07: Random Variables II

Lisa Yan
April 22, 2020

Binomial RV

Binomial Random Variable

Consider an experiment: n independent trials of Ber(p) random variables.

def A **Binomial** random variable X is the number of successes in n trials.

$$X \sim \text{Bin}(n, p)$$

Support: $\{0, 1, \dots, n\}$

PMF $k = 0, 1, \dots, n$:

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

Expectation $E[X] = np$

Variance $\text{Var}(X) = np(1-p)$

Examples:

- # heads in n coin flips
- # of 1's in randomly generated length n bit string
- # of disk drives crashed in 1000 computer cluster (assuming disks crash independently)


SLOW
DOWN

Reiterating notation

1. The random variable

$$X \sim \text{Bin}(n, p)$$

2. is distributed as a

3. Binomial

4. with parameters

The parameters of a Binomial random variable:

- n: number of independent trials
- p: probability of success on each trial

Reiterating notation

$$X \sim \mathrm{Bin}(n, p)$$

If X is a binomial with parameters n and p, the PMF of X is

$$\underbrace{P(X = k)}_{\text{Probability that } X \text{ takes on the value } k} = \underbrace{\binom{n}{k} p^k (1 - p)^{n-k}}_{\textbf{Probability Mass Function} \text{ for a Binomial}}$$

Three coin flips

$X \sim \text{Bin}(n, p) \quad p(k) = \binom{n}{k} p^k (1-p)^{n-k}$

Three fair ("heads" with $p = 0.5$) coins are flipped.

- X is number of heads
- $X \sim \text{Bin}(3, 0.5)$

Compute the following event probabilities:

$$P(X = 0)$$

$$P(X = 1)$$

$$P(X = 2)$$

$$P(X = 3)$$

$$P(X = 7)$$

P(event)

Three coin flips

$X \sim \text{Bin}(n, p) \quad p(k) = \binom{n}{k} p^k (1-p)^{n-k}$

Three fair ("heads" with $p = 0.5$) coins are flipped.

- X is number of heads
- $X \sim \text{Bin}(3, 0.5)$

Compute the following event probabilities:

$$P(X=0) \quad = p(0) \quad = \binom{3}{0} p^0 (1-p)^3 = \frac{1}{8}$$

$$P(X=1) \quad = p(1) \quad = \binom{3}{1} p^1 (1-p)^2 = \frac{3}{8}$$

$$P(X=2) \quad = p(2) \quad = \binom{3}{2} p^2 (1-p)^1 = \frac{3}{8}$$

$$P(X=3) \quad = p(3) \quad = \binom{3}{3} p^3 (1-p)^0 = \frac{1}{8}$$

$$P(X=7) \quad = p(7) \quad = 0$$

P(event) PMF

Extra math note:
By Binomial Theorem,
we can prove
$\sum_{k=0}^{n} P(X=k) = 1$

Binomial Random Variable

Consider an experiment: n independent trials of Ber(p) random variables.

def A **Binomial** random variable X is the number of successes in n trials.

$X \sim \text{Bin}(n, p)$

Range: $\{0, 1, \ldots, n\}$

PMF $k = 0, 1, \ldots, n$:

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

Expectation $E[X] = np$

Variance $\text{Var}(X) = np(1 - p)$

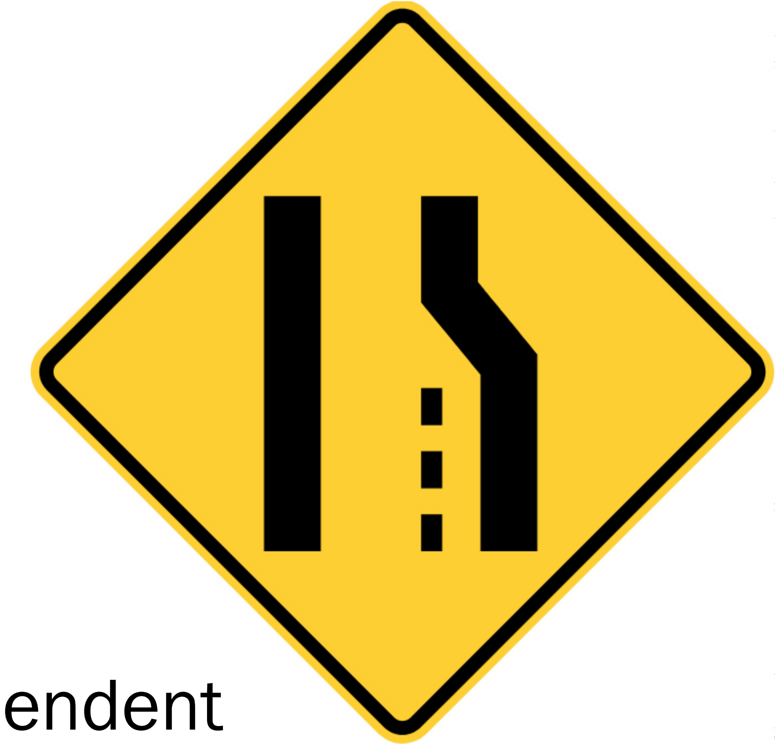
Examples:

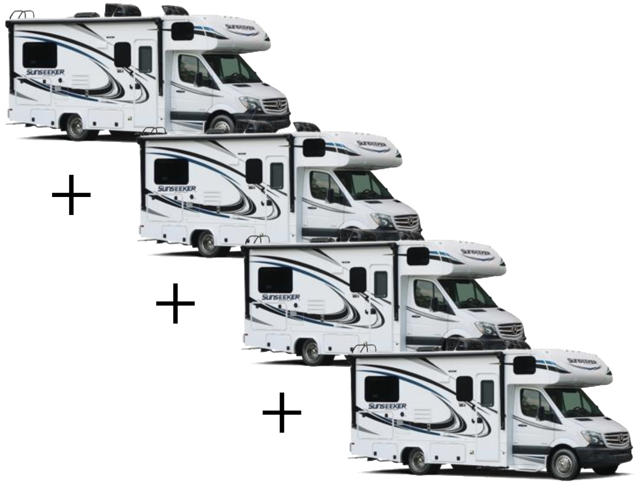
- # heads in n coin flips
- # of 1's in randomly generated length n bit string
- # of disk drives crashed in 1000 computer cluster (assuming disks crash independently)

Binomial RV is sum of Bernoulli RVs

Bernoulli

- $X \sim \text{Ber}(p)$





Binomial

- $Y \sim \text{Bin}(n, p)$
- The sum of n independent Bernoulli RVs

$$Y = \sum_{i=1}^{n} X_i, \qquad X_i \sim \text{Ber}(p)$$

$\text{Ber}(p) = \text{Bin}(1, p)$

Binomial Random Variable

Consider an experiment: n independent trials of Ber(p) random variables.

def A **Binomial** random variable X is the number of successes in n trials.

$X \sim \text{Bin}(n, p)$

Range: $\{0, 1, \ldots, n\}$

PMF $k = 0, 1, \ldots, n$:

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

Expectation $E[X] = np$

Variance $\text{Var}(X) = np(1 - p)$

Examples:

- # heads in n coin flips
- # of 1's in randomly generated length n bit string
- # of disk drives crashed in 1000 computer cluster (assuming disks crash independently)

Proof:

Binomial Random Variable

Consider an experiment: n independent trials of Ber(p) random variables.

def A **Binomial** random variable X is the number of successes in n trials.

$$X \sim \text{Bin}(n, p)$$

Range: $\{0, 1, \ldots, n\}$

PMF $k = 0, 1, \ldots, n$:

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

Expectation $E[X] = np$

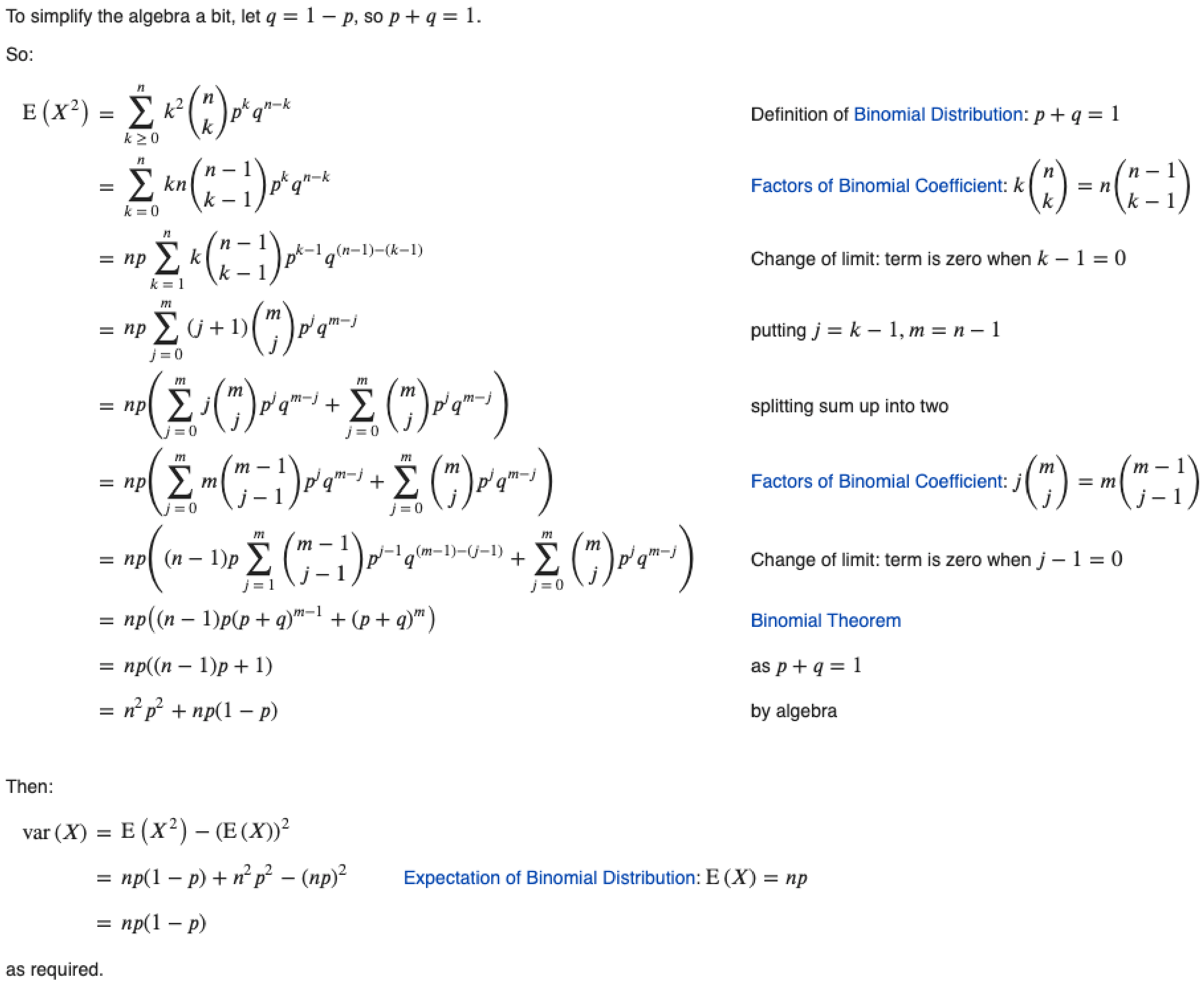
Variance $\text{Var}(X) = np(1-p)$

We'll prove this later in the course

Examples:

- # heads in n coin flips
- # of 1's in randomly generated length n bit string
- # of disk drives crashed in 1000 computer cluster (assuming disks crash independently)

No, give me the variance proof right now



To simplify the algebra a bit, let $q = 1 - p$, so $p + q = 1$.

So:

$$
\begin{aligned}
\operatorname{E}\left(X^2\right) &= \sum_{k \ge 0}^{n} k^2 \binom{n}{k} p^k q^{n-k} && \text{Definition of Binomial Distribution: } p + q = 1 \\
&= \sum_{k = 0}^{n} kn \binom{n-1}{k-1} p^k q^{n-k} && \text{Factors of Binomial Coefficient: } k\binom{n}{k} = n\binom{n-1}{k-1} \\
&= np \sum_{k = 1}^{n} k \binom{n-1}{k-1} p^{k-1} q^{(n-1)-(k-1)} && \text{Change of limit: term is zero when } k - 1 = 0 \\
&= np \sum_{j = 0}^{m} (j+1) \binom{m}{j} p^j q^{m-j} && \text{putting } j = k - 1, m = n - 1 \\
&= np\left(\sum_{j = 0}^{m} j \binom{m}{j} p^j q^{m-j} + \sum_{j = 0}^{m} \binom{m}{j} p^j q^{m-j}\right) && \text{splitting sum up into two} \\
&= np\left(\sum_{j = 0}^{m} m \binom{m-1}{j-1} p^j q^{m-j} + \sum_{j = 0}^{m} \binom{m}{j} p^j q^{m-j}\right) && \text{Factors of Binomial Coefficient: } j\binom{m}{j} = m\binom{m-1}{j-1} \\
&= np\left((n-1)p \sum_{j = 1}^{m} \binom{m-1}{j-1} p^{j-1} q^{(m-1)-(j-1)} + \sum_{j = 0}^{m} \binom{m}{j} p^j q^{m-j}\right) && \text{Change of limit: term is zero when } j - 1 = 0 \\
&= np\left((n-1)p(p+q)^{m-1} + (p+q)^m\right) && \text{Binomial Theorem} \\
&= np((n-1)p + 1) && \text{as } p + q = 1 \\
&= n^2 p^2 + np(1-p) && \text{by algebra}
\end{aligned}
$$

Then:

$$
\begin{aligned}
\operatorname{var}(X) &= \operatorname{E}\left(X^2\right) - (\operatorname{E}(X))^2 \\
&= np(1-p) + n^2 p^2 - (np)^2 && \text{Expectation of Binomial Distribution: } \operatorname{E}(X) = np \\
&= np(1-p)
\end{aligned}
$$

as required.

proofwiki.org

Poisson

Before we start

The natural exponent e:

$$\lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^n = e^{-\lambda}$$

https://en.wikipedia.org/wiki/E_(mathematical_constant)

Jacob Bernoulli
while studying
compound interest
in 1683

Algorithmic ride sharing

Probability of k requests from this area in the next 1 minute?

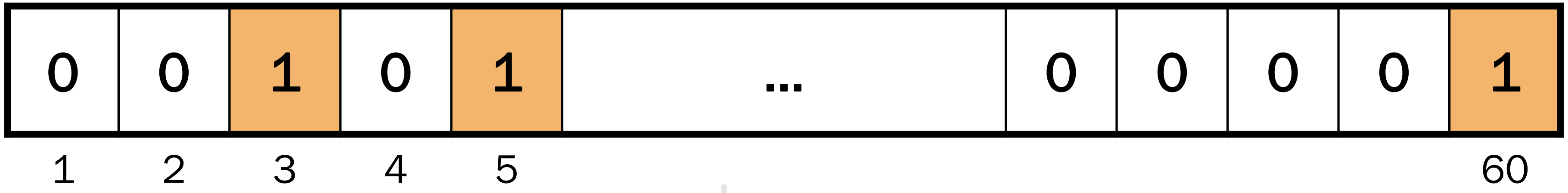
Suppose we know: On average, $\lambda = 5$ requests per minute

Algorithmic ride sharing, approximately

Probability of k requests from this area in the next 1 minute?

On average, $\lambda = 5$ requests per minute

Break a minute down into 60 seconds:



0	0	1	0	1	…	0	0	0	0	1
1	2	3	4	5						60

At each second:

- Independent trial
- You get a request (1) or you don't (0).

Let X = # of requests in minute.

$E[X] = \lambda = 5$

$X \sim \text{Bin}(n = 60,\ p = 5/60)$

$$P(X = k) = \binom{60}{k} \left(\frac{5}{60}\right)^k \left(1 - \frac{5}{60}\right)^{n-k}$$

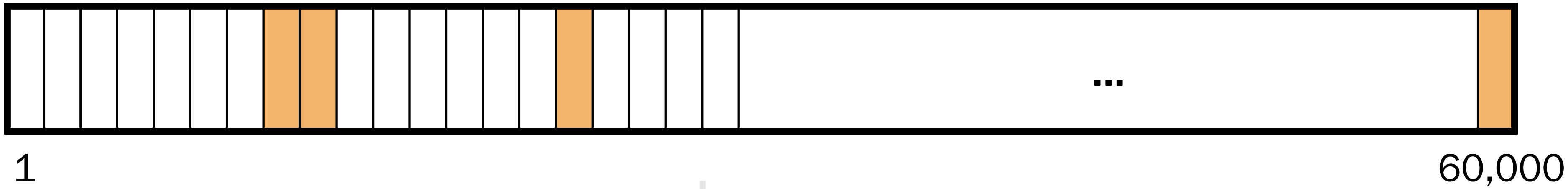
But what if there are *two* requests in the same second?

Algorithmic ride sharing, approximately

Probability of k requests from this area in the next 1 minute?

On average, $\lambda = 5$ requests per minute

Break a minute down into 60,000 milliseconds:



At each millisecond:

- Independent trial
- You get a request (1) or you don't (0).

Let $X =$ # of requests in minute.

$E[X] = \lambda = 5$

$$X \sim \text{Bin}(n = 60000,\ p = \lambda/n)$$

$$P(X = k) = \binom{n}{k}\left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k}$$

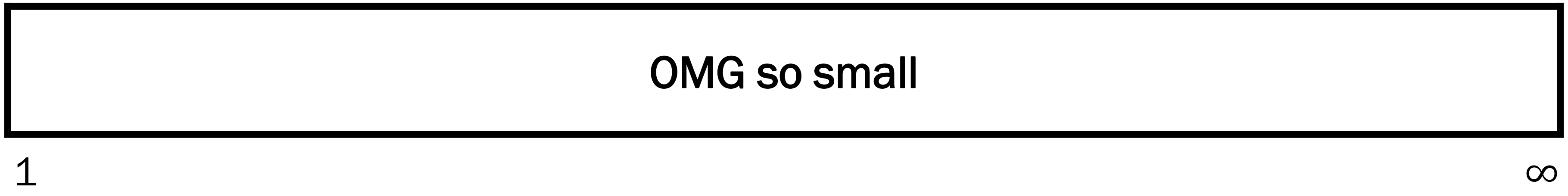
But what if there are *two* requests in the same millisecond?

Algorithmic ride sharing, approximately

Probability of k requests from this area in the next 1 minute?

On average, $\lambda = 5$ requests per minute

Break a minute down into infinitely small buckets:



OMG so small

1

∞

For each time bucket:

- Independent trial
- You get a request (1) or you don't (0).

Let X = # of requests in minute.

$E[X] = \lambda = 5$

$$X \sim \text{Bin}(n,\ p = \lambda/n)$$

$$P(X = k)$$

$$= \lim_{n \to \infty} \binom{n}{k} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k}$$

Who wants to see some cool math?

Binomial in the limit

$$\lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^n = e^{-\lambda}$$

$$P(X=k)$$

$$= \lim_{n\to\infty} \binom{n}{k}\left(\frac{\lambda}{n}\right)^k\left(1-\frac{\lambda}{n}\right)^{n-k}$$

Rearrange

$$= \lim_{n\to\infty} \frac{n!}{n^k(n-k)!}\,\frac{\lambda^k}{k!}\,\frac{\left(1-\frac{\lambda}{n}\right)^n}{\left(1-\frac{\lambda}{n}\right)^k}$$

Expand

$$= \lim_{n\to\infty} \frac{n!}{k!(n-k)!}\,\frac{\lambda^k}{n^k}\,\frac{\left(1-\frac{\lambda}{n}\right)^n}{\left(1-\frac{\lambda}{n}\right)^k}$$

Def natural exponent

$$= \lim_{n\to\infty} \frac{n!}{n^k(n-k)!}\,\frac{\lambda^k}{k!}\,\frac{e^{-\lambda}}{\left(1-\frac{\lambda}{n}\right)^k}$$

Expand

$$= \lim_{n\to\infty} \frac{n(n-1)\cdots(n-k+1)}{n^k}\,\frac{(n-k)!}{(n-k)!}\,\frac{\lambda^k}{k!}\,\frac{e^{-\lambda}}{\left(1-\frac{\lambda}{n}\right)^k}$$

Limit analysis + cancel

$$= \lim_{n\to\infty} \frac{n^k}{n^k}\,\frac{\lambda^k}{k!}\,\frac{e^{-\lambda}}{1}$$

Simplify

$$= \frac{\lambda^k}{k!}e^{-\lambda}$$

Algorithmic ride sharing

Probability of k requests from this area in the next 1 minute?

On average, $\lambda = 5$ requests per minute

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

Simeon-Denis Poisson

French mathematician (1781 – 1840)

- Published his first paper at age 18
- Professor at age 21
- Published over 300 papers

"Life is only good for two things: doing mathematics and teaching it."

Poisson Random Variable

Consider an experiment that lasts a fixed interval of time.

def A **Poisson** random variable X is the number of successes over the experiment duration.

$$X \sim \text{Poi}(\lambda)$$

Support: $\{0,1,2,\ldots\}$

PMF $$P(X = k) = e^{-\lambda}\frac{\lambda^k}{k!}$$

Expectation $E[X] = \lambda$

Variance $\text{Var}(X) = \lambda$

Examples:

- # earthquakes per year
- # server hits per second
- # of emails per day

Yes, expectation == variance for Poisson RV! More later.

Earthquakes

$X \sim \text{Poi}(\lambda)$

$E[X] = \lambda$

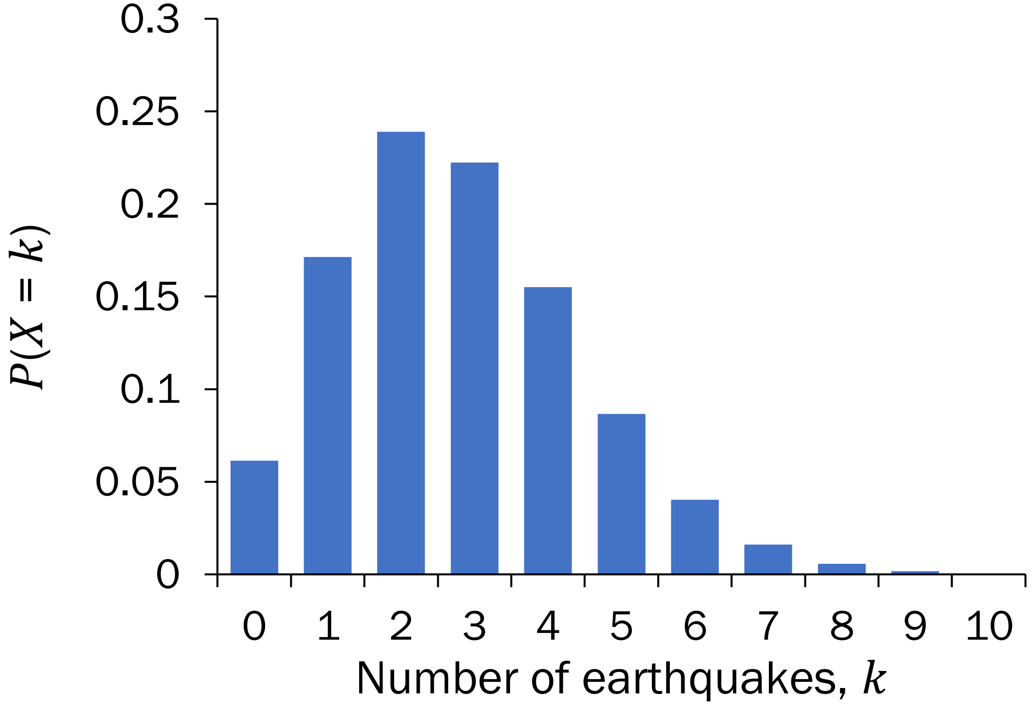
$p(k) = e^{-\lambda}\frac{\lambda^k}{k!}$

There are an average of 2.79 major earthquakes in the world each year.

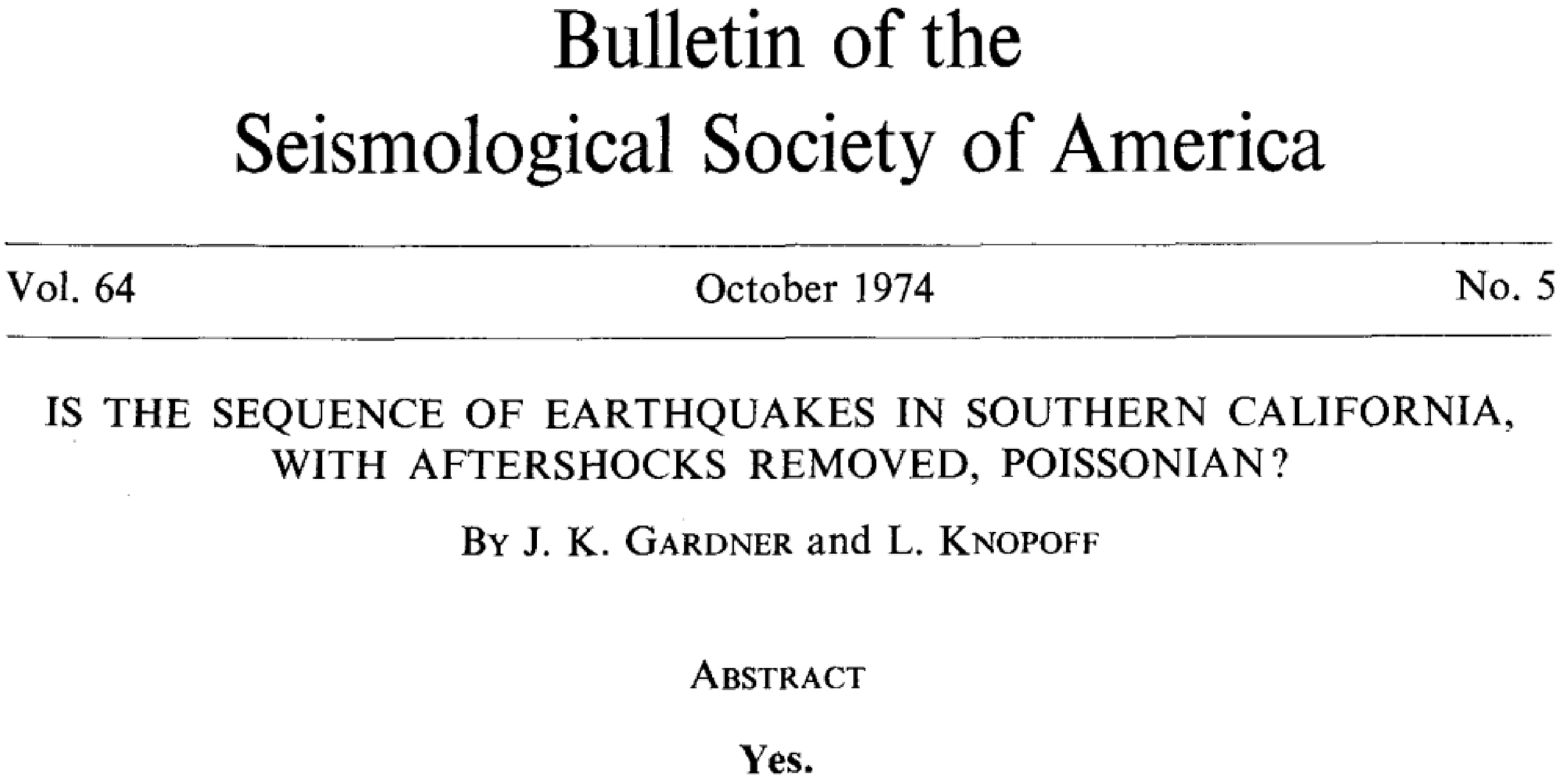
What is the probability of 3 major earthquakes happening next year?

1. Define RVs

2. Solve



Are earthquakes really Poissonian?



Bulletin of the
Seismological Society of America

Vol. 64 October 1974 No. 5

IS THE SEQUENCE OF EARTHQUAKES IN SOUTHERN CALIFORNIA, WITH AFTERSHOCKS REMOVED, POISSONIAN?

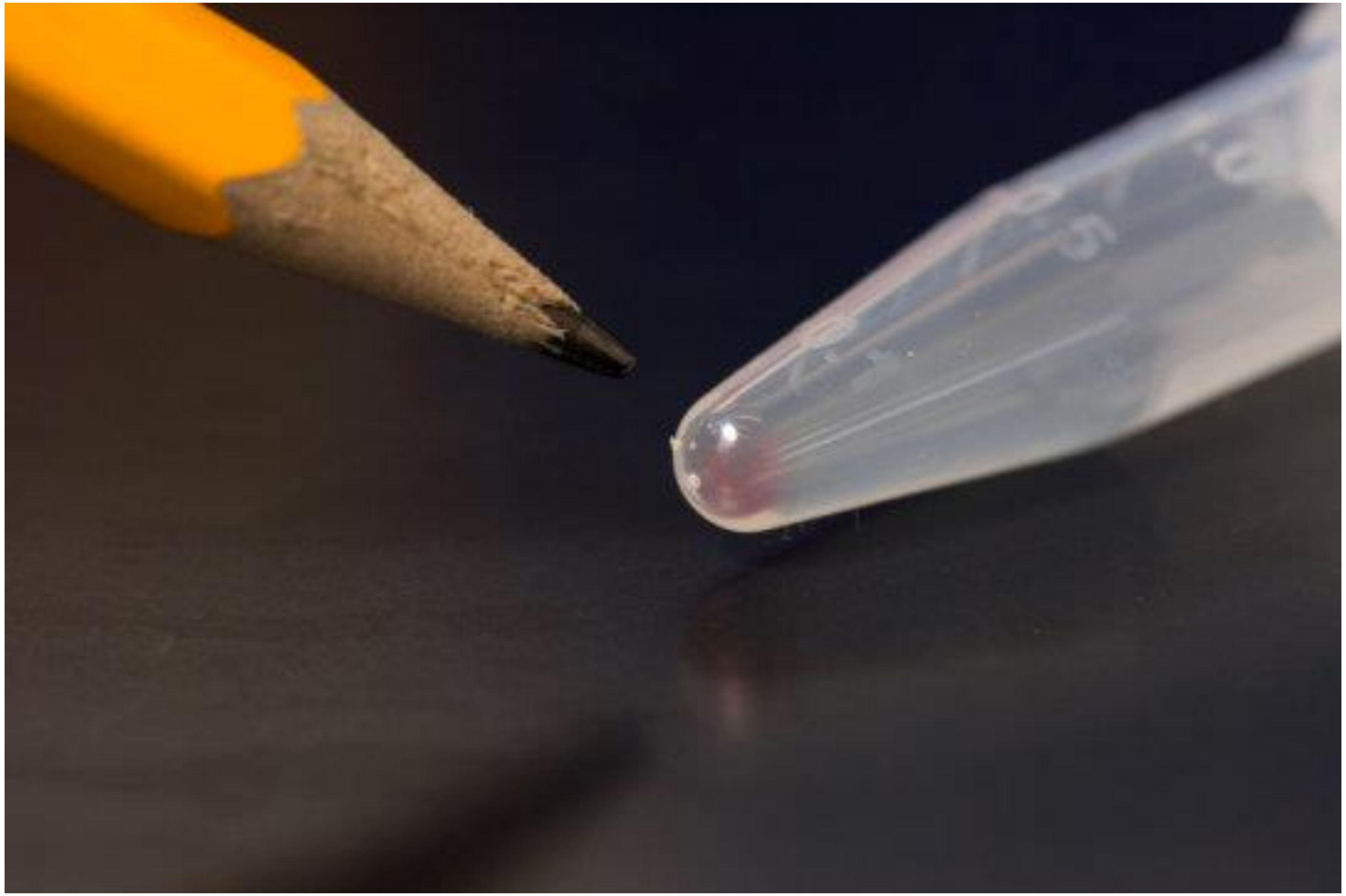
By J. K. Gardner and L. Knopoff

Abstract

Yes.

Poisson Paradigm

DNA



All the movies, images, emails and other digital data from more than 600 smartphones (10,000 GB) can be stored in the faint pink smear of DNA at the end of this test tube.

What is the probability that DNA storage stays uncorrupted?

DNA

What is the probability that DNA storage stays uncorrupted?

- In DNA (and real networks), we store large strings.
- Let string length be long, e.g., $n \approx 10^4$
- Probability of corruption of each base pair is very small, e.g., $p = 10^{-6}$
- Let $X =$ # of corruptions.

What is P(DNA storage is uncorrupted) = $P(X = 0)$?

1. Approach 1:

$$X \sim \text{Bin}(n = 10^4, p = 10^{-6})$$

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

unwieldy! ⚠

$$= \binom{10^4}{0} 10^{-6 \cdot 0} (1 - 10^{-6})^{10^4 - 0}$$

$$\approx 0.99049829$$

2. Approach 2:

$$X \sim \text{Poi}(\lambda = 10^4 \cdot 10^{-6} = 0.01)$$

$$P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!} = e^{-0.01} \frac{0.01^0}{0!}$$

$$= e^{-0.01}$$

$$\approx 0.99049834$$

a good approximation! ✓

The Poisson Paradigm, part 1

$X \sim \text{Poi}(\lambda)$	$Y \sim \text{Bin}(n, p)$
$E[X] = \lambda$	$E[Y] = np$

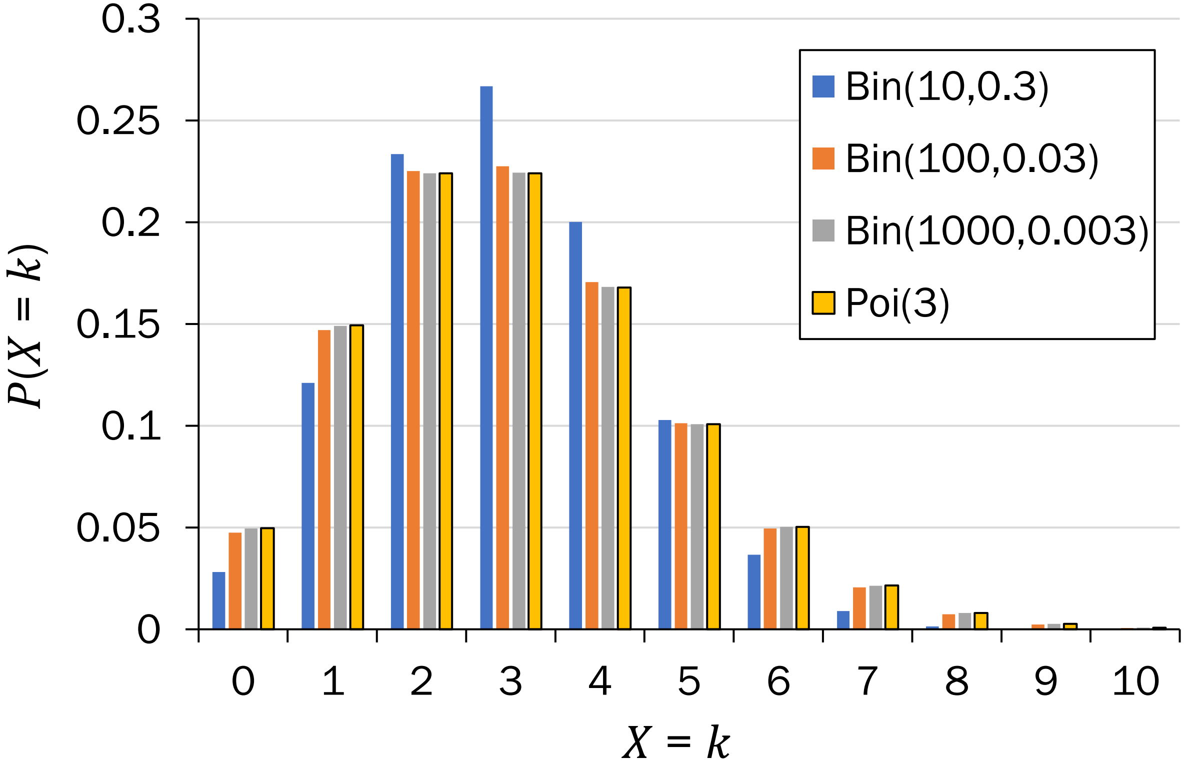
Poisson approximates Binomial when n is large, p is small, and $\lambda = np$ is “moderate.”

Different interpretations of “moderate”:

- $n > 20$ and $p < 0.05$
- $n > 100$ and $p < 0.1$

Poisson is Binomial in the limit:

- $\lambda = np$, where $n \to \infty, p \to 0$



Poisson can approximate Binomial.

Poisson Random Variable

Consider an experiment that lasts a fixed interval of time.

def A **Poisson** random variable X is the number of occurrences over the experiment duration.

$X \sim \text{Poi}(\lambda)$

Support: $\{0, 1, 2, \ldots\}$

PMF $P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}$

Expectation $E[X] = \lambda$

Variance $\text{Var}(X) = \lambda$

Examples:

- # earthquakes per year
- # server hits per second
- # of emails per day

Time to show intuition for why expectation == variance!

Properties of Poi(λ) with the Poisson paradigm

Recall the Binomial:

$Y \sim \text{Bin}(n, p)$	Expectation	$E[Y] = np$
	Variance	$\text{Var}(Y) = np(1-p)$

Consider $X \sim \text{Poi}(\lambda)$, where $\lambda = np$ $(n \to \infty, p \to 0)$:

$X \sim \text{Poi}(\lambda)$	Expectation	$E[X] = \lambda$
	Variance	$\text{Var}(X) = \lambda$

Proof:

$$E[X] = np = \lambda$$
$$\text{Var}(X) = np(1-p) \to \lambda(1-0) = \lambda$$

A Real License Plate Seen at Stanford

No, it’s not mine…
but I kind of wish it was.

Poisson Paradigm, part 2

Poisson can still provide a good approximation of the Binomial, even when assumptions are “mildly” violated.

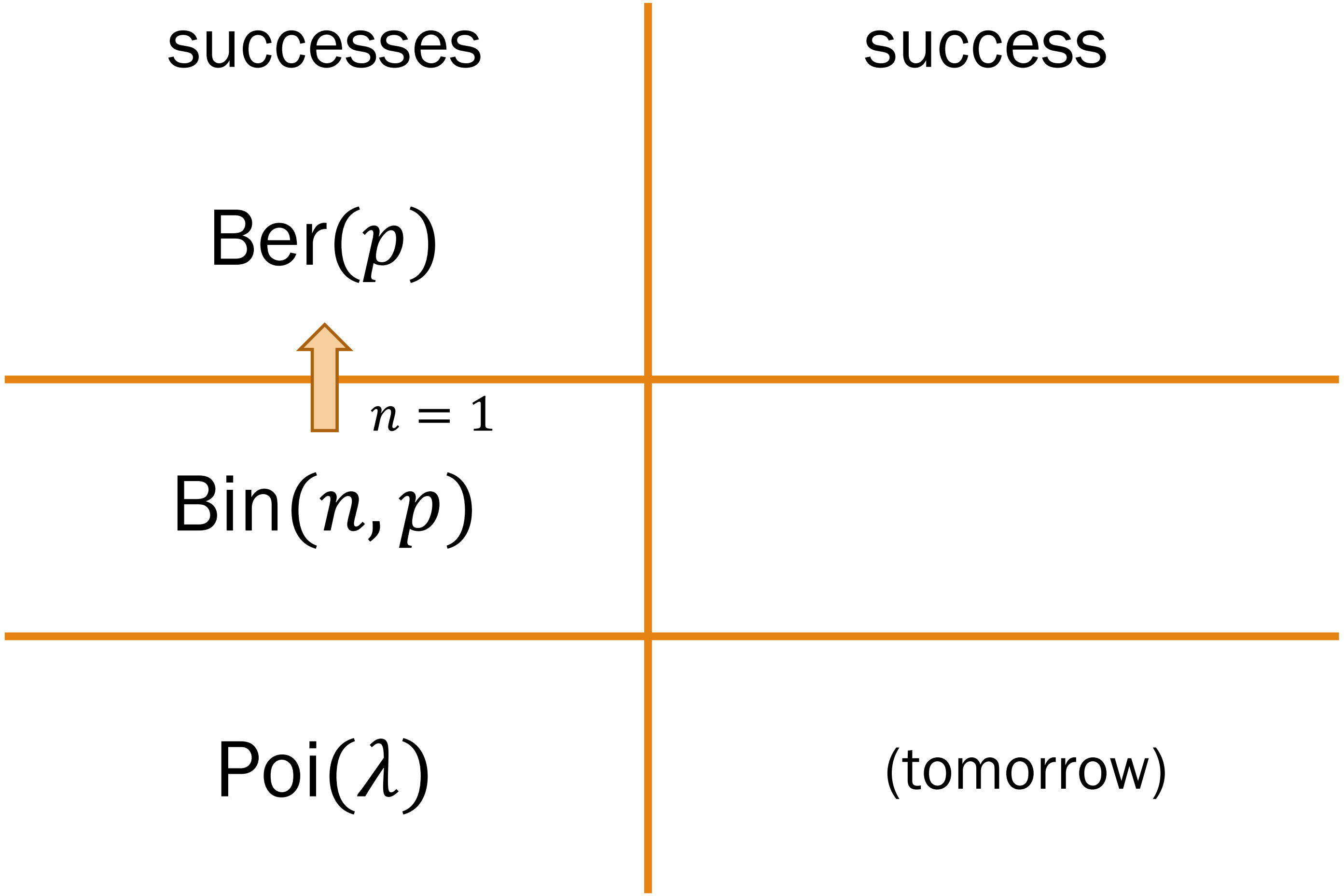
You can apply the Poisson approximation when:

- ”Successes” in trials are not entirely independent
  e.g.: # entries in each bucket in large hash table.
- Probability of “Success” in each trial varies (slightly), like a small relative change in a very small p
  e.g.: Average # requests to web server/sec may fluctuate slightly due to load on network

We won’t explore this too much, but I want you to know it exists.

Other Discrete RVs

Grid of random variables



	Number of successes	Time until success	
One trial	$\text{Ber}(p)$		One success
Several trials	$\text{Bin}(n, p)$		Several successes
Interval of time	$\text{Poi}(\lambda)$	(tomorrow)	Interval of time to first success

Focus on understanding <u>how</u> and <u>when</u> to use RVs, not on memorizing PMFs.

Geometric RV

Consider an experiment: independent trials of Ber(p) random variables.

def A **Geometric** random variable X is the # of trials until the first success.

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

Support: $\{1, 2, \dots\}$

PMF $P(X = k) = (1 - p)^{k-1} p$

Expectation $E[X] = \frac{1}{p}$

Variance $\text{Var}(X) = \frac{1-p}{p^2}$

Examples:

- Flipping a coin ($P(\text{heads}) = p$) until first heads appears
- Generate bits with $P(\text{bit} = 1) = p$ until first 1 generated

Negative Binomial RV

Consider an experiment: independent trials of Ber(p) random variables.

def A **Negative Binomial** random variable X is the # of trials until r successes.

(fixed lecture error)

$$X \sim \text{NegBin}(r, p)$$

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

Expectation $E[X] = \frac{r}{p}$

Variance $\text{Var}(X) = \frac{r(1-p)}{p^2}$

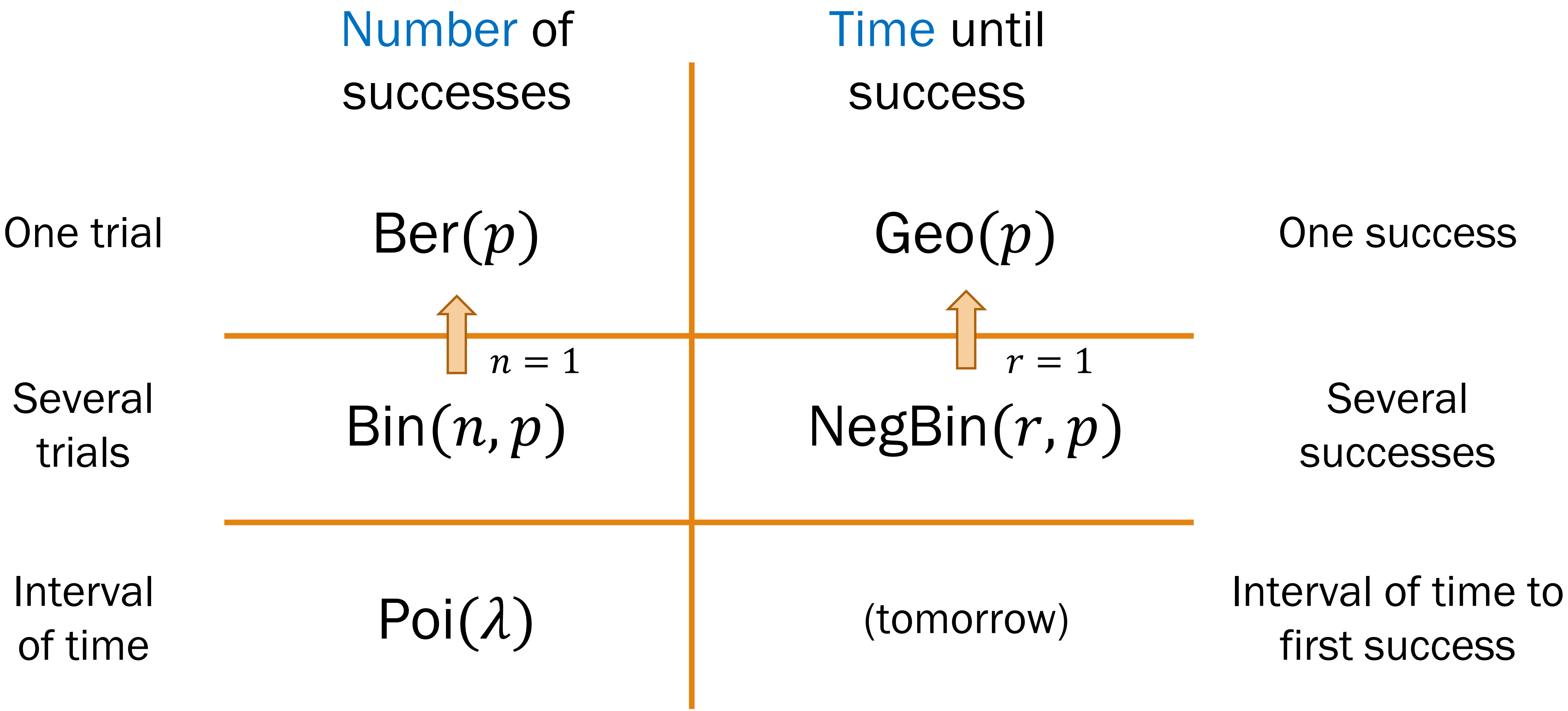
Support: $\{r, r+1, \ldots\}$

Examples:

- Flipping a coin until r^{th} heads appears
- # of strings to hash into table until bucket 1 has r entries

$\text{Geo}(p) = \text{NegBin}(1, p)$

Grid of random variables



	Number of successes	Time until success	
One trial	Ber(p)	Geo(p)	One success
	$n = 1$	$r = 1$	
Several trials	Bin(n, p)	NegBin(r, p)	Several successes
Interval of time	Poi(λ)	(tomorrow)	Interval of time to first success

Catching Pokemon

Wild Pokemon are captured by throwing Pokeballs at them.

- Each ball has probability *p* = 0.1 of capturing the Pokemon.
- Each ball is an independent trial.

What is the probability that you catch the Pokemon on the 5th try?

1. Define events/ RVs & state goal

   $X \sim$ some distribution

   Want: $P(X = 5)$

2. Solve

   A. $X \sim \text{Bin}(5, 0.1)$
   B. $X \sim \text{Poi}(0.5)$
   C. $X \sim \text{NegBin}(5, 0.1)$
   D. $X \sim \text{NegBin}(1, 0.1)$
   E. $X \sim \text{Geo}(0.1)$
   F. None/other

Catching Pokemon

Wild Pokemon are captured by throwing Pokeballs at them.

- Each ball has probability p = 0.1 of capturing the Pokemon.
- Each ball is an independent trial.

What is the probability that you catch the Pokemon on the 5th try?

1. Define events/ RVs & state goal

$X \sim$some distribution

Want: $P(X = 5)$

2. Solve

A. $X \sim \text{Bin}(5, 0.1)$
B. $X \sim \text{Poi}(0.5)$
C. $X \sim \text{NegBin}(5, 0.1)$
D. $X \sim \text{NegBin}(1, 0.1)$
E. $X \sim \text{Geo}(0.1)$
F. None/other

Catching Pokemon

$X \sim \text{Geo}(p) \quad p(k) = (1-p)^{k-1}p$

Wild Pokemon are captured by throwing Pokeballs at them.

- Each ball has probability p = 0.1 of capturing the Pokemon.
- Each ball is an independent trial.

What is the probability that you catch the Pokemon on the 5th try?

1. Define events/ RVs & state goal
2. Solve

$X \sim \text{Geo}(0.1)$

Want: $P(X = 5)$

(live) 08: Random Variables II II

Oishi Banerjee and Cooper Raterink
Adapted from Lisa Yan
July 8, 2020

Our first common RVs

Review

1. The random variable

$$X \sim \mathrm{Ber}(p)$$

2. is distributed as a

3. Bernoulli

4. with parameter

Example: Heads in one coin flip,
P(heads) = 0.8 = p

$$Y \sim \mathrm{Bin}(n, p)$$

Example: # heads in 40 coin flips,
P(heads) = 0.8 = p

otherwise

Identify PMF, or
identify as a function of an
existing random variable

Think

The next slide has a matching question to go over by yourself. We'll go over it together afterwards.

Post any clarifications here!

https://us.edstem.org/courses/667/discussion/84212

Think by yourself: 2 min

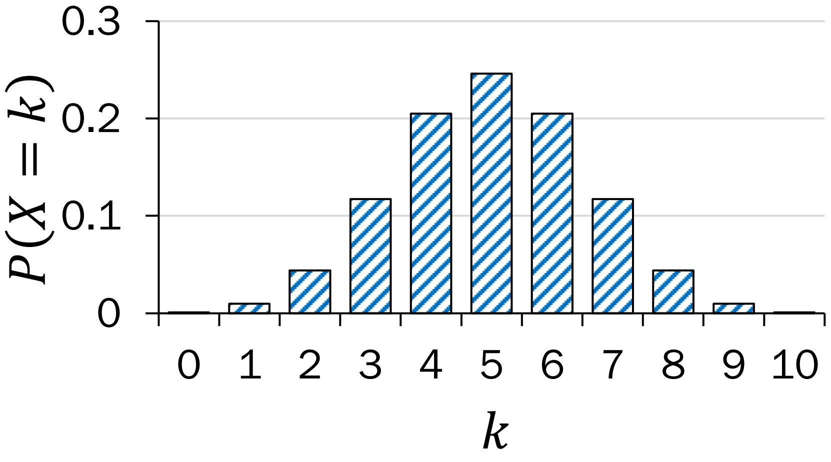
(by yourself)

Visualizing Binomial PMFs

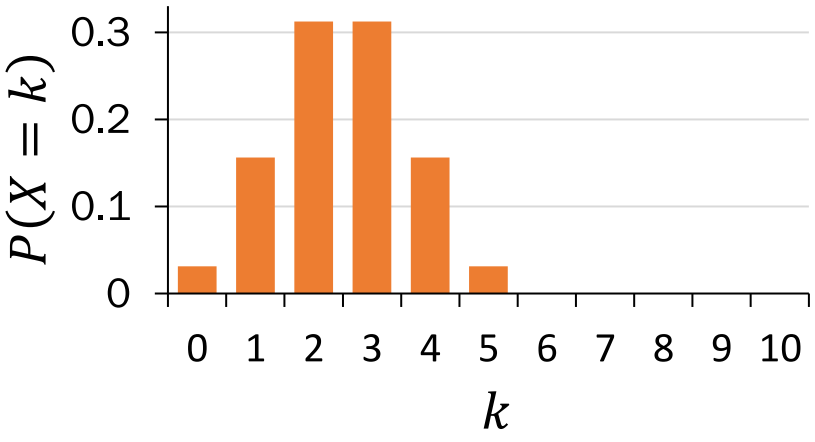
$$E[X] = np$$

$$X \sim \text{Bin}(n, p) \qquad p(i) = \binom{n}{k} p^k (1-p)^{n-k}$$

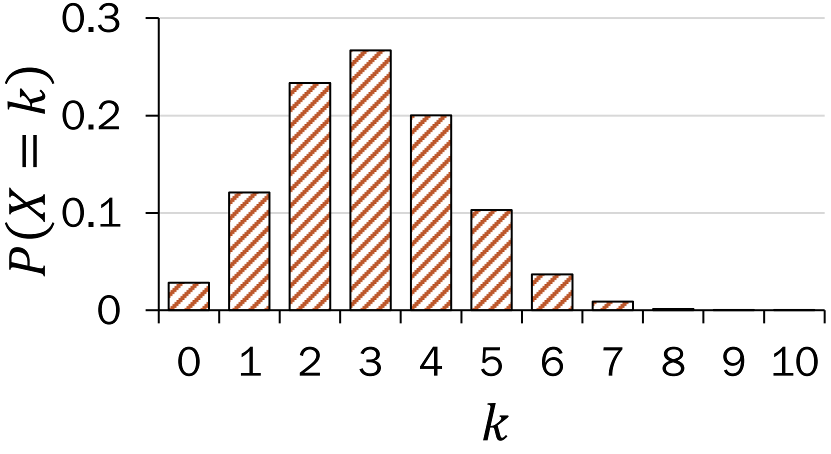
A.



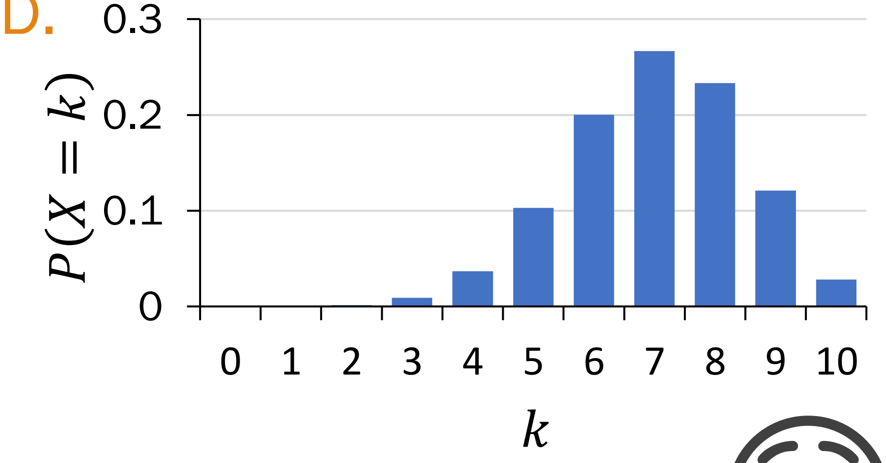
B.



C.



D.



Match the distribution to the graph:

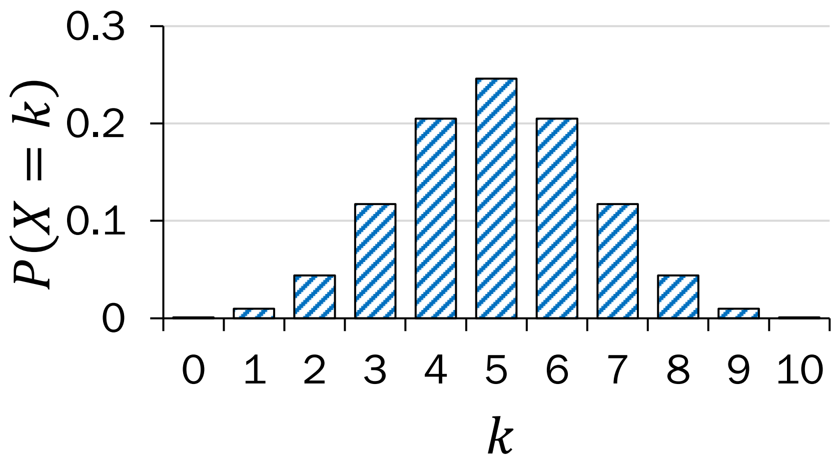
1. $\text{Bin}(10, 0.5)$
2. $\text{Bin}(10, 0.3)$
3. $\text{Bin}(10, 0.7)$
4. $\text{Bin}(5, 0.5)$

(by yourself)

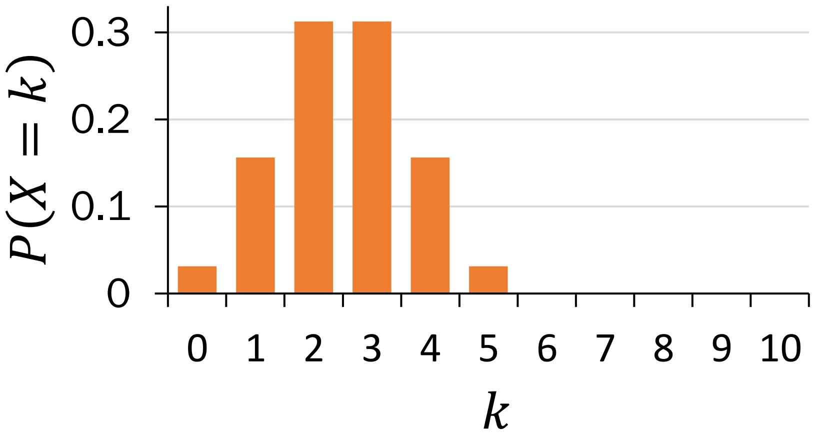
Visualizing Binomial PMFs

$$X \sim \text{Bin}(n, p) \qquad E[X] = np \qquad p(i) = \binom{n}{k} p^k (1-p)^{n-k}$$

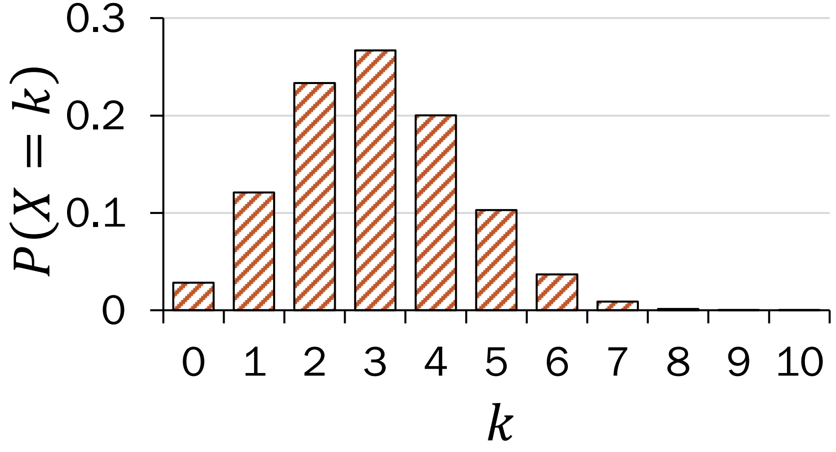
A.



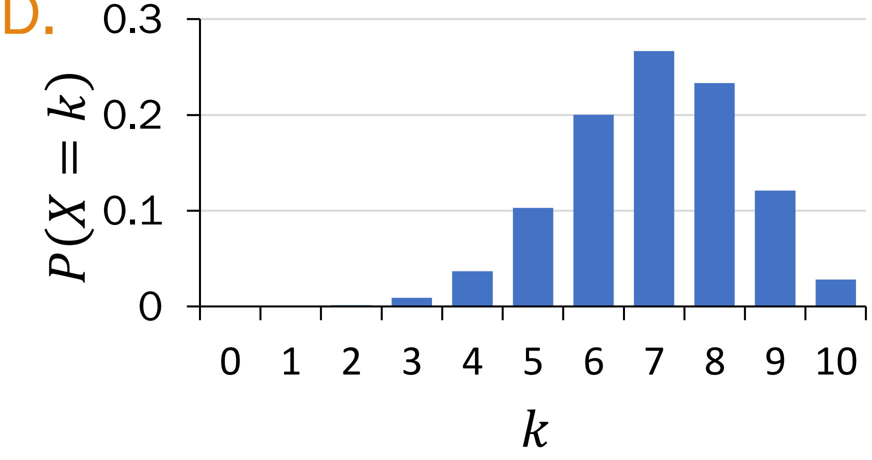
B.



C.



D.



Match the distribution to the graph:

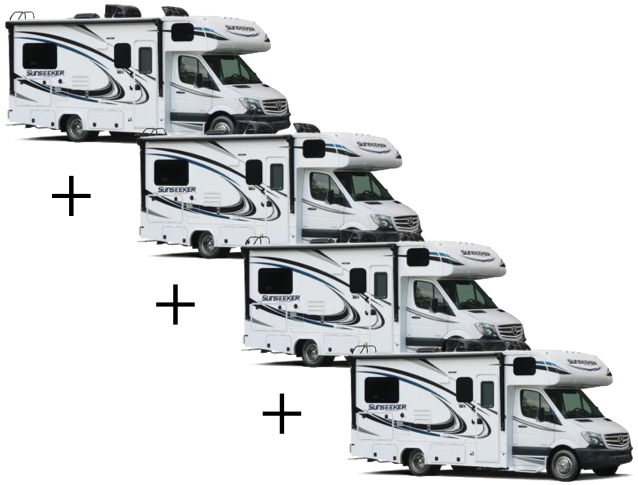
1. Bin(10,0.5)
2. Bin(10,0.3)
3. Bin(10,0.7)
4. Bin(5,0.5)

Binomial RV is sum of Bernoulli RVs

Review

Bernoulli

- $X \sim \text{Ber}(p)$



Binomial

- $Y \sim \text{Bin}(n, p)$
- The sum of n independent Bernoulli RVs

$$Y = \sum_{i=1}^{n} X_i, \qquad X_i \sim \text{Ber}(p)$$

NBA Finals and genetics

Think, then Breakout Rooms

Check out the questions on the next slide.
Post any clarifications here!

https://us.edstem.org/courses/667/discussion/84212

By yourself: 2 min

Breakout rooms: 5 min.

NBA Finals and genetics

1. The Golden State Warriors are going to play the Toronto Raptors in a 7-game series during the 2019 NBA finals.
   - The Warriors have a probability of 58% of winning each game, independently.
   - A team wins the series if they win at least 4 games (we play all 7 games).

What is P(Warriors winning)?

2. Each person has 2 genes per trait (e.g., eye color).

- Child receives 1 gene (equally likely) from each parent
- Brown is “dominant”, blue is ”recessive”:
  - Child has brown eyes if either (or both) genes are brown
  - Blue eyes only if both genes are blue.
- Parents each have 1 brown and 1 blue gene.

A family has 4 children. What is P(exactly 3 children with brown eyes)?

NBA Finals

$X \sim \text{Bin}(n, p) \quad p(k) = \binom{n}{k} p^k (1-p)^{n-k}$

The Golden State Warriors are going to play the Toronto Raptors in a 7-game series during the 2019 NBA finals.

- The Warriors have a probability of 58% of winning each game, independently.
- A team wins the series if they win at least 4 games (we play all 7 games).

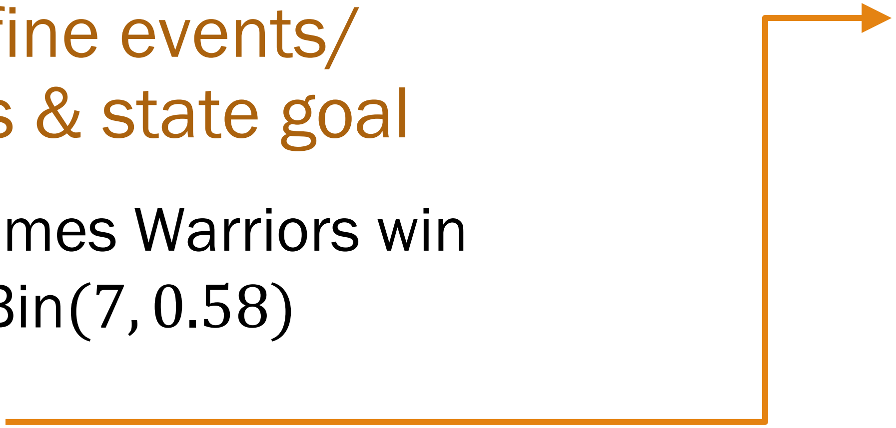
What is P(Warriors winning)?

1. Define events/ RVs & state goal

X: # games Warriors win

$X \sim \text{Bin}(7, 0.58)$

Want:



Desired probability? (select all that apply)

A. $P(X > 4)$

B. $P(X \geq 4)$

C. $P(X > 3)$

D. $1 - P(X \leq 3)$

E. $1 - P(X < 3)$

NBA Finals

$X \sim \text{Bin}(n, p) \quad p(k) = \binom{n}{k} p^k (1-p)^{n-k}$

The Golden State Warriors are going to play the Toronto Raptors in a 7-game series during the 2019 NBA finals.

- The Warriors have a probability of 58% of winning each game, independently.
- A team wins the series if they win at least 4 games (we play all 7 games).

What is P(Warriors winning)?

1. Define events/ RVs & state goal

X: # games Warriors win

$X \sim \text{Bin}(7, 0.58)$

Want:

Desired probability? (select all that apply)

A. $P(X > 4)$
[✓ B.] $P(X \geq 4)$
[✓ C.] $P(X > 3)$
[✓ D.] $1 - P(X \leq 3)$
E. $1 - P(X < 3)$

NBA Finals

$X \sim \text{Bin}(n, p) \quad p(k) = \binom{n}{k} p^k (1-p)^{n-k}$

The Golden State Warriors are going to play the Toronto
game series during the 2019 NBA finals.

- The Warriors have a probability of 58% of winning each game, independently.
- A team wins the series if they win at least 4 games (we play all 7 games).

What is P(Warriors winning)?

1. Define events/ RVs & state goal
2. Solve

X: # games Warriors win
$X \sim \text{Bin}(7, 0.58)$

Want: $P(X \geq 4)$

$$P(X \geq 4) = \sum_{k=4}^{7} P(X = k) = \sum_{k=4}^{7} \binom{7}{k} 0.58^k (0.42)^{7-k}$$

Cool Algebra/Probability Fact: this is identical to the probability of winning if we define winning = first to win 4 games

Genetic inheritance

$X \sim \text{Bin}(n, p) \quad p(k) = \binom{n}{k} p^k (1-p)^{n-k}$

Each person has 2 genes per trait (e.g., eye color).

- Child receives 1 gene (equally likely) from each parent
- Brown is “dominant”, blue is ”recessive”:
    - Child has brown eyes if either (or both) genes are brown
    - Blue eyes only if both genes are blue.
- Parents each have 1 brown and 1 blue gene.

A family has 4 children. What is P(exactly 3 children with brown eyes)?

Subset of ideas:

A. Product of 4 independent events

B. Probability tree

C. Bernoulli, success $p = 3$ children with brown eyes

D. Binomial, $n = 3$ trials, success $p =$ brown-eyed child

E. Binomial, $n = 4$ trials, success $p =$ brown-eyed child

Genetic inheritance

$X \sim \text{Bin}(n, p) \quad p(k) = \binom{n}{k} p^k (1-p)^{n-k}$

Each person has 2 genes per trait (e.g., eye color).

- Child receives 1 gene (equally likely) from each parent
- Brown is “dominant”, blue is ”recessive”:
  - Child has brown eyes if either (or both) genes are brown
  - Blue eyes only if both genes are blue.
- Parents each have 1 brown and 1 blue gene.

A family has 4 children. What is P(exactly 3 children with brown eyes)?

1. Define events/ RVs & goal
2. Identify known probabilities
3. Solve

X: # brown-eyed children,

$X \sim \text{Bin}(4, p)$

p: P(brown−eyed child)

Want: $P(X = 3)$

Interlude for jokes/announcements

Announcements

Midterm Quiz

Time frame:	Mon-Tues, July 20-21 5pm-5pm PT
Covers:	Up to and including Lecture 11

Info and practice:
http://web.stanford.edu/class/archive/cs/cs109/cs109.1208/exams/quizzes.html

Interesting probability news

Polly knows probability: this parrot can predict the chances of something happening

March 3, 2020 2.05pm EST

https://theconversation.com/polly-knows-probability-this-parrot-can-predict-the-chances-of-something-happening-132767

LIVE

Discrete RVs

The hardest part of problem-solving is determining what is a random variable .

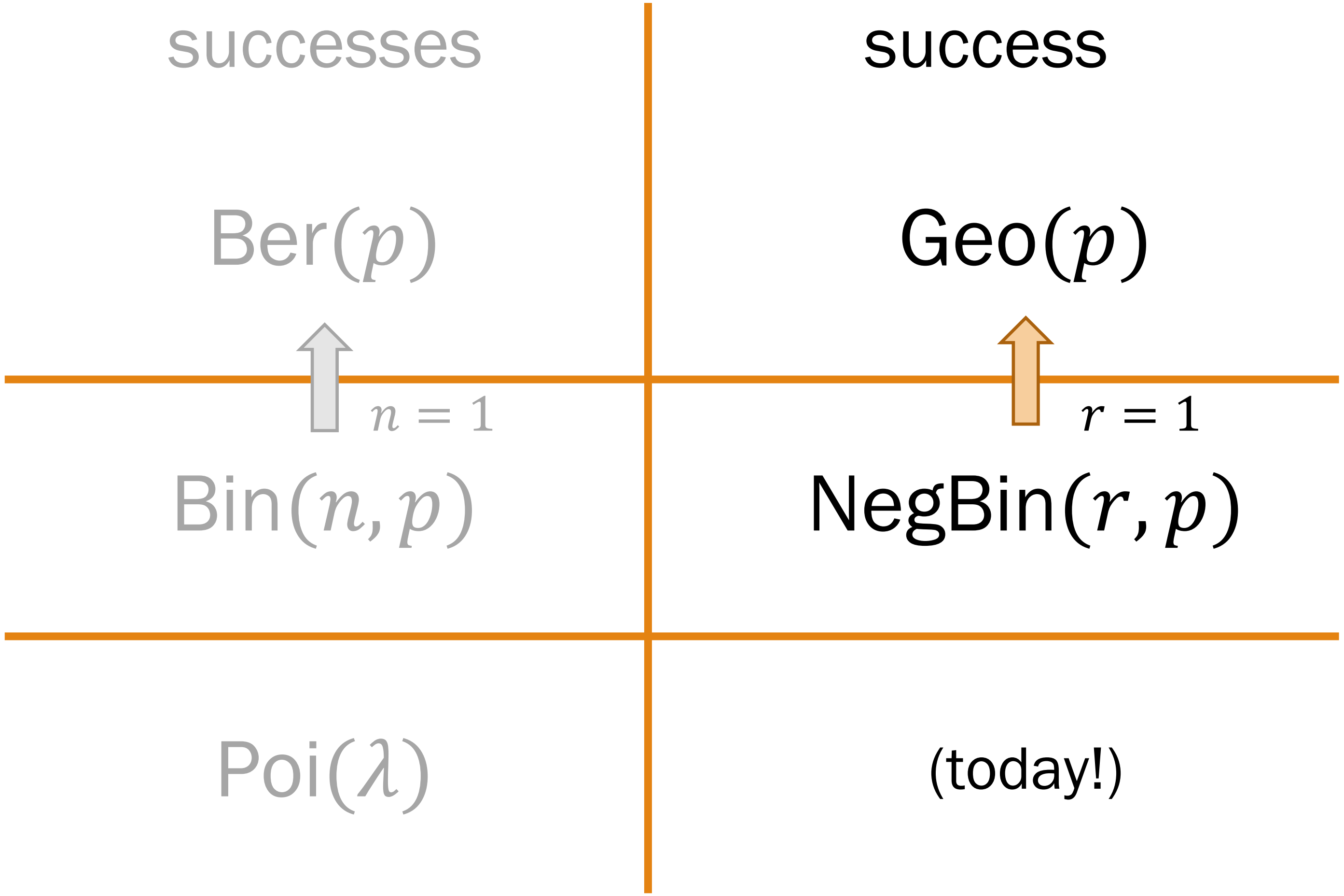
Grid of random variables

Review

	Number of successes	Time until success	
One trial	$\text{Ber}(p)$	$\text{Geo}(p)$	One success
Several trials	$\uparrow\ n = 1$ $\text{Bin}(n, p)$	$\uparrow\ r = 1$ $\text{NegBin}(r, p)$	Several successes
Interval of time	$\text{Poi}(\lambda)$	(today!)	Interval of time to first success

Grid of random variables

Review



	Number of successes	Time until success	
One trial	Ber(p)	Geo(p)	One success
Several trials	Bin(n, p)	NegBin(r, p)	Several successes
Interval of time	Poi(λ)	(today!)	Interval of time to first success

Breakout Rooms

Check out the question on the next slide.
Post any clarifications here!

https://us.edstem.org/courses/667/discussion/84212

Breakout rooms: 5 min. Introduce yourself!

An RV Tour

Choose from:

A. $\text{Ber}(p)$
B. $\text{Bin}(n, p)$
C. $\text{Poi}(\lambda)$
D. $\text{Geo}(p)$
E. $\text{NegBin}(r, p)$

How would you model the following?

1. # of snapchats you receive in a day
2. # of children until the first one with brown eyes (same parents)
3. Whether stock went up or down in a day
4. # of probability problems you try until you get 5 correct (if you are randomly correct)
5. # of years in some decade with at least 6 Atlantic hurricanes

An RV Tour

Choose from:
A. Ber(p)
B. Bin(n, p)
C. Poi(λ)
D. Geo(p)
E. NegBin(r, p)

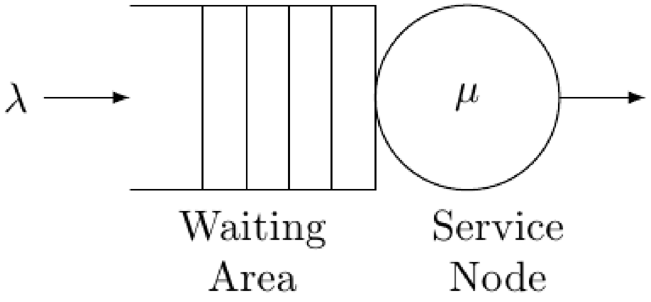
How would you model the following?

1. # of snapchats you receive in a day	C. Poi(λ)
2. # of children until the first one with brown eyes (same parents)	D. Geo(p) or E. NegBin$(1, p)$
3. Whether stock went up or down in a day	A. Ber(p) or B. Bin$(1, p)$
4. # of probability problems you try until you get 5 correct (if you are randomly correct)	E. NegBin$(r = 5, p)$
5. # of years in some decade with at least 6 Atlantic hurricanes	B. Bin$(n = 10, p)$, where $p = P(\geq$ 6 hurricanes in a year) calculated from C. Poi(λ)

CS109 Learning Goal: Use new RVs

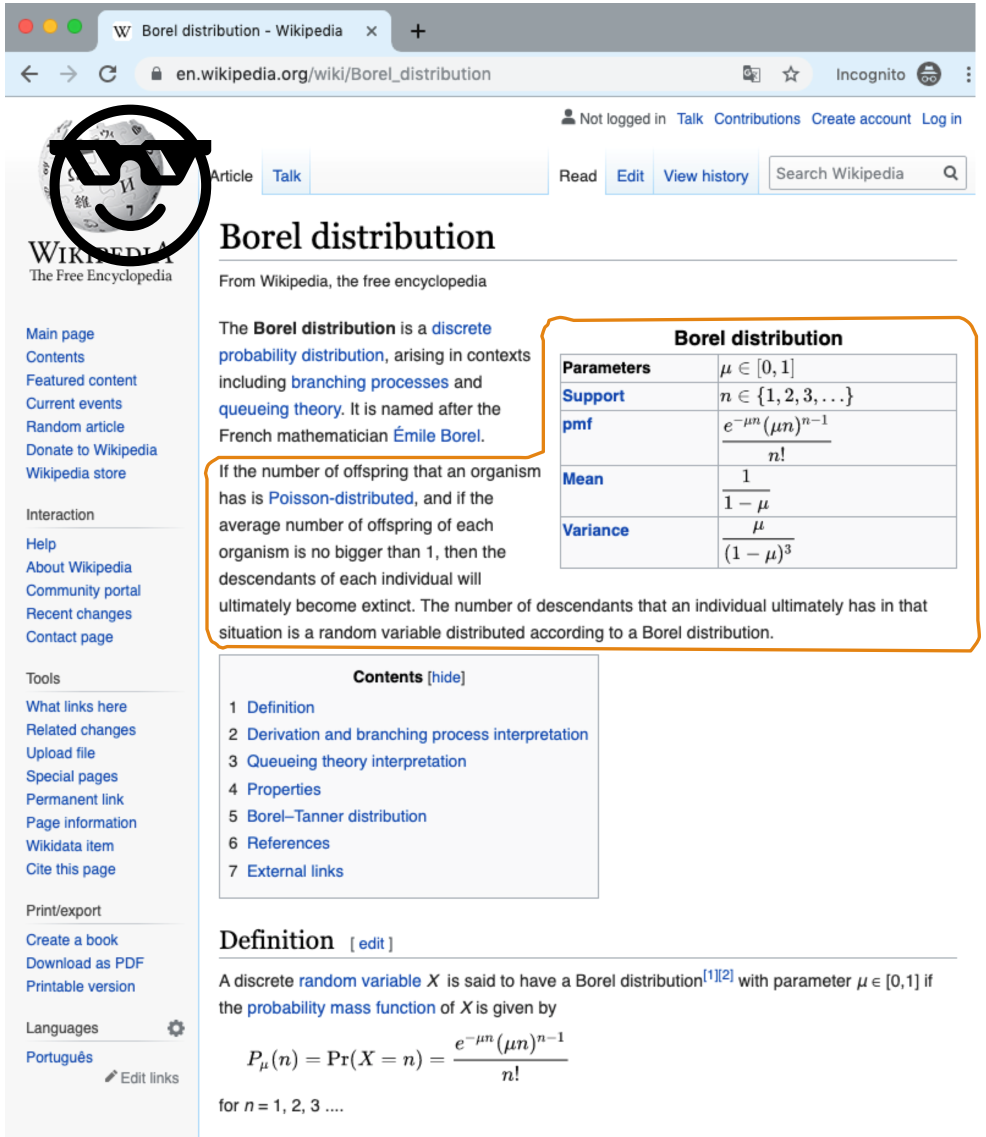
Let's say you are learning about servers/networks.

You read about the M/D/1 queue:



"The service time busy period is distributed as a Borel with parameter $\mu = 0.2$."

Goal: You can recognize terminology and understand experiment setup.



Borel distribution

From Wikipedia, the free encyclopedia

The **Borel distribution** is a discrete probability distribution, arising in contexts including branching processes and queueing theory. It is named after the French mathematician Émile Borel.

If the number of offspring that an organism has is Poisson-distributed, and if the average number of offspring of each organism is no bigger than 1, then the descendants of each individual will ultimately become extinct. The number of descendants that an individual ultimately has in that situation is a random variable distributed according to a Borel distribution.

Borel distribution

Parameters	$\mu \in [0, 1]$
Support	$n \in \{1, 2, 3, \ldots\}$
pmf	$\frac{e^{-\mu n}(\mu n)^{n-1}}{n!}$
Mean	$\frac{1}{1-\mu}$
Variance	$\frac{\mu}{(1-\mu)^3}$

Contents [hide]

Definition [edit]

A discrete random variable X is said to have a Borel distribution[1][2] with parameter $\mu \in [0,1]$ if the probability mass function of X is given by

$$P_\mu(n) = \Pr(X = n) = \frac{e^{-\mu n}(\mu n)^{n-1}}{n!}$$

for $n = 1, 2, 3$ ....

LIVE

Poisson RV

Poisson Random Variable

Review

$X \sim \text{Poi}(\lambda)$

Support: $\{0, 1, 2, \ldots\}$

PMF $$P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}$$

Expectation $E[X] = \lambda$

Variance $\text{Var}(X) = \lambda$

In CS109, a Poisson RV $X \sim \text{Poi}(\lambda)$ most often models

- \# of successes over a fixed interval of time.
  $\lambda = E[X]$, average success/interval
- Approximation of $Y \sim \text{Bin}(n, p)$ where n is large and p is small.
  $\lambda = E[Y] = np$
- Approximation of Binomial even when success in trials are not entirely independent.

} (explored in problem set 3)

Breakout Rooms

The next slide has two questions to go over in groups.

Post any clarifications here!

https://us.edstem.org/courses/667/discussion/84212

Breakout rooms: 5 mins

Web server load

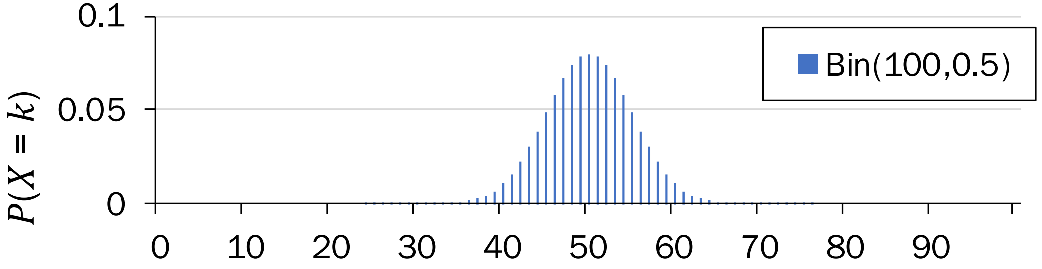
$X \sim \text{Poi}(\lambda)$ $\quad p(k) = e^{-\lambda}\frac{\lambda^k}{k!}$

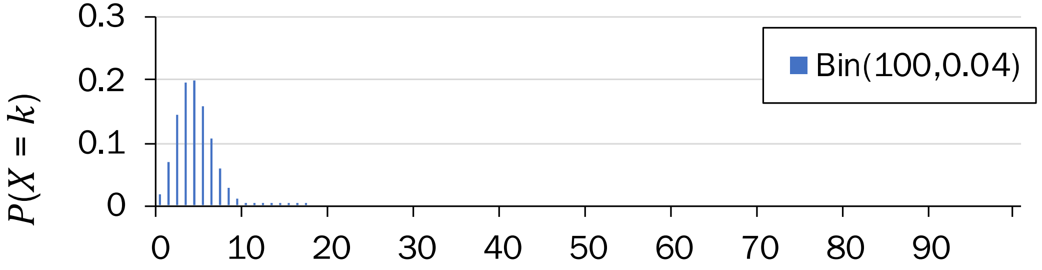
$E[X] = \lambda$

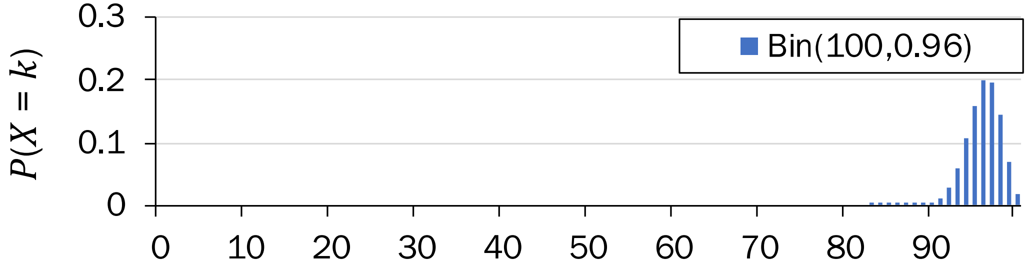
1. Consider requests to a web server in 1 second.
   - In the past, server load averages 2 hits/second.
   - Let X = # hits the server receives in a second.

   What is $P(X < 5)$?

2. Can the following Binomial RVs be approximated with Poisson?







1. Web server load

$X \sim \text{Poi}(\lambda)$ $\quad p(k) = e^{-\lambda}\frac{\lambda^k}{k!}$

$E[X] = \lambda$

Consider requests to a web server in 1 second.

- In the past, server load averages 2 hits/second.
- Let X = # hits the server receives in a second.

What is $P(X < 5)$?

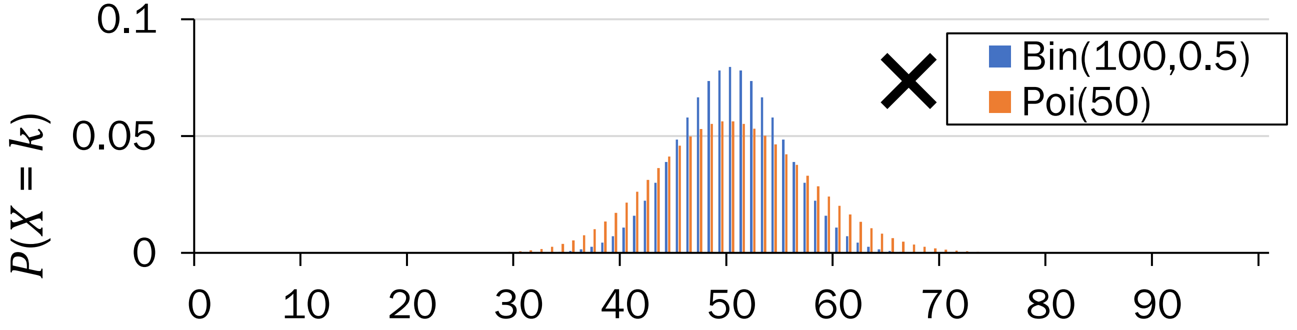
1. Define RVs
2. Solve

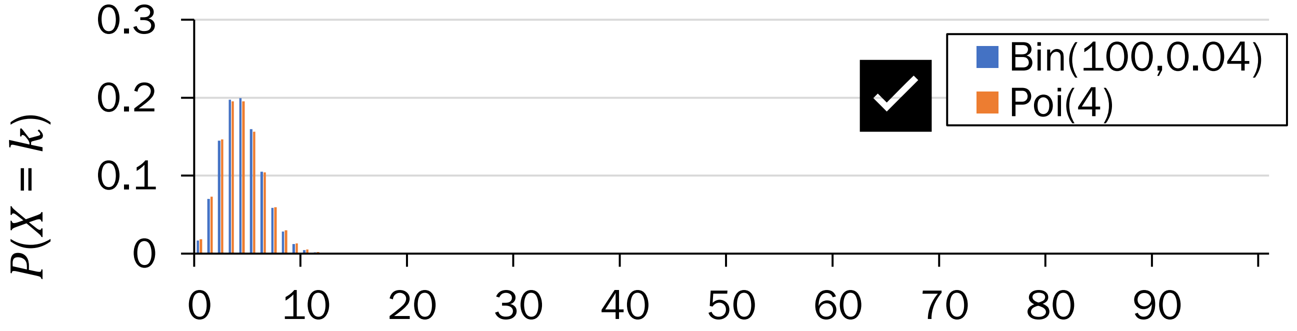
2. Can these Binomial RVs be approximated?

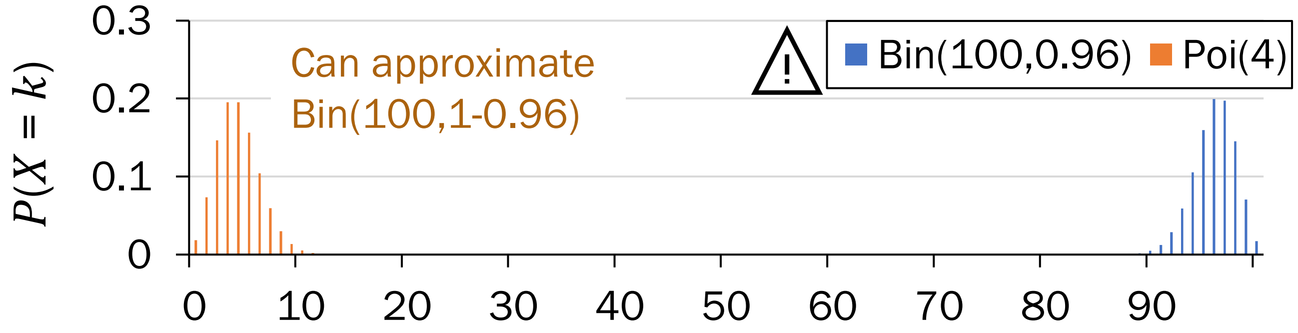
Poisson approximates Binomial when n is large, p is small, and $\lambda = np$ is "moderate."

Different interpretations of "moderate":

- $n > 20$ and $p < 0.05$
- $n > 100$ and $p < 0.1$






Can approximate Bin(100,1-0.96)