Math 3339

Section 27204
MWF 10-11:00am AAAud 2

Bekki George
bekki@math.uh.edu
639 PGH

Office Hours:
M & Th noon – 1:00 pm & T 1:00 – 2:00 pm
and by appointment

Popper 06

1. A furniture store is having a sale on sofas and you're going to buy one. The advertisers know that buyers get to the store and that 1 out of 4 buyers change to a more expensive sofa than the one in the sale advertisement. Let X be the cost of the sofa. What is the average cost of a sofa if the advertised sofa is $200 and the more expensive sofa is $375?

a) 243.42

b) 331.25

c) 243.75

d) 235.00

e) 287.50

cost X	\$200	\$375
$P(X=x)=f(x)$		

$$E[X] = \sum_i x_i f(x_i)$$

A.3 #1 Toss a coin until H or 6T

# of tosses X	1	2	3	4	5	6
$P(X=x)$	$\frac{1}{2}$	$\frac{1}{4}$				

either 6T or 5T + 1H

The Geometric Distribution

The **geometric distribution** is the distribution produced by the random variable X defined to count the number of trials needed to obtain the first success.

For example: Flipping a coin until you get a head
Rolling a die until you get a 5

A random variable X is geometric if the following conditions are met:

1. Each observation falls into one of just two categories, "success" or "failure."
2. The probability of success is the same for each observation.
3. The observations are all independent.
4. The variable of interest is the number of trials required to obtain the first success. (not given n)

Notice that this is different from the binomial distribution in that the number of trials is unknown. With geometric distributions we are trying to determine how many trials are needed in order to obtain a success.

The probability that the first success occurs on the n^{th} trial is

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

where p is the probability of success.

The probability that it takes *more* than n trials to see the first success is
$\boxed{P(X > n) = (1-p)^n}$

R commands:

$$\left.\begin{aligned} P(X = n) &= \text{dgeom}(n-1, \text{p}) \\ P(X \le n) &= \text{pgeom}(n-1, \text{p}) \\ P(X > n) &= 1\text{-pgeom}(n-1, p) \end{aligned}\right\} *$$

The mean, or expected number of trials to get a success in a geometric distribution is $E[X] = \mu = \dfrac{1}{p}$ and the variance is $\sigma^2 = \dfrac{1-p}{p^2}$.

$\sigma = \dfrac{\sqrt{1-p}}{p}$

$$\sum_{k=1}^{n} r^k = \frac{r(1-r^n)}{1-r}$$

$$\begin{aligned}
P(X>n) &= 1 - P(X \le n) \\
&= 1 - \sum_{k=1}^{n} (1-p)^{k-1} \cdot p \\
&= 1 - p \sum_{k=1}^{n} (1-p)^{k-1} = 1 - \frac{p}{1-p} \underbrace{\sum_{k=1}^{n} (1-p)^{k}} \\
&= 1 - \frac{p}{1-p} \left(\frac{(1-p)\left(1-(1-p)^n\right)}{1-(1-p)} \right) \\
&= 1 - \frac{\cancel{p}}{\cancel{(1-p)}} \, \frac{\cancel{(1-p)}\left(1-(1-p)^n\right)}{\cancel{p}} \\
&= 1 - 1 + (1-p)^n \\
&= \underline{(1-p)^n}
\end{aligned}$$

$p = .44$

Examples:

A quarter back completes 44% of his passes. We want to observe this quarterback during one game to see how many pass attempts he makes before completing one pass.

a. What is the probability that the quarterback throws 3 incomplete passes before he has a completion?

$$P(X=4) = (1-.44)^3(.44) = \text{dgeom}(3, .44)$$
$$= .0773$$

b. How many passes can the quarterback expect to throw before he completes a pass?

$$E[X] = \frac{1}{p} = \frac{1}{.44} = 2.72$$

c. Determine the probability that it takes more than 5 attempts before he completes a pass.

$$P(X>5) = (1-.44)^5 = .0551$$
$$1 - P(X \leq 5) = 1 - \text{pgeom}(4, .44)$$

Ex: Newsweek in 1989 reported that 60% of young children have blood lead levels that could impair their neurological development. Assuming a random sample from the population of all school children at risk, find:

* a. The probability that at least 5 children out of 10 in a sample taken from a school may have a blood lead level that may impair development. $n=10$

Binomial $p=.6$ $n=10$ $\sim \text{Binom}(10, .6)$

$$P(X \geq 5) = 1 - P(X \leq 4) = 1 - \text{pbinom}(4, 10, .6)$$
$$= .8338$$

b. The probability you will need to test 10 children before finding a child with a blood lead level that may impair development. $\sim \text{Geom}(.6)$

$$P(X=10) = (1-.6)^9(.6) = \text{dgeom}(9, .6)$$
$$= .00016$$

c. The probability you will need to test no more than 10 children before finding a child with a blood lead level that may impair development.

$$P(X \leq 10) = \text{pgeom}(9, .6) = .9999$$

Popper 06

$p = .6$

2. Joe has an 60% probability of passing his statistics quiz 4 each time he takes it. What is the probability he will take no more than 7 tries to pass it?

$P(X \le 7)$ (TI 84 geomcdf)

$1 - P(X > 7) = 1 - (1 - .6)^7$

a) 0.9990

b) 0.0009

c) 0.0229

d) 0.9983

e) 0.0016

The Poisson Probability Distribution

A ***Poisson random variable***, X, represents the number of occurrences of a rare event during some fixed time period, where the expected number of occurrences is λ $= \mu$. (For example: The number of customers to enter a furniture store between 1:00pm and 2:00pm)

Conditions under which X is Poisson:

1. There is a fixed time period during which we are counting occurrences.
2. The expected number of arrivals during that time period is known to be λ.
3. The occurrences are independent of one another (the number of occurrences that have happened at any given point does not affect future occurrences).

If X has a Poisson distribution with parameter λ, then

$$P\left(X=x\right)=\frac{e^{-\lambda}\lambda^{x}}{x!}, \quad x=0,1,2,\ldots$$ where λ is a positive real number.

$\lambda = 5$ $\quad P(X=2) = \dfrac{e^{-5}(5)^2}{2!} = \dfrac{5^2}{e^5(2!)}$

R commands:

$$P(X=x)=dpois(x,\lambda)$$

$$P(X\le x)=ppois(x,\lambda)$$

$$P(X>x)=1-ppois(x,\lambda)$$

For a Poisson random variable, X, with parameter λ

$E[X] = \lambda$, and $Var(X) = \sigma^2 = \lambda$

Example:

Let X be the number of flaws on the surface of a randomly selected boiler of a certain type that as a Poisson distribution with parameter $\mu = 4$. Find $P(4 \leq X \leq 7)$

1 2 3 $\boxed{4\ 5\ 6\ 7}$ 8 $\qquad \lambda$

$$P(4 \leq X \leq 7) = P(X \leq 7) - P(X \leq 3)$$

$$= \text{ppois}(7,4) - \text{ppois}(3,4)$$

$$= .5153$$

$\lambda = 2$

Calls to a toll-free telephone hotline service are made randomly and independently at an expected rate of two per minute. The hotline service has five customer service representatives, none of whom is currently busy. Using a Poisson distribution, determine the probability that the hotline receives fewer than five calls in the next minute.

$$P(X < 5) = P(X \le 4) = \text{ppois}(4, 2)$$

$$= .9473$$

Example (Using Poisson to approximate Binomial):
A computer hardware company manufactures a particular type of microchip. There is a 0.1% chance that any given microchip of this type is defective, independent of the other microchips produced.
Using the Poisson distribution, approximate the probability that there are at least 2 defective microchips in a shipment of 1000. Then use the Binomial distribution to determine the exact probability for comparison.

$p = .001 \quad n = 1000 \quad \sim \text{Binom}(1000, .001)$

$P(X \geq 2) = 1 - P(X \leq 1) = 1 - \text{pbinom}(1, 1000, .001)$

$= .2642$

$E[X] = np = 1000(.001)$

Poisson: $\lambda = 1$

$P(X \geq 2) = 1 - P(X \leq 1) = 1 - \text{ppois}(1, 1)$

$= .2642$

Popper 06

3. Suppose the number X of tornadoes observed in Kansas during a 1 year period has a Poisson distribution with $\lambda = 9$. Compute $P(6 \le X \le 9)$

a. 0.578

b. 0.116

c. 0.471

d. 0.413

~~e. none of these~~

```
> ppois(9,9)-ppois(6,9)
[1] 0.3806274
> ppois(9,9)-ppois(5,9)
[1] 0.4717177
> ppois(9,9)
[1] 0.5874082
> dpois(9,9)-dpois(6,9)
[1] 0.04066532
> dpois(9,9)-dpois(5,9)
[1] 0.07102876
```

4. If we plotted the heights of the team members on the cartoon, the graph would be

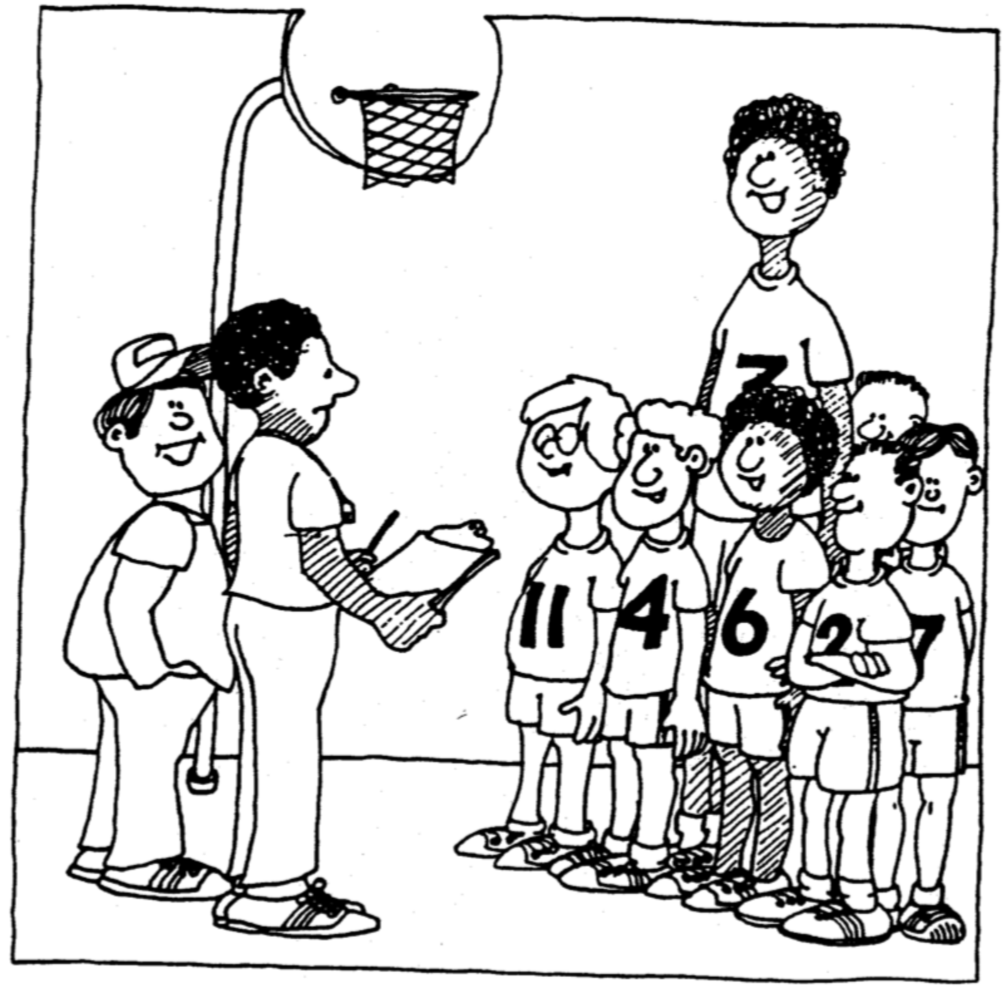
a. skewed right

b. skewed left

c. roughly symmetrical

d. there is no way to tell without more data

Heights



"Should we scare the opposition by announcing our mean height or lull them by announcing our median height?"

Chebyshev's Inequality:

Theorem 4.3. If X is a random variable with mean μ and standard deviation σ and if k is a positive constant, then

$$Pr(|X-\mu| > k\sigma) \le 1/k^2.$$

Ex 4.3 #2

Verify Chebyshev's inequality for k = 2 and k = 3 when X is the total number of spots on two rolls of a fair 6-sided die.

Let's do this for k=~~3~~: 1.5

Example 4.1. Roll a 6-sided die twice. Assume that all 36 outcomes are equally likely. Let T denote the total number of spots on the two rolls. A table of values of T and their probabilities is given below.

<3.4 $\qquad$ >10.6

t	2	3	4	5	6	7	8	9	10	11	12
f(t)	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36

$E[X]=7$ $\quad$ $\sigma = 2.415$

$P(|X-7| > 1.5(2.415)) \le \frac{1}{1.5^2}$

$P(X>10.6)$ or $P(X<3.4) \le \frac{4}{9}$

$\frac{2}{36}+\frac{1}{36} + \frac{1}{36}+\frac{2}{36}$

$\frac{1}{6} \le \frac{4}{9}$ ✓

$|X-7| > 3.6225$

$X-7 > 3.6225$ OR $X-7 < -3.6225$

$X > 10.6225$ or $X < 3.3775$

Popper 06

#5 choose B