

# CSCI 3022-002 Intro to Data Science

## Discrete pdfs

$$f(x) = P(X=x)$$

$$F(x) = P(X \leq x)$$

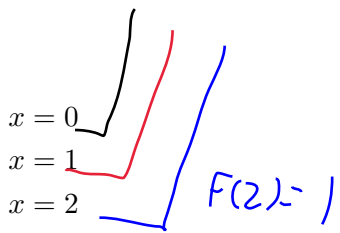
**Example:** Suppose we are given the following pmf:

$$P(X=x) = f(x) = \begin{cases} .5 & x=0 \\ .167 & x=1 \\ .333 & x=2 \\ 0 & \text{else} \end{cases}$$

1. Calculate:  $F(0), F(1), F(2)$ .
2. What is  $F(1.5)$ ?  $F(20.5)$ ?
3. Is  $P(X < 1) = P(X \leq 1)$ ?

## Opening Example; Soln

**Example:** Suppose we are given the following pmf:

$$P(X = x) = f(x) = \begin{cases} .5 & x = 0 \\ .167 & x = 1 \\ .333 & x = 2 \\ 0 & \text{else} \end{cases}$$


$F(2) = 1$

1. Calculate:  $F(0)$ ,  $F(1)$ ,  $F(2)$ .

$$F(0) = P(X \leq 0) = .5$$

2. What is  $F(1.5)$ ?  $F(20.5)$ ?

$$P(X \leq 1.5) = P(X \leq 1) + P(1 < X \leq 1.5) = .667$$

3. Is  $P(X < 1) = P(X \leq 1)$ ?

$$F(1) = P(X \leq 1) = P(X=0) + P(X=1) = .5 + \frac{1}{6} = \frac{2}{3}$$

$$P(X \leq 20.5) = 1$$

## Opening Example; Soln

**Example:** Suppose we are given the following pmf:

$$P(X = x) = f(x) = \begin{cases} .5 & x = 0 \\ .167 & x = 1 \\ .333 & x = 2 \\ 0 & \text{else} \end{cases}$$

1. Calculate:  $F(0), F(1), F(2)$ . *non-decreasing*

$$F(0) = P(X \leq 0) = .5; F(1) = P(X \leq 1) = .667; F(2) = P(X \leq 2) = 1$$

2. What is  $F(1.5)$ ?  $F(20.5)$ ?

$$F(1.5) = P(X \leq 1.5) = P(X \leq 1) = .667; F(20.5) = P(X \leq 20.5) = 1$$

3. Is  $P(X < 1) = P(X \leq 1)$ ?

Most certainly not!

$$P(X \leq 1) = P(X < 1) + P(X = 1) \quad .167$$

# Announcements and To-Dos

## Announcements:

1. HW 3, posted.
2. Another nb day this Friday.

## Last time we learned:

1. Bayes Theorem, and introduced pdfs and cdfs.

## To do:

1. Start that HW!

## Last Time...

We got a cool formula that was secretly just the definition of conditional probability rewritten!

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(B|A)P(A)}{P(B)}$$

We got two functions to play with:

- ▶ A *Probability density function* (pdf) is a function  $f(x)$  that describes the probability distribution of a random variable  $X$ . Discrete case:  $f(x) = P(X = x)$ .
- ▶ The *cumulative density function* (cdf), denoted  $F(x)$ , is defined for every real number  $x$  to be the probability that the observed value of  $X$  will be at most  $x$ , or  $F(x) = P(X \leq x)$ . Discrete case: a sum of values of  $f(x)$ !

$$F(a) = P(X \leq a)$$

pdf to cdf

P.V.  $X$  w/ pdf  $f$  ; cdf  $F$ 

The relationship between pdf and cdf is very important!

$$f(x) = \begin{cases} 1/36 & x=1 \\ 2/36 & x=2 \\ \vdots & \vdots \\ 0 & x=6 \end{cases}$$

$$P(X=1) + P(X=2) + \dots + P(X=6)$$

$$F(a) = P(X \leq a) = \sum_{x \leq a} P(X=x)$$

**Example:** What is the probability that if I roll two dice, they add up to at least 9. Write in terms of  $F(x)$ , then compute.

$X$ : sum of 2 dice

$x$

$X = \{9, 10, \dots, 12\}$

$$\begin{aligned} P(X \geq 9) &= 1 - \underbrace{P((X \geq 9)^c)}_{\text{complement}} = 1 - P(X < 9) \\ &= 1 - P(X \leq 8) \\ &= 1 - F(8) \end{aligned}$$

## pdf to cdf

The relationship between pdf and cdf is very important!

$$F(a) = P(X \leq a) = \sum_{x \leq a} P(X = x)$$

**Example:** What is the probability that if I roll two dice, they add up to at least 9. Write in terms of  $F(x)$ , then compute. *Fair!*

$X$  := the sum of the two dice, we want

$$P(X \geq 9) = 1 - P(X < 9) = 1 - P(X \leq 8) = 1 - F(8).$$

The easier probability is probably the

$$P(X \geq 9) = P(X = 9) + P(X = 10) + P(X = 11) + P(X = 12) = 10/36.$$

*4/36*

*3/36*

*2/36*

*1/36*

2d6;  $\Omega$  and  $X$ 

Suppose we roll two fair, 6-sided dice. Let  $X :=$  the value representing the maximum of the two dice.

1. What are the possible values of  $X$ ?  $= X \in \{1, 2, 3, 4, 5, 6\}$
2. Which elements of the sample space map to which values of  $X$ ?
3. What is the pmf of  $X$ ?  $P(\frac{1}{6}) \rightarrow (2, 3) \Rightarrow X=3$

Red

	1	2	3	4	5	6
1	1	2	3	4	5	6
2	2	2	3	4	5	6
3	3	3	3	4	5	6
4	4	4	4	4	5	6
5	5	5	5	5	5	6
6	6	6	6	6	6	6

Blue (4)

36 possibilities, each tile equally likely



## 2d6; $\Omega$ and $X$

Suppose we roll two fair, 6-sided dice. Let  $X :=$  the value representing the maximum of the two dice.

1. What are the possible values of  $X$ ?
2. Which elements of the sample space map to which values of  $X$ ?
3. What is the pmf of  $X$ ?

1.  $X \in \{1, 2, 3, 4, 5, 6\}$

2.

		Die 2					
Die 1		1	2	3	4	5	6
	1	1	2	3	4	5	6
	2	2	2	3	4	5	6
	3	3	3	3	4	5	6
	4	4	4	4	4	5	6
	5	5	5	5	5	5	6
	6	6	6	6	6	6	6

3. The pmf is:  $P(X = x)$ ; or

$$f(x) = \begin{cases} 1/36 & X = 1 \\ 3/36 & X = 2 \\ 5/36 & X = 3 \\ 7/36 & X = 4 \\ 9/36 & X = 5 \\ 11/36 & X = 6 \end{cases}$$

$f(x) = 0$   
 $\Rightarrow$  for other  $x$ .

## 2d6; The Max

Now we have

~~X~~ the max of 2 die rolls,  
 $F(2) = P$  that the max of  
 2 die rolls is exactly 2.  $P(X \text{ is 3 or less})?$

$$P(X \leq 3) = F(3)$$

$$f(x) = \begin{cases} 1/36 & X = 1 \\ 3/36 & X = 2 \\ 5/36 & X = 3 \\ 7/36 & X = 4 \\ 9/36 & X = 5 \\ 11/36 & X = 6 \end{cases}$$

↑  
lower case!

What are:

1.  $P(X \text{ is even})$ ?

$$P(X \text{ is 2 or 4 or 6}) \\ = P(X=2) + P(X=4) + P(X=6)$$

3. What is the cdf for  $X$ ?

$$F(x) = \begin{cases} 0 & x < 1 \\ P(X=1) = 1/36 & 1 \leq x < 2 \\ P(X \leq 2) = 1/36 + 3/36 & 2 \leq x < 3 \end{cases}$$

$$F_X(a) \\ = P(X \leq a)$$

$$= \frac{3 + 3 + 11}{36} = \frac{21}{36}$$

## 2d6; The Max

Now we have

lower-case!

↓

$$f(x) = \begin{cases} 1/36 & X = 1 \\ 3/36 & X = 2 \\ 5/36 & X = 3 \\ 7/36 & X = 4 \\ 9/36 & X = 5 \\ 11/36 & X = 6 \end{cases}$$

What are:

1.  $P(X \text{ is even})?$ 

$$\frac{3 + 5 + 7}{36}$$

2.  $P(X \text{ is 3 or less})?$ 

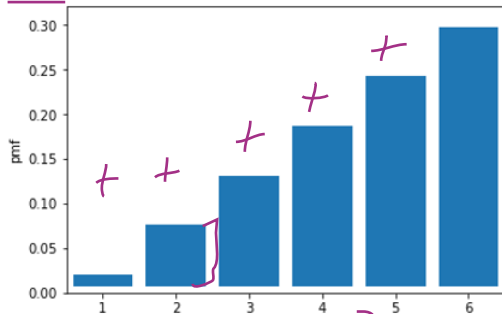
$$\frac{1 + 3 + 5}{36}$$

3. What is the cdf for  $X$ ?

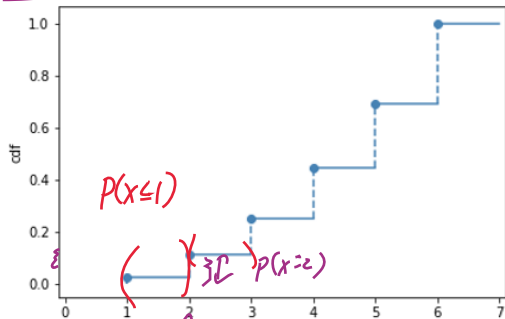
F(a)

$$F(x) = \begin{cases} 0 & x < 1 \\ 1/36 & 1 \leq x < 2 \\ 4/36 & 2 \leq x < 3 \\ 9/36 & 3 \leq x < 4 \\ 16/36 & 4 \leq x < 5 \\ 25/36 & 5 \leq x < 6 \\ 36/36 & x > 6 \end{cases}$$

pmf &amp; pdf

A picture denoting the pdf and cdf of our  $X$ :~~PDF:~~ pmf; histogram $\uparrow P(X \geq 2) \approx .08$ 

pdf (cont's): line

CDF:(jumps)  
always a line $P(X \leq 1)$  $P(X \leq 2)$  $P(X \leq 3)$  $= P(X \leq 2)$  $= P(X < 2) + P(X = 2)$

# Discrete Random Variables

Discrete random variables can be categorized into different types or classes. Each type/class models many different real-world situations. They can loosely be broken down into a few groups:

1. The Discrete Uniform for modeling  $n$  equally likely (*fair*) outcomes np.random.choice
2. Distributions built on counting trials-until-event (how rolls until I get a 6, etc.) when the trials are identical and repeated binary outcomes  
 Examples: Binomial, Geometric, etc.
3. Counting occurrences of an event over fixed areas of time/space.  
 Example: Poisson

# The Bernoulli : ~~X~~

A random variable whose only possible values are 0 or 1.

This is a discrete random variable – why?

This distribution is specified by a single parameter:  $p(\{H\})$   
 The probability of a heads/“success”  $p$ ! This gives the pdf:

$$f(a) = \begin{cases} p & a=1 \\ 1-p & a=0 \end{cases}$$

“is distributed as”  
 $\downarrow$

$$X \sim \text{bern}(p)$$

We denote the Bernoulli random variable  $X$  by

## The Bernoulli

A random variable whose only possible values are 0 or 1.

This is a discrete random variable – why?

### Countable outcomes

This distribution is specified by a single parameter:

The probability of a heads/ “success”  $p$ ! This gives the pdf:

$$P(X = x) = f(x) = \begin{cases} p & x = 1 \\ (1 - p) & x = 0 \\ 0 & \text{else} \end{cases}$$

We denote the Bernoulli random variable  $X$  by  $X \sim \text{Bern}(p)$

## The Bernoulli

A random variable whose only possible values are 0 or 1.

This is a discrete random variable – why?

This distribution is specified by a single parameter:

The probability of a heads/“success”  $p$ ! This gives the pdf:

$$P(X = x) = f(x) = \begin{cases} p & x = 1 \\ (1 - p) & x = 0 \\ 0 & \text{else} \end{cases}$$

awkward!

It turns out, it's nice to write the pdf as a single line whenever possible. The nicest way to do so for the Bernoulli:

$$f(x) = p^x (1 - p)^{1-x}$$

which works as long as we remember  $x$  can only be 0 or 1.

We denote the Bernoulli random variable  $X$  by  $X \sim \text{Bern}(p)$

$\rightarrow x=0$

$x=1$

$$\cancel{p}^x (1 - \cancel{p})^{1-0} = 1 - p$$

$$p^1 (\cancel{1-p})^{1-1} = p$$



## Overview

The Bernoulli random variable is the building block for numerous important probability distributions that reflect **repeated** measurements.

Statistics and data science on repeated measurements requires us understand principles of **counting**!

# Overview

The Bernoulli random variable is the building block for numerous important probability distributions that reflect **repeated** measurements.

Statistics and data science on repeated measurements requires us understand principles of **counting**!

1. Some counting is easy: how many integers are there in  $[0, 9]$ ?

## Overview

The Bernoulli random variable is the building block for numerous important probability distributions that reflect **repeated** measurements.

Statistics and data science on repeated measurements requires us understand principles of **counting**!

2. Zach, Felix, Rachel, and Ioana line up at a coffee stand. How many different orders could they stand in?

## Overview

The Bernoulli random variable is the building block for numerous important probability distributions that reflect **repeated** measurements.

Statistics and data science on repeated measurements requires us understand principles of **counting**!

2. Zach, Felix, Rachel, and Ioana line up at a coffee stand. How many different orders could they stand in?

This is a permutation: it counts distinct orderings

## Overview

The Bernoulli random variable is the building block for numerous important probability distributions that reflect **repeated** measurements.

Statistics and data science on repeated measurements requires us understand principles of **counting**!

3. There are 10 problems on an exam, and you need 7 correct to pass. How many different ways are there to pass?

## Overview

The Bernoulli random variable is the building block for numerous important probability distributions that reflect **repeated** measurements.

Statistics and data science on repeated measurements requires us understand principles of **counting**!

3. There are 10 problems on an exam, and you need 7 correct to pass. How many different ways are there to pass?

This is a *combination*: it counts ways a set can be split into subsets

# Permutations

How many ways can you order a set of one object; e.g.  $\{A\}$ ?

$1: \{A\}$

How many ways can you order a set of two objects; e.g.  $\{A, B\}$ ?

$A, B$  or  $BA$

How many ways can you order a set of three objects; e.g.  $\{ABC\}$ ?

$CA, B$  or  $A(B$  or

$CB, A, B, CA, BA, ABC$

What's the pattern? How many ways could you order  $n$  objects?

# Permutations

How many ways can you order a set of one object; e.g.  $\{A\}$ ?

**A:** 1 way.  $\{A\}$ .

How many ways can you order a set of two objects; e.g.  $\{A, B\}$ ?

**A:** 2 ways.  $\{AB, BA\}$ .

How many ways can you order a set of three objects; e.g.  $\{ABC\}$ ?

**A:** 6 ways.  $\{ABC, ACB, BAC, BCA, CBA, CAB\}$ . = 3!

What's the pattern? How many ways could you order  $n$  objects?

**A:**  $n!$



## Permutations; General

What if you have  $n$  objects, but only want to permute  $r$  of them?

How many 3-character strings can we make if each character is a distinct letter from the English alphabet?

$\{S, A, D\} \neq \{A, P, S\}$   
order matters!

What is the general form for an  $r$ -permutation of  $n$  objects?

## Permutations; General

What if you have  $n$  objects, but only want to permute  $r$  of them?

How many 3-character strings can we make if each character is a distinct letter from the English alphabet?

**A:** There are 24 that start with  $\{AB\}$ . There are 25 letters (including  $B$ ) that could have followed an  $A$ . There are 26 options to start with. That multiplies to  $26 \cdot 25 \cdot 24$ .

What is the general form for an  $r$ -permutation of  $n$  objects?

**A:**  $P(n, r) = \frac{n!}{(n-r)!}$

This should feel a lot like **sampling without replacement**.. because it is, only without probabilities.

## Combinations

Counting *combinations* means counting the number of ways an object can be sliced into subsets. The big difference: **order doesn't matter**.

How many 3-character *combinations* can we make if each character is a distinct letter from the English alphabet?

$$\{S, A, D\} \stackrel{\text{set}}{=} \{D, S, A\}$$

## Combinations

Counting *combinations* means counting the number of ways an object can be sliced into subsets. The big difference: **order doesn't matter**.

How many 3-character *combinations* can we make if each character is a distinct letter from the English alphabet?

Start with the number of permutations:  $P(n, r) = 26 \cdot 25 \cdot 24$ , then ask how many times we "overcounted," because now we don't want subsets with the same elements.

**Ex:** How many times did we include a subset with  $\{A, B, C\}$ ?

$$6 \leftarrow 3!$$

## Combinations

Counting *combinations* means counting the number of ways an object can be sliced into subsets. The big difference: **order doesn't matter**.

How many 3-character *combinations* can we make if each character is a distinct letter from the English alphabet?

Start with the number of permutations:  $P(n, r) = 26 \cdot 25 \cdot 24$ , then ask how many times we "overcounted," because now we don't want subsets with the same elements.

**Ex:** How many times did we include a subset with  $\{A, B, C\}$ ?

Our permutation set had  $\{ABC\}, \{ACB\}, \{BAC\}, \{BCA\}, \{CBA\}$ , and  $\{CAB\}$  as distinct... or all 6 orderings of those 3 elements! So:

$$C(n, r) = \frac{n!}{(n-r)!(r!)}$$

↗  
extra overcounting

## Combinations; Example

Combinations often use a variety of notations, including

$$C(n, r) = \binom{n}{k} = \frac{n!}{(n-r)!r!} := \text{"n choose k"}$$

**Example:** If there are 10 problems on an exam, and you need 7 correct to pass, how many different ways are there to pass?

## Combinations; Example

Combinations often use a variety of notations, including

$$C(n, r) = \binom{n}{r} = \frac{n!}{(n-r)!r!} := \text{"n choose k"}$$

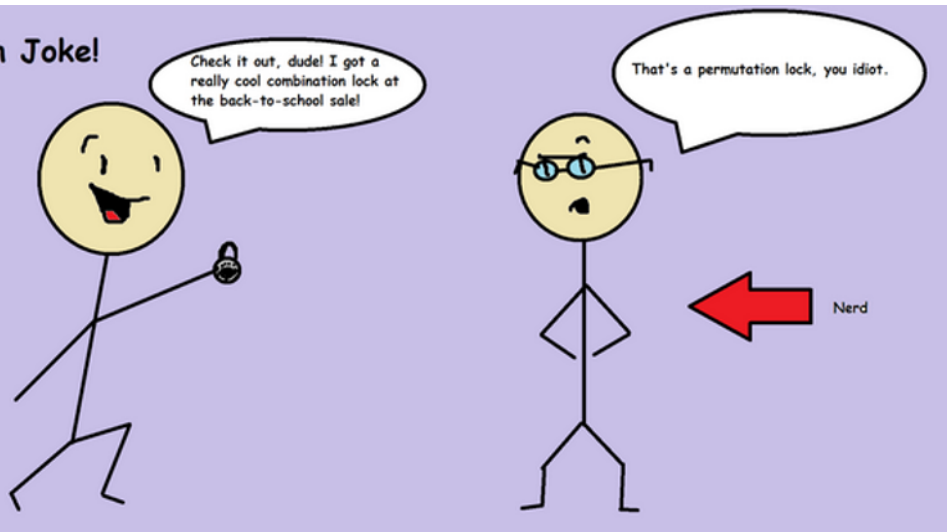
**Example:** If there are 10 problems on an exam, and you need 7 correct to pass, how many different ways are there to pass?

**Answer:**  $C(10, 7) + C(10, 8) + C(10, 9) + C(10, 10)$

$\nearrow$   
 # of ways to choose  $\nearrow$  correct +  $\dots$  =  $\frac{10!}{(10-0)!(0!)} = \frac{10!}{1 \cdot 10!} = 1$   
 $\nearrow$  correct out of 10

## Perms and Combs; Summary

### Cheesy Math Joke!





# Binomials

Exponents are useful, and pretty common: a lot of both data science and computational problems often involve solutions that look like polynomials. For example, let's consider:

1. Expand  $(x + y)^1$

2. Expand  $(x + y)^2$

3. Expand  $(x + y)^3$

4. Expand  $(x + y)^4$

# Binomials

Exponents are useful, and pretty common: a lot of both data science and computational problems often involve solutions that look like polynomials. For example, let's consider:

1. Expand  $(x + y)^1$  *↓ binomial: 2 things to a power*

**Solution:**  $(x + y)^1 = x + y$

2. Expand  $(x + y)^2$

**Solution:**  $(x + y)^2 = x^2 + 2xy + y^2$

3. Expand  $(x + y)^3$

**Solution:**  $(x + y)^3 = (x + y)(x^2 + 2xy + y^2) = x^3 + 3x^2y + 3xy^2 + 1$

4. Expand  $(x + y)^4$

**Solution:**  $(x + y)^4 = (x + y)(x^3 + 3x^2y + 3xy^2 + 1) = x^4 + 4x^3y + 6x^2y^2 + 4xy^3 + 1$

# Binomials

Exponents are useful, and pretty common: a lot of both data science and computational problems often involve solutions that look like polynomials. For example, let's consider:

1. Expand  $(x + y)^1$

**Solution:**  $(x + y)^1 = x + y$

2. Expand  $(x + y)^2$

**Solution:**  $(x + y)^2 = x^2 + 2xy + y^2$

3. Expand  $(x + y)^3$

**Solution:**  $(x + y)^3 = (x + y)(x^2 + 2xy + y^2) = x^3 + 3x^2y + 3xy^2 + 1$

4. Expand  $(x + y)^4$

**Solution:**  $(x + y)^4 = (x + y)(x^3 + 3x^2y + 3xy^2 + 1) = x^4 + 4x^3y + 6x^2y^2 + 4xy^3 + 1$

What are some patterns? It's definitely symmetric - the coefficients are palindromic - and it seems to always start with 1 and then  $n$  (the power)

## Binomials

One way to think about a binomial (two term) expansion is using "choose." Think about foiling:

$$(x_0 + x_1)(a + b) = \underbrace{ax_0}_{\text{first}} + \underbrace{bx_0}_{\text{outer}} + \underbrace{ax_1}_{\text{inner}} + \underbrace{bx_1}_{\text{last}}$$

There are 4 terms, but these are the same 4 terms as we would get from a multiplication rule: "choose" one of the first 2 terms and "choose" one of the second 2 terms for  $2 \cdot 2$  total.

For our problem, we have to worry about repeating terms, though! If we think about:

$$(x + y)^4 = (\underline{\bar{x}} + \underline{\bar{y}})(\underline{\bar{x}} + \underline{\bar{y}})(\underline{\bar{x}} + \underline{\bar{y}})(\underline{\bar{x}} + \underline{\bar{y}})$$

it's making 4 choices: "choose  $x$  or  $y$ ," then "choose  $x$  or  $y$ ," then "choose  $x$  or  $y$ ," then "choose  $x$  or  $y$ ." The coefficient of the  $x^2y^2$  term is the number of ways we could "choose  $x$  or  $y$ " 4 times and end up with 2  $x$ 's and 2  $y$ 's.

## Binomials, Cont'd

So we're expanding

$$\begin{aligned}
 (x + y)^4 &= (x + y)(x + y)(x + y)(x + y) \\
 &= (x + y)(x^3 + 3x^2y + 3xy^2 + 1) \\
 &= x^4 + 4x^3y + 6x^2y^2 + 4xy^3 + 1
 \end{aligned}$$

and the coefficient of the  $x^2y^2$  term is the number of ways we could “choose  $x$  or  $y$ ” 4 times and end up with 2  $x$ 's and 2  $y$ 's.

Let's check. We're looking for all of the ways you could get e.g.  $xxyy$ ,  $yyxx$ ,  $xyyx$ , etc. This is the same as asking for the number of ways to choose 2 of the 4 “slots” to be  $x$  or choosing 2 of the 4 slots to be  $y$ , or  $C(4, 2) = \frac{4!}{2!}$ .

# Binomial Theorem

**Theorem:** Let  $x$  and  $y$  be variables and  $n$  be a non-negative integer. Then

$$(x + y)^n = \sum_{k=0}^n C(n, k)x^{n-k}y^k = C(n, 0)x^ny^0 + C(n, 1)x^{n-1}y^1 + \cdots + C(n, n)x^0y^n$$

In other words,  $C(n, k)$  is the coefficient of  $x^ky^{n-k}$  and  $x^{n-k}y^k$ . We usually write the  $C$  numbers in choose notation:

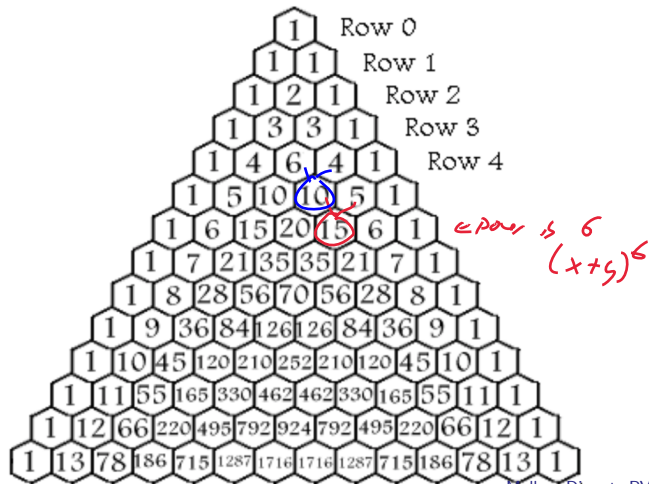
$$(x + y)^n = \sum_{k=0}^n \binom{n}{k} x^{n-k} y^k = \binom{n}{0} x^n y^0 + \binom{n}{1} x^{n-1} y^1 + \cdots + \binom{n}{n} x^0 y^n$$

## Pascal's Triangle

For small expansions, an easy trick to find the binomial coefficients is Pascal's triangle. Each entry of the triangle is the sum of the two entries above it:

## Pascal's Triangle

For small expansions, an easy trick to find the binomial coefficients is Pascal's triangle. Each entry of the triangle is the sum of the two entries above it:



Finding/using:

$$\binom{n}{k} = \frac{n!}{(n-k)!(k!)}$$



# The Binomial

$$P(\{H\}) = p$$

$$\Rightarrow P(X=1) = p$$

**Example:** A fair Bernoulli coin is tossed eight times. A "successful toss" is defined to be the coin landing on heads.

Let  $X = \#$  of successes or heads in 8 tosses.

1. How many ways in  $\Omega$  can  $X = 3$ ?

counting:  $\left\{ \begin{matrix} 1100000 \\ 10110000 \end{matrix} \right\} \subset \binom{8}{3}$  *equally likely*

2. What is  $P(X=3)$  for each *one* of those ways?

$$P(\{HHHTTT\}) = P(\{TTTTTH\})$$

*and and*

$$= P(\{H\})^3 \cdot P(\{T\})^5$$

*and and and*

3. What is  $P(X=3)$ ?

# The Binomial

**Example:** A fair Bernoulli coin is tossed eight times. A "successful toss" is defined to be the coin landing on heads.

Let  $X = \#$  of successes or heads in 8 tosses.

1. How many ways in  $\Omega$  can  $X = 3$ ?

$$C(8, 3) \text{ OR } C(8, 5)$$

2. What is  $P(X = 3)$  for each *one* of those ways?

One such way is  $\{HHHTTTTT\}$  which has probability  $P(\{H\})^3 \cdot P(\{T\})^5$ .

3. What is  $P(X = 3)$ ? The product of these two things!

## The Binomial

Lets generalize those ideas to derive the Binomial pdf for  $n$  trials of an underlying Bern( $p$ ).

Let  $X :=$  the number of successes of  $n$  trials of a Bern( $p$ ). Then:

*$p$  of  $X$  successes in  $n$  tries.*

NOTATION: We write  $\overset{\substack{\text{\# of tries} \\ \downarrow}}{\text{bin}(n, p)}$  to indicate that  $X$  is a Binomial rv with success probability  $p$  and  $n$  trials.

## The Binomial

Lets generalize those ideas to derive the Binomial pdf for  $n$  trials of an underlying Bern( $p$ ).

Let  $X :=$  the number of successes of  $n$  trials of a Bern( $p$ ). Then:

$$P(X = i) = (\# \text{ of ways that } X = i) \cdot P(\text{of one such outcome})$$

NOTATION: We write  $X \sim \text{bin}(n, p)$  to indicate that  $X$  is a Binomial rv with success probability  $p$  and  $n$  trials.

## The Binomial

Lets generalize those ideas to derive the Binomial pdf for  $n$  trials of an underlying Bern( $p$ ).

Let  $X :=$  the number of successes of  $n$  trials of a Bern( $p$ ). Then:

$$P(X = i) = (\# \text{ of ways that } X = i) \cdot P(\text{of one such outcome})$$

$$P(X = i) = \binom{n}{i} \cdot P(n \text{ successes}) \cdot P(n - i \text{ failures}).$$

$$P(X = i) = \binom{n}{i} p^i (1 - p)^{(n-i)}$$

$$f(x) = P(X = x) = \binom{n}{x} p^x (1 - p)^{(n-x)}; \quad x \in \{0, 1, 2, \dots, n\}$$

NOTATION: We write  $X \sim \text{bin}(n, p)$  to indicate that  $X$  is a Binomial rv with success probability  $p$  and  $n$  trials.

# The Binomial

The Binomial r.v. counts the total number of successes out of  $n$  trials, where  $X$  is the number of successes.

Important Assumptions:

1. Each trial must be *independent* of the previous experiment.
2. The probability of success must be *identical* for each trial.

The binomial is often defined and derived as the sum of  $n$  *independent, identically distributed* Bernoulli random variables.

In practice, any time we try to study a proportion on an underlying population, we gather a smaller sample where the observed proportion can often be thought of as a binomial random variable.

# Daily Recap

Today we learned

1. Bernoullis, Binomials!

Moving forward:

- nb day Friday!

Next time in lecture:

- More special and common pdfs!