# 9-20: More Discrete RVs
# CSCI 3022 Fall 19

Today:
100% less microsoft.

Opening **Example**: Suppose the Detroit Lions are secretly a coin with a $25\%$ chance per game (flip) of winning (landing on heads). Define $X =$ the number of wins (heads) the Lions achieve in a ~~16~~ 4 game season. What is $P(X = 0)$? What is $P(X = 1)$? What is $P(X = 2)$?

# Opening Example Sol'n

Suppose the Detroit Lions are secretly a coin with a $25\%$ chance per game (flip) of winning (landing on heads). Define $X$ = the number of wins (heads) the Lions achieve in a 4 game season. What is $P(X=0)$? What is $P(X=1)$? What is $P(X=2)$?

$$P(X=0) = P(\{LLLL\}) = P(\{L\}) \cdot P(\{L\}) \cdots$$

$$= (.75)^4$$

Wins $X \sim \text{bin}(4, .25)$

$$P(X=1) = \binom{4}{1} \cdot \underbrace{(1-p)^3 \quad p^1}_{\cdot \text{ that prob.}}$$

orders of WLLL

# Opening Example Sol'n

Suppose the Detroit Lions are secretly a coin with a $25\%$ chance per game (flip) of winning (landing on heads). Define $X =$ the number of wins (heads) the Lions achieve in a 4 game season. What is $P(X = 0)$? What is $P(X = 1)$? What is $P(X = 2)$?

1. $P(X = 0) = P(\{LLLL\}) = .75^4$.
2. $P(X = 1) = P(\{WLLL\} \cup \{LWLL\} \cup \{LLLW\} \cup \{LLLW\}) = \boxed{4} \cdot .25^1 \cdot .75^3$.
3. $P(X = 2) = P(\{WWLL\} \cup \{WLWL\} \cup \{WLLW\} \cup \{LWWL\} \cup \{LWLW\} \cup \{LLWW\}) = \boxed{6} \cdot .25^2 \cdot .75^2$.
4. What's the pattern? $\binom{n}{x}$

# Announcements and Reminders

- Homework due Friday the 27th.
- Notebook day on Monday over the discrete rvs.

# Last Time...: Repeated Trials

Counting!

1. Combinations: choose $k$ things out of $n$; $C(n,k) = \binom{n}{k} = \frac{n!}{k!(n-k)!}$
2. Permutations: order all $n$ things: $n!$; order $r$ things out of $n$:

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

.

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.

# Binomials and Bernoullis

A Bernoulli rv is a single trial with success (or "1") with probability $p$, or:

$$P(X = 1) = p; \quad P(X = 0) = (1 - p); \quad \textbf{OR } \underbrace{f(x) = p^x (1 - p)^{(1-x)}}_{n=1}$$

The last way to write it is the same thing as a *binomial* with $n = 1$.

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

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

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

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

$P(x=0)+P(x=1)+P(x=2)=P(X\leq 2)$

CDF

# Binomial $n = 10, 50$

$p=.10$

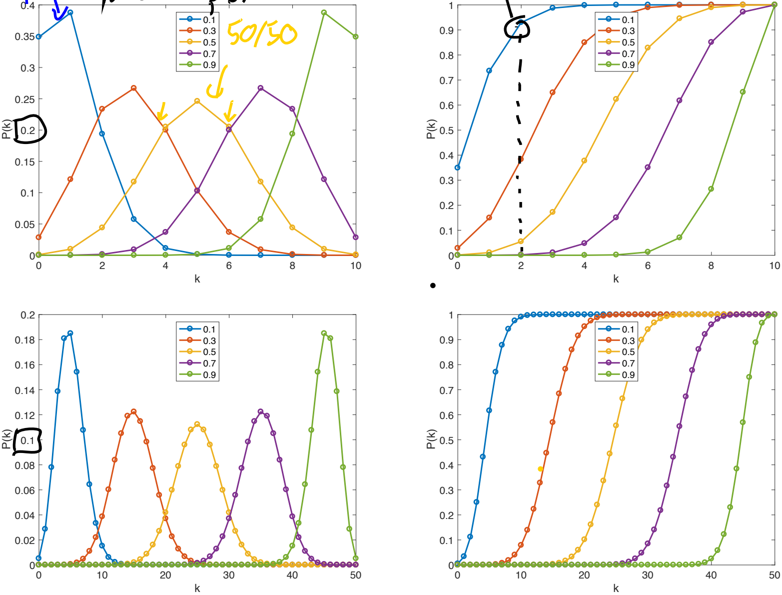
$n=10$

PDF

50/50

$n=10$

$n=50$



**"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? Prior solution: **conditioning**.

**"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? Prior solution: **conditioning**. Idea: each song is equally likely sample from the unplayed songs, so we can use conditionals:

$$
\begin{aligned}
P(\text{songs 1-4 are N, song 5 is B}) &= P(\text{5=B GIVEN 1-4=N})P(\text{1-4=N}) \\
&= \frac{10}{96} \cdot P(\text{1-4=N}) \\
&\cdots \\
&= \frac{10}{96} \cdot \frac{87}{97} \cdot \frac{88}{98} \cdot \frac{89}{99} \cdot \frac{\overset{90}{\cancel{99}}}{100}
\end{aligned}
$$

<u>or</u> Counting

$$\frac{\text{\# of outcomes with NNNNB}}{\text{\# all possible shuffles}}$$

**"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?

Prior solution: **conditioning**.

$$26 \cdot 25 \cdot 24 = \frac{26!}{23!}$$

$$26 \cdot 26 \cdot 26 = 26^3$$

New solution: **counting**!

$$P(E) = \frac{|E|}{|\Omega|} = \frac{\text{Ways for this outcome}}{\text{all shuffles}}$$

all possible first 4 = NNNN

$$P(E) = \frac{\boxed{P(90,4)} \cdot P(10,1)}{\boxed{P(100,5)}} \cdot 10$$

→ all possible first 5

**"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?

Prior solution: **conditioning**. Idea: each song is equally likely sample from the unplayed songs, so we can use conditionals:

$$\begin{aligned} P(\text{songs 1-4 are N, song 5 is B}) &= P(\text{5=B GIVEN 1-4=N})P(\text{1-4=N}) \\ &= \frac{10}{96} \cdot P(\text{1-4=N}) \\ &\dots \\ &= \frac{10}{96} \cdot \frac{87}{97} \cdot \frac{88}{98} \cdot \frac{89}{99} \cdot \frac{\overset{90}{\cancel{99}}}{100} \end{aligned}$$

New solution: **counting**!

$$P(E) = \frac{|E|}{|\Omega|} = \frac{\text{Ways for this outcome}}{\text{all shuffles}}$$

$$P(E) = \frac{P(90,4) \cdot P(10,1)}{P(100,5)} = \frac{90 \cdot 89 \cdot 88 \cdot 87 \cdot 10}{100 \cdot 99 \cdot 98 \cdot 97 \cdot 96}$$

# 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!)

$P(X=1) = p = .1$

$P(X=2) = P(\text{fail then succeed})$

$= P(\text{fail}) \cdot P(\text{succeed})$

$= .9 \cdot .1$

# The Geometric pdf

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

X: # of trials/ patients

$P(X=1) = P$

$P(X=2) = (1-P)\cdot P$

$P(X=x) = P($ fail $x-1$ times, then succeed$)$.

$= (1-p)^{x-1} \cdot P.$ for $x = \{1, 2, \ldots\}$

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 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 $p$.

# 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! $\Rightarrow \quad P(X=x) = (1-p)^{x} \cdot p \quad x \in \{0, 1, \ldots\}$

# 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.

$X=2$: {FFTT} or {FTFT} or {TFFT}

$X=1$: {FTT} or {TFT}

$X=0$: {TT}

end in T; have 1x T prior to that 2

How is this related to the geometric distribution? The binomial distribution? First, NB for 2 success is summing 2 geometrics.

binomial!

# 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 $\underbrace{\text{trial } \#x?}_{\overset{\frown}{-}T}$ The previous $x - 1$ trials?

# The Negative Binomial

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

$$\{\underbrace{\cdots\cdots}_{\text{length } x-1;\ r-1 \times T} T\}$$

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

= (# of ways to get $r-1 \times T$ in $x-1$ tries) · (that Prob) · P( success in trial #$x$).

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

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

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 $r$th success. The pdf/pmf is:

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

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

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

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

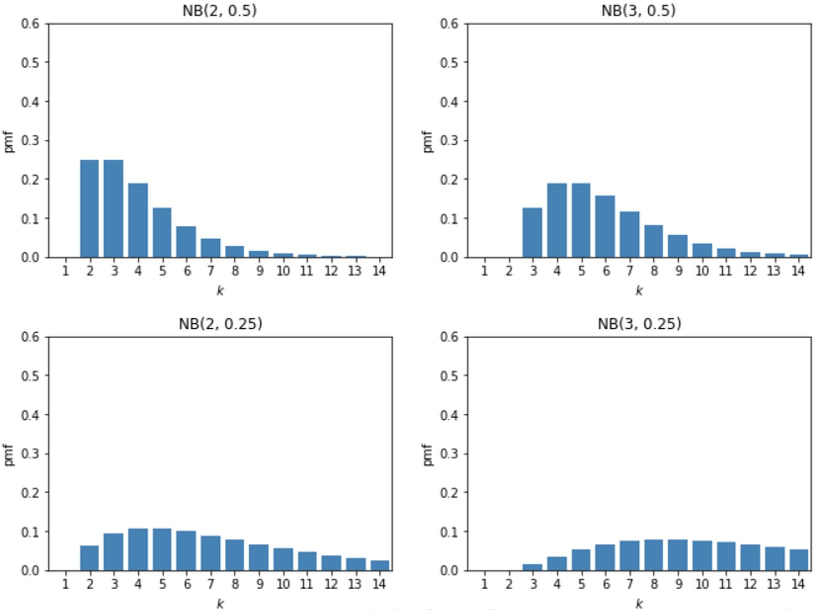
$$P(X = x) = \binom{x-1}{r-1} p^{r} (1-p)^{(x-r)}$$

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

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

$r=2$ $p=.5$


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

# The Negative Binomial

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

**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?

$$X \sim NB(5, .2)$$

want: $P(X=15) = \binom{15-1}{5-1} p^5 (1-p)^{10}$

or $P(X \geq 15) = \sum_{n=15}^{\infty} \binom{n-1}{5-1} p^5 (1-p)^{n-5}$

$$= 1 - \sum_{n=5}^{14} f(x)$$

# 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. ; rate of occurrence is $\lambda$

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

$$P(X=x) = f(x) = \frac{e^{-\lambda} \lambda^{x}}{x!}$$

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

$$\sum_{x=0}^{\infty} \frac{\lambda^{x}}{x!} = e^{\lambda}$$

↑

Taylor Series for $e^{x}$ at $x=\lambda$.

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

# 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 $\lambda$ $(\lambda > 0)$ 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 $\lambda$.

# 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?

$4.5 \frac{\text{mosquitos}}{\text{time}}$

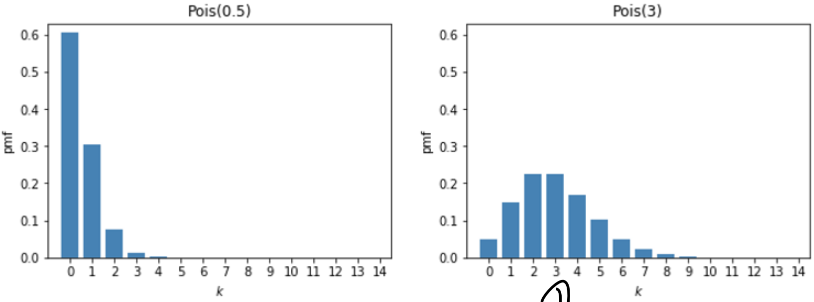
$$P(X=5) = \frac{e^{-4.5}(4.5)^5}{5!}$$

# 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

$\lambda=3$



# 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?

$\lambda$ 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?

$\lambda$ 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 $\left(\ddot\frown\right)$ 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?

> $\lambda$ 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 $(\ddot{\frown})$ 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?

$\lambda$ 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.) $P(X = 2)$:

(iv.) Assumptions:

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

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

(iv.) Assumptions: Parts are... *Poisson*?