MAT 240 – Week 3a

Section(s) covered: 6.3
Subject: The critical value method and the p-method method (hypothesis tests)

Fundamental ideas:

- There are two equivalent ways to assess the significance of hypothesis tests
  - You can define a "critical value" … in other words, if you are () steps away or more, that's too far. A typical critical value is 1.96 (we've been using 2) which corresponds to a specific point on the normal distribution
  - You can define a "probability level" … in other words, if this event has a small probability of happening, then it's too unusual. A typical probability level is .05 (something happens 5% of the time or less is too rare)
- The two are exactly equivalent
  - For a critical value of 1.96, P[Z > 1.96] = .025 and P[Z < –1.96] = .025.
    - The chance of being 1.96 steps or closer … P[–1.96 < Z < 1.96] is .95
    - The chance of being 1.96 steps or further … 0.025 + 0.025 = 0.05
  - A z-value of more than 1.96 or less than –1.96 means that it's in this 5% range
  - A z-value in this 5% range has to be either more than 1.96 or less than –1.96
  - This sounds like double-talk, and it is double-talk.
- This is the relationship between the critical value method and the p-value method.

How to reach a conclusion for a hypothesis test:

Example 1: The same as Example 1 in Week2b … the first hypothesis test week. We're measuring weights of elephants. We guess that the average weight of a certain type of elephant is 15,000 pounds.

- We go find 100 of them and weigh them, their average weight is 15,500 pounds and the standard deviation of their weights is 1200 pounds.
- We're going to use the $s / \sqrt{n}$ formula … big surprise … the standard deviation of the sample mean of 15,500 is $1200 / \sqrt{100} = 120$.
- 15,500 is 500 away … that's 500/120 = 4.2 steps or standard errors away. This is the "z–score" formula … (15,500 – 15,000) / 120 = 4.2
- The critical value method
  - How far are we allowed to be? Either 1.96 steps above or closer … or 1.96 steps below or closer.
  - Here we are 4.2 steps above … way more
  - Reject the null hypothesis … we're too far
- The p-value method
  - What's the chance of being 4.2 or more away from the mean? We first calculate P[Z > 4.2] = .000013. However, we would consider Z < –4.2 as being that far away too, and P[Z < –4.2] = .000013 also. So the chance that

we're 4.2 away (either positive or negative) is .00026. This is way less than a .05 (or 5%) event, so we reject the null hypothesis … this is too unlikely.

- Conclusion: We're wrong. The average weight isn't 15,000 … the data is inconsistent with that hypothesis.

Example 2: The same as Example 2 from Week 2b. We're guessing that the average amount of soda in a 2.0 liter bottle is 2.0 liters. Let's say that we want to be really sure … 99% sure instead of just 95%. So instead of 1.96, we use 2.576 (a normal distribution calculation. Instead of P[Z > 1.96] = .025 … half of .05 … we have P[Z > 2.576] = .005 … half of .01. The stats way of saying this is that here the significance level $\alpha = .01$.

- We have a sample size of 64 bottles, a sample mean of 1.989 liters and a sample standard deviation of .032 liters.
- What's the next step? Of course it's $.032 / \sqrt{64} = .004$, the standard error.
- The z-score is $(2 - 1.989) / .004 = 2.75$
- The critical value way
  - 2.75 is more than 2.576 … so we reject
- The p-value way
  - P[Z > 2.75] = .003 … so P[Z < –2.75] = .003 also, and the probability of being 2.75 or more away (either greater 2.75 or less than –2.75) is .006.
  - .006 is less than .01 … it's more unlikely … so we reject

Comparing the methods:

So why do some people use the critical value method and others use the p-value method? Actually, more people use the p-value method. Here's a comparison. Remember that both give the exact same answer (reject or do not reject).

Advantages of the critical value method

- It's easier to understand. 2 steps or more away is too far away … that seems to be a pretty easy concept for me.
- From the z-score, you can tell immediately whether it's a reject or not. If the z-score is more than 2 or less than –2, it's a reject. End of story.

Advantages of the p-value method

- The critical value method only says "reject" … the p-value method says "reject by how much". A p-value of .049 is a barely reject … a p-value of .0009 is a really strong reject. It's not possible to be as precise as this using critical values.

Since most everything is computed with software programs, and software can do either one just as well, and since the p-value method is more informative, almost all (all?) software programs output the p-value for you to use for your conclusion.