# Economics 103 – Statistics for Economists

Francis J. DiTraglia

University of Pennsylvania

Lecture # 10

# Discrete RVs – Part III

# Definition of Conditional PMF

How does the distribution of $y$ change with $x$?

$$p_{Y|X}(y|x) = P(Y = y | X = x)$$

# Which these is the formula for $p_{Y|X}(y|x)$?

You can figure this out from what you already know about probability, using the definition $p_{Y|X}(y|x) = P(Y = y | X = x)$

(a) $p_X(x)/p_Y(y)$

(b) $p_{XY}(x, y)/p_X(x)$

(c) $p_X(x)p_{XY}(x, y)$

(d) $p_{XY}(x, y)/p_Y(y)$

(e) $p_Y(y)/p_X(x)$

# Conditional PMF from Joint and Marginal

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

Hence,

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

and similarly,

$$\boxed{p_{X|Y}(x|y) = \frac{p_{XY}(x, y)}{p_Y(y)}}$$

# Conditional PMF of $Y$ given $X = 2$

| | | $Y$ | | | |
|---|---|---|---|---|---|
| | | 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 |

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

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

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

# 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 |   |

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

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

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

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

# Independent RVs

### Definition

We say that two discrete RVs are <span style="color:red">independent</span> if and only if their joint pmf equals the product of their marginal pmfs:

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

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

### In Terms of Conditional PMF

From the previous slide, it follows that an equivalent definition of independence is that both conditional pmfs equal the corresponding marginal pmfs: $p_{Y|X}(y|x) = p_Y(y)$ and $p_{X|Y}(x|y) = p_X(x)$ for all $(x, y)$ in the support.

# 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

$E[Y|X]$ is our "best guess" of the realization that $Y$ will take on having observed the realization of $X$.

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

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

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

To get a "best guess" for $Y$, we plug in the realization we observed for $X$: $E[Y|X = x]$ is a constant, our guess of the realization of $Y$.

### 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 | |
| | 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 | |

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 |   |

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

In this particular example we have seen that:

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

But from above we know the marginal distribution of $X$ :

$$P(X = 0) = 1/8 \qquad P(X = 1) = 3/8$$
$$P(X = 2) = 3/8 \qquad P(X = 3) = 1/8$$

Therefore, $E[Y|X]$ is a RV that takes on the value 1 with probability 1/4 and the value 7/3 with probability 3/4.

# The Law of Iterated Expectations

Since $E[Y|X]$ is a random variable, we can ask what its expectation is. It turns out that, for any RVs $X$ and $Y$

$$E\left[E\left[Y|X\right]\right] = E[Y]$$

and this is called the Law of Iterated Expectations. I've posted a proof $\boxed{\text{HERE}}$ for those who are interested.

This will be helpful in Econ 104...

# Law of Iterated Expectations for Our Example

### Marginal pmf of $Y$

$$
\begin{aligned}
P(Y = 1) &= 1/4 \\
P(Y = 2) &= 1/2 \\
P(Y = 3) &= 1/4
\end{aligned}
$$

$$
\begin{aligned}
E[Y] &= 1 \times 1/4 + 2 \times 1/2 + 3 \times 1/4 \\
&= 2
\end{aligned}
$$

### $E[Y|X]$

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

$$
\begin{aligned}
E\left[E\left[Y|X\right]\right] &= 1 \times 1/4 + 7/3 \times 3/4 \\
&= 2
\end{aligned}
$$

# 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

Same For Continuous Random Variables

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:

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

We'll talk more about this in an upcoming lecture...

# 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 |   |

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

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

$$
\begin{aligned}
E[XY] &= 1/4 \times (2 + 4) + 1/8 \times (3 + 6 + 3) \\
&= 3
\end{aligned}
$$

$$
\begin{aligned}
Cov(X, Y) &= E[XY] - E[X]E[Y] \\
&= 3 - 3/2 \times 2 = 0
\end{aligned}
$$

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

Hence, zero covariance (correlation) does *not* imply independence!

However, independence *does* imply zero covariance (correlation)

I've posted a proof ⟨ HERE ⟩ for those who are interested.

# 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:

$$E[aX + bY + c] = aE[X] + bE[Y] + c$$

where $X, Y$ are random variables and $a, b, c$ are constants. I've posted a proof $\boxed{\text{HERE}}$ for those who are interested.

# 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

Joint pmf equals product of marginal pmfs (see last lecture):
Knowing the 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) $p(x) = p^n(1 - p)^x$

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

# 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:

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

Extremely Important:

Variance of Sum $\neq$ Sum of Variances!

# Variance of a Sum

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

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

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

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

As explained above, if $X$ and $Y$ are independent, $Cov(X, Y) = 0$.
Hence, independence implies

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

This is also true for more than two 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

It is always true that
$$E[X_1 + X_2 + \ldots + X_n] = E[X_1] + E[X_2] + \ldots + E[X_n]$$

### Variance

It is not true in general that
$$Var[X_1 + X_2 + \ldots + X_n] = Var[X_1] + Var[X_2] + \ldots + Var[X_n]$$
but it is true in the special case where $X_1, \ldots X_n$ are independent.

# 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

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