MACHAKOS UNIVERSITY

University Examinations for 2020/2021 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 MATHEMATICS

BACHELOR OF EDUCATION

BACHELOR OF ARTS

SMA 464: DESIGN AND ANALYSIS OF EXPERIMENTS

DATE: 19/08/2021 **TIME: 8:30 – 10:30 AM**

Instructions to the Candidate:

1. *Answer **Question 1** and any other **two** questions.*
2. *You must have Statistical Tables and a Scientific Calculator.*

QUESTION ONE (COMPULSORY) (30 MARKS)

(a) (i) Explain the term *treatment* as used in the design of experiments, illustrating with *two* examples from real life situations. (2 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 each of the following techniques of error control as used in the design of experiments, illustrating with a real life example where necessary:

(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 one-way ANOVA with the model given by $y_{ij} = \mu + t_i + e_{ij}$, and taking the following sums of squares, prove the following:

(i) Sum of squares due to treatment (SS_t):

$$SS_t = \sum_{i=1}^{k} \sum_{j=1}^{n} (\bar{y}_{i.} - \bar{y}_{..})^2 = \sum_{i=1}^{k} \frac{T_i^2}{n_i} - \frac{G^2}{N} \qquad \text{(3 marks)}$$

(ii) Sum of squares due to error (SS_e):

$$SS_e = \sum_{i=1}^{k} \sum_{j=1}^{n} (y_{ij} - \bar{y}_{i.})^2 = \sum_{i=1}^{k} \sum_{j=1}^{n} y_{ij}^2 - \sum_{i=1}^{k} \frac{T_i^2}{n_i} \qquad \text{(3 marks)}$$

(e) A poultry nutrition researcher carried out a study on the production of 4 feed varieties A, B, C and D in which a random sample of 32 chicks were used, with each feed given to a random sample of 8 chicks for a period of six months. The increase in weight was then recorded in grams. After statistical analysis, he derived the following ANOVA table.

Source of Variation	Sum of Squares (SS)	Degrees of Freedom	Mean Sum of Squares (MSS)	Variance Ratio - F	P–value
Feed	682.00				0.004216605
Error	1,156.00				
Total					

Given the treatment means: $\bar{y}_{A.} = 25.0$, $\bar{y}_{B.} = 28.0$, $\bar{y}_{C.} = 30.50$, $\bar{y}_{D.} = 37.50$

(i) Complete the ANOVA table above, and hence using the F test statistic, evaluate whether there is a significant difference in weight increase between the poultry feeds. (5 marks)

(iii) Compute the critical difference between the feeds, and hence identify which pair of animal feeds differ significantly. (4 marks)

QUESTION TWO (20 MARKS)

(a) Given a two-way ANOVA with the 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, and hence show that:

$$SS_T = SS_t + SS_b + SS_e \qquad (5\ marks)$$

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

An animal breeding researcher carried out a study on the effect of breeds of dairy cows on milk production. Three breeds of dairy cows were tested: local breed, cross breed and exotic breed. A random sample of 8 dairy cows was taken from each breed. The 24 cows were then kept under the same conditions and given the same feedstuff for a period of one year. The mean monthly production of milk for each cow was recorded in litres as shown in the table below:

Cow Breed	Observations (milk yield in litres)							
Local	65	75	64	68	80	64	78	72
Cross	62	76	68	72	78	68	80	74
Exotic	82	80	78	86	90	84	92	88

Using a complete randomised design (CRD) assuming the ANOVA model $y_{ij} = \mu + t_i + e_{ij}$, do the following:

(i) State *three* assumptions in the model. (3 marks)

(ii) Evaluate whether there is a significant difference between the breeds (treatments) in the milk production among the dairy cows. Use the 5% level of significance. (12 marks)

(iii) Compute the *critical difference* in milk production between the breeds of cows, and hence, identify which pair of breeds has a statistically significant difference. (4 marks)

(iv) Which breed of dairy cows would you recommend, justifying your choice? (1 mark)

QUESTION FOUR (20 MARKS)

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. Three types of drugs were tested with 5 different drug dosages in millilitres. A random sample of 15 patients was taken, and 5 patients randomly assigned to each drug, and 3 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	48	50	42	56	50
Drug B	52	48	30	40	46
Drug C	20	25	18	24	36

(a) By using a randomised block design (RBD) with drugs as treatments and dosage as blocks, and corresponding two-way ANOVA model, evaluate each of the following at the 5% level of significance:

(i) Evaluate whether there is a significant difference between the drug varieties in their effectiveness as pain killers. (9 marks)

(ii) Evaluate whether there is a significant difference between the dosages of the drugs in their effectiveness as pain killers. (9 marks)

(b) Outline *four* assumptions in the ANOVA model used above. (2 marks)

QUESTION FIVE (20 MARKS)

(a) Differentiate between the terms *variation within* and a *variation between* as used in the analysis of variance. (2 marks)

(a) Given a one-way ANOVA with its model $y_{ij} = \mu + t_i + e_{ij}$, split or decompose the total sum of squares given by $SS_T = \sum_{i=1}^{k} \sum_{j=1}^{n} (y_{ij} - \bar{y}_{..})^2$, 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	$T_3 = 15$	$T_2 = 10$	$T_1 = 15$	$T_4 = 20$
Row 2	$T_2 = 10$	$T_4 = 22$	$T_3 = 20$	$T_1 = 16$
Row 3	$T_1 = 12$	$T_3 = 24$	$T_4 = 24$	$T_2 = 18$
Row 4	$T_4 = 25$	$T_1 = 16$	$T_2 = 18$	$T_3 = 30$

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:

(i) between the maize varieties; (5 marks)

(ii) between the fertilizers; (5 marks)

(iii) between the spacing gaps. (5 marks)