

2405/303

STATISTICAL METHODS

Oct./Nov. 2009

Time: 3 hours

THE KENYA NATIONAL EXAMINATIONS COUNCIL

DIPLOMA IN APPLIED STATISTICS

STATISTICAL METHODS

3 hours

INSTRUCTIONS TO CANDIDATES

You should have the following for this examination:

Answer booklet;

Mathematical tables/scientific calculator.

Answer any FIVE of the EIGHT questions in this paper.

All questions carry equal marks.

Maximum marks for each part of a question are as indicated.

This paper consists of 5 printed pages.

Candidates should check the question paper to ascertain that all the pages are printed as indicated and that no questions are missing.

1. (a) Explain each of the following terms:

- (i) correlation;
- (ii) regression.

(4 marks)

(b) The following data were collected by a researcher on the weight of some products versus their volumes.

Table 1

Weight (kg)	Volume (cubic meters)
63	66
65	68
66	65
67	67
67	69
68	70

- (i) Using the data above, determine the regression line equation $y = a + bx$.
- (ii) From the equation in b(i) determine the volume of a product whose weight is 70 kg.
- (iii) Calculate Karl Pearson's coefficient of correlation and interpret your answer.

(6 marks)

2. (a) Explain each of the following types of random sampling:

- (i) systematic sampling;
- (ii) stratified sampling;
- (iii) cluster sampling.

(6 marks)

(b) State the difference between Type 1 error and Type 2 error.

(4 marks)

(c) A company employs a large number of workers to make a certain type of car components. The average number of units produced per hour for all employees is 100 units with standard deviation of 20 units. It is suspected that one worker's performance is below the expected standard. The management tests this particular worker on 36 separate random occasions. It is discovered that his average hourly output is 90 units. Test the workers efficiency at 99% significant level.

(10 marks)

3. (a) Differentiate between:

- (i) mutually exclusive events and independent events;
- (ii) discrete random variable and continuous random variable.

(4 marks)

(b) The personnel department of a company has records which show the following analysis of its engineers:

Table 2

Age under 30	Bachelor's degree	Master's degree	Total
Under 30	90	10	100
30 to 40	20	30	50
Over 40	40	10	50
Total	150	50	200

If one engineer is selected at random from the company, determine the probability that:

- (i) he has a master's degree given that he is over 40;
- (ii) he is under 30 given that he has only a bachelor's degree. (8 marks)
- (c) Table 3 shows data for the national income of a country for three years, in millions of shillings.

Table 3

Year	Agriculture (sh. million)	Manufacture (sh. million)	Others (sh. million)	Total
2007	374	170	136	680
2008	480	240	80	800
2009	414	322	184	920

Represent the above information on a percentage component bar chart.

(8 marks)

4. (a) A study was done to assess the awareness of professionals in accounting, teaching and medicine about new developments in science and technology. A sample of 995 professionals was taken and given a test. Some 410 of the professionals passed the test. Among those who passed the test, 155 were accountants. Among those who failed, 180 were accountants and 195 teachers. The number of teachers in the study was 345. Present the information in tabular form. (6 marks)
- (b) Explain each of the following terms as used in statistics:
- (i) central limit theorem;
- (ii) point estimator;
- (iii) interval estimator. (6 marks)
- (c) Describe the **four** characteristics of a good statistical estimator. (8 marks)

5. (a) A manufacturer sets up the following sampling for accepting or rejecting crates of identical items of raw material received. He takes a random sample of 20 items from a crate. If he finds more than two defective items in the sample, he rejects the crate. It is known that approximately 5% of these types of items received are defective.

Calculate:

- (i) the proportion of crates that will be rejected;
 (ii) the mean and variance of the number of defectives in a sample of 20. (12 marks)

- (b) Explain **four** characteristics of a binomial distribution. (8 marks)

6. During the year 2008, a college ran 70 different courses of which 44 were 'science' and 26 were 'arts'. The following is the frequency distribution of the size of the courses.

Table 4

No. of course	No. of students taking science course	No. of students taking arts course
1 - 6	4	0
7 - 12	15	3
13 - 18	11	10
19 - 24	8	8
25 - 30	5	4
31 - 36	1	1

- (a) Calculate the mean number of courses taken by:

- (i) science students;
 (ii) arts students. (10 marks)

- (b) Determine the coefficient of variation of the number of courses taken by:

- (i) science students;
 (ii) arts students. (8 marks)

- (c) Comment on results in 6(b). (2 marks)

7. (a) Explain **four** circumstances under which a census may be preferred to a sample. (8 marks)

- (b) The following data shows the life span of a number of bulbs in an industry.

Table 5

Life span (hours)	No. of bulbs
300 - 400	14
400 - 500	46
500 - 600	58
600 - 700	76
700 - 800	68
800 - 900	62
900 - 1000	48
1000 - 1100	22
1100 - 1200	6

Calculate:

- the average life span in hours;
- median life span in hours;
- Karl Pearson's coefficient of skewness. (12 marks)

8. (a) The following are the sales figures of a retail store.

Table 6

Year	Sales ('000)
1997	200
1998	240
1999	-
2000	350
2001	400

Interpolate the sales for 1999 using parabolic method. (8 marks)

- (b) The resistance, $R(\Omega)$ of a length of a wire depends on its temperature, $T(^{\circ}\text{C})$. The relationship is of the form $R = aT + b$. The following values were found during an experiment on the resistance of a wire.

Table 7

Temperature $T(^{\circ}\text{C})$	20	40	60	80	100
Resistance $R(\Omega)$	10.8	11.2	11.9	13.2	14.0

- Plot the points on a graph.
- Draw a line of best fit and find the relationship. (12 marks)