Homework #5 – Due Monday, Nov. 3
COR1-GB.1305 – Statistics and Data Analysis

Problem 1

For each of the following values of α, and n, find $t_{\alpha/2,n-1}$. Round the answer to two digits after the decimal point.

(a) $\alpha = 0.10$, $n = 25$.

$$t_{.050,24} = 1.711 \approx 1.71.$$

(b) $\alpha = 0.02$, $n = 10$.

$$t_{.010,9} = 2.821 \approx 2.82.$$

. . . .

Problem 2

Consider the time it takes for a call center to answer its calls. A random sample of 7 *calls revealed a sample mean time of* 191 *seconds and a sample standard deviation of* 11.4 *seconds.*

(a) *What is the sample?*

The times to answer the $n = 7$ sampled calls.

(b) *What is the population?*

The times to answer all calls at the call center.

(c) *Explain what the population mean represents in this problem.*

The expected time it takes the call center to answer a call. (Equivalently, the average of all of the times taken to answer all calls at the call center.)

(d) *Construct a 95% confidence interval for the population mean.*

For a 95% confidence interval and 7 observations, we have $\alpha = .05$ and

$$t_{\alpha/2,n-1} = t_{.025,6} = 2.447.$$

The 95% confidence interval for the population mean is

$$\begin{aligned}\bar{x} \pm t_{\alpha/2,n-1}\frac{s}{\sqrt{n}} &= 191 \pm 2.447\frac{11.4}{\sqrt{7}} \\ &= 191 \pm 10.5 \\ &= (180.5, 201.5).\end{aligned}$$

(e) *Construct a 99% confidence interval for the population mean.*

For a 99% confidence interval and 7 observations, we have $\alpha = .01$ and

$$t_{\alpha/2,n-1} = t_{.005,6} = 3.707.$$

The 99% confidence interval for the population mean is

$$\begin{aligned}\bar{x} \pm t_{\alpha/2,n-1}\frac{s}{\sqrt{n}} &= 191 \pm 3.707\frac{11.4}{\sqrt{7}} \\ &= 191 \pm 16 \\ &= (175, 207).\end{aligned}$$

(f) *State any assumptions you needed to do this problem. Do you think that the assumptions are reasonable? Why or why not?*

Since the sample size is small ($n < 30$), we need to assume that the population is normal. That is, we need to assume that the histogram of the times to answer all calls looks like a bell curve. Personally, I do not think that this is reasonable; I expect that the time to answer a call has a high skew to the right. (You will still get full credit if you think that the assumption is reasonable, as your reasoning makes sense.)

. . . .

Problem 3

A study of 80 *students who used a private tutor to help them improve their* SAT *scores revealed that their score on the mathematical section improved by an average of* 11 *points, with a sample standard deviation of* 65 *points.*

(a) *What is the sample?*

The SAT improvements (in points) for the $n = 80$ students who used the private tutor.

(b) *What is the population?*

The SAT improvements (in points) for all students who use the private tutor.

(c) *Explain what the population mean represents in this problem.*

The expected improvement for a student who uses the tutor. (Equivalently, the average improvement of all students who use the tutor.)

(d) *Provide a 95% confidence interval for the population mean.*

Since $n \geq 30$, we can use the approximation $t_{\alpha/2,n-1} \approx z_{\alpha/2} = 1.96 \approx 2$. A rough 95% confidence interval is

$$\begin{aligned} \bar{x} \pm t_{\alpha/2,n-1} \frac{s}{\sqrt{n}} &= 11 \pm 2 \frac{65}{\sqrt{80}} \\ &= 11 \pm 14.5 \\ &= (-3.5, 25.5) \end{aligned}$$

A more precise interval uses the value $t_{.025,79} \approx t_{.025,75} = 1.992$; this gives essentially the same answer.

(e) *Does the interval suggest that the tutoring is beneficial? Why or why not?*

Our 95% confidence interval for the mean improvement includes both positive and negative values. In light of this, we do not have compelling evidence that tutoring is beneficial.

. . . .

Problem 4

Consider the excess market returns (the variable called `MarketReturn` *in the file* `MARKET.CSV`*).*

(a) *Use* Stat ⇒ Basic Statistics ⇒ 1-Sample t*, and choose* Samples in Columns: `MarketReturn` *to create a 95% confidence interval for the population mean.*

```
Variable          N     Mean    StDev  SE Mean          95% CI
MarketReturn  11537  0.01782  0.97422  0.00907  (0.00004, 0.03560)
```

(b) *What does this population mean represent, and why might it be of interest to a stock investor?*

The population mean is the expected daily return for this investment (i.e., the long run average daily return over all days). This quantity determines the long-term profitability of the investment.

(c) *Use the Minitab output to check the calculation of the confidence interval in (a). Also verify the calculation of* `SE Mean`*, the (estimated) standard error for the sample mean.*

According to Minitab, the sample mean is $\bar{x} = 0.01782$, the sample standard deviation is $s = 0.97422$, and the sample size is $n = 11537$. The estimated standard error of the mean is

$$\begin{aligned}\mathrm{se}(\bar{x}) &= \frac{s}{\sqrt{n}} \\ &= \frac{0.97422}{\sqrt{11537}} \\ &= 0.00907.\end{aligned}$$

The 95% confidence interval for the population mean (using $t_{.025,11537} \approx 1.96$ is

$$\begin{aligned}\bar{x} \pm 1.96\,\mathrm{se}(\bar{x}) &= 0.01782 \pm (1.96)(0.00907) \\ &= 0.01782 \pm 0.01777 \\ &= (0.00005, 0.03559).\end{aligned}$$

This answers agree with Minitab (up to rounding error).

(d) *Get a 99% confidence interval for the population mean, proceeding as in (a) but adding* Options ⇒ Confidence Level: *99.0.*

```
Variable          N     Mean    StDev  SE Mean           99% CI
MarketReturn  11537  0.01782  0.97422  0.00907  (-0.00555, 0.04119)
```

The 99% confidence interval is $(-0.00555, 0.04119)$.

(e) *What do the confidence intervals from (a) and (d) suggest about the long-term value of investing in the US stock market?*

At the 95% confidence level, there is evidence that the investment yields a positive return in the long run; at the 99% confidence level, we cannot be sure that the investment yields a positive return.

....

Problem 5

Recall the data set `NormTemp.CSV`*, which you studied in Homework 1. This gives data on body temperatures for* 130 *randomly selected subjects. Using Minitab, get a 95% confidence interval for the population mean temperature. Are the results of the confidence interval surprising, in view of the fact that the population mean temperature is supposed to be* 98.6 *degrees?*

Here is the Minitab output:

```
Variable    N      Mean   StDev  SE Mean          95% CI
Temp      130  98.2492  0.7332   0.0643  (98.1220, 98.3765)
```

The 95% confidence interval for the population mean is $(98.1220, 98.3765)$. It is somewhat surprising that 98.6 is not in this interval. However, we should note that only 95% of all 95% confidence intervals contain their parameters; 5% of the time, they do not.

....