NATIONAL UNIVERSITY OF SCIENCE AND TECHNOLOGY

FACULTY OF APPLIED SCIENCE

DEPARTMENT OF STATISTICS AND OPERATIONS RESEARCH

SORS 2110 :INTRODUCTION TO APPLIED STATISTICS

BSc Sports Science & Coaching: Part 2

DECEMBER 2024 EXAMINATION

Time : 3 hours

Candidates may attempt **ALL** Questions in Section **A** and **THREE** Questions in Section **B**. Show all working where necessary. You may use a calculator, Graph paper and Statistical Tables will be provided.

SECTION A: Answer all questions in this section (40 marks).

A1. Using relevant examples, differentiate the following terms:-

(a) Probability Sampling and Non-Probability Sampling, [4]

(b) Parameter and Statistic, [2]

(c) Confidence Level and Significance Level. [4]

A2. The 100 metre race time for 12 athletes is listed below .
1.03 ; 1.32 ; 1.05 ; 1.00 ; 1.13 ; 1.06 ; 1.12 ; 1.03 ; 1.10 ; 1.24 ; 1.42 ; 1.37.

Determine

(a) the Median time. [2]

(b) the Modal time. [2]

(c) the Variance. [4]

(d) the Range. [2]

A3. Discuss any four (4) probability sampling techniques. [20]

SECTION B: Answer any THREE questions in this section (60 marks).

B4. A random sample of 201 Zimbabwe Premier Soccer players was used to collect data on the amount of time spent training per day . The results are displayed in Table 1

Table 1: Time spent at the gym

Time spent (t minutes)	0-15	15-30	30-45	45-60	60-75	75-90
Number of players	15	33	62	41	34	16

(a) Construct a suitable Frequency Distribution Table for the data. [6]

(b) Calculate the mean time spent training per day. [2]

(c) Calculate the modal time spent training per day. [4]

(d) Calculate the median time spent training per day. [4]

(e) Calculate the standard deviation for the data. [4]

B5. Table 2 shows information on the number of games played and the number of injuries per Rugby player.

Table 2: Number of Rugby games and injuries

Games	37	43	58	49	70	28	65	32	27	14	20	23
Injuries	6	7	5	6	4	9	6	7	6	8	7	8

(a) Construct a scatter plot for the data. [3]

(b) Fit a linear regression model for the data. [7]

(c) Calculate and interpret the Pearson's Correlation Coefficient for the data. [5]

(d) Calculate and interpret the Spearman's Rank Correlation Coefficient for the data. [5]

B6. (a) Table 3 displays results of a shooting competition between seven Police and Army officers.

Table 3: Scores (%)

Police	64.2	58.7	62.1	62.5	59.8	59.2
Army	57.8	56.2	60.9	54.4	53.6	53.2

Use an appropriate t-test at the 5% level of significance to test whether there is any significant difference on the mean scores for Police and Army Officers. [10]

(b) Seven athletes were recorded blood pressure readings. Table 4 shows systolic blood pressure readings (in mmHg) recorded over two consecutive days for each athlete.

Table 4: Blood Pressure Readings (mmHg)

Athlete	1	2	3	4	5	6	7
Day 1	107	130	121	145	149	166	125
Day 2	120	130	115	143	156	148	130

Use an appropriate t-test at the 5% level of significance to test whether there is any significant difference in systolic blood pressure readings recorded over the two days. [10]

B7. (a) The Cricket Team Statistician recorded scores over 10 games. Table 5 shows the actual and expected number of boundaries for each game.

Table 5: Actual and Expected Number of Boundaries

Game	1	2	3	4	5	6	7	8	9	10
Actual	14	17	11	10	6	3	5	9	12	13
Expected	10	10	10	10	10	10	10	10	10	10

Use the Chi-Square Test to assess whether there is any significant difference between the expected and actual number of boundaries recorded per game. [10]

(b) Table 6 is a partially completed ANOVA table for a Completely Randomised Design.

Table 6: Incomplete Summary ANOVA Table

Source	df	SS	MS	F
Treatments	**	2709.20	**	35.81
Error	36	**	**	
Total	39	3617.06		

(i) Copy and complete the table. [4]

(i) Test if there is sufficient evidence to indicate a difference among the treatment means given that $\alpha = 0,05$. [6]

LIST OF SELECTED FORMULAE

Population Proportion	Two Population Proportions	Population Mean
Parameter p	Parameter p_1-p_2	Parameter μ
Statistic $\hat{p}$	Statistic $\hat{p}_1$-$\hat{p}_2$	Statistic $\hat{\mu}$
Standard Error $s.e(\hat{p}) = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$	Standard Error $s.e(\hat{p}_1 - \hat{p}_2) = \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$	Standard Error $s.e(\bar{x}) = \frac{s}{\sqrt{n}}$
Confidence Interval	Confidence Interval	Confidence Interval
$\hat{p} \pm z * s.e(\hat{p})$	$(\hat{p_1} - \hat{p_2}) \pm z * s.e(\hat{p_1} - \hat{p_2})$	$\bar{x} \pm t^* s.e(\bar{x})$ df=$n-1$ Paired Confidence Interval $\bar{d} \pm t^* s.e(\bar{d})$ df=$n-1$
Large-Sample z-Test $z_0 = \frac{\hat{p}-p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$	Large-Sample z-Test $z_0 = \frac{\hat{p_1}-\hat{p_2}}{\sqrt{\hat{p}(1-\hat{p})}\sqrt{\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}$ where $\hat{p} = \frac{x_1 + x_2}{n_1 + n_2}$	One-Sample t-test $t_0 = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}}$ df=$n-1$ paired t-Test $t_0 = \frac{\bar{d}}{\frac{s_d}{\sqrt{n}}}$ df=$n-1$

Table 7: Two Population Means

General	Pooled
Parameter $\mu_1 - \mu_2$	Parameter $\mu_1 - \mu_2$
Statistic $\mu_1 - \mu_2$	Statistic $\mu_1 - \mu_2$
Standard Error $s.e(\bar{x}_1 - \bar{x}_2) = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$	Standard Error pooled s.e$(\bar{x}_1 - \bar{x}_2) = s_p\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$ where $s_p^2 = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}$
Confidence Interval $(\bar{x}_1 - \bar{x}_2) \pm t * s.e(\bar{x}_1 - \bar{x}_2)$ df=min$(n_1 - 1, n_2 - 1)$	Confidence Interval $(\bar{x}_1 - \bar{x}_2) \pm t *$ (pooled s.e$(\bar{x}_1 - \bar{x}_2)$) df=$n_1 + n_2 - 2$
Two-Sample t-Test $t_0 = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$ df=min$(n_1 - 1, n_2 - 1)$	Pooled Two-Sample t-Test $t_0 = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{s_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$ df=$n_1 + n_2 - 2$ df=min$(n_1 - 1, n_2 - 1)$

END OF QUESTION PAPER