Review for Exam 2

Tables Provided: Z table (Table III from the book), t table (Table V from the book), Studentized Range Distribution (Tukey q values) (Table VIII from the book)

Chapter 1.3

- Be able to distinguish between observational studies and experiments and what different information can be determined from them.
 a) Be able to determine if a study is anecdotal.
- 2. Be able to identify the experimental units, response variable, explanatory variable, treatments or factors, the number of levels and the response variable in a given situation.
- 3. Be able to determine if a study uses simple random sampling or stratified sampling.
- 4. Be able to determine if a standard sampling problem occurs (bias, convenience sample, undercoverage, nonresponse).
- 5. Be able to draw the experimental design for a random experiment, matched pair and block
- 6. Be able to recognize what factors are important in designing an experiment (control, randomization, replication)
- 7. Be able to define lurking and confounding variables.
- 8. Be able to determine whether a variable is confounding or has a common response in a specific situation.

Chapters 8, 9, and 10

- 9. Be able to determine what affects the width of the confidence interval and what the effect is (increase or decrease).
- 10. Be able to state what is random about a confidence interval, that is, what is the meaning of a confidence interval.
- 11. Be able to state the assumptions (and determine if they are valid) that are required for inference to be performed. All graphs will be provided if required.

12. Be able to calculate and interpret the CI in the following cases:

	one-sample z	one-sample t	2-sample z(t) independent	2-sample t(z) pair
μ	$\overline{x} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$	$\overline{x} \pm t_{\alpha/2,n-1} \frac{s}{\sqrt{n}}$	$\overline{x}_1 - \overline{x}_2 \pm t_{\alpha/2,\nu} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$	$\overline{d} \pm t_{\alpha/2,n-1} \frac{s_D}{\sqrt{n}}$
df	N/A	n – 1	will be given to you	n - 1

a) Be able to determine which column you should use.

- b) Be able to calculate z or t critical value.
- c) Note: for the two-sample cases, we will provide you with enough information so that you can do the calculation either via the independent or pairs methodology.
- d) If the question says 'interpret', this means that the conclusion needs to be stated in words; otherwise only the interval is required. The words needs to be what is on the SLIDES.
- 13. Be able to calculate and interpret the lower and upper bound for each of the cases in 12. The table below is only for the one-sample z. Notice the different critical value

lower bound	upper bound
$\mu > \overline{x} - z_{\alpha} \frac{\sigma}{\sqrt{n}}$	$\mu < \overline{x} + z_{\alpha} \frac{\sigma}{\sqrt{n}}$

14. Be able to calculate the sample size, n, needed for a particular margin of error, ME (half-width). $t'_{\alpha/2,n-1}$ is calculated for the preliminary study (the study that obtained s)

one-sample z	one-sample t	
$n = \left(\frac{z_{\alpha/2}\sigma}{ME}\right)^2$	$n = \left(\frac{t'_{\alpha/2, n-1}s}{m}\right)^2$	

15. Be able to use the 4 steps in hypothesis testing:

- 1. Identify the parameter(s) of interest and describe it (them) in the context of the problem situation.
- 2. State the Hypotheses.
- 3. Calculate the appropriate test statistic, df (if necessary), find the P-value (for z tests, you will need to be able to calculate this, for t tests, the value will be provided.)
- 4. Decision, reason for your decision including numbers, conclusion in the problem context (see SLIDEs for the wording).
- 16. The following are the test statistics (Step 3):

	one-sample		two-sample	
			independent	pairs
μ	$z_{ts} = \frac{\overline{x} - \mu_o}{\sigma / \sqrt{n}}$	$t_{ts} = \frac{\overline{x} - \mu_o}{s/\sqrt{n}}$	$t_{ts}' = \frac{\overline{x}_1 - \overline{x}_2 - \Delta_0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$	$t_{ts} = \frac{\overline{d} - \Delta_0}{s_D / \sqrt{n}}$
df	N/A	n - 1	will be given to you	n - 1

a) Be able to determine which column to use.

For the two-sample, be able to explain your answer.

b) Δ_0 is the symbol that is used for the null value for the two-sample cases.

c) Note: for the two-sample cases, we will provide you with enough information so that you can do the calculation either via the independent or pairs methodology.

17. The following are methods for calculating the P-values (Step 3)

	Alternative Hypothesis	z (by hand)	t (know equation only)	
upper tailed*	H _a : $\mu > \mu_0 (\Delta_0)$	$P(Z > z_{ts}) = 1 - P(Z \le z_{ts})$	$P(T > t_s)$	
lower tailed*	$H_a: μ < μ_0 (Δ_0)$	$P(Z < z_{ts})$	$P(T < t_s)$	
two-tailed	H _a : μ ≠ μ₀ (∆₀)	$2(P(Z > z_{ts})) = 2(1 - P(Z < z_{ts})) = 2 P(Z < - z_{ts})$	$2(P(T > t_s))$	
For the type blocks, if you need the Divelve, it will be given to you				

For the t problems, if you need the P-value, it will be given to you.

- 18. Be able to state what the meaning of the P-value is.
- 19. Be able to determine if a certain error is Type I or Type II and which error is possible in a certain circumstance.
- 20. Define and interpret the power (Power = 1β) and state what factors affect the power and what the effects are.
- 21. Be able to calculate the power given the H_a, α and the true value of the mean, μ_a .
- 22. Be able to compare the results of the CI with the hypothesis test. This includes how they are similar and how they are different.a) Given the results of the CI or hypothesis test, be able to predict the other one.
- 23. Be able to state the meaning of statistical significance and be able to differentiate between practical significance and statistical significance.
 - a) Given enough information, be able to determine what the practical significance of the results of the inference is.

Chapter 11

- 24. Be able to explain why variances are used to determine if the means are different (background of ANOVA).
- 25. Be able to determine the factor, the levels of the factor and the response variable, the populations to be compared and the total number of observations for ANOVA problems.
- 26. State the assumptions required for ANOVA to be valid, state how you would check them, and draw conclusions from the resulting graphs or provided information (you will not be required to make any graphs).
 - a) Normality: QQ plot, histogram
 - b) Constant standard deviation: the ratio of standard deviations.

27. Be able to calculate the one-way ANOVA table via hand (fill in the boxes). You will not have to calculate any summations directly; they will be able to be calculated via addition, subtraction, multiplication and/or division.

Source	df	SS (Sum of Squares)	MS (Mean Square)
Factor A	k - 1	$\sum_{i=1}^{k} n_i (\bar{x}_{i.} - \bar{x}_{})^2$	SSA dfa
Error	N - k	$\sum_{i=1}^{k} \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)^2 = \sum_{i=1}^{k} (n_i - 1) s_i^2$	SSE dfe
Total	N - 1	$\sum_{i=1}^{k} \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_{})^2$	

a) SST = SSA + SSE

a) dft = dfa + dfe

28. Be able to calculate the appropriate test statistic:

$$F = \frac{MSA}{MSE}, df1 = dfa, df2 = dfe$$

- 29. Be able to perform the 4-step hypothesis test for one-way ANOVA (the P-value will be given to you).
- 30. Be able to state the advantages for using a t distribution versus an F distribution and why an F distribution would be preferred.
- 31. Be able to explain why you have to use a special technique to perform multiple comparison analysis.
- 32. Be able to determine when to use Tukey's multiple comparison analysis.
- 33. Be able to perform Tukey's multiple comparison analysis indicating the factors that are different, presenting a visual display and presenting the final results in 'layman's' terms. Note because of the time involved in the calculations, you might just be required to do a), b), c) below and determine the visual display from the resulting confidence intervals.
 a) x
 _{i.} x
 {j.} ± t^{**}{column,df} · SE

b)
$$t_{column,df}^{**} = \frac{Q_{\alpha,k,n-k}}{\sqrt{2}}$$

c) $SE = \sqrt{MSE\left(\frac{1}{n_i} + \frac{1}{n_j}\right)} = \sqrt{\frac{2MSE}{n_i}}$
d) Example visual display:
0 mg (control) 20 mg 40 mg

0 mg (control)	20 mg	40 mg
57.60	69.28	75.70
		-