Special discrete distributions (cont'd)

– Math 161a, Spring 2019, San Jose State University

Prof. Guangliang Chen

March 12, 2019

Outline

Review and overview

So far we have covered three discrete distributions:

- Bernoulli
- Binomial
- Hypergeometric

They can be understood as counting successes in a fixed number of trials:

- *Flip coins*:
  - Bernoulli (1 trial)
  - Binomial (n trials)
- *Draw balls from an urn with N balls, r of which are red*:
  - Bernoulli (one trial)
  - Binomial (n trials with replacement)
  - Hypergeometric (n trials without replacement)

Today we cover the following three new distributions:

- Geometric
- Negative Binomial
- Poisson

They can also be associated to the two experiments (flip coins, or draw balls), but they are performed indefinitely until a pre-specified number of certain outcome has been obtained.

- **Flip coins for a fixed number of heads**:
  - Geometric (1 head)
  - Negative Binomial ($2+$ heads)
- **Draw balls with replacement from an urn with N balls (r of which are red) until a fixed number of red balls are obtained**:
  - Geometric (1 red ball)
  - Negative Binomial ($2+$ red balls)

Geometric

The following two experiments are identical in nature:

Ex 0.1. Consider the experiment of repeatedly flipping a coin until the first head appears. Let $X = \#$flips needed to get a head.

$$\underbrace{\underset{1}{\underline{\mathrm{T}}}\quad \underset{2}{\underline{\mathrm{T}}}\quad \underline{\mathrm{T}}\quad \cdots\quad \underline{\mathrm{T}}\quad \underline{\mathrm{H}}}_{X = \#\text{flips}}$$

Ex 0.2. Consider the experiment of repeatedly drawing balls, with replacement, from an urn containing N balls (r of which are red), until the first red ball has been selected. Let $X =$ total $\#$ trials needed.

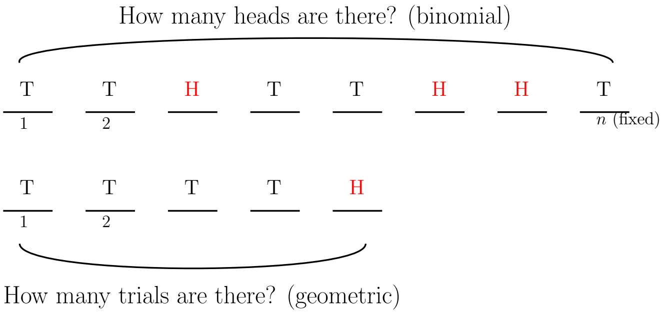
We make the following abstraction:

- The experiment consists of a sequence of **repeated Bernoulli trials** (i.e., each trial has only two outcomes: "success" and "failure");
- The **probability p of getting successes is always fixed**
- The Bernoulli trials are all **independent**;
- The experiment is stopped as soon as **one success has occurred**;
- X denotes the **total number of trials** that have been performed.

In short, X counts the total number of independent trials (with fixed probability of success) that are needed for the first success to occur.

Def 0.1. We say that the previous random variable X has a geometric distribution with parameter p, and write $X \sim \mathrm{Geom}(p)$.

Remark. Binomial (fixed #trials), geometric (fixed #successes):



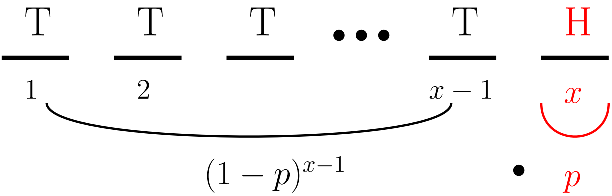
Ex 0.3. The following random variables both have geometric distributions:

- (Repeatedly flip a fair coin until the first head appears) $X =$ total number of flips needed. Geom($\frac{1}{2}$)
- (Repeatedly draw balls with replacement from an urn containing 3 red and 7 blue balls) $X = \#$selections required to obtain a red ball for the first time. Geom($\frac{3}{10}$)

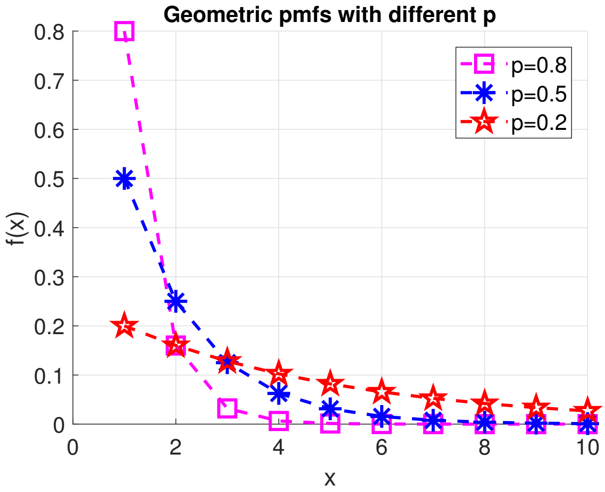
Theorem 0.1. *The pmf of of* $X \sim Geom(p)$ *is*

$$f(x) = (1-p)^{x-1}p, \quad x = 1, 2, \ldots$$

Proof. See the following figure. $\square$



Ex 0.4. Suppose X has a geometric distribution with $p = \frac{1}{2}$. Find $P(X = 4)$ and $F(X \geq 4)$.


Geometric pmfs with different p
p=0.8
p=0.5
p=0.2
f(x)
x
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0
2
4
6
8
10

Infinitely many mathematicians walk into a bar.

The first says, "I'll have a beer."

The second says, "I'll have half a beer."

The third says, "I'll have a quarter of a beer."

The barman pulls out just two beers.

The mathematicians are all like, "That's all you're giving us? How drunk do you expect us to get on that?"

The bartender says, "Come on guys. Know your limits."

An infinite number of mathematicians walk into a bar.

The first one orders a beer.

The second orders half a beer.

The third orders a third of a beer.

The bartender bellows, "Get the hell out of here, are you trying to ruin me?"

Theorem 0.2. *Let $X \sim$ Geom(p). Then*

$$E(X) = \frac{1}{p}, \quad Var(X) = \frac{1-p}{p^2}$$

Proof. We prove only the formula for expected value. By definition,

$$\mathrm{E}(X) = \sum_{x=1}^{\infty} x \cdot (1-p)^{x-1} p = p \sum_{x=1}^{\infty} x(1-p)^{x-1}$$

This series is of the form $\sum_{x=1}^{\infty} x a^{x-1}$, which is equal to $\frac{1}{(1-a)^2}$. Applying this formula gives that

$$\mathrm{E}(X) = p \cdot \frac{1}{(1-(1-p))^2} = \frac{1}{p}.$$

$\square$

Negative Binomial

Briefly speaking, negative binomial distributions are generalizations of geometric distributions by waiting for $r > 1$ successes.

Ex 0.5. Consider the more general experiment of repeatedly tossing a coin until a total of r heads have been obtained. Let $X = \#$ flips needed to get the r heads. Then X follows a *negative binomial* distribution with parameters p and r. We denote it by $X \sim NB(p, r)$.

$$\underbrace{\underset{1}{\underline{\mathrm{T}}}\quad \underset{2}{\underline{\mathrm{T}}}\quad \underline{\mathrm{T}}\quad \underline{\mathrm{H}}\quad \underline{\mathrm{T}}\quad \underline{\mathrm{H}}\quad \underline{\mathrm{T}}\quad \underline{\mathrm{T}}\quad \underline{\mathrm{H}}}_{X = \#\mathrm{flips}}$$

Special discrete distributions (cont'd)

The pmf of X is

$$f_X(x) = \binom{x-1}{r-1} p^r (1-p)^{x-r}, \quad x = \underbrace{r, r+1, r+2, \ldots}_{\text{range}}$$

in which

- $\binom{x-1}{r-1}$: #ways of getting $r-1$ heads in first $x-1$ trials
- p^r: probability of getting r heads (including last head)
- $(1-p)^{x-r}$: probability of getting $x-r$ tails

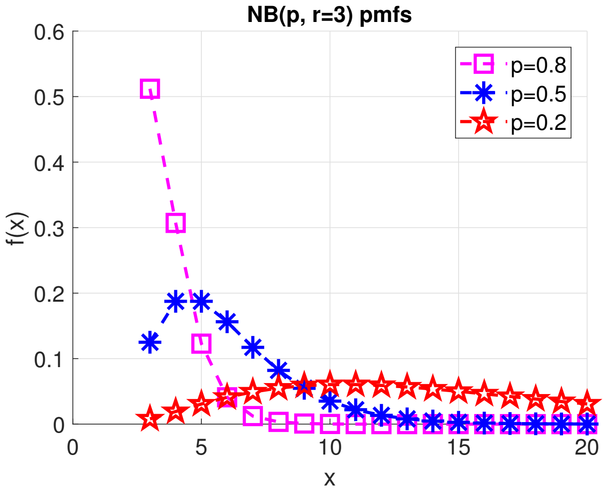
$$\underbrace{\underset{1}{\underline{\mathrm{T}}}\quad \underset{2}{\underline{\mathrm{T}}}\quad \underline{\mathrm{T}}\quad \underline{\color{red}{\mathrm{H}}}\quad \underline{\mathrm{T}}\quad \underline{\color{red}{\mathrm{H}}}\quad \underline{\mathrm{T}}\quad \underset{x-1}{\underline{\mathrm{T}}}}_{x-1\text{ trials, } r-1\text{ heads}}\quad \underbrace{\underset{x}{\underline{\color{red}{\mathrm{H}}}}}_{\text{last head}}$$

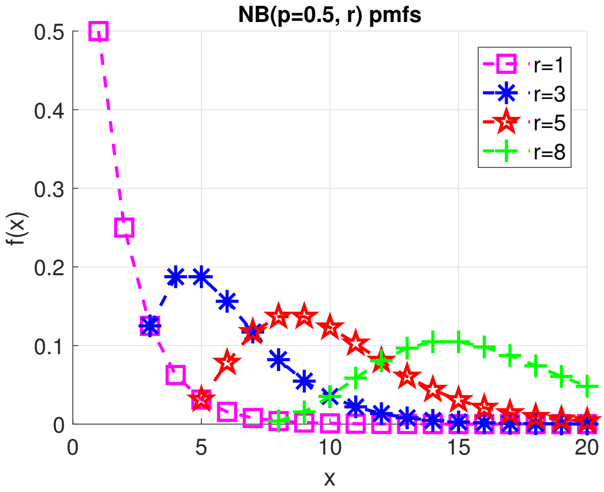
Ex 0.6. Suppose $X \sim \mathrm{NB}(p = \frac{1}{2}, r = 3)$. Find the following probabilities:

- $P(X = 2) =$
- $P(X = 3) =$
- $P(X = 4) =$
- $P(X = 5) =$
- $P(X \geq 5) =$

Answers:

- $P(X = 2) = 0$,
- $P(X = 3) = 0.125$
- $P(X = 4) = 0.1875$
- $P(X = 5) = 0.1875$
- $P(X \geq 5) = 0.6875$.


NB(p, r=3) pmfs
p=0.8
p=0.5
p=0.2
f(x)
x
0
0.1
0.2
0.3
0.4
0.5
0.6
0
5
10
15
20


NB(p=0.5, r) pmfs
0.5
0.4
0.3
0.2
0.1
0
f(x)
0
5
10
15
20
x
r=1
r=3
r=5
r=8

Theorem 0.3. *Let* $X \sim \mathrm{NB}(p, r)$*. Then*

$$E(X) = \frac{r}{p}, \qquad Var(X) = \frac{r(1-p)}{p^2}$$

Proof. Let

$$X = \underbrace{\underset{1}{\underline{\mathrm{T}}}\quad \underset{2}{\underline{\mathrm{T}}}\quad \underline{\mathrm{T}}\quad \underline{\mathrm{H}}}_{X_1} + \underbrace{\underline{\mathrm{T}}\quad \underline{\mathrm{H}}}_{\cdots} + \underbrace{\underline{\mathrm{T}}\quad \underline{\mathrm{T}}\quad \underline{\mathrm{H}}}_{X_r}$$

Then each single X_i has a Geom(p) distribution:

$$\mathrm{E}(X_i) = \frac{1}{p}, \quad \mathrm{Var}(X_i) = \frac{1-p}{p^2}, \quad \text{for all } i = 1, \ldots, r$$

By linearity (and independence),

$$\mathrm{E}(X) = \mathrm{E}(X_1) + \cdots + \mathrm{E}(X_n) = \frac{1}{p} + \cdots + \frac{1}{p} = \frac{r}{p}$$

$$\mathrm{Var}(X) = \mathrm{Var}(X_1) + \cdots + \mathrm{Var}(X_n) = \frac{1-p}{p^2} + \cdots + \frac{1-p}{p^2} = \frac{r(1-p)}{p^2}.$$

Ex 0.7. Find the expected value and variance of the X in the last example.

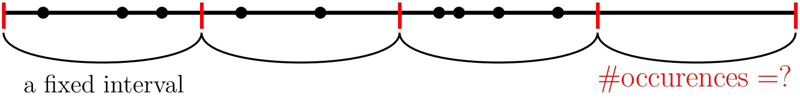
Poisson

Consider the following random variables:

- #hurricanes that hit a region each year
- #earthquakes occurring in certain country in a year
- #car accidents on a certain highway per week
- #phone calls received by a call center per minute
- #customers arriving at a bank counter in every hour
- #typos on each page of certain book

These examples have the following common characteristics:

- X counts the **total number of certain event**
- ... that is **rare** (with a small rate of occurrence λ)
- ... and occurs **independently** of each other
- ... over an **fixed interval of time or space**



Such random variables are modeled by **Poisson distributions**.

Def 0.2 **(X ~Pois(λ).** We say that a discrete random variable has a Poisson distribution with parameter λ, if the pmf has the form

$$f_X(x) = \frac{\lambda^x}{x!}e^{-\lambda}, \quad x = 0, 1, 2, \ldots$$

Remark. The Poisson pmf is based on the following power series:

$$e^\lambda = \sum_{x=0}^{\infty} \frac{\lambda^x}{x!} = \frac{1}{0!} + \frac{\lambda}{1!} + \frac{\lambda^2}{2!} + \frac{\lambda^3}{3!} + \cdots$$

which implies that

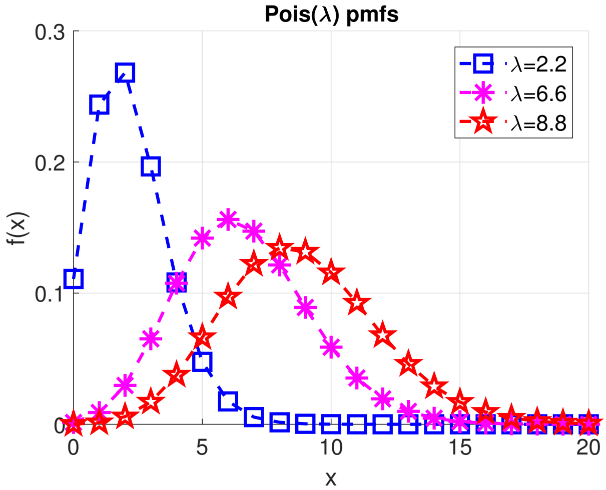
$$1 = \sum_{x=0}^{\infty} \frac{\lambda^x}{x!}e^{-\lambda}$$

Ex 0.8. Suppose, on average, 2.2 hurricanes hit a region each year. Let $X = \#$ hurricanes next year. Find the following probabilities:

- $P(X = 0) =$
- $P(X = 1) =$
- $P(X = 2) =$
- $P(X \geq 2) =$

Answers:

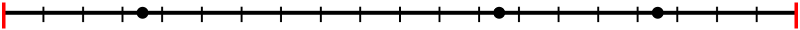
- $P(X = 0) = e^{-2.2} = 0.1108$
- $P(X = 1) = 2.2e^{-2.2} = 0.2438$
- $P(X = 2) = \frac{2.2^2}{2} e^{-2.2} = 0.2681$
- $P(X \geq 2) = 1 - 3.2e^{-2.2} = 0.6454$


Pois(λ) pmfs
0.3
0.2
0.1
0
0
5
10
15
20
f(x)
x
λ=2.2
λ=6.6
λ=8.8

Special discrete distributions (cont'd)

Theorem 0.4. *If n is large and p is small, then $B(n,p) \approx$ Pois($\lambda = np$).*

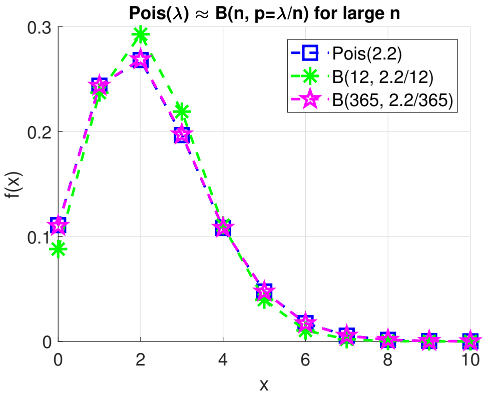
Why this result is true: Consider the hurricane example again (where the number of hurricanes that hit a region in a year has a Poisson distribution with rate λ). We divide a year into n equal subintervals of time (e.g., 12 months, or 365 days).



When n is large, the number of hurricanes that occur in each subinterval is at most 1, thus a Bernoulli random variable with $p = \frac{\lambda}{n}$.

Therefore, the total number of hurricanes during the year is (approximately) binomial, with parameters n, p.

Ex 0.9. We have calculated that for $X \sim \text{Pois}(2.2)$, $P(X = 2) = 0.2681$. In contrast, if $X \sim B(n = 365, p = \frac{2.2}{365})$, then $P(X = 2) = 0.2689$.



Ex 0.10. The first draft of a probability textbook has 600 pages. Assume that the probability of any given page containing at least one typographical error is 0.015 and the numbers of errors on all the pages are mutually independent. Let T be the total number of pages which have at least one typographical error. Find the probability that $T \leq 3$.

Answer: .1328 (exact), or .1318 (approx)

Theorem 0.5. *If $X \sim$ Pois(λ), then $E(X) = \lambda$ and $Var(X) = \lambda$.*

Proof. We only prove the formula for the expected value:

$$\mathrm{E}(X) = \sum_{x=0}^{\infty} x \cdot \frac{\lambda^x}{x!} e^{-\lambda} = \sum_{x=1}^{\infty} \frac{\lambda^x}{(x-1)!} e^{-\lambda} = \lambda \sum_{y=0}^{\infty} \frac{\lambda^y}{y!} e^{-\lambda} = \lambda \cdot 1 = \lambda.$$

Remark. We can also use the binomial approximation to understand why this theorem is true. Treating Pois(λ) as $B(n,p)$ for some large n (and small $p = \frac{\lambda}{n}$), we have

$$\mathrm{E}(X) \approx np = n \cdot \frac{\lambda}{n} = \lambda;$$
$$\mathrm{Var}(X) \approx np(1-p) = n \cdot \frac{\lambda}{n} \cdot (1 - \frac{\lambda}{n}) \approx \lambda.$$

The above approximations become exact when n goes to infinity.