9-18 Discrete RVs

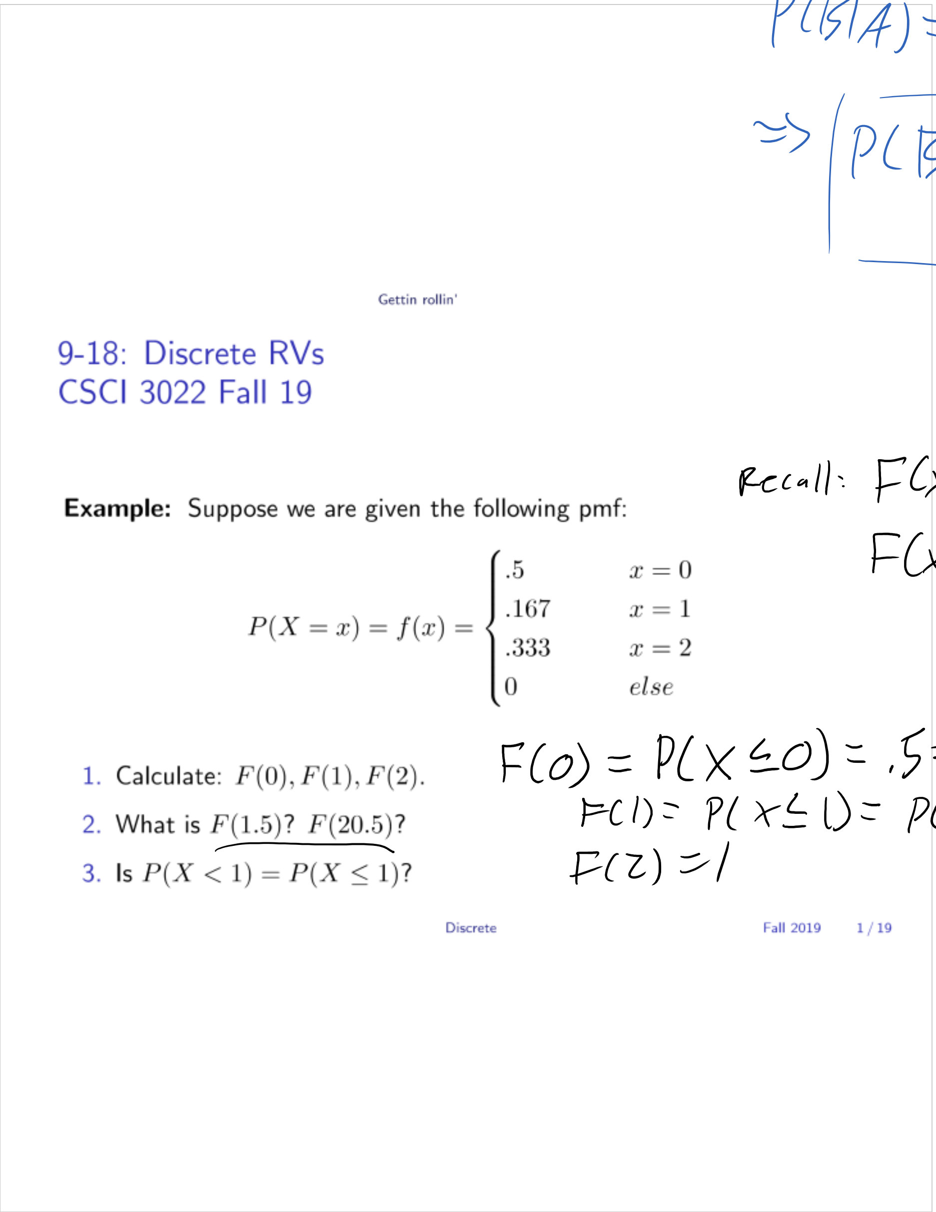
Wednesday, September 18, 2019 12:25 PM

9-18 Discrete RVs

Cond: $P(A|B) = \frac{P(A\cap B)}{P(B)} = \frac{\overset{\text{data}}{\overset{\downarrow}{P(B|A)}} \cdot \overset{\text{Prior}}{\overset{\downarrow}{P(A)}}}{P(B)}$

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

$\Rightarrow \boxed{P(B|A)\cdot P(A) = P(A|B)P(B) = P(A\cap B)}$



Gettin rollin'

9-18: Discrete RVs
CSCI 3022 Fall 19

Example: Suppose we are given the following pmf:

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

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

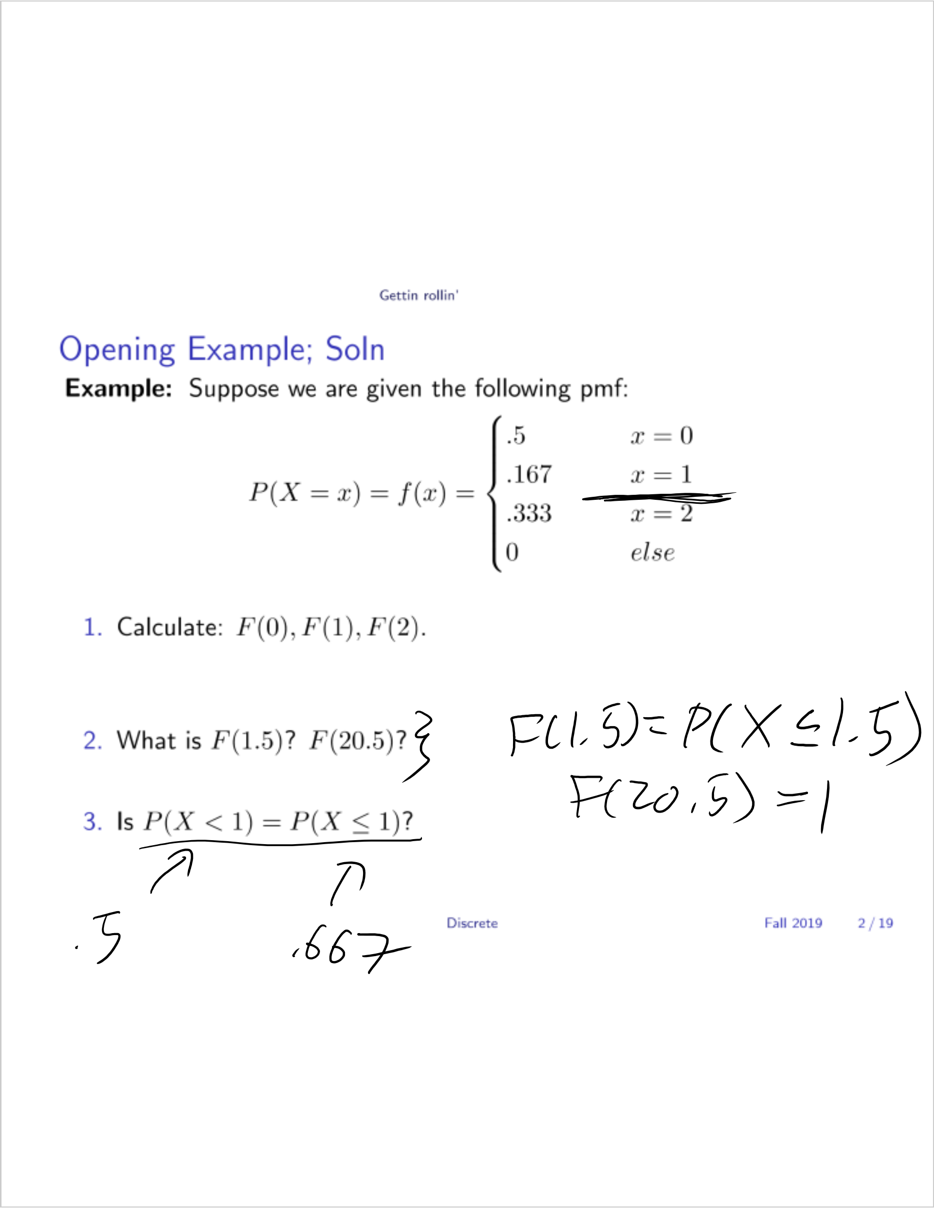
Recall: $F(x)$ is the cdf of X,

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

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

$F(1) = P(X\leq 1) = P(X=0) + P(X=1) = .667$

$F(2) = 1$



Opening Example; Soln

Example: Suppose we are given the following pmf:

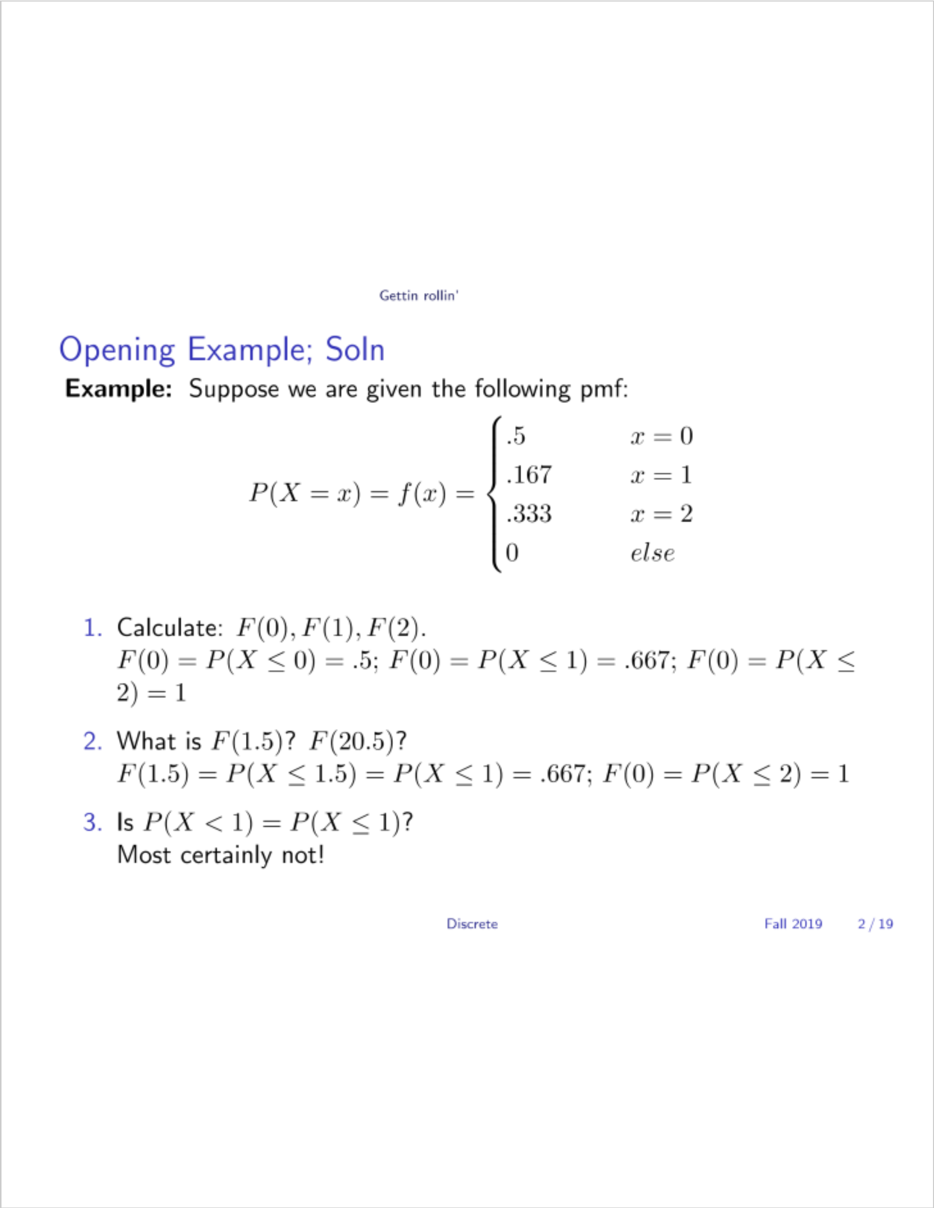
$$P(X = x) = f(x) = \begin{cases} .5 & x = 0 \\ .167 & x = 1 \\ .333 & x = 2 \\ 0 & else \end{cases}$$

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

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

$F(20.5) = 1$

$.5$ $\quad$ $.667$



Opening Example; Soln

Example: Suppose we are given the following pmf:

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

1. Calculate: $F(0), F(1), F(2)$.
   $F(0) = P(X \leq 0) = .5$; $F(0) = P(X \leq 1) = .667$; $F(0) = P(X \leq 2) = 1$
2. What is $F(1.5)$? $F(20.5)$?
   $F(1.5) = P(X \leq 1.5) = P(X \leq 1) = .667$; $F(0) = P(X \leq 2) = 1$
3. Is $P(X < 1) = P(X \leq 1)$?
   Most certainly not!

Announcements and Reminders

- Homework due Friday the 27th.
- Exam 1: Tuesday Oct 8.
- Homework recap! Why were you asked about...:

1. Sampling → biases.
2. Means and Medians
3. Rolling Mean/Variance → memory efficient
4. Titanic } Data is good
5. Weather
6. Normalization

Last Time...

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

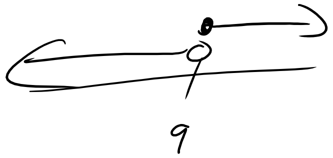
$$P(A|B) \overset{\text{Cond. Prob.}}{=} \frac{P(A \cap B)}{P(B)} \underset{\text{Bayes}}{=} \frac{P(B|A)P(A)}{P(B)}$$

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*, 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 $f(x)$!

pdf to cdf

$X \in \{\boxed{2, 3, 4, \cdots, 8,} \mid 9, \cdots 12\}$



The relationship between pdf and cdf is very important!

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

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.

$$P(X \ge 9) = 1 - P(X < 9)$$
$$= 1 - P(X \le 8)$$
$$= 1 - F(8)$$

<u>or</u> $P(X=9) + P(X=10) + P(X=11) + P(X=12)$

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

2d6; Ω 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: $\Omega = \{Die_1, Die_2\}$

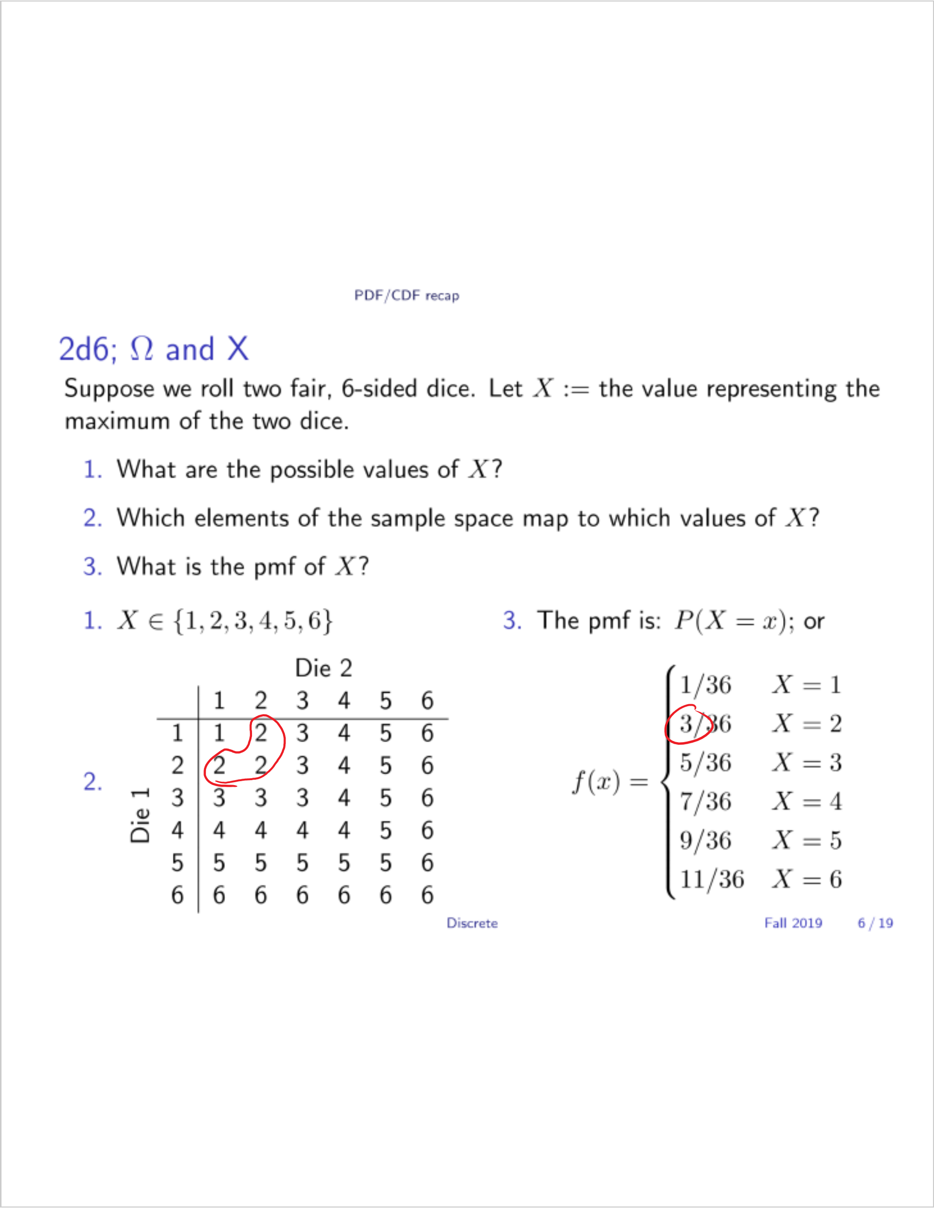
$$P(X=1) = \frac{|\{\Omega \mid X=1\}|}{|\Omega|} = \frac{1}{36}$$

36 pieces of Ω

D2

D1

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



2d6; Ω 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

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

2d6; The Max

Now we have

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

What are:

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

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

$= P(X=2) + P(X=4) + P(X=6)$

$= 3/36 + 7/36 + 11/36$

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

$P(X \leq 3) = F(3) = 9/36$

3. What is the cdf for X?

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

$$= \begin{cases} 0 & x < 1 \\ 1/36 & 1 \leq x < 2 \\ 4/36 & 2 \leq x < 3 \\ \vdots & \end{cases}$$

2d6; The Max

Now we have

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

What are:

1. $P(X$ is even)?

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

2. $P(X$ is 3 or less)?

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

3. What is the cdf for X?

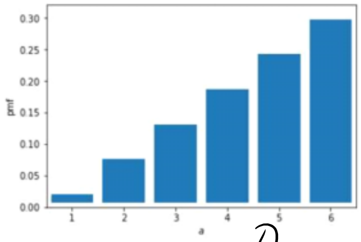
$$F(x) = \begin{cases} 0 & X < 1 \\ 1/36 & 1 \leq X < 2 \\ 4/36 & 2 \leq X < 3 \\ 9/36 & 3 \leq X < 4 \\ 16/36 & 4 \leq X < 5 \\ 25/36 & 5 \leq X < 6 \\ 36/36 & X > 6 \end{cases}$$

A picture denoting the pdf and cdf of our X:

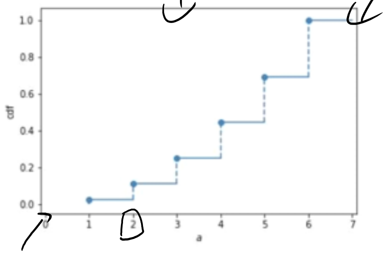
always increasing

ends at 1

PDF:

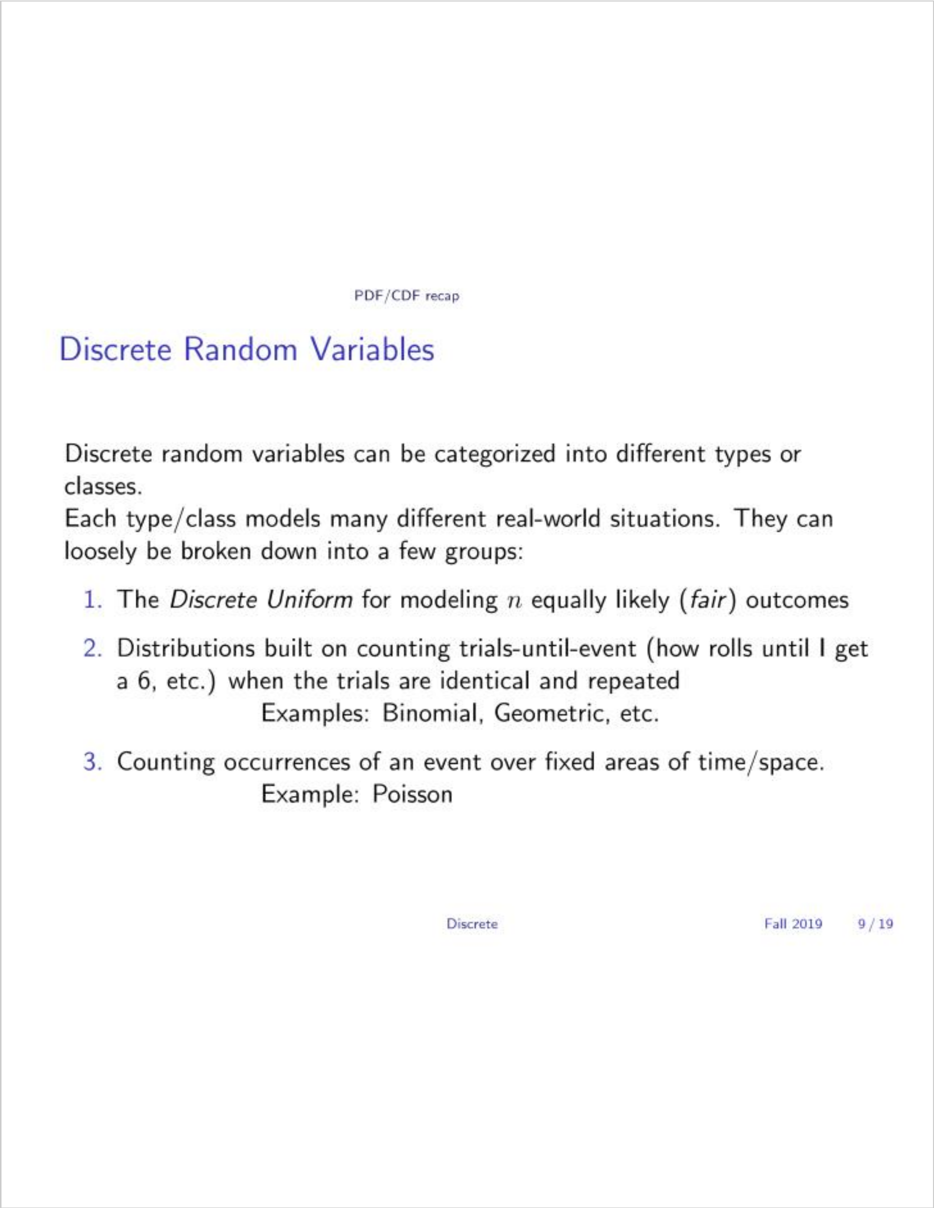


CDF:



starts at 0

coincidentally increasing



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
2. Distributions built on counting trials-until-event (how rolls until I get a 6, etc.) when the trials are identical and repeated
   Examples: Binomial, Geometric, etc.
3. Counting occurrences of an event over fixed areas of time/space.
   Example: Poisson

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:

$$P(X=1)=p$$
$$P(X=0)=1-p$$

$$\Rightarrow f(x)=\begin{cases} p & x=1 \text{ (heads)} \\ 1-p & x=0 \text{ (tails)} \\ 0 & \text{else.} \end{cases}$$

$X \sim Bern(p)$

We denote the Bernoulli random variable X by ____________

The Bernoulli

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

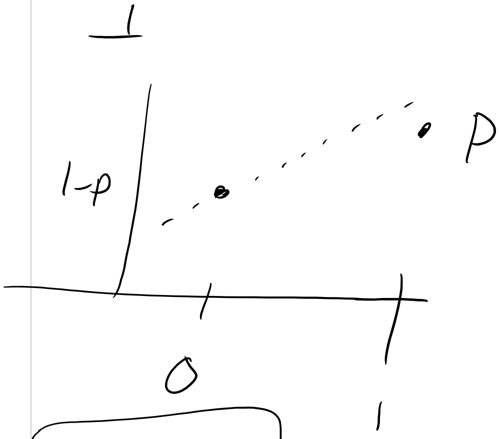
This is a discrete random variable – why?

Countable outcomes

This distribution is specified by a single parameter:

The probability of a heads/"success" p! This gives the pdf:

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



or $f(x) = p^x (1-p)^{1-x}$ $\boxed{x \in \{0,1\}}$

$x=0$ is $1 \cdot (1-p)$

if $x=1$ $(p)\cdot(1-p)^0 = p.$

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

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:

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

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 $\underline{X \sim Bern(p)}$

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**!

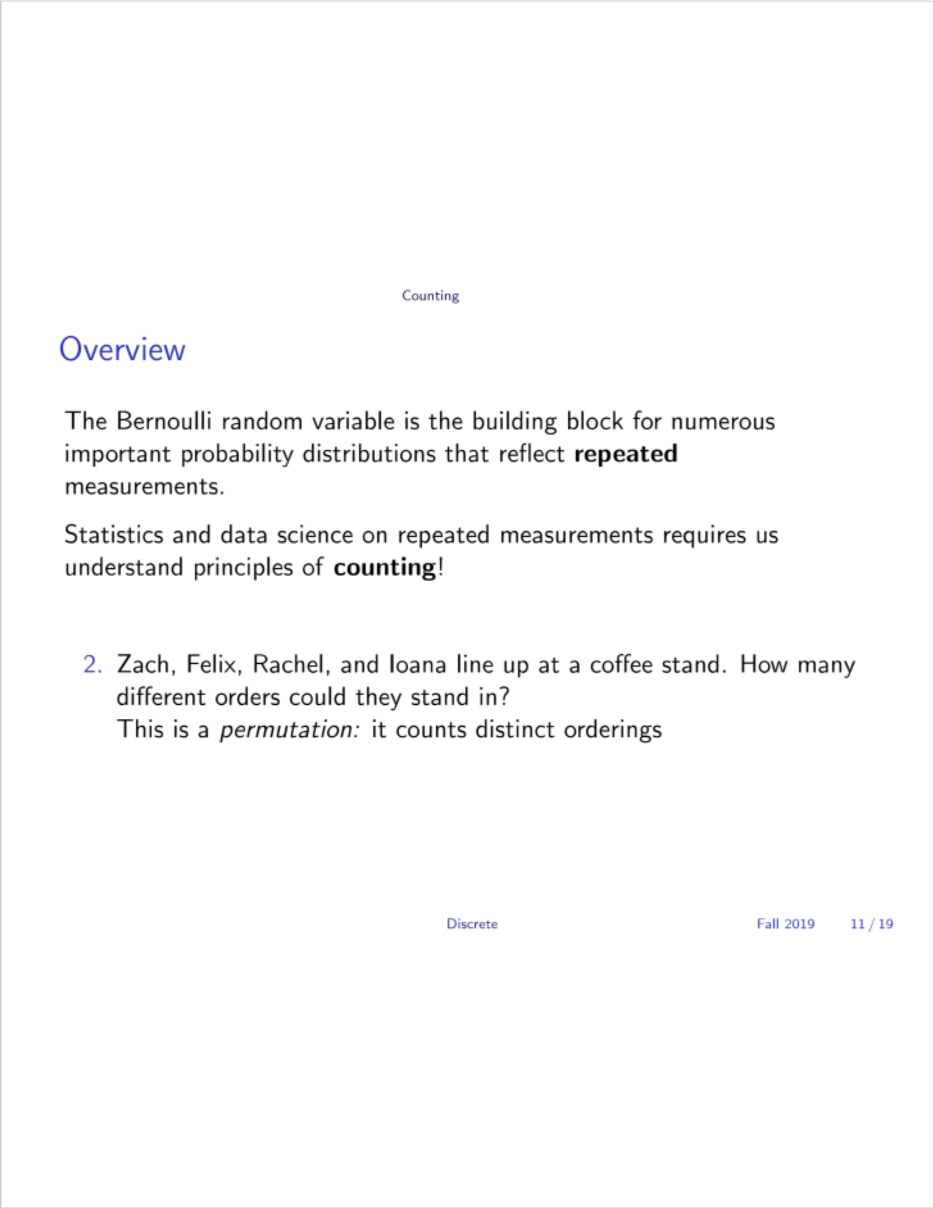
1. Some counting is easy: how many integers are there in $[0, 9]$?

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**!

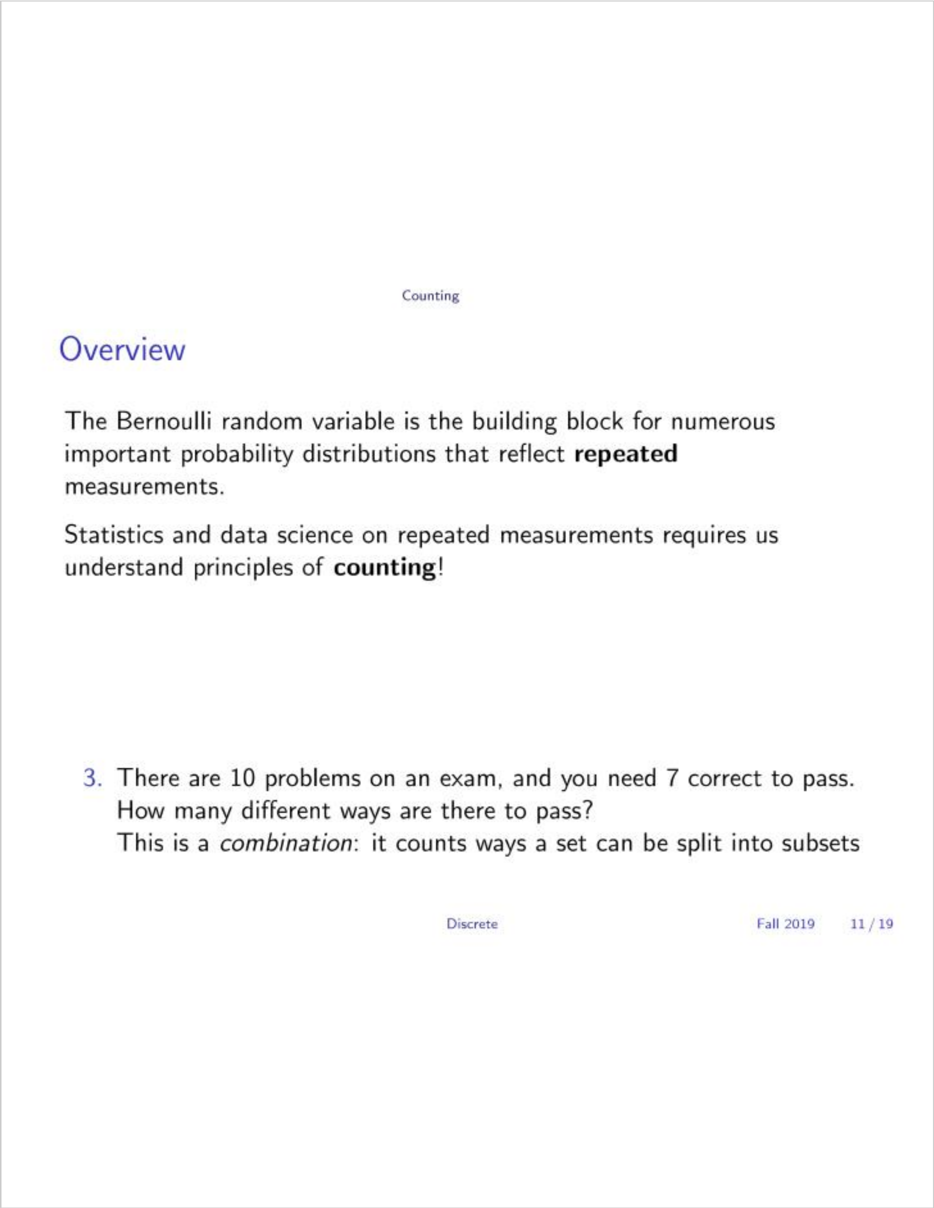
2. Zach, Felix, Rachel, and Ioana line up at a coffee stand. How many different orders could they stand in?
   This is a *permutation:* it counts distinct orderings

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

$\{A\}$.

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

1

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

AB, BA $= 2$

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

ABC, ACB, BAC, BCA, CAB, CBA $= 6$

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

$\boxed{n!}$

$n = 4$

$\boxed{4} \cdot \boxed{3} \cdot \boxed{2} \cdot \boxed{1}$

any of them

any other

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. $\{ABC, ACB, BAC, BCA, CBA, CAB\}$.

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?

$$\boxed{26} \cdot \boxed{25} \cdot \boxed{24}$$

$$26 \cdot 25 \cdot 24 = \frac{26!}{23!} = \frac{26!}{(26-3)!}$$

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

$$\frac{n!}{(n-r)!}$$

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

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

$\to \{A,B,C\} = \{B,C,A\}$

and

1) $P(26,3) = \frac{26!}{23!}$

includes

keep

throw away

$\boxed{ABC}, \boxed{ACB, BCA, BAC, CAB, CBA} = \boxed{3!}$

$$\Rightarrow C(26,3) = \frac{P(26,3)}{\boxed{3!}}$$

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:

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

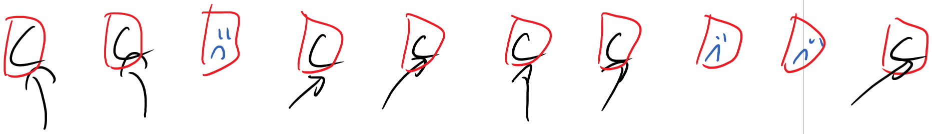
Combinations; Example

Combinations often use a variety of notations, including

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

choose subsets of size k from a pool (unique) of size n.

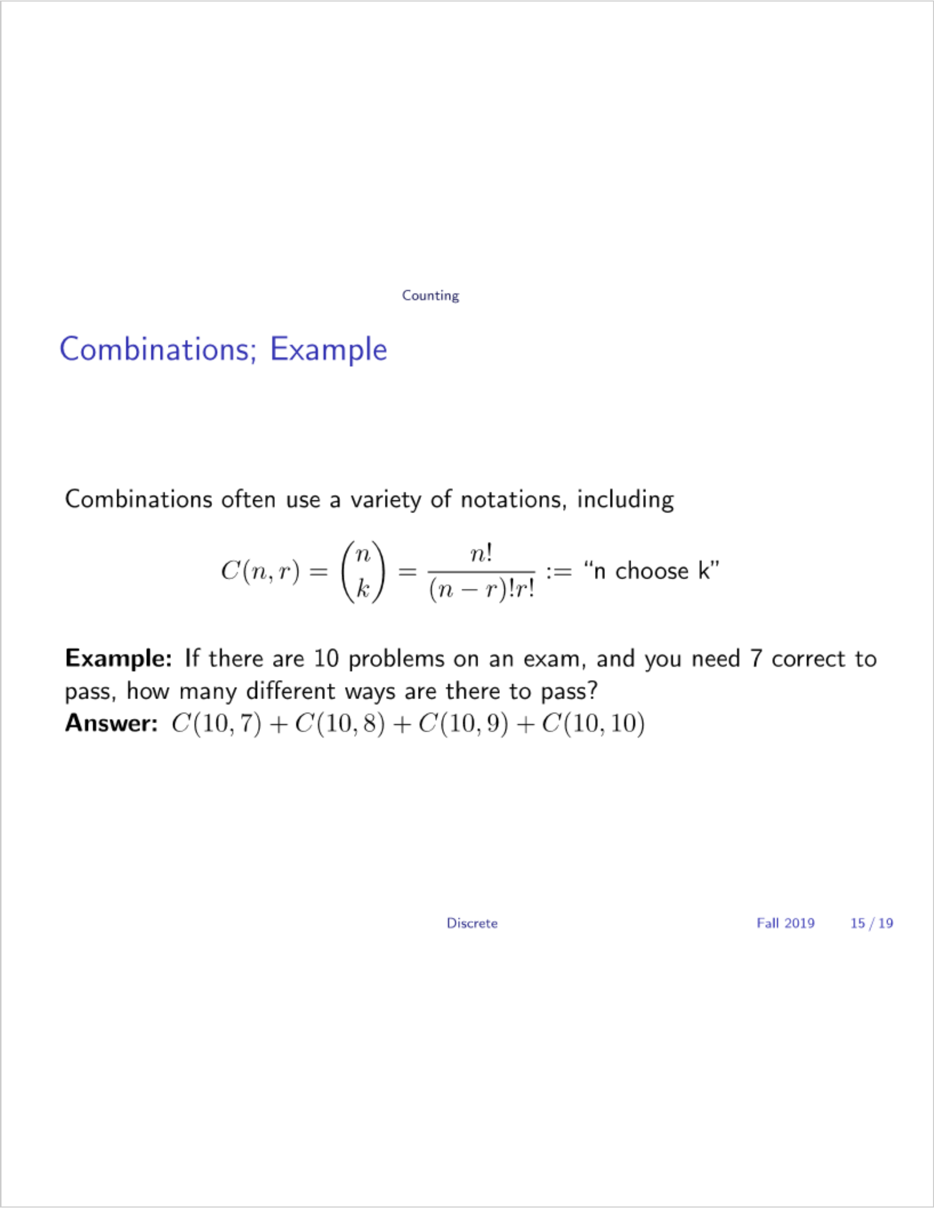
Example: If there are 10 problems on an exam, and you need 7 correct to pass, how many different ways are there to pass?



#ways: C(10,7)

A: C(10,7)

$$\binom{10}{7} + \binom{10}{8} + \binom{10}{9} + \binom{10}{10}$$



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

The Binomial

$\underline{X}$

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 Ω can $X = 3$?

$$\binom{8}{3} = \binom{8}{5}$$

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

$$P(\{HHHTTTTT\}) = (P(H))^3 \cdot (P(T))^5 \overset{\text{fair}}{=} \left(\tfrac{1}{2}\right)^3 \left(\tfrac{1}{2}\right)^8$$

$$= 256.$$

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 Ω can $X = 3$?

   $C(8, 3)$ OR $C(8, 5)$

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

   One such way is $\{HHHTTTTT\}$ which has probability

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 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:

$$P(X=k) = (\text{\# of ways } X=k \text{ can happen}) \cdot P(X=k, \text{ once specifically})$$

$$\uparrow$$

$$= \binom{n}{k} \cdot P(\{k \text{ successes } \underline{\text{then}} \text{ } n-k \text{ failures}\})$$

$$= \binom{n}{k} \cdot p^k (1-p)^{n-k} \qquad x \in \{0,1,2,3,\ldots,n\}$$

NOTATION: We write $X \sim Bin(n,p)$ 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 $\underline{X \sim bin(n, p)}$ 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:

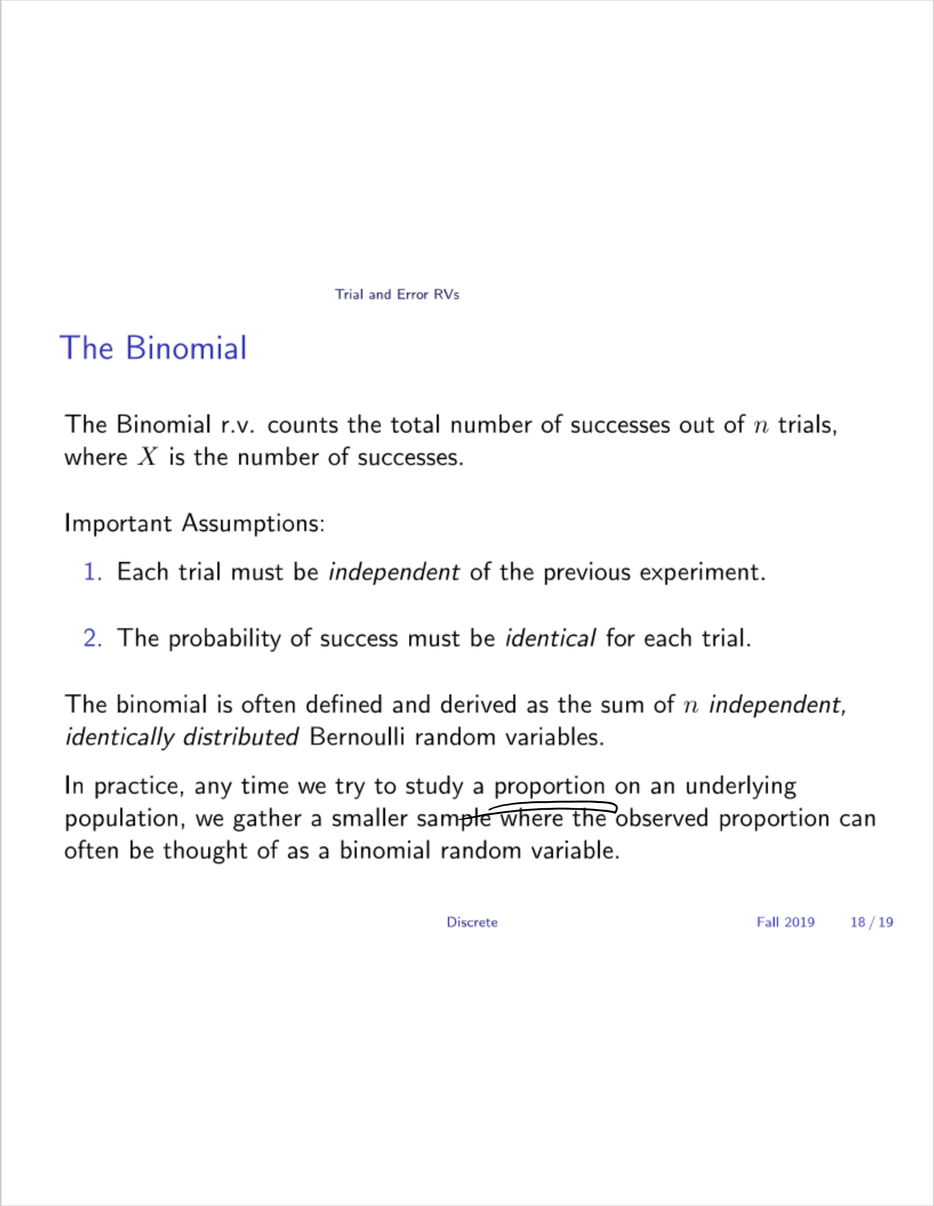
$$P(X = i) = (\# \text{ of ways that } X = i) \cdot P(\text{of one such outcome})$$

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

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

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

NOTATION: We write $\underline{X \sim bin(n, p)}$ 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.

"Counting" Example 1: Suppose that your music app contains 100 songs, 10 of which are by the Beatles. Using the shuffle feature, what is the probability that the first Beatles song heard is the fifth song played?

1. Prior solution: conditioning

2. New solution: counting!