Conditional Expectation and Prediction 7.5-7.6

Eva Gao, Abby Williams, Jeffery Huang

November 25th

1 Pre-recorded Lecture

1.1 Conditional Expectation

For Continuous Random Variable:

$$E[X|Y=y] = \int x f_{x|y}(x|y)\ dx = \int x\ \frac{f(x,y)}{f_Y(y)}\ dx$$

For Discrete Random Variable:

$$E[X|Y=y] = \sum_i x_i p_{X|Y}(x_i|y) = \sum_i x_i P[X=x_i|Y=y]$$

For the Conditional Expectation of a function of a random variable:

$$E[g(x)|Y=y] = \int g(x) f_{X|Y}(x|y)\ dx$$

Denote by $E[X|Y] = g(Y)$ which is a function of the random variable Y, and itself is a random variable here
Whereas: $E[X|Y=y]$ is not a random variable

$$E[E[X|Y]] = \int E[X|Y=y] F_Y(y)\ dy$$

$$= \int_y \int_x x f_{X|Y}(x|y)\ dx f_Y(y)\ dy$$

$$= \int \int f(x, y)\ dx\ dy = E[X]$$

This calculation shows that $E[E[X|Y]] = E[X]$ (**Note:** $E[E[X|Y]]$is a iterated expectation)

1.2 Conditional Variance

$$Var(X|Y) = E[(X - E[X|Y])^2|Y] = E[X^2|Y] - (E[X|Y])^2$$

Since $Var(X|Y)$is a random variable of function Y, we can do this:

$$E[Var(X|Y)] = E[E[X^2|Y] - (E[X|Y])^2]$$

$$= E[X^2] - E[E[X|Y]^2] \quad (*)$$

Also we know that E[E[X—Y]]=E[X], so:

$$Var(E[X|Y]) = E[(E[X|Y])^2] - (E[E[X|Y]])^2$$

Here$(E[E[X|Y]])^2 = E[X]$

Hence, $Var(E[X|Y]) = E[(E[X|Y])^2] - (E[X])^2 \quad (**)$

$$(*) + (**) = E[Var(X|Y)] + Var(E[X|Y])$$

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

We have the following result:

Conditional Variance Formula:

$$Var(X) = Var(E[X|Y]) + E[Var(X|Y)]$$

Example: Let $x_1, x_2, ..., x_n$ be independent random variables, and $N > 0$ is an intervalued random variable. What is the $Var(\sum_{i=1}^{N} X_i)$? (Hint: The sum depends on several random variables including all the x's and the N, so we have to conditioned on N first to see with fixed values of N what's the variance of the sum)

1.3 Conditional Expectation and Prediction

Consider two random variables X,Y. We observe X and predict what value Y will take. Let g(x) be our predictor (i.e. Observe X=x, predict y=g(x)) Once choice for a "good" predictor is to minimize the square error:

$$\min_g E[(Y - g(x))^2] = \min_g \int \int (y - g(x))^2 f(x, y)\ dx\ dy$$

Proposition: $E[(Y - g(x))^2] \geqslant E[(Y - E[Y|X])^2]$

Example: the best linear predictor of Y

i.e. Find the best a,b to minimize $E[(Y - (a + bx))^2]$
sol: Because $a = \mu_Y - \frac{\rho\sigma_Y\mu_X}{\sigma_X}$, $b = \rho\frac{\sigma_Y}{\sigma_X}$

$$g(x) = \mu_Y + \frac{\rho\sigma_Y}{\sigma_X}(x - \mu_x)$$

And using these values of a,b, we can compute the mean square error:

$$E[(Y - (a + bx))^2] = \sigma_Y^2(1 - \rho^2)$$

This means that if $\rho = \pm 1$ then the mean square error is zero

2 In-class Lecture

2.1 Conditional Expectation

$$E[X|Y = y] = \int x \cdot f_{X|Y}(x, y)\ dx$$

Note: The above equation depends on y.

$$E[E[X|Y]] = \int E[X|Y = y] \cdot f_Y(y)\ dy$$

Note: $E[X|Y]$ is a random variable

$$= \int \int x \cdot f_{X|Y}(x, y)\ dx f_Y(y)\ dy$$

$$= \int\int x \cdot f_Y(y)\frac{f(x,y)}{f_Y(y)}\ dx\ dy$$
$$= \int\int x \cdot f(x,y)\ dx\ dy = E[X]$$

Example: Let N be the number of customers entering a certain store and X_i be the amount of money that the customer i spends. The two random variables are independent to each other. (X_i is an independent and identically distributed random variable.) What is the expected value of the following expression?

$$T = \sum_{i=1}^{N} X_i$$

Solution:

$$E[T] = E[\sum_{i=1}^{N} X_i] = E[E[T/N]] = E[E[\sum_{i=1}^{N} X_i/N]]$$

From the property we concluded above:

$$E[\sum_{i=1}^{N} X_i/N = n] = E[\sum_{i=1}^{n} X_i] = n \cdot E[X_i]$$

Therefore, we conclude that

$$E[T] = E[N \cdot E[X_i]] = E[N] \cdot E[X_i]$$

2.2 Conditional Variance

$Var(X|Y)$ is similar to $Var(X)$ but all the expectations are conditional on the fact that Y is given.

$$Var(X|Y) = E[(X - E[X|Y])^2|Y]$$

$$Var(X|Y) = E[X^2|Y] - (E[X|Y])^2$$

So

$$E[Var(X|Y)] = E[E[X^2|Y]] - E[(E[X|Y])^2]$$

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

Since

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

We have

$$Var(E[X|Y]) = E[(E[X|Y])^2] - (E[X])^2$$

Thus, the conditional variance formula is

$$Var(X) = E[Var(X|Y)] + Var(E[X|Y])$$

Example: Following the previous example, we want to calculate the conditional variance.

$$Var(T) = Var(\sum_{i=1}^{N} X_i)$$

$$= E[Var(\sum_{i=1}^{N} X_i|N)] + Var(E[\sum_{i=1}^{N} X_i|N])$$

$$= E[N \cdot Var(X_i) + Var[N \cdot E[X_i]]$$

$$= Var[X_i] \cdot E[N] + (E[X_i]^2) \cdot Var[N]$$

3 Additional Examples

1. **Theoretical Exercise 7.26**

   Show that if X and Y are independent, then

   $$E[X|Y = y] = E[X] \text{ for all } y$$

   a. In the discrete case:

   From the definition of conditional expectation we have

   $$E[X|Y = y] = \sum_i x_i P[X = x_i|Y = y] = \sum_i x_i \frac{P[X = x_i, Y = y]}{P[Y = y]}$$

Since X and Y are independent,

$$\frac{P[X = x_i, Y = y]}{P[Y = y]} = \frac{P[X = x_i]P[Y = y]}{P[Y = y]} = P[X = x_i]$$

So,

$$\begin{aligned} E[X|Y = y] &= \sum_i x_i P[X = x_i|Y = y] \\ &= \sum_i x_i P[X = x_i] \\ &= E[X] \text{ for all } y \end{aligned}$$

b. In the continuous case:

Similarly from the definition of conditional expectation,

$$E[X|Y = y] = \int x \; \frac{f(x, y)}{f_Y(y)} \; dx$$

Since X and Y are independent,

$$\frac{f(x, y)}{f_Y(y)} = \frac{f_X(x)f_Y(y)}{f_Y(y)} = f_X(x)$$

Therefore

$$\begin{aligned} E[X|Y = y] &= \int x \; \frac{f(x, y)}{f_Y(y)} \; dx \\ &= \int x \; f_X(x) \; dx \\ &= E[X] \text{ for all } y \end{aligned}$$

2. **Theoretical Exercise 7.30**

Let $X_1, ..., X_n$ be independent and identically distributed random variables. Find

$$E[X_1|X_1 + ... + X_n = x]$$

We know that

$$E[X_1 + ... + X_n|X_1 + ... + X_n = x] = x$$

And, by properties of conditional expectation,

$$
\begin{aligned}
E[X_1 + ... + X_n | X_1 + ... + X_n] &= E\left[\sum_{i=1}^{n} X_i | X_1 + ... + X_n = x\right] \\
&= \sum_{i=1}^{n} E[X_i | X_1 + ... + X_n = x] \\
x &= n\ E[X_1 | X_1 + ... + X_n = x]
\end{aligned}
$$

Rearranging the above equation,

$$
E[X_1 | X_1 + ... + X_n = x] = \frac{x}{n}
$$