CSCI 3022 Intro to Data Science
Discrete Random Variables

Opening:

What is the difference between a *permutation* and a *combination*?

Announcements and To-Dos

Announcements:

1. Another nb day this Friday.
2. No HW this week Next HW3: Feb 22.

Last time we learned:

1. about pdfs and cdfs

To do:

1. Check out the next set of notebooks!

Probability Distributions

Definition: *Probability Density Function*
A *Probability density function* (pdf) is a function f that describes the probability distribution of a random variable X.

If X is discrete, the pdf or probability mass function (pmf) f gives us
$\boxed{f(x) = P(X = x)}$.

In the continuous case, the pdf instead gives probability to *intervals*. integrals

Definition: *Cumulative Density Function*
For a discrete r.v. X with pdf $f(x) = P(X = x)$, 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.
Mathematically: $F(x) = P(X \leq x)$

Making a pdf

Recall: we did an opening **example**: Suppose we flip a coin with a p chance per flip of landing on heads. Define $X =$ the number of tails flips before we see a heads. What is $P(X = 0)$? $P(X = 1)$? $P(X = i)$? Verify that $P(X) = 1$ over all of Ω.

- State space:
- Associated r.v. possible values or *support:*
- pdf $P(X = x) =$ probability of x tails before a heads:

Making a pdf

Recall: we did an opening **example**: Suppose we flip a coin with a p chance per flip of landing on heads. Define X = the number of tails flips before we see a heads. What is $P(X = 0)$? $P(X = 1)$? $P(X = i)$? Verify that $P(X) = 1$ over all of Ω.

- State space: $\{H, TH, TTH, TTTH, \dots\}$
- Associated r.v. possible values or *support*: $\{0, 1, 2, 3, \dots\}$ valid #3
- pdf $P(X = x)$ = probability of x tails before a heads:

$$P(X = x) = P(\{T \dots TH\}) = \boxed{P(\{T\})^x} P(\{H\}) = (1-p)^x \cdot p$$

So we report $f(x) = (1-p)^x \cdot p$

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:

np.random.choice([list], p=None).

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?

Fair or unfair

how unfair?

This distribution is specified by a single parameter:
The probability of a heads/“success” p! This gives the pdf:

$$\overbrace{P(X=0)}^{\text{"failure"}} = 1-p = f(0)$$

$$\underbrace{P(X=1)}_{\text{success!}} = p = f(1)$$

$$\boxed{f(x) = \begin{cases} p & \text{if } x=1 \\ 1-p & \text{if } x=0 \end{cases}}$$

"is distributed as"

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?

Countable outcomes

This distribution is specified by a single parameter:

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

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

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?

$f(0) = p^{0}(1-p)^{1-0} = 1-p$

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

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:

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

which works as long as we remember x can only be 0 or 1.
We denote the Bernoulli random variable X by $X \sim Bern(p)$

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

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

A

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

{A B} or {B A} = 2 ways

2x more

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

{A B C}, ACB, BAC, BCA, CAB, CBA. = 6 or

3x more

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

Permutations

3

/ n(L)

any of the $\quad$ remaining $\quad$ loser.

$$\underset{1st}{\underline{3}} \cdot \underset{2nd}{\underline{2}} \cdot \underset{3rd}{\underline{1}}$$

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? ABC, ABD, ABE....

$26 \cdot 25 \cdot 24$

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

Permutations; General

$$\underbrace{(n)\ (n-1)\ \cdots\ (n-r+1)}_{r \text{ things.}}$$

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: $P(n, r) = \frac{n!}{(n-r)!}$ $= \dfrac{n(n-1)(n-2)\cdots\cancel{(2)(1)}}{(n-r)(n-r-1)(n-r-2)\cdots\cancel{(2)(1)}}$

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?

ABC is same subset as CBA

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 $\{A, B, C\}$?

ABC
ACB
BAC
BCA
CAB
CBA

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:

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

total

not chosen   # chosen

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

$$\frac{26!}{(23!)\ (3!)}$$

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?

4 ways: we get exactly 7 correct
8
9
10

Choose which 7/8/9/10 are correct.

Combinations; Example

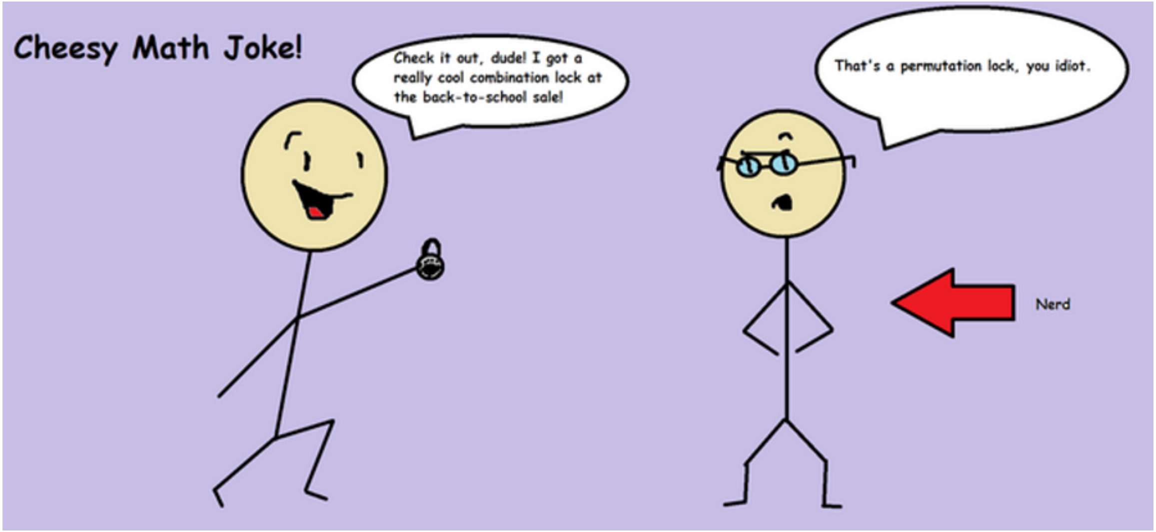
Combinations often use a variety of notations, including

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

$r \leftrightarrow k$

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

Perms and Combs; Summary



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

What are some patterns? It's definitely symmetric - the coefficient 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:

$$(x + y)^4 = (x + y)(x + y)(x + y)(x + y)$$

$\ldots + 4xy^3 + \ldots$

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.

$(a+b)^n$

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

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

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

\binom{upper}{lower}

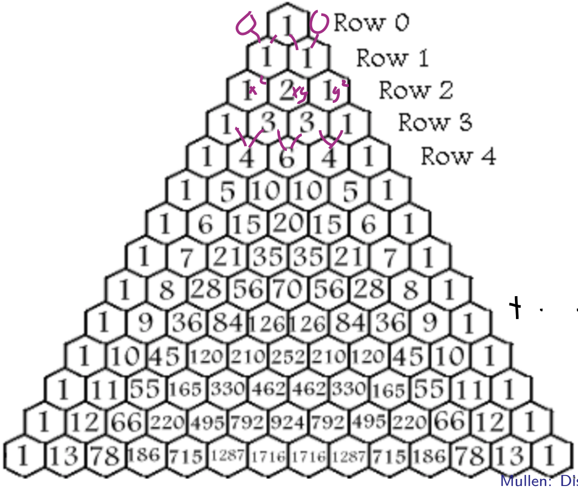
$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^{n-k}y^k = \boxed{\binom{n}{0}} x^n y^0 + \binom{n}{1} x^{n-1}y^1 + \cdots + \binom{n}{n} x^0 y^n$$

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:



$$(2x-3y)^8$$

$$\left([2x]+[-3y]\right)^8$$

$$+\cdots+\binom{8}{5}[2x]^5(-3y)^3+\cdots$$

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.

8 "slots" ; choose 5 x T, 3 x H

1. How many *ways* in Ω can $X = 3$?

   of 8 "flips" how many ways 3 x H, 5 x T

   HHH TTTTT

   ex: HHTTTHTT

   ⋮ etc.

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

   $P(HHHTTTTT) = P(H)P(H)P(H)\cdot P(T)\cdot P(T)\cdot P(T)\cdot P(T)\cdot P(T)$

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

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

The Binomial

$P(WWL) + P(WLW) + P(LWW)$
$= 3 \cdot P(WWL)$

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:

$P(X = x) = P(x$ "wins" and rest "failures"$)$

$= \left(\begin{array}{c}\text{Count ways/orders to get} \\ x \text{ "W"s and } n-x \text{ "Fails"}\end{array}\right) \cdot P(\text{each one}).$

NOTATION: We write ____________ to indicate that X is a Binomial rv with success probability p and n trials.

The Binomial

1100
1001 0011
1010
0110 =6
0101

1 0 1 0

4 slots; choose 2 to be H
choose 2 to be T

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:

$\binom{4}{2} = \frac{4!}{2!2!} = \frac{4 \cdot 3}{2} = 6$

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

4 tries, P(success) = .2

what is P(X=2)? $= \boxed{6 \cdot .2^2 \cdot .8^2}$

$\underbrace{.2 \cdot .2}_{P(\text{success}) \cdot P(\text{success}) = P(2x \text{ success})} \cdot \underbrace{.8 \cdot .8}_{P(2x \text{ fails})} = P(SSFF) = P(1100) = P(HHTT)$

· # of orders of 1100; HHTT

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:

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

choose "i" successes or choose "n - i" failures.

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

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

don't memorize.

n - x copies of (1-p).

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

x copies of "p"

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

tries prob. random variable

The Binomial $\binom{100}{4} = \frac{100!}{96!\,4!}$

$\binom{5}{2} = \frac{5!}{2!(5-2)!} = \frac{5!}{3!\,2!} = \frac{5\cdot 4\cdot \cancel{3\cdot 2\cdot 1}}{\cancel{3\cdot 2\cdot 1}\cdot 2\cdot 1} = 5\cdot 4/2 = 10$

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

$\boxed{= \frac{100\cdot 99\cdot 98\cdot 97}{4\cdot 3\cdot 2\cdot 1}}$

$\frac{543\cancel{21}}{321\;\cancel{21}} = \frac{5\cdot 4\cdot 3}{3\cdot 2\cdot 1} = \frac{60}{6} = 10$

Important Assumptions:

1. Each trial must be *independent* of the previous experiment.

   (gave us products).

2. The probability of success must be *identical* for each trial.

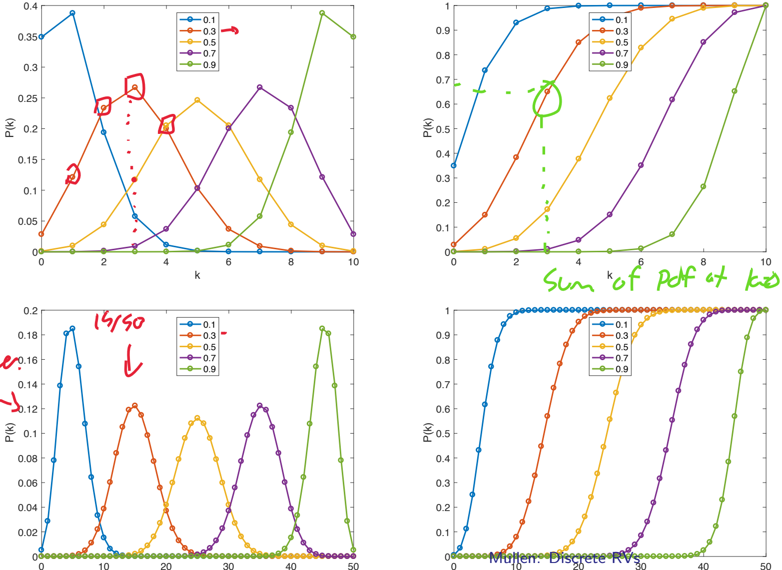
The binomial is often defined and derived as the sum of n *independent, identically distributed* Bernoulli random variables.

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.

Pdfs

cdfs : sum of the left-side of the pdf.

Binomial pdf and cdf. Top: $n = 10$; bottom: $n = 50$.



Prob(3 x H 7 x T) = .26

on 10 flips

where P(H) = .3

The Geometric

Motivating example: A patient is waiting for a suitable matching kidney donor for a transplant. The probability that a randomly selected donor is a suitable match is 0.1.

What is the probability the first donor tested is the first matching donor? Second? Third?

(The per-donor probability checks are independent and identically distributed!)

The Geometric pdf

Continuing in this way, a general formula for the pmf emerges:

The parameter p can assume any value between 0 and 1.
Depending on what parameter p is, we get different members of the geometric distribution.

NOTATION: We write _____________ to indicate that X is a Geometric rv with success probability p.

The Geometric pdf

Continuing in this way, a general formula for the pmf emerges:

$$P(X = x) = P(\text{failed x-1 times}) \cdot P(\text{then success!})$$

$$P(X = x) = (1 - p)^{x-1} p; \quad x \in \{1, 2, 3, \ldots, \infty\}$$

The parameter p can assume any value between 0 and 1.
Depending on what parameter p is, we get different members of the geometric distribution.

NOTATION: We write $\underline{X \sim geom(p)}$ to indicate that X is a Geometric rv with success probability p.

The Geometric pdf

Continuing in this way, a general formula for the pmf emerges:

$$P(X = x) = P(\text{failed x-1 times}) \cdot P(\text{then success!})$$

$$P(X = x) = (1-p)^{x-1}p; \quad x \in \{1, 2, 3, \ldots, \infty\}$$

The parameter p can assume any value between 0 and 1.
Depending on what parameter p is, we get different members of the geometric distribution.

NOTATION: We write $\underline{X \sim geom(p)}$ to indicate that X is a Geometric rv with success probability p.
Important **note:** sometimes the geometric is counting the number of total *trials*; sometimes it's counting the number of *failures*. Know which one your software is doing!

The Negative Binomial

Motivating example:
A "successful toss" is defined to be the coin landing on heads. Let $X = \#$ of failures/tails before the *second* success/heads.

How is this related to the geometric distribution? The binomial distribution?

The Negative Binomial

Motivating example:
A "successful toss" is defined to be the coin landing on heads. Let $X = \#$ of failures/tails before the *second* success/heads.
Events in $X = 2$: $\{HTH, THH\}$
Events in $X = 3$: $\{HTTH, THTH, TTHH\}$
Events in $X = 4$: $\{HTTTH, THTTH, TTHTH, TTTHH\}$

How is this related to the geometric distribution? The binomial distribution?
It's like adding two geometrics.
The relationship to the binomial is a little harder, but if we know this random variables equals x, what do we know about trial $\#x$? The previous $x - 1$ trials?

The Negative Binomial

In general, let $X = \#$ of trials before the rth success. The pdf/pmf is:

NOTATION: We write ______________ to indicate that X is a Negative Binomial rv with success probability p and r successes until completion.

The Negative Binomial

In general, let $X = \#$ of trials before the rth success. The pdf/pmf is:

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

NOTATION: We write $\underline{X \sim NB(r, p)}$ to indicate that X is a Negative Binomial rv with success probability p and r successes until completion.

The Negative Binomial

In general, let $X = \#$ of trials before the rth success. The pdf/pmf is:

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

$$(\# \text{ of ways that } x-1 \text{ trials contain exactly } r-1 \text{ successes})$$

$$\cdot P(\text{r successes and } (x-1)-(r-1) \text{ failures}).$$

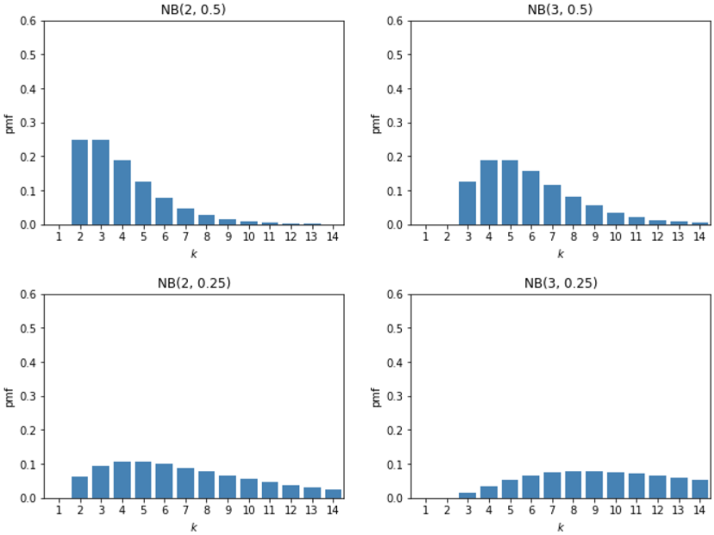
$$= \binom{x-1}{r-1} p^{r-1}(1-p)^{(x-1)-(r-1)} p$$

$$P(X = x) = \binom{x-1}{r-1} p^{r}(1-p)^{(x-r)}$$

for $x = \{r, r+1, r+2, \ldots \infty\}$.

NOTATION: We write $\underline{X \sim NB(r,p)}$ to indicate that X is a Negative Binomial rv with success probability p and r successes until completion.

NB pdfs


NB(2, 0.5)
NB(3, 0.5)
NB(2, 0.25)
NB(3, 0.25)
pmf
k
0.0
0.1
0.2
0.3
0.4
0.5
0.6
1 2 3 4 5 6 7 8 9 10 11 12 13 14

The Negative Binomial

Example:
A physician wishes to recruit 5 people to participate in a new health regimen. Let $p = .2$ be the probability that a randomly selected person agrees to participate. What is the probability that 15 people must be asked before 5 are found who agree to participate?

The Negative Binomial

Example:
A physician wishes to recruit 5 people to participate in a new health regimen. Let p = .2 be the probability that a randomly selected person agrees to participate. What is the probability that 15 people must be asked before 5 are found who agree to participate?

For $X \sim NB(5, .2)$, find $P(X = 15)$:

The Negative Binomial

Example:
A physician wishes to recruit 5 people to participate in a new health regimen. Let p = .2 be the probability that a randomly selected person agrees to participate. What is the probability that 15 people must be asked before 5 are found who agree to participate?

For $X \sim NB(5, .2)$, find $P(X = 15)$:

$$P(X = 15) = \binom{15 - 1}{5 - 1} .2^5 (.8)^{(15-5)}$$

The Poisson Distribution/RV

A Poisson r.v. describes the total number of events that happen in a certain time period.

Examples:
# of vehicles arriving at a parking lot in one week
# of gamma rays hitting a satellite per hour
# of cookies sold at a bake sale in 1 hour

The Poisson Distribution/RV

A Poisson r.v. describes the total number of events that happen in a certain time period.

A discrete random variable X is said to have a Poisson distribution with parameter λ $(\lambda > 0)$ if the pdf of X is

NOTATION: We write ____________ to indicate that X is a Poisson r.v. with parameter λ.

The Poisson Distribution/RV

A Poisson r.v. describes the total number of events that happen in a certain time period.

A discrete random variable X is said to have a Poisson distribution with parameter λ $(\lambda > 0)$ if the pdf of X is

$$P(X = x) = f(x) = \frac{e^{-\lambda}\lambda^x}{x!}; \quad x \in 0, 1, 2, \infty$$

NOTATION: We write $\underline{X \sim Pois(\lambda)}$ to indicate that X is a Poisson r.v. with parameter λ.

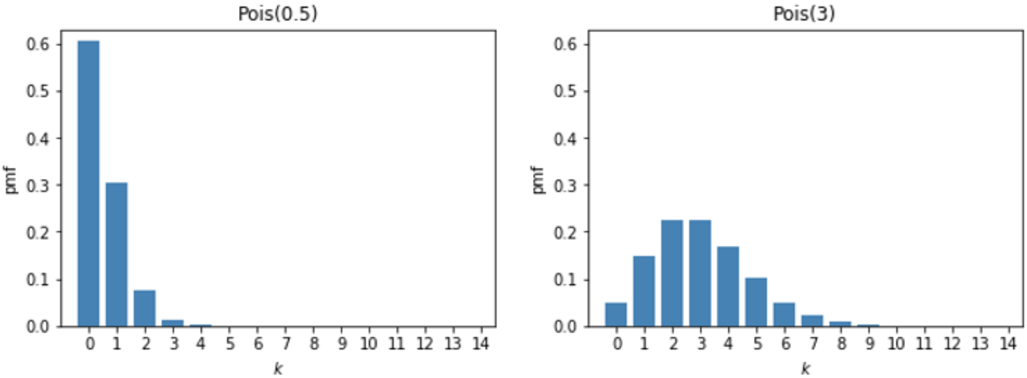
The Poisson Distribution/RV

Example:
Let X denote the number of mosquitoes captured in a trap during a given time period. Suppose that X has a Poisson distribution with $\lambda = 4.5$. What is the probability that the trap contains 5 mosquitoes?

The Poisson Distribution/RV

Example:
Let X denote the number of mosquitoes captured in a trap during a given time period. Suppose that X has a Poisson distribution with $\lambda = 4.5$. What is the probability that the trap contains 5 mosquitoes? $P(X = 5) =$

Poisson pdfs



Poisson and... binomial?

One way to generate the Poisson is to take limits of a binomial: suppose you get texts during class $\left(\frown\right)$ at a rate of 29 texts per hour. What is the probability that you get 29 texts in an hour? 12 texts in an hour? 107 texts in an hour?

λ is the *rate* of the Poisson.

Poisson and... binomial?

One way to generate the Poisson is to take limits of a binomial: suppose you get texts during class $\left(\ddot{\frown}\right)$ at a rate of 29 texts per hour. What is the probability that you get 29 texts in an hour? 12 texts in an hour? 107 texts in an hour?

λ is the *rate* of the Poisson.

Think about a Bernoulli that represents your friends asking "should I text...?" then flipping a coin with probability p. Then:

Poisson and... binomial?

One way to generate the Poisson is to take limits of a binomial: suppose you get texts during class ($\frown$) at a rate of 29 texts per hour. What is the probability that you get 29 texts in an hour? 12 texts in an hour? 107 texts in an hour?

λ is the *rate* of the Poisson.

Think about a Bernoulli that represents your friends asking "should I text...?" then flipping a coin with probability p. Then:

$\lambda = \frac{texts}{hour} \approx \frac{flips}{hour} \cdot \frac{texts}{flip} = np$ for the same n and p as a *binomial*.

Poisson and... binomial?

One way to generate the Poisson is to take limits of a binomial: suppose you get texts during class $\left(\frown\!\!\!\!\!\!\ddot{}\;\right)$ at a rate of 29 texts per hour. What is the probability that you get 29 texts in an hour? 12 texts in an hour? 107 texts in an hour?

λ is the *rate* of the Poisson.

Think about a Bernoulli that represents your friends asking "should I text...?" then flipping a coin with probability p. Then:

$\lambda = \frac{texts}{hour} \approx \frac{flips}{hour} \cdot \frac{texts}{flip} = np$ for the same n and p as a *binomial*.

...but n might vary a bit from hour to hour, so these are only equivalent *in the limit* (n large, p small)!

Discrete Distributions Example

Example:
A factory makes parts for a medical device company. 6% of those parts are defective. For each one of the problems below:

(i.) Define an appropriate random variable for the experiment.

(ii.) Give the values that the random variable can take on.

(ii.) Find the probability that the random variable equals 2.

(iv.) State any assumptions you need to make.

Problems:

1. Out of 10 parts, X are defective.
2. Upon observing an assembly line, X non-defective parts are observed before finding a defective part.
3. X is the number of defective parts made per day, where the rate of defective parts per day is 10.

Discrete Distributions Example

6% of those parts are defective.

1. Out of 10 parts, X are defective.

(i.) r.v.:

(ii.) Values of r.v.:

(iii.) $P(X = 2)$:

(iv.) Assumptions:

Discrete Distributions Example

6% of those parts are defective.

1. Out of 10 parts, X are defective.

(i.) r.v.:

$$X \sim bin(10, .06)$$

(ii.) Values of r.v.:

$$X \in \{0, 1, 2, \ldots, 10\}$$

(iii.) $P(X = 2)$:

$$\binom{10}{2} .06^2 .94^8$$

(iv.) Assumptions: Parts are *i.i.d.*

Discrete Distributions Example

6% of those parts are defective.

2. Upon observing an assembly line, X non-defective parts are observed before finding a defective part.

(i.) r.v.:

(ii.) Values of r.v.:

(iii.) $P(X = 2)$:

(iv.) Assumptions:

Discrete Distributions Example

6% of those parts are defective.

2. Upon observing an assembly line, X non-defective parts are observed before finding a defective part.

(i.) r.v.:

$$X + 1 \sim Geom(.06)$$

(ii.) Values of r.v.:

$$X \in \{0, 1, 2, \dots, \infty\}$$

(iii.) $P(X = 2)$:

$$.94^2 .06^1$$

(iv.) Assumptions: Parts are *i.i.d.*

Discrete Distributions Example

6% of those parts are defective.

3. X is the number of defective parts made per day, where the rate of defective parts per day is 10.

(i.) r.v.:

(ii.) Values of r.v.:

(iii.) $P(X = 2)$:

Discrete Distributions Example

6% of those parts are defective.

3. X is the number of defective parts made per day, where the rate of defective parts per day is 10.

(i.) r.v.:

$$X \sim Pois(10)$$

(ii.) Values of r.v.:

$$X \in \{0, 1, 2, \ldots, \infty\}$$

(iii.) $P(X = 2)$:

$$\frac{e^{-10} \cdot 10^2}{2!}$$

Daily Recap

Today we learned

1. Discrete pdfs!

Moving forward:

- nb day Friday!
- HW 3 due Feb 22 (a week off!)

Next time in lecture:

- *Continuous* pdfs.