STT 667 sample exam, winter 2009, Dr. Dan Voss

*Instructions:* SAS output is provided for the first two problems. In answering questions, refer to the SAS output in your answers as appropriate.

1. An experiment was conducted to determine the effects of four different pesticides (factor $A$) on the yield of fruit ($y$) from three different varieties of a citrus tree (factor $B$). Eight trees from each variety were randomly selected from an orchard. The four pesticides were then randomly assigned to trees of a particular variety and applications were made according to recommended levels. Yields of fruit, in bushels per tree, were obtained after the test period. These data appear in the SAS program "fruit.sas" and its output, given in the first six of the eight SAS pages provided with this exam.

   (a) Use the residual plots in the SAS output to check the model assumptions.

   (b) Page 7 of the SAS output contains a plot of the estimated treatment means. Use this plot to assess the factor effects $A$ and $B$.

   (c) Discuss the results of the analysis of variance. Are the results consistent with your interpretation in part (b)?

   (d) Suppose the experimenter anticipated that the factors might interact so decided in advance to use Tukey's method for all pairwise comparisons of the 12 cell means $\mu_{ij}$. Determine the minimum significant difference for 90% confidence intervals for the pairwise comparisons. For which treatment combination was the estimated cell mean largest? Which treatment combinations were significantly worse that this one, based on Tukey's method? (You need not actually construct the confidence intervals.)

   (e) Suppose the experimenter decided in advance to use Scheffé's method to construct simultaneous 90% confidence intervals for all possible treatment contrasts. Determine the critical value, $S$, to be used. Determine the corresponding minimum significant difference for comparing any two levels of factor $A$—namely, for main-effect-of-$A$ pairwise differences. What can you conclude about which level or levels of $A$ are best?

   (f) Suppose the experimenter decided just to examine pairwise main effect comparisons for each factor, in view of the analysis of variance results. The experimenter decided to use the method of Tukey for each factor, with a 95% confidence level for each set of main effect comparisons. Construct the corresponding confidence intervals for the main effect of $B$, and interpret the results.

   (g) Briefly discuss the relative merits of the methods of analysis proposed by the experimenter in parts (e) and (f).

2. A second experiment like that of problem 1 was subsequently conducted in the same orchard. The experimenter again studied the effects of four pesticides and three varieties, but this time they were willing to assume based on the results of the first experiment that the factors would not significantly interact. Thus, a single replicate experiment was conducted—namely, with only one observation per cell. To analyze the data, they ran the SAS program "fruit2.sas" provided with this exam—the program and output are on the last two of the eight SAS pages provided with this exam. The experimenter was dismayed by the resulting analysis of variance table. Indicate what is wrong with the program, and give the results of the correct analysis of variance, using a significance level of 5% for each test.

3. An experiment was conducted to study the effects of two factors on a response variable $y$. Two levels of each factor were included in the experiment. The SAS output reproduced below was generated by the following SAS statements.

```
proc glm;
  class A B;
  model y = A | B;
  means A*B;
```

The experimenter wanted to construct an individual 95% confidence interval for the main effect comparison of the two levels of factor $A$. Do this at the bottom of this page, and interpret the results.

```
=============================================================================
General Linear Models Procedure

Dependent Variable: Y
                                  Sum of          Mean
Source                    DF       Squares        Square  F Value    Pr > F
Model                      3     38.898620    12.966207    11.05    0.0016
Error                     10     11.731715     1.173172
Corrected Total           13     50.630335

                  R-Square          C.V.     Root MSE             Y Mean
                  0.768287      3.933571       1.0831             27.536

Source                    DF     Type I SS  Mean Square  F Value    Pr > F
A                          1     31.001889    31.001889    26.43    0.0004
B                          1      6.236964     6.236964     5.32    0.0438
A*B                        1      1.659768     1.659768     1.41    0.2617

Source                    DF   Type III SS  Mean Square  F Value    Pr > F
A                          1     32.431557    32.431557    27.64    0.0004
B                          1      6.236964     6.236964     5.32    0.0438
A*B                        1      1.659768     1.659768     1.41    0.2617
=============================================================================
General Linear Models Procedure

Level of   Level of          --------Y--------
A          B           N     Mean        SD
1          1           4     26.92       0.349
1          2           3     24.88       1.227
2          1           4     29.30       1.432
2          2           3     28.65       1.048
=============================================================================
```

4. A study of the relationship between facility conditions at gasoline stations and aggressiveness in the pricing of gasoline ("An Analysis of Price Aggressiveness in Gasoline Marketing." *J. Marketing Research*, 1970: 36–42) reports the accompanying data based on a sample of $n = 441$ stations. Do the data suggest that facility conditions and pricing policy are independent of one another? Test using a 1% significance level.

| | Observed pricing policy | | |
|---|---|---|---|
| Condition | Aggressive | Neutral | Nonaggressive |
| Substandard | 24 | 15 | 17 |
| Standard | 52 | 73 | 80 |
| Modern | 58 | 86 | 36 |

5. The accompanying data resulted from an experiment to compare the effects of vitamin C in orange juice and in synthetic ascorbic acid on the length of odontoblasts in guinea pigs over a 6-week period ("The Growth of the Odontoblasts of the Incisor Tooth as a Criterion for the Vitamin C Intake of the Guinea Pig," *J. Nutrition*, 1947: 491–504). Use the Wilcoxon rank-sum test at level 0.01 to decide whether true average length differs for the two types of vitamin C intake. Compute also an approximate $p$-value.

Orange juice: 8.2 9.4 9.6 10.0 14.5 15.2 16.1 17.6 21.5
Ascorbic acid: 4.2 5.2 5.8 6.4 7.0 7.3 10.1 11.2 11.3 11.5