

Lecture #8: Conditional Expectation

Example:

Let X and Y be two discrete independent random variables with

$$P(X=1)=\frac{1}{2}, \quad P(Y=1)=\frac{1}{2}$$

$$P(X=0)=\frac{1}{2}, \quad P(Y=2)=\frac{1}{2}$$

Let $Z=X+Y$

$$\begin{aligned} \bullet E[Z|X=0] &= 1 \cdot P(Z=1|X=0) + 2 \cdot P(Z=2|X=0) + 3 \cdot P(Z=3|X=0) \\ &= \frac{1 \cdot P(Z=1, X=0) + 2 \cdot P(Z=2, X=0) + 3 \cdot P(Z=3, X=0)}{P(X=0)} \end{aligned}$$

$$= \frac{1}{P(X=0)} \sum_{j=1}^3 j \cdot P(Z=j|X=0)$$

$$= \frac{1}{\frac{1}{2}} \left(\frac{1}{4} + 2 \cdot \frac{1}{4} \right)$$

$$= \frac{3}{2}$$

$$\bullet E[Z|X=1] = \frac{1}{P(X=1)} \sum_{j=1}^3 j \cdot P(Z=j|X=1) = \frac{1}{\frac{1}{2}} \left(2 \cdot \frac{1}{4} + 3 \cdot \frac{1}{4} \right) = \frac{5}{2}$$

Therefore, we define the random variable

$$E[Z|X](\omega) = \begin{cases} \frac{3}{2} & \text{if } X(\omega)=0 \\ \frac{5}{2} & \text{if } X(\omega)=1 \end{cases}$$

Definition - Let X, Y be continuous random variables with joint PDF $f(x, y)$ then

$$E[Y|X] = h(X) \leftarrow \text{New random variable}$$

where

$$h(x) \sim \frac{\int_{-\infty}^{\infty} y f(x, y) dy}{\int_{-\infty}^{\infty} f(x, y) dy}$$

probability of x and y occurring
probability of x -occurring

Observation 1: For random variables X, Y with joint density:
 $E[g(X)Y] = E[g(X)E[Y|X]]$

proof:

$$E[g(X)Y] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x)y f(x,y) dy dx$$

$$= \int_{-\infty}^{\infty} g(x) \left(\int_{-\infty}^{\infty} y f(x,y) dy \right) dx$$

$$= \int_{-\infty}^{\infty} \frac{g(x) \left(\int_{-\infty}^{\infty} f(x,z) dz \right) \left(\int_{-\infty}^{\infty} y f(x,y) dy \right)}{\int_{-\infty}^{\infty} f(x,z) dz} dx$$

$$= \int_{-\infty}^{\infty} g(x) \left(\int_{-\infty}^{\infty} f(x,z) dz \right) \left(\int_{-\infty}^{\infty} \frac{y f(x,y) dy}{\int_{-\infty}^{\infty} f(x,z) dz} \right) dx$$

$$= \int_{-\infty}^{\infty} g(x) \left(\int_{-\infty}^{\infty} f(x,z) dz \right) E[Y|X] dx$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x) E[Y|X] f(x,z) dz dx$$

$$= E[g(X)E[Y|X]]$$

Observation 2: If X and Y are independent with joint density then
 $E[Y|X] = E[Y]$.

proof:

Since X, Y are independent the joint density can factor:

$$(X, Y) \sim f(x,y) = g(x)h(y)$$

$$\Rightarrow E[Y|X] = \frac{\int_{-\infty}^{\infty} y f(x,y) dy}{\int_{-\infty}^{\infty} f(x,y) dy} = \frac{\int_{-\infty}^{\infty} y g(x) h(y) dy}{\int_{-\infty}^{\infty} g(x) h(y) dy} = \frac{\int_{-\infty}^{\infty} y h(y) dy}{\int_{-\infty}^{\infty} h(y) dy} = E[Y]$$

Observation #1 motivates the following definition.

Definition: Let X, Y be integrable random variables, i.e., $E[|X|], E[|Y|] < \infty$, on $(\Omega, \mathcal{F}, P)$. The conditional expectation $E[Y|X] = h(X)$ is the random variable satisfying

$$E[g(X)Y] = E[g(X)E[Y|X]] = E[g(X)h(X)]$$

for all bounded functions g .

← average over joint density

→ average over X only.

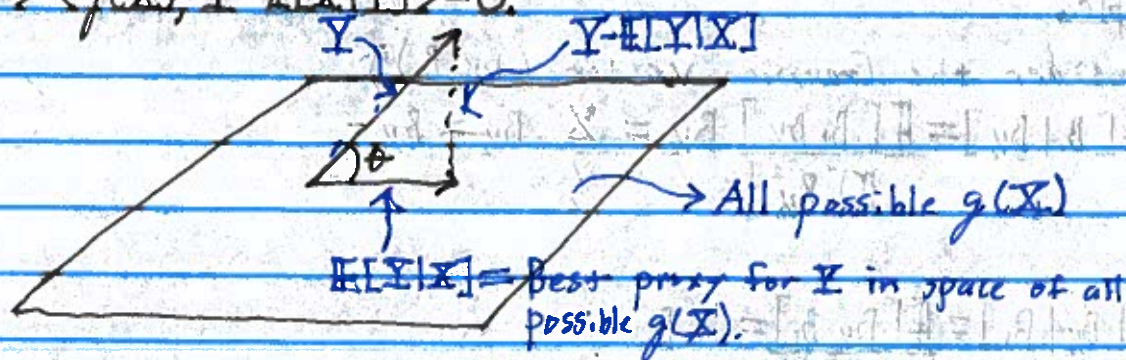
Note: $E[g(X)Y]$ is well defined since

$$|E[g(X)Y]| \leq E[|g(X)| \cdot |Y|] = E[\max_X |g(X)| \cdot |Y|] = \max_X |g(X)| E[|Y|] < \infty$$

Corollary: The random variable $Y - E[Y|X]$ is orthogonal to $g(X)$, for all bounded g .

proof:

$$\begin{aligned} \langle g(X), Y \rangle &= E[g(X)Y] = E[g(X)E[Y|X]] = \langle g(X), E[Y|X] \rangle \\ \Rightarrow \langle g(X), Y - E[Y|X] \rangle &= 0. \end{aligned}$$



Corollary: $E[Y] = E[E[Y|X]]$

proof:

Letting $g(X) = 1$ we have that

$$E[1 \cdot Y] = E[1 \cdot E[Y|X]]$$

$$\Rightarrow E[Y] = E[E[Y|X]]$$

Theorem - Let (X, Y) be a Gaussian vector mean 0. Then,

$$E[Y|X] = \frac{E[XY]}{E[X^2]} X = \frac{\langle X, Y \rangle}{\langle X, X \rangle} X = \frac{\|Y\| \cos \theta}{\|X\|} X \leftarrow \text{Projection onto } X!$$

proof

$$\begin{aligned} \text{Cov}\left(X, Y - \frac{\langle X, Y \rangle}{\|X\|^2} X\right) &= \left\langle X, Y - \frac{\langle X, Y \rangle}{\|X\|^2} X \right\rangle \\ &= \langle X, Y \rangle - \frac{\langle X, Y \rangle \langle X, X \rangle}{\|X\|^2} \\ &= 0 \end{aligned}$$

Therefore, since these random variables are Gaussian they are independent and thus

$$\begin{aligned} E[g(X)(Y - \frac{\langle X, Y \rangle}{\|X\|^2} X)] &= E[g(X)] E[Y - \frac{\langle X, Y \rangle}{\|X\|^2} X] = 0 \\ \Rightarrow E[Y|X] &= \frac{E[XY]}{E[X^2]} X \end{aligned}$$

Example:

Consider the Gaussian vector (B_1, B_2)

$$1. E[B_1 | B_2] = \frac{E[B_1 B_2]}{E[B_2^2]} B_2 = \frac{1/2}{1/2} B_2 = B_2$$

$$2. E[B_2 | B_1] = \frac{E[B_2 B_1]}{E[B_1^2]} B_1 = \frac{1/2}{1/2} B_1$$