Chapter 10: Comparing Two Populations or Groups

Key Vocabulary:

- difference between two proportions
- two sample z interval for proportions
- two sample z test for difference between two proportions
- two sample z statistic
- two sample t statistic
- pooled combined sample proportion
- standard error
- randomization distribution
- paired t-test
- two sample t test for means
- two sample t interval for means
- difference between two means
- pooled two sample t statistic

10.1 Comparing Two Proportions

1. Summarize the three properties of a sampling distribution of a sample proportion:
    - Shape
    - Center
    - Spread

2. What are the shape, center, and spread of the sampling distribution of $p_1 - p_2$? Provide the formulas for the mean and standard deviation.
   - Shape
   - Center
   - Spread

3. What conditions need to be met for the sampling distribution of $p_1 - p_2$?

4. Give the formula for the *standard error* when calculating a confidence interval for $p_1 - p_2$, and define each variable in the equation.

5. What is the confidence interval for $p_1 - p_2$?

6. What conditions must be met in order to use the Two-sample z Interval for a Difference between Two Proportions?
   - Random
   - Normal
   - Independent

8. Use the example, *Teens and Adults on Social Networking Sites*, to outline how to construct and interpret a confidence interval for the difference between two proportions, $p_1 - p_2$.

9. State the null hypothesis for a *two proportion significance test.*

10. What does p_c represent, and how is it calculated?

11. Why do we *pool* the sample proportions?

12. Give the formula for the *two-proportion z-statistic*, and define each variable in the equation.
    a. Is this on the formula sheet? What does the test statistic measure?

13. State and use diagrams to illustrate the three possible alternative hypotheses for a *two proportion z-test.*

14. What are the *conditions* for conducting a two-sample *z* test for a difference between proportions?

15. How are these *different* than the conditions for a one-sample *z* interval for *p*?

16. Describe the *randomization distribution*.

17. What must you be careful about when *defining parameters* in experiments? How can this be avoided?

18. Can you use your calculator for the *Do* step? Are there any drawbacks?
    a. What are the calculator commands for the two-sample z test and interval for $p_1 - p_2$?

10.2 Comparing Two Means (pp.627-648)

1. Summarize the three properties of a sampling distribution of a *sample mean*:
   - Shape
   - Center
   - Spread

2. What are the shape, center, and spread of the sampling distribution of $\bar{x}_1 - \bar{x}_2$? Give the formula for the mean and standard deviation.
   - Shape
   - Center
   - Spread

3. What are the conditions for the sampling distribution of $\bar{x}_1 - \bar{x}_2$?

4. Give the formula for the *two-sample t-statistic*, and define each variable in the equation.
   a. Is this on the formula sheet? What does it measure?

5. What is the standard error of $\bar{x}_1 - \bar{x}_2$? Is this on the formula sheet?

6. What distribution does the two-sample *t* statistic have?

7. Why do we use a *t* statistic rather than a *z* statistic?

8. ~~Without using technology, how do you estimate the degrees of freedom when using two-sample t-procedures?~~ **Do not need to know.**

9. How do you calculate the confidence interval for $\mu_1 - \mu_2$?

10. In a *two-sample t interval* problem, what conditions must be met for comparing two means?

11. What are the conditions for conducting a two-sample *t* test for $\mu_1 - \mu_2$?

12. Draw a sketch of the three possible scenarios for the alternative hypothesis.

13. Describe the *Normal Condition* when using the two sample t procedures.

14. What calculator commands are used for a two-sample t test and interval for $\mu_1 - \mu_2$?

15. How do you proceed when using two-sample t procedures to check the Normal Condition in the following cases:
    - Sample size less than 15
    - Sample size at least 15
    - Large samples

16. In a two-sample problem, must/should the two sample sizes be equal?

17. When doing two-sample *t* procedures, should we pool the data to estimate a common standard deviation? Is there any benefit? Are there any risks?