Lecture 5 — Mon Feb/05/2024

Last time
- Poisson Distribution
- Law of rare events

Today
- General Poisson Convergence
- Conditional Expectation

Generalization Poisson convergence

Theorem 2.0 For each n, let $X_{n,m} \in \mathbb{N}$ $\forall m \in [n]$ be ind. r.v.'s such that

$$X_{n,m} = \begin{cases} 1 & \text{with prob } p_{n,m} \\ 0 & \text{with prob } 1 - p_{n,m} - \varepsilon_{n,m} \\ \geq 2 & \text{with prob } \varepsilon_{n,m} \end{cases}$$

Moreover, suppose that

1) $\sum_{m=1}^{n} p_{n,m} \to \lambda \in (0,\infty)$,
2) $\max_{m \in [n]} p_{n,m} \to 0$ as $n \to \infty$.
3) $\sum_{m=1}^{n} \varepsilon_{n,m} \to 0$ as $n \to \infty$.

Let $S_n = \sum_{k=1}^{n} X_{nk}$, then,

$$S_n \xrightarrow{w} Z \quad \text{with} \quad Z \sim \text{Poisson}(\lambda).$$

$\dashv$

See proof in lecture 4 notes.

Conditional Expectation

Motivation

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a prob. space and X and Z be discrete r.v. taking values in

$$\mathcal{X} = \{x_1, \dots, x_m\},$$
$$\mathcal{Z} = \{z_1, \dots, z_n\}.$$

The elementary conditional prob:

$$\mathbb{P}(X = x_i \mid Z = z_j) = \frac{\mathbb{P}(X = x_i ; Z = z_j)}{\mathbb{P}(Z = z_j)}$$ ← Not so nice for general measures

The elementary conditional expectation:

$$(\star)\ \mathbb{E}(X \mid Z = z_j) = \sum_i x_i\, \mathbb{P}(X = x_i \mid Z = z_j).$$

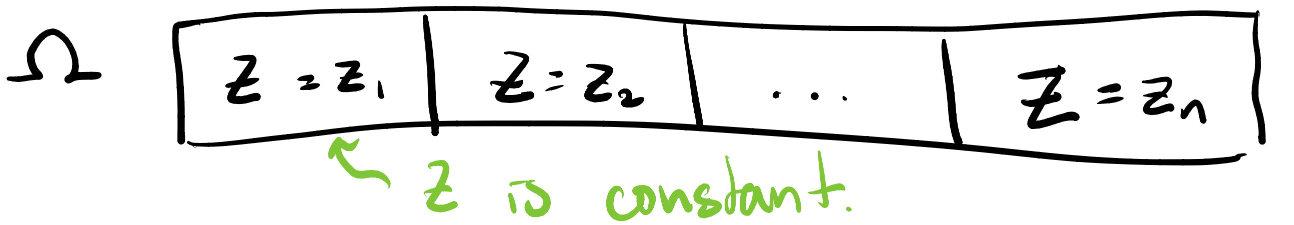
are likely familiar to you.

Question: How do we extend this definition for more general measures?

Let us interpret $(\star)$ as a random variable $Y = \mathbb{E}[X \mid Z]$ ← random

$$\text{if } Z(\omega) = z_j \Rightarrow Y(\omega) := \mathbb{E}[X \mid Z = z_j] = y_j.$$

Let's consider $\mathcal{G} = \sigma(Z)$ the σ-algebra generated by Z. Notice that Z can take n values, thus we can partition Ω



If sample $\omega \in \{Z = z_i\}$, then $Y(\omega) = y_i$. Thus, we obtain that

1) Y is $\mathcal{G}$-measurable.

Moreover,

2) $$\int_{\{Z = z_j\}} Y \, d\mathbb{P} = y_j \, \mathbb{P}(Z = z_j) = \sum_i x_i \, \mathbb{P}(X = x_i \mid Z = z_j) \, \mathbb{P}(Z = z_j)$$

$$= \sum_i x_i \, \mathbb{P}(X = x_i ; Z = z_j) = \int_{\{Z = z_j\}} X \, d\mathbb{P}.$$

This motivates the definition.

Theorem / Definition (Kolmogorov, 1933)

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, and

$\mathbb{E}|X| < \infty$

$X \in \mathcal{L}^1$ a r.v. Let $\mathcal{G} \subseteq \mathcal{F}$ be a sub-σ-algebra. Then, there exists a r.v. Y s.t.

0) $Y \in \mathcal{L}^1$.

1) Y is $\mathcal{G}$-measurable.

2) For all $G \in \mathcal{G}$

$$\int_G Y \, d\mathbb{P} = \int_G X \, d\mathbb{P}$$

Definition.

Moreover, Y is a.s. unique. $\dashv$

Notation: We write $\mathbb{E}[X | Z]$ for $\mathbb{E}[X | \sigma(Z)]$, and $\mathbb{E}[X | Z_1, \ldots)$ for $\mathbb{E}[X | \sigma(Z_1, \ldots)]$.

Before we prove this let's see a couple of examples:

Example 1: Discrete variables ✓

Even countable support

Example 2: Continuous densities

Let X, Y be random variables with joint density $f(x,y)$. Then,

$$\mathbb{E}[X \mid Y] = \frac{\int x \, f(x,y)\, dx}{\underbrace{\int f(x,y)\, dx}_{Z}}$$

1) Since Z is a measurable function of y, Z is $\sigma(Y)$-measurable

2) Let $A \in \sigma(y)$

$$\int_A \left. \frac{\int x \, f(x,y)\, dx}{\int f(x,y)\, dx} \right|_{y = Y(w)} dP(w) = \int_B \frac{\int x \, f(x,y)\, dx}{\cancel{\int f(x,y)\, dx}} \cancel{\int f(x,y)\, dx}\, dy$$

$$= \int_B \int_{\mathbb{R}} x \, f(x,y)\, dx\, dy = \int_{\substack{(x,y) \\ y \in B}} X \, dP(w).$$

$A = \{w : Y(w) \in B\}$

Example 3: If X is measurable w.r.t. $\mathcal{F}$

$\Rightarrow \mathbb{E}[X \mid \mathcal{F}] = X.$

To see this note that 1) follows trivially, and

$$\int_A X \, dP = \int_A X \, dP \quad \text{trivially } \forall A \in \mathcal{F}.$$

Trivial σ-algebra

Example 4: If $\mathcal{G} = \{\emptyset, \Omega\}$, then

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

Since constants are measurable w.r.t. $\mathcal{G}$, 1) follows.

Moreover

$$\int_\Omega \mathbb{E}[X]\, d\mathbb{P} = \mathbb{E}[X] = \int_\Omega X\, d\mathbb{P} \Rightarrow 2).$$

Proof of Theorem:

1) and 2) $\Rightarrow$ 0) Assume Y satisfies 1) and 2). Then, for $A = \{Y > 0\}$,

$$\int_A Y\, dP = \int_A X\, dP \leq \int_A |X|\, dP$$

$$\int_{A^c} -Y\, dP = \int_{A^c} -X\, dP \leq \int_{A^c} |X|\, dP$$

$$\Rightarrow \int_\Omega |Y|\, dP = \int_A Y\, dP + \int_{A^c} -Y\, dP \leq \int_\Omega |X|\, dP < \infty.$$

Uniqueness: suppose that there exists another $\tilde{Y}$ satisfying 0), 1), 2)
Then for all $G \in \mathcal{G}$

$$\mathbb{E}[(Y-\tilde{Y})\mathbb{1}_G] = 0$$

Searching contradiction assume that $\mathbb{P}(Y>\tilde{Y})>0$. Since

$$\underbrace{\{Y > \tilde{Y} + n^{-1}\}}_{A_n} \uparrow \overbrace{\{Y > \tilde{Y}\}}^{A}$$

monotonic convergence

$A_1 \subseteq A_2 \subseteq \dots \subseteq A$ and $A = \bigcup A_n$.

We see that for large enough n, $\mathbb{P}(A_n)>0$. Notice that such an event $A_n \in \mathcal{G}$, so

$$0 = \mathbb{E}(Y-\tilde{Y})\mathbb{1}_{A_n} \geq n^{-1}\mathbb{P}(A_n) > 0 \quad \lightning$$

TO BE CONTINUED...