

2026

STATISTICS — MDC

Paper : CC-4

(Design of Experiments - I and Sample Survey - I)

Full Marks : 75

The figures in the margin indicate full marks.

*Candidates are required to give their answers in their own words
as far as practicable.*

1. Answer **any five** questions :

2×5

- (a) Why is the Latin Square Design considered restrictive in its application? Explain briefly.
- (b) State the objective of a uniformity trial experiment. Why is it important in the comparison of experimental designs?
- (c) Why is sampling preferred to complete enumeration in large-scale surveys? Mention any two valid reasons.
- (d) Distinguish between sampling error and non-sampling error, clearly stating the nature of each.
- (e) "Randomization alone does not ensure precision in an experiment." — Justify this statement with reference to experimental error.
- (f) Explain why a Randomized Block Design (RBD) generally yields a smaller error variance than a Completely Randomized Design (CRD) when block effects are present.
- (g) Why does stratified random sampling usually provide more precise estimates than Simple Random Sampling (SRS)? State one condition under which this improvement occurs.
- (h) What is meant by two-stage sampling? Mention one practical situation where it is referred over single-stage sampling.

2. Answer **any four** of the following questions :

5×4

- (a) Discuss the three basic principles of experimental design proposed by Fisher and explain how they contribute to improving the validity and precision of experimental results.
- (b) In an agricultural field experiment laid out in a Randomized Block Design (RBD), observation in one experimental plot is lost due to unforeseen natural causes such as flooding. Discuss the statistical issues arising from the missing observation and explain how such a situation is handled in the analysis of variance.
- (c) A large-scale household survey aims to estimate the average monthly expenditure, but several selected households fail to respond. Critically discuss the statistical problems caused by non-response and explain how they may affect the reliability of survey estimates.

Please Turn Over

(6709)

- (d) Within the fixed effects ANOVA framework, explain why the total sum of squares can be uniquely decomposed into treatment and error components. Discuss the mathematical role of orthogonality in ensuring this decomposition.
- (e) In survey sampling, unbiasedness is often regarded as a desirable property of an estimator. Critically examine situations in which a biased estimator may be preferred over an unbiased one.
- (f) The finite population correction (FPC) appears in the variance of the sample mean under Simple Random Sampling Without Replacement (SRSWOR). Explain its mathematical origin and interpret its limiting behaviour.

Answer *any three* questions from the following.

3. (a) In an analysis of variance, an observed F-statistic is sometimes found to be less than one. Explain how such a result should be interpreted from a statistical and practical point of view. Discuss what this indicates about the relative magnitudes of treatment and error mean squares and comment on the implications for hypothesis testing and experimental conclusions.
(b) Distinguish between fixed-effects, random-effects and mixed-effects ANOVA models, highlighting their conceptual foundations and mathematical formulations in the context of a two-way ANOVA with one observation per cell. Briefly explain how the expectations of mean squares and the estimation of variance components differ among these models. Illustrate your answer with a suitable practical example, showing how the nature of the factors determines the appropriate modelling framework. 8+7
4. (a) Compare the relative efficiencies of the Completely Randomized Design (CRD), Randomized Block Design (RBD) and Latin Square Design (LSD). Explain how the presence of one or two sources of extraneous variation influences the choice of design and discuss the conditions under which LSD becomes more efficient than CRD and RBD.
(b) Present a complete statistical analysis of a RBD in the presence of missing observations. Discuss the estimation of missing values, the necessary adjustments in sums of squares and degrees of freedom and the implications of missing observations for the validity of ANOVA-based inference. 7+8
5. (a) Distinguish between stratified random sampling and two-stage sampling, highlighting their objectives, sampling structure and typical applications.
(b) In two-stage sampling with primary sampling units (PSUs) of equal size and equal selection probability at each stage, derive an unbiased estimator of the population mean.
(c) Obtain the variance of the estimator derived in part (b), clearly identifying the between-PSU and within-PSU components.
(d) Explain how the variance is estimated from the sample data. 4+3+4+4
6. (a) "In stratified random sampling, the variance of the unbiased estimator of the population mean depends not only on the total sample size but also on how the sample is distributed among the strata". — Justify with necessary derivation.
(b) In practical survey situations the stratum variances are usually unknown before data collection. Discuss how this uncertainty affects the implementation of optimum allocation and explain the alternative approaches that may be used by a statistician in such cases. 8+7

7. (a) Discuss how sampling errors and non-sampling errors jointly influence the accuracy and reliability of survey results, with a suitable example.
- (b) In a two-way ANOVA with one observation per cell, explain why the interaction effect cannot be separately estimated and is therefore omitted from the model formulation. Discuss the case of a two-way ANOVA with three observations per cell and present the structure of the model along with the basic outline of its analysis.

3+12