



AMREF INTERNATIONAL UNIVERSITY
SCHOOL OF PUBLIC HEALTH
DEPARTMENT OF COMMUNITY HEALTH
MASTER OF PUBLIC HEALTH
END OF SEMESTER EXAMINATION APRIL 2025

MPH 712: BIostatistics

DATE: April 2025

TIME: Three Hours **Start:** 5H30 PM **Finish:** 8H30 PM

INSTRUCTIONS

1. This exam is marked out of 100 marks
2. This Examination comprises TWO Sections
3. Section A: Compulsory Question (25 marks)
4. Section B: Long Answer Questions (75 marks)
5. All questions in Section A are compulsory and Answer any THREE questions in Section B
6. This online exam shall take 3 Hours
7. Late submission of the answers will not be accepted
8. Ensure your web-camera is on at all times during the examination period
9. No movement is allowed during the examination
10. Idling of your machine for 5 min or more will lead to lock out from the exam
11. The Learning Management System (LMS) has inbuilt integrity checks to detect cheating
12. Any aspect of cheating detected during and or after the exam administration will lead to nullification of your exam
13. In case you have any questions call the invigilator for this exam on Tel. +254725984499 and write an email to michel.mutabazi@mcampus.amref.ac.ke
14. For adverse incidences please write an email to: amiu.examinations@amref.ac.ke

SECTION A: ANSWER ALL QUESTIONS

1. A researcher surveys a group of people about their daily water consumption in liters. The data is presented as the following frequency distribution:

Daily Water Consumption (liters)	Frequency (f)
1	3
1.5	5
2	8
2.5	10
3	4
Total	30

Calculate

- The mode (1 mark)
 - The mean (1 marks)
 - The median (1 marks)
 - The Standard deviation (1 marks)
 - The coefficient of variation (1 mark)
2. A researcher is studying the relationship between the number of hours of study and the scores on a biology test among a group of students. The following data shows the number of study hours and the corresponding test scores for each student:

Student	Hours of Study (X)	Test Score (Y)
1	4	65
2	2	55
3	6	75
4	5	70
5	1	50
6	3	60

- Compute the Spearman's Rank Correlation Coefficient (ρ) between the number of hours studied and test scores. (3 marks)
- Interpret the result (2 marks)

3. Below is a dataset of age (in years) and systolic blood pressure (in mmHg) for a sample of 10 individuals:

Age (years) X	Systolic Blood Pressure (mmHg) Y
25	120
30	125
35	130
40	135
45	140
50	145
55	150
60	155
65	160
70	165

a) Determine the regression equation that best describes the relationship between age and systolic blood pressure in the form $Y = a + b X$ (5 marks)

b) Estimate the blood pressure for a person who is 38 years old (2 marks)

3. A researcher is studying the number of people who develop a certain disease in a population. The disease occurs with an average rate of 5 cases per 1,000 people per year, with a standard deviation of 2 cases per 1,000 people per year. Assuming the number of cases follows a normal distribution, calculate the probability that between 3 and 7 cases will occur in a random sample of 1,000 people in one year (3 marks).

4. A study was conducted to determine if there is a relationship between smoking status and the presence of a particular lung disease. The following contingency table was constructed based on the results from 200 participants:

	Disease Present	Disease Absent	Total
Smoker	50	70	120
Non-Smoker	30	50	80
Total	80	120	200

Test the hypothesis at 95% confidence level (5 marks)

SECTION B: ANSWER ANY THREE (3) QUESTIONS

5. A researcher wants to study whether a new medication improves blood pressure in patients. The blood pressure of 15 patients is measured before taking the medication and then again after 8 weeks of using the medication. The researcher wants to test whether there is a significant difference in blood pressure before and after the medication.

The data (in mmHg) for blood pressure before and after the medication for 15 patients is as follows:

Before	After
130	125
140	135
150	145
160	155
155	150
165	160
170	165
140	138
130	128
145	140
155	150
160	158
150	148
140	136
135	130

Conduct a paired samples t-test to determine if there is a significant difference in the blood pressure of the patients before and after using the medication at 95% confidence level (25 marks)

6. a) Consider the information given below.

- **Plant-based diet group:**

Sample size $n_1 = 30$, sample mean $\bar{X}_1 = 2200$, sample standard deviation $s_1 = 250$.

- **Non-plant-based diet group:**

Sample size $n_2 = 30$, sample mean $\bar{X}_2 = 2400$, sample standard deviation $s_2 = 300$.

Test whether there is a significant difference in the mean calorie intake between two groups of people: those who consume a plant-based diet and those who consume a non-plant-based diet at 95% confidence level. Assume equality of variance. (20 marks)

b) A hospital records that on average, 3 patients per day arrive in the emergency department with a particular type of injury. Assume the number of patients arriving per day follows a Poisson distribution. Determine the probability that exactly 5 patients will arrive at the emergency department in a single day. (5 marks)

7. a) We want to compare the mean cholesterol levels between the two groups and calculate a 95% confidence interval for the difference between the means.

Data:

- Mean cholesterol level in the intervention group = 180 mg/dL
- Mean cholesterol level in the control group = 200 mg/dL
- Standard deviation of cholesterol levels in the intervention group = 15 mg/dL
- Standard deviation of cholesterol levels in the control group = 20 mg/dL
- Sample size for each group = 30

Estimate the 95% confidence interval for the difference between two means and interpret the result (10 marks)

b) A public health campaign aims to increase vaccination rates. We want to compare the vaccination rates in two different communities, one with the campaign and one without.

Data:

- **Community 1 (with campaign):**
 - Number of vaccinated individuals = 400
 - Total population in community = 500
 - Proportion vaccinated = $\frac{400}{500} = 0.8$
- **Community 2 (without campaign):**
 - Number of vaccinated individuals = 250
 - Total population in community = 500
 - Proportion vaccinated = $\frac{250}{500} = 0.5$

Calculate the 95% confidence interval for the difference in vaccination proportions between the two communities and interpret the results (10 marks).

c) Explain how you can use the point estimation in research giving one example (5 marks)

8. In a company, the HR department wants to assess if employee satisfaction is independent of the department they work in. Employees were surveyed about their level of satisfaction with the work environment and were asked to rate it as either "Satisfied" or "Dissatisfied". The company has three departments: Marketing, Sales, and Operations. The data collected from the survey is as follows:

- 60 employees from Marketing
- 50 employees from Sales
- 40 employees from Operations

The number of employees in each department who reported being "Satisfied" or "Dissatisfied" are given below:

Department	Satisfied	Dissatisfied	Total
Marketing	30	30	60
Sales	35	15	50
Operations	25	15	40
Total	90	60	150

Using chi-square test at 95% confidence level, test the hypothesis that employee satisfaction is independent of the department they work in. (25 marks)

9. a) In a reproductive health clinic, a new fertility treatment has a success rate of 70% for women trying to conceive. If 6 women undergo the treatment, determine the probability that exactly 4 of them will conceive successfully. Use the binomial distribution (10 marks)

b) A study from 10 years ago found that the average age for starting contraception was 21 years. A researcher has a sample of 100 women from this year's data, and he wants to determine whether the average age at which women are starting contraception has changed.

Data:

- Sample size (n) = 100
- Sample mean (\bar{x}) = 22.5 years
- Population mean (μ) = 21 years (from the study 10 years ago)
- Population standard deviation (σ) = 3 years (from historical data)

Compare this year's average age with the previous decade's average to see if there is a significant difference using z-test at 95% confidence level (15 marks)
