Stat 371-003, Solutions to Homework #5

1. **6.10 (pg 194)**

   (a) The standard error of the mean (what I would call our *estimate* of the standard error of the sample mean) is $s/\sqrt{n}$, where s is the sample SD and n is the sample size.

   So, the SEM is $8.7/\sqrt{5} \approx$ **3.9**.

   In R:

   ```
   x <- c(29.6, 21.5, 28.0, 34.6, 44.9)
   sd(x)/sqrt(length(x))
   ```

   (b) To construct a 90% confidence interval, we need the 95th percentile of a t distribution with 4 degrees of freedom (4 = 5 – 1, and the sample size is 5). From the table at the back of the book, this is **2.132**. In R, we would type `qt(0.95, 4)`.

   The 90% confidence interval for the population average is then 31.7 $\pm$ 2.132 $\times$ 3.9 $\approx$ 31.7 $\pm$ 8.3, or the interval **(23.4, 40.0)**.

   In R, we could type

   ```
   t.test(x, conf.level=0.9)
   ```

2. **6.19 (pg 196)**

   We first calculate the estimated standard error of the sample mean, $1.84/sqrt36 \approx 0.31$.

   We next find the 97.5 percentile of a t distribution with $36 - 1 = 35$ degrees of freedom. From the table in the back of the book, it is between the values 2.021 (for 40 d.f.) and 2.042 (for 30 d.f.); let's call it 2.03. (In R, we would type `qt(0.975, 35)`, which gives 2.030108, to be overly precise.)

   So our 95% confidence interval for the population mean is 6.21 $\pm$ 2.03 $\times$ 0.31 $\approx$ 6.21 $\pm$ 0.63, or the interval **(5.58, 6.84)**.

3. **7.11 (pgs 231–232)**

   Our estimate of the standard error of the difference between the two sample averages is

   $$\sqrt{\frac{13^2}{4} + \frac{13^2}{4}} \approx 9.19$$

   (a) To calculate a 95% confidence interval for the difference between the two population averages (Dark – Photoperiod), we first need the 97.5 percentile of a t distribution with 6 degrees of freedom. From the table at the back of the book, this is **2.447**. In R, we would type `qt(0.975, 6)`.

   Our confidence interval is then $(92 - 115) \pm 2.447 \times 9.19 \approx -23 \pm 22.5$, or the interval **(–45.5, –0.5)**.

(b) To calculate a 90% confidence interval for the difference between the two population averages (Dark – Photoperiod), we need the 95th percentile of a t distribution with 6 degrees of freedom. From the table at the back of the book, this is **1.943**. In R, we would type `qt(0.95, 6)`.

Our confidence interval is then $(92 - 115) \pm 1.943 \times 9.19 \approx -23 \pm 17.9$, or the interval **(–40.9, –5.1)**.

4. **7.19 (pgs 233–234)**

We first calculate our estimate of the standard error of the difference between the two sample means,

$$\sqrt{\frac{11.1^2}{9} + \frac{11.2^2}{11}} \approx 5.01$$

We then find the 95th percentile of a t distribution with 17.3 degrees of freedom. Interpolating between the values in the table for 17 and 18 d.f., we might say **1.738**.

The 90% confidence interval for the difference between the two population means is then $(7.3 - 5.9) \pm 1.738 \times 5.01 \approx 1.4 \pm 8.71$, or the interval **(–7.3, 10.1)**

In R, we could type:

```
x <- c(28,11,-3,14,-2,-4,18,2,2)
y <- c(26,1,0,-4,-4,14,16,8,0,18,-10)
t.test(x, y, conf.level=0.9)
```

5. **[R problem]**

We first read in the data.

```
dat <- read.csv("http://www.biostat.wisc.edu/%7Ekbroman/teaching/stat371/hw05data.csv'
```

Let's assign the values for p53 +/– to `x` and `y`.

```
x <- dat$activity[dat$p53=="++"]
y <- dat$activity[dat$p53=="--"]
```

(a) We can calculate a 95% confidence interval for the average activity in p53 positive cells with the code `t.test(x)`.

We obtain the interval **(41.6, 85.2)**.

(b) We can calculate a 95% confidence interval for the average activity in p53 negative cells with the code `t.test(y)`.

We obtain the interval **(54.1, 63.0)**.

(c) We can calculate a 95% confidence interval for the difference between the two averages with the code `t.test(x, y)`.

We obtain the interval **(–15.8, 25.5)**.