Ch 3.4 Practice: Named Discrete Distributions

Solutions

Last revised: 2022-04-17 22:30:25

1. Suppose you take a 20 question multiple choice test, where each question has four choices. You guess randomly on each question. You are interested in your overall score on the test.

Let X be the number of correct questions, $n = 20$ and $p = .25$. $X \sim Binomial(20, .25)$

a. What is your expected score?

Find: $E(X)$

```
# Theoretical
20*.25
```

```
## [1] 5
```

```
# Simulation
x <- rbinom(10000, 20, .25)
mean(x)
```

```
## [1] 5.0246
```

b. What is the probability you get 10 or more questions correct?

Find: $P(X \geq 10)$

```
# Theoretical
1 - pbinom(9,20,0.25)
```

```
## [1] 0.01386442
```

```
# Simulation
correct <- rbinom(10000,20,.25)
mean(correct>9)
```

```
## [1] 0.0124
```

2. A recent national study showed that approximately 45% of college students binge drink. Let X equal the number of students in a random sample of size $n = 12$ who binge drink.

$$X \sim Binomial(12, .45)$$

a. What is the probability that X is at most 2?

Theoretical using the pmf

$$P(X \leq 2) = P(X=0)+P(X=1)+P(X=2) = \binom{12}{0}(.45)^0(.55)^{12}+\binom{12}{1}(.45)^1(.55)^{11}+\binom{12}{2}(.45)^2(.55)^{10}$$

```
choose(12, 0)*(0.45)^(0)*(0.55)^(12) +
choose(12, 1)*(0.45)^(1)*(0.55)^(11) +
choose(12, 2)*(0.45)^(2)*(0.55)^(10)
```

```
## [1] 0.04214198
```

Theoretical using R commands

```
pbinom(2, 12, .45)
```

```
## [1] 0.04214198
```

Simulation

```
binge <- rbinom(10000, 12, .45)
mean(binge<=2)
```

```
## [1] 0.0404
```

b. What is the probability that X is at least 1?

Theoretical using the pmf

$$P(X \geq 1) = 1 - P(X=0)1 - \binom{12}{0}(.45)^0(.55)^{12}$$

```
1-choose(12, 0)*(0.45)^(0)*(0.55)^(12)
```

```
## [1] 0.9992338
```

Theoretical using R commands

```
1-pbinom(0, 12, .45)
```

```
## [1] 0.9992338
```

```
# or
1 - dbinom(0, 12, .45)
```

```
## [1] 0.9992338
```

Simulation

```
mean(binge >= 1)
```

```
## [1] 0.9995
```

c. Use simulation to obtain the mean and variance of X.

```r
x <- rbinom(100000, size=12, prob = .45)
mean(x)
```

```
## [1] 5.39521
```

```r
var(x)
```

```
## [1] 2.970429
```

3. Stephen Curry is a 91% free throw shooter. He decides to shoot free throws until his first miss.

Let X be the number of shots Stephen Curry makes before he misses, $p(miss) = .09$.

$$X \sim Geom(.09)$$

```
p <- .09
```

a. What is the probability that he shoots exactly 20 free throws (including the one he misses)?

Find: $P(X = 19)$

Theoretical using the pmf

```
(1-p)^19*p
```

```
## [1] 0.01499785
```

Theoretical using R commands

```
dgeom(19, p)
```

```
## [1] 0.01499785
```

Simulation

```
shots <- rgeom(10000, p)
mean(shots==19)
```

```
## [1] 0.014
```

4. A couple decides to have children until a daughter is born. Assume the probability of a daughter is 0.5.

Let X be the number of sons a couple has before having a daughter. $X \sim Geom(.5)$

What is the expected number of children this couple will have?

Since the question is asking about the total number of children, and X is only the number of sons, you have to add 1 to the following answers.

Theoretical using the formula

```
p.girl <- 0.5
(E_n.kids <- (1-p.girl)/p.girl)+1
```

```
## [1] 2
```

Simulation

```
n.kids <- rgeom(10000,p.girl)
mean(n.kids)+1
```

```
## [1] 1.9855
```

5. Yruc Nepets is not nearly as good as Curry and only can make a free throw 19% of the time.

Let X be the number of shots Nepets *misses* before they successfully make five free throws?

$X \sim NegBinom(5, .19)$

a. What is the expected number of free throws they will take until they make 5 baskets?

The expected number of free throws made until the 5th basket is $E(X+5) = E(X) + 5$ (expected number of failures + the 5 successes)

Theoretical using the formula

$$E(X) = \frac{n(1-p)}{p} = \frac{5(.81)}{.19} + 5$$

```
5*.81/.19+5
```

```
## [1] 26.31579
```

Simulation

```
Nepets.shots <- rnbinom(10000, 5, .19)
mean(Nepets.shots)+5
```

```
## [1] 26.2387
```

b. What is the variance in the number of free throws?

Since the number of total free throws is $X + 5$, $Var(X+5) = Var(X) + Var(5) = Var(X)$

Theoretical using the formula

$$Var(X) = \frac{n(1-p)}{p^2} = \frac{5(.81)}{.19^2}$$

```
5*.81/.19^2
```

```
## [1] 112.1884
```

Simulation

```
var(Nepets.shots)
```

```
## [1] 112.2385
```

c. What is the probabilty that Yruc will make their 5th successful free throw in at most 10 shots?

Don't write out the pmf equation here, use R functions and simulation only.

At most 10 shots, with 5 successes means at most 5 failures. Find: $P(X \leq 5)$

Theoretical using R commands

```r
pnbinom(5, 5, .19)
```

```
## [1] 0.02663246
```

Simulation

```r
mean(Nepets.shots <= 5)
```

```
## [1] 0.0239
```

6. There is a .85 probabilty that Sanjeev will show up to work exactly on time (not 1 minute late), independent of any other day. The company gives a $20 bonus for every employee that shows up to work exactly on time at least 5 days in a week.

Let X be the number of days Sanjeev doesn't show up to work on time before he arrives exactly on time 5 times. $X \sim NegBinom(5, .85)$

Note: If you let X be the number of days Sanjeev shows up to work in a week, then $X \sim Binomial(7, .85)$. You can answer part b and c using this distribution, but not part a because I specified that the 5th success is on thursday (the 5th day)

a. Assuming the work week starts on Sunday (as day 1), what is the probability that Sanjeev gets this bonus by thursday?

If X is the number of tardy days, then to get his bonus by Thursday he can't be tardy. That is $X = 0$.
Theoretical using the pmf

$$P(X = 0) = \binom{0 + 5 - 1}{0}(.85)^5(.15)^0$$

```
choose(4, 0)*.85^5*.15^0
```

```
## [1] 0.4437053
```

Theoretical using the formula

```
dnbinom(0, 5, .85)
```

```
## [1] 0.4437053
```

Simulation

```
something.need.doing <- rnbinom(10000, 5, .85)
mean(something.need.doing == 0)
```

```
## [1] 0.4402
```

b. What day on average will Sanjeev get this bonus?

Since X is the number of days he will *not* show up on time before the 5th time he does, to get the day of when the bonus is awarded we get to add 5 to X. For example if X=1, then he missed 1 day, and gets the bonus on the 6th day.

Theoretical using the formula $E(X) = n\frac{1-p}{p} = 5\frac{.15}{.85}$

```
5*.15/.85+5
```

```
## [1] 5.882353
```

Simulation

```
mean(something.need.doing)+5
```

```
## [1] 5.8946
```

On average, since he has such a high probability of getting to work on time, he will get the bonus on Friday.

c. What is the probability that Sanjeev won't get this bonus?

Find $P(X > 2)$ If he isn't on time for more than 2 days in the week he can't be on time for 5 days that week.

Theoretical using the formula $P(X > 2) = 1 - P(X \leq 2) = 1 - [P(X = 0) + P(X = 1) + P(X = 2)]$

I don't want to type out that pmf, so here's the **Theoretical using R commands**

```
1-pnbinom(2, 5, .85)
```

```
## [1] 0.07376516
```

simulation

```
mean(something.need.doing>2)
```

```
## [1] 0.0741
```

7. Flaws in a certain type of drapery material appear on the average of two in 150 square feet. If we assume a Poisson distribution, find the probability of at most 2 flaws in 450 square feet.

Let X be the number of flaws in $450ft^2$ of material. $X \sim Poison(6)$. Find $.

```
lambda <- 2*3
```

Theoretical using the pmf

$$P(X \leq 2) = e^{-6}\frac{6^0}{0!} + e^{-6}\frac{6^1}{1!} + e^{-6}\frac{6^2}{2!}$$

```
exp(-6)*(6^{0})/factorial(0)+exp(-6)*(6^{1})/factorial(1)+exp(-6)*(6^{2})/factorial(2)
```

```
## [1] 0.0619688
```

Theoretical using R commands

```
ppois(2,lambda)
```

```
## [1] 0.0619688
```

Simulation

```
mybad <- rpois(10000, lambda)
mean(mybad<=2)
```

```
## [1] 0.0584
```

8. The Brown's Ferry incident of 1975 focused national attention on the ever-present danger of fires breaking out in nuclear power plants. The Nuclear Regulatory Commission has estimated that with present technology there will be, on average, one fire for every 10 years. Assume the incident of fires can be described by a Poisson distribution.

Let X be the number of fires in 10 years. $X \sim Poisson(1)$

a. What is the expected value and variance of the number of fires over 10 years?

Theoretical $E(X) = Var(X) = \lambda = 1$

Simulation

```
fires <- rpois(10000, 1)
var(fires)
```

```
## [1] 1.035428
```

b. Suppose that a certain state put three reactors on line in 2000. What is the probability that at least two fires will have occurred by 2015. *Careful! The time frame has changed. Be sure to adjust λ accordingly.*

Let X be the number of fires in 16 years. $X \sim Poisson(1.6)$. Find $P(X \geq 2)$

Theoretical using the pmf

$$P(X \geq 2) = 1 - P(X \leq 1) = 1 - e^{-1.6}\frac{1.6^0}{0!} + e^{-1.6}\frac{1.6^1}{1!}$$

Theoretical using R commands

```
1 - ppois(1, 1.6)
```

```
## [1] 0.4750691
```

Simulation

```
fire.bad <- rpois(10000, 1.6)
mean(fire.bad >=2)
```

```
## [1] 0.4811
```

9. In the game of Scrabble, players make words using letter tiles. The tiles consist of 42 vowels and 58 non-vowels.

a. If a player draws 7 tiles, what is the probability of getting 7 vowels?

Let X be the number of vowels drawn out of a sample of $k = 7$ tiles, from a population of $m = 42$ vowels and $n = 58$ non-vowels. $X \sim Hypergeometric(100, 58, 7)$

Find $P(X = 7)$.

```r
m <- 42
n <- 58
k <- 7
total <- m+n
```

Theoretical using the pmf

$$P(X = 7) = \frac{\binom{58}{0}\binom{42}{7}}{\binom{100}{7}}$$

```r
choose(n, 0)*choose(m, 7)/choose(total, 7)
```

```
## [1] 0.001685349
```

Theoretical using R commands

```r
dhyper(7,m,n,k)
```

```
## [1] 0.001685349
```

Simulation

```r
id.like.to.buy.a.vowel <- rhyper(10000, m,n,k)
mean(id.like.to.buy.a.vowel == 7)
```

```
## [1] 0.0019
```

b. If a player draws 7 tiles, what is the probability of 2 or fewer vowels?

Find $P(X \leq 2)$

Theoretical using the pmf

$$P(X \leq 2) = \frac{\binom{42}{0}\binom{58}{7}}{\binom{100}{7}} + \frac{\binom{42}{1}\binom{58}{6}}{\binom{100}{7}} + \frac{\binom{42}{2}\binom{58}{5}}{\binom{100}{7}}$$

```r
choose(m, 0)*choose(n, 7)/choose(total, k) +
choose(m, 1)*choose(n, 6)/choose(total, k) +
choose(m, 2)*choose(n, 5)/choose(total, k)
```

```
## [1] 0.3714395
```

Theoretical using R commands

```r
phyper(2,m,n,k)
```

```
## [1] 0.3714395
```

Simulation

```r
mean(id.like.to.buy.a.vowel <=2)
```

```
## [1] 0.3657
```

c. What is the expected number of vowels drawn when drawing 7 tiles?

$E(X) = 7\frac{58}{100}$

```r
k*m/total
```

```
## [1] 2.94
```

Simulation

```r
mean(id.like.to.buy.a.vowel)
```

```
## [1] 2.9516
```

d. What is the standard deviation of the number of vowels drawn when drawing 7 tiles?

$$SD(X) = \sqrt{Var(X)} = 7\frac{m}{m+n}\frac{n}{m+n}\frac{m+n-k}{m+n-7}$$

```r
k*(m/total)*(n/total)*((total-k)/(total-1))
```

```
## [1] 1.601855
```

Simulation

```r
sd(id.like.to.buy.a.vowel)
```

```
## [1] 1.269259
```

10. One of the popular tourist attractions in Alaska is watching black bears catch salmon swimming upstream to spawn. Not all "black" bears are black, some are tan-colored. Suppose that six black bears and three tan-colored bears are working the rapids of a salmon stream. Over the course of an hour, four different bears are sighted.

What is the probability that there were more than twice as many black bears as tan-colored bears in those 4 sighted?

Let X be the number of black bears sighted out of the $k = 4$ from a total population of $m = 6$ black bears, $n = 3$ tan bears. $X \sim Hypergeometric(9, 3, 4)$

```
m.black <- 6
n.tan <- 3
k.seen <- 4
total.bears <- m.black + n.tan
```

When only observing 4 bears, "twice as many" means 3 black and 1 tan, or 4 black and 0 tan. So Find: $P(X \geq 3) = P(X = 3) + P(X = 4)$

Theoretical using the pmf

$$P(X \geq 3) = \frac{\binom{6}{3}\binom{3}{1}}{\binom{9}{4}} + \frac{\binom{6}{4}\binom{3}{0}}{\binom{9}{4}}$$

```
choose(m.black, 3)*choose(n.tan, 1) / choose(total.bears, k.seen) +
choose(m.black, 4)*choose(n.tan, 0) / choose(total.bears, k.seen)
```

```
## [1] 0.5952381
```

Theoretical using R commands

```
dhyper(3,m.black, n.tan, k.seen) + dhyper(4,m.black, n.tan, k.seen)
```

```
## [1] 0.5952381
```

```
## or
1 - phyper(2,m.black, n.tan, k.seen)
```

```
## [1] 0.5952381
```

Simulation

```
da.bears <- rhyper(10000,m.black, n.tan, k.seen)
mean(da.bears >= 3)
```

```
## [1] 0.5956
```