Section 6.5 Summary and Review Exercises

6.66 b  6.67 a  6.68 d  6.69 a  6.70 c  6.71 a  6.72 b

6.73 Because it is the hypothesis that the researcher wishes to support.

6.74 $H_0: \pi = .70$ versus $H_a: \pi > .70$.

6.75 $H_0: \mu = 5.5$ versus $H_a: \mu \neq 5.5$.

6.76 $H_0: \mu = 2,000$ versus $H_a: \mu > 2,000$.

> tsum.test(mean.x=2080,s.x=240,n.x=60,mu=2000,alternative="greater")

One-sample t-Test

data: Summarized x
t = 2.582, df = 59, p-value = 0.006163
alternative hypothesis: true mean is greater than 2000
95 percent confidence interval:
  2028.223       NA
sample estimates:
mean of x
  2080

There is significant evidence that the county agent is right and his county needs more than 2000 pounds a lime per acre.

6.77 $H_0: \mu = 15$ versus $H_a: \mu > 15$.

> tsum.test(mean.x=17.8,s.x=6.2,n.x=24,mu=15,alternative="greater")

One-sample t-Test

data: Summarized x
t = 2.2124, df = 23, p-value = 0.01857
alternative hypothesis: true mean is greater than 15
95 percent confidence interval:
  15.63098       NA
sample estimates:
mean of x
  17.8

There is significant evidence that the politicians scored above the norm.
6.78 \quad H_0: \pi = .70 \text{ versus } \quad H_a: \pi \neq .70.

> prop.test(x=154, n=200, p=.70)

         1-sample proportions test with continuity correction

data:  154 out of 200, null probability 0.7
X-squared = 4.3393, df = 1, p-value = 0.03724
alternative hypothesis: true p is not equal to 0.7
95 percent confidence interval: 
  0.7042503 0.8251428
sample estimates:
  p
0.77

Based on a random sample of 200 convicted felons there is significant evidence to contradict the claim that 70% have a history of juvenile delinquency.

6.79

> data(Step)
> str(Step)
'data.frame': 12 obs. of 1 variable:
$ score: int 58 60 82 80 67 70 65 73 75 77 ...
> attach(Step)
> ntester(score)

Simulated Normal Data on Perimeter - Actual Data in Center

The normal probability plot shows that the data are close to normal and therefore a t-test is permissible.
\( H_0: \mu = 80 \) versus \( H_a: \mu < 80. \)

\[
> \text{t.test(score, mu=80, alternative="less")}
\]

One Sample t-test

data: score
t = -3.6526, df = 11, p-value = 0.001902
alternative hypothesis: true mean is less than 80
95 percent confidence interval:
    -Inf 75.63684
sample estimates:
mean of x
    71.41667

Based on a random sample of 12 ability-grouped students there is highly significant evidence that the group falls below the national median (80).

6.80 \( H_0: \pi = .58 \) versus \( H_a: \pi < .58. \)

\[
> \text{prop.test(x=17,n=32,p=.58, alternative="less")}
\]

1-sample proportions test with continuity correction
data: 17 out of 32, null probability 0.58
X-squared = 0.1441, df = 1, p-value = 0.3521
alternative hypothesis: true p is less than 0.58
95 percent confidence interval:
0.000000 0.682258
sample estimates:
p
0.53125

Based on a random sample of 32 awarded PhDs there is insignificant evidence that the percent awarded to US students is unusually low.

6.81 \( H_0: \mu = 4 \) versus \( H_a: \mu < 4. \)

\[
> \text{tsum.test(mean.x=3.6, s.x=1.5, n.x=32, mu=4, alternative="less")}
\]

One-sample t-Test
data: Summarized x
t = -1.5085, df = 31, p-value = 0.07078
alternative hypothesis: true mean is less than 4
95 percent confidence interval:
    NA 4.049592
sample estimates:
mean of x
    3.6

Moderately significant evidence exists that the mean life is less than 4 years.
6.82 $H_0: \mu = 40$ versus $H_a: \mu > 40$.

6.83 A type I error will occur if the psychologist decides the subjects average memorizing more than 40 correct phrases when in fact they average 40 or less correct phrases. A type II error will occur if the psychologist decides the subjects average memorizing 40 or less phrases correctly when in fact they average more than 40.

6.84 $H_0: \mu = 800$ versus $H_a: \mu > 800$.

```r
> zsum.test(mean.x=848,sigma.x=200,n.x=45,mu=800,alternative="greater")

One-sample z-Test
data:  Summarized x
z = 1.61, p-value = 0.0537
alternative hypothesis: true mean is greater than 800
95 percent confidence interval:  
 798.96     NA
sample estimates:  
mean of x
  848
```

Based on 45 randomly selected property owners there is moderately significant evidence to reject the government's claim.

6.85 $H_0: \mu = 65$ versus $H_a: \mu > 65$.

```r
> zsum.test(mean.x=68.2,sigma.x=10,n.x=81,mu=65,alternative="greater")

One-sample z-Test
data:  Summarized x
z = 2.88, p-value = 0.001988
alternative hypothesis: true mean is greater than 65
95 percent confidence interval:  
66.37238       NA
sample estimates:  
mean of x
  68.2
```

Based on a sample of 81 freshmen there is highly significant evidence that State University freshmen are more intelligent than the average freshman.

6.86 $H_0: \mu = 36$ versus $H_a: \mu < 36$.

```r
> tsum.test(mean.x=33.6,s.x=4.8,n.x=36,mu=36,alternative="less")

One-sample t-Test
data:  Summarized x
t = -3, df = 35, p-value = 0.002474
alternative hypothesis: true mean is less than 36
95 percent confidence interval: 
NA 34.95166
sample estimates: 
mean of x
  33.6
```
Based on a random sample of 36 women there is highly significant evidence that women on the pill have a smaller maximal oxygen uptake than women not on the pill.

6.87 $H_0: \mu = 50$ versus $H_a: \mu < 50$.

> tsum.test(mean.x=42,s.x=15,n.x=9,mu=50,alternative="less")

One-sample t-Test
data:  Summarized x
t = -1.6, df = 8, p-value = 0.07413
alternative hypothesis: true mean is less than 50
95 percent confidence interval:
  NA 51.29774
sample estimates:
mean of x
  42

Based on a random sample of 9 slow-ability children there is moderately significant evidence that their mean score is less than 50. For the t-test to be valid it must be assumed that the distribution of scores for slow learners is approximately normal.

6.88 $H_0: \mu = 65.42$ versus $H_a: \mu < 65.42$.

> tsum.test(mean.x=64.82,s.x=2.32,n.x=144,mu=65.42,alternative="less")

One-sample t-Test
data:  Summarized x
t = -3.1034, df = 143, p-value = 0.001153
alternative hypothesis: true mean is less than 65.42
95 percent confidence interval:
  NA 65.14008
sample estimates:
mean of x
  64.82

Based on a random sample of 144 adults there is highly significant evidence that the mean adult height of residents in the depressed area is below that of all residents.

6.89 $H_0: \mu = 250$ versus $H_a: \mu > 250$.

> tsum.test(mean.x=263,s.x=72,n.x=100,mu=250,alternative="greater")

One-sample t-Test
data:  Summarized x
t = 1.8056, df = 99, p-value = 0.03701
alternative hypothesis: true mean is greater than 250
95 percent confidence interval:
  251.0452       NA
sample estimates:
mean of x
  263

Based on a random sample of 100 accounts there is significant evidence to cast doubt on the utility company's claim.
6.90 \( H_0: \mu = 15.9 \) versus \( H_a: \mu > 15.9 \)

\[
> \text{zsum.test(mean.x=16.2,\sigma.x=2.4,n.x=64,\mu=15.9,alternative="greater")}
\]

One-sample z-Test

data:  Summarized x
z = 1, p-value = 0.1587
alternative hypothesis: true mean is greater than 15.9
95 percent confidence interval:
     15.70654       NA
sample estimates:
mean of x
    16.2

Based on a random sample of 64 strings there is insufficient evidence that the new string has a greater breaking strength than the old.

6.91 \( H_0: \pi \geq 0.7 \) versus \( H_a: \pi < 0.7 \)

\[
> \text{prop.test(x=132,n=200,\text{p}=.70,alternative="less")}
\]

1-sample proportions test with continuity correction

data:  132 out of 200, null probability 0.7
X-squared = 1.3393, df = 1, p-value = 0.1236
alternative hypothesis: true p is less than 0.7
95 percent confidence interval:
     0.000000 0.7150033
sample estimates:
    \ p
    0.66

Based on a random sample of 200 people there is insufficient evidence to conclude that less than 70% favor raising the drinking age.

6.92 \( H_0: \mu = 1.5 \) versus \( H_a: \mu \neq 1.5 \).

\[
> \text{tsum.test(mean.x=1.56,\sigma.x=.09,n.x=16,\mu=1.5)}
\]

One-sample t-Test

data:  Summarized x
t = 2.6667, df = 15, p-value = 0.01760
alternative hypothesis: true mean is not equal to 1.5
95 percent confidence interval:
     1.512042 1.607958
sample estimates:
mean of x
    1.56

Based on a random sample of 16 parts there is significant evidence that the shipment should be rejected.
6.93  \( H_0: \theta \leq 98.5 \quad \text{versus} \quad H_a: \theta > 98.5 \)

```r
> data(Stable)
> str(Stable)
'data.frame': 9 obs. of 1 variable:
$ time: num 104.6 98.8 101.4 98.2 99.7 ...
> attach(Stable)
> sign.test(time,md=98.5,alternative="greater")
$rval

One-sample Sign-Test
data:  time
s = 8, p-value = 0.01953
alternative hypothesis: true median is greater than 98.5
95 percent confidence interval:
 98.74333      Inf
sample estimates:
mожет of x
101.4

$Confidence.Intervals

<table>
<thead>
<tr>
<th>Conf.Level</th>
<th>L.E.pt</th>
<th>U.E.pt</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lower Achieved CI</td>
<td>0.9102</td>
<td>98.8000 Inf</td>
</tr>
<tr>
<td>Interpolated CI</td>
<td>0.9500</td>
<td>98.7433 Inf</td>
</tr>
<tr>
<td>Upper Achieved CI</td>
<td>0.9805</td>
<td>98.7000 Inf</td>
</tr>
</tbody>
</table>

Based on 9 trials there is significant evidence that the median time is more than 98.5 seconds, therefore you reject the horse. Because his lifetime median time is 100 seconds you did not make an error.

6.94  \( H_0: \mu = 200 \quad \text{versus} \quad H_a: \mu \neq 200 \)

```r
> zsum.test(mean.x=215,sigma.x=50,n.x=100,mu=200)

One-sample z-Test
data:  Summarized x
z = 3, p-value = 0.0027
alternative hypothesis: true mean is not equal to 200
95 percent confidence interval:
 205.2002 224.7998
sample estimates:
mean of x
215

Based on a sample of 100 students there is highly significant evidence that students from this high school score unusual high on the pre-admission algebra exam.
> data(Phone)
> str(Phone)
'data.frame': 20 obs. of 1 variable:
$ time: num 12.8 3.5 2.9 9.4 8.7 3.5 4.8 7.7 5.9 6.2 ...
> attach(Phone)
> sign.test(time,md=5,alternative="greater")
$rval

One-sample Sign-Test

data:  time
s = 9, p-value = 0.7483
alternative hypothesis: true median is greater than 5
95 percent confidence interval:
 3.058559      Inf
sample estimates:
median of x
  4.75

$Confidence.Intervals

Conf.Level L.E.pt U.E.pt
Lower Achieved CI  0.9423 3.1000 Inf
Interpolated CI    0.9500 3.0586 Inf
Upper Achieved CI  0.9793 2.9000 Inf

a. $H_0$: $\theta \leq 5$ versus $H_a$: $\theta > 5$
b. This is a right-tailed test.
c. The test statistic $s = 9$.
d. This is a small sample test.
e. Because this is a right-tailed test we expect a large number of observations greater than 5, the hypothesized value of the median. Less than half, however, are greater than 5. In other words the observed value of the test statistic has fallen on the left tail of the sampling distribution instead of the right tail, making the p-value greater than 50%.
f. The null hypothesis should not be rejected.
g. Based on a random sample of 20 times there is insufficient evidence to conclude that the length of long-distance phone calls for the small business firm exceeds 5 minutes.

6.96

> data(Counsel)
> str(Counsel)
'data.frame': 18 obs. of 1 variable:
$ score: int 68 71 75 65 61 70 70 64 71 73 ...
> attach(Counsel)
> EDA(score)
[1] "score"
Size (n)  Missing  Minimum   1st Qu     Mean   Median
18.000    0.000   61.000   66.500   69.500   70.000
TrMean  3rd Qu     Max.   Stdev.     Var. SE Mean
69.500   72.250   78.000    4.579   20.971   1.079
I.Q.R. Range Kurtosis Skewness SW p-val
5.750 17.000  -0.742 -0.111  0.907
EXPLORATORY DATA ANALYSIS

The EDA analysis support the assumptions for a t-test.

\[ H_0: \mu = 70 \quad \text{versus} \quad H_a: \mu < 70. \]

\[
> \text{t.test(score,mu=70,alternative="less")}
\]  

One Sample t-test

data: score  
t = -0.4632, df = 17, p-value = 0.3245  
alternative hypothesis: true mean is less than 70  
95 percent confidence interval:   
   -Inf 71.37767  
sample estimates:   
   mean of x  
   69.5

Based on the results obtained from 18 volunteers there is insignificant evidence that the counseling process reduces one’s score on the exam.

6.97

\[
> \text{data(Earthqk)}
\]  

\[
> \text{str(Earthqk)}
\]  

`data.frame`: 100 obs. of 2 variables:

\$ year : int 1770 1771 1772 1773 1774 1775 1776 1777 1778 1779 ...  
\$ severity: int 66 62 66 197 63 0 121 0 113 27 ...

\[
> \text{attach(Earthqk)}
\]  

\[
> \text{hist(severity,col="pink",main="Histogram for part a."})
\]
a. The histogram looks reasonably symmetric and possibly normally distributed.

```r
> par(mfrow=c(1,2)) # Splitting screen for two graphs
> boxplot(severity,col="green",main="Boxplot for part b")
> qqnorm(severity,col="blue",main="QQ-plot for part b")
> qqline(severity,col="red")
```
b. The boxplot and QQ-plot show that the distribution is slightly skewed right.

c. The deviation from symmetry is not enough to warrant a study of the population median. There should be very little difference between the mean and median.

d.

```r
> t.test(severity,mu=100,alternative="greater")

One Sample t-test

data: severity
t = -0.4541, df = 99, p-value = 0.6746
alternative hypothesis: true mean is greater than 100
95 percent confidence interval:
 90.54745      Inf
sample estimates:
mean of x
 97.97

The sample mean is less than 100 and therefore we certainly would not reject the null hypothesis in favor of the alternative that says that $\mu$ is greater than 100. This is also indicated by the p-value = .67.

6.98

```r
> par(mfrow=c(1,2)) # Splitting the screen for two graphs
> data(Grnriv2)
> str(Grnriv2)
'data.frame': 101 obs. of 1 variable:
$ thick: num 6 7.2 7.1 7.1 7.2 7.4 8 8.6 10 11.4 ...
> attach(Grnriv2)
> boxplot(thick,col="blue",main="Histogram of Varve Thickness")
> qqnorm(thick,col="green",main="QQ-plot of Varve Thickness")
> qqline(thick)

The histogram and QQ-plot show that the distribution is definitely skewed right. A good option is to use the function EDA in the BSDA library.
EXPLORATORY DATA ANALYSIS

b. The rules for applying the t-test say that when the sample size is large, moderate departures from normality are permissible. Because of the outliers in the right tail of the distribution, however, a test of the median thickness may be more appropriate.

c.

> sign.test(thick, md=8, alternative="less")

$rval$

One-sample Sign-Test

data: thick
s = 38, p-value = 0.01049
alternative hypothesis: true median is less than 8
95 percent confidence interval:
-Inf 7.8
sample estimates:
median of x
 7.3
Based on these results there is significant evidence that the median varve thickness is less than 8 millimeters. This is an instance where a test of the population median found a significant difference when the test of the population mean failed to detect a difference.

6.99

> data(Schizop2)
> str(Schizop2)
'data.frame': 17 obs. of 1 variable:
  $ score: int 36 29 30 32 37 15 34 23 32 5 ...
> attach(Schizop2)
> par(mfrow=c(1,1))
> hist(score,col="gold",main="Histogram of Scores")

![](https://example.com/histogram.png)

a. The histogram is highly skewed left.

> par(mfrow=c(1,2))
> boxplot(score,col="brown",main="Histogram for part b.")
> qqnorm(score,col="red",main="QQ-plot part b.")
> qqline(score,col="blue")

![](https://example.com/boxplot.png)

![](https://example.com/qqplot.png)

b. The boxplot and normal QQ-plot also show a skewed left distribution.
c. The median exam best represents a typical measurement.

d. $H_0: \theta \leq 22$ versus $H_a: \theta > 22$

e.

```r
> sign.test(score,md=22,alternative="greater")
$rval

One-sample Sign-Test

data: score
s = 13, p-value = 0.02452
alternative hypothesis: true median is greater than 22
95 percent confidence interval:
  23.53969      Inf
sample estimates:
  median of x
    30

$Confidence.Intervals
Conf.Level  L.E.pt U.E.pt
Lower Achieved CI  0.9283 24.0000    Inf
Interpolated CI  0.9500 23.5397    Inf
Upper Achieved CI  0.9755 23.0000    Inf
```

Based on the test scores of these 17 patients there is evidence that the tranquilizer significantly improved the amount of learning exhibited by schizophrenics.

6.100

```r
> data(Marked)
> str(Marked)
'data.frame': 65 obs. of 1 variable:
$ percent: int 61 74 49 65 62 62 73 71 37 54 ... 
> attach(Marked)
> hist(percent,col="aquamarine",main="Histogram of Percentage of Unmarked Police Cars")
```

a. The histogram is skewed right.
> EDA(percent)
[1] "percent"
Size (n)  Missing  Minimum    1st Qu    Mean    Median
   65.000    0.000    37.000    54.000   61.108   60.000
TrMean   3rd Qu    Max.    Stdev.    Var.    SE Mean
  60.932   68.000   92.000    9.980   99.598    1.238
I.Q.R.    Range Kurtosis Skewness SW p-val
   14.000   55.000    0.273    0.443    0.072

EXPLORATORY DATA ANALYSIS

b. The boxplot indicates the distribution is slightly skewed to the right since the mean is further to the right than is the median as well as one outlier on the right tail. The normal probability plot exhibits some sinusoidal behavior that suggest that the data are not normally distributed.

c. Since there is only a slight skewness and the sample size is relatively large, the parameter of interest would be the population mean $\mu$.

d. $H_0: \mu = 60$ versus $H_a: \mu > 60$.

e. Because their are no extreme outliers and the data are reasonably symmetric a test based on $\bar{y}$ would be appropriate.

> t.test(percent,mu=60,alternative="greater")

One Sample t-test
data:  percent
t = 0.8949, df = 64, p-value = 0.1871 alternative hypothesis: true mean is greater than 60 95 percent confidence interval:
  59.0417  Inf sample estimates:
mean of x
  61.10769

Based on information provided by 65 police departments there is insignificant evidence that the mean percent of marked cars exceeds 60.