

# Chapter 5: Statistics with R - 2nd Edition

Robert Stinerock

Student Exercises

1. Suppose we have an experiment which consists of flipping a coin three times.

(a) List the elements of the sample space.

$$S = \{(HHH), (HHT), (HTH), (THH), (HTT), (THT), (TTH), (TTT)\}$$

(b) Define the random variable  $x$  that describes the number of tails.

Answer: Let  $x$  be the number of tails occurring on three coin flips

(c) Since a random variable is a quantitative characterization of the experimental outcome, what is the value of  $x$  for each sample point in the sample space?

Answer:

Sample Point	HHH	HHT	HTH	THH	HTT	THT	TTH	TTT
$x$	0	1	1	1	2	2	2	3

2. Referring to the previous exercise, provide the entire probability distribution of the random variable  $x$  along with the associated probability function  $f(x)$ .

Answer: when  $x = 0$ ,  $f(0) = 0.1250$ ; when  $x = 1$ ,  $f(1) = 0.3750$ ; when  $x = 2$ ,  $f(2) = 0.3750$ ; and when  $x = 3$ ,  $f(3) = 0.1250$

3. Again, referring to the previous exercise, is this a valid probability distribution?

Answer: Yes, this is a valid probability distribution since for all values  $i$  of  $f(x_i)$

$$0 \leq f(x_i) \leq 1$$

and

$$\sum_{i=0}^3 f(x_i) = 1$$

4. Referring again to the previous exercise, what is the probability that the random variable  $x$  is less than or equal to 2?

Answer:

$$p(x \leq 2) = p(x = 0) + p(x = 1) + p(x = 2) = 0.1250 + 0.3750 + 0.3750 = 0.8750$$

5. Referring once more to the previous exercise, what is  $E(x)$ ?

Answer:  $E(x) = 1.50$

$$E(x) = \sum_{i=0}^3 x_i f(x_i) = (0)(0.1250) + (1)(0.3750) + (2)(0.3750) + (3)(0.1250) = 1.50$$

6. For each binomially-distributed random variable  $x$ , provide the expected value, the variance, and the standard deviation.

(a)  $n = 75$  and  $p = 0.40$

Answer:

$$E(x) = \mu = np = (75)(0.40) = 30$$

$$\sigma^2 = np(1 - p) = (75)(0.40)(0.60) = 18$$

$$\sigma = \sqrt{np(1 - p)} = \sqrt{18} = 4.2426$$

(b)  $n = 40$  and  $p = 0.70$

Answer:

$$E(x) = \mu = np = (40)(0.70) = 28$$

$$\sigma^2 = np(1 - p) = (40)(0.70)(0.30) = 8.4$$

$$\sigma = \sqrt{np(1 - p)} = \sqrt{8.4} = 2.8983$$

(c)  $n = 107$  and  $p = 0.30$

Answer:

$$E(x) = \mu = np = (107)(0.30) = 32.1$$

$$\sigma^2 = np(1 - p) = (107)(0.30)(0.70) = 22.47$$

$$\sigma = \sqrt{np(1 - p)} = \sqrt{22.47} = 4.7403$$

(d)  $n = 4121$  and  $p = 0.50$

Answer:

$$E(x) = \mu = np = (4121)(0.50) = 2060.5$$

$$\sigma^2 = np(1 - p) = (4121)(0.50)(0.50) = 1030.25$$

$$\sigma = \sqrt{np(1 - p)} = \sqrt{1030.25} = 32.0975$$

(e)  $n = 11$  and  $p = 0.60$

Answer:

$$E(x) = \mu = np = (11)(0.60) = 6.6$$

$$\sigma^2 = np(1 - p) = (11)(0.60)(0.40) = 2.64$$

$$\sigma = \sqrt{np(1-p)} = \sqrt{2.64} = 1.6248$$

(f)  $n = 37$  and  $p = 0.65$

Answer:

$$E(x) = \mu = np = (37)(0.65) = 24.05$$

$$\sigma^2 = np(1-p) = (37)(0.65)(0.35) = 8.4175$$

$$\sigma = \sqrt{np(1-p)} = \sqrt{8.4175} = 2.9013$$

7. If 85% of vehicles arriving at the Lincoln Tunnel (connecting New Jersey and New York City) have either New York or New Jersey license plates, what is the probability that, of the next 20 vehicles, 2 or fewer (that is, 0, 1, or 2) will bear license plates from states other than New Jersey or New York?

Answer: Using  $f(x) = p(x|n, p) = \binom{n}{x} p^x (1-p)^{n-x}$ , setting  $n = 20$  and  $p = 0.15$ , find  $f(0)$ ,  $f(1)$ , and  $f(2)$ , and sum. The answer is 0.4049.

$$\begin{aligned} \sum_{x=0}^2 \binom{n}{x} p^x (1-p)^{n-x} &= \sum_{x=0}^2 \binom{20}{x} (0.15)^x (1-0.15)^{20-x} \\ &= \binom{20}{0} (0.15)^0 (0.85)^{20} + \binom{20}{1} (0.15)^1 (0.85)^{19} + \binom{20}{2} (0.15)^2 (0.85)^{18} \\ &= 0.0388 + 0.1368 + 0.2293 = 0.4049 \end{aligned}$$

Note: if we prefer using R to answer this question (and we almost certainly will do), three approaches can be used to find this probability:

```
# (1) Use the dbinom(x, n, p) function for x = 0, 1, and 2 and sum.
dbinom(0, 20, 0.15) + dbinom(1, 20, 0.15) + dbinom(2, 20, 0.15)

## [1] 0.4048963
```

```

# (2) Use pbinom(x, n, p) function to find the cumulative binomial
# probabilities.

pbinom(2, 20, 0.15)

## [1] 0.4048963

# (3) Use the (nested) sum(dbinom()) function.

sum(dbinom(0 : 2, 20, 0.15))

## [1] 0.4048963

```

8. Find the binomial probability for each exercise below. Label each parameter as  $x$ ,  $n$ , and  $p$ , and submit to `dbinom(x, n, p)`. Hint: remember that we can answer each question by simply hitting the up arrow (to recover the `dbinom()` function) and entering the required  $n$ ,  $x$ , and  $p$  values as needed. Any function, not only the `dbinom()` function, can be recycled repeatedly in this way.

(a)  $n = 50$ ,  $x = 10$ , and  $p = 0.20$

Answer: 0.1398

```

# Use function dbinom() and n, x, and p.

n = 50
x = 10
p = 0.20

dbinom(x, n, p)

## [1] 0.139819

```

(b)  $n = 40$ ,  $x = 20$ , and  $p = 0.40$

Answer: 0.05541

```

# Use function dbinom() and n, x, and p.

n = 40
x = 20
p = 0.40

dbinom(x, n, p)

## [1] 0.05541415

```

(c)  $n = 30$ ,  $x = 10$ , and  $p = 0.30$

Answer: 0.1416

```
# Use function dbinom() and n, x, and p.  
  
n = 30  
x = 10  
p = 0.30  
  
dbinom(x, n, p)  
  
## [1] 0.1415617
```

(d)  $n = 50$ ,  $x = 30$ , and  $p = 0.70$

Answer: 0.03704

```
# Use function dbinom() and n, x, and p.  
  
n = 50  
x = 30  
p = 0.70  
  
dbinom(x, n, p)  
  
## [1] 0.03703876
```

(e)  $n = 40$ ,  $x = 20$ , and  $p = 0.60$

Answer: 0.05541

```
# Use function dbinom() and n, x, and p.  
  
n = 40  
x = 20  
p = 0.60  
  
dbinom(x, n, p)  
  
## [1] 0.05541415
```

(f)  $n = 30$ ,  $x = 10$ , and  $p = 0.40$

Answer: 0.1152

```
# Use function dbinom() and n, x, and p.  
  
n = 30  
x = 10  
p = 0.40  
  
dbinom(x, n, p)
```

```
## [1] 0.1151854
```

9. At a large public university, half the students enrolled in a statistics class take the course on a pass-fail basis; the other half take it for a normal grade.

- (a) If 20 students are selected at random, what is the probability that 12 (of the 20) are taking the course pass-fail?

Answer: 0.1201

```
# Use the dbinom(x, n, p) function.
```

```
dbinom(12, 20, 0.50)
```

```
## [1] 0.1201344
```

- (b) What is the probability that no more than 5 students are taking the course for a normal grade?

Answer: 0.02069

```
# (1) Use pbinom(x,n,p) for cumulative probabilities.
```

```
pbinom(5, 20, 0.50)
```

```
## [1] 0.02069473
```

```
# (2) Use the sum(dbinom()) function.
```

```
sum(dbinom(0 : 5, 20, 0.50))
```

```
## [1] 0.02069473
```

- (c) What is the expected number who are taking the course on a pass-fail basis?

Answer:  $E(x) = np = (20)(0.50) = 10$  students.

10. Suppose  $x$  is a Poisson-distributed random variable with an expected value of 3 occurrences per interval.

- (a) What is the Poisson probability function  $f(x)$ ?

Answer:

$$f(x) = p(x|\mu) = \frac{\mu^x e^{-\mu}}{x!} =$$

$$f(x) = p(x|\mu = 3) = \frac{3^x e^{-3}}{x!}$$

(b) What is the probability of 2 occurrences over the same interval?

Answer: 0.224

$$f(2) = p(x = 2|\mu = 3) = \frac{3^2 e^{-3}}{2!} = 0.224$$

```
# Use the dpois() function.  
  
x <- 2  
  
mu <- 3  
  
dpois(x, mu)  
  
## [1] 0.2240418
```

(c) What is the probability of 3 occurrences over the same interval?

Answer: 0.224

$$f(3) = p(x = 3|\mu = 3) = \frac{3^3 e^{-3}}{3!} = 0.224$$

```
# Use the dpois() function.  
  
x <- 3  
  
mu <- 3  
  
dpois(x, mu)  
  
## [1] 0.2240418
```

(d) What are the Poisson probabilities of  $x = 0, 1, \dots, 4$  when  $\mu = 3$ ?

```
# Use the dpois() function.  
  
x1 <- 0
```



```
x2 <- 4

mu <- 3

dpois(x1 : x2, mu)

## [1] 0.04978707 0.14936121 0.22404181 0.22404181 0.16803136
```

(e) What are the Poisson probabilities of  $x = 5, 6, 7, 8$  when  $\mu = 3$ ?

```
# Use the dpois() function.

x1 <- 5

x2 <- 8

mu <- 3

dpois(x1 : x2, mu)

## [1] 0.100818813 0.050409407 0.021604031 0.008101512
```

11. A Poisson process has an expected value of 5 occurrences per interval.

(a) What is the Poisson probability function  $f(x)$ ?

Answer:

$$f(x) = p(x|\mu) = \frac{\mu^x e^{-\mu}}{x!} =$$

$$f(x) = p(x|\mu = 5) = \frac{5^x e^{-5}}{x!}$$

(b) What is the expected number of occurrences in 5 intervals?

Answer: The expected number of occurrences is 25 for 5 intervals.

(c) What is the new probability function for 5 intervals?

Answer:

$$f(x) = p(x|\mu) = \frac{\mu^x e^{-\mu}}{x!} =$$

$$f(x) = p(x|\mu = 25) = \frac{25^x e^{-25}}{x!}$$

(d) What is the expected number of occurrences in 0.25 intervals?

Answer: The expected number of occurrences is 1.25 for 0.25 intervals.

(e) What is the new probability function for 0.25 intervals?

Answer:

$$f(x) = p(x|\mu) = \frac{\mu^x e^{-\mu}}{x!} =$$

$$f(x) = p(x|\mu = 1.25) = \frac{1.25^x e^{-1.25}}{x!}$$

12. If  $\mu = 1.25$ , what is the probability of  $x = 0, 1, 2, 3$  occurrences? Use  $f(x)$  and R.

Answer:  $f(0) = 0.2865$ ;  $f(1) = 0.35813$ ;  $f(2) = 0.22383$ ; and  $f(3) = 0.09326$

$$f(x) = p(x|\mu = 1.25) = \frac{1.25^x e^{-1.25}}{x!}$$

$$f(0) = p(x = 0|\mu = 1.25) = \frac{1.25^0 e^{-1.25}}{0!} = 0.2865$$

$$f(1) = p(x = 1|\mu = 1.25) = \frac{1.25^1 e^{-1.25}}{1!} = 0.35813$$

$$f(2) = p(x = 2|\mu = 1.25) = \frac{1.25^2 e^{-1.25}}{2!} = 0.22383$$

$$f(3) = p(x = 3|\mu = 1.25) = \frac{1.25^3 e^{-1.25}}{3!} = 0.09326$$

```

# Use the dpois() function.

x1 <- 0

x2 <- 3

mu <- 1.25

dpois(x1 : x2, mu)

## [1] 0.28650480 0.35813100 0.22383187 0.09326328

```

13. If  $\mu = 25$ , what is probability of  $x = 23, 24, 25, 26$  occurrences? Use  $f(x)$  and R.

Answer:  $f(23) = 0.07634$ ;  $f(24) = 0.07952$ ;  $f(25) = 0.07952$ ; and  $f(26) = 0.07646$

$$f(x) = p(x|\mu = 25) = \frac{25^x e^{-25}}{x!}$$

$$f(23) = p(x = 23|\mu = 25) = \frac{25^{23} e^{-25}}{23!} = 0.07634$$

$$f(24) = p(x = 24|\mu = 25) = \frac{25^{24} e^{-25}}{24!} = 0.07952$$

$$f(25) = p(x = 25|\mu = 25) = \frac{25^{25} e^{-25}}{25!} = 0.07952$$

$$f(26) = p(x = 26|\mu = 25) = \frac{25^{26} e^{-25}}{26!} = 0.07646$$

```

# Use the dpois() function.

x1 <- 23

x2 <- 26

mu <- 25

dpois(x1 : x2, mu)

## [1] 0.07634203 0.07952295 0.07952295 0.07646438

```

14. Calls arrive at a customer-service number at the rate of 20 per hour. Use R to answer the following questions.

(a) What is the probability of 5 or more calls in a 15 fifteen minute period?

Answer: 0.5595

```
# (1) Convert 20 calls/60 minutes to 5 calls/15 minutes;  
# use 1 minus the probability of 4 or fewer calls.  
  
1 - ppois(4, 5)  
## [1] 0.5595067  
  
# (2) Or subtract from 1 the probabilities of  
# x = 0, 1, 2, 3 and 4 calls.  
  
1 - (dpois(0, 5) + dpois(1, 5) + dpois(2, 5) + dpois(3, 5) +  
      dpois(4, 5))  
## [1] 0.5595067
```

(b) What is the probability that between 7 and 12 calls, inclusive, will arrive in the next 30 minutes?

Answer: 0.6614

```
# (1) Convert 20 calls/60 minutes to 10 calls/30 minutes;  
# subtract ppois(6,10) from ppois(12,10).  
  
ppois(12, 10) - ppois(6, 10)  
## [1] 0.6614151  
  
# (2) Or sum probabilities of x = 7, 8, 9, 10, 11, 12.  
  
sum(dpois(7 : 12, 10))  
## [1] 0.6614151
```

(c) If one particular caller has a difficult problem to resolve and requires 15 minutes of the customer-service representative's time, how many callers would we expect to be waiting on hold once the call has been completed?

Answer: Since 20 calls every 60 minutes is equivalent to 5 calls every 15 minutes, we would expect 5 callers to be waiting in the queue.

(d) If the customer-service representative wishes to take a 5 minute coffee break, what is the probability that no calls will be waiting once he returns?

Answer: 0.1882

```
# Convert 20 calls/60 minutes into 1.67 calls/5 minutes.
x <- 0
mu <- 1.67
dpois(x, mu)
## [1] 0.1882471
```

15. Small aircraft land at a private airport at an average rate of 1 arrival/10 minutes.

- (a) What is the probability that there will be exactly 4 arrivals in the next hour? Use the general form of the Poisson probability function in which the parameter is expressed as  $\mu = \lambda t$  where  $\lambda$  is the rate of occurrence per interval, and  $t$  is the number of intervals.

Answer: 0.1339

To find  $\mu = \lambda t$ , we set  $\lambda = 1$  arrival per interval and  $t = 6$  intervals (since there are 6 10-minute intervals per hour). Thus,  $\mu = \lambda t = (1 \text{ arrival per 10-minute interval})(6 \text{ 10-minute intervals})$ , or  $\mu = \lambda t = (1)(6) = 6$ .

$$p(x|\mu) = p(x|\lambda t) = \frac{(\lambda t)^x e^{-\lambda t}}{x!}$$

$$p(x|\lambda t) = p(x = 4|\lambda t = 6) = \frac{(6)^4 e^{-6}}{4!} = 0.1339$$

```
# Use the dpois() function.
x <- 4
mu <- 6
dpois(x, mu)
## [1] 0.1338526
```

- (b) What is the probability that there will be more than 4 arrivals in the next hour?

Answer: 0.7149

```

# Use 1 minus probability of 4 or fewer arrivals.

x <- 4

mu <- 6

1 - ppois(x, mu)

## [1] 0.7149435

```

- (c) If the airport is open 12 hours a day, what is the probability that fewer than 75 aircraft will arrive in the next day of business (i.e., over the next 12 hours)?

Answer: 0.6227

In this instance, we must adjust the value of the parameter  $\mu = \lambda t$ . Since  $\lambda = 1$  arrival per interval and there are now  $t = 72$  10-minute intervals, it should be clear that  $\lambda t = 72$ . That is, what is  $p(x \leq 74 | \lambda t = 72)$ ?

```

# (1) Use the ppois() function.

ppois(74, 72)

## [1] 0.6226727

# (2) Use the sum(dpois()) function.

sum(dpois(0 : 74, 72))

## [1] 0.6226727

```

16. The number of defects associated with a plate glass manufacturing process is Poisson-distributed with a rate of 0.002 defects per square foot. Suppose the manufacturer has received an order for a large plate glass window of dimensions 10 feet by 10 feet.

- (a) What is the probability the window will have no defects? 1 defect? 2 defect?

Note that  $\mu = \lambda t = (0.002 \text{ defects per square foot})(100 \text{ square feet}) = 0.20$ .

Answer: 0.81873; 0.16375; and 0.01637

```

# Use the dpois() function.

dpois(0 : 2, 0.20)

## [1] 0.81873075 0.16374615 0.01637462

```

(b) What is the probability there will be no more than 2 defects?

Answer: 0.9989

```
# (1) Use the sum(dpois()) function.  
sum(dpois(0 : 2, 0.20))  
## [1] 0.9988515  
# (2) Or use the ppois() function.  
ppois(2, 0.20)  
## [1] 0.9988515
```

(c) What is the probability there will be more than 2 defects?

Answer: 0.001148

```
# Use 1 minus probability of 2 or fewer defects.  
1 - ppois(2, 0.20)  
## [1] 0.001148481
```

17. As soon as the spring season arrives, many businesses contract for construction jobs they have been delaying over the previous cold winter months. One of the major Canadian ice hockey clubs has decided to resurface a large parking area adjacent to their arena. The dimensions of the surface are 300 feet by 200 feet, or 60000 square feet. Unfortunately, since asphalt is usually imperfectly applied, there is an average of 0.06 air bubbles per square foot.

(a) What is the probability there will be between 3500 and 3700 bubbles, inclusive, in the entire surface?

Note that  $\mu = \lambda t = (0.06 \text{ bubbles per square foot})(60000 \text{ square feet}) = 3600$ .

Answer: 0.9061

```
# From ppois(3700,3600) subtract ppois(3499,3600).  
ppois(3700, 3600) - ppois(3499, 3600)  
## [1] 0.9060769  
# or  
sum(dpois(3500 : 3700, 3600))  
## [1] 0.9060769
```

- (b) What is the probability there will be between 3550 and 3650 bubbles, inclusive, in the entire surface?

Answer: 0.6000

```
# From ppois(3650,3600) subtract ppois(3549,3600).  
ppois(3650, 3600) - ppois(3549, 3600)  
## [1] 0.6000281  
  
# or  
sum(dpois(3550 : 3650, 3600))  
## [1] 0.6000281
```

- (c) What is the probability there will be more than 3650 bubbles?

Answer: 0.1998

```
# Use 1 minus probability of 3650 or fewer bubbles.  
1 - ppois(3650, 3600)  
## [1] 0.1997592
```

18. Suppose the number of radioactive particles identified by a sensitive counting device is Poisson-distributed. The particles pass through a certain threshold at a rate of 60 per millisecond. (A millisecond is one-thousandth of a second, or 1/1000 second.)

- (a) What is the expected number of occurrences—the count of radioactive particles—in a second?

Answer:  $\mu = \lambda t = (60 \text{ particles per millisecond})(1000 \text{ milliseconds}) = 60000$

- (b) What is the expected number of occurrences—the count of radioactive particles—in a minute?

Answer:  $\mu = \lambda t = (60 \text{ particles per millisecond})(60000 \text{ milliseconds}) = 3,600,000$

- (c) What is the expected number of occurrences—the count of radioactive particles—in one-third a millisecond?

Answer:  $\mu = \lambda t = (60 \text{ particles per millisecond})(1/3 \text{ milliseconds}) = 20$



- (d) What is the probability of between 18 and 21 particles, inclusive, in one-third of a millisecond?

Answer: 0.3467

```
ppois(21, 20) - ppois(17, 20)
## [1] 0.3466693
# or
sum(dpois(18 : 21, 20))
## [1] 0.3466693
```

19. If  $N = 12$  and  $r = 4$ , what are the hypergeometric probabilities  $f(x)$  for the following values of  $x$  and  $n$ ?

Note: if  $N = 12$  and  $r = 4$ , then  $(N - r) = (12 - 4) = 8$

Recall: the hypergeometric probability function is

$$f(x) = \frac{\binom{r}{x} \binom{N-r}{n-x}}{\binom{N}{n}}$$

which in this case is

$$f(x) = \frac{\binom{4}{x} \binom{8}{n-x}}{\binom{12}{n}}$$

- (a)  $x = 0$  and  $n = 4$

$$f(0) = \frac{\binom{4}{0} \binom{8}{4}}{\binom{12}{4}} = \frac{(1)(70)}{495} = 0.1414$$

(b)  $x = 1$  and  $n = 4$

$$f(1) = \frac{\binom{4}{1} \binom{8}{3}}{\binom{12}{4}} = \frac{(4)(56)}{495} = 0.4525$$

(c)  $x = 2$  and  $n = 4$

$$f(2) = \frac{\binom{4}{2} \binom{8}{2}}{\binom{12}{4}} = \frac{(6)(28)}{495} = 0.3394$$

(d)  $x = 3$  and  $n = 4$

$$f(3) = \frac{\binom{4}{3} \binom{8}{1}}{\binom{12}{4}} = \frac{(4)(8)}{495} = 0.06465$$

(e)  $x = 4$  and  $n = 4$

$$f(4) = \frac{\binom{4}{4} \binom{8}{0}}{\binom{12}{4}} = \frac{(1)(1)}{495} = 0.00202$$

20. Use the `dhyper(x,r,N-r,n)` function to answer all parts of the previous exercise. Hint: copy-and-paste this function along with values for  $N$ ,  $r$ , and  $n$ , and then change the value of  $x$  for each question. All other values remain the same for all parts.

```
# Use the dhyper(x, r, N-r, n) function.
```

```
N = 12
```

```
r = 4
```

```
n = 4
```

```
x = 0
```

```
dhyper(x, r, N - r, n)
```

```
## [1] 0.1414141
```

```
N = 12
r = 4
n = 4
x = 1
dhyper(x, r, N - r, n)
```

```
## [1] 0.4525253
```

```
N = 12
r = 4
n = 4
x = 2
dhyper(x, r, N - r, n)
```

```
## [1] 0.3393939
```

```
N = 12
r = 4
n = 4
x = 3
dhyper(x, r, N - r, n)
```

```
## [1] 0.06464646
```

```
N = 12
r = 4
n = 4
x = 4
dhyper(x, r, N - r, n)
```

```
## [1] 0.002020202
```

21. Referring to the previous exercise, is there another way to find the five probabilities? Do they sum to 1?

```
# (1) Use the dhyper(0 : 4, 4, 8, 4) function to find probabilities.
```

```
dhyper(0 : 4, 4, 8, 4)
```

```
## [1] 0.141414141 0.452525253 0.339393939 0.064646465 0.002020202
```

```
# (2) Use the sum() function to add five probabilities.
```

```
sum(dhyper(0 : 4, 4, 8, 4))
```

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

Yes, the sum of the five probabilities add to 1.

22. If  $N = 24$ ,  $r = 12$ , and  $n = 6$ , what are the hypergeometric probabilities  $f(x)$  for the following values of  $x$ ? Note: if  $N = 24$  and  $r = 12$ , then  $(N-r) = (24-12) = 12$ .

(a) What is  $p(x \leq 3)$ ?

Answer: 0.6798

Summing the probabilities for  $x = 0, 1, 2$ , and  $3$ , and using

$$f(x) = \frac{\binom{r}{x} \binom{N-r}{n-x}}{\binom{N}{n}}$$

$$p(x \leq 3) = \frac{\binom{12}{0} \binom{12}{6}}{\binom{24}{6}} + \frac{\binom{12}{1} \binom{12}{5}}{\binom{24}{6}} + \frac{\binom{12}{2} \binom{12}{4}}{\binom{24}{6}} + \frac{\binom{12}{3} \binom{12}{3}}{\binom{24}{6}} =$$

$$p(x \leq 3) = \frac{(1)(924)}{134596} + \frac{(12)(792)}{134596} + \frac{(66)(495)}{134596} + \frac{(220)(220)}{134596} =$$

$$p(x \leq 3) = 0.0069 + 0.0706 + 0.2427 + 0.3596 = 0.6798$$

(b) What is  $p(x \geq 4)$ ?

Answer: 0.3202

Summing the probabilities for  $x = 4, 5$ , and  $6$  using

$$f(x) = \frac{\binom{r}{x} \binom{N-r}{n-x}}{\binom{N}{n}}$$

$$p(x \geq 4) = \frac{\binom{12}{4} \binom{12}{2}}{\binom{24}{6}} + \frac{\binom{12}{5} \binom{12}{1}}{\binom{24}{6}} + \frac{\binom{12}{6} \binom{12}{0}}{\binom{24}{6}} =$$

$$p(x \geq 4) = \frac{(495)(66)}{134596} + \frac{(792)(12)}{134596} + \frac{(924)(1)}{134596} =$$

$$p(x \geq 4) = 0.2427 + 0.0706 + 0.0069 = 0.3202$$

(c) What is  $p(2 \leq x \leq 4)$ ?

Answer: 0.8450

Since  $p(2 \leq x \leq 4) = p(x = 2) + p(x = 3) + p(x = 4)$ , and since we know from parts (a) and (b) of this exercise,  $p(x = 2) = 0.2427$ ,  $p(x = 3) = 0.3596$ , and  $p(x = 4) = 0.2427$ , we need only sum these probabilities. That is

$$p(2 \leq x \leq 4) = p(x = 2) + p(x = 3) + p(x = 4) = 0.2427 + 0.3596 + 0.2427 = 0.8450$$

23. Use R to confirm the answers to the previous exercise.

```
# (1) Use the phyper(x, r, N - r, n) function.

N = 24
r = 12
n = 6
x = 3
phyper(x, r, N - r, n)

## [1] 0.6797973

# (2) Use 1 minus the phyper(x, r, N - r, n) function.

N = 24
r = 12
n = 6
x = 3
1 - phyper(x, r, N - r, n)
```

```
## [1] 0.3202027

# (3) From phyper(4, 12, 12, 6) subtract phyper(1, 12, 12, 6).

N = 24
r = 12
n = 6
phyper(4, r, N - r, n) - phyper(1, r, N - r, n)

## [1] 0.8450474
```

24. Referring again to the previous exercise, what is  $E(x)$ ,  $\sigma^2$ , and  $\sigma$ ?

(a)

$$E(x) = n\left(\frac{r}{N}\right) = (6)\left(\frac{12}{24}\right) = 3$$

(b)

$$\text{Var}(x) = \sigma^2 = n\left(\frac{r}{N}\right)\left(1 - \frac{r}{N}\right)\left(\frac{N-n}{N-1}\right) = (6)\left(\frac{12}{24}\right)\left(1 - \frac{12}{24}\right)\left(\frac{24-6}{24-1}\right) = 1.1739$$

(c)

$$\sigma = \sqrt{\sigma^2} = \sqrt{n\left(\frac{r}{N}\right)\left(1 - \frac{r}{N}\right)\left(\frac{N-n}{N-1}\right)} = \sqrt{1.1739} = 1.0835$$

25. A small manufacturer is concerned with the quality of AA batteries being delivered by a certain supplier. Recently, the average battery life has fallen below the manufacturer's standard, and so the manufacturer has begun to test the life of the batteries from this particular supplier on a regular basis. To do this, the quality control inspector randomly selects small samples of batteries from each shipping carton of 100. If even 1 of the batteries fails the test, the entire box of 100 is rejected and returned to the supplier. Suppose the current carton of 100 batteries contains 15 defective batteries.

(a) What is the probability the carton of 100 batteries will be returned if the sample size is  $n = 5$ ?

Answer: 0.5643

One way to approach this question is to re-express it as 1 minus the probability the shipment will not be returned, or 1 minus the probability that no defective batteries are detected. Using

$$f(x) = \frac{\binom{r}{x} \binom{N-r}{n-x}}{\binom{N}{n}}$$

for

$$f(0) = \frac{\binom{15}{0} \binom{85}{5}}{\binom{100}{5}} = \frac{32801517}{75287520} = 0.4357$$

to find the probability that no defectives are found, we need only subtract this probability from 1.

That is, the probability the shipment will be returned is one minus the probability that no defective batteries are detected.

$$= 1 - 0.4357 = 0.5643$$

```
# Use 1 minus dhyper(0, 15, 85, 5).
1 - dhyper(0, 15, 85, 5)
## [1] 0.5643167
```

- (b) What is the probability the carton of 100 batteries will be returned if the sample size is  $n = 10$ ?

Answer: 0.8192

The probability the shipment will not be returned is the probability no defective batteries are detected.

$$f(0) = \frac{\binom{15}{0} \binom{85}{10}}{\binom{100}{10}} = 0.1808$$

The probability the shipment will be returned is one minus the probability that no defective batteries are detected.

$$= 1 - 0.1808 = 0.8192$$

```
# Use 1 minus dhyper(0,15,85,10).
```

```
1 - dhyper(0, 15, 85, 10)
```

```
## [1] 0.8192313
```

- (c) If the quality control department has decided that it wants roughly a 0.85 probability of correctly identifying at least 1 of the 15 defective batteries out of the 100 in the shipping carton, what sample size should they use?

Answer: A sample size of  $n = 11$  results in nearly 0.85 probability that at least one defective item is identified.

Subtract from 1 the probability that no defective items are identified in a sample of size  $n = 11$ .

$$f(0) = \frac{\binom{15}{0}\binom{85}{11}}{\binom{100}{11}} = 0.1506$$

$$1 - 0.1506 = 0.8494$$

```
# Use 1 minus dhyper(0,15,85,11).
```

```
1 - dhyper(0, 15, 85, 11)
```

```
## [1] 0.8493594
```

This result is intuitive in that whenever we use larger sample sizes, the probability that we will uncover a defective battery increases.

26. Suppose we select a random sample of 5 balls from an urn containing 4 red, 3 white, and 2 blue balls. What is the probability that we draw 2 blue and at least 1 red ball? Please use R.

Answer: 0.2698

Sum the probabilities of the different ways one could get 2 blue and at least 1 red ball from the urn.

$$\frac{\binom{3}{2}\binom{4}{1}\binom{2}{2}}{\binom{9}{5}} + \frac{\binom{3}{1}\binom{4}{2}\binom{2}{2}}{\binom{9}{5}} + \frac{\binom{3}{0}\binom{4}{3}\binom{2}{2}}{\binom{9}{5}} = 0.2698$$



```
# Sum the probabilities of 2 blues and 1, 2, and 3 reds.

(choose(3, 2) * choose(4, 1) * choose(2, 2) +
 choose(3, 1) * choose(4, 2) * choose(2, 2) +
 choose(3, 0) * choose(4, 3) * choose(2, 2)) / (choose(9, 5))

## [1] 0.2698413
```