Major Exam

MAL2040: Design and Analysis of Experiments Date: 2026-04-27

Duration: 180 minutes Total: 40 points

Name: ____________________ **Roll No.:** ____________________

Instructions:

- Hand calculators are allowed.
- State all assumptions used for full credit.
- Unnecessary explanations can result in loss of credit.

Problem 1. **[2 points]**

State whether each of the following statements is True or False. No justification is required.

(i) If the block effect is significant in an RCBD, then the treatment F-test is more likely to detect a true treatment effect compared to in a completely randomized design.

(ii) In a $p \times p$ Latin square design, removing one row still allows the treatment effects to be estimated independently of the row and column effects.

(iii) In the intrablock analysis of a BIBD, the adjusted treatment totals Q_i satisfy $\sum_{i=1}^{a} Q_i = 0$.

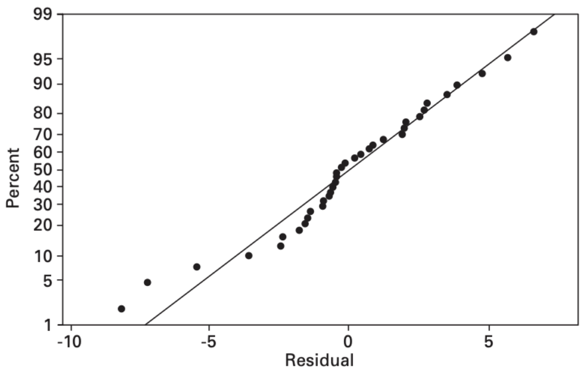
(iv) In response surface methodology, the method of steepest ascent is used to explore the region near a suspected optimum where a second-order model is to be fitted.

Problem 2. A textile manufacturer studies the effect of four dye formulations (1, 2, 3, 4) on the colourfastness of a fabric. Since fabric bolts may differ in quality, a randomized complete block design is used with five bolts as blocks. After fitting the RCBD model, the following diagnostic plots are obtained.

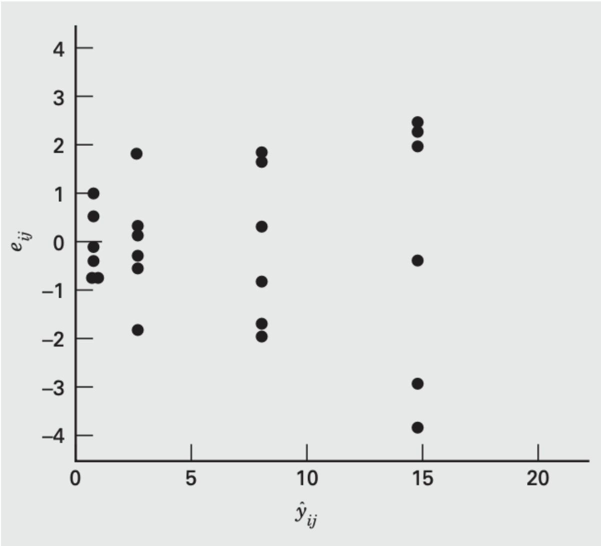
(a) Write down the statistical model for this experiment, clearly defining every term and parameter, and state all assumptions. **[2 points]**

(b) State the null and alternative hypotheses for the treatment F-test, and write down the F-statistic and its degrees of freedom. **[1 point]**

(c) The figure below is a normal probability plot of the residuals. What conclusions do you draw? **[1 point]**



(d) The figure below shows residuals plotted against fitted values. What conclusions do you draw? **[1 point]**



Problem 3. An experiment is conducted to compare the yields (in %) of four chemical formulations (A, B, C, D). Since batch of raw material and operator may introduce variability, a 4×4 Latin square design is used with batches as rows and operators as columns. The observed yields, along with row and column totals, are:

	Op. 1	**Op. 2**	**Op. 3**	**Op. 4**	**Row Total**
Batch 1	$A = 21$	$B = 17$	$C = 26$	$D = 19$	83
Batch 2	$B = 14$	$C = 21$	$D = 17$	$A = 25$	77
Batch 3	$C = 19$	$D = 16$	$A = 21$	$B = 15$	71
Batch 4	$D = 18$	$A = 22$	$B = 14$	$C = 21$	75
Col. Total	72	76	78	80	306

You are given that $\sum_{\text{all obs}} y_{ijk}^2 = 6046$.

(a) Compute the correction factor $\text{CF} = y_{\cdots}^2/N$ and SS_{Total}. **[1 point]**

(b) Compute $\text{SS}_{\text{Batches}}$ and $\text{SS}_{\text{Operators}}$. **[1 point]**

(c) Compute the treatment totals for A, B, C, D, and hence compute $\text{SS}_{\text{Treatments}}$. **[1 point]**

(d) Compute SS_{Error}. **[1 point]**

(e) Construct the ANOVA table. **[1 point]**

(f) Test whether there is a significant difference in yield among the four formulations at $\alpha = 0.05$. The critical value is $F_{0.05,3,6} = 4.76$. **[1 point]**

Problem 4.

(a) A BIBD is proposed with $a = 6$ treatments, block size $k = 4$, and each treatment replicated $r = 5$ times.

Compute λ and b. Find a BIBD with these parameters if it exists, or show that such a BIBD cannot exist. **[2 points]**

(b) Consider a BIBD with a treatments and b blocks of size k, in which each treatment appears in r blocks and each pair of treatments appears together in λ blocks. The model is:

$$y_{ij} = \mu + \tau_i + \beta_j + \varepsilon_{ij}, \qquad \varepsilon_{ij} \sim \mathrm{NID}(0, \sigma^2).$$

The adjusted treatment total for treatment i is defined as

$$Q_i = y_{i.} - \frac{1}{k}\sum_{j=1}^{b} n_{ij}\, y_{.j},$$

where $n_{ij} = 1$ if treatment i appears in block j, and 0 otherwise.

(i) Write down the normal equations by minimising the error sum of squares:

$$L = \sum_{i=1}^{a}\sum_{j=1}^{b} n_{ij}(y_{ij} - \mu - \tau_i - \beta_j)^2.$$

[2 points]

(ii) Apply the usual constraints and solve the resulting system of linear equations to find a formula for the intrablock estimator $\hat{\tau}_i$ of the treatment effect τ_i in terms of the adjusted treatment totals Q_i. **[2 points]**

(c) In the interblock analysis of the BIBD, the block effects β_j are assumed to behave like random noise. The model is:

$$y_{.j} = k\mu + \sum_{i=1}^{a} n_{ij}\tau_i + \text{error}, \qquad j = 1, \ldots, b,$$

where the error term includes the block effects.

(i) Write down the normal equations by minimising the error sum of squares:

$$L = \sum_{j=1}^{b}\left(y_{.j} - k\mu - \sum_{i=1}^{a} n_{ij}\tau_i\right)^2.$$

[2 points]

(ii) Apply the usual constraint and solve the resulting system of linear equations to find a formula for the interblock estimator $\tilde{\mu}$ of the grand mean μ. **[2 points]**

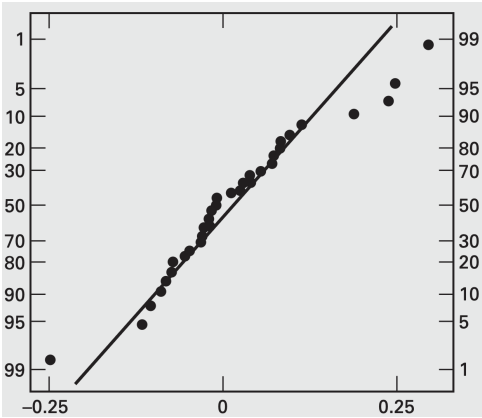
Problem 5.

(a) In a 2^k factorial design, state the total number of (i) treatment combinations, (ii) main effects, and (iii) two-factor interaction effects. Give your answers in terms of k. **[1 point]**

(b) State the *sparsity of effects principle.* Explain how it is used to identify estimate the error in an unreplicated 2^k design. **[1 point]**

(c) A 2^4 factorial experiment is run with factors A, B, C, D. After analysis, it is found that B and all interactions involving B are negligible. What does the original 2^4 design reduce to, and how many effective replicates does the experiment have? **[1 point]**

(d) The figure below shows a normal probability plot of the estimated effects from an unreplicated 2^4 experiment. How many effects appear to be active? Justify your answer. **[2 points]**



Problem 6. A chemical engineer studies three process factors: reaction temperature (A), pressure (B), and catalyst concentration (C), each at two levels. A 2^3 factorial experiment is run with $n = 2$ replicates. The observed yields (in %) are:

Treatment	Replicate 1	Replicate 2	Treatment Total
(1)	28	25	53
a	36	32	68
b	18	20	38
ab	30	29	59
c	31	28	59
ac	40	37	77
bc	20	22	42
abc	34	36	70

You are given: $\sum_{\text{all obs}} y^2 = 14244$, grand total $y_{...} = 466$, $\text{CF} = y_{...}^2/N = 13572.25$.

(a) Compute SS_{Total}. **[1 point]**

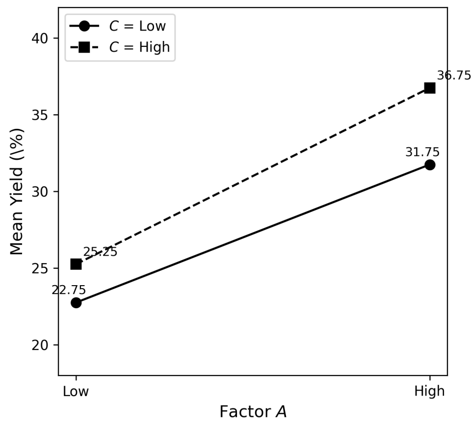
(b) Write down the sign table for a 2^3 factorial design in standard order. **[1 point]**

(c) Compute the contrast and sum of squares for each of the seven effects (A, B, AB, C, AC, BC, ABC). **[2 points]**

(d) Compute SS_{Error} and construct the complete ANOVA table. **[1 point]**

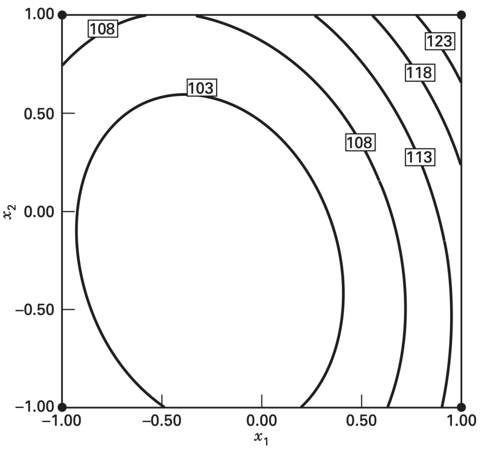
(e) Using $F_{0.05,1,8} = 5.32$, state which effects are significant. What factor-level combination maximises yield? **[1 point]**

(f) The figure below shows an interaction plot for factors A and C (averaged over the levels of B). Interpret this plot. Does it agree with your conclusion from part (d)? **[1 point]**



Problem 7.

(a) The contour plot below shows the fitted response surface $\hat{y}(x_1, x_2)$ from a second-order model. Based on the plot, determine whether the stationary point of this surface is a local maximum, local minimum, or saddle point. Justify your answer by referring to specific features of the plot. **[2 points]**



(b) A first-order response surface model is fitted in two coded variables, giving

$$\hat{y} = 52.3 + 1.8\,x_1 + 0.9\,x_2,$$

where $x_1 = (\xi_1 - 25)/5$ and $x_2 = (\xi_2 - 80)/10$, with ξ_1 = reaction time (minutes) and ξ_2 = reaction temperature (°C).

(i) State the direction of steepest ascent in the coded variable space.

(ii) Taking a step size of $\Delta\xi_1 = 5$ minutes, compute $\Delta\xi_2$.

(iii) Starting from $(\xi_1, \xi_2) = (25, 80)$, tabulate the first three points along the path of steepest ascent. **[2 points]**

(c) After moving along the steepest ascent path, a second-order model is fitted. The stationary point is found at $\mathbf{x}_0 = (1.2, -0.8)^T$, and the eigenvalues of the matrix of second-order coefficients are $\lambda_1 = -3.2$ and $\lambda_2 = -1.5$. What type of stationary point is this? Compute the optimal operating conditions in the natural variables ξ_1 and ξ_2. **[1 point]**

End of Exam