MACHAKOS UNIVERSITY

University Examinations 2018/2019

SCHOOL OF PURE AND APPLIED SCIENCES

DEPARTMENT OF MATHEMATICS, STATISTICS AND ACTUARIAL SCIENCE

FOURTH YEAR SPECIAL/SUPPLEMENTARY EXAMINATION FOR

BACHELOR OF SCIENCE (PROGRAMMIN & STATISTICS)

BACHELOR OF SCIENCE IN (MATHEMATICS)

BACHELOR OF EDUCATION

SMA 464 : DESIGN AND ANALYSIS OF EXPERIMENTS

DATE: 24/9/2019 **TIME: 2:00 – 4:00 PM**

INSTRUCTION: *Answer question one and any other two Questions*

Instructions to the Candidate:

1. *Answer* ***Question* 1** *and any other* ***two*** *questions.*
2. *You need a Scientific Calculator and Statistical Tables for this paper:*

1. (a) 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. (6 *marks*)

   (b) Given a one-way classification in ANOVA with the model given by $y_{ij} = \mu + t_i + e_{ij}$, and taking the various sums of squares, prove the following: (6 *marks*)

   (i) Total sum of squares: $SS_T = \sum_{i=1}^{k} \sum_{j=1}^{n} \left(y_{ij} - \bar{y}_{..} \right)^2 = \sum_{i=1}^{k} \sum_{j=1}^{n_i} y_{ij}^2 - N\,\bar{y}_{..}^2$

   (ii) Treatment sum of squares: $SS_t = \sum_{i=1}^{k} \sum_{j=1}^{n} (y_{i.} - \bar{y}_{..})^2 = \sum_{i=1}^{k} n_i\,\bar{y}_{i.}^{\,2} - N\,\bar{y}_{..}^2$

(c) A study was carried out to assess if there is a difference in academic performance in Design and Analysis of Experiments between three B.Sc. courses – K63, S09, and S10. A random sample of 10 students was taken from each course. The marks for the 30 students were then recorded as shown in the table below:

Course	Observations (marks scored)									
K63	56	28	64	36	72	36	54	42	32	48
S09	37	56	74	73	45	32	65	68	42	74
S10	96	79	54	85	76	52	48	74	82	42

Using a complete randomised design (CRD) with the linear model $y_{ij} = \mu + t_i + e_{ij}$, evaluate whether there is a significant difference between the three courses in academic performance of the students in the unit. Test at both the 5% level and 1% level of significance. (8 *marks*)

(d) A wheat breeding researcher carried out a study on the yields of 4 wheat varieties A, B, C and D. A random sample of 20 plots was used with each variety planted randomly in 5 plots. The crop yields were then observed and recorded. After 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
Variety	26,234.95				0.00021802
Error	11,558.80				
Total					

Given the treatment means: $\bar{y}_{A.}$ = 43.80 , $\bar{y}_{B.}$ = 71.00 , $\bar{y}_{C.}$ = 81.40 , $\bar{y}_{D.}$ = 142.80

(i) Complete the ANOVA table above, and hence evaluate whether there is a significant difference in yield between the wheat varieties. (4 *marks*)

(iii) Compute the critical difference between the wheat varieties, and hence identify which pair of wheat varieties differ significantly. (6 *marks*)

2. (a) Differentiate between the terms *assignable causes* and *chance causes* as used in the design and analysis of experiments, illustrating with a suitable example in a real life situation. (4 *marks*)

(b) Given a one way ANOVA with the model given by $y_{ij} = \mu + t_i + e_{ij}$ show that the sum of squares due to error S_e^2 is an unbiased estimator of the error variance σ_e^2 under each of the following:

(i) fixed effects model;

(ii) random effects model. (10 *marks*)

(c) Given a two-way ANOVA with its model $y_{ij} = \mu + t_i + b_j + e_{ij}$ and its corresponding total sum of squares given by $SS_T = \sum_{i=1}^{k} \sum_{j=1}^{n} \left(y_{ij} - \bar{y}_{..} \right)^2$, split or decompose the total sum of squares into its various components, and hence show that $SS_T = SS_t + SS_b + SS_e$ (6 *marks*)

3. An animal researcher carried out a study on the effect of exotic 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 of the 24 cows was recorded in litres as shown in the table below:

Cow Breed	Observations (milk yield in litres)							
Local	66	75	65	66	80	64	78	70
Cross	77	76	74	73	78	68	80	74
Exotic	96	80	90	86	82	74	82	88

Using a complete randomised design (CRD) assuming the linear additive model $y_{ij} = \mu + t_i + e_{ij}$, and using a 5% level of significance:

(i) Evaluate whether there is a significant difference between the breeds (treatments) in the milk production among the dairy cows. (12 *marks*)

(ii) Compute the critical difference between the breeds of cows and identify which pair of breeds has a statistically significant difference. (6 *marks*)

(iii) Which breed of dairy cows would you recommend justifying your choice? (2 *marks*)

4. A medical researcher carried out a study on the effect of anti-malarial drugs on patients to assess the duration it takes for a patient to show signs of improvement. Four types of anti-malarial drugs were tested. A random sample of 20 malaria patients of varying age groups was taken and 4 patients randomly selected from 5 age groups. For each age group, the 4 patients were each given a different drug. The duration in hours a patient took to show signs of improvement was recorded as shown in the table below:

Age in years	01 – 09	10 – 19	20 – 39	40 – 59	60 – 89
Drug A	20	18	23	26	38
Drug B	34	30	34	32	46
Drug C	42	38	40	34	42
Drug D	48	42	44	36	44

By using a randomised block design (RBD) taking the age groups as blocks, and using a two-way ANOVA having the linear additive model given by $y_{ij} = \mu + t_i + b_j + e_{ij}$, and using a 5% level of significance:

(i) Evaluate whether there is a significant difference between the drugs in their effectiveness in treating malaria. (9 *marks*)

(ii) Evaluate whether there is a significant difference between the age groups in the drug effectiveness in treating malaria. (9 *marks*)

(iii) Which drug would you recommend justifying your choice? (2 *marks*)

5. An automotive engineer carried out a study on the efficiency of break fluids on the breaking distance of a vehicle. In the study, 4 types of break fluids were considered. A sample of 4 types of vehicle engines was taken and 4 fluid transmission systems taken and tested on a random sample of 16 vehicles. Taking the break fluids to represent the treatments, the rows to represent the type of engine, and columns to represent the fluid transmission system, the design forms a 4 X 4 Latin square design. The yield in terms of the distance in metres a vehicle moving at a speed of 50 km per hour takes to stop immediately after the break pedal is pressed 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 = 10	T_4 = 25
Row 2	T_2 = 12	T_4 = 22	T_3 = 20	T_1 = 16
Row 3	T_1 = 10	T_3 = 24	T_4 = 24	T_2 = 18
Row 4	T_4 = 25	T_1 = 12	T_2 = 10	T_3 = 25

(a) Using the Latin square design (LSD) and the ANOVA model $y_{ijk} = \mu + r_i + c_j + t_k + e_{ijk}$. Evaluate whether there is a significant difference in the yield at 5% level of significance:

(i) between the types of break fluids;

(ii) between the types of vehicle engines;

(iii) between the fluid transmission systems. (15 *marks*)

(b) What conclusions would you draw from the findings in (a)? (5 *marks*)