STAT 252

Midterm Review Problems (proposed by Jeff Andrews)

DISCLAIMER:
These problems are intended to give you practice of the material in an 'unsorted' manner. They are NOT intended as an all-encompassing list of the types of questions I can or will ask you. Please review your assignments and in-class exercises to prepare for the test as well.

SUGGESTION:
Attempt these problems using only the formula sheet and statistical tables.

1. Find the p-value associated with a test statistic of $t = -1.97$ with $\nu = 24$ if it's...
   (a) a lower-tailed test.
   (b) a two-tailed test.
   (c) an upper-tailed test.
2. Give the model for a two-way ANOVA with factorial design. Explain each term.
3. Give the assumptions for the error term in all the ANOVA models we have discussed thus far. What are the implications of this, with respect to performing an ANOVA test?
4. When carrying out a hypothesis test at $\alpha = 0.01$ to see if a new drug is more effective than an established drug, we have the following null and alternative hypotheses

$$H_0 : \mu_{\text{new}} \leq \mu_{\text{old}}$$

$$H_a : \mu_{\text{new}} > \mu_{\text{old}}$$

   and after much calculation, we determined a p-value of 0.03.
   (a) Carefully and briefly define a p-value.
   (b) Based on this information, make a decision and interpret that decision.
5. Describe the Central Limit Theorem and its importance with regards to inferential statistics.
6. A sample of 45 people were taken, 15 from each of three levels of training (beginner, intermediate, advanced) and asked to perform the same task. Time to completion was recorded, and The following calculations were made:

$$SSTR = 180.1 \quad SST = 701.56$$

   Carry out an appropriate hypothesis test at $\alpha = 0.05$ to investigate if training level has an effect on average completion time.
7. Set up a contrast to investigate the difference between the beginner and intermediate groups.

8. A data set contains weight measurements on 30 plants grown under 3 different experimental conditions (10 in each treatment). You may assume the data is normally distributed. Here is some summary information:

	Control	Treatment 1	Treatment 2
T_j	50.32	46.61	55.26
n_j	10	10	10
s_j^2	0.34	0.63	0.20

and $\sum x_i^2 = 786.32$

   (a) Create an ANOVA table for this data.

   (b) Perform a test at $\alpha = 0.05$ to see if there exists a difference in plant growth among the three treatments.

   (c) Find all pairwise confidence intervals Bonferroni's procedure. Summarize your findings using the method from class.

9. If your confidence interval contains your hypothesized parameter, state whether the following statements are true, false, or if you cannot be certain with the information provided. For a hypothesis test with the same confidence level...

   (a) you would reject H_0: $\mu = \mu_0$

   (b) you would infer that $\mu \neq \mu_0$

   (c) you would fail to reject H_0: $\mu = \mu_0$

   (d) you would reject H_0: $\mu \leq \mu_0$

   (e) you would fail to reject H_0: $\mu \geq \mu_0$

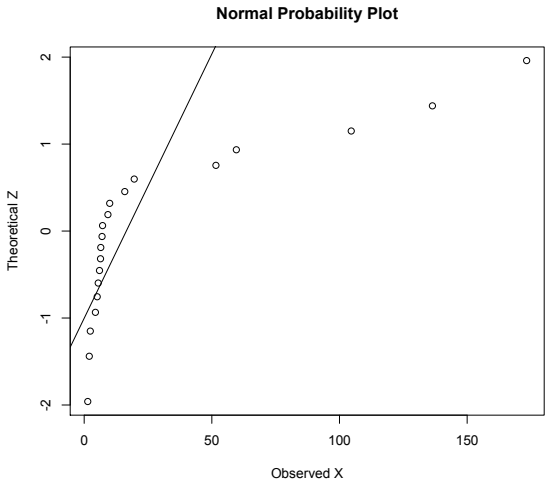
10. I take a random sample of size 20 from a population and calculate a 90% confidence interval for the true population mean. If I repeat this whole process 130 times, how many times would I expect the confidence interval to contain the true population mean?

11. In a randomized block design ANOVA, the F-value calculated by $\frac{MSB}{MSE}$ had a p-value associated with it that is smaller than α. Explain the implications of this result.

12. Provide a potential diagram to show the following effects in a two-way ANOVA with factorial design. Assume Factor A has 4 levels and Factor B has 2 levels.

    - Factor A: No significant effect
    - Factor B: Significant effect
    - Interaction: Significant effect

13. Below is a normal probability plot for sample of 20 observations.



(a) Interpret this plot.

(b) If we created a histogram from this sample, how would the shape best be described.

(c) If we perform a one-sample t-test on this data, should we be concerned about interpreting the results? Why or why not?