ESMA 6787: Exam 1

Due on Octubre 27, 2025

Israel Almodovar

Alejandro Ouslan

Contents

1 problem 1

Define the followings:

(a) p-value:

(b) Projection Matrix:

(c) Complete randomized Design:

(d) Complete Randomized Block Design:

2 Problem 2

In the experiment to study the effects of the amount of baking powder in a biscuit dough upon the rise heights of the biscuits, four levels of baking powder were tested and four replicate biscuits were made with each level in a random order.

Table 1: Training Methods Data

0.25 tsp	**0.5 tsp**	**0.75 tsp**	**1 tsp**
11.4	27.8	47.6	61.6
11.0	29.2	47.0	62.4
11.3	26.8	47.3	63.0
9.5	26.0	45.5	63.9

(a) What is the experimental unit?

(b) Under this model show why each of the following is estimable or non-estimable:

$$\tau_1, \tau_2, \tau_3, \tau_4, \tau_1, \quad \tau_1 + \tau_2 - (\tau_3 + \tau_4), \mu + \tau_1 + \tau_2$$

(c) Preform the analysis of variance to test the hypothesis of no treatment effect.

(d) Formulate a contrast to test the hypothesis that increase in rise height is a linear function of the increase in baking powder in the dough, and test this hypothesis.

(e) If the dough were made in batches and the four replicate biscuit rise heights in each column (Table 1) were all from the same batch, would your answer to (a) be different? How could the data be analyzed if this were the case?

3 Problem 3

Consider a completely randomized design with four treatment groups, in which a total of $N = 100$ units are to be used. Although it won't be explicitly used in the analysis model, treatments 1 through 5 actually represent increasing concentrations of one component in an otherwise standard chemical compound, and the primary purpose of the experiment is to understand whether certain measurable properties of the compound change with this concentration. The investigator decides to address these questions by estimating these quantities

$$\tau_2 - \tau_1, \tau_2 - \tau_3, \tau_2 - \tau_4, \tau_3 - \tau_4, \tau_3 - \tau_1$$

where each τ_i is a parameter in the standard effects model. Find the optimal allocation for the 100 available units (i.e., values for $n1, \cdots, n5$) that minimizes the average variance of estimates of the five contrasts of interest. Do this as a constrained, continuous optimization problem, then round the solution to integer values that are consistent with the required constraints.

4 Problem 4

The effect of plant growth regulators and spear bud scales on spear elongation in asparagus was investigated by Yang-Gyu and Woolley (2006). Elongation rate of spears is an important factor determining final yield of asparagus in many temperate climatic conditions. Spears were harvested from 6-year-old Jersey Giant asparagus plants grown in a commercial planting at Bulls (latitude 40.2S, longitude 175.4E), New Zealand. Spears were harvested randomly and transported from field to lab for investigation. After trimming to 80mm length, spears were immersed completely for 1 h in aqueous solutions of 10 mg l-1 concentration of indole-3-acetic acid (IAA), abscisic acid (ABA), GA3, or CPPU (Sitofex EC 2.0%; SKW, Trostberg, Germany) in test tubes. Control spears were submerged in distilled water for 1 h. The experiment was a completely randomized design with five replications (spears) per treatment

Table 2: Spear length in mm

Control	**IAA**	**ABA**	**GA3**	**CPPU**
94.7	89.9	96.8	99.1	104.4
96.1	94.0	87.8	95.3	98.9
86.5	99.1	89.1	94.6	98.9
98.5	92.8	91.1	93.1	106.5
94.9	99.4	89.4	95.7	104.8

(a) What are the experimental units?

(b) Proposed a linear model in this case. Explain each of the variable carefully.

(c) Test the hypothesis of no treatment effect. State the alternative, decision rule, and conclusion.

(d) Test all pairwise comparisons of treatment means. You can use either Bonferroni, Turkeyk, Scheffe or Dunnett, asssume that $\alpha = 0.01$.

(e) Can we test if there's difference between the control group with any of the treatment group? Discuss your results bases on part (c) and (d).

(f) Perform a non-parametric approach to test the hypothesis of no treatment effect. State the alternative, decision rule, conclusion and asuume that $\alpha 0.01$.

(g) Preform non-parametric pairwise comparisons of treatment effect. Discuss your results comparing the control gropu wiht the treatment groups. Further, asume that $\alpha = 0.01$.

5 Problem 5

An accounting firm prior to introducing in the firm widespread training in statistical sampling for auditing tested three training methods: (1) study at home with programmed training materials, (2) training sessions at local offices conducted by local staff, and (3) training sessions in Chicago con- ducted by national staff. Thirty auditors were grouped into 10 blocks of three according to time elapsed since college graduation and the auditors in each block were randomly assigned to the three training methods. At the end of the training each auditor was asked to analyze a complex case involving statistical applications proficiency measure based on this analysis was obtained for each auditor.

Table 3: Training Methods

Block	**1**	**2**	**3**
1	73	81	92
2	76	78	89
3	75	76	87
4	74	77	90
5	76	71	88
6	73	75	86
7	68	72	88
8	64	74	82
9	65	73	81
10	62	69	78

(a) Why do you think the blocking variable "time elapsed since college graduation" was employed?

(b) Obtain the residuals for the randomized block model and plot them against the fitted values. Also, prepare a normal probability plot of the residuals. What are your findings?

(c) Plot the responses Y_{ij} by blocks. What does this plot suggest about the appropriateness of the no-interaction assumption here?

Assume that the randomized block model is appropriate.

(a) Test whether or not the mean proficiency is the same for the three training methods. Use a level of significance $\alpha = 0.05$. State the alternatives, decision rule, and conclusion. What is the p-value of the test?

(b) Make all pairwise comparisons between the training method means; use the Tukey procedure with a 90% confidence level coefficient. State your findings.

(c) Test whether or not blocking effects are present; use $\alpha = 0.05$. State the alternatives, decision rule, and conclusion. What is the p-value of the test?