ECSE 2500
Lec 22
April 20

Topic: Conditional expectation
— Law of iterated expectation
Sum/Average of several RVs
— Law of large number

□ Conditional expectation

Recall the definition in the last class

Q: What is the expected value of the conditional expectation of Y given X (e.g. a function of RV X)?

$$E[g(X)] = \int_{-\infty}^{+\infty} g(x) f_X(x)\, dx$$

$$= \int_{-\infty}^{+\infty} E[Y|X=x] f_X(x)\, dx$$

$$= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \underbrace{y f_{Y|X}(y|x)\, dy\, f_X(x)}\, dx$$

$$= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \underbrace{y f_{XY}(x,y)}\, dx\, dy$$

$$= \int_{-\infty}^{+\infty} y \cdot f_Y(y)\,dy$$

$$= E[Y]$$

To summarize, we have

$$\boxed{\underset{X}{E}[\underset{Y|X}{E}[Y|X]] = E[Y]}$$

The law of iterated expectation!!

Composition/Nest of two Expectations

Reduce to the simplified case if $h(Y) = Y$.

Slight generalization:

If we have a function $h(\cdot)$ of a RV Y,

$$E[\,E[h(Y)|X]] = E[h(Y)]$$

That is, we can think of $E[h(Y)]$ in two steps: First, compute $E[h(Y)|x]$ for each possible value x of RV X. Secondly, we average over the distributions of X.

Example (Discrete RVs)

Consider the coin toss example in the last week. Flip 3 coins. and let

$$X = \#\text{ of heads}$$

$$Y = \text{position of first head}$$

In the last week, we have computed $P_{Y|X}$ as

$P(Y\|X)$	$Y=0$	1	2	3
$X=0$	1	0	0	0
1	0	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$
2	0	$\frac{2}{3}$	$\frac{1}{3}$	0
3	0	1	0	0

Given the conditional PMF in the table, we calculate

$$E[Y|X=0] = \sum_{y_j} y_j P(y_j|X=0)$$

$$= 1\cdot 0 + 1\cdot 0 + 2\cdot 0 + 3\cdot 0 = 0$$

$$E[Y|X=1] = 0\cdot 0 + (1+2+3)\cdot \frac{1}{3} = 2$$

$$E[Y|X=2] = 0\cdot 0 + \frac{2}{3} + 2\cdot \frac{1}{3} + 3\cdot 0 = \frac{4}{3}$$

$$E[Y|X=3] = 1 \cdot 1 = 1$$

Then we use the law of iterated expectation, we have

$$\begin{aligned}
E[Y] &= E\big[E[Y|X]\big] \\
&= E[Y|X=0]P(X=0) + E[Y|X=1]P(X=1) \\
&\quad + E[Y|X=2]P(X=2) + E[Y|X=3]P(X=3) \\
&= 0 \cdot \frac{1}{8} + 2 \cdot \frac{3}{8} + \frac{4}{3} \cdot \frac{3}{8} + 1 \cdot \frac{1}{8} = \frac{11}{8}
\end{aligned}$$

To do a sanity check,

$$\begin{aligned}
E[Y] &= \sum_{y_j} y_j P(Y=y_j) \\
&= 0 \cdot \frac{1}{8} + 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{4} + 3 \cdot \frac{1}{8} = \frac{11}{8}
\end{aligned}$$

$\rightarrow$ The law of iterated expectation is verified!

Example (Continuous RVs)

Consider X is uniformly drawn from $[0,1]$, and Y is uniformly drawn from $[0,x]$.

Q: What is the expected value of Y?

The problem structure facilitates us to use the law of iterated expectation!

① What is $E[Y|x]$?

$$E[Y|x] = \int_{-\infty}^{+\infty} y \underbrace{f_{Y|x}(y|x)}_{\frac{1}{x}} dy$$

$$= \int_0^x \frac{y}{x} dy$$

$$= \frac{1}{x} \frac{y^2}{2} \Big|_{y=0}^{y=x} = \frac{x}{2}$$

② What is $E[Y]$?

$$E[Y] = E[E[Y|x]]$$

$$= \int_0^1 \frac{x}{2} \cdot f_X(x) dx$$

$$= \int_0^1 \frac{x}{2} dx = \frac{x^2}{4} \Big|_{x=0}^{x=1} = \frac{1}{4}$$

To do a sanity check, from previous notes, we have calculated $f_Y(y) = -\log y$, and then

$$E[Y] = \int_0^1 -y \log y \, dy = \frac{y^2}{4} - \frac{y^2}{2} \log y \Big|_{y=0}^{y=1} = \frac{1}{4}$$

✓

□ Sum / Mean of several random variables

In many situations, we observe a sequence of RVs (e.g., a series of coin flips) and want to process these to extract information from those RVs.
A natural way to do so is via the sum / mean of RVs.

Sum : $$S_n = X_1 + X_2 + \cdots + X_n = \sum_{i=1}^{n} X_i$$

number of RVs

Mean : $$M_n = \frac{1}{n} S_n = \frac{1}{n}\sum_{i=1}^{n} X_i$$

First we compute the expected value of these S_n and M_n

$$E[S_n] = E[X_1 + X_2 + \cdots + X_n]$$
$$= \sum_{i=1}^{n} E[X_i]$$

$$E[M_n] = \frac{1}{n} E[X_1 + X_2 + \cdots + X_n]$$
$$= \frac{1}{n}\sum_{i=1}^{n} E[X_i]$$

That is, the expected value of Sum of RVs is sum of Expected RVs.

Second, what about their variance?

$$Var(S_n) = E[(S_n - E[S_n])^2]$$

$$= E[S_n^2] - (E[S_n])^2$$

$$= E[(X_1+X_2+\cdots+X_n)(X_1+X_2+\cdots+X_n)]$$

$$- (E[X_1+X_2+\cdots+X_n])^2$$

$$= E\left[\sum_{i=1}^{n}\sum_{j=1}^{n} X_i X_j\right] - \sum_{i=1}^{n}\sum_{j=1}^{n} E[X_i]E[X_j]$$

$$= \sum_{j=1}^{n}\left[E[X_j^2] - (E[X_j])^2\right]$$

$$+ \sum_{i=1}^{n}\sum_{j\neq i}\left[\underbrace{E[X_i X_j] - E[X_i]E[X_j]}_{=(X_i - E[X_i])(X_j - E[X_j])}\right]$$

$$= \sum_{j=1}^{n} Var(X_j) + \sum_{i=1}^{n}\sum_{j\neq i} Cov(X_i, X_j)$$

That is, the variance of sum of RVs is the sum of all the possible covariances.

$\boxed{\text{Special cases}}$ 1) $Cov(X_j, X_i) = 0$, for all $i \neq j$

all uncorrelated

$$Var(S_n) = \sum_{j=1}^{n} Var(X_j)$$

2) All RVs X_i are independent and identically distributed (iid). That is, each X_i has the same PDF $f_X(x)$ and their covariance are zero,

$$(\text{iid}) \xrightarrow{\text{implication}} \begin{cases} E[X_i] = \mu, & \text{for all } i \\ Var(X_i) = \sigma^2, & \text{for all } i \\ Cov(X_i, X_j) = 0, & \text{for } i \neq j \end{cases}$$

Sum S_n
$$\left[\begin{array}{l} E[S_n] = \sum_{j=1}^{n} E[X_j] = n\mu \\ Var(S_n) = n\sigma^2 \end{array} \right.$$

Mean M_n
$$\left[\begin{array}{l} E[M_n] = \frac{1}{n} \sum_{j=1}^{n} E[X_j] = \mu \\ Var(M_n) = \frac{1}{n^2} Var(S_n) = \frac{\sigma^2}{n} \quad (*) \end{array} \right.$$

$Var(aX) = a^2 Var(X)$

The equation $(*)$ is important since it says that the variance of M_n gets smaller as the number of RVs increases!!

In the limit, for n iid. RVs,

$$\lim_{n \to \infty} Var(M_n) = 0,$$

$$\lim_{n \to \infty} E[M_n] = \mu.$$

So we get closer and closer to the right mean of X_i when $n \to \infty$.