

Lecture/text homework assignment # 6

Note: Please circle your answers when appropriate! Show your work!

0) Problems 6 & 7 from last time are due; note that your recitation instructor may ask you to present them.

1) Let's assume that the average blood urea levels for women is about 13.7 mg/dL with a standard deviation of about 3.67 mg/dL ($\mu = 13.7$ mg/dL, $\sigma = 3.67$ mg/dL)

a) Calculate $Pr\{Y < 8 \text{ mg/dL}\}$

b) Now assume you have a sample of $n = 3$ women. Calculate $Pr\{\bar{Y} < 8 \text{ mg/dL}\}$.

c) Which probability is lower? **Why?**

2) Refer to problem 1. Now assume you take a sample of $n = 6$ women

a) What is the probability of the sample average being within 2 mg/dL of the population mean (13.7 mg/dL)?

In other words, you need to figure out $Pr\{\mu - 2 < \bar{Y} < \mu + 2\}$

b) What is the probability of the sample average being within 2 mg/dL of the population mean if the population mean is actually $\mu = 15.8$ mg/dL?

Are you surprised? ***Why or why not?*** You should make sure you understand what happened here.

c) Now assume a sample of $n = 11$ women and repeat (a).

What is the effect of sample size?

3) Let's go back to our coquí (*Eleutherodactylus coqui*) densities in Puerto Rico. You measure 8 plots, each plot being a hectare in size. You get the following results for number of frogs/hectare:

18528 18587 19270 19886 20142 20513 20668 21242

These data yield the following: $\bar{y} = 19854.5$, and $s = 986.4806$

a) Construct an 90% CI for these data.

b) Construct a 99.8% CI for these data.

c) Suppose someone told you that they thought the average number of frogs/hectare was 23,000. Is their estimate consistent with your data? ***Why or why not?***

4) Consider the results of 3(b). Notice that all the data fit within the 99.8% confidence interval. Is this usually the case (in other words, will a 99.8% CI contain most of the observations)?

Caution: a lot of people get this wrong! Here's a hint: suppose you had measured the worms in 6000 acres (instead of just 8). What happens to the confidence interval? If you're not sure, substitute 6,000 for 8 in your calculation for 3(b) to see what happens.

5) Suppose the (sample) average length of great white sharks is $\bar{y} = 4.75\text{m}$ for females and 3.7m for males. Let's assume that for both the sample standard deviation is about $s = 0.45\text{m}$.

(a) Suppose you take a sample of 5 females and 7 males. Construct a 99% CI for both. Do the confidence intervals overlap?

(b) Now repeat using a sample of 67 females and 83 males. Do the confidence intervals overlap?

(c) Can you explain what happened? Why is sample size important if you're trying to find differences between groups?

*(Statistical comment: comparing CI's is **not** the correct way of comparing two groups, but it can give you an idea if there's a difference).*

6) Use the data for heights that you collected in recitation last week and construct an 80% CI. Use R to do this.

While your lecture may not have learned about the t-test yet, a CI is just the reverse of a one sample t-test. To get Confidence Intervals in R see the following page.

Remember: Copying answers from other students, web sites, previous classes or using AI is a violation of the honor code!!

Be prepared to discuss these problems in recitation the week of March 15th. As usual, problems not discussed in recitation are due at the end of class.

R notes:

To get 95% Confidence intervals in R/Rstudio:

```
t.test(height)
```

or

```
t.test(dataset$height)
```

Depending on how you've arranged your data.

R defaults to a 95% CI, *but if you want an 80%CI*, you need to modify the above:

```
t.test(height,conf.level = .80)
```

or

```
t.test(dataset$height, conf.level = .80)
```

The confidence interval will be near the bottom of the printout from R.