

# Economics 103 – Statistics for Economists

Francis J. DiTraglia

University of Pennsylvania

Lecture # 9

# Discrete RVs – Part II

# Linearity of Expectation

Holds for Continuous RVs as well, but proof is different.

Let  $X$  be a RV and  $a, b$  be constants. Then:

$$E[a + bX] = a + bE[X]$$

This is one of the most important facts in the course: the special case in which  $E[g(X)] = g(E[X])$  is  $g = a + bX$ .

## Example: Linearity of Expectation



Let  $X \sim \text{Bernoulli}(1/3)$  and define  $Y = 3X + 2$

1. What is  $E[X]$ ?

## Example: Linearity of Expectation



Let  $X \sim \text{Bernoulli}(1/3)$  and define  $Y = 3X + 2$

1. What is  $E[X]$ ?  $E[X] = 0 \times 2/3 + 1 \times 1/3 = 1/3$

## Example: Linearity of Expectation



Let  $X \sim \text{Bernoulli}(1/3)$  and define  $Y = 3X + 2$

1. What is  $E[X]$ ?  $E[X] = 0 \times 2/3 + 1 \times 1/3 = 1/3$
2. What is  $E[Y]$ ?

## Example: Linearity of Expectation



Let  $X \sim \text{Bernoulli}(1/3)$  and define  $Y = 3X + 2$

1. What is  $E[X]$ ?  $E[X] = 0 \times 2/3 + 1 \times 1/3 = 1/3$
2. What is  $E[Y]$ ?  $E[Y] = E[3X + 2] = 3E[X] + 2 = 3$

## Proof: Linearity of Expectation For Discrete RV

$$\begin{aligned}E[a + bX] &= \sum_{\text{all } x} (a + bx)p(x) \\&= \sum_{\text{all } x} p(x) \cdot a + \sum_{\text{all } x} p(x) \cdot bx \\&= a \sum_{\text{all } x} p(x) + b \sum_{\text{all } x} x \cdot p(x) \\&= a + bE[X]\end{aligned}$$



# Variance and Standard Deviation of a RV

The Defs are the same for continuous RVs, but the method of calculating will differ.

## Variance (Var)

$$\sigma^2 = \text{Var}(X) = E[(X - \mu)^2] = E[(X - E[X])^2]$$

## Standard Deviation (SD)

$$\sigma = \sqrt{\sigma^2} = \text{SD}(X)$$

## Key Point

Variance and std. dev. are *expectations of functions of a RV*

It follows that:

1. Variance and SD are constants
2. To derive facts about them you can use the facts you know about expected value

# How To Calculate Variance for Discrete RV?

Remember: it's just a function of  $X$ !

$$\text{Recall that } \mu = E[X] = \sum_{\text{all } x} xp(x)$$

$$\text{Var}(X) = E[(X - \mu)^2] = \sum_{\text{all } x} (x - \mu)^2 p(x)$$

## Shortcut Formula For Variance

This is *not* the definition, it's a shortcut for doing calculations:

$$\text{Var}(X) = E[(X - \mu)^2] = E[X^2] - (E[X])^2$$

We'll prove this in an upcoming lecture.

## Example: The Shortcut Formula



Let  $X \sim \text{Bernoulli}(1/2)$ .

1. What is  $E[X]$ ?

## Example: The Shortcut Formula



Let  $X \sim \text{Bernoulli}(1/2)$ .

1. What is  $E[X]$ ?  $E[X] = 0 \times 1/2 + 1 \times 1/2 = 1/2$

## Example: The Shortcut Formula



Let  $X \sim \text{Bernoulli}(1/2)$ .

1. What is  $E[X]$ ?  $E[X] = 0 \times 1/2 + 1 \times 1/2 = 1/2$
2. What is  $E[X^2]$ ?

## Example: The Shortcut Formula



Let  $X \sim \text{Bernoulli}(1/2)$ .

1. What is  $E[X]$ ?  $E[X] = 0 \times 1/2 + 1 \times 1/2 = 1/2$
2. What is  $E[X^2]$ ?  $E[X^2] = 0^2 \times 1/2 + 1^2 \times 1/2 = 1/2$



## Example: The Shortcut Formula



Let  $X \sim \text{Bernoulli}(1/2)$ .

1. What is  $E[X]$ ?  $E[X] = 0 \times 1/2 + 1 \times 1/2 = 1/2$
2. What is  $E[X^2]$ ?  $E[X^2] = 0^2 \times 1/2 + 1^2 \times 1/2 = 1/2$
3. What is  $\text{Var}(X)$ ?

## Example: The Shortcut Formula



Let  $X \sim \text{Bernoulli}(1/2)$ .

1. What is  $E[X]$ ?  $E[X] = 0 \times 1/2 + 1 \times 1/2 = 1/2$
2. What is  $E[X^2]$ ?  $E[X^2] = 0^2 \times 1/2 + 1^2 \times 1/2 = 1/2$
3. What is  $\text{Var}(X)$ ?  $E[X^2] - (E[X])^2 = 1/2 - (1/2)^2 = 1/4$

## Variance of Bernoulli RV – via the Shortcut Formula

Step 1 –  $E[X]$

$$\mu = E[X] = \sum_{x \in \{0,1\}} p(x) \cdot x = (1-p) \cdot 0 + p \cdot 1 = p$$

Step 2 –  $E[X^2]$

$$E[X^2] = \sum_{x \in \{0,1\}} x^2 p(x) = 0^2(1-p) + 1^2 p = p$$

Step 3 – Combine with Shortcut Formula

$$\sigma^2 = \text{Var}[X] = E[X^2] - (E[X])^2 = p - p^2 = p(1-p)$$

## Variance of Bernoulli RV – Without Shortcut

You will fill in the missing steps on Problem Set 5.

$$\begin{aligned}\sigma^2 &= \text{Var}(X) = \sum_{x \in \{0,1\}} (x - \mu)^2 p(x) \\ &= \sum_{x \in \{0,1\}} (x - p)^2 p(x) \\ &\vdots \\ &= p(1 - p)\end{aligned}$$

## Variance of a Linear Function



Suppose  $X$  is a random variable with  $\text{Var}(X) = \sigma^2$  and  $a, b$  are constants. What is  $\text{Var}(a + bX)$ ?

- (a)  $\sigma^2$
- (b)  $a + \sigma^2$
- (c)  $b\sigma^2$
- (d)  $a + b\sigma^2$
- (e)  $b^2\sigma^2$

## Variance and SD are *NOT* Linear

$$\text{Var}(a + bX) = b^2\sigma^2$$

$$\text{SD}(a + bX) = |b|\sigma$$

These should look familiar from the related results for sample variance and std. dev. that you worked out on an earlier problem set.

## Variance of a Linear Transformation

$$\begin{aligned}\text{Var}(a + bX) &= E \left[ \{(a + bX) - E(a + bX)\}^2 \right] \\&= E \left[ \{(a + bX) - (a + bE[X])\}^2 \right] \\&= E \left[ (bX - bE[X])^2 \right] \\&= E[b^2(X - E[X])^2] \\&= b^2 E[(X - E[X])^2] \\&= b^2 \text{Var}(X) = b^2 \sigma^2\end{aligned}$$

The key point here is that variance is defined in terms of expectation and expectation is linear.

# Binomial Random Variable

What we get if we sum a bunch of indep. Bernoulli RVs



# Binomial Random Variable

Let  $X$  = the sum of  $n$  independent Bernoulli trials, each with probability of success  $p$ . Then we say that:  $X \sim \text{Binomial}(n, p)$

## Parameters

$p$  = probability of “success,”  $n$  = # of trials

## Support

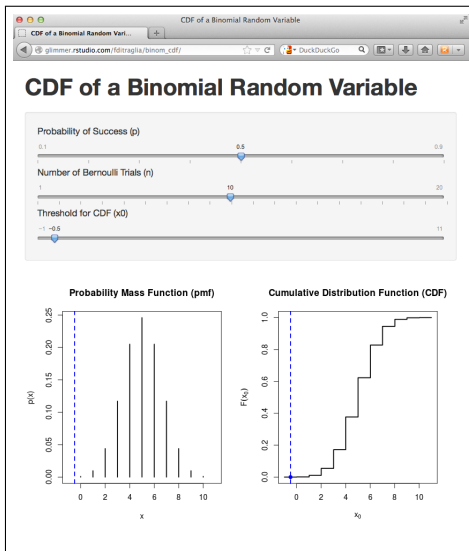
$\{0, 1, 2, \dots, n\}$

## Probability Mass Function (pmf)

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

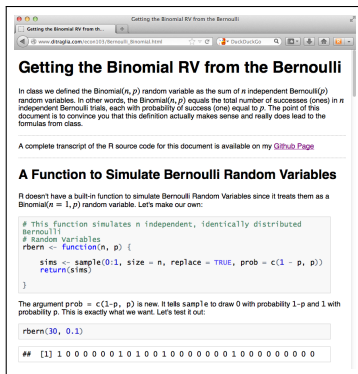
[http://fditraglia.shinyapps.io/binom\\_cdf/](http://fditraglia.shinyapps.io/binom_cdf/)

Try playing around with all three sliders. If you set the second to 1 you get a Bernoulli.



[http://fditraglia.github.com/Econ103Public/Rtutorials/Bernoulli\\_Binomial.html](http://fditraglia.github.com/Econ103Public/Rtutorials/Bernoulli_Binomial.html)

Source Code on my [Github Page](#)



**Getting the Binomial RV from the Bernoulli**

In class we defined the  $\text{Binomial}(n, p)$  random variable as the sum of  $n$  independent  $\text{Bernoulli}(p)$  random variables. In other words, the  $\text{Binomial}(n, p)$  equals the total number of successes (ones) in  $n$  independent Bernoulli trials, each with probability of success (one) equal to  $p$ . The point of this document is to convince you that this definition actually makes sense and really does lead to the formulas from class.

A complete transcript of the R source code for this document is available on my [Github Page](#).

**A Function to Simulate Bernoulli Random Variables**

R doesn't have a built-in function to simulate Bernoulli Random Variables since it treats them as a  $\text{Binomial}(n = 1, p)$  random variable. Let's make our own:

```
# This function simulates n independent, identically distributed
# Bernoulli
# Random Variables
rbern <- function(n, p) {
  sims <- sample(0:1, size = n, replace = TRUE, prob = c(1 - p, p))
  return(sims)
}
```

The argument `prob = c(1-p, p)` is new. It tells `sample` to draw 0 with probability  $1-p$  and 1 with probability  $p$ . This is exactly what we want. Let's test it out:

```
rbern(30, 0.1)
```

```
## [1] 1 0 0 0 0 0 1 0 1 0 0 1 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0
```

Don't forget this!

Binomial RV counts up the *total* number of successes (ones) in  $n$  indep. Bernoulli trials, each with prob. of success  $p$ .

# Where does the Binomial pmf come from?

## Question

Suppose we flip a fair coin 3 times. What is the probability that we get exactly 2 heads?

# Where does the Binomial pmf come from?

## Question

Suppose we flip a fair coin 3 times. What is the probability that we get exactly 2 heads?

## Answer

Three basic outcomes make up this event:  $\{HHT, HTH, THH\}$ , each has probability  $1/8 = 1/2 \times 1/2 \times 1/2$ . Basic outcomes are mutually exclusive, so sum to get  $3/8 = 0.375$

# Where does the Binomial pmf come from?

## Question

Suppose we flip an *unfair* coin 3 times, where the probability of heads is  $1/3$ . What is the probability that we get exactly 2 heads?

## Answer

No longer true that *all* basic outcomes are equally likely, but those with exactly two heads *still are*

$$P(HHT) = (1/3)^2(1 - 1/3) = 2/27$$

$$P(THH) = 2/27$$

$$P(HTH) = 2/27$$

Summing gives  $2/9 \approx 0.22$

# Where does the Binomial pmf come from?

Starting to see a pattern?

Suppose we flip an unfair coin 4 times, where the probability of heads is  $1/3$ . What is the probability that we get exactly 2 heads?

HHTT    TTHH

HTHT    THTH

HTTH    THTT

Six equally likely, mutually exclusive  
basic outcomes make up this event:

$$\binom{4}{2} (1/3)^2 (2/3)^2$$

## Multiple RVs *at once* - Definition of Joint PMF

Let  $X$  and  $Y$  be discrete random variables. The joint probability mass function  $p_{XY}(x, y)$  gives the probability of each pair of realizations  $(x, y)$  in the support:

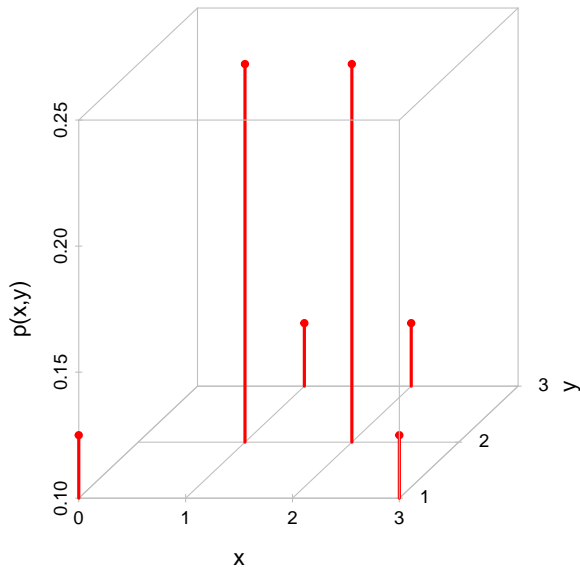
$$p_{XY}(x, y) = P(X = x \cap Y = y)$$



## Example: Joint PMF in Tabular Form

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

## Plot of Joint PMF



What is  $p_{XY}(1, 2)$ ?



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

What is  $p_{XY}(1, 2)$ ?



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

$$p_{XY}(1, 2) = P(X = 1 \cap Y = 2) = 1/4$$

What is  $p_{XY}(2, 1)$ ?



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

What is  $p_{XY}(2, 1)$ ?



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

$$p_{XY}(2, 1) = P(X = 2 \cap Y = 1) = 0$$

# Properties of Joint PMF

1.  $0 \leq p_{XY}(x, y) \leq 1$  for any pair  $(x, y)$
2. The sum of  $p_{XY}(x, y)$  over all pairs  $(x, y)$  in the support is 1:

$$\sum_x \sum_y p(x, y) = 1$$

Does this satisfy the properties of a joint pmf?



(A = YES, B = NO)

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



Does this satisfy the properties of a joint pmf?



(A = YES, B = NO)

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

1.  $p(x, y) \geq 0$  for all pairs  $(x, y)$
2.  $\sum_x \sum_y p(x, y) = 1/8 + 1/4 + 1/8 + 1/4 + 1/8 + 1/8 = 1$

# Joint versus Marginal PMFs

## Joint PMF

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

## Marginal PMFs

$$p_X(x) = P(X = x)$$

$$p_Y(y) = P(Y = y)$$

You can't calculate a joint pmf from marginals alone but you *can* calculate marginals from the joint!

## Marginals from Joint

$$p_X(x) = \sum_{\text{all } y} p_{XY}(x, y)$$

$$p_Y(y) = \sum_{\text{all } x} p_{XY}(x, y)$$

Why?

$$\begin{aligned} p_Y(y) &= P(Y = y) = P\left(\bigcup_{\text{all } x} \{X = x \cap Y = y\}\right) \\ &= \sum_{\text{all } x} P(X = x \cap Y = y) = \sum_{\text{all } x} p_{XY}(x, y) \end{aligned}$$

To get the marginals sum “into the margins” of the table.

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

To get the marginals sum “into the margins” of the table.

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

$$p_X(0) = 1/8 + 0 + 0 = 1/8$$

To get the marginals sum “into the margins” of the table.

		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	1/8	0	0	

$$p_X(0) = 1/8 + 0 + 0 = 1/8$$

$$p_X(1) = 0 + 1/4 + 1/8 = 3/8$$

To get the marginals sum “into the margins” of the table.

		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	

$$p_X(0) = 1/8 + 0 + 0 = 1/8$$

$$p_X(1) = 0 + 1/4 + 1/8 = 3/8$$

$$p_X(2) = 0 + 1/4 + 1/8 = 3/8$$

To get the marginals sum “into the margins” of the table.

		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

$$p_X(0) = 1/8 + 0 + 0 = 1/8$$

$$p_X(1) = 0 + 1/4 + 1/8 = 3/8$$

$$p_X(2) = 0 + 1/4 + 1/8 = 3/8$$

$$p_X(3) = 1/8 + 0 + 0 = 1/8$$



What is  $p_Y(2)$ ?



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

What is  $p_Y(2)$ ?



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

$$p_Y(1) = 1/8 + 0 + 0 + 1/8 = 1/4$$

What is  $p_Y(2)$ ?



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

$$p_Y(1) = 1/8 + 0 + 0 + 1/8 = 1/4$$

$$p_Y(2) = 0 + 1/4 + 1/4 + 0 = 1/2$$

What is  $p_Y(2)$ ?



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

$$p_Y(1) = 1/8 + 0 + 0 + 1/8 = 1/4$$

$$p_Y(2) = 0 + 1/4 + 1/4 + 0 = 1/2$$

$$p_Y(3) = 0 + 1/8 + 1/8 + 0 = 1/4$$