Math 408 - Mathematical Statistics

Lecture 7. Conditional Expectation and Conditional Variance

February 1, 2013

Definition

Suppose that X and Y are random variables.

<u>Q:</u> What is the mean of X among those times when $Y = y$?

<u>A:</u> It is the mean of X as before, but instead of $f_X(x)$ we use $f_{X|Y}(x|y)$.

Definition

The **conditional expectation** of X given $Y = y$ is

$$\mathbb{E}[X|Y = y] = \begin{cases} \sum_x x f_{X|Y}(x|y), & \text{discrete case;} \\ \int x f_{X|Y}(x|y)dx, & \text{continuous case.} \end{cases}$$

If $Z = r(X, Y)$ is a new random variable, then

$$\mathbb{E}[Z|Y = y] = \begin{cases} \sum_x r(x, y) f_{X|Y}(x|y), & \text{discrete case;} \\ \int r(x, y) f_{X|Y}(x|y)dx, & \text{continuous case.} \end{cases}$$

<u>Important Remark:</u>

- $\mathbb{E}[X]$ is a number
- $\mathbb{E}[X|Y = y]$ is a function of y

Conditional Expectation

Question: What is $\mathbb{E}[X|Y=y]$ before we observe the value y of Y?

Answer: Before we observe Y, we don't know the value of $\mathbb{E}[X|Y=y]$, it is uncertain, so it is a random variable which we denote $\mathbb{E}[X|Y]$.

$\mathbb{E}[X|Y]$ is the random variable whose value is $\mathbb{E}[X|Y=y]$ when $Y=y$.

Example 1:
Suppose we draw

$$X \sim U(0,1)$$

After we observe $X=x$, we draw

$$Y|X=x \sim U(x,1)$$

Find $\mathbb{E}[Y|X=x]$.
Answer:

$$\mathbb{E}[Y|X=x] = \frac{x+1}{2}, \qquad \text{as intuitively expected}$$

Note that $\mathbb{E}[Y|X] = \frac{X+1}{2}$ is a random variable whose value is the number $\mathbb{E}[Y|X=x] = \frac{x+1}{2}$ once $X=x$ is observed.

The Rule of Iterated Expectations

Theorem

For random variables X and Y, assuming the expectations exist, we have

$$\mathbb{E}[\mathbb{E}[Y|X]] = \mathbb{E}[Y] \quad \textit{and} \quad \mathbb{E}[\mathbb{E}[X|Y]] = \mathbb{E}[X]$$

More generally, for any function $r(x, y)$ we have

$$\mathbb{E}[\mathbb{E}[r(X,Y)|X]] = \mathbb{E}[r(X,Y)] \quad \textit{and} \quad \mathbb{E}[\mathbb{E}[r(X,Y)|Y]] = \mathbb{E}[r(X,Y)]$$

Example 2: Compute $\mathbb{E}[Y]$ in Example 1.

Answer:

$$\mathbb{E}[Y] = \mathbb{E}[\mathbb{E}[Y|X]] = \mathbb{E}\left[\frac{X+1}{2}\right] = \frac{1/2+1}{2} = \frac{3}{4}$$

Conditional Variance

Recall, that "unconditional" variance of random variable Y is

$$\mathbb{V}[Y] = \mathbb{E}[(Y - \mathbb{E}[Y])^2]$$

Therefore, it is natural to define **conditional variance** of Y given that $X = x$ as follows (replace all expectations by conditional expectations):

$$\mathbb{V}[Y|X = x] = \mathbb{E}[(Y - \mathbb{E}[Y|X = x])^2|X = x]$$

Denote $\mathbb{E}[Y|X = x]$ by $\mu_Y(x)$. Then

$$\mathbb{V}[Y|X = x] = \int (y - \mu_Y(x))^2 f_{Y|X}(y|x)dy$$

- $\mathbb{V}[Y]$ is a number, $\mathbb{V}[Y|X = x]$ is a function of x

Theorem

For random variables X and Y

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

Example: Statistical Analysis of a Disease

- Draw a state at random from the US.
- Let Q be the proportion of people in that state with a certain disease. Q is a random variable since it varies from state to state, and state is picked at random.
  - Suppose that Q has a uniform distribution on $(0, 1)$, $Q \sim U(0, 1)$.
  - This assumption is natural if we don't have any information about the disease.
- Draw n people at random from the state, and let X be the number of those people who have the disease.
  - Given $Q = q$, it is natural to model X as a Binomial variable, $X|Q = q \quad \sim \quad \mathrm{Bin}(n, q)$.

Problem: Find $\mathbb{E}[X]$ and $\mathbb{V}[X]$

Answer:

$$\mathbb{E}[X] = \frac{n}{2}$$

$$\mathbb{V}[X] = \frac{n}{6} + \frac{n^2}{12}$$

Summary

- The conditional expectation of X given $Y=y$ is

$$\mathbb{E}[X|Y=y] = \begin{cases} \sum_x x f_{X|Y}(x|y), & \text{discrete case;} \\ \int x f_{X|Y}(x|y)dx, & \text{continuous case.} \end{cases}$$

  - $\mathbb{E}[X]$ is a number
  - $\mathbb{E}[X|Y=y]$ is a function of y
  - $\mathbb{E}[X|Y]$ is the random variable whose value is $\mathbb{E}[X|Y=y]$ when $Y=y$
- The Rule of Iterated Expectations

$$\mathbb{E}\mathbb{E}[Y|X] = \mathbb{E}[Y] \quad \text{and} \quad \mathbb{E}\mathbb{E}[X|Y] = \mathbb{E}[X]$$

- The conditional variance of X given $Y=y$ is

$$\mathbb{V}[X|Y=y] = \mathbb{E}[(X-\mathbb{E}[X|Y=y])^2|Y=y]$$

  - $\mathbb{V}[X]$ is a number
  - $\mathbb{V}[X|Y=y]$ is a function of y
  - $\mathbb{V}[X|Y]$ is the random variable whose value is $\mathbb{V}[X|Y=y]$ when $Y=y$
- For random variables X and Y

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