Homework 8

Brian Cervantes Alvarez

June 4, 2024

ST 553 Statistical Methods

Problem 1

Solution

Given:

- $a = 2$ machines
- $b = 4$ operators
- $n = 2$ boxes per machine-operator combination

We can write the parameter vector and design matrix as follows,

$$\gamma = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \\ \alpha\beta_{1,1} \\ \alpha\beta_{2,1} \\ \alpha\beta_{1,2} \\ \alpha\beta_{2,2} \\ \alpha\beta_{1,3} \\ \alpha\beta_{2,3} \\ \alpha\beta_{1,4} \\ \alpha\beta_{2,4} \end{pmatrix} \quad Z = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Problem 2

Solution

Model Description

The experiment described involves multiple factors and nested structures. The model for the response variable, Y_{ijklm}, representing the time spent by the doctor with each patient, can be described as follows,

$$Y_{ijklm} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \gamma_k(i) + \delta_l + (\delta\gamma)_{lk} + \eta_m + \epsilon_{ijklm}$$

where:

- μ is the overall mean.
- α_i is the fixed effect of the i^{th} clinic ($i = 1, 2, 3, 4$).
- β_j is the random effect of the j^{th} doctor nested within the i^{th} clinic ($j = 1, 2, 3$).
- $(\alpha\beta)_{ij}$ is the interaction effect between the clinic and doctor.
- $\gamma_{k(i)}$ is the fixed effect of the k^{th} checkup type ($k = 1, 2$).
- δ_l is the fixed effect of the l^{th} insurance type ($l = 1, 2$).
- $(\delta\gamma)_{lk}$ is the interaction effect between the checkup type and insurance type.
- η_m is the random effect of the m^{th} group within each combination of checkup type and insurance type ($m = 1, 2, \ldots, 12$).
- ϵ_{ijklm} is the random error term.

Model Assumptions

1. The observations are independent.
2. The random effects and errors are normally distributed:

$$\beta_j \sim N(0, \sigma_\beta^2), \quad \eta_m \sim N(0, \sigma_\eta^2), \quad \epsilon_{ijklm} \sim N(0, \sigma^2)$$

3. The variance of the errors is constant across all levels of the factors.
4. Sum-to-zero constraints for the following fixed effects,

$$\sum_{i=1}^{4} \alpha_i = 0 \quad \sum_{i=1}^{2} \gamma_{k(i)} = 0 \quad \sum_{l=1}^{2} \delta_l = 0$$

Problem 3

Solution

Report

While soybean variety does not impact seed weight, the crop rotation pattern also does not show an effect within the context of this study. This implies that neither the choice of soybean variety nor the rotation pattern is critical for influencing seed weight within the studied parameters.

Appendix

ANOVA Results

- Variety: No effect on mean seed weight $F(1, 24) = 0.089$, $p = 0.768$.
- Rotation Pattern: No effect $F(3, 24) = 0.422$, $p = 0.739$.
- Interaction (Variety:Rotation): No interaction effect $F(3, 24) = 0.164$, $p = 0.920$.

Tukey's HSD Test The Tukey's HSD test showed no differences between the rotation patterns. In other words, the mean weights for the different rotation patterns are suggested to be very similar.

Mean Weights by Rotation Pattern

Rotation Pattern	Mean Weight (g)	Group
Rotation 1	158.375	a
Rotation 2	153.250	a
Rotation 3	151.375	a
Rotation 4	155.250	a

Problem 4

To achieve the same precision using a CRD for the experiment in Problem 13.5, we need to replicate each treatment combination equally. The original experiment has 8 treatment combinations (2 varieties $\times$ 4 rotation patterns) and 32 experimental units. Thus, to maintain the same precision, the total number of experimental units should remain the same. So, each treatment combination in the CRD should be replicated 4 times ($\frac{32 \text{ units}}{8 \text{ combinations}}$). Therefore, we would need 32 experimental units for the CRD, ensuring each treatment combination is replicated 4 times.