigma(Z)](\omega) = \begin{cases} \dfrac{X(a)+X(b)+X(c)+X(d)}{4} = \dfrac{1+3+3+5}{4} = 3 \\[2ex] \dfrac{X(e)+X(f)}{2} = \dfrac{5+7}{2} = 6 \end{cases}$$

$$= \begin{cases} 3 & \text{if } \omega \in \{a,b,c,d\} \\ 6 & \text{if } \omega \in \{e,f\} \end{cases}$$



Note the differences between $E[X|\sigma(Y)]$ & $E[X|\sigma(Z)]$ for $\omega \in \{a,b,c,d\}$. They are equiv for $\omega \in \{e,f\}$. (due to different groupings of ω's in $\sigma(Y)$ & $\sigma(Z)$).

Martingales

(Ref: §14 R, §35 B)

def: A stochastic process $X_0, X_1, X_2, \ldots$ is a martingale if $E[|X_n|] < \infty$ $\forall n$ and with probability 1

$$\boxed{E[X_{n+1} \mid X_0, X_1, \ldots, X_n] = X_n}.$$

i.e. $E[X_{n+1} \mid \mathcal{F}_n] = X_n$ where $\mathcal{F}_n = \sigma(X_0, X_1, \ldots, X_n)$

In words, this says that on average, the value of X_{n+1} is the same as that of X_n.

- Markov chains often provide good examples of martingales

  although $\exists$ non-Markovian martingales

- Simple symmetric RW on $\mathbb{Z}$ where $X_0 = 0$ &

  $P_{i,i-1} = P_{i,i+1} = \frac{1}{2}$ $\forall i \in \mathbb{Z}$ ← is a martingale.

  (actually this is true in any dimension, $\mathbb{Z}^d$)