STA 4211 – Spring 2002 – Exam 1

Print Name: **SSN:**

1) An epidemiologist studies the concentration of lead in children's blood in $r = 6$ sections of her city. She sets up an Analysis of Variance, wishing to test whether mean lead levels vary among the city's five sections. Samples of 5 children are obtained from each of the sections and lead concentrations are assayed from samples of blood. The between and within section sums of squares for lead levels are $SSTR = 1200$ and $SSE = 4000$, respectively.

Give the test statistic, rejection region, and conclusion for testing (under the usual normality, constant variance, and independence assumptions among the error terms):

$$H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5 = \mu_6 \qquad \text{vs} \qquad H_A : \text{ Not all } \mu_i \text{ are } =$$

where μ_i is the true (population) mean lead concentration among children in section i. Conduct the test at the $\alpha = 0.05$ significance level.

a) Test Statistic:

b) Rejection Region: (also sketch it)

c) Conclusion:

d) Based on your test, give a range for your P–value.

2) An HMO accountant has determined that the mean monthly expenditures on prescription medications varies among three groups of patients (those aged 25-44 (group 1), 45-64 (2), and 65+ (3)), based on the F-test for the Completely Randomized Design. The model fit is:

$$Y_{ij} = \mu_i + \varepsilon_{ij} \qquad \varepsilon_{ij} \sim NID(0, \sigma^2)$$

She wishes to make comparisons among all three groups, simultaneously, at the $\alpha = 0.05$ significance level. She obtains the following information from an extensive search of the records:

$$n_1 = n_2 = n_3 = 40 \quad MSE = 720.0 \quad \overline{Y}_{1.} = 160 \quad \overline{Y}_{2.} = 125 \quad \overline{Y}_{3.} = 175$$

a) Use Bonferroni's method to compare the true means among all pairs of groups with an overall experimentwise error rate of $\alpha = 0.05$. Clearly state your conclusions.

b) Repeat part a) for Tukey's method.

3) When a Completely Randomized Design has been conducted, for the treatment sum of squares in the ANOVA to be 0, which of the following conditions would have to occur?

a) Every observation within a given treatment would have to be the same.
b) All observations in the entire sample would have to be the same.
c) The treatment means would all have to be the same.
d) The treatment standard deviations would all have to be the same.
e) None of the above.

4) When a Completely Randomized Design has been conducted, for the error sum of squares in the ANOVA to be 0, which of the following conditions would have to occur?

a) All observations within common treatments would be the same.
b) All observations in the entire sample would be the same.
c) The treatment means would all be the same.
d) The treatment standard deviations would all be the same.
e) None of the above.

5) A model is fit to compare four brands of hair color for color intensity. A hairdresser samples 12 of his friends randomly assigning three to brand A, three to brand B, three to brand C, and three to brand D. He then measures the color intensity after coloring each subject's hair. He fits the cell means model:

$$Y_{ij} = \mu_i + \varepsilon_{ij}$$

and obtains the following data with $\overline{Y}_{..}$:

	Brand			
j	A ($i = 1$)	B ($i = 2$)	C ($i = 3$)	D ($i = 4$)
1	72	76	92	64
2	75	80	88	61
3	78	72	84	58

a) Give the Analysis of Variance.

b) Give the $\mathbf{X}$, $\mathbf{X'X}$, $\mathbf{X'Y}$ and $\mathbf{b} = (\mathbf{X'X})^{-1}\mathbf{X'Y}$ matrices and vectors for the cell means model, where:

$$\boldsymbol{\beta} = \begin{bmatrix} \boldsymbol{\mu_1} \\ \boldsymbol{\mu_2} \\ \boldsymbol{\mu_3} \\ \boldsymbol{\mu_4} \end{bmatrix}$$

6) A plant manager is interested in comparing three production line configurations in terms of daily output in her factory. She randomly selects $n_T = 12$ days in her present quarter, observing $n_1 = n_2 = n_3 = 4$ days under each configuration. The data are below:

j	$i = 1$	$i = 2$	$i = 3$
1	40	50	70
2	35	45	65
3	45	50	75
4	40	55	70
$\overline{Y}_{i.}$	40	50	70

Use the general linear test approach to set up the test to determine whether the true means of the first two configurations are equal by completing the following steps (utilize the entire Analysis of Variance, not just the first two treatments):

a) H_0 : H_A :

b) Under H_0 : $\hat{\mu}_1 =$ $\hat{\mu}_2 =$ $\hat{\mu}_3 =$

c) Under H_A : $\hat{\mu}_1 =$ $\hat{\mu}_2 =$ $\hat{\mu}_3 =$

d) **Set up** SSE for reduced model and give its degrees of freedom:
$SSE(R) =$ $df_E(R) =$

e) **Set up** SSE for full model and give its degrees of freedom:
$SSE(F) =$ $df_E(F) =$

f) Give the test statistic for testing the null and alternative hypotheses in a). You can use symbols for $SSE(R)$ and $SSE(F)$.

g) Clearly state the values for the test statistic in d) for which you'd reject the null hypothesis in a) for the altenative hypothesis (based on $\alpha = 0.05$ significance level).