MAT 240 – Week 2c

Section(s) covered: 5.1, 5.2
Subject: The basis of confidence intervals

Fundamental ideas:

This is very similar to hypothesis tests (sections 6.1, 6.2) but is looking at the situation from the opposite side. I think that hypothesis tests are easier to explain (the question to answer there: is it too far or is it close enough … an easier question to grasp) and confidence intervals are a little harder. But the two are very closely tied together.

- The basic question is this: I really don't know what the average "." is. So I measure some examples and get a guess as to what the average is. What's the likely accuracy of my guess?
  - For example, say I want to know the average weight of a green turtle. Who knows? So, I go out and catch 100 of them … they weigh, on the average, say 470 pounds. So how accurate is this? Can I say that I think that the average weight of all green turtles is between 460 and 480? or between 420 and 520? or between 370 and 570?
  - For example, say that I want to know the percent of people who will vote for Mr. X this coming election. So I go out and poll 1000 people … 53% of them say that they will be voting for Mr. X. How accurate is that? Is that 53% with a 1% margin of error (i.e. the actual percent will be between 52 and 54)? With a 3% margin of error (TYPICAL!). With a 10% margin of error?

- How can we <u>quantify</u> the question "how confident are we in our estimate, or "what's the margin of error"? We use the normal distribution again … surprise, surprise. Say that Z is the standard normal (mean = 0, std dev = 1). What do we think as an area that includes "most" of the cases? Let's say it's between –2 and +2 … –2 standard deviations below to 2 standard deviations above. We know that includes 95% of the cases, so let's say that this is "confident enough".

- So for a general normal distribution, the range from the mean – 2 std dev's to the mean + 2 std dev's contains "enough" of the possibilities, we're "confident enough" that the data values are mainly there.

- That's a confidence interval! Notice how the calculations are very similar to the hypothesis test, like $s / \sqrt{n}$ and the 2 standard deviations.

How to do a confidence interval:

<u>Example 1</u>: We want to know the average weight of all green turtles. We don't really have any idea what that is, so we go do an experiment.

- We go find 100 of them and weigh them, their average weight is 470 pounds. How confident are we in how accurate this is? (Answer so far – we can't tell.)
- In doing our 100 measurements, we find that the standard deviation of their weights is 170 pounds. (Start to sound familiar with hypothesis tests?)

- You shouldn't be surprised that we're going to use the s / √n formula … the standard deviation of the 470 measurement is 170 / √100 = 17. (Section 4.11!)
- So our estimate of the average weight of all green turtles is 470, and this has a standard deviation of 17.
- We're "confident enough" that the real average is between 470 – 2*17 = 436 and 470 + 2*17 = 504 (because we're confident enough that the data is between –2 std devs and +2 std dev's away from the mean).
- Thus our "confidence interval" can be written as either
  - (436,504) or
  - 470 ± 34 or
  - 470 with a margin of error of 34
- Written officially, this is

$$\bar{X} - 2\,(s/\sqrt{n}) \quad \text{to} \quad \bar{X} + 2\,(s/\sqrt{n})$$

$$470 - 2\,(170/\sqrt{100}) \quad \text{to} \quad 470 + 2\,(170/\sqrt{100})$$

or

(436, 504)

- Conclusion: We're reasonably confident that the population mean (the average weight of all green turtles) is between 436 and 504 pounds.

Example 2: This one I'll go through quicker. It's exactly the same though (as you saw in section 4.11, a lot of this is doing the same process over and over again, repeatedly). We're interested in the average number of cashews there over the population of all 16 ounce jars.

- We go find 64 jars, pour out the cashews, and count them. The average number is 127 cashews.
- In eating, uh, counting our 64 jars, we find that the standard deviation of the numbers of cashews is 12. So what's the next step? Of course it's 12 / √64 = 1.5, the standard error. (I'm on a diet where every few days I can eat as many cashews as I want to … wonderful! It seems to work too, strangely enough.)
- Now, how confident are we in the 127? We're pretty sure that the average number over all 16 ounce jars is between 127 – 2*1.5 = 124 and 127 + 2*1.5 = 130.
- Our confidence interval for the average number of cashews per jar is (124, 130).

Example 3: The process of doing a confidence interval:

- We want to estimate the mean of the total population
- We have a measured sample mean (Xbar, the 470 or the 127).
- We have a measured standard deviation (σ or s, the 160 or the 12).
- We do the (s / √n) calculation (the 17 or 1.5)
- We're 95% confident that the mean of the total population is between the sample mean minus (s / √n) to the sample mean plus (s / √n)

Why 2? And Why Sometimes Other Numbers?

Where do we get the number 2?

Hypothesis tests

- For the normal distribution the probability of being 2 or more away from the mean is 5%
  - $P[Z < -2] + P[Z > 2] = .05$

Confidence intervals

- For the normal distribution, 95% of the values are 2 or closer to the mean
  - $P[-2 < Z < 2] = 95\%$

I think of hypothesis testers are pessimists (they look at the 5% that can go wrong) while the confidence intervalers are optimists (they look at the 95% that can go right).

So … instead of being a 5% pessimist, what about being a 1% pessimist?
So … instead of being a 95% optimist, what about being a 99% optimist?

For the HT at .01 or CI at 99%, we need to use a different normal distribution calculation:

- $P[Z < -2.576] = .005$
- $P[-2.576 < Z < 2.576] = .99$
- $P[Z > 2.576] = .005$

So the 1% pessimists, a hypothesis test at a .01 level of significance, or $\alpha = .01$, we reject if the z score is more than 2.576 or less than −2.576 … i.e. if we are more than 2.576 steps away (either too much lower or too much higher).

So the 99% optimists, a 99% confidence interval we would use the sample mean – 2.576 standard errors to the sample mean + 2.576 standard errors.

The 10%/90% would use 1.645 … the 20%/80% would use 1.28 … that's why we spent time on these normal distribution calculations.

Example 4 (Problem 5.10) : Latex allergies in health care workers (my wife is a nurse and she has latex allergies, that's why she couldn't pursue what she really wanted– pathology)

a. The point estimate is just a fancy name for the average we measured – 19.3 here (the Xbar that we're given)
b. A 95% confidence interval
   - Sample mean = 19.3
   - Sample standard deviation = 11.9
   - Calculation – standard error is … what else … $11.9 / \sqrt{46} = 1.75$
   - The lower limit … $19.3 - 2*1.75 = 15.8$
   - The upper limit … $19.3 + 2*1.75 = 22.8$
   - Official answer: a 95% confidence interval is (15.8, 22.8)

c. We're 95% confident that the true percent of latex gloves used per week is (15.8, 22.8). In other words, if we took another sample of 46 and computed the confidence interval for that, and another 46, and another 46, and another 46, … then 95% of the time the true population average will be in those intervals

This isn't quite the answer that a book would give because the number 2 from the normal distribution is actually 1.96, so a textbook would give a slightly different answer (using 1.96 instead of 2). From my point of view, that's a pretty trivial point to argue.