

Section 10.3

Conditional expectation

Ex) Roll a die, set the result to be Y .

Then, roll Y dice, and let X = sum of the Y dice.

What is $\mathbb{E}[X]$?

First, what is $\mathbb{E}[X|Y = y]$?

we roll 1 die

$$\mathbb{E}[X|Y = 1] = \frac{7/2}{1}$$

$$\mathbb{E}[X|Y = 2] = \frac{2 \cdot 7/2}{2}$$

$$\mathbb{E}[X|Y = y] = \frac{y \cdot 7/2}{y}, \quad y \in \{1, 2, 3, 4, 5, 6\}$$

of dice rolled *Avg of a die roll*

$$\mathbb{E}X = \frac{\sum_y \mathbb{E}[X|Y=y] \cdot P_Y(y)}{6}$$

$$\mathbb{E}X = \sum_y \mathbb{E}[X|Y=y] \cdot P_Y(y) = \sum_{y=1}^6 y \cdot 7/2 \cdot 1/6 = 7/2 \cdot 7/2$$

we will find a "cute" way to solve this below.

Note: $\mathbb{E}[X|Y = y]$ takes a different value for each choice of y , thus it defines a function: $v(y) := \mathbb{E}[X|Y = y]$.

In example above, $v(y) = \underline{y/2}$.

Then, we may consider $v(Y) = 3.5 Y$, which is a function of a random variable, and is thus a random variable (as opposed to $v(y)$).

This defines $\mathbb{E}[X|Y]$, i. e. $\mathbb{E}[X|Y] := v(Y)$.

It is different than $\mathbb{E}[X|Y = y] = v(y)$.

Def: Let X and Y be discrete or jointly continuous r.v.s. Let $v(y) := E[X|Y = y]$.
The *conditional expectation* of X given Y is
 $E[X|Y] := v(Y)$.

$$E[X|Y]$$

vs

$$E[X|Y = y]$$

“Conditional expectation of X given Y ”

vs

“Conditional expectation of X given $Y = y$ ”

$E[X|Y]$ is a r.v.

vs $E[X|Y = y]$ is a scalar

Finding expectation from conditional expectation

Fact: $\mathbb{E}[\overbrace{\mathbb{E}[X|Y]}^{f.v.}] = \mathbb{E}[X]$

Pf (continuous case): $v(y) = \mathbb{E}[X|Y = y]$

$$\begin{aligned}\mathbb{E}[\mathbb{E}[X|Y]] &= \mathbb{E}[v(Y)] = \int_{-\infty}^{\infty} v(y) f_Y(y) dy = \int_{-\infty}^{\infty} \mathbb{E}[X|Y=y] \cdot f_Y(y) dy \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x \cdot \underbrace{f_{X|Y}(x|y) \cdot f_Y(y)}_{f_{X,Y}(x,y)} dx dy = \int \int x f_{X,Y}(x,y) dy dx \\ &= \int_{-\infty}^{\infty} x \int_{-\infty}^{\infty} f_{X,Y}(x,y) dy dx = \int_{-\infty}^{\infty} x \cdot f_X(x) dx = \mathbb{E} X \quad \checkmark\end{aligned}$$

Conditional expectation respects linearity

Fact: For all y such that $\mathbb{P}[Y = y] > 0$ in the discrete case and $f_Y(y) > 0$ in the continuous case:

$$\mathbb{E}[X_1 + X_2 | Y = y] = \mathbb{E}[X_1 | Y = y] + \mathbb{E}[X_2 | Y = y].$$

Hence,

$$\mathbb{E}[X_1 + X_2 | Y] = \mathbb{E}[X_1 | Y] + \mathbb{E}[X_2 | Y]$$

Pf: (Discrete case) The first equality in the discrete case is the usual linearity of expectation (recalling that the conditional expectation given $Y = y$ is just some regular expectation w.r.t. a new distribution corresponding to the conditioning on $Y = y$)

The second equality is immediate from the first.

Back to example above

Ex) Y = die roll, X = sum of Y dice. What is $\mathbb{E}[X]$?

Recall, $\mathbb{E}[X|Y] = \underline{Y \cdot \frac{7}{2}}$,

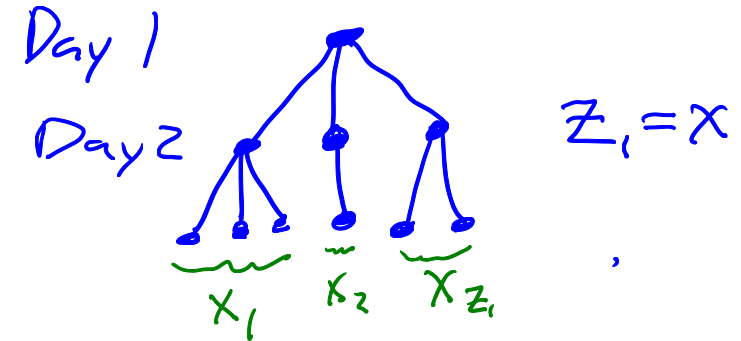
$$\mathbb{E}[X|Y] = v(Y) = Y \cdot \frac{7}{2}$$

$$\text{So } \mathbb{E}X = \mathbb{E}[\mathbb{E}[X|Y]] = \underline{\mathbb{E}\left[Y \cdot \frac{7}{2}\right]} = \frac{7}{2} \mathbb{E}Y = \boxed{\frac{7}{2} \cdot \frac{7}{2}} \checkmark$$

Ex) Virus spread (simple model). Suppose that on Day 1, Person 1 has a certain virus. During that day, Person 1 infects X people, and then gets better. This repeats, with each newly infected person infecting an iid copy of X people in the following day after the day they got infected. Suppose $\mathbb{E}X = r$. Let Z_n be the number of new infections on Day n . What is $\mathbb{E}[Z_n]$?

$$\mathbb{E}[Z_1] = \underline{\mathbb{E}X = r}$$

iid copies of X



$$\mathbb{E}[Z_{n+1} | Z_n = z] = \underline{\mathbb{E}[X_1 + X_2 + \dots + X_z]}$$

$$= \mathbb{E}X_1 + \dots + \mathbb{E}X_z = z \cdot \mathbb{E}X = z \cdot r$$

$$\Rightarrow \mathbb{E}[Z_{n+1} | Z_n] = Z_n \cdot r$$

$$\Rightarrow \mathbb{E}[Z_{n+1}] = \mathbb{E}[\mathbb{E}[Z_{n+1}|Z_n]] = \mathbb{E}[Z_n \cdot r] = r \cdot \mathbb{E}[Z_n]$$

$$\mathbb{E}Z_1 = r$$

$$\mathbb{E}Z_2 = r \cdot \mathbb{E}Z_1 = r \cdot r = r^2$$

$$\vdots$$

$$\mathbb{E}[Z_n] = r^n$$

$$\text{If } r > 1, \quad \mathbb{E}Z_n \rightarrow \infty$$

$$\text{If } r < 1, \quad P(Z_n \neq 0) = P(Z_n \geq 1) \stackrel{\text{Markov} \neq}{\leq} \frac{\mathbb{E}Z_n}{1} = r^n \rightarrow 0$$

i.e. virus will go extinct eventually.

Conditioning and independence: Independence means conditioning has no effect

Fact: R.v.s X, Y are independent if and only if

Discrete case: $p_{X|Y}(x|y) = \underline{p_X(x)}$

Continuous case: $f_{X|Y}(x|y) = \underline{f_X(x)}$

For all x, y
where they
are defined

Conditional expectation and independence

Fact: If r.v.s X, Y are independent, then

$$\mathbb{E}[X|Y = y] = \underline{\mathbb{E}X}$$

and

$$\mathbb{E}[X|Y] = \underline{\mathbb{E}X}$$

Conditioning $g(X)$ on X

Fact:

- $\mathbb{E}[X|X = x] = \underline{x}$ $\leftarrow$ scalar x
- $\mathbb{E}[g(X)|X = x] = \underline{g(x)}$
- $\mathbb{E}[X|X] = \underline{X}$ $\leftarrow$ r.v. X
- $\mathbb{E}[g(X)|X] = \underline{g(X)}$ $\leftarrow$ r.v.s
- $\mathbb{E}[g(X)f(Z)|X] = \underline{g(X)\mathbb{E}[f(Z)|X]}$

Multivariate conditional distributions: Everything generalizes naturally

- $p_{X_1, \dots, X_n | Y}(x_1, \dots, x_n | y) := \frac{p_{X_1, \dots, X_n, Y}(x_1, \dots, x_n, y)}{p_Y(y)}$
- $f_{X_1, \dots, X_n | Y}(x_1, \dots, x_n | y) := \frac{f_{X_1, \dots, X_n, Y}(x_1, \dots, x_n, y)}{f_Y(y)}$
- $\mathbb{E}[g(X_1, \dots, X_n) | Y = y] =$
 - $\sum_{x_1, \dots, x_n} g(x_1, \dots, x_n) p_{X_1, \dots, X_n | Y}(x_1, \dots, x_n | y)$ (discrete case)
 - $\int_{\mathbb{R}^n} g(x_1, \dots, x_n) f_{X_1, \dots, X_n | Y}(x_1, \dots, x_n | y) dx_1 \cdots dx_n$ (continuous case)

Below is not tested: Conditional prediction

Question: Let $W, X \sim N(0,1)$ be independent. Let $Y = \underbrace{W}_{\text{blue circle}} + \underbrace{X}_{\text{blue underline}}$.

What is your best estimate of X given Y ?

Estimate: Some function $h(Y)$.

Best estimate?

We wish to minimize *expected squared error*: $E[(h(Y) - X)^2]$

Conditional expectation is best predictor

Thm: $\mathbb{E}[(X - h(Y))^2] \geq \mathbb{E}[(X - \mathbb{E}[X|Y])^2]$.

Equality only holds if $h(Y) = \mathbb{E}[X|Y]$.

$$\text{Pf: } \mathbb{E}(X - h(Y))^2 = \mathbb{E}\left[\mathbb{E}[(X - h(Y))^2 | Y]\right]$$

Inner expectation: $v(Y) = \mathbb{E}[X|Y]$

$$\mathbb{E}[(X - h(Y))^2 | Y] = \mathbb{E}[(X - v(Y) + v(Y) - h(Y))^2 | Y]$$

$$= \mathbb{E}[(X - v(Y))^2 | Y] + \mathbb{E}[(v(Y) - h(Y))^2 | Y] + 2 \mathbb{E}[(X - v(Y))(v(Y) - h(Y)) | Y]$$

$$= I + II + 2 \cdot III$$

$$II \geq 0$$

$$III = E[(X - v(Y)) \cdot (v(Y) - h(Y)) | Y] = (v(Y) - h(Y)) \cdot E[X - v(Y) | Y]$$

$$\Rightarrow E(X - h(Y))^2 = E[E[(X - v(Y))^2 | Y]] + E \underbrace{II}_{\geq 0}$$

$$\geq E(X - E[X|Y])^2$$

$$\begin{aligned} & E[X|Y] - E[v(Y)|Y] \\ &= v(Y) - v(Y) = 0 \end{aligned}$$

