Economics 103 – Statistics for Economists

Minsu Chang

University of Pennsylvania

Lecture # 10

Definition of Conditional PMF

How does the distribution of y change with x?

$$p_{Y|X}(y|x) = P(Y = y|X = x) = \frac{P(Y = y \cap X = x)}{P(X = x)} = \frac{p_{XY}(x, y)}{p_X(x)}$$

Conditional PMF of Y given $X = 2$

			Y		
		1	2	3	
	0	1/8	0	0	1/8
X	1	0	1/4	1/8	3/8
	2	0	1/4	1/8	3/8
	3	1/8	0	0	1/8

$$p_{Y|X}(1|2) = \frac{p_{XY}(2,1)}{p_X(2)} = \frac{0}{3/8} = 0$$

$$p_{Y|X}(2|2) = \frac{p_{XY}(2,2)}{p_X(2)} = \frac{1/4}{3/8} = 2/3$$

$$p_{Y|X}(3|2) = \frac{p_{XY}(2,3)}{p_X(2)} = \frac{1/8}{3/8} = 1/3$$

What is $p_{X|Y}(1|2)$?

		|	Y			|
		|	1	2	3	|
	0	|	1/8	0	0	|
X	1	|	0	1/4	1/8	|
	2	|	0	1/4	1/8	|
	3	|	1/8	0	0	|
		|	1/4	1/2	1/4	|

What is $p_{X|Y}(1|2)$?

		Y			
		1	2	3	
	0	1/8	0	0	
X	1	0	1/4	1/8	
	2	0	1/4	1/8	
	3	1/8	0	0	
		1/4	1/2	1/4	

$$p_{X|Y}(1|2) = \frac{p_{XY}(1,2)}{p_Y(2)} = \frac{1/4}{1/2} = 1/2 \qquad (1)$$

What is $p_{X|Y}(1|2)$?

		Y			
		1	2	3	
	0	1/8	0	0	
X	1	0	1/4	1/8	
	2	0	1/4	1/8	
	3	1/8	0	0	
		1/4	1/2	1/4	

$$p_{X|Y}(1|2) = \frac{p_{XY}(1,2)}{p_Y(2)} = \frac{1/4}{1/2} = 1/2 \tag{1}$$

Similarly:

$$p_{X|Y}(0|2) = 0, \quad p_{X|Y}(2|2) = 1/2, \quad p_{X|Y}(3|2) = 0$$

Independent RVs: Joint Equals Product of Marginals

Definition

Two discrete RVs are independent if and only if

$$p_{XY}(x, y) = p_X(x)p_Y(y)$$

for all pairs (x, y) in the support.

Equivalent Definition

$$p_{Y|X}(y|x) = p_Y(y) \text{ and } p_{X|Y}(x|y) = p_X(x)$$

for all pairs (x, y) in the support.

Are X and Y Independent?

(A = YES, B = NO)

		Y			
		1	2	3	
X	0	1/8	0	0	1/8
	1	0	1/4	1/8	3/8
	2	0	1/4	1/8	3/8
	3	1/8	0	0	1/8
		1/4	1/2	1/4	

Are X and Y Independent?

(A = YES, B = NO)

		Y			
		1	2	3	
	0	1/8	0	0	1/8
X	1	0	1/4	1/8	3/8
	2	0	1/4	1/8	3/8
	3	1/8	0	0	1/8
		1/4	1/2	1/4	

$$p_{XY}(2,1) = 0$$
$$p_X(2) \times p_Y(1) = (3/8) \times (1/4) \neq 0$$

Therefore X and Y are *not* independent.

Conditional Expectation

Intuition

$E[Y|X] = $ "best guess" about realization of Y after observing realization of X.

$E[Y|X]$ is a Random Variable

While $E[Y]$ is a constant, $E[Y|X]$ is a function of X, hence a Random Variable.

$E[Y|X = x]$ is a Constant

The constant $E[Y|X = x]$ is the "guess" of Y if we see $X = x$.

Calculating $E[Y|X = x]$

Take the mean of the conditional pmf of Y given $X = x$.

Conditional Expectation: $E[Y|X=2]$

		Y			
		1	2	3	
	0	1/8	0	0	1/8
X	1	0	1/4	1/8	3/8
	2	0	1/4	1/8	3/8
	3	1/8	0	0	1/8
		1/4	1/2	1/4	

We showed above that the conditional pmf of $Y|X=2$ is:

$$p_{Y|X}(1|2) = 0 \qquad p_{Y|X}(2|2) = 2/3 \qquad p_{Y|X}(3|2) = 1/3$$

Hence

$$E[Y|X=2] = 2 \times 2/3 + 3 \times 1/3 = 7/3$$

Conditional Expectation: $E[Y|X = 0]$

		Y			
		1	2	3	
	0	1/8	0	0	1/8
X	1	0	1/4	1/8	3/8
	2	0	1/4	1/8	3/8
	3	1/8	0	0	1/8
		1/4	1/2	1/4	

The conditional pmf of $Y|X = 0$ is

$$p_{Y|X}(1|0) = 1 \qquad p_{Y|X}(2|0) = 0 \qquad p_{Y|X}(3|0) = 0$$

Hence $E[Y|X = 0] = 1$

Calculate $E[Y|X = 3]$

			Y		
		1	2	3	
X	0	1/8	0	0	1/8
	1	0	1/4	1/8	3/8
	2	0	1/4	1/8	3/8
	3	1/8	0	0	1/8
		1/4	1/2	1/4	

The conditional pmf of $Y|X = 3$ is

$$p_{Y|X}(1|3) = 1 \qquad p_{Y|X}(2|3) = 0 \qquad p_{Y|X}(3|3) = 0$$

Hence $E[Y|X = 3] = 1$

Calculate $E[Y|X = 1]$

		Y			
		1	2	3	
	0	1/8	0	0	1/8
X	1	0	1/4	1/8	3/8
	2	0	1/4	1/8	3/8
	3	1/8	0	0	1/8
		1/4	1/2	1/4	

Calculate $E[Y|X=1]$

		Y			
		1	2	3	
	0	1/8	0	0	1/8
X	1	0	1/4	1/8	3/8
	2	0	1/4	1/8	3/8
	3	1/8	0	0	1/8
		1/4	1/2	1/4	

The conditional pmf of $Y|X=1$ is

$$p_{Y|X}(1|1) = 0 \qquad p_{Y|X}(2|1) = 2/3 \qquad p_{Y|X}(3|1) = 1/3$$

Hence

$$E[Y|X=1] = 2 \times 2/3 + 3 \times 1/3 = 7/3$$

$E[Y|X]$ is a Random Variable

For this example:

$$E[Y|X] = \begin{cases} 1 & X = 0 \\ 7/3 & X = 1 \\ 7/3 & X = 2 \\ 1 & X = 3 \end{cases}$$

From above the marginal distribution of X is:

$$P(X = 0) = 1/8 \qquad P(X = 1) = 3/8$$

$$P(X = 2) = 3/8 \qquad P(X = 3) = 1/8$$

$E[Y|X]$ takes the value 1 with prob. 1/4 and 7/3 with prob. 3/4.

The Law of Iterated Expectations

$E[Y|X]$ is an RV so what is its expectation?

For any RVs X and Y

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

Law of Iterated Expectations for Our Example

Marginal pmf of Y

$$P(Y = 1) = 1/4$$
$$P(Y = 2) = 1/2$$
$$P(Y = 3) = 1/4$$

$$E[Y] = 1 \times 1/4 + 2 \times 1/2 + 3 \times 1/4$$
$$= 2$$

$E[Y|X]$

$$E[Y|X] = \begin{cases} 1 & \text{w/ prob. } 1/4 \\ 7/3 & \text{w/ prob. } 3/4 \end{cases}$$

$$E[E[Y|X]] = 1 \times 1/4 + 7/3 \times 3/4$$
$$= 2$$

Expectation of Function of Two Discrete RVs

$$E[g(X, Y)] = \sum_x \sum_y g(x, y) p_{XY}(x, y)$$

Some Extremely Important Examples

Let $\mu_X = E[X], \mu_Y = E[Y]$

Covariance

$$\sigma_{XY} = Cov(X, Y) = E[(X - \mu_X)(Y - \mu_Y)]$$

Correlation

$$\rho_{XY} = Corr(X, Y) = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}$$

Shortcut Formula for Covariance

Much easier for calculating:

$$Cov(X, Y) = E[XY] - E[X]E[Y]$$

I'll mention this again in a few slides...

Calculating $Cov(X, Y)$

		Y			
		1	2	3	
	0	1/8	0	0	1/8
X	1	0	1/4	1/8	3/8
	2	0	1/4	1/8	3/8
	3	1/8	0	0	1/8
		1/4	1/2	1/4	

$$E[X] = 3/8 + 2 \times 3/8 + 3 \times 1/8 = 3/2$$

$$E[Y] = 1/4 + 2 \times 1/2 + 3 \times 1/4 = 2$$

Calculating $Cov(X, Y)$

		Y			
		1	2	3	
	0	1/8	0	0	1/8
X	1	0	1/4	1/8	3/8
	2	0	1/4	1/8	3/8
	3	1/8	0	0	1/8
		1/4	1/2	1/4	

$$E[X] = 3/8 + 2 \times 3/8 + 3 \times 1/8 = 3/2$$

$$E[Y] = 1/4 + 2 \times 1/2 + 3 \times 1/4 = 2$$

$$E[XY] = 1/4 \times (2 + 4) + 1/8 \times (3 + 6 + 3)$$
$$= 3$$

Calculating $Cov(X, Y)$

		Y			
		1	2	3	
	0	1/8	0	0	1/8
X	1	0	1/4	1/8	3/8
	2	0	1/4	1/8	3/8
	3	1/8	0	0	1/8
		1/4	1/2	1/4	

$$E[X] = 3/8 + 2 \times 3/8 + 3 \times 1/8 = 3/2$$

$$E[Y] = 1/4 + 2 \times 1/2 + 3 \times 1/4 = 2$$

$$E[XY] = 1/4 \times (2 + 4) + 1/8 \times (3 + 6 + 3)$$
$$= 3$$

$$Cov(X, Y) = E[XY] - E[X]E[Y]$$

Calculating $Cov(X, Y)$

		Y			
		1	2	3	
	0	1/8	0	0	1/8
X	1	0	1/4	1/8	3/8
	2	0	1/4	1/8	3/8
	3	1/8	0	0	1/8
		1/4	1/2	1/4	

$$E[X] = 3/8 + 2 \times 3/8 + 3 \times 1/8 = 3/2$$

$$E[Y] = 1/4 + 2 \times 1/2 + 3 \times 1/4 = 2$$

$$E[XY] = 1/4 \times (2 + 4) + 1/8 \times (3 + 6 + 3)$$
$$= 3$$

$$Cov(X, Y) = E[XY] - E[X]E[Y]$$
$$= 3 - 3/2 \times 2 = 0$$

Calculating $Cov(X, Y)$

		Y			
		1	2	3	
	0	1/8	0	0	1/8
	1	0	1/4	1/8	3/8
X	2	0	1/4	1/8	3/8
	3	1/8	0	0	1/8
		1/4	1/2	1/4	

$$E[X] = 3/8 + 2 \times 3/8 + 3 \times 1/8 = 3/2$$

$$E[Y] = 1/4 + 2 \times 1/2 + 3 \times 1/4 = 2$$

$$E[XY] = 1/4 \times (2+4) + 1/8 \times (3+6+3)$$
$$= 3$$

$$Cov(X, Y) = E[XY] - E[X]E[Y]$$
$$= 3 - 3/2 \times 2 = 0$$

$$Corr(X, Y) = Cov(X, Y)/[SD(X)SD(Y)] = 0$$

Zero Covariance versus Independence

- From this example we learn that zero covariance (correlation) *does not* imply independence.

- However, it turns out that independence *does* imply zero covariance (correlation).

Linearity of Expectation, Again

Holds for Continuous RVs as well, but different proof.

In general, $E[g(X, Y)] \neq g(E[X], E[Y])$. The key exception is when g is a linear function:

$$\boxed{E[aX + bY + c] = aE[X] + bE[Y] + c}$$

where X, Y are random variables and a, b, c are constants.

Application: Shortcut Formula for Variance

By the Linearity of Expectation,

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

Application: Shortcut Formula for Variance

By the Linearity of Expectation,

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

Application: Shortcut Formula for Variance

By the Linearity of Expectation,

$$
\begin{aligned}
Var(X) &= E[(X - \mu)^2] = E[X^2 - 2\mu X + \mu^2] \\
&= E[X^2] - 2\mu E[X] + \mu^2 \\
&=
\end{aligned}
$$

Application: Shortcut Formula for Variance

By the Linearity of Expectation,

$$
\begin{aligned}
Var(X) &= E[(X - \mu)^2] = E[X^2 - 2\mu X + \mu^2] \\
&= E[X^2] - 2\mu E[X] + \mu^2 \\
&= E[X^2] - 2\mu^2 + \mu^2 \\
&=
\end{aligned}
$$

Application: Shortcut Formula for Variance

By the Linearity of Expectation,

$$
\begin{aligned}
Var(X) &= E[(X - \mu)^2] = E[X^2 - 2\mu X + \mu^2] \\
&= E[X^2] - 2\mu E[X] + \mu^2 \\
&= E[X^2] - 2\mu^2 + \mu^2 \\
&= E[X^2] - \mu^2
\end{aligned}
$$

We saw in a previous lecture that it's typically much easier to calculate variances using the shortcut formula.

Another Application: Shortcut Formula for Covariance

Similar to Shortcut for Variance: in fact $Var(X) = Cov(X, X)$

$$
\begin{aligned}
Cov(X, Y) &= E[(X - \mu_X)(Y - \mu_Y)] \\
&= E[XY - \mu_X Y - \mu_Y X + \mu_X \mu_Y] \\
&\vdots \\
&= E[XY] - E[X]E[Y]
\end{aligned}
$$

You'll fill in the details for homework...

Expected Value of Sum = Sum of Expected Values

Repeatedly applying the linearity of expectation,

$$E[X_1 + X_2 + \ldots + X_n] = E[X_1] + E[X_2] + \ldots + E[X_n]$$

regardless of how the RVs $X_1, \ldots, X_n$ are related to each other. In particular it doesn't matter if they're dependent or independent.

Independent and Identically Distributed (iid) RVs

Example

$X_1, X_2, \ldots X_n \sim$ iid Bernoulli(p)

Independent

Realization of one of the RVs gives no information about the others.

Identically Distributed

Each X_i is the same kind of RV, with the same values for any parameters. (Hence same pmf, cdf, mean, variance, etc.)

Binomial(n, p) Random Variable

Definition

Sum of n independent Bernoulli RVs, each with probability of "success," i.e. 1, equal to p

Parameters

p = probability of "success," n = # of trials

Support

$\{0, 1, 2, \ldots, n\}$

Using Our New Notation

Let $X_1, X_2, \ldots, X_n \sim$ iid Bernoulli(p), $Y = X_1 + X_2 + \ldots + X_n$.
Then $Y \sim$ Binomial(n, p).

Which of these is the PMF of a Binomial(n, p) RV?

(a) $p(x) = p^x(1-p)^{n-x}$

(b) $p(x) = \binom{n}{x}p^x(1-p)^{n-x}$

(c) $p(x) = \binom{x}{n}p^x$

(d) $p(x) = \binom{n}{x}p^{n-x}(1-p)^x$

(e) $p(x) = p^n(1-p)^x$

Which of these is the PMF of a Binomial(n, p) RV?

(a) $p(x) = p^x(1-p)^{n-x}$

(b) $p(x) = \binom{n}{x}p^x(1-p)^{n-x}$

(c) $p(x) = \binom{x}{n}p^x$

(d) $p(x) = \binom{n}{x}p^{n-x}(1-p)^x$

(e) $p(x) = p^n(1-p)^x$

$$p(x) = \binom{n}{x}p^x(1-p)^{n-x}$$

Expected Value of Binomial RV

Use the fact that a Binomial(n, p) RV is defined as the sum of n iid Bernoulli(p) Random Variables and the Linearity of Expectation:

$$
\begin{aligned}
E[Y] &= E[X_1 + X_2 + \ldots + X_n] = E[X_1] + E[X_2] + \ldots + E[X_n] \\
&= p + p + \ldots + p \\
&= np
\end{aligned}
$$

Variance of a Sum $\neq$ Sum of Variances!

$$Var(aX + bY) = E\left[\{(aX + bY) - E[aX + bY]\}^2\right]$$

Variance of a Sum $\neq$ Sum of Variances!

$$
\begin{aligned}
Var(aX + bY) &= E\left[\{(aX + bY) - E[aX + bY]\}^2\right] \\
&= E\left[\{a(X - \mu_X) + b(Y - \mu_Y)\}^2\right]
\end{aligned}
$$

Variance of a Sum $\neq$ Sum of Variances!

$$
\begin{aligned}
Var(aX + bY) &= E\left[\{(aX + bY) - E[aX + bY]\}^2\right] \\
&= E\left[\{a(X - \mu_X) + b(Y - \mu_Y)\}^2\right] \\
&= E\left[a^2(X - \mu_X)^2 + b^2(Y - \mu_Y)^2 + 2ab(X - \mu_X)(Y - \mu_Y)\right]
\end{aligned}
$$

Variance of a Sum $\neq$ Sum of Variances!

$$
\begin{aligned}
Var(aX + bY) &= E\left[\{(aX + bY) - E[aX + bY]\}^2\right] \\
&= E\left[\{a(X - \mu_X) + b(Y - \mu_Y)\}^2\right] \\
&= E\left[a^2(X - \mu_X)^2 + b^2(Y - \mu_Y)^2 + 2ab(X - \mu_X)(Y - \mu_Y)\right] \\
&= a^2 E[(X - \mu_X)^2] + b^2 E[(Y - \mu_Y)^2] + 2ab E[(X - \mu_X)(Y - \mu_Y)]
\end{aligned}
$$

Variance of a Sum $\neq$ Sum of Variances!

$$
\begin{aligned}
Var(aX + bY) &= E\left[\{(aX + bY) - E[aX + bY]\}^2\right] \\
&= E\left[\{a(X - \mu_X) + b(Y - \mu_Y)\}^2\right] \\
&= E\left[a^2(X - \mu_X)^2 + b^2(Y - \mu_Y)^2 + 2ab(X - \mu_X)(Y - \mu_Y)\right] \\
&= a^2 E[(X - \mu_X)^2] + b^2 E[(Y - \mu_Y)^2] + 2ab E[(X - \mu_X)(Y - \mu_Y)] \\
&= a^2 Var(X) + b^2 Var(Y) + 2ab Cov(X, Y)
\end{aligned}
$$

Since $\sigma_{XY} = \rho \sigma_X \sigma_Y$, this is sometimes written as:

$$
\boxed{Var(aX + bY) = a^2 \sigma_X^2 + b^2 \sigma_Y^2 + 2ab\rho\sigma_X\sigma_Y}
$$

Independence $\Rightarrow Var(X + Y) = Var(X) + Var(Y)$

X and Y independent $\implies Cov(X, Y) = 0$. Hence, independence implies

$$
\begin{aligned}
Var(X + Y) &= Var(X) + Var(Y) + 2Cov(X, Y) \\
&= Var(X) + Var(Y)
\end{aligned}
$$

Also true for three or more RVs

If $X_1, X_2, \ldots, X_n$ are independent, then

$$Var(X_1 + X_2 + \ldots X_n) = Var(X_1) + Var(X_2) + \ldots + Var(X_n)$$

Crucial Distinction

Expected Value

Always true that

$$E[X_1 + X_2 + \ldots + X_n] = E[X_1] + E[X_2] + \ldots + E[X_n]$$

Variance

Not true in general that

$$Var[X_1 + X_2 + \ldots + X_n] = Var[X_1] + Var[X_2] + \ldots + Var[X_n]$$

except in the special case where $X_1, \ldots X_n$ are independent (or at least uncorrelated).

Variance of Binomial Random Variable

Definition from Sequence of Bernoulli Trials

If $X_1, X_2, \ldots, X_n \sim$ iid Bernoulli(p) then

$$Y = X_1 + X_2 + \ldots + X_n \sim \text{Binomial}(n, p)$$

Using Independence

$$
\begin{aligned}
Var[Y] &= Var[X_1 + X_2 + \ldots + X_n] \\
&= Var[X_1] + Var[X_2] + \ldots + Var[X_n] \\
&=
\end{aligned}
$$

Variance of Binomial Random Variable

Definition from Sequence of Bernoulli Trials

If $X_1, X_2, \ldots, X_n \sim$ iid Bernoulli(p) then
$$Y = X_1 + X_2 + \ldots + X_n \sim \text{Binomial}(n, p)$$

Using Independence

$$
\begin{aligned}
Var[Y] &= Var[X_1 + X_2 + \ldots + X_n] \\
&= Var[X_1] + Var[X_2] + \ldots + Var[X_n] \\
&= p(1-p) + p(1-p) + \ldots + p(1-p) \\
&=
\end{aligned}
$$

Variance of Binomial Random Variable

Definition from Sequence of Bernoulli Trials

If $X_1, X_2, \ldots, X_n \sim$ iid Bernoulli(p) then
$$Y = X_1 + X_2 + \ldots + X_n \sim \text{Binomial}(n, p)$$

Using Independence

$$
\begin{aligned}
Var[Y] &= Var[X_1 + X_2 + \ldots + X_n] \\
&= Var[X_1] + Var[X_2] + \ldots + Var[X_n] \\
&= p(1-p) + p(1-p) + \ldots + p(1-p) \\
&= np(1-p)
\end{aligned}
$$