Statistics 5100 Exam #1 March 3rd, 2025 Instructor: Dr. Genschel

Good Luck! Name Sabrina Morse - Student A

Partner: Sam Olson

Please answer each question in the space provided, and **show all your work**. Credit cannot be given if work is not shown. Ask for extra paper if you need it. Good luck.

1. If I asked your best friend about one of your character strengths, what would they say?

____

2. **Technical and Conceptual.** A study was conducted to see if there are differences in the quality of steel produced by $m = 3$ machines. It is also felt that there may be differences in the feedstock obtained from $l = 3$ different suppliers. Nine samples of feedstock were selected from each supplier, and $n = 3$ samples were randomly assigned to each machine.

   In class we talked extensively about two different types of models to analyze data from this type of experiment. Most recently you encountered both models on Homework 5, Problem 1.

   (a) Name each statistical model. The order in which you name the models that does not matter!

   Model 1: Cell Means Model Model 2: Additive Model

   (b) For each, write out the statistical model to describe the quality y_{ijk} of the k^{th} piece of steel produced by the i^{th} machine using feedstock from supplier j.

   - Do **not** write the model in matrix or vector form!
   - Be sure to properly define all parameters, use clear and complete notation and state the **necessary assumptions** such that the model is also a **Gauss-Markov Model** with **Normal Errors**.
   - Your answer should look like one of the slides in the notes when we introduce a new statistical model.

   Model 1: $y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \epsilon_{ijk}$

μ = combined mean from both factors

α_i = machine main effect

β_j = supplier main effect

γ_{ij} = interaction effect between machine & supplier

ϵ_{ijk} = random error

Assumptions:

- $\epsilon_{ijk} \sim N(0, \sigma^2)$
- independence of random errors

Name ______________________________

Model 2: $y_{ijk} = \mu + \alpha_i + \beta_j + \epsilon_{ijk}$

μ = overall mean

α_i = macnine main effect

β_j = supplier main effect

ϵ_{ijk} = random error

Assumptions:

- $\epsilon_{ijk} \sim N(0, \sigma^2)$
- independence of random error
- no interaction between treatments

(c) **In the context of the data**, what is the main difference between both types of statistical models in terms of model complexity?

Cell means is more complex than additive model. Since it includes a term for potential interaction between macnines and suppliers

Name ______________________________

3. Suppose that $\boldsymbol{y} \sim \mathcal{N}(\boldsymbol{\mu}, \sigma^2, \boldsymbol{I})$, and let $q_i = \boldsymbol{y}^\top \boldsymbol{A}_i \boldsymbol{y}, i = 1, 2$ where

$$\boldsymbol{A}_1 = \frac{1}{3}\mathbf{1}\mathbf{1}^\top \qquad \text{and} \qquad \boldsymbol{A}_2 = \frac{1}{2}\begin{bmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Note that $\mathbf{1}$ is a vector of three ones, i.e., $\mathbf{1} = (1\ \ 1\ \ 1)^\top$.

(a) Determine the distribution of each quadratic form and explain how you know that this indeed is the distribution of $q_i, i = 1, 2$.

quadratic form eq: $y_T A y \sim \chi^2_{\underset{\text{rank of } A}{m}}\left(\frac{\mu^T A \mu}{2}\right)$

$A_1 = \frac{1}{3}\begin{bmatrix}1\\1\\1\end{bmatrix}[1\ 1\ 1] = \frac{1}{3}\begin{bmatrix}1&1&1\\1&1&1\\1&1&1\end{bmatrix} = \begin{bmatrix}1/3&1/3&1/3\\1/3&1/3&1/3\\1/3&1/3&1/3\end{bmatrix} \Rightarrow \text{rank} = 1$

$A_2 = \frac{1}{2}\begin{bmatrix}1&-1&0\\-1&1&0\\0&0&0\end{bmatrix} = \begin{bmatrix}1/2&-1/2&0\\-1/2&1/2&0\\0&0&0\end{bmatrix} \Rightarrow \text{rank} = 1$

$\therefore \quad y^T A_i y \sim \chi^2_1\left(\frac{\mu^T A \mu}{2}\right)$ now find ncp for each A_i

$\mu^T = [\mu_1\ \ \mu_2\ \ \mu_3]$

$A_1: \ A_1\mu = \begin{bmatrix}1&1&1\\1&1&1\\1&1&1\end{bmatrix}\begin{bmatrix}\mu_1\\\mu_2\\\mu_3\end{bmatrix} = \begin{bmatrix}\mu_1+\mu_2+\mu_3\\\mu_1+\mu_2+\mu_3\\\mu_1+\mu_2+\mu_3\end{bmatrix}$

$\mu^T A_1 \mu = [\mu_1\ \mu_2\ \mu_3]\begin{bmatrix}\mu_1+\mu_2+\mu_3\\\mu_1+\mu_2+\mu_3\\\mu_1+\mu_2+\mu_3\end{bmatrix}$

$= (\mu_1+\mu_2+\mu_3)^2$

so $y^T A_1 y \sim \chi^2_1\left(\frac{(\mu_1+\mu_2+\mu_3)^2}{6}\right)$

$A_2: \ A_2\mu = \begin{bmatrix}1&-1&0\\-1&1&0\\0&0&0\end{bmatrix}\begin{bmatrix}\mu_1\\\mu_2\\\mu_3\end{bmatrix} = \begin{bmatrix}\mu_1-\mu_2\\\mu_2-\mu_1\\0\end{bmatrix}$

$\mu^T A_2 \mu = [\mu_1\ \mu_2\ \mu_3]\begin{bmatrix}\mu_1-\mu_2\\\mu_2-\mu_1\\0\end{bmatrix} = \mu_1(\mu_1-\mu_2) + \mu_1(\mu_2-\mu_1) + \mu_2(\mu_1-\mu_2) + \mu_2(\mu_2-\mu_1)$

$= \cancel{\mu_1^2} - \cancel{\mu_1\mu_2} + \cancel{\mu_1\mu_2} - \cancel{\mu_1^2} + \cancel{\mu_2\mu_1} - \cancel{\mu_2^2} + \cancel{\mu_2^2} - \cancel{\mu_2\mu_1}$

$= 0$ so $y^T A_2 y \sim \chi^2_1(0)$

we know these are distributions of q_i for $i = 1,2$ since the normality of y ensures the quadratic form will follow χ^2 distribution of rank m & ncp of $\frac{\mu^T A_i \mu}{2}$

(b) Show that q_1 and q_2 are independent.

show $A_1 \perp A_2$ = independence of q_1 & $q_2 \Rightarrow A_1 \cdot A_2 = 0$ means orthogonal

$\begin{bmatrix}1/3&1/3&1/3\\1/3&1/3&1/3\\1/3&1/3&1/3\end{bmatrix} \cdot \begin{bmatrix}1&-1&0\\-1&1&0\\0&0&0\end{bmatrix} = 1/3 - 1/3 + 0 - 1/3 + 1/3 + 0 + 0 + 0 + 0 = 0$

since $A_1 \cdot A_2 = 0 \Rightarrow q_1$ & q_2 are independent ✓