Announcements

Exam 2 will be returned on Wednesday

THERE IS NO SECTION THIS WEEK

HOMEWORK 5 IS DUE WEDNESDAY

- Office Hours This Week
  - MONDAY 2pm-4pm
  - WEDNESDAY 8am-9:30am
- Unclaimed 1st exams & homeworks (1, 2, 3, & 4) are up front
- Did you accidently pick up an exam that wasn't yours? Please return it to the front table, no questions asked.
- Please Remove Chapter 22 from your outlines, we will skip 22 and move straight to Chapter 23

Chapter 23. Inferences about Means

- Recall: Proportions (0,1; success/failure) vs. Quantitative Data (potentially has different values for each individual, wider range)
- Quantitative Data is summarized with means and standard deviations (like Chapter 5)
- Chapter 23 involves INFERENCE (making generalizations from samples back to the pouplation) so it involves thinking about Chapter 18 (sampling distribution/distribution of all possible samples)

SD vs. SE for a sampling distribution

	Parameters Known (SD)	Parameters Unknown (SE)
Binary outcomes (0,1)	$SD_{(p)} = \sqrt{\frac{pq}{n}}$	$SE_{(\hat{p})} = \sqrt{\frac{\hat{p}\hat{q}}{n}}$
Quantitative Outcomes (has a range of values)	$SD_{(\bar{y})} = \frac{\sigma}{\sqrt{n}}$	$SE_{(\bar{y})} = \frac{s}{\sqrt{n}}$

The problem with only knowing the SE for quantitative data

- For small samples (< 50) statisticians noticed that the SE of y-bar varied quite a bit and the sampling distribution was not really quite normal.
- Gosset, a quality control engineer at Guiness Brewery in Ireland solved the problem.
- He developed a new distribution, it looks normal, but it is not, it is "t".

The t-distribution: Properties

- Has mean “mu”
- Has standard error
- Is bell-shaped
- Is symmetric around mu
- Has DEGREES OF FREEDOM (df) n-1
- Has fatter “tails” than the normal
- See Table A-84 (back of textbook)
- Basically used when sigma is unknown and you must use s from the sample.

Standard Normal Distribution (Z) vs. t distribution (t)

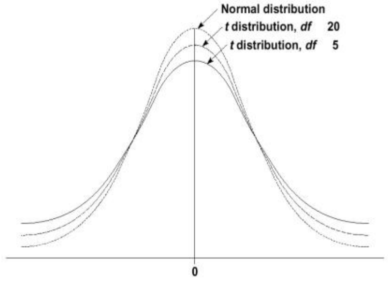

Fig. 9.1 Degrees of Freedom and the Shape of the t-Distribution

Notice there is only one standard normal but the distribution of t is determined by its “degrees of freedom” (df) which is n-1

The t-distribution: formula

- For the same distance between the mean y-bar and mu and same size denominator (SD for Z distribution, SE for t-distribution), a t-distribution will produce larger p-values and wider confidence intervals
- Formula

$$t = \frac{\bar{y} - \mu}{\frac{s}{\sqrt{n}}} \quad \text{VS.} \quad Z = \frac{\bar{y} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

- Upshot – having to use t instead of normal (Z) has a “cost” – i.e. a larger margin of error.

Z vs. t (p. 433)

- If sigma is known, use the standard normal Z distribution and it’s formulas.
- When sigma is unknown and your sample is small (n size 50 or less) you should probably use t.
- Your textbook has various criteria for sample size, they basically use t for all cases where sigma is unknown. When n is large enough, t and Z are effectively the same.

Assumptions and Conditions

- Random Samples
- Independence – check to make sure the sample is no larger than 10% of the population
- (p. 435) for very small samples, n < 15 or so your data should be relatively normal for this to work.
- (p. 435) for samples between 15- about 50, t works well as long as there is symmetry (no skewness)
- (p. 435) for samples over size 50 you can use t even if it's non-normal or skewed (you could probably use the normal after size 50)

Confidence Interval Using t for the mean AKA One-Sample t-interval

- Problem 10, p. 448. On 44 weekdays in a public parking garage, the daily fees collected had a mean of $126 with a standard deviation of $15. The population is not known, but not thought to be highly skewed.
- Please construct a 90% confidence interval for the daily income for this parking garage

Solution

- First the formula $\bar{y} \pm t^*_{n-1} * SE_{(\bar{y})} \Rightarrow \bar{y} \pm t^*_{n-1} * \left(\frac{s}{\sqrt{n}}\right)$
- Second, we need to find a t* for 90% confidence and n-1=43 (look up in the t-table), the closest, which is the next lowest is n=40 or t=1.684
- The resulting 90% confidence interval is

$$126 \pm 1.684 * \left(\frac{15}{\sqrt{44}}\right) \Rightarrow 126 \pm 3.808$$

Hypothesis Testing using t or a one-sample t-test for the mean

- Recall Chapter 20 (this was for proportion)
- Same method, different test and slightly different null and alternative hypotheses.
- P-value is found by looking up the appropriate value in the t-table
- Alpha Level is still generally equal to .05

Example: #22 from the text (p.450)

- A company would like their fleet of vehicles to have an average MPG of at least 26 miles per gallon. The population SD is unknown.
- A random sample of size 50 has a mean of 25.02 mpg with a standard deviation of 4.83mpg.
- Does this sample provide sufficient evidence to show that they have failed?

Solution

- State a null hypothesis $H_0 : \mu = 26$
- State an alternative hypothesis $H_1 : \mu < 26$
- Perform a t-test because we don't know sigma (σ) and we only know s (sample SD) it looks like:

$$t = \frac{\bar{y} - \mu}{\frac{s}{\sqrt{n}}} \Rightarrow t = \frac{25.02 - 26}{\frac{4.83}{\sqrt{50}}} = -1.435$$

- The p-value is t_{n-1} or t with 49 degrees of freedom, if you can't find the exact value, use the next lowest so df=45
- If alpha=.05 and this is a one-sided (< or >) test so the one-tailed probability is: 1.679

Solution (continued)

- The t you calculated from the test is -1.435
- The one-tailed probability (alpha) for t with 45 degrees of freedom (closest to 50-1=49) is 1.679 for alpha=.05
- Take the absolute value of your calculated t, so 1.435 and compare it to the table t of 1.679.
- If the calculated t < table t, do not reject the null
- If the calculated t > table t, reject the null
- In our case, we WOULD NOT REJECT the null hypothesis. Our results are not statistically significant, this means it is quite possible that the true mean is really 26 and the company is doing OK.

Announcements

Exam 2 will be returned on Wednesday

THERE IS NO SECTION THIS WEEK

HOMEWORK 5 IS DUE WEDNESDAY

- Office Hours This Week
  - MONDAY 2pm-4pm
  - WEDNESDAY 8am-9:30am
- Unclaimed 1st exams & homeworks (1, 2, 3, & 4) are up front
- Did you accidently pick up an exam that wasn't yours? Please return it to the front table, no questions asked.
- Please Remove Chapter 22 from your outlines, we will skip 22 and move straight to Chapter 23