# Homework #6

*Derek Sonderegger*

```
# load my usual libraries
library(dplyr)
library(ggplot2)
```

1. One way the amount of sewage and industrial pollutants dumped into a body of water affects the health of the water is by reducing the amount of dissolved oxygen available for aquatic life. Over a 2-month period, 8 samples were taken from a river at a location 1 mile downstream from a sewage treatment plant. The amount of dissolved oxygen in the samples was determined and is reported in the following table. Current research suggests that the mean dissolved oxygen level must be at least 5.0 parts per million (ppm) for fish to survive.

   a) Use R to calculate the sample mean and standard deviation.

| Sample | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| Oxygen (ppm) | 5.1 | 4.9 | 5.6 | 4.2 | 4.8 | 4.5 | 5.3 | 5.2 |

```
data <- data.frame(O2 = c(5.1, 4.9, 5.6, 4.2, 4.8, 4.5, 5.3, 5.2))
xbar <- mean( data$O2 )
s    <- sd( data$O2 )
cbind( xbar, s)
```

```
##      xbar         s
## [1,] 4.95 0.4503967
```

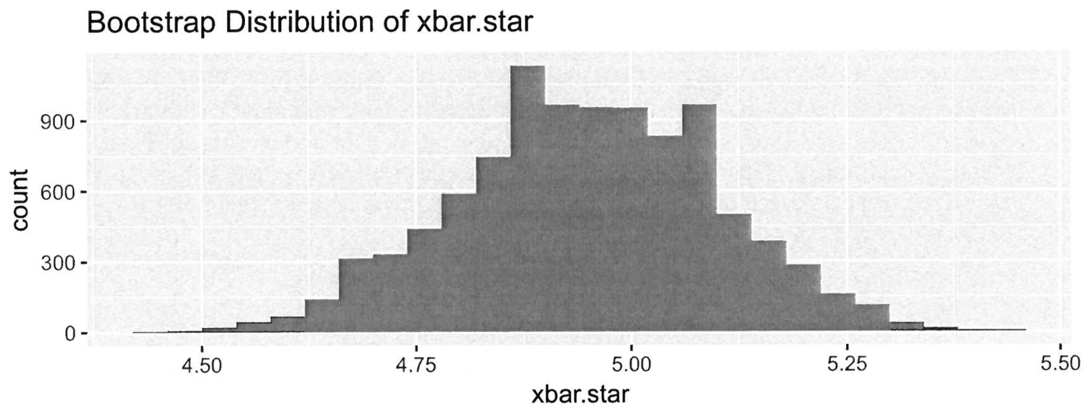
   b) Using the asymptotic results and the quantities you calculated, create a 95% two-sided confidence interval for the mean dissolved oxygen level during the 2-month period. What assumption is being made for this calculation to be valid?

   *In order for the following calculations to be appropriate, the sample mean the distribution of $\bar{X}$ must be approximately normal. That will be the case if the original data came from a normal distribution or if the sample size is large enough for the Central Limit Theorem to guarantee that the distribution of the sample mean is approximately normally distributed. Because the sample size is small, the we have to hope that the originial population isn't too far from normal and the CLT helps a bit.*

$$4.95 \pm t_7^{0.975}\left(\frac{0.4504}{\sqrt{8}}\right)$$
$$4.95 \pm 2.3646\left(\frac{0.4504}{\sqrt{8}}\right)$$
$$4.95 \pm 0.37654$$
$$(4.57,\ 5.33)$$

   c) Calculate a 95% two-sided confidence interval using the bootstrap method. Examine the histogram of bootstrap means, does it appear normal? If so, what does that imply about the assumption you made in the calculation in the previous part?

```
BootDist <- mosaic::do(10000)*{ data %>%
    mosaic::resample() %>%
    summarise(xbar.star = mean(O2))
}
ggplot(BootDist, aes(x=xbar.star)) + geom_histogram(binwidth=0.04) +
  ggtitle('Bootstrap Distribution of xbar.star')
```




```
quantile(BootDist$xbar.star, probs=c(0.025, 0.975))
```

```
##   2.5%  97.5%
## 4.6500 5.2375
```

*Because the bootstrap distribution of $\bar{X}$ looks quite normal, the assumption made in part (b) is likely true and we see that the confidence intervals are quite similar.*

d) Using the confidence interval calculated in part (b), do the data support the hypothesis that the mean dissolved oxygen level is equal to 5 ppm?

*Because the 95% CI also includes 5 ppm, then that is a reasonable value for $\mu$ given the data we've seen.*

e) Perform a 1-sided hypothesis test that the mean oxygen level is less that 5 ppm with a significance level of $\alpha = 0.05$.

$$t_{n-1} = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}}$$

$$t_7 = \frac{4.95 - 5.0}{0.4504/\sqrt{8}}$$

$$= -0.3140$$

$$\text{p.value} = P\left(T_7 \leq -0.3140\right)$$

```
p.value = pt(-0.314, df=7)
p.value
```

```
## [1] 0.381336
```

f) Use the function t.test in R to repeat the calculations you made in parts (b) and (e).

```
t.test( data$O2, mu=5 )   # for the 95% confidence interval
```

```
##
##  One Sample t-test
##
## data:  data$O2
## t = -0.31399, df = 7, p-value = 0.7627
## alternative hypothesis: true mean is not equal to 5
## 95 percent confidence interval:
##  4.573459 5.326541
## sample estimates:
## mean of x
##      4.95
```

```
t.test( data$O2, mu=5, alternative='less')  # for the p.value
```

```
##
##  One Sample t-test
##
## data:  data$O2
## t = -0.31399, df = 7, p-value = 0.3813
## alternative hypothesis: true mean is less than 5
## 95 percent confidence interval:
##      -Inf 5.251691
## sample estimates:
## mean of x
##      4.95
```

2. We are interested in investigating how accurate radon detectors sold to homeowners are. We take a randomly selection of $n = 12$ detectors and expose to 105 pico-curies per liter (pCi/l) of radon. The following values were given by the radon detectors. Do all of the following calculations by hand (except for the calculations of the mean and standard deviation).

| | | | | | |
|---|---|---|---|---|---|
| 91.9 | 97.8 | 111.4 | 122.3 | 105.4 | 95.0 |
| 103.8 | 99.6 | 96.6 | 119.3 | 104.8 | 101.7 |

a) Calculate a 90% confidence interval using the asymptotic method.

```
data <- data.frame(
  radon = c( 91.9, 97.8, 111.4, 122.3, 105.4,  95.0,
            103.8, 99.6,  96.6, 119.3, 104.8, 101.7  ))
xbar <- mean(data$radon)
s    <- sd(data$radon)
cbind(xbar, s)
```

```
##          xbar        s
## [1,] 104.1333 9.397421
```

$$\bar{x} \pm t_{n-1}^{1-\alpha/2} \left( \frac{s}{\sqrt{n}} \right)$$

$$104.1 \pm t_{11}^{.95} \left( \frac{9.397}{\sqrt{12}} \right)$$

```
qt(.95, df=11)
```

```
## [1] 1.795885
```

$$104.1 \pm 1.796 \left( \frac{9.397}{\sqrt{12}} \right)$$

$$104.1 \pm 4.872 \quad \text{or} \quad (99.23, 108.97)$$

b) State an appropriate null and alternative hypothesis for a two-sided t-test. Why is a two-sided test appropriate here?

$$H_0 : \mu = 105$$
$$H_a : \mu \neq 105$$

*A two sided test is appropriate because homeowners would be upset if the detectors under-reported the amount of radon but if the detectors over-reported the radon levels, then costly renovations might be undertaken without cause.*

c) Calculate an appropriate test statistic.

$$t_{n-1} = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

$$t_{11} = \frac{104.13 - 105}{9.38/\sqrt{12}}$$

$$= -0.321$$

d) Calculate a p-value.

$$\text{p.value} = P\left(T_{11} \leq -0.321\right)$$

```
2 * pt(-0.321, df=11)

## [1] 0.7542266
```

e) At an $\alpha = 0.10$ level, what is your conclusion. Be sure to state your conclusion in terms of the problem.

*Because the p-value is large compared to $\alpha$, we will fail to reject the null hypothesis and conclude that there is insufficient evidence to reject they null hypothesis that the detectors are accurate.*

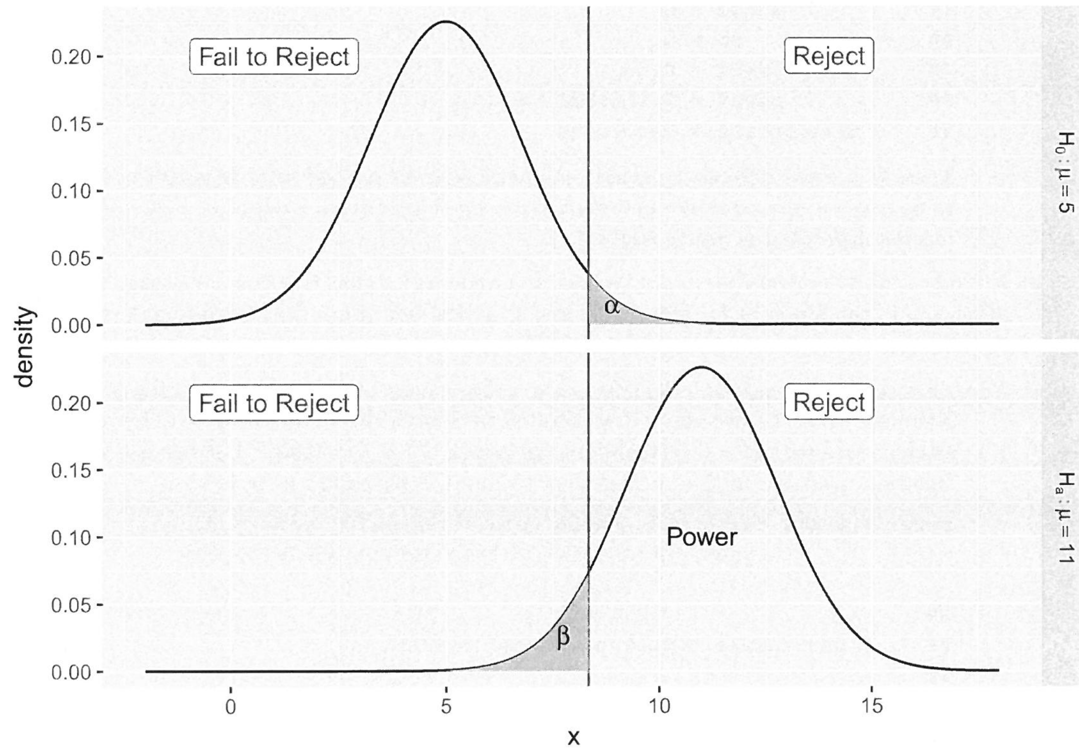
f) Use the function `t.test()` to redo the the hand calculations you did in parts (a), (c), (d).

```
t.test(data$radon, mu=105, conf.level=0.90)

##
##  One Sample t-test
##
## data:  data$radon
## t = -0.31947, df = 11, p-value = 0.7554
## alternative hypothesis: true mean is not equal to 105
## 90 percent confidence interval:
##   99.26145 109.00521
## sample estimates:
## mean of x
##  104.1333
```

3. Given data such that $X_i \sim N\left(\mu, \sigma^2 = 5^2\right)$, the following graph shows the distribution of a sample mean of $n = 8$ observations under the null hypothesis $H_0 : \mu = 5$. We are interested in testing the alternative $H_a : \mu > 5$ at the $\alpha = 0.05$ level and therefore the cut off point for rejecting the null hypothesis is $t_{crit} = 1.895$ and $\bar{x}_{crit} = 1.895 * 5 + 5 = 8.35$.

a) Add the plot of the distribution of the sample mean if $\mu = 11$ and denote which areas represent $\alpha$, $\beta$, and the power in the figure below.




b) Under the same alternative value of $\mu = 11$, find the probability of a Type II error. That is, calculate the value of $\beta = P(\bar{X} < 8.35 | \mu = 11)$

$$\beta = P\left(\bar{X} < 8.35\right)$$

$$= P\left(\frac{\bar{X} - \mu}{s/\sqrt{n}} < \frac{8.35 - 11}{5/\sqrt{8}}\right)$$

$$= P\left(T_7 < -1.50\right)$$

$$= 0.08865$$

*So we then can calculate the power as*

$$\text{Power} = 1 - \beta = 1 - 0.08865 = 0.9113$$

*Suppose that we wanted to actually use the* `power.t.test()` *function to calculate this for us. This isn't too bad, but we must remember to specify that it is a one-sample test and a one-sided alternative.*

```
power.t.test( delta=6, sd=5, n=8, sig.level=0.05,
              alternative = 'one.sided',type = 'one.sample')
```

```
##
##      One-sample t test power calculation
```

```
##
##                 n = 8
##             delta = 6
##                sd = 5
##         sig.level = 0.05
##             power = 0.9189822
##       alternative = one.sided
```

*There is a slight difference between my hand calculation and what R gives me because I really ought to be using a non-central t-distribution in my calculations, instead of just shifting the whole thing, but the difference is pretty slight.*

4. A study is to be undertaken to study the effectiveness of connective tissue massage therapy on the range of motion of the hip joint for elderly clients. Practitioners think that a reasonable standard deviation of the differences (post - pre) would be $\sigma = 20$ degrees.

    a) Suppose an increase of 5 degrees in the range would be a clinically significant result. How large of a sample would be necessary if we wanted to control the Type I error rate by $\alpha = 0.1$ and the Type II error rate with $\beta = 0.1$ (therefore the power is $1 - \beta = 0.90$)? Use the use the `power.t.test()` function available in the package `pwr` to find the necessary sample size.

    ```
    power.t.test( delta = 5, sd=20, sig.level=0.10, power=.9,
                  type='one.sample', alternative='one.sided')
    ```

    ```
    ##
    ##      One-sample t test power calculation
    ##
    ##                 n = 105.9429
    ##             delta = 5
    ##                sd = 20
    ##         sig.level = 0.1
    ##             power = 0.9
    ##       alternative = one.sided
    ```

    *So we would need to have around* $n = 106$ *observations.*

    b) Suppose we were thought that only increases greater than 10 degrees were substantive. How large must our minimum sample size be in this case? What about for 15, 20, 25 and 30 degrees? Sketch a graph of $n$ versus the difference to be detected and comment on how much larger a sample size must be to detect a difference half as small.

    ```
    power.t.test( delta = 10, sd=20, sig.level=0.10, power=.9,
                  type='one.sample', alternative='one.sided')
    ```

    ```
    ##
    ##      One-sample t test power calculation
    ##
    ##                 n = 27.13847
    ##             delta = 10
    ##                sd = 20
    ##         sig.level = 0.1
    ##             power = 0.9
    ##       alternative = one.sided
    ```

    *For a difference twice as large, we need approximately* $n = 28$ *subjects. Notice that this is approximately 1/4 the first sample size. Recall for figuring out how large of* $n$ *we needed for a* *particular margin of error for a CI, we had used the formula*

$$n = \left[z_{1-\alpha/2}\left(\frac{\sigma}{ME}\right)\right]^2$$

*so if we want the margin of error to be twice as small, that "twice" part gets squared so that the $n$ needed is 4 times as large.*

*Notice for $\delta = 10$ we needed approximately $n = 28$ observations but for $\delta = 5$ we needed $n = 106$. Notice that $106/28 \approxeq 4$ and the only reason this ratio isn't exactly four is because of the t-distribution changing as $n$ increases. So the general rule is...*

*If you want to be twice as accurate, you have to have four times the samples.*

*To make a nice graph of the relationship between the largest non-interesting difference between $\mu_0$ and $\mu_a$, we need to call the* `power.t.test()` *multiple times. Instead of doing this by hand, I'll write a little loop that will call* `power.t.test()` *repeatedly with different values for delta.*

```
results <- data.frame( d=5:30, n=NA )              # Consider d= 5, 6, 7, 8, ... 30
for( i in 1:nrow(results) ){                       # for each row, i = 1,2,3, ... 25
  temp <- power.t.test(
              delta = results$d[i],                  # use the d in row i
              sd=20, sig.level=0.10, power=.9,       # to do the power analysis
              type='one.sample',                     #
              alternative='one.sided')               #

  results[i,'n'] <- temp$n        # save calculated 'n' into row i of results data frame
}

ggplot(results, aes(x=d, y=n)) +
  geom_line() +
  geom_point() +
  ylab('Sample size (n)') +
  xlab('delta: Largest negligible effect size')
```