MACHAKOS UNIVERSITY

University Examinations 2016/2017

SCHOOL OF PURE AND APPLIED SCIENCES

DEPARTMENT OF MATHEMATICS AND STATISTICS

FOURTH YEAR FIRST SEMESTER EXAMINATION FOR

BACHELOR OF SCIENCE (MATHEMATICS AND COMPUTER SCIE,NCE)

BACHELOR OF SCIENCE (STATISTICS AND PROGRAMMING)

BACHELOR OF SCIENCE (MATHEMATICS)

BACHELOR OF EDUCATION

SMA 464: DESIGN AND ANALYSIS OF EXPERIMENTS

DATE: 31/7/2017 **TIME: 2:00 – 4:00 PM**

<u>*Instructions to the Candidate*</u>**:**

1. ***Answer Question* 1 *and any other two questions.***
2. ***Out of the three questions answered, each question must start on a new page.***
3. ***You need a Scientific Calculator and Statistical Tables for this paper.***

1. (a) (i) Explain the term *ANOVA* in the context of analysis of data from an experiment
. (2 marks)

(ii) Explain the importance of *error control* in an experiment. (2 marks)

(b) Explain *any two* of the following techniques of error control in the design of an experiment:

(i) Randomisation;
(ii) Replication;
(iii) Blocking. (4 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. (6 marks)

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

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

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

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

(e) A wheat breeding researcher carried out a study on the yields of 4 wheat varieties A, B, C and D in which a random sample of 20 plots were 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	16,646.95	3			0.00053974
Error	8,712.00	16			
Total					

Given the treatment means: $\bar{y}_{1.} = 43.8$, $\bar{y}_{2.} = 71.0$, $\bar{y}_{3.} = 81.4$, $\bar{y}_{4.} = 124.0$

(i) Complete the ANOVA table above, and hence evaluate whether there is a significant difference 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) With the aid of suitable examples in each case, differentiate between each of the following terms as used in the design and analysis of experiments:

(i) Assignable causes and chance causes;

(ii) Fixed effects model and random effects model. (10 marks)

(b) Given a one-way ANOVA with the model given by $y = \mu + t_i + e_{ij}$, 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)

3. An animal nutrition researcher carried out a study on the effects of feeds for dairy cows on milk production. Three types of feeds were considered. A random sample of 18 dairy cows of the same breed and age was taken and 3 cows randomly allocated to 6 farmers in different regions. For each farmer, the 3 cows were each given a different feed for a period of one year. The mean monthly production of milk in litres was recorded as shown in the table below:

Farmer	A	B	C	D	E	F
Feed A	63	72	68	65	80	60
Feed B	78	76	74	72	78	66
Feed C	96	80	90	86	84	68

(i) Using a one-way ANOVA with the model given by $y = \mu + t_i + e_{ij}$, test whether there is a significant difference in the milk production between the feeds among the dairy cows. Use the 5% level of significance. (12 marks)

(ii) Determine the critical difference in milk production between the feeds. Hence identify which of the feeds differ significantly in their milk production. (6 marks)

(iii) Which feed would you recommend based on the findings in (ii) above? (2 marks)

4. A medical researcher carried out a study on the effect of anti-malarial drugs on patients to assess how long it takes for a patient to show signs of improvement. Four types of anti-malarial drugs were tested. A random sample of 15 malaria patients of the same age group and state was taken and 3 patients randomly selected from 5 regions in Kenya. For each region, the 3 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:

Region	Block 1	Block 2	Block 3	Block 4	Block 5
Drug A	20	18	25	26	18
Drug B	24	20	26	28	24
Drug C	34	20	30	36	33

By taking the regions as blocks and using a two-way ANOVA with the linear additive model given by $y = \mu + t_i + b_j + e_{ij}$, and using a 5% level of significance:

(i) Test whether there is a significant difference in the drug effectiveness in treating malaria. (8 marks)

(ii) Evaluate whether there is a significant difference between regions in the drug effectiveness in treating malaria. (8 marks)

(iii) Which conclusions would you draw based on the findings in (ii) above? (4 marks)

5. (a) State the general ANOVA table for a Latin square design. (3 marks)

(b) An agricultural scientist carried out a study on the effect of varieties of mangoes on yield. Four types of mango varieties were tested. A random sample of 16 plots having similar characteristics was taken and 4 mango varieties randomly planted in rows and columns with each row and each column having only one of each mango variety forming a 4 X 4 Latin square design. The yield in terms of the number of fruits in hundreds per tree was recorded as shown in the table below:

	Column 1	Column 2	Column 3	Column 4
Row 1	$T_3 = 15$	$T_2 = 12$	$T_1 = 10$	$T_4 = 25$
Row 2	$T_2 = 12$	$T_4 = 22$	$T_3 = 14$	$T_1 = 16$
Row 3	$T_1 = 10$	$T_3 = 18$	$T_4 = 20$	$T_2 = 18$
Row 4	$T_4 = 18$	$T_1 = 12$	$T_2 = 12$	$T_3 = 18$

Using a Latin square design (LSD) and corresponding ANOVA having a linear additive model given by $y = \mu + r_i + c_i + b_k + e_{ijk}$, and using a 5% level of significance.

Test whether there is a significant difference in the yield:

(i) between the mango varieties. (6 marks)

(ii) between rows. (4 marks)

(iii) between columns. (4 marks)

Which conclusions would you draw based on the findings in (i) to (iii) above? (3 marks)