Review for Exam 2

Tables Provided: t-table, Table 5 (parts of), Table 6 (parts of)

Confidence Intervals (Chapter 6)

1. For a quantitative variable, be able to calculate the confidence interval for μ:
   $\overline{y} \pm t(n-1)_{\alpha/2} SE_{\overline{y}}$ *where* $SE_{\overline{y}} = \frac{s}{\sqrt{n}}$
   and be able to write the interpretation of the interval in words.

2. For a quantitative variable be able to calculate and interpret the appropriate confidence bound for μ (upper or lower)
   $\overline{y} + t(n-1)_{\alpha} SE_{\overline{y}}$ *OR* $\overline{y} - t(n-1)_{\alpha} SE_{\overline{y}}$ *where* $SE_{\overline{y}} = \frac{s}{\sqrt{n}}$

3. Be able to differentiate between the standard deviation of the population mean and the standard error of the sample mean.

4. Be able to determine how the confidence interval changes if the confidence level changes, the sample size changes and the sample mean changes.

5. Given a guess for the standard deviation, be able to determine an appropriate sample size for a given w ($\overline{y} \pm$ w) for the quantitative case.
   $n = \left(\frac{\mathbf{2.5}\, SD_{Guessed}}{w}\right)^2$

6. For a quantitative variable, be able to calculate the confidence interval for μ_1 - μ_2:
   $\overline{y}_1 - \overline{y}_2 \pm t(df)_{\alpha/2} SE_{\overline{y}_1 - \overline{y}_2}$ *where df will be given to you on the exam*
   unspooled $SE_{\overline{y}_1 - \overline{y}_2}$
   $$SE_{\overline{y}_1 - \overline{y}_2} = \sqrt{SE_1^2 + SE_2^2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$
   and be able to write the interpretation of the interval in words.

7. Given a 2-sample confidence interval, be able to determine if there is a difference between the two populations.

8. Be able to verify that the conditions are appropriate to use the C.I. method.
   1. The data are random and independent.
   2. The population distribution (or sampling distribution) is normal.

Hypothesis Testing (Chapter 7)

9. Be able to determine if two populations are different given the results of the randomization test.

10. Be able to write down what H_o and H_A refer to in a given situation.

11. Be able to state the interpretation of the P-value (E.g., “The P-value for a hypothesis test is, assuming that the null hypothesis is true, the probability of the test statistic being at least as extreme as the value of the test statistic that was actually obtained.”) (MC)

12. Given a P-value, determine if you should reject or fail to reject the null hypothesis.

13. Know how to perform hypothesis testing using the following procedure.

	Steps	**Comments**
1	State the scientific question to be answered	Use a complete sentence with no symbols
2	Define the parameters of interest	μ, μ_1, etc.
3	State the null hypothesis H_0 and the alternative hypothesis H_A in terms of parameters	The null hypothesis is usually an equality The alternative hypothesis can be $\neq$, > or <
4	State the significance level α.	If not given, then $\alpha = 0.05$.
5	Calculate the test statistic from the data.	The distribution is usually a t test, but there are cases when F tests, and WMW are used Include other parameters, like df, for the distribution if appropriate
6	Calculate the critical value.	Find the critical value assuming that the distribution of the test statistic assuming that H_0 is true. This defines the rejection region. You may also calculate the P-value here.
7	Compare the test static to the rejection region or compare the P-value to α.	Is test statistic ≤ or ≥ the critical value? Is the P-value ≤ α?
8	Make a decision about the null hypothesis a) if the test statistic is in the rejection region, or the P-value is smaller than α, state “reject H_0” b) if not, state “do not reject H_0”	The word ‘accept’ should not be used
9	Form a scientific conclusion based on that decision. if 8a) (reject H_0), then start with “This study provides evidence … “ if 8b) (fail to reject H_0), then start with “This study does not provide evidence …” followed by “[(P=x)] at the ___ significance level that “ followed by the verbal statement of H_A	Use complete sentences with no symbols (except possibly P) Use similar working in step 3 You must not affirm H_0 It is important to NOT be creative here.

14. Be able to perform a hypothesis test using the Student’s t method:
    a) Two-sample t test: directional and nondirectional
    b) One-sample t test: directional and nondirectional
    c) Be able to determine which method is appropriate
        one or two sample t test, directional or nondirectional

The following table summarizes the different t-tests

T - tests	Null Hypothesis	Alternative Hypothesis				df	t_s	(1-α) CI
		non-directional		directional				
	H_0	H_A	rejection region	H_A	rejection region			
One Sample	$\mu = \mu_o$	$\mu \neq \mu_o$	$t_s<-t_{\alpha/2}$ $t_s>t_{\alpha/2}$	$\mu < \mu_o$	$t_s< -t_\alpha$	n-1	$\frac{\overline{y} - \mu_0}{SE_{\overline{y}}}$	for μ: $\overline{y} \pm t_{\alpha/2} SE_{\overline{y}}$
				$\mu > \mu_o$	$t_s > -t_\alpha$			
Two Independent Samples	$\mu_1 - \mu_2 = \mu_0$	$\mu_1-\mu_2\neq\mu_0$	$t_s<-t_{\alpha/2}$ $t_s>t_{\alpha/2}$	$\mu_1-\mu_2<\mu_0$	$t_s < -t_\alpha$	*	$\frac{\overline{y}_1 - \overline{y}_2 - \mu_0}{SE_{\overline{y}_1-\overline{y}_2}}$	for $\mu_1 - \mu_2$: $\overline{y}_1 - \overline{y}_2 \pm$ $t_{\alpha/2} SE_{\overline{y}_1-\overline{y}_2}$
				$\mu_1-\mu_2>\mu_0$	$t_s > -t_\alpha$			

(Area in grey not required)

* $\min(n_1 - 1, n_2 - 1) < df = \frac{(SE_1^2 + SE_2^2)^2}{SE_1^4/(n_1-1) + SE_2^4/(n_2-1)} < n_1 + n_2 - 2$

this will be given the df on the exam

You can use Table 4 (t-table) to obtain the critical values.

15. Be able to identify which errors are Type I and Type II. (See table that I gave in class.)
    a) be able to determine in a given situation whether a Type I or Type II error can occur.

16. Be able to compare the results of a hypothesis test and the confidence interval. That is, given the results of the hypothesis test, be able to state whether 0 is in the confidence interval and vice versa.

17. Be able to state the definition of power. (Power = 1 - β)

18. Be able to state the difficulties in interpreting observation studies (MC)

19. Be able to state if there is confounding or spurious association in an observational study and how the association might affect the utility of a study (extra question).

20. Be able to state the difference between statistically significant and scientifically important. (MC)

21. Be able to state the effect of α, σ, n, $\mu_1 - \mu_2$ on power. (MC)

22. Be able to calculate the estimated effect size.

$$estimated\ effect\ size = \frac{|\overline{y}_1 - \overline{y}_2|}{s}$$

23. Given the wanted power, effect size, α, determine the necessary sample size (Table 5)

24. Determine if a hypothesis test (t test) is valid for a particular situation (MC)
    1. The data are random and independent.
    2. The population distribution (or sampling distribution) is normal.

25. For the Wilcoxon-Mann-Whitney Test
    a) Be able to determine when the method is valid (MC).
    b) Given a particular situation, be able to determine whether a t-test or WMW test is preferred (MC).
    c) Be able to perform the test (Table 6).