CONDITIONAL DISTRIBUTIONS AND MOMENTS

EXAMPLE

Let X and Y have the joint probability mass function specified in the following table.

Joint Probability Mass Function of X and Y

		x			
		0	1	2	
	0	0	.2	.1	.3
y	1	.2	.4	0	.6
	2	.1	0	0	.1
		.3	.6	.1	1.0

The conditional distribution of Y given $X = 1$ is obtained by restricting attention to the column of values $f_{X,Y}(1, y)$ corresponding to $X = 1$. This needs to be rescaled to distribute probability one.

Conditional Probability Mass Function of Y given $X = 1$

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

The **conditional probability mass function** of Y given $X = x$ is defined as

$$f_{Y|X=x}(y) = \frac{f_{X,Y}(x, y)}{f_X(x)}$$

provided $f_X(x) > 0$.

This distribution has moments(conditional) just as any other distribution. For any function $g(\cdot)$, the conditional expectation of $g(Y)$ given $X = x$ is defined as

$$E[Y|X = x] = \sum_y g(y) f_{Y|X=x}(y) \qquad \text{discrete case}$$

$$\int_{-\infty}^{\infty} g(y) f_{Y|X=x}(y) \qquad \text{continuous case}$$

In particular, with $g(y) = y$,

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

In general, the conditional variance is defined as

$$var(Y|X = x) = \sum_y (y - E[Y|X = x])^2 f_{Y|X=x}(y)$$

$$= \sum_y (y^2 - 2yE[Y|X=x] + (E[Y|X=x])^2) f_{Y|X=x}(y)$$

$$= \sum_y y^2 f_{Y|X=x}(y) \; - \; 2\sum_y yE[Y|X=x] f_{Y|X=x}(y) + (E[Y|X=x])^2 \sum_y f_{Y|X=x}(y)$$

$$= \sum_y y^2 f_{Y|X=x}(y) - (E[Y|X=x])^2 = E[Y^2|X=x] - (E[Y|X=x])^2$$

Since, in our example,

$$E(Y^2|X=1) = 0^2 \times \frac{1}{3} + 1^2 \times \frac{2}{3} \times .6 + 2^2 \times 0 = \frac{2}{3}$$

we have $var(Y|X=1) = \frac{2}{3} - (\frac{2}{3})^2 = \frac{2}{9}$

More Properties of Conditional Expectation

Above, we have shown that the conditional variance

$$var(Y|X=x) = \sum_y (y - E[Y|X=x])^2 f_{Y|X=x}(y) = E[Y^2|X=x] - (E[Y|X=x])^2$$

Notice that the conditional expectation of Y given $x = x$,

$$E[Y|X=x] = \int_{-\infty}^{\infty} y \frac{f_{X,Y}(x,y)}{f_X(x)} dy$$

is a function of the observed value x of X. Let us now consider X to be random and write $E[Y|X]$ for the random quantity that takes value $E[Y|X=x]$ when $X = x$. We establish two important properties regarding conditional expected values.

Property 1. $E[E[Y|X]] = E[Y]$

Property 2. $E[E[XY|X]] = E[XY]$

First, $E[Y|X]$ has expected value

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

Canceling $f_X(x)$, over the range where it is positive, this last double integral becomes

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y f_{X,Y}(x,y) dx dy = \int_{-\infty}^{\infty} y f_Y(y) dy = E[Y]$$

(When $E[|Y|] < \infty$, the integration can be done in any order).

To establish the second property, we notice that

$$E[XY|X=x] = \int_{-\infty}^{\infty} xy \frac{f_{X,Y}(x,y)}{f_X(x)} dy = xE[Y|X=x]$$

Considering X as random, we write $E[E[XY|X]]$ which has expected value

$$E[E[XY|X]] = \int_{-\infty}^{\infty} E[XY|X=x] f_X(x) dx = \int_{-\infty}^{\infty} (\int_{-\infty}^{\infty} xy \frac{f_{X,Y}(x,y)}{f_X(x)} dy) f_X(x) dx$$

$$= \int_{-\infty}^{\infty} (\int_{-\infty}^{\infty} xy f_{X,Y}(x,y) dx dy = E[XY]$$

as claimed.

EXAMPLE Suppose the joint distribution of X and Y is such that

$$E[X] = 5 \quad E[X^2] = 30 \quad E[Y|X] = 2 + 3X \quad \mathrm{Var}(Y) = 81$$

Find (a) Cov(X, Y) and (b) Corr (X, Y).

The answer requires the two properties above.

Solution to Example Using the properties above

$$E[Y] = E[E[Y|X]] = E[2 + 3X] = 2 + 3E[X] = 2 + 3(5) = 17$$

$$E[XY] = E[E[XY|X]] = E[2X + 3X^2]$$
$$= 2E[X]) + 3E[X^2] = 10 + 3(30) = 100$$

Consequently, Cov($X, Y) = E[XY] - E[X]E[Y] = 100 - (17)(5) = 15$ and $\mathrm{Var}(X) = E[X^2] - (E[X])^2 = 30 - 5^2 = 5$ so

$$\mathrm{Corr}(X, Y) = \frac{\mathrm{Cov}(X, Y)}{\sqrt{\mathrm{Var}(X)}\sqrt{\mathrm{Var}(Y)}} = \frac{15}{\sqrt{5}\sqrt{81}}$$

The next decomposition of variance is useful in many statistical applications.

Property 3. $var(Y) = E[var(Y|X)] + var(E[Y|X])$

Proof:

Write $Y - \mu_Y = Y - E[Y|X] + E[Y|X] - \mu_Y$ and square

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

Take expected values with respect to the joint distribution of (X, Y). The left hand side gives $var(Y)$. For the first term on the right hand side, we take the iterated expected value by conditioning on X. and using

$$E[(Y - E[Y|X])^2 \; |X = x] = \mathrm{var}(Y|X = x)$$

Next, taking the expected value over X produces the first term in Property 3. By Property 1 $E[E[Y|X]\ ] = E[Y] = \mu_Y$ so the second term is the conditional variance of $E[Y|X]$. Finally, the conditional expected value of the cross-term

$$E\left[[2(Y - E[Y|X])(E[Y|X] - \mu_Y)|X\ \right]]$$
$$= 2(E[Y|X] - \mu_Y)E\left[(Y - E[Y|X])|X\ \right]] = 0$$

so the unconditional expected value is also 0 and Property 3 follows.