6.6 Conditional expectation given a random variable

Unit 6: Joint Distributions and Conditional Expectation

Adapted from Blitzstein-Hwang Chapters 7 and 9.

In this section we introduce conditional expectation given a random variable. That is, we want to understand what it means to write $E(Y|X)$ for an r.v. X. We will see that $E(Y|X)$ is a *random variable* that is, in a certain sense, our best prediction of Y, assuming we get to know X.

The key to understanding $E(Y|X)$ is first to understand $E(Y|X=x)$. Since $X=x$ is an event, $E(Y|X=x)$ is just the conditional expectation of Y given this event, and it can be computed using the conditional distribution of Y given $X=x$.

If Y is discrete, we use the conditional PMF $P(Y=y|X=x)$ in place of the unconditional PMF $P(Y=y)$:

$$E(Y|X=x) = \sum_y yP(Y=y|X=x).$$

Analogously, if Y is continuous, we use the conditional PDF $f_{Y|X}(y|x)$ in place of the unconditional PDF:

$$E(Y|X=x) = \int_{-\infty}^{\infty} yf_{Y|X}(y|x)dy.$$

Notice that because we sum or integrate over y, $E(Y|X=x)$ is a function of x only. We can give this function a name, like g: let $g(x) = E(Y|X=x)$. We define $E(Y|X)$ as the random variable obtained by finding the form of the function $g(x)$, then *plugging in* X *for* x.

> **Definition 6.6.1 (Conditional expectation given an r.v.).**
>
> Let $g(x) = E(Y|X=x)$. Then the *conditional expectation of* Y *given* X, denoted $E(Y|X)$, is defined to be the random variable $g(X)$. In other words, if after doing the experiment X crystallizes into x, then $E(Y|X)$ crystallizes into $g(x)$.

> **Warning 6.6.2.**

The notation in this definition sometimes causes confusion. It does *not* say "$g(x) = E(Y|X = x)$, so $g(X) = E(Y|X = X)$, which equals $E(Y)$ because $X = X$ is always true". Rather, we should first compute the function $g(x)$, *then* plug in X for x. For example, if $g(x) = x^2$, then $g(X) = X^2$.

☣ **Warning 6.6.3.**

By definition, $E(Y|X)$ is a function of X, so it is a random variable. Thus it makes sense to compute quantities like $E(E(Y|X))$ and $\mathrm{Var}(E(Y|X))$, the mean and variance of the r.v. $E(Y|X)$. It is easy to be ensnared by category errors when working with conditional expectation, so we should always keep in mind that conditional expectations of the form $E(Y|A)$ are numbers, while those of the form $E(Y|X)$ are random variables.

Example 6.6.4.

Suppose we have a stick of length 1 and break the stick at a point X chosen uniformly at random. Given that $X = x$, we then choose another breakpoint Y uniformly on the interval $[0, x]$. Find $E(Y|X)$, and its mean and variance.

Solution

From the description of the experiment, $X \sim \mathrm{Unif}(0, 1)$ and $Y|X = x \sim \mathrm{Unif}(0, x)$. Then $E(Y|X = x) = x/2$, so by plugging in X for x, we have

$$E(Y|X) = X/2.$$

The expected value of $E(Y|X)$ is

$$E(E(Y|X)) = E(X/2) = 1/4.$$

(We will show in the next section that a general property of conditional expectation is that $E(E(Y|X)) = E(Y)$, so it also follows that $E(Y) = 1/4$.) The variance of $E(Y|X)$ is

$$\mathrm{Var}(E(Y|X)) = \mathrm{Var}(X/2) = 1/48.$$

Learn About Verified Certificates

© All Rights Reserved