Lecture 4
Conditional Expectation

Nils Detering

January 17th

Conditional Expectation

- We learned about conditional density (continuous RVs) and conditional probability mass function (discrete RVs)
- They define a probability distribution by

$$\mathbb{P}(Y \in R \mid X = x) = \begin{cases} \sum_{y_i \in R} p_{Y|X}(y_i|x) & X, Y \text{ discrete} \\ \int_R f_{Y|X}(y|x)dy & X, Y \text{ continuous} \end{cases}$$

  for $R \subset \mathbb{R}$.
- We can define expectation with respect to this distribution.

Conditional Expectation

We denote by $\mathbb{E}[Y|X = x]$ the conditional expectation of Y given $X = x$.
It calculates as

$$\mathbb{E}[Y|X = x] = \begin{cases} \sum_{y_i} y_i \, p_{Y|X}(y_i|x) & \text{X,Y discrete} \\ \int_{-\infty}^{\infty} y \, f_{Y|X}(y|x) dy & \text{X,Y continuous} \end{cases}$$

Observe that $\mathbb{E}[Y|X = x]$ is a function of x.

Example (Discrete case)

See whiteboard.

Example (Continuous case)

See whiteboard.

Properties Conditional Expectation

Let X, Y and Z be random variables, then

1. **Linearity:** For constants $a, b \in \mathbb{R}$
$$\mathbb{E}[aY + bZ|X = x] = a\mathbb{E}[Y|X = x] + b\mathbb{E}[Z|X = x]$$
2. If $g : Y(\Omega) \to \mathbb{R}$
$$\mathbb{E}[g(Y)|X = x] = \begin{cases} \sum_{y_i} g(y_i) p_{Y|X}(y_i|x) & \text{X,Y discrete} \\ \int_{-\infty}^{\infty} g(y) f_{Y|X}(y|x) dy & \text{X,Y continuous} \end{cases}$$
3. **Independence:** If X and Y are independent, then
$$\mathbb{E}[Y|X = x] = \mathbb{E}[Y]$$
4. If $Y = g(X)$ is a fct. of X, then
$$\mathbb{E}[Y|X = x] = g(x)$$

Proof.

See whiteboard. □

Example

RVs X, Y and U, where U is uniformly distributed on $(0, 1)$. What is

$$\mathbb{E}[UX^2 + (1 - U)Y^2 | U = u] \ ?$$

Extension Conditional Expectation given event

- Observe that so far we conditioned on events of the form $A = \{\omega : X(\omega) = x\}$.
- Can we do more general, say on events such as $A = \{\omega : X(\omega) \geq 0\}$? Yes!

But first we need the following

Definition (Indicator)

Let $A \in \mathcal{F}$. Define the indicator function I_A for A by

$$I_A(\omega) = \begin{cases} 1 & \text{if } \omega \in A \\ 0 & \text{if } \omega \notin A \end{cases}$$

One gets: $\mathbb{E}[I_A] = 1 \cdot \mathbb{P}(A) + 0 \cdot \mathbb{P}(A^c) = \mathbb{P}(A)$.

Extension Conditional Expectation given event

Definition (Conditional expectation given event)

Let $A \in \mathcal{F}$ with $\mathbb{P}(A) > 0$. The conditional expectation of Y given A is

$$\mathbb{E}[Y|A] = \frac{\mathbb{E}[YI_A]}{\mathbb{P}(A)}.$$

Observe that in the discrete case

$$\mathbb{E}[Y|A] = \frac{1}{\mathbb{P}(A)} \sum_y y\mathbb{P}(\{Y = y\} \cap A) = \sum_y y\mathbb{P}(\{Y = y\}|A).$$

Choosing $A = \{\omega : X(\omega) = x\}$ retrieves $\mathbb{E}[Y|X = x]$.

Example (Continuous Case)

Joe's phone talking. See whiteboard.

Law of total expectation

Let $A_1, A_2, \ldots$ be a sequence of events that partition the state space i.e. $\Omega = \cup_{i=1}^{\infty} A_i$ and $A_i \cap A_j = \emptyset$ for $i \neq j$.
Then

$$\sum_{i=1}^{\infty} I_{A_i} = 1 \qquad \text{why?}$$

and therefore

$$Y = \sum_{i=1}^{\infty} Y I_{A_i}.$$

Taking expectation yields

$$\mathbb{E}[Y] = \sum_{i=1}^{\infty} \mathbb{E}[Y I_{A_i}] = \sum_{i=1}^{\infty} \frac{\mathbb{E}[Y I_{A_i}]}{\mathbb{P}(A_i)} \mathbb{P}(A_i) = \sum_{i=1}^{\infty} \mathbb{E}[Y | A_i] \mathbb{P}(A_i) \quad (1)$$

Law of total expectation

Let's look at two special cases of the law of total expectation!

1. For the random variable $Y = I_B$ we know that $\mathbb{E}[I_B] = \mathbb{P}(B)$ and similarly $\mathbb{E}[I_B|A_i] = \mathbb{P}(B|A_i)$. Thus from (1)

$$\mathbb{P}(B) = \mathbb{E}[I_B] = \sum_{i=1}^{\infty} \mathbb{E}[I_B|A_i]\mathbb{P}(A_i) = \sum_{i=1}^{\infty} \mathbb{P}(B|A_i)\mathbb{P}(A_i).$$

2. For the event $A_i = \{\omega : X(\omega) = x_i\}$ for $i = 1, 2, \ldots$ we get from (1) that

$$\mathbb{E}[Y] = \sum_{i=1}^{\infty} \mathbb{E}[Y|X = x_i]\mathbb{P}(X = x_i).$$

Example

See whiteboard.

Conditional expectation as a random variable

Recall: The conditional expectation of X given $Y = y$ ($\mathbb{E}[X|Y = y]$).

- $\mathbb{E}[X|Y = y]$ is just a deterministic number for every y.
- Can define $f : Y(\Omega) \to \mathbb{R}$ by $f(y) = \mathbb{E}[X|Y = y]$
- Allows us to define RV by plugging Y into f, $f(Y) : \Omega \to \mathbb{R}$ defined by $\omega \mapsto f(Y(\omega))$.

Definition

Let X and Y be discrete or jointly continuous RVs. The conditional expectation of X given Y, denoted by $\mathbb{E}[X|Y]$ is the random variable $f(Y)$ with f defined by $f(y) = \mathbb{E}[X|Y = y]$.

Conditional expectation as a random variable

Observation: $\mathbb{E}[X|Y=y]$ is just a number while $\mathbb{E}[X|Y]$ is a random variable. All possible values of $\mathbb{E}[X|Y]$ are given by $\mathbb{E}[X|Y=y]$ as y varies.

Example

See whiteboard.

Properties of $\mathbb{E}[X|Y]$

1. **Linearity:** $\mathbb{E}[aX + bZ|Y] = a\mathbb{E}[X|Y] + b\mathbb{E}[Z|Y]$
2. **Independence:** X, Y indep. $\mathbb{E}[X|Y] = \mathbb{E}[X]$
3. **Taking out what is known:** $\mathbb{E}[Xh(Y)|Y] = h(Y)\mathbb{E}[X|Y]$
4. **Tower property:** $\mathbb{E}[\mathbb{E}[X|Y]] = \mathbb{E}[X]$
5. $\mathbb{E}[(X - h(Y))^2] \geq \mathbb{E}[(X - \mathbb{E}[X|Y])^2]$