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:

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

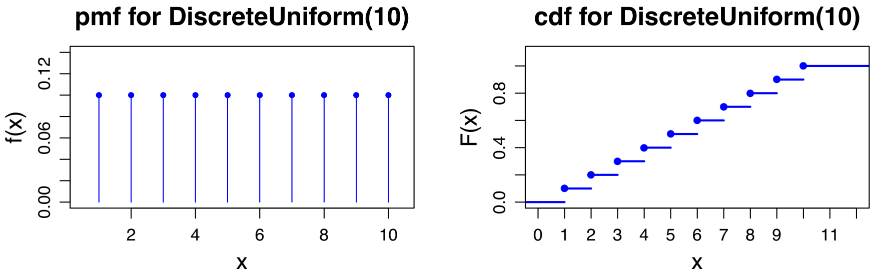
- parameter space Θ and support $\mathcal{X} = \{x : f(x) > 0\}$
- $E(X)$, $\mathrm{Var}(X)$, $M(t)$
- special features and connections between distributions

See tables p. 621-627 in the Textbook

Discrete Uniform Distributions

Setting:

- Have N possible outcomes
- Each outcome is equally likely



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

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 - $\mathrm{Bernoulli}(p)$

- 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:

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

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

Bernoulli Distributions - Bernoulli(p)

$=$ Binomial ($n=1$, p)

- cdf:

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

- 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

Y = # successes in n trials, Support: $\{0, 1, \dots, n\}$

- 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

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

- Are we sure that this $f(y)$ is a ~~pdf~~ pmf?
  - 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...

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

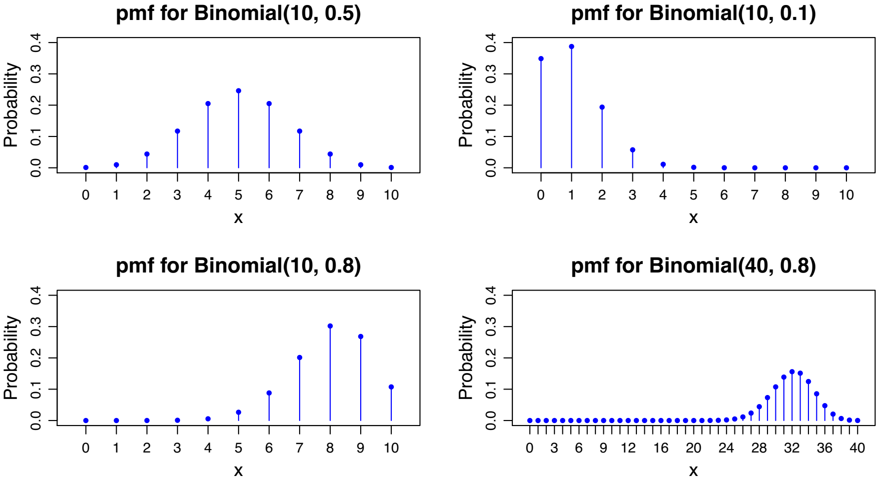
Mean and Variance

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

Moment generating function

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

Binomial pmfs



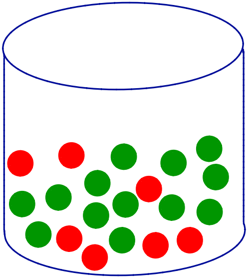
Hypergeometric distributions

$N = 20$ 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

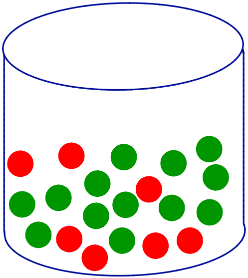


Hypergeometric distribution - Example

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

$$X = \text{ number of red balls in the sample}$$

- What is $f(x)$?



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

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

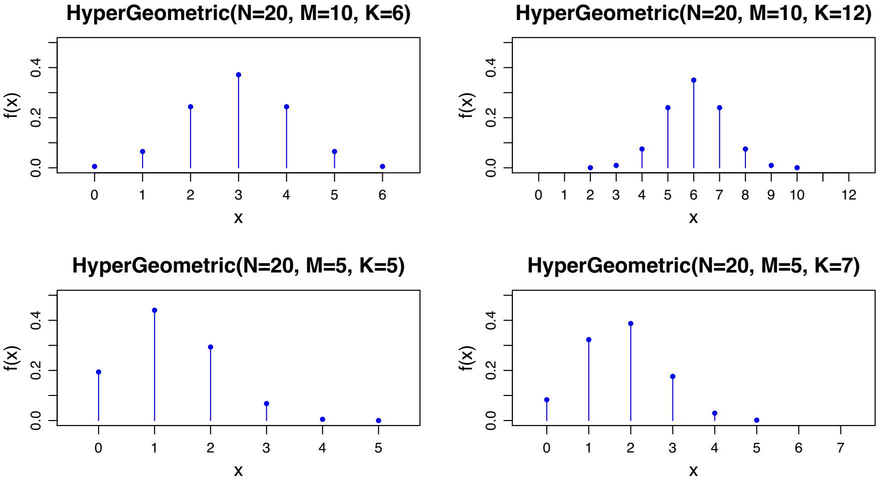
- Parameter space: $N, M, K \in \{1, 2, 3, \ldots\}$, $M \leq N$, $K \leq N$
- Implied: $M - (N - K) \leq x \leq M$
- Showing $\sum_x f(x \mid M, N, K) = 1$ is not trivial

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

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 $p = \frac{M}{N}$:

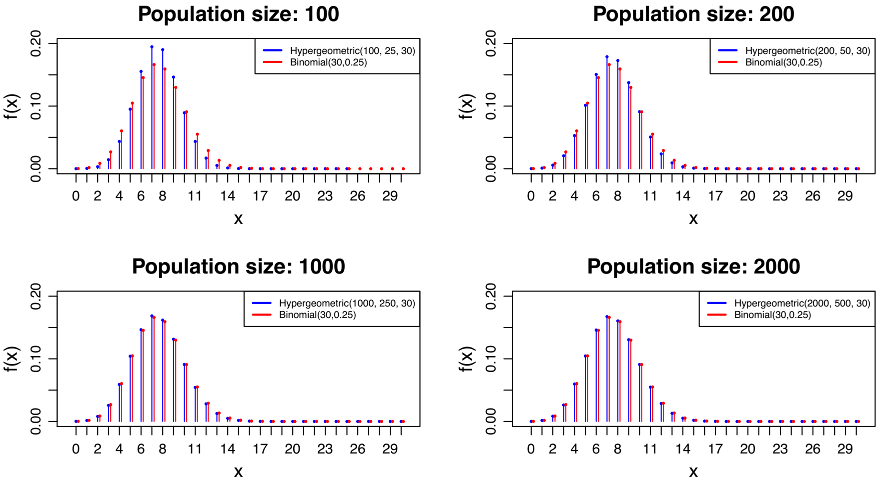
$$E(X) = \frac{KM}{N} = np$$

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

  - 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

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

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

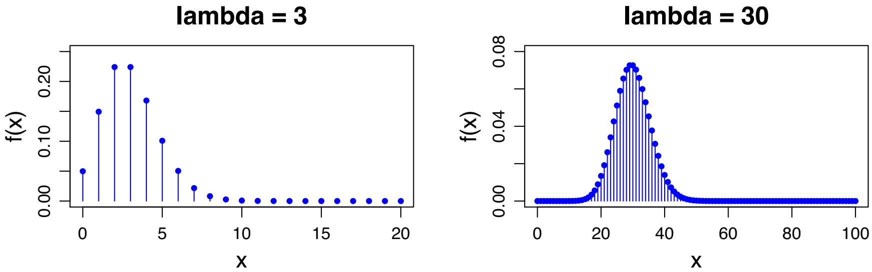
Moment generating function

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

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 $s = \textit{success}$, $f = \textit{failure}$

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	
⋮	⋮	⋮
f $\times(x-1)$, s	x	$(1-p)^{x-1} p$

pmf?

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: $0 \leq p \leq 1$
- cdf: $F(x) = 1 - (1-p)^x$ ← for $x = 1, 2, 3, \ldots$

Mean and Variance

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

Moment generating function

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

In Lecture 9 we found

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

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

Y : # failures before success

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

Geometric distribution - memoryless

- Geometric distribution has a memoryless property

Theorem

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

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

Negative Binomial

p = prob. of success

- - - - - - s ← r^{th} success

$\underbrace{\quad\quad\quad}_{\text{\# failures and } r-1 \text{ successes}}$

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 = *success*, f = *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	$p^3(1-p)^x\binom{x+r-1}{x}$

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

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

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:

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

more later...

Negative Binomial pmfs