LECTURE 6

Two "elementary" examples of conditional probability/expectation:

1. (X,Y) have finitely or countably many values, $p(x,y) = \mathbb{P}[X=x, Y=y]$

$$\mathbb{P}[X=x \mid Y=y] := \frac{p(x,y)}{\sum_{x'} p(x',y)}$$

$$\mathbb{E}[X \mid Y=y] := \sum_x x \cdot \mathbb{P}[X=x \mid Y=y] = f(y)$$

2. $(X,Y) \in \mathbb{R}^2$ have continuous density $p(x,y) > 0$

$$p(x \mid y) := \frac{p(x,y)}{\int p(x',y)\,dx'}$$

$$\mathbb{E}[X \mid Y=y] := \int x\, p(x \mid y)\,dx = f(y)$$

In these cases, $\mathbb{E}[X \mid Y=y] = f(y)$ is a function.

3. A, B events, $\mathbb{P}[B] > 0 \rightsquigarrow \mathbb{P}[A \mid B] := \frac{\mathbb{P}[A \cap B]}{\mathbb{P}[B]}$.
X r.v. $\rightsquigarrow \mathbb{E}[X \mid B] := \frac{\mathbb{E}[X \mathbb{1}_B]}{\mathbb{P}[B]} \in \mathbb{R}$.

<u>Goal</u>: A general framework including all of these, + more situations.
First step: unify discrete 1 and continuous 2.
View $\mathbb{E}[X \mid Y] = f(Y)$ as <u>random variable</u> in 1, 2.

<u>Prop</u>: In 1, 2, if $(\Omega, \mathcal{F}, \mathbb{P})$ underlying prob. space, then $\mathbb{E}[X \mid Y]$ is $\sigma(Y)$-measurable $(\sigma(Y) := \{Y^{-1}(B) : B \text{ Borel}\} \subseteq \mathcal{F})$

<u>Pf</u>: By construction, $\mathbb{E}[X \mid Y] = f(Y)$ for $f: \mathbb{R} \to \mathbb{R}$ measurable.

<u>Prop</u>: If E is $\sigma(Y)$-measurable, then $E = f(Y)$ for some measurable f.

Rk: In Kolmogorov philosophy, σ-algebras describe information.
$\mathcal{G} \subseteq \mathcal{F}$ sub-algebra = "events whose occurrence can be decided by ..."
$\mathcal{G} = \{\emptyset, \Omega\} \rightsquigarrow$ no info $\qquad \mathcal{G} = \sigma(Y) \rightsquigarrow$ value of Y.

Prop: If E, E' $\mathcal{G}$-m'ble r.v.'s and $\mathbb{E}\, E \mathbb{1}_A = \mathbb{E}\, E' \mathbb{1}_A \ \forall A \in \mathcal{G}$, then $E = E'$ a.s. (i.e., values $(\mathbb{E}\, E \mathbb{1}_A)_{A \in \mathcal{G}}$ determine $\mathcal{G}$-m'ble r.v.)

Pf:
$$\mathbb{E}|E - E'| = \mathbb{E}\,(E - E')\mathbb{1}_{\{E - E' > 0\}} + \mathbb{E}\,(E' - E)\mathbb{1}_{\{E' - E > 0\}} = 0. \qquad \{\cdot\} \in \mathcal{G}$$

How does this "sufficient data" look for $E = \mathbb{E}[X|Y]$ in ①, ②?

①: $Y \in \{y_1, y_2, \dots\}$ $\quad A \in \sigma(Y) \rightsquigarrow A = \{Y \in B\}$

$$\mathbb{E}[E\mathbb{1}_A] = \mathbb{E}[f(Y)\mathbb{1}\{Y \in B\}]$$
$$= \sum_{y \in B} f(y) \underbrace{\sum_{x'} p(x', y)}_{\mathbb{P}[Y=y]} = \sum_{y \in B} \sum_x x \cdot p(x,y)$$
$$= \mathbb{E}[X\mathbb{1}_A]$$

Rk: As usual, measure theory language unites discrete + continuous.

②: sums $\rightarrow$ integrals, with same conclusion.

Findings: in both cases, $E = \mathbb{E}[X|Y]$ satisfies:
- E $\sigma(Y)$-measurable
- $\forall A \in \sigma(Y)$, $\mathbb{E}[E\mathbb{1}_A] = \mathbb{E}[X\mathbb{1}_A]$

which together (by Prop) determine E a.s.

Thm: $(\Omega, \mathcal{F}, \mathbb{P})$ prob. space, X r.v. $\mathbb{E}|X| < \infty$, $\mathcal{G} \subseteq \mathcal{F}$ subalg.
Then, $\exists$ E r.v. such that:
1. E is $\mathcal{G}$-measurable
2. $\forall A \in \mathcal{G}$, $\mathbb{E}[E \mathbb{1}_A] = \mathbb{E}[X \mathbb{1}_A]$.
3. $\mathbb{E}|E| < \infty$.

For any E, E' w/ 1, 2, 3, $E = E'$ a.s. (by Prop.)

Def: Such $E =: \mathbb{E}[X | \mathcal{G}]$, $\mathbb{E}[X|Y] := \mathbb{E}[X | \sigma(Y)]$.

Rk: $\mathcal{G}$ smaller $\rightsquigarrow$ $\mathcal{G}$-measurable E "less random" (more restrictions)

Ex: Recovers discrete + cts. definitions.

Ex: If $\mathbb{P}[B] > 0$, $\mathbb{E}[X | \{\emptyset, B, B^c, \Omega\}]$ = r.v. depending only on whether $\omega \in A$. $= E(\omega)$

$$E(\omega) = \begin{cases} \frac{\mathbb{E}[X \mathbb{1}_A]}{\mathbb{P}(A)} = \mathbb{E}[X|A] & \text{if } \omega \in A \\ \frac{\mathbb{E}[X \mathbb{1}_{A^c}]}{\mathbb{P}[A^c]} = \mathbb{E}[X|A^c] & \text{if } \omega \in A^c \end{cases}$$

$\rightsquigarrow$ definition also captures $\mathbb{E}$ conditional on events.

Ex: $\mathbb{E}[X | \{\emptyset, \Omega\}]$ = r.v. "depending on nothing" = const. = $\mathbb{E}[X]$.

Ex: X is $\mathcal{G}$-measurable $\rightsquigarrow$ $\mathbb{E}[X | \mathcal{G}] = X$.

Ex: Suppose X, Y indep. Then $\mathbb{E}[X|Y] = \mathbb{E}[X] =: E$

Pf: Say $A \in \sigma(Y) \longrightarrow A = \{Y \in B\}$.

$\mathbb{E}[E \mathbb{1}_A] = \mathbb{E}[X] \mathbb{E}[\mathbb{1}_A] = \mathbb{E}[X] \cdot \mathbb{P}[Y \in B]$.

$\mathbb{E}[X \mathbb{1}_A] = \mathbb{E}[X \mathbb{1}\{Y \in B\}] = \mathbb{E}[X] \cdot \mathbb{P}[Y \in B]$.

indep. (fn. of indep. X, Y)

Pf: (of Thm) Uniqueness: by Prop. 1 + 2 $\Longrightarrow$ 3:

$$E|E| = E[E \mathbb{1}\{\underbrace{E>0}_{\in \mathcal{G} \text{ by } 1}\}] - E[E \mathbb{1}\{E<0\}]$$

$$\overset{2}{=} E[X \mathbb{1}\{E>0\}] - E[X \mathbb{1}\{E<0\}] \leq E|X|.$$

Tool for existence:

Thm: (Radon-Nikodym) μ, ν finite measures on $(\Omega, \mathcal{G})$ with $\nu \ll \mu$, i.e. $\mu(A)=0 \Longrightarrow \nu(A)=0$

Then $\exists$ $\mathcal{G}$-measurable $f \geq 0$ s.t. $\int_A f \, d\mu = \nu(A) \ \forall A \in \mathcal{G}$.

$f =: \frac{d\nu}{d\mu}$ ("relative density" of measures.)

Ex: Some more work $\rightsquigarrow$ any prob. measure on $\mathbb{R}$ = mixture of density and point masses.

Rk: Given μ, ν on $(\Omega, \mathcal{F})$, can apply RN to $(\Omega, \mathcal{G})$ $\rightsquigarrow$ gives different $f = f_{\mathcal{G}}$: $\mathcal{G}$-measurable cond. stricter but $\forall A \in \mathcal{G}$ cond. looser.

Case $X \geq 0$: $\nu(A) := E X \mathbb{1}_A = \int_A X(\omega) \, d\mathbb{P}(\omega)$

Exc: ν is finite measure on $(\Omega, \mathcal{F})$, $\nu \ll \mathbb{P}$. $\left(X = \frac{d\nu}{d\mathbb{P}}\right)$.

RN on $(\Omega, \mathcal{G})$ $\rightsquigarrow$ $E(\omega)$ $\mathcal{G}$-measurable, $\int_A E \, d\mathbb{P} = \nu(A) = \int_A X \, d\mathbb{P}$ $\forall A \in \mathcal{G}$.

General case: $X = X^+ - X^-$, $E[X|\mathcal{G}] := E[X^+|\mathcal{G}] - E[X^-|\mathcal{G}]$