Practice Final Exam
STAT 705 Online Name _____

**Question 1.**

The owner of a pet shop breeds guinea pigs to sell in his shop. Female guinea pigs typically give birth to numerous newborns at a time (called a litter), and the owner is concerned about the weight of the newborn pigs. He measured the weights (in ounces) of newborns in seven separate litters, and conducted an ANOVA analysis to compare the mean weights of the litters. Refer to the SAS output to answer the following questions.

(a) (4 points) How many newborn guinea pigs are represented in the data? **47**

(b) (5 points) Does it appear that any of the assumptions are violated? Explain.

**The assumptions appear to be satisfied. The normal probability plot looks a bit curved, but it is not severe enough for us to declare that the errors are not normal. The Brown-Forsythe test has p-value 0.3829, so we do not reject the hypothesis that the variances are equal.**

(c) (8 points) Is there a significant difference between mean weights of the litters? If so, which litter(s) have significantly different means? Justify your answer.

**Yes, there is a significant difference between the mean weights of the litters. The overall ANOVA F test has p-value less than 0.0001.**

**Next… compare all pairs of treatment means using Tukey-adjusted p-values. The following pairs of litters have different means:**
**Litter 1 & Litter 4**
**Litter 1 & Litter 5**
**Litter 1 & Litter 6**
**Litter 2 & Litter 5**
**Litter 3 & Litter 5**
**Litter 5 & Litter 7**
**Litter 6 & Litter 7**

(d) (4 points) Would it be appropriate to test for linear or quadratic trends in these data? Explain.

**No. The litters are numbered, but these numbers are simply designations (i.e. categorical). We cannot treat the litters as numeric data. Linear and quadratic trends make sense only when the treatments are numeric.**

**Question 2.**

(8 points) After analyzing the data in Question 1, the owner of the pet shop added an eighth litter to his data set and re-ran the ANOVA analysis. Use the provided SAS output to answer the same 4 questions that are in Question 1.

**There are 57 observations in the data set.**
**The assumption of equal variance has been violated (Brown-Forsythe p-value = 0.0020).**
**We cannot answer the remaining questions because the assumptions have been violated.**

**Question 3.**

(12 points) An experiment was conducted to examine the effect of music on the productivity of factory workers. For three consecutive weeks, on Mondays, Wednesdays and Fridays, one of three types of music were played in the factory and the overall productivity of the factory was measured. (Higher values indicate a more productive workforce.) The three levels of music are pop/rock (level A), classical (level B) and no music (level C).

Use the SAS output to analyze these data, and provide a recommendation to the factory supervisor which level of music, if any, generates significantly higher productivity.

(Note: Your answer should be complete and concise. Cite relevant SAS output to justify your conclusions.)

**No assumptions appear to be violated.**

**The overall F test has p-value 0.0246, so there are significant differences among the means.**

**Day*Music interaction is not significant (p=0.2356), so we can compare marginal means.**

**Main effect of Day is not significant (p = 0.2938)**

**Main effect of Music is significant (p = 0.0038)**

**Music A (pop/rock) has significantly lower productivity than Music B (classical), p = 0.0028).**

**There is not a significant difference between A & C or between B & C.**

**To achieve the greatest productivity, they should play either classical music or no music.**

**Question 4.**

In a study of reaction time under the influence of alcohol, age is thought to be another factor that could affect the time. Thirty six test subjects (individuals) were classified into three age groups: 20-39, 40-59, and 60+, with 12 people per group. In each age group, alcohol (0 oz., 1 oz., or 2 oz.) was randomly assigned to 4 individuals. Thirty minutes after consuming the alcohol, the reaction time (in seconds) was measured.

Use the provided SAS output to answer these questions.

(a) (4 points) One of the assumptions is that the errors follow a normal distribution with mean 0 and variance $\sigma^2$. What is the estimate for $\sigma^2$?

**MSE = 0.1362**

(b) (4 points) Is alcohol a treatment factor or a blocking factor? **treatment**

(c) (4 points) Is age group a treatment factor or a blocking factor? **blocking**

(d) (5 points) Does the difference in reaction time for the different amounts of alcohol depend on the age group? Provide the test statistic and the p-value of the test you use to answer this question.

**Yes. The interaction is significant. The test statistic is F = 34.96 and the p-value is < .0001.**

(e) (6 points) Calculate Fisher's LSD for these data. (Use $\alpha$ = 0.05.)

$$LSD = t_{\alpha/2}\sqrt{MSE\left(\frac{1}{n_i}+\frac{1}{n_j}\right)} = 2.052\sqrt{0.1362037\left(\frac{1}{4}+\frac{1}{4}\right)} = 0.5355$$

**Notes:**
**2.052 is from the t table with $\alpha/2$ = 0.025 and df = dfError = 27.**
**The two n's are the number of observations in each treatment.**

(f) (5 points) Use Fisher's LSD to determine if there is a significant difference in mean reaction time between the middle age group (40-59) who have consumed no alcohol to the younger age group (20-39) who have consumed 1 oz. of alcohol.

**Age = 2, alcohol = 0 $\Rightarrow$ sample mean = 5.225**
**Age = 1, alcohol = 1 $\Rightarrow$ sample mean = 4.825**
**| 5.225 – 4.825 | = 0.4. This is less than the LSD, so we conclude the difference is NOT significant.**

**Question 5.**

Officials of a marketing research corporation were interested in studying the effect of a new promotional campaign for an improved brand of D-cell batteries. They randomly selected four Standard Metropolitan Statistical Areas (SMSAs) in which to conduct their study. Within each of their SMSAs, they randomly selected three stores and recorded the weekly sales volumes (in dollars) for each store for two weeks following the promotional campaign. The following table shows the ANOVA degrees of freedom, sums of squares, and mean squares for the various sources of variability.

| Source | DF | SS | MS |
|---|---|---|---|
| **Store** | 2 | 62973.0 | 31486.5 |
| **SMSA** | 3 | 136644.1 | 45548.0 |
| **Store*SMSA** | 6 | 54127.0 | 9021.2 |
| **Error** | 12 | 4913.5 | 409.6 |

(a) (5 points) What is the value of the F statistic for testing the main effect of Store, assuming SMSA is a random effect?

$$\boldsymbol{F = \frac{MS(Store)}{MS(Store * SMSA)} = \frac{31486.5}{9021.2} = 3.49}$$

(b) (5 points) What is the value of the F statistic for testing the main effect of Store, assuming SMSA is a fixed effect?

$$\boldsymbol{F = \frac{MS(Store)}{MSE} = \frac{31486.5}{409.6} = 76.87}$$

(c) (6 points) Give two reasons why SMSA should be a random effect.

**(1) The SMSA's were randomly selected from a large population of all SMSA's.**

**(2) We want the results to apply to the population of all SMSA's.**

**Question 6.**

(7 points) A pharmaceutical company would like to examine the potency of a liquid medication mixed in large vats. To do this, a random sample of five vats from a month's production was obtained, and four separate vials were selected from each vat. The potency of the medication in each vial was measured. In the ANOVA table given below, Vat is assumed to be a fixed effect. If we assume Vat is a random effect, which contributes more to the variability in potency: the vats or the vials?

| Source | DF | Sum of Squares | Mean Square | F Value | Pr > F |
|---|---|---|---|---|---|
| **Model** | 4 | 181.2830 | 45.32075 | 1.02 | 0.4274 |
| **Error** | 15 | 664.9025 | 44.32683 | | |
| **Corrected Total** | 19 | 846.1855 | | | |

$$\widehat{\sigma}^2_{VAT} = \frac{MS(Vat) - MSE}{4} = \frac{45.32075 - 44.32683}{4} = 0.25$$

$MSE > \widehat{\sigma}^2_{VAT}$

**The vials contribute more than the vats to the variability in potency.**

**Question 7.**

Suppose we want to test the effectiveness of two different headache remedies (A and B) and we have eight people who are willing to serve as experimental units.

(a) (4 points) A representation of this experiment is shown below. For each person, write either A or B to indicate one possible random assignment of headache remedies to individuals.

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| **A** | **A** | **A** | **B** | **B** | **B** | **B** | **A** |
| Person 1 | Person 2 | Person 3 | Person 4 | Person 5 | Person 6 | Person 7 | Person 8 |

**(Need to have 4 A's and 4 B's in any order.)**

(b) (4 points) We are still considering the same experiment, but now assume that persons 1, 2, 3 and 4 are male, while the remaining people are female. We want to block on Gender. For each person, write either A or B to indicate one possible random assignment of headache remedies to individuals.

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| **A** | **B** | **B** | **A** | **B** | **B** | **A** | **A** |
| Person 1 | Person 2 | Person 3 | Person 4 | Person 5 | Person 6 | Person 7 | Person 8 |

**(Need to have 2 A's and 2 B's in Persons 1 to 4, and 2 A's and 2 B's in Persons 5 to 8.)**