Lecture 6 Wed Feb/07/2024

Last time
- General Poisson Convergence
- Conditional Expectation

Today
- Continuation of the proof
- Properties
- Regular Conditional probability

Existence:
Consider two cases:

Case 1: $X \geq 0$. Then, we can define

$$\nu(A) = \int_A X \, d\mathbb{P} \qquad \forall A \in \mathcal{G}$$

DCT

Exercise: Show that ν is a measure on $(\Omega, \mathcal{G})$ [just like $\mathbb{P}$]. Now we would like to find a $\mathcal{G}$-measurable function f such that

$$\nu(A) = \int_A f \, d\mathbb{P}.$$

Density.

Question: When do such densities exist?

We use a hammer from measure Theory:

Theorem (Radon-Nikodym) Let μ and ν be σ-finite measures on $(\Omega, \mathcal{F})$. If $\nu \ll \mu$, i.e.,

$$\mu(A) = 0 \Rightarrow \nu(A) = 0 \quad \text{for } A \in \mathcal{G},$$

then, there exists a measurable function f such that $\forall A \in \mathcal{G}$

$$\nu(A) = \int_A f \, d\mu .$$

The function f is often denoted $\frac{d\nu}{d\mu}$. $\dashv$

This result is immediate for $X \geq 0$, $E[X|\mathcal{G}] = d\nu / d\mathbb{P}$.

Case 2. Decompose $X = X^+ - X^-$. Let $Y^+ = \mathbb{E}[X^+|\mathcal{G}]$ and $Y^- = \mathbb{E}[X^-|\mathcal{G}]$. Then, $Y = Y^+ - Y^-$ is measurable and for $A \in \mathcal{G}$

$$\int_A X \, d\mathbb{P} = \int_A X^+ \, d\mathbb{P} - \int_A X^- \, d\mathbb{P}$$

$$= \int_A Y^+ \, d\mathbb{P} - \int Y^- \, d\mathbb{P}$$

$$= \int_A Y \, d\mathbb{P}.$$

$\square$

Properties

In order to work with conditional Expectations is important to have a list of "valid operations".

Theorem: Let $X_1, X_2, \dots$ be r.v. with $\mathbb{E}|X_i| < \infty$. Let $\mathcal{G}$ and $\mathcal{H}$ be sub-σ-algebras of $\mathcal{F}$. Then

a) $\mathbb{E}[\mathbb{E}[X \mid \mathcal{G}]] = \mathbb{E}[X]$.

b) If X is $\mathcal{G}$-measurable $\Rightarrow$ $X = \mathbb{E}[X \mid \mathcal{G}]$ a.s.

c) (Linearity) Let $a_1, a_2 \in \mathbb{R}$, then
$\mathbb{E}[a_1 X_1 + a_2 X_2 \mid \mathcal{G}] = a_1 \mathbb{E}[X_1 \mid \mathcal{G}] + a_2 \mathbb{E}[X_2 \mid \mathcal{G}]$. almost surely.

d) (Positivity) If $X \geq 0 \Rightarrow \mathbb{E}[X \mid \mathcal{G}] \geq 0$.

e) (C. Montone) If $0 < X_n \uparrow X$, then

$$\mathbb{E}[X_n \mid \mathcal{G}] \uparrow \mathbb{E}[X \mid \mathcal{G}] \quad \text{a.s.}$$

f) (C. Fatou) If $X_n \geq 0$ then

lower semiconti nuity.

$$\mathbb{E}[\liminf X_n \mid \mathcal{G}] \leq \liminf \mathbb{E}[X_n \mid \mathcal{G}]$$

g) (C. Dominated) If $|X_n| < V$ $\forall n$ and $\mathbb{E}|V| < \infty$ and $X_n \to X$ a.s. Then,

$$\mathbb{E}[X_n|\mathcal{G}] \to \mathbb{E}[X|\mathcal{G}] \quad \text{a.s.}$$

h) (C. Jensen) Let $f: \mathbb{R} \to \mathbb{R}$ be convex.

Then, $f(\mathbb{E}[X|\mathcal{G}]) \leq \mathbb{E}[f(X)|\mathcal{G}]$ a.s.

Useful corollary: $\|X\|_p \geq \|\mathbb{E}[X|\mathcal{G}]\|_p \quad \forall p \geq 1$. a.s.

i) (Tower Law) If $\mathcal{H}$ is a sub-algebra of $\mathcal{G}$, then

$$\mathbb{E}[\mathbb{E}[X|\mathcal{G}]|\mathcal{H}] = \mathbb{E}[X|\mathcal{H}] \quad \text{a.s.}$$

What happens if we swap these two here?

j) (Taking out what is known) If Z is $\mathcal{G}$ measurable and bounded

$$\mathbb{E}[ZX|\mathcal{G}] = Z\,\mathbb{E}[X|\mathcal{G}] \quad \text{a.s.}$$

k) (Independence) If $\sigma(X)$ is ind. of $\mathcal{G}$

$$\Rightarrow \quad \mathbb{E}[X|\mathcal{G}] = \mathbb{E}[X].$$

Proof:

a) It's a consequence of i) with $\mathcal{H} = \{\emptyset, \Omega\}$.

b) Proved in Example 3 of previous lecture.

c) We prove that $a_1\mathbb{E}[X_1|\mathcal{G}] + a_2\mathbb{E}[X_2|\mathcal{G}]$ satisfies 1) and 2), $\mathcal{G}$-measurable functions form a linear subspace $\Rightarrow$ 1) holds. On the other hand, $\forall E \in \mathcal{G}$

$$\int_A a_1 \mathbb{E}[X_1|\mathcal{G}] + a_2 \mathbb{E}[X_2|\mathcal{G}]\, d\mathbb{P}$$

$$= a_1 \int_A \mathbb{E}[X_1|\mathcal{G}]\, d\mathbb{P} + a_2 \int \mathbb{E}[X_2|\mathcal{G}]\, d\mathbb{P}$$

$$= a_1 \int_A X_1\, d\mathbb{P} + a_2 \int_A X_2\, d\mathbb{P}$$

$$= \int_A a_1 X_1 + a_2 X_2\, d\mathbb{P}.$$

d) Consider the sets $A_n = \{-\mathbb{E}[X|\mathcal{G}] \geq n^{-1}\}$.

Then, $n^{-1}\, \mathbb{P}(\mathbb{E}[X|\mathcal{G}] \leq -n^{-1}) \leq \int_{A_n} -\mathbb{E}[X|\mathcal{G}]\, d\mathbb{P}$

$$= \int_{A_n} X\, d\mathbb{P} = 0$$

Thus $\mathbb{P}(A_n) = 0 \quad \forall n.$

e) Let $Y_n = \mathbb{E}[X_n|\mathcal{G}]$. By d) we have

$$0 \leq Y_n \uparrow.$$

Prob 1.

Define $Y = \limsup Y_n$, which is $\mathcal{G}$-measurable. Then

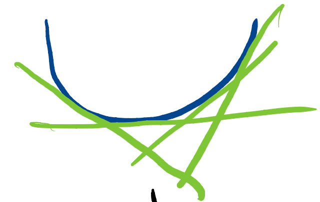
$$\int_A Y_n\, d\mathbb{P} = \int_A X_n\, d\mathbb{P} \Rightarrow \int_A Y\, d\mathbb{P} = \int_A X\, d\mathbb{P} \quad \forall A \in \mathcal{G}.$$

Monotone Convergence Thm.

f) + g) Left as an exercise.

h) Any convex function can be written as

$$\varphi(x) = \sup_{(a,b)\in\Delta} \{ax+b\}$$



where $\Delta = \{(a,b) \mid a,b\in\mathbb{R},\ ax+b \le \varphi(x)\}$.

Then, for any (a,b) we have

$$a\,\mathbb{E}[X|\mathcal{G}] + b \underset{d}{\le} \varphi(\mathbb{E}[X|\mathcal{G}])$$

Taking sup over Δ finishes the proof.
The corollary follows by taking $\varphi(x) = |x|^p$, and using a).

i) Let $A\in\mathcal{H}$, thus $A\in\mathcal{G}$ as well.

$$\Rightarrow \int_A \mathbb{E}[\mathbb{E}[X|\mathcal{G}]|\mathcal{H}]\,d\mathbb{P} = \int_A \mathbb{E}[X|\mathcal{G}]\,d\mathbb{P} = \int_A X\,d\mathbb{P}$$

j) Notice that $Z\,\mathbb{E}[X|\mathcal{G}]$ is $\mathcal{G}$-measurable.
Typical proof in measure theory. Assume $Z\ge 0$, otherwise decompose $Z = Z^+ - Z^-$. First we prove it for indicators $X = \mathbb{1}_B$ with $B\in\mathcal{G}$.

$$\Rightarrow \int_A \mathbb{1}_B\,\mathbb{E}[X|\mathcal{G}]\,d\mathbb{P} = \int_{A\cap B} \mathbb{E}[X|\mathcal{G}]\,d\mathbb{P} = \int_{A\cap B} X\,d\mathbb{P} = \int_A \mathbb{1}_B X\,d\mathbb{P}.$$

Then we extend it to simple X using c). Finally we extend it to general X taking $X_n \uparrow X$ and applying MCT.

k) Recall that $\sigma(X)$ and $\mathcal{G}$ are $\mathbb{P}$-ind if

iff $\mathbb{P}(\{X \in B\} \cap A) = \mathbb{P}(X \in B)\, \mathbb{P}(A)$.

Clearly $\mathbb{E}[X]$ is $\mathcal{G}$ measurable. Let $A \in \mathcal{G}$,

$$\int_A X \, d\mathbb{P} = \int X \mathbb{1}_A \, d\mathbb{P} = \mathbb{E}[X]\, \mathbb{P}(A)$$

$$= \int_A \mathbb{E}[X] \, d\mathbb{P}.$$

$\square$

Regular Conditional Probabilities

Question: Can we always use the conditional expectation to define well-defined conditional probability via $\mathbb{P}(A|Y) = \mathbb{E}[\mathbb{1}_A|Y]$?

Not always, but it is often the case.

Def: Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a prob. space. Let $\mathcal{G}$ be a sub-σ-algebra of $\mathcal{F}$.

Given $X : (\Omega, \mathcal{F}) \to (\mathcal{X}, \mathcal{B})$ a r.v.

A function $\mu : \Omega \times \mathcal{B} \to [0, 1]$ is called a **regular conditional probability** for X given $\mathcal{G}$

a) For $B \in \mathcal{B}$, the function $\omega \to \mu(\omega, B)$ is a version of $\mu(B \mid \mathcal{G})$.

b) For any fixed ω, $\mu(\omega, \cdot)$ is a probability measure on $(\mathcal{X}, \mathcal{B})$.

The reason that we like reg. cond. prob. is that they allow us to compute cond. expectations for all functions of X.

Theorem: Let μ be a r.c.p. for X given $\mathcal{G}$. If $f : \mathcal{X} \to \mathbb{R}$ has $\mathbb{E}|f(X)| < \infty$, then,

$$\mathbb{E}[f(X) \mid \mathcal{F}] = \int f(x)\, \mu(\omega, dx).$$