

STA 4211 – Spring 2002 – Exam 3

Print Name:

SSN:

1) A clothing manufacturer has three types of machines that create rolls of fabric from yarn. Further, there are three operators for each machine. For instance, for machine A, Adam, Anna, and Arnold are the operators. For machine B, Bill, Brenda, and Barbara are the operators. For machine C, Caren, Charlie, and Chris are the operators. The plant engineer is interested in studying the variation in the strength of the fabric produced across machines and operators. She samples $n = 4$ pieces of fabric from each operator and measures the breaking strength of each sampled fabric piece. **35 Points**

a) Write out the statistical model, clearly stating all parameters, and ranges of subscripts.

b) The means (and standard deviations) are given in the following table for each machine/operator. Give the parameter estimates for all model parameters.

Operator	Machine A	Machine B	Machine C
1	25.0 (2.0)	31.0 (3.0)	18.0 (1.0)
2	28.0 (3.0)	30.0 (2.0)	25.0 (3.0)
3	22.0 (1.0)	29.0 (2.0)	17.0 (2.0)

c) Give the Analysis of Variance, and conduct all relevant tests (for machine and operator effects)

2) A chemical company scientist is interested in the relation of potency of a chemical compound (Y) to two factor: temperature that the compound was prepared at (X_1 , in degrees Fahrenheit) and the preparation time (X_2 , in minutes). They are interested in estimating the response surface:

$$E(Y) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_{12} X_1 X_2 + \beta_{11} X_1^2 + \beta_{22} X_2^2$$

over the range: $200 \leq X_1 \leq 400$ and $30 \leq X_2 \leq 90$. **12 Points**

a) Give the design points (settings of X_1 and X_2) for the central composite design with star points located at a (scaled) distance of $\alpha = 1.50$ from the center point.

b) How many center points would you run if you can afford to make a total 12 observations?

3) 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. 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:

30 Points

$$n_1 = n_2 = n_3 = 10 \quad MSE = 640.0 \quad \bar{y}_1 = 240 \quad \bar{y}_2 = 200 \quad \bar{y}_3 = 280$$

a) Use Tukey's method to compare the true means among all pairs of groups by setting up simultaneous 95% confidence intervals for $\mu_i - \mu_j$ and interpreting them.

b) Give the Analysis of Variance table.

4) Order the following three configurations of the model parameters of the Completely Randomized Design (with 3 treatments, and n reps per treatment) in terms of the expected value of the Treatment mean square ($MSTR$) from 1=smallest to 3=largest (assume σ^2 is constant across configurations) **6 Points**.

- $\tau_1 = 10, \tau_2 = -5, \tau_3 = -5$ _____
- $\tau_1 = 10, \tau_2 = 0, \tau_3 = -10$ _____
- $\tau_1 = 0, \tau_2 = 0, \tau_3 = 0$ _____

5) Coke and Pepsi each hire three advertising agencies. Due to corporate secrecy, Coke uses agencies $C_1, C_2,$ and C_3 ; while Pepsi uses three different agencies: $P_1, P_2,$ and P_3 . Coke and Pepsi consider 4 possible strategies: (1) “feel good”, (2) celebrity endorsement, (3) humor, and (4) comparison (like a taste test). They ask each of their ad agencies to produce one ad of each type, then observe relative sales changes in randomly selected test markets (each ad produced is observed in two test markets). The statistical model used for analysis is (where i represents soda manufacturer, j represents ad agency, k represents ad type and l represent replicate):

$$Y_{ijkl} = \mu + \alpha_i + \beta_{j(i)} + \gamma_k + (\alpha\gamma)_{ik} + (\beta\gamma)_{jk(i)} + \varepsilon_{ijkl} \quad i = 1, \dots, a \quad j = 1, \dots, b \quad k = 1, \dots, c \quad l = 1, \dots, n$$

Fill in the degrees of freedom for the Analysis of Variance table **8 Points**:

Source	df
Soda Brand	
Agency (Brand)	
Ad Type	
Brand \times Ad Type	
Ad Type \times Agency (Brand)	
Error	
Total	

6) A large timber area is made up of a population of parcels of land (each consisting of one acre of land). On each parcel of land, is a large number of trees. A forest researcher is interested in the numbers of termites (per cubic foot) of trees in the area. She samples r parcels, and n trees per parcel, measuring the density of termites per cubic foot per tree. She makes the usual assumptions regarding distributions and independence among model components (this is a random effects model). She fits the model:

$$y_{ij} = \mu + \alpha_i + \varepsilon_{ij} \quad i = 1, \dots, r \quad j = 1, \dots, n$$

Across all trees in the timber area, the average density of termites per cubic foot is 100. Across the parcels of land, approximately 95% have mean termite densities between 50 and 150 termites per cubic foot. Among trees with a parcel of land, approximately 95% have a density within 20 of the parcel mean **7 Points**.

- a) μ = _____
- b) σ_ε^2 = _____
- c) σ_α^2 = _____