NATIONAL UNIVERSITY OF SCIENCE AND TECHNOLOGY

FACULTY OF SCIENCE AND TECHNOLOGY EDUCATION

DEPARTMENT OF SCIENCE, MATHEMATICS AND TECHNOLOGY EDUCATION

STATISTICAL INFERENCE (PST2034)

Main Examination Paper **November 2024**

This Examination Paper consists of 5 printed pages

Time Allowed: 3 hours

Total Marks: 100

Special Requirements: Statistical tables

Internal Examiner: Mr T Moyo

External Examiner: Dr G Sunzuma

INSTRUCTIONS

1. Answer all questions in section A and any three questions in section B.
2. Each question should start on a fresh page
3. Marks will be allocated as indicated below

MARK ALLOCATION

QUESTION	MARKS
A1	4
A2	8
A3	9
A4	9
A5	10
B6	20
B7	20
B8	20
B9	20
TOTAL	100

SECTION A (40 Marks)

Candidates should attempt ALL questions in this section

A1. Distinguish between

a) Descriptive statistics and inferential statistics (2)

b) Type I error and type II error (2)

A2. Let $X_1, X_2, X_3, \dots, X_{10}$ be a random sample from a Bernoulli distribution with parameter θ where $\theta \in (0,1)$. It is desired to test $H_0 : \theta = 0.3 \quad H_1 : \theta = 0.5$

Decision rule: Reject H_0if $H_0 : Y = \sum_{i=1}^{10} X_i > 4$

a) Calculate α and β for the test (6)

b) Find power of the test (3)

A3. Let $X_1, X_2, X_3, \dots, X_n$ be a random sample from a normal distribution with mean μ and variance σ^2 i.e. $X \sim N(\mu, \sigma^2)$. Find the distribution of

a) $T_1 = \frac{\bar{x}-\mu}{\sigma}$ (3)

b) $T_2 = n\frac{(\bar{x}-\mu)^2}{\sigma^2}$ (3)

c) $T_3 = \sqrt{n}\left(\frac{\bar{x}-\mu}{\sigma}\right)$ (3)

A4: A professor carried out an investigation to prove the claim that performance of disabled children at a school tends to increase as they get older. Raw scores on a task involving basic knowledge concepts are shown below.

7 year olds	15	20	24	26	34	37	43			
8 year olds	23	28	27	34	23	47				
10 year olds	22	23	34	32	63	36	38	43	48	50

Use Kruskal-Wallis H test to determine if the claim is true. (8)

A5. a) Let $X_1, X_2, \ldots, X_n$ be a random sample from a population with variance σ^2. Prove that

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2 \qquad x = 1,2,\ldots$$

is an unbiased estimator of σ^2. (6)

b) Suppose that 175 heads and 225 tails were obtained from 400 tosses of a coin. Find a 90% confidence interval for the probability of a tail. (4)

Section B (60 Marks)

Answer any THREE questions being careful to number them B6 to B9.

B6. Let $X_1, X_2, X_3, \ldots, X_n$ be a random sample from a distribution with pdf

$$p(x) = \frac{1}{\theta}\left(1 - \frac{1}{\theta}\right)^x \qquad x = 1,2,\ldots$$

a) Find the maximum likelihood estimator of θ. (8)

b) Show that the estimator is (i) unbiased (4)

(ii) consistent (4)

(iii) sufficient (4)

B7. a) State and explain the Central Limit Theorem. (5)

b) Explain the properties /characteristics of a chi-square test. (4)

c) An educational researcher wanted to determine if there is an association between teacher's gender and attitude towards online teaching. After undertaking a survey on a sample of 400 teachers, the following results were obtained

		Attitude towards online teaching		
		In favour of	**Neutral**	**Negative**
Gender	**Male**	91	58	51
	Female	79	57	64

(i) Carry out a test at 1% significance level to ascertain if teacher's gender and attitude towards online teaching are independent. (10)

(ii) Comment on the results obtained (1)

B8. a) The relationship between X and Y is given by the regression model

$Y_i = \beta_0 + \beta_1 X_i + e_i$ for $i = 1, 2, \ldots, n$ where $e_i \sim N(0, \sigma^2)$.

Show that the Least Squares Estimator of β_0 is given by $\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X}$ (8)

b) A sugar manufacturer sells sugar in bags with a stated mass of 500g. If bags are sold which consistently weigh more than 500g or less than 500g then this has implications for the manufacturer.

(i) State the implications to the manufacturer in both cases. (4)

(ii) A consumer refutes the claim of the manufacturer that the mass of the sugar is 500g. The data below shows the masses for a random sample of 20 bags which were weighed to verify the consumer's remarks.

489.6	493.7	503.1	499.3	507.8	502.5	498.2	495.3	497.4	488.6
507.7	501.0	491.1	498.2	492.3	510.8	496.6	495.5	504.0	493.4

Use 5% significance level to test if the manufacturer's claim is true. (8)

B9. a) Let $X_1, X_2, X_3, \ldots, X_n$ be a random sample from a Gamma distribution

$$f(x, r, \lambda) = \frac{\lambda^r x^{r-1}}{\Gamma(r)} e^{-\lambda x} \quad ; x > 0, \qquad r > 0, \qquad \lambda > 0$$

i) Show that $E(X) = \frac{r}{\lambda}$ and $E(X^2) = \frac{r(r+1)}{\lambda^2}$ (3,3)

ii) Use method of moments to find estimators of r and λ. (4,4)

b) Independent random samples from normal populations produced the following results.

Sample 1	0.1	2.4	1.4	3.2	2.5	3.0	4.0	4.3	3.1	6.4
Sample 2	0.2	2.3	1.2	3.2	2.4	3.1	4.0	4.2	2.0	6.3

i) Calculate the pooled estimate, S_p^2 , of the variance, σ^2. (2)

ii)Find a 95% confidence interval for the difference of the means. (4)

END OF EXAMINATION PAPER

LIST OF SELECTED FORMULAE

1. a) $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} X_i$

   b) $s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$

2. a) $\bar{x} \pm t_{\frac{\alpha}{2}}(n-1) \times \frac{s}{\sqrt{n}}$

   b) $\hat{p} \pm Z_{\frac{\alpha}{2}} \times \sqrt{\frac{p(1-p)}{n}}$

   c) $\bar{d} \pm t_{\frac{\alpha}{2}}(n-1) \times \frac{s_d}{\sqrt{n}}$

3. a) $(\bar{x}_1 - \bar{x}_2) \pm t_{\frac{\alpha}{2}}(n_1 + n_2 - 2) \times \sqrt{S_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$ where $S_p^2 = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2 +}{n_1 + n_2}$

   b) $(\bar{x}_1 - \bar{x}_2) \pm t_{\frac{\alpha}{2}}(k) \times \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$ where $k = \left[\frac{\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)^2}{\frac{(S_1^2/n_1)^2}{n_1-1} + \frac{(S_2^2/n_2)^2}{n_2-1}}\right]$

4. a) $(p_1 - p_2) \pm Z_{\frac{\alpha}{2}} \times \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}$

   b) $(p_1 - p_2) \pm Z_{\frac{\alpha}{2}} \times \sqrt{\hat{p}\hat{q}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$ where $\hat{p} = \frac{n_1\hat{p}_1 + n_2\hat{p}_2}{n_1 + n_2}$

5. a) $Z_{cal} = \frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}}$ and $t_{cal} = \frac{\bar{x}-\mu}{\frac{S}{\sqrt{n}}}$

   b) $t_{cal} = \frac{\hat{p}-p}{\sqrt{\frac{p(1-p)}{n}}}$

6. a) $Z_{cal} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$

   b) $t_{cal} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{S_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$

   c) $t_{cal} = \frac{\bar{d}}{\frac{s_d}{\sqrt{n}}}$

7. $H = \frac{12}{N(N+1)}\sum_{j=1}^{k}\frac{R_j^2}{N_j} - 3(N+1)$ and $H_c = \frac{H}{c}$ where $c = 1 - \frac{\sum(T^3 - T)}{N^3 - N}$