Stat 2470, 4/7 Discussion Questions Name ______________________________

Instructions: Attempt to answer these questions by reading the textbook or with online resources before coming to class on the date above.

1. What does the term ANOVA stand for?

Analysis of Variance

2. What is the assumption for the null hypothesis for ANOVA?

$H_0: \mu_1 = \mu_2 = \mu_3 = \cdots = \mu_n$ all means are equal

3. What is the alternative hypothesis for the ANOVA test?

$H_a: \mu_i \neq \mu_j$ for some $i \neq j$ at least one mean is different

4. When do we use ANOVA? How is it different than the two-sample T-test?

we use it when we have more than two samples.
It tests all the means at once for equality

5. What is the grand mean? What is the formula for it and what does it mean in words?

$$\bar{X}_{\cdot\cdot} = \frac{\sum_{i=1}^{I}\sum_{j=1}^{J} x_{ij}}{IJ}$$

add up all the values over all the treatments & divide by the total # of values; it's the average of all the data combined.

6. What is the formula for the mean of each treatment?

$$\bar{X}_{i\cdot} = \frac{\sum_{j=1}^{J} x_{ij}}{J}$$

7. *I* represents what in the ANOVA test? What does *J* represents?

I: the # of treatments; J: the # of samples in each treatment

8. What is the formula for the standard deviation of each treatment? What does the formula say to do in words?

$$S_i^2 = \frac{\sum_{j=1}^{J}(X_{ij} - \bar{X}_{i\cdot})^2}{J-1}$$

Square root for S_i

find the variance of each treatment separately (conventially)

find the square of the distance from each value to its treatment mean; add up & divide by sample size minus one.

9. What is the formula for the mean square for treatments?

$$MSTr = \frac{J}{I-1}\left[(\bar{X}_{1\cdot}-\bar{X}_{\cdot\cdot})^2 + (\bar{X}_{2\cdot}-\bar{X}_{\cdot\cdot})^2 + \dots\right] = \frac{J}{I-1}\sum(\bar{X}_{i\cdot}-\bar{X}_{\cdot\cdot})^2$$

treatment mean minus grand mean, square results, add up and divide by $I-1$ multiply by J.

10. What is the formula for the mean square for error?

$$MSE = \frac{S_1^2+S_2^2+\dots+S_I^2}{I} = \frac{\sum S_i^2}{I}$$

11. What is the formula for the F test statistic?

$$F = \frac{MSTr}{MSE}$$

12. Which of these values come out of the ANOVA function in the calculator?

MSTr, MSE, F

13. The F test statistic has two measures of degrees of freedom (they are sometimes referred to by "numerator" and "denominator" degrees of freedom from the formula for F). What is the formula for each?

numerator degrees of freedom is $I-1$
denominator degrees of freedom is $I(J-1)$

14. How do we find the P-value from the calculator given the test statistic and the degrees of freedom?

Fcdf(F statistic, E99, $I-1$, $I(J-1)$) from DISTR menu

15. What are the formulas for SST, SSTr and SSE, and what does each mean?

$$SST = \sum_{i=1}^{I}\sum_{j=1}^{J}(x_{ij}-x_{\cdot\cdot})^2 = \sum_i\sum_j x_{ij}^2 - \frac{1}{IJ}x_{\cdot\cdot}^2$$

$$SSTR = \sum\sum(\bar{x}_{i\cdot}-\bar{x}_{\cdot\cdot})^2 = \frac{1}{J}\sum x_{i\cdot}^2 - \frac{1}{IJ}x_{\cdot\cdot}^2$$

$$SSE = \sum\sum(x_{ij}-x_{i\cdot})^2$$

16. How are SST, SSTr and SSE related?

$$SST = SSTr + SSE$$

17. Run an ANOVA test in your calculator using data in problem #4 in section 10.1 and use that data to fill in the ANOVA table shown at the bottom of page 399 (Table 10.2 in the 8th edition).

enter data in L_1, L_2, L_3, L_4 run Stat → Tests → ANOVA(L_1, L_2, L_3, L_4)

$F = 2.3144$ $p = .2175$

Factor/treatment
$df = 3$
$SSTr = 15.605$
$MSTr = 5.20$

Error
$df = 4$
$SSE = 8.99$
$MSE = 2.2475$

$S_{xp} = 1.499$

18. When do we use Tukey's procedure?

when we reject Ho under the ANOVA test

19. What is the formula for Tukey's procedure?

$$\bar{X}_{i.} - \bar{X}_{j.} \pm \underbrace{Q_{\alpha, I, I(J-1)} \sqrt{MSE/J}}_{\text{margin of error}}$$

20. How do we calculate Q? What are the values for the degrees of freedom we need?

use the Q table in back of book

I (# of treatments) = V or n, $I(J-1) = m$ ← typically the larger value

21. Describe the steps for conducting Tukey's procedure.

1. reject Ho w/ standard ANOVA
2. Calculate The means for each treatment
3. Sort the means by size.
4. Calculate $Q_{\alpha, I, I(J-1)} \sqrt{MSE/J}$
5. use this value to determine how the means are grouped (you can subtract all the pairs of means and determine if any values are larger than #4; or use it to create an interval around each mean and see which others are inside it).
6. underline the means that group together; groups may be disjoint or overlap