**AMREF INTERNATIONAL UNIVERSITY**

**SCHOOL OF PUBLIC HEALTH**

**DEPARTMENT OF COMMUNITY HEALTH**

**MASTER OF PUBLIC HEALTH**

**SPECIAL/SUPPLEMENTARY EXAMINATION - MARCH 2024**

**MPH 712**: **BIOSTATISTICS**

**DATE:** March 2024

**TIME:** Three Hours **Start:** 1600 Hours **Finish**: 1900 Hours

**INSTRUCTIONS**

1. This exam is marked out of 100 marks
2. This Examination comprises TWO Sections
   **Section A: Compulsory Question (25 marks)**
   **Section B: Long Answer Questions (75 marks)**
3. All questions in Section A are compulsory and Answer any THREE questions in Section B
4. This online exam shall take 3 Hours
5. Late submission of the answers will not be accepted
6. Ensure your web camera is on at all times during the examination period
7. No movement is allowed during the examination
8. Idling of your machine for 5 min or more will lead to lock out from the exam
9. The Learning Management System (LMS) has inbuilt integrity checks to detect cheating
10. Any aspect of cheating detected during and or after the exam administration will lead to nullification of your exam.
11. In case you have any questions call the invigilator on Head of Department on Tel +254 720 573 449 or Ag Head of Department +254 723 742 370
12. For adverse incidences please write an email to: amiu.examinations@amref.ac.ke

**Question 1**

Consider the data below

| Blood pressure (mm Hg) | Number of patients |
|---|---|
| 115-124 | 4 |
| 125-134 | 5 |
| 135-144 | 5 |
| 145-154 | 7 |
| 155-164 | 5 |
| 165-174 | 4 |
| 175-184 | 5 |
| Total | 35 |

Determine and interpret the:

a) Mode (3 marks)

b) Median (3 marks)

c) Mean (3 marks)

d) Semi-interquartile range (4 marks)

e) Sample variance (3 marks)

f) Sample standard deviation (3 marks)

g) Coefficient of variation (3 marks)

h) Skewness (3 marks)



**SECTION B: ANSWER ANY THREE (3) QUESTIONS**

**Question 2**

A sample of 200 people with a particular disease was selected. Out of these, 100 were given a drug and others were not given any drug. The results are as follows.

| | Drug | No drug | Total |
|---|---|---|---|
| **Cured** | 65 | 55 | 120 |
| **Not cured** | 35 | 45 | 80 |
| **Total** | 100 | 100 | 200 |

Test whether the drug is effective or not using Chi-square test. (25 marks)

**Question 3**

a) A Hospital records the weights of every new born child at the hospital. The distribution of weights is normally shaped, has mean of 2.9 Kg, and has a standard deviation of 0.45 Kg. Determine the following:

i. The percentage of new borns who weighed under 2.1 Kg (7 marks)
ii. The percentage of new borns who weighed between 1.8 Kg and 4.0 Kg (7 marks)
iii. If 1500 babies have been born at the hospital, how many weighed less than 2.5 Kg? (5 marks)

b) Suppose that it is known that 30 percent of a certain population is immune to some disease. If a random sample of size 10 is selected from this population, what is the probability that it will contain exactly four immune persons? Use the Binomial Distribution. (6 marks)

**Question 4**

The following data shows the marks obtained by 10 students in Anatomy and Physiology.

| % Marks in Anatomy | % Marks in Physiology |
|---|---|
| 78 | 84 |
| 45 | 55 |
| 36 | 50 |
| 78 | 60 |

| | |
|---|---|
| 62 | 82 |
| 90 | 86 |
| 65 | 58 |
| 75 | 60 |
| 39 | 47 |
| 41 | 51 |

a) Determine the Pearson's correlation coefficient. (15 marks)
b) Determine the rank correlation coefficient. (10 marks)

**Question 5**

For a random sample of 31 individuals selected from the population of 12-40 years with foetal alcohol syndrome, the mean height is 147.4cm. If the true mean height is 160.0cm with a standard deviation of 6cm. Test, the claim that the mean height of this group is 160.0cm. Use 5% level of significance and show all the steps. (25 Marks).

**Question 6**

Derive the formula for:

a) The test statistic for testing the null hypothesis of equal means when population variances are known. (6 marks)
b) The test statistic for testing the null hypothesis of equal means when the population variances are unknown. (6 marks)
c) Confidence interval for the difference between two population means. (6 marks)
d) Confidence interval for the difference between two proportions. (7 marks)

---

Statistical formulas and tables may be used.

1. $z = \frac{(x-\mu)}{\sigma}$

2. $\chi^2 = \sum \frac{(o-e)^2}{e}$

3. $\phi^2 = \frac{\chi^2}{N}$

4. $b = \frac{\sum_{i=1}^{n} X_i Y_i - n\bar{X}\bar{Y}}{\sum_{i=1}^{n} X_i^2 - n\bar{X}^2}$ and $a = \bar{Y} - b\bar{X}$ for the line $y = a + bx$

5. $r_s = 1 - \frac{6\sum_{1}^{n} d_i^2}{n(n^2-1)}$

6. $r_{xy} = \frac{\frac{1}{n}\sum_{i=1}^{n} X_i Y_i - \bar{X}\bar{Y}}{\sqrt{\left(\frac{1}{n}\sum_{i=1}^{n} X_i^2 - \bar{X}^2\right)\left(\frac{1}{n}\sum_{i=1}^{n} Y_i^2 - \bar{Y}^2\right)}}$

[redacted]

**Formulas**

- $Z = \frac{\overline{X} - \mu_o}{\delta / \sqrt{n}}$ $\qquad Z = \frac{X - \mu}{\delta}$
- $s = \sqrt{\frac{\sum_{i=1}^{n} \left(X_i - \overline{X}\right)^2}{n-1}}$ $\qquad \overline{X} = \frac{\sum X}{n}$
- $Z = \frac{\left(\overline{X}_1 - \overline{X}_2\right) - \left(\mu_1 - \mu_2\right)}{\sqrt{\frac{\delta_1^{\,2}}{n_1} + \frac{\delta_2^{\,2}}{n_2}}}$;
- $t = \frac{\left(\overline{X}_1 - \overline{X}_2\right) - \left(\mu_1 - \mu_2\right)}{\sqrt{s_p^{\,2}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$, $\qquad s_p^{\,2} = \frac{(n_1 - 1)s_1^{\,2} + (n_2 - 1)s_2^{\,2}}{(n_1 - 1) + (n_2 - 1)}$
- $t = \frac{\overline{X} - \mu_o}{s / \sqrt{n}}$; $\qquad t = \frac{\overline{X} - \mu}{s_{\bar{x}}}, \quad s_{\bar{x}} = \frac{s}{\sqrt{n-1}}$
-
- $t = \frac{\left(\overline{X}_1 - \overline{X}_2\right)}{S\left(\overline{X}_1 - \overline{X}_2\right)}$; where, $S(\overline{X}_1 - \overline{X}_2) = \sqrt{\frac{S_p^{\,2}}{n_1} + \frac{S_p^{\,2}}{n_2}}$; with, $S_p = \sqrt{\frac{(n_1 - 1)s^2{}_1 + (n_2 - 1)s^2{}_2}{n_1 + n_2 - 2}}$
- $P(X = r) = \frac{n!}{(n-r)!r!} p^r q^{n-r}$ OR $P(X = r) = \binom{n}{r} p^r q^{n-r}$
- $P(x) = \frac{e^{-\lambda} \lambda^x}{x!}$
- $S_B^{\,2} = SSB = \sum \frac{n(\overline{X} - \overline{X}_{GM})^2}{k-1}$ $\qquad S_W^{\,2} = SSW = \frac{\sum (n_i - 1) S_i^{\,2}}{\sum (n_i - 1)}$
- $\overline{X}_{GM} = \frac{\sum X}{N}$ $\qquad F = \frac{SSB}{SSW} = \frac{S_B^{\,2}}{S_W^{\,2}}$
- $T = \sum \frac{T_i^2}{n_i} - \overline{X}_{GM}$ $\qquad T_{SS} = \sum y^2{}_{ij} - \overline{X}_{GM}$
- $C.I = \overline{X} \pm Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$ (large sample) or $C.I = \overline{X} \pm t_{df, n-1} \frac{s}{\sqrt{n}}$ (small sample)
- $C.I = p_s \pm Z_{\alpha/2} \sqrt{\frac{p_s(1 - p_s)}{n}}$ or $\left(p_{s_1} - p_{s_2}\right) \pm Z \sqrt{\overline{p}(1 - \overline{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$
- $Expected\ Value = \frac{Row\ Sum \times Column\ Sum}{Grand\ Total}$; $\qquad \chi^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}$