Chapter 4

4.1

If the possible values for x are 0, 1, 2, 3, 4, 5, and the corresponding values for $P(x)$ are 0.2, 0.2, 0.15, 0.3, 0.35, 0.2, 0.1, respectively, does $P(x)$ qualify as a probability function?

```
x <- c(0, 1, 2, 3, 4, 5)
p <- c(0.2, 0.2, 0.15, 0.3, 0.35, 0.2, 0.1)

sum(p)
```

```
[1] 1.5
```

No. P > 1.

4.2

If the possible values for x are 2, 3, 4, and the corresponding values for $P(x)$ are 0.2, -0.1, 0.9, respectively, does $P(x)$ qualify as a probability function?

No. $0 \leq x \leq 1$

4.3

If the possible values for x are 1, 2, 3, 4, and the corresponding values for $P(x)$ are 0.1, 0.15, 0.5, 0.25, respectively, does $P(x)$ qualify as a probability function?

```
x <- c(1, 2, 3, 4)
p <- c(0.1, 0.15, 0.5, 0.25)

sum(p)
```

```
[1] 1
```

Yes, this is a valid probability function.

4.4

If the possible values for x are 2, 3, 4, 5, and the corresponding values for $P(x)$ are 0.2, 0.3, 0.4, 0.1, respectively, what is the probability of observing a value less than or equal to 3.4?

```r
x <- c(2, 3, 4, 5)
p <- c(0.2, 0.3, 0.4, 0.1)

stopifnot( sum(p) == 1 )

prob <- sum( p[ x <= 3.4 ] )
```

Probability: **50%**

4.5

For the previous distribution, what is the probability of observing a 1?

Zero.

4.6

For the previous distribution, what is the probability of observing a value greater than 3?

```r
prob <- sum( p[ x > 3 ])
```

Probability: **50%**

4.7

For the previous distribution, what is the probability of observing a value greater than or equal to 3?

```r
prob <- sum( p[ x >= 3 ])
```

Probability: **80%**

4.8

If the probability of observing a value less than or equal to 6 is 0.3, what is the probability of observing a value greater than 6?

```r
prob <- 1 - .3
```

Probability: **70%**

4.9

For the probability function:

$x \ : \ 0, 1$

$P(x) \ : 0.7, 0.3$

Verify that the mean and variance are 0.3 and 0.21, respectively.

```
x <- c(0, 1)
p <- c(0.7, 0.3)

mu <- sum( x * p )
variance <- sum( (x - mu)^2 * p)
```

$\mu = 0.3, \sigma^2 = 0.21$

What is the probability of getting a value less than the mean?

50%

4.10

Imagine that an auto manufacturer wants to evaluate how potential customers will rate handling for a new car being considered for production. Also, suppose that if all potential customers were to rate handling on a four-point scale, 1 being poor and 4 being excellent, the corresponding probabilities associated with these ratings would be:

$P(1) = 0.2, P(2) = 0.4, P(3) = 0.3, P(4) = 0.1$

Determine the population mean, variance and standard deviation.

```
x <- 1:4
p <- c(0.2, 0.4, 0.3, 0.1)

stopifnot(sum(p) == 1)

mu <- sum(x * p)
variance <- sum( (x - mu)^2 * p)
stdDev <- sqrt(variance)
```

$\mu = 2.3, \sigma^2 = 0.81, \sigma = 0.9$

4.11

If the possible values for x are 1, 2, 3, 4, 5, with probabilities 0.2, 0.1, 0.1, 0.5, 0.1, respectfully, what are the population mean, variance, and standard deviation?

```
x <- 1:5
p <- c(0.2, 0.1, 0.1, 0.5, 0.1)

mu <- sum( x * p )
variance <- sum( (x - mu)^2 * p )
sigma <- sqrt(variance)
```

$\mu = 3.2$, $\sigma^2 = 1.76$, $\sigma = 0.9$

4.12

In the previous exercise, determine the probability of getting a value within one standard deviation of the mean.

That is, $\mu - \sigma \leq x \leq \mu + \sigma$

```
vals <- mu + c(-1, 1)*sigma
round(vals, 4)
```

```
[1] 1.8734 4.5266
```

```
sum( p[ x >= vals[1] & x <= vals[2] ] )
```

```
[1] 0.7
```

4.13

If the possible values for x are 1, 2, 3, with probabilities 0.2, 0.6, and 0.2, respectively, what is the mean and standard deviation?

```
x <- 1:3
p <- c(0.2, 0.6, 0.2)

mu <- sum( x * p )
variance <- sum( (x - mu)^2 * p)
stdDev <- sqrt(variance)
```

$\mu = 2$, $\sigma^2 = 0.4$, $\sigma = 0.6324555$

4.14

In the previous excersize, suppose the possible values for x are now 0, 2, 4 with the same probabilities as before.

Will the standard deviation increase, decrease or stay the same?

Increase.

```
x <- c(0, 2, 4)
p <- c(0.2, 0.6, 0.2)

mu <- sum( x * p )
variance <- sum( (x - mu)^2 * p)
stdDev <- sqrt(variance)
```

$\mu = 2, \sigma^2 = 1.6, \sigma = 1.2649111$

4.15

For the probability function:

$x : 1, 2, 3, 4, 5 \; P(x) : 0.15, 0.2, 0.3, 0.2, 0.15$

Determine the mean, the variance, and the probability that a value is less than the mean.

```
x <- 1:5
p <- c(0.15, 0.2, 0.3, 0.2, 0.15)

mu <- sum( x * p)
variance <- sum( ( x - mu)^2 * p )
stdDev <- sqrt(variance)

sum( p[x < mu] )
```

```
[1] 0.35
```

$\mu = 3, \sigma^2 = 1.6, \sigma = 1.2649111$

4.16

For the probability function:

$x : 1, 2, 3, 4, 5 \; P(x) : 0.1, 0.25, 0.3, 0.25, 0.1$

Would you expect the variance to be larger or smaller than the previous pdf?

Larger.

4.17

For the probability function:

$x : 1, 2, 3, 4, 5 \; P(x) : 0.2, 0.2, 0.2, 0.2, 0.2$

Would you expect the variance to be larger or smaller than the previous pdf?

Smaller.

4.18

For the following probabilities:

Income			
Age	High	Medium	Low
< 30	0.030	0.180	0.090
30-50	0.052	0.312	0.156
Over 50	0.018	0.108	0.054

a.) The probability that someone is under 30.

.03 + 0.18 + 0.09 = .30

b.) The probability that someone has a high income given that they are under 30.

.03 / .3 = .01

c.) The probability of someone having a low income given that they are under 30.

0.09 / .3 = 0.3

d.) The probability of a medium income given that they are over 50.

0.018 + 0.108 + 0.054 = .18

.108 / .18 = .6

4.19

For the previous data, are income and age independent?

Yes.

4.20

Attitude		
Member	1	0
Yes	757	496
No	1,071	1,074

```r
d <- matrix(c(757, 496, 1071, 1074), nrow = 2, byrow = T)

prop.table(data.table(d))
```

```
         V1        V2
1 0.2227781 0.1459682
2 0.3151854 0.3160683
```

a.) Probability of boy choosing “yes”.

.4

b.) P(yes|1)

.22

c.) P(1|yes)

.41

d.) is yes independent of attitude?

No, the probabilities are disproportionate

4.21

Let Y be the cost of a home and let X be a measure of the crime rate. If the variance of the cost of a home changes with X, does this mean that the cost of a home and the crime rate are dependent?

Yes, this can only happen when the conditional probabilites change when told X.

4.22

If the probability of Y < 6 is .4 given that X = 2, and if the probability of Y < 6 is .3 given that X = 4, does this mean that X and Y are dependent?

Yes.

4.23

If the range of possible Y values varies with X, does this mean that X and Y are dependent?

Absolutely.

4.24

For a binomial with n = 10 and p = .4, determine:

a.) $P(0)$

```
dbinom(0, size = 10, prob = .4)
```

```
[1] 0.006046618
```

b.) $P(X \leq 3)$

```
pbinom(3, size = 10, prob = .4)
```

```
[1] 0.3822806
```

c.) $P(X < 3)$

```
pbinom(2, size = 10, prob = .4)
```

```
[1] 0.1672898
```

d.) $P(X > 4)$

```
1 - pbinom(4, size = 10, prob = .4)
```

```
[1] 0.3668967
```

e.) $P(2 \leq X \leq 5)$

```
pbinom(5, size = 10, prob = .4) - pbinom(1, size = 10, prob = .4)
```

```
[1] 0.787404
```

4.25

For a binomial with n = 15 and p = 0.3, determine.

a.) $P(0)$

```
dbinom(x = 0, prob = .3, size = 15)
```

```
[1] 0.004747562
```

b.) $P(X \leq 3)$

```
pbinom( q = 3, prob = .3, size = 15)
```

```
[1] 0.2968679
```

c.) $P(X < 3)$

```
pbinom(2, size = 15, prob = .3)
```

```
[1] 0.1268277
```

d.) $P(X > 4)$

```
pbinom(4, size = 15, prob = .3, lower.tail = F)
```

```
[1] 0.4845089
```

e.) $P(2 \leq X \leq 5)$

```
pbinom(5, size = 15, prob = .3) - pbinom(1, size = 15, prob = .3)
```

```
[1] 0.6863538
```

4.26

For a binomial with n = 15, p = 0.6 determine the probability of exactly 10 successes.

```
dbinom(10, size = 15, prob = .6)
```

```
[1] 0.1859378
```

4.27

For a binomial with n = 7 and p = 0.35, what is the probability of exactly 2 successes?

```r
dbinom(2, size = 7, p = .35)
```

```
[1] 0.2984848
```

4.28

For a binomial with n = 18 and p = 0.6, determine the mean, variance of X, the total number of successes.

```r
n <- 18
p <- 0.6
q <- 1 - p

mu <- n * p
variance <- mu * q
```

$\mu = 10.8, \sigma^2 = 4.32$

4.29

For a binomial with n = 22 and p = .2, determine the mean and variance of X, the total number of successes.

```r
n <- 22
p <- .2
q <- 1 - p

mu <- n * p
variance <- mu * q
```

$\mu = 4.4, \sigma^2 = 3.52$

4.30

For a binomial with n = 20 and p = .7, determine the mean and variance of $\hat{p}$, the proportion of observed success.

```r
n <- 20
p <- .7
q <- 1 - p

mu <- n * p
variance <- mu * q
```

4.31

For a binomial with n = 30 and p = 0.3, determine the mean and variance of $\hat{p}$.

```
n <- 30
p <- .3
q <- 1 - p

phat <- p / n
variance <- p*q / n
```

$\hat{p} = 0.01, \sigma^2 = 0.007$

4.32

For a binomial with n = 10 and p = 0.8, determine:

```
n <- 10
p <- 0.8
q <- 1 - p

variance <- p*q / n
```

a.) the probability that $\hat{p}$ is less than or equal to 0.7.

b.) the probability that $\hat{p}$ is greater than or equal to 0.8.

c.) the probability that $\hat{p}$ is exactly equal to 0.8.

4.33

A coin is rigged so that when it is flipped, the probability of a head is 0.7. If the coin is flipped three times, which is the more likely outcome, exactly three heads or two heads and a tail?

```
dbinom(3, 3, .7) # 3 heads
```

```
[1] 0.343
```

```
dbinom(2, 3, .7) # 2 heads 1 tail
```

```
[1] 0.441
```

Two heads, 1 tail.

4.34

Imagine that the probability of heads when flipping a coin is given by the binomial probability function with p = 0.5.

If you flip the coin nine times and get nine heads, what is the probability of a head on the 10th flip?

```
# independent events.
dbinom(1, 1, .5)
```

```
[1] 0.5
```

4.35

The Department of Agriculture of the United States reports that 75% of all people who invest in the futures market lose money. Based on the binomial probability function, with n = 5, determine:

a.) the probability that all 5 lose money.

$P(x) = 0$

```
dbinom(5, size = 5, prob = .75)
```

```
[1] 0.2373047
```

b.) the probability that all five make money.

```
dbinom(5, size = 5, prob = .25)
```

```
[1] 0.0009765625
```

c.) the probability that at least two lose money.

```
pbinom(q = 3, size = 5, prob = .25)
```

```
[1] 0.984375
```

4.36

If for a binomial distribution p = 0.4 and n = 25, determine:

```
n <- 25
p <- .4
q <- 1 - p
```

a.) $P(X < 11)$

```
pbinom(10, size = n, prob = p)
```

```
[1] 0.585775
```

b.) $P(X \leq 11)$

```
pbinom(11, size = n, prob = p)
```

```
[1] 0.7322822
```

c.) $P(X > 9)$

```
pbinom(9, size = n, prob = p, lower.tail = F)
```

```
[1] 0.575383
```

d.) $P(X \geq 9)$

```
pbinom(8, size = n, prob = p, lower.tail = F)
```

```
[1] 0.7264685
```

4.37

In the previous problem, determine the mean of X, the variance of X, the mean of $\hat{p}$, and the variance of $\hat{p}$.

```
mu <- n * p
variance <- mu * q

phat <- p
v <- p*q/n
```

$\mu = 10$, $\sigma^2 = 6$, $\hat{p} = 0.4$, $\sigma^2 = 0.0096$

4.38

Given that Z has a standard normal distribution, determine:

a.) $P(Z \geq 1.5)$

```
pnorm(1.5, lower.tail = F)
```

```
[1] 0.0668072
```

b.) $P(Z \leq -2.5)$

```
pnorm(-2.5)
```

```
[1] 0.006209665
```

c.) $P(Z < -2.5)$

```
pnorm(-2.5)
```

```
[1] 0.006209665
```

d.) $P(-1 \leq Z \leq 1)$

```
# P(Z > -1) - P(Z > 1)
pnorm(-1, lower.tail = F) - pnorm(1, lower.tail = F)
```

```
[1] 0.6826895
```

```
# 1 - 2*tail_area
1 - 2 * pnorm(1, lower.tail = F)
```

```
[1] 0.6826895
```

4.39

If Z has a standard normal distribution, determine:

a.) $P(Z \leq 0.5)$

```r
pnorm(0.5)
```

```
[1] 0.6914625
```

b.) $P(Z > -1.25)$

```r
pnorm(-1.25, lower.tail = F)
```

```
[1] 0.8943502
```

c.) $P(-1.2 < Z < 1.2)$

```r
1 - 2 *pnorm(1.2, lower.tail = F)
```

```
[1] 0.7698607
```

d.) $P(-1.8 \leq Z \leq 1.8)$

```r
pnorm(1.8, lower.tail = T) - pnorm(-1.8)
```

```
[1] 0.9281394
```

4.40

If Z has a standard normal distribution, determine:

a.) $P(Z < -.5)$

```r
1 - pnorm(-.5, lower.tail = F)
```

```
[1] 0.3085375
```

b.) $P(Z < 1.2)$

```r
1 - pnorm(1.2, lower.tail = F)
```

```
[1] 0.8849303
```

c.) $P(Z > 2.1)$

```r
pnorm(2.1, lower.tail = F)
```

```
[1] 0.01786442
```

d.) $P(-.28 < Z < 0.28)$

```r
pnorm(-.28, lower.tail = F) - pnorm(.28, lower.tail = F)
```

```
[1] 0.2205225
```

4.41

If Z has a standard normal distribution, find *c* such that:

a.) $P(Z \leq c) = 0.0099$

```r
qnorm(0.0099)
```

```
[1] -2.330116
```

b.) $P(Z < c) = .9732$

```r
qnorm(.9732)
```

```
[1] 1.930055
```

c.) $P(Z > c) = 0.5691$

```r
qnorm(.5691, lower.tail = F)
```

```
[1] -0.1740833
```

d.) $P(-c \leq Z \leq c) = 0.2358$

```r
qnorm( (1 + 0.2358) / 2)
```

```
[1] 0.2999701
```

4.42

If Z has a standard normal distribution with, determine:

a.) $P(Z > c) = 0.0764$

```r
qnorm(0.0764, lower.tail = F)
```

```
[1] 1.429711
```

b.) $P(Z > c) = 0.5040$

```r
qnorm(0.504, lower.tail = F)
```

```
[1] -0.01002668
```

c.) $P(-c \leq Z \leq c) = 0.9108$

```r
qnorm((1 + 0.9108)/2)
```

```
[1] 1.699633
```

d.) $P(-c \leq Z \leq c) = 0.8$

```r
qnorm((1+.8)/2)
```

```
[1] 1.281552
```

4.43

If X has a normal distribution with mean $\mu = 50$ and standard deviation $\sigma = 9$

a.) $P(X \leq 40)$

```r
pnorm(40, mean = 50, sd = 9)
```

```
[1] 0.1332603
```

b.) $P(X < 55)$

```r
1 - pnorm(55, mean = 50, sd = 9, lower.tail = F)
```

```
[1] 0.7107426
```

c.) $P(X > 60)$

```r
pnorm(60, mean = 50, sd = 9, lower.tail = F)
```

```
[1] 0.1332603
```

d.) $P(40 \leq X \leq 60)$

```r
pnorm(60, mean = 50, sd = 9) - pnorm(40, mean = 50, sd = 9)
```

```
[1] 0.7334795
```

4.44

If X has a normal distribution with $\mu = 20$ and $\sigma = 9$, determine:

a.) $P(X < 22)$

```r
1 - pnorm(22, mean = 20, sd = 9, lower.tail = F)
```

```
[1] 0.5879296
```

b.) $P(X > 17)$

```r
pnorm(17, mean = 20, sd = 9, lower.tail = F)
```

```
[1] 0.6305587
```

c.) $P(X > 15)$

```r
pnorm(15, mean = 20, sd = 9, lower.tail = F)
```

```
[1] 0.7107426
```

d.) $P(2 \leq X \leq 38)$

```r
pnorm(38, mean = 20, sd = 5) - pnorm(2, mean = 20, sd = 9)
```

```
[1] 0.9770908
```

4.45

If X has a normal distribution with mean $\mu = .75$ and standard deviation $\sigma = 0.5$, determine:

a.) $P(0.5 < X < 1)$

```
pnorm(1, mean = .75, sd = .5) - pnorm(.5, mean = .75, sd = .5)
```

```
[1] 0.3829249
```

b.) $P(0.25 < X < 1.25)$

```
pnorm(1.25, mean = .75, sd = .5) - pnorm(.25, mean = .75, sd = .5)
```

```
[1] 0.6826895
```

4.46

If X has a normal distribution, determine *c* such that:

$P(\mu - c\sigma < X < \mu + c\sigma) = .95$

```
qnorm((1 +.95)/2)
```

```
[1] 1.959964
```

4.47

If X has a normal distribution, determine *c* such that:

$P(\mu - c\sigma < X < \mu + c\sigma) = .8$

```
qnorm((1 +.8)/2)
```

```
[1] 1.281552
```

4.48

Assuming that the scores on a math achievement test are normally distributed with $\mu = 68$ and standard deviation $\sigma = 10$, what is the probability of getting a score greater than 78?

```
pnorm(78, mean = 68, sd = 10, lower.tail = F)
```

```
[1] 0.1586553
```

4.49

In the previous problem, how high must someone score to be in the top 5%?

That is, determine *c* such that $P(X > c) = 0.05$

```
qnorm(1 - 0.05, mean = 68, sd = 10)
```

```
[1] 84.44854
```

4.50

A manufacturer of car batteries claims that the life of their batteries is normally distributed with mean $\mu = 58$ and $\sigma = 3$.

Determine the probability that a randomly selected battery will last at least 62 months.

```
pnorm(62, mean = 58, sd = 3, lower.tail = F) # more than 62
```

```
[1] 0.09121122
```

4.51

Assume that the income of pediatricials is normally distributed with mean $\mu = \$100,000$ and $\sigma = 10,000$.

Determine the probability of observing an income between $85,000 and $115,000.

```
pnorm(1.15, mean = 1, sd = .1) - pnorm(.85, mean = 1, sd = .1)
```

```
[1] 0.8663856
```

4.52

Suppose the winnings of gamblers at Las Vegas are normally distributed with $\mu = -300$ and $\sigma = 100$.

Determine the probability that a gambler does not lose any money.

```
pnorm(0, mean = -300, sd = 100, lower.tail = F)
```

```
[1] 0.001349898
```

4.53

A large computer company claims that their salaries are normally distributed with $\mu = \$50,000$ and $\sigma = \$10,000$.

What is the probability of observing an income between $40,000 and $60,000?

```
pnorm(6, mean = 5, sd = 1) - pnorm(4, mean = 5, sd = 1)
```

```
[1] 0.6826895
```

4.54

Suppose the daily amount of solar radiation in Los Angeles is normally distributed with mean 450 and sd 50.

Determine the probability that for a given day the radiation is between 350 and 550.

```
pnorm(5.5, mean = 4.5, sd = .5) - pnorm(3.5, mean = 4.5, sd = .5)
```

```
[1] 0.9544997
```

4.55

If the cholesterol levels of adults are normally distributed with mean 230 and standard deviation 25, what is the probability that a randomly sampled adult has a cholesterol level greater than 260?

```
pnorm(2.6, mean = 2.3, sd = .25, lower.tail = F)
```

```
[1] 0.1150697
```

4.56

If after 1 year, the annual mileage of privately owned cars is normally distributed with mean 14,000 miles and sd 3,500, what is the probability a car has greater than 20,000 miles?

```
pnorm(20, mean = 14, sd = 3.5, lower.tail = F)
```

```
[1] 0.04323813
```

4.57

Can small changes in the tails of a distribution result in large changes in the population mean, μ, relative to the changes in median?

Yes, the mean is heavly influenced by the tails, where as the median is not.

4.58

Explain in what sense the population variance is sensitive to small changes in a distribution.

The variance is sensetive to small changes in the tail.

4.59

For normal random variables, the probability of being within one standard deviation of the mean is .68. That is, $P(\mu - \sigma \leq X \leq \mu + \sigma) = .68$, if X has a normal distribution.

For nonnormal distribution, is it safe to assume that this probability is again .68?

No. The AUC (and therefore the mean/variance relationship) for a distribution is defined by its pdf, $P(X) = F_x$, *which will be unique per distribution.*

4.60

If a distribution appears to be bell-shaped and symmetric about its mean, can we assume that the probability of being within one sd of the mean is .68?

No.

4.61

Can two distribution differ by a large amount yet have equal means and variances?

Yes.

4.62

If a distribution is skewed, is it possible that the mean exceedes the .85 quantile?

Yes.