MACHAKOS UNIVERSITY

University Examinations 2021/2022 Academic Year

SCHOOL OF PURE AND APPLIED SCIENCES

DEPARTMENT OF MATHEMATICS AND STATISTICS

FOURTH YEAR FIRST SEMESTER EXAMINATION FOR

BACHELOR OF SCIENCE IN MATHEMATICS AND COMPUTER SCIENCE
BACHELOR OF SCIENCE IN STATISTICS AND PROGRAMMING
BACHELOR OF SCIENCE IN ACTUARIAL SCIENCE
BACHELOR OF SCIENCE IN MATHEMATICS
BACHELOR OF EDUCATION

SMA 464: DESIGN AND ANALYSIS OF EXPERIMENTS

DATE: 30/8/2022 **TIME: 2.00-4.00 PM**

INSTRUCTIONS:
Answer Question One and Any Other Two Questions

You must have Statistical Tables and a Scientific Calculator.

QUESTION ONE (30 MARKS)

(a) (i) Explain the term *treatment* as used in the design of experiments, illustrating with *two* examples from real life situations. (3 *marks*)

(ii) Explain the term *double blind* as used in the design of experiments, illustrating with an example from a real life situation. (3 *marks*)

(b) Explain the *meaning* and *importance* of each of the following techniques of error control as used in the design of experiments, illustrating with a real life example:

(i) randomisation;

(ii) blocking. (6 *marks*)

(c) Differentiate between an *experimental group* and a *control group* as used in the design of experiments, illustrating with an example from a real life situation. (4 *marks*)

(d) Given a two-way ANOVA with the model given by $y_{ij} = \mu + t_i + b_j + e_{ij}$, and taking the sum of squares due to treatment (SS_t), show that:

$$SS_t = \sum_{i=1}^{k}\sum_{j=1}^{b} (\bar{y}_{i.} - \bar{y}_{..})^2 = \frac{1}{b}\sum_{i=1}^{k} T_i^{\,2} - \frac{G^2}{bk}$$ (6 *marks*)

(e) An educational researcher carried out a study on three methods of teaching: lecture, handouts and e-learning. A random sample of 30 students was selected, with 10 students randomly assigned to each teaching method for a period of one academic year. The students were taught using the same curriculum, and finally sat for the same examination. The mean mark scored by each student was recorded. After statistical analysis, the ANOVA table below was derived.

Source of Variation	Sum of Squares (SS)	Degrees of Freedom	Mean Sum of Squares (MSS)	Variance Ratio - F	P–value
Method	2429.600				0.0208880123
Error	7321.600				
Total					

(i) Complete the ANOVA table above, and hence, using the *F* test statistic, evaluate whether there is a significant difference in academic performance between the three methods of teaching. Test at the 5% level and 1% level of significance. (4 *marks*)

(ii) Given the treatment means: $\bar{y}_{1.} = 46.8$ $\bar{y}_{2.} = 56.6$ $\bar{y}_{3.} = 68.8$
Compute the critical difference between the methods of teaching, and hence, identify which pair of methods differ significantly. (4 *marks*)

QUESTION TWO (20 MARKS)

(a) Given a two-way ANOVA with its model $y_{ij} = \mu + t_i + b_j + e_{ij}$ and its total sum of squares given by $(SS_T) = \sum_{i=1}^{k} \sum_{j=1}^{b} (y_{ij} - \bar{y}_{..})^2$, split or decompose this total sum of squares into its various components sums of squares, and hence show that:

$SS_T = SS_t + SS_b + SS_e$ (5 *marks*)

(b) Given a one-way ANOVA with the model given by $y_{ij} = \mu + t_i + e_{ij}$ and assuming the *random effects* model, show that:

(i) the mean sum of squares due to treatment (MSS_t) is a biased estimator of the error variance σ_e^2 ; (8 *marks*)

(ii) the mean sum of squares due to error (MSS_e) is an unbiased estimator of the error variance σ_e^2 . (7 *marks*)

QUESTION THREE (20 MARKS)

A researcher in insects carried out a study on three different insecticides on cockroaches in terms of the duration it takes for a cockroach to die after application. A random sample of 24 cockroaches was taken, and 8 cockroaches randomly assigned to each insecticide. The duration in seconds each cockroach took before death was recorded as shown in the table below:

	Observations (duration in seconds for cockroach to die)							
Insecticide A	84	70	76	80	56	74	62	68
Insecticide B	75	72	54	62	65	68	56	74
Insecticide C	50	56	45	58	55	42	50	48

Using a complete randomised design (CRD), and corresponding one-way ANOVA model, and the 5% level of significance where necessary, do the following:

a) Taking the model $y_{ij} = \mu + t_i + e_{ij}$, outline *three* assumptions in the model. (3 *marks*)

b) Evaluate whether there is a significant difference between the insecticides (treatments) in the duration it takes for a cockroach to die. (13 *marks*)

c) Compute the *critical difference* between the insecticides in terms of the duration before death, and hence, identify which pair of insecticides has a statistically significant difference. (4 *marks*)

QUESTION FOUR (20 MARKS)

(a) Differentiate between the terms *experimental studies* and a *observational studies* as used in the analysis of variance. (4 *marks*)

(b) A medical researcher carried out a study on the effect of pain killer drugs on patients to assess the duration it takes for a patient to feel well. Four drugs were tested with 5 different dosages in millilitres. A random sample of 20 patients was taken, and 5 patients randomly assigned to each drug, with 4 patients randomly assigned to each dosage. The duration in minutes a patient took for the pain to subside was recorded as shown in the table below:

Dosage	10 ml	15 ml	20 ml	25 ml	30 ml
Drug A	12	14	16	15	18
Drug B	28	30	32	30	36
Drug C	30	28	34	32	38

Drug D	34	36	28	36	42

Using a randomised block design (RBD) with drugs as treatments and dosage as blocks, and corresponding two-way ANOVA model, evaluate whether there is a significant difference in the duration it took for pain to subside for each of the following, at the 5% level of significance:

(i) between the drug varieties;

(ii) between the dosages. (16 *marks*)

QUESTION FIVE (20 MARKS)

a) Given a one-way ANOVA with its model $y_{ij} = \mu + t_i + e_{ij}$, decompose the total sum of squares given by $SS_T = \sum_{i=1}^{k} \sum_{j=1}^{n} \left(y_{ij} - \bar{y}_{..} \right)^2$, into its various component sums of squares, and hence show that: $SS_T = SS_t + SS_e$ (3 *marks*)

b) An agricultural scientist carried out a study on the effects of three factors on maize yield: varieties of maize, fertilizer type and spacing gap. A random sample of 16 plots was taken and 4 maize varieties randomly planted in the plots in such a way that each row and each column has only one of each maize variety forming a 4×4 Latin square design. Four types of fertilizers were applied in each row at random, and four different spacing gaps were applied in each column at random. The yield in terms of the number of sacks per plot was recorded as shown in the table below:

	Column 1	Column 2	Column 3	Column 4
Row 1	$T_3 = 58$	$T_2 = 44$	$T_1 = 50$	$T_4 = 36$
Row 2	$T_2 = 22$	$T_4 = 28$	$T_3 = 48$	$T_1 = 20$
Row 3	$T_1 = 36$	$T_3 = 54$	$T_4 = 30$	$T_2 = 24$
Row 4	$T_4 = 30$	$T_1 = 30$	$T_2 = 28$	$T_3 = 60$

Using a Latin square design (LSD), and corresponding ANOVA model given by $y_{ijk} = \mu + r_i + c_j + t_k + e_{ijk}$, and using a 5% level of significance:

Evaluate whether there is a significant difference in the maize yield for each of the following:

(i) between the varieties of maize;

(ii) between the types of fertilizers;

(iii) between the spacing gaps. (17 *marks*)