STAT 345/445 Lecture 10

Families of Discrete Distributions – Sections 3.1 and 3.2

Families of Discrete Distributions

We will learn about some of the most commonly used discrete distributions, including their

- $f(x)$ (usually $F(x)$ is not available in closed form)
  - Notation for pmf that emphasizes the parameters:

$$f(x \mid \theta)$$

- parameter space Θ and support $\mathcal{X} = \{x : f(x) > 0\}$
- $E(X)$, $\mathrm{Var}(X)$, $M(t)$
- special features and connections between distributions

See tables p. 621-627 in the Textbook

Discrete Uniform Distributions

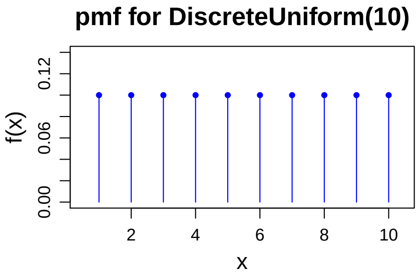
Setting:

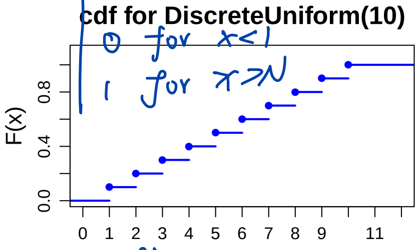
$\mathcal{X} = \{1, 2, 3, \dots N\}$

- Have N possible outcomes
- Each outcome is equally likely

$f(x) = \begin{cases} \frac{1}{N} & \text{for } x = 1, 2, \dots N \end{cases}$

$F(x) = \begin{cases} \frac{\lfloor x \rfloor}{N} & \text{for } 1 \le x \le N, \\ 0 & \text{for } x < 1 \\ 1 & \text{for } x > N \end{cases}$





$\lfloor x \rfloor$ = floor of x

- Determine the pmf, cdf, mean and variance ...

= largest integer that is $\le x$

$$\mathcal{E}(x) = \sum_{x=1}^{N} x \frac{1}{N} = \frac{1}{N} \sum_{x=1}^{N} x = \frac{1}{N} \frac{(N+1)N}{2} = \frac{N+1}{2}$$

$$\mathcal{E}(x^2) = \sum_{x=1}^{N} x^2 \frac{1}{N} = \frac{1}{N} \frac{N(N+1)(2N+1)}{6}$$

$$\Rightarrow V(x) = \frac{(N+1)(2N+1)}{6} - \frac{(N+1)^2}{4}$$

$$= \frac{2(N+1)(2N+1) - 3(N+1)^2}{12} = \frac{(N+1)(N-1)}{12}$$

Useful sums

- Finite sums of powers

$$\sum_{i=1}^{n} i = \frac{n(n+1)}{2} \qquad \sum_{i=1}^{n} i^2 = \frac{n(n+1)(2n+1)}{6}$$

- **Binomial formula:** For all $x, y \in \mathbb{R}$ and $n \in \mathbb{N}$

$$\sum_{i=0}^{n} \binom{n}{i} x^i y^{n-i} = (x+y)^n$$

- **Geometric series:** For $-1 < r < 1$

$$\sum_{k=0}^{\infty} r^k = \frac{1}{1-r}$$

Discrete Uniform Distributions – summary

Probability mass function

$$f(x \mid N) = \frac{1}{N} \qquad \text{for } x \in \{1, 2, 3, \ldots, N\}$$

- Parameter space: $N \in \{1, 2, 3, \ldots\}$

Mean and Variance

$$E(X) = \frac{N+1}{2} \qquad \mathrm{Var}(X) = \frac{N^2-1}{12}$$

Moment generating function

$$M_X(t) = \sum_{x=1}^{N} e^{tx} \frac{1}{N}$$

NO simplification available.

Bernoulli Distributions - $\text{Bernoulli}(p)$

- Two possible outcomes:

$$\begin{aligned} \text{success:} \quad & X = 1 \\ \text{failure:} \quad & X = 0 \end{aligned}$$

- Think: Games (win or lose), coin toss, etc
- Probability of success: $p = P(X = 1)$
- pmf:

$$f(x) = \begin{cases} 1-p & \text{if } x = 0 \\ p & \text{if } x = 1 \\ 0 & \text{otherwise} \end{cases} = p^x(1-p)^{1-x} \quad \text{for } x = 0, 1$$

Parameter space: $p \in [0, 1]$, support: $\mathcal{X} = \{0, 1\}$

Bernoulli Distributions - $\mathrm{Bernoulli}(p)$

- cdf:

"Binomial (n=1, p)"

$$F(x) = \begin{cases} 0 & x < 0 \\ 1 - p & 0 \le x < 1 \\ 1 & x \ge 1 \end{cases}$$

- mean and variance:

$$E(X) = \sum_x xf(x) = 0 * (1-p) + 1 * p = p$$

$$\mathrm{Var}(X) = \sum_x x^2 f(x) - (E(X))^2 = 0^2 * (1-p) + 1^2 * p - p^2$$

$$= p - p^2 = p(1-p)$$

- mgf:

$$M(t) = \sum_x e^{tx} f(x) = e^{t*0}(1-p) + e^{t*1} p = 1 - p + pe^t$$

Binomial distributions

- **Bernoulli trial:** n independent Bernoulli random variables
  - $X_i =$ outcome of trial i (0 or 1), $i = 1, 2, \ldots, n$
  - Same probability of success (p) for all i

Support for Y:

$\mathcal{Y} = \{0, 1, 2, \ldots, n\}$

- $Y =$ total number of successes in n trials

$$Y = X_1 + X_2 + \cdots + X_n$$

- What is $f(y) = P(Y = y)$?
  - We haven't yet covered distributions of functions of more than one random variable (Chapter 4) but we can approach this differently ...

Binomial distributions

- What is $P(Y = y)$?
  - $Y = y$ means we had y successes and $n - y$ failures
  - By independence, the probability of any one such outcome is

multiply prob

$$p^y(1-p)^{n-y}$$

  - Number of ways we could get y successes in n trials: $\binom{n}{y}$
  - These are disjoint events so we add up the probabilities and get

$$f(y) = \binom{n}{y} p^y (1-p)^{n-y} \quad \text{for } y = 0, 1, \ldots, n$$

to prove $\sum_{y=0}^{n} f(y) = 1$?
$= (p+(1-p))^n = 1$ by the Binomial formula

- Are we sure that this $f(y)$ is a pdf?
  - Parameter space? Support?

Mean, Variance and mgf for the Binomial Distribution

- Finding $E(X)$ and $E(X^2)$ directly involves evaluating the sums

$$E(X) = \sum_{y=0}^{n} y \frac{n!}{y!(n-y)!} p^y (1-p)^{n-y}$$

$$E(X^2) = \text{ and } \sum_{y=0}^{n} y^2 \frac{n!}{y!(n-y)!} p^y (1-p)^{n-y}$$

not impossible but a bit tedious. The mgf route is a bit easier in this case... mgf: $M(t) = E(e^{tY}) = \sum_{y=0}^{n} e^{ty} \binom{n}{y} p^y (1-p)^{n-y}$

$$= \sum_{y=0}^{n} \binom{n}{y} (e^t p)^y (1-p)^{n-y} = (e^t p + 1 - p)^n$$

Compare to

$M(t) = e^t p + 1 - p$, for Bernoulli (p)

$$E(y) = \frac{d}{dt} M(t) \Big|_{t=0} = n(e^t p + 1 - p)^{n-1} e^t p \Big|_{t=0} = np$$

$$E(y^2) = \frac{d^2}{dt^2} M(t) \Big|_{t=0} = n(n-1)(e^t p + 1 - p)^{n-2} e^t p \, e^t p + n(e^t p + 1 - p)^{n-1} e^t p \Big|_{t=0}$$

$$= n(n-1)p^2 + np$$

$$Var(y) = E(y^2) - E(y)^2 = np(1-p)$$

Binomial Distributions – $\mathrm{Binomial}(n, p)$

Probability mass function

$$f(x \mid n, p) = \binom{n}{x} p^x (1-p)^{n-x} \qquad \text{for } x \in \{0, 1, 2, \ldots, n\}$$

- Parameter space: $0 \leq p \leq 1$, $n \in \{1, 2, 3, \ldots\}$
- Special case: **Bernoulli distribution** if $n = 1$

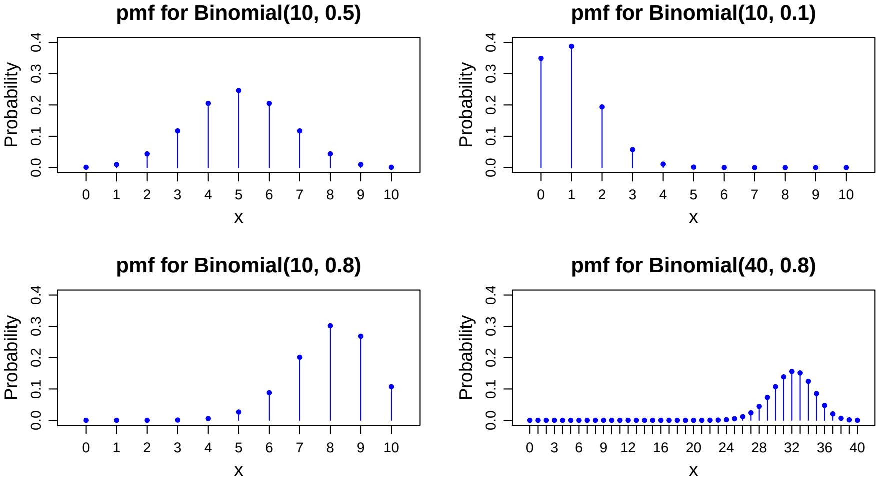
Mean and Variance

$$E(X) = np \qquad \mathrm{Var}(X) = np(1-p)$$

Moment generating function

$$M_X(t) = \left(pe^t + 1 - p\right)^n$$

Binomial pmfs



Hypergeometric distributions

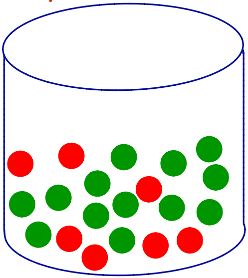
$N = 20$ balls

$M = 7$ red

- Sampling from a finite population, *without replacement*
- Have a population of N items, M of which are of the type of interest
  - Think: N balls in an urn, M of which are red
- Randomly pick K items, without replacement
  - and unordered
- Random variable of interest:

  $X =$ number of items of type 1 in the sample

  red balls

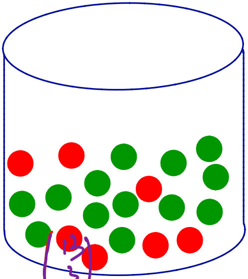


Hypergeometric distribution - Example

- $N = 20$ balls in an urn, $M = 7$ are red
- Randomly pick $K = 3$ balls, without replacement
- Random variable of interest:

  $X =$ number of red balls in the sample

- What is $f(x)$?



$$f(x) = \frac{\binom{M}{x}\binom{N-M}{K-x}}{\binom{N}{K}}$$

$x \in \{0, 1, 2 \cdots K\}$

$X = 0: \ f(x=0) = P(X=0) = \frac{\binom{13}{3}}{\binom{20}{3}}$

$X = 1, \ f(x=1) = P(x=1) = \frac{\binom{7}{1}\binom{13}{2}}{\binom{20}{3}}$

$X = 2, \ f(x=2) = P(x=2) = \frac{\binom{7}{2}\binom{13}{1}}{\binom{20}{3}}$

$X = 3, \ f(x=3) = P(x=3) = \frac{\binom{7}{3}}{\binom{20}{3}}$

Hypergeometric distribution - Example

- Opinion poll
  - $N = 5150$ CWRU undergraduate students (Fall 2018)
  - $M = 2550$ Engineering students
  - $K = 300$ students randomly selected for a survey
- $X =$ number of engineering students selected for the survey
- Probability that $X = x$:

$$f(x) = \frac{\binom{2550}{x}\binom{5150-2550}{300-x}}{\binom{5150}{300}} \qquad x = 0, 1, 2, \ldots, 300$$

Or: $X \sim \mathrm{HyperGeometric}(N = 5150, M = 2550, K = 300)$

Hypergeometric Distributions – $\mathrm{HyperGeo}(N, M, K)$

Probability mass function

$$f(x \mid M, N, K) = \frac{\binom{M}{x}\binom{N-M}{K-x}}{\binom{N}{K}} \qquad \text{for } x \in \{0, 1, 2, \ldots, K\}$$

- Parameter space: $N, M, K \in \{1, 2, 3, \ldots\}$, $M \leq N$, $K \leq N$
- Implied: $M - (N - K) \leq x \leq M$
- Showing $\sum_x f(x \mid M, N, K) = 1$ is not trivial

Mean and Variance

$$E(X) = \frac{KM}{N} \qquad \mathrm{Var}(X) = \frac{KM}{N}\frac{(N-M)}{N}\frac{(N-K)}{N-1}$$

- mgf: no simplification available

$$E(X) = \sum_{x=0}^{K} x \frac{\binom{M}{x}\binom{N-M}{K-x}}{\binom{N}{K}} = \sum_{x=1}^{K} \frac{\frac{M!}{(x-1)!(M-x)!}\binom{N-M}{K-x}}{\binom{N}{K}}$$

$$= \sum_{t=0}^{K-1} \frac{\frac{M(M-1)!}{t!(M-t-1)!}\binom{N-M}{K-t-1}}{\binom{N}{K}}$$

$t = x-1$

$(x = t+1)$

$$= M \sum_{t=0}^{K-1} \frac{\binom{M-1}{t}\binom{N-1-(M-1)}{K-1-t}}{\frac{N}{K}\binom{N-1}{K-1}}$$

$$= \frac{KM}{N}$$

$$\binom{N}{K} = \frac{N!}{K!(N-K)!}$$

$$= \frac{N(N-1)!}{K(K-1)!(N-1-(K-1))!}$$

$$= \frac{N}{K}\binom{N-1}{K-1}$$

Hypergeometric

$$E(X^2) = \sum_{x=0}^{k} x^2 \frac{M!}{x!(M-x)!} \binom{N-M}{k-x} \frac{1}{\binom{N}{k}} = \text{etc.}$$

Another way:

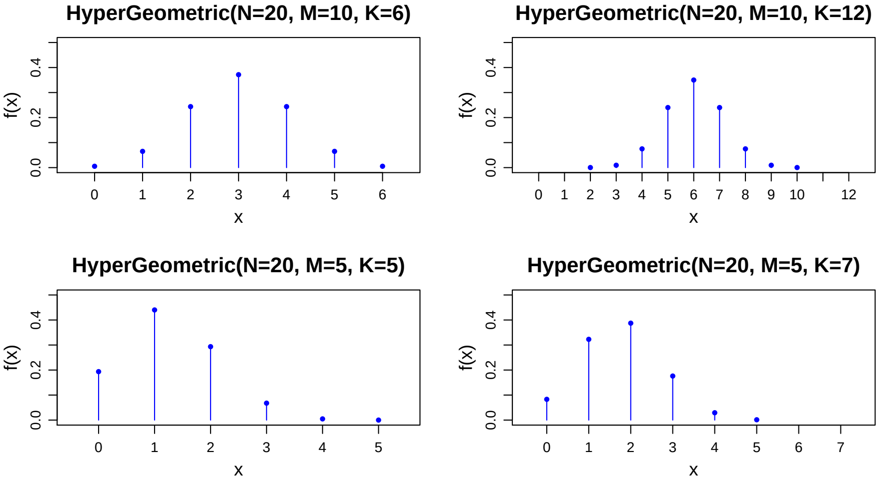
$$E(X(X-1)) = \sum_{x=0}^{k} x(x-1) \frac{M!}{x!(M-x)!} \binom{N-M}{k-x} \frac{1}{\binom{N}{k}}$$

$$= \sum_{x=2}^{k} \frac{M!}{(x-2)!(M-x)!} \cdots \text{etc.}$$

and Note that $E(X(X-1)) = E(X^2-X) = E(X^2) - E(X)$

$\therefore \quad E(X^2) = E(X) + E(X^2-X) = E(X) + E(X(X-1))$

Hypergeometric pmfs



Hypergeometric distributions - Example

- Opinion poll
  - $N = 5150$ CWRU students, $M = 2550$ Engineering students
  - $K = 300$ students randomly selected for a survey
- R.v. $X =$ number of engineering students selected for the survey
- If 300 students are randomly selected from an *infinite* population then $X \sim \mathrm{Binomial}(300, p)$ where p is the probability that a student is in engineering.
- Sampling from a finite population is trickier than an infinite population
- Real world survey, e.g. a random sample of US adults, usually assume that the population is infinite

Comparing Hypergeometric and Binomial

- Mean and variance of $X \sim \mathrm{HyperGeo}(N, M, K)$:

$$E(X) = \frac{KM}{N} \qquad \mathrm{Var}(X) = \frac{KM}{N}\underbrace{\frac{(N-M)}{N}}_{\left(1-\frac{M}{N}\right)}\frac{(N-K)}{N-1}$$

  - Compare to Binomial with $n = K$ and $p = \frac{M}{N}$:

$$E(X) = \frac{KM}{N} = np$$

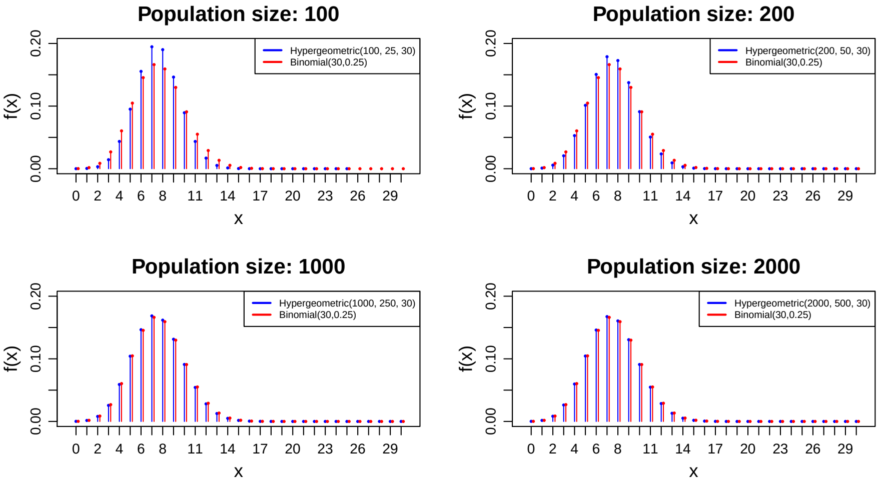
Var of Binomial

$$\mathrm{Var}(X) = \underbrace{np(1-p)}\frac{(N-K)}{N-1} \leq np(1-p)$$

  - Same mean but smaller variance than a Binomial random variable
  - Variance similar if N is big and $K << N$

Hypergeometric pmfs

Hypergeometric with $M/N = 0.25$, $K = 30$
Binomial with $n = 30$, $p = 0.25$



Poisson distributions

$$= \frac{\lambda e^{-\lambda} \lambda^{x-1}}{x(x-1)!} = \frac{\lambda}{x} f(x-1)$$

$$= \frac{\lambda}{x} P(X = x-1)$$

Probability mass function

$$f(x \mid \lambda) = \frac{e^{-\lambda}\lambda^x}{x!} \qquad x = 0, 1, 2, 3, \ldots$$

- Parameter space: $\lambda > 0$

Have show all of this before

Mean and Variance

$$E(X) = \lambda \qquad \mathrm{Var}(X) = \lambda$$

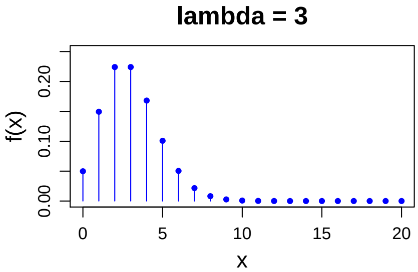
Moment generating function

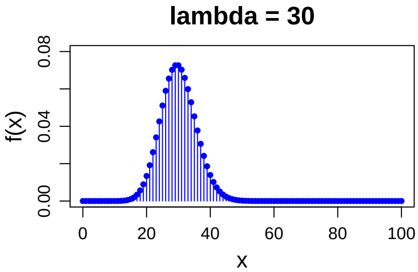
$$M_X(t) = \exp\left(\lambda\left(e^t - 1\right)\right)$$

Poisson distributions

- Useful to model counts
  - E.g. radiation, calls per minute, etc.
- Recursive property: If $X \sim \mathrm{Poisson}(\lambda)$ then

$$P(X = x) = \frac{\lambda}{x} P(X = x - 1)$$





Negative Binomial and Geometric distributions

- Binomial: n Bernoulli trials
- Geometric: Bernoulli trials until we get a success
- Negative binomial: Bernoulli trials until we get r successes

Examples:

- Randomly trying keys to open door, but don't keep track of which key has already been checked
- Randomly select fish from a catch until we have r juveniles.

Geometric distributions

- X = number of trials until (and including) we get the first success
- Let p be the probability of success for each trial
- What is $f(x)$? Let s = *success*, f = *failure*

Outcome	x	$P(X = x)$
s	1	p
f, s	2	$(1-p)p$
f, f, s	3	$(1-p)^2 p$
f, f, f, s	4	$(1-p)^3 p$
f, f, f, f, s	5	$(1-p)^4 p$
⋮	⋮	⋮
f $\times(x-1)$, s	x	$(1-p)^{x-1} p$

$f(x) = (1-p)^{x-1} p$, for $x = 1, 2, 3, \dots$

$$\sum_{x=1}^{\infty} (1-p)^{x-1} p = p \sum_{x=1}^{\infty} (1-p)^{x-1}$$

$$= p \sum_{x=0}^{\infty} (1-p)^{t} = p \frac{1}{1-(1-p)}$$

$$= 1$$

Cdf:

$$F(x) = \sum_{v=1}^{x} (1-p)^{v-1} p = p \sum_{t=0}^{x-1} (1-p)^{t} = p \frac{1-(1-p)^x}{1-(1-p)}$$

$$= \frac{p}{p}\left(1-(1-p)^x\right) = 1-(1-p)^x$$

$$M(t) = \sum_{x=1}^{\infty} e^{xt} p(1-p)^{x-1} = p \sum_{k=0}^{\infty} (e^t)^{k+1} (1-p)^k$$

$(k = x-1)$

$$= pe^t \sum_{k=0}^{\infty} (e^t(1-p))^k = pe^t \frac{1}{1-e^t(1-p)}$$

if $e^t(1-p) < 1 \Rightarrow t + \log(1-p) < 0 \Rightarrow t < -\log(1-p)$

In lecture 9 we found

$$E(Y) = \frac{1-\theta}{\theta}$$

$f(y) = \theta(1-\theta)^y \quad y = 0, 1, 2, \ldots$

Y = # failures before success

or $X = Y + 1 \quad E(X) = E(Y) + 1 = \frac{1}{\theta}$

Useful sums

- **Geometric series:** For $-1 < r < 1$

$$\sum_{k=0}^{\infty} r^k = \frac{1}{1-r}$$

- **Geometric sum:** For $r \neq 1$

$$\sum_{k=0}^{n} r^k = \frac{1-r^{n+1}}{1-r}$$

Geometric distributions

X = number of trials until (and including) we get the first success

Probability mass function

$$f(x \mid p) = p(1-p)^{x-1} \qquad x = 1, 2, 3, \ldots$$

- Parameter space: $0 \le p \le 1$ for
- cdf: $F(x) = 1 - (1-p)^x \longleftarrow X = 1, 2, 3 \ldots$

Mean and Variance

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

Moment generating function

$$M_X(t) = \frac{pe^t}{1-(1-p)e^t} \qquad t < -log(1-p)$$

Geometric distribution - memoryless

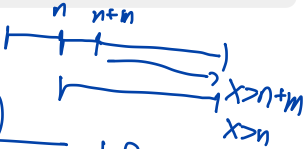
- Geometric distribution has a memoryless property

Theorem

If $X \sim \mathrm{Geometric}(p)$ then

$$P(X > n + m \mid X > n) = P(X > m)$$

Geometric Memoryless Property:



$$P(X>n+m \mid X>n) = \frac{P(X>n+m,\ X>n)}{P(X>n)}$$

definition of condi prob.

$$= \frac{P(X>n+m)}{P(X>n)} = \frac{1-P(X\le n+m)}{1-P(X\le n)}$$

$$= \frac{1-F(n+m)}{1-F(n)} = \frac{1-(1-(1-p)^{n+m})}{1-(1-(1-p)^{n})} = \frac{(1-p)^{n+m}}{(1-p)^{n}}$$

$$= (1-p)^{m} = 1-F(m) = P(X>m)$$

$X>n$: more than n trials

knowing how many trials, we already had does not change the prob that there will be more than m trial until success.

 ....

Negative Binomial

p = prob of success

. . - .. s ← r-th success

failures and r-1 success

Bernoulli trials until we have r successes

- X = number of failures until we get r successes
- Let p be the probability of success for each trial
- What is $f(x)$? Let $s = \textit{success}$, $f = \textit{failure}$ and $r = 3$

Outcome	x	$P(X = x)$
sss	0	p^3
fsss or sfss or ssfs	1	$p^3(1-p)\binom{3}{1}$
2 f and 2 s in the first 4 trials, then s	2	$p^3(1-p)^2\binom{4}{2}$
3 f and 2 s in the first 5 trials, then s	3	$p^3(1-p)^3\binom{5}{3}$
4 f and 2 s in the first 6 trials, then s	4	$p^3(1-p)^4\binom{6}{4}$
⋮	⋮	
x f and 2 s in the first $x+2$ trials, then s	x	$p^r(1-p)^x\binom{x+r-1}{x}$

Negative Binomial

$X =$ number of failures before the r success

Probability mass function

$$f(x \mid r, p) = \binom{r+x-1}{x} p^r (1-p)^x \qquad x = 0, 1, 2, 3, \ldots$$

- Parameter space: $0 \leq p \leq 1$, $r \in \mathbb{N}$

Mean and Variance

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

Moment generating function

$$M_X(t) = \left(\frac{p}{1-(1-p)e^t} \right)^r \qquad t < -log(1-p)$$

Show $\sum_{x=0}^{\infty} f(x) = 1$ is tricky.

The known sum (or integral) trick.

We know that for any $p \in [0,1]$ and $r = 1, 2, 3 \ldots$

$$\sum_{x=0}^{\infty} \binom{x+r-1}{x} p^r (1-p)^x = 1$$

$$E(X) = \sum_{x=0}^{\infty} x \frac{(x+r-1)!}{x!\,(r-1)!} p^r (1-p)^x$$

$$= \sum_{x=1}^{\infty} \frac{(x+r-1)!}{(x-1)!\,(r-1)!} p^r (1-p)^x \qquad (*)$$

Set $t = x-1 \Rightarrow x = t+1$, $x = 1, 2, 3, \ldots$

$, \ t = 0, 1, 2, 3, \ldots$

$$(*) = \sum_{t=0}^{\infty} \frac{(t+r)!}{t!\,(r-1)!} p^r (1-p)^{t+1}$$

$\tilde{r} = r+1$

$$= \sum_{t=0}^{\infty} \frac{(t+r-1+1)!}{t!\,(r-1)!} p^r (1-p)^{t+1}$$

$$= \sum_{t=0}^{\infty} \frac{(t+\tilde{r}-1)!}{t!\,(\tilde{r}-2)!} p^{\tilde{r}-1} (1-p)^{t+1}$$

$$= \sum_{t=0}^{\infty} \frac{(t+\tilde{r}-1)!\,(\tilde{r}-1)}{t!\,(\tilde{r}-1)!} p^{\tilde{r}-1} (1-p)^{t+1}$$

$$= \sum_{t=0}^{\infty} \frac{(t+\tilde{r}-1)!\,(\tilde{r}-1)}{t!\,(\tilde{r}-1)!\,p} p^{\tilde{r}}(1-p)^{t+1}$$

$$= (1-p)\frac{(\tilde{r}-1)}{p} \underbrace{\sum_{t=0}^{\infty} \frac{(t+\tilde{r}-1)!}{t!\,(\tilde{r}-1)!} p^{\tilde{r}}(1-p)^{t}}_{=1}$$

$$= (1-p)\frac{\tilde{r}-1}{p} = \frac{r}{p}(1-p) = E(X)$$

Negative Binomial

- Negative Binomial is often used to model counts as an alternative to Poisson
  - Can be helpful for *over-dispersed* data

- Negative binomial can be written as a *mixture distribution* of a Poisson and a Gamma:

$$Y \mid \lambda \sim \text{Poisson}(\lambda) \qquad \text{and} \qquad \lambda \sim \text{Gamma}(\alpha, \beta)$$
$$\Rightarrow Y \sim \text{NegativeBinomial}(\cdot, \cdot)$$

more later...

Negative Binomial pmfs