

**FINAL EXAMINATION**

**BACHELOR OF COMMUNICATION (HONS) IN CORPORATE COMMUNICATION
BACHELOR OF BUSINESS ADMINISTRATION (HONOURS)
HUMAN RESOURCE MANAGEMENT
BACHELOR OF BUSINESS ADMINISTRATION (HONOURS)**

COURSE : STATISTICS FOR SOCIAL SCIENCES

COURSE CODE : STA2113

DURATION : 3 HOURS

INSTRUCTIONS TO CANDIDATES:

1. This question paper consists of **FOUR (4)** questions.
2. Answer **ALL** questions in the Answer Booklet provided.
3. Please check to make sure that this examination pack consists of:
 - i. The Question Paper
 - ii. an Answer Booklet
 - iii. Appendix 1(1) and 1(2)
4. Do not bring any material into the examination hall. Electronic calculator is allowed.
5. Please write your answer using permanent ink.

MYKAD/PASSPORT NO : _____

ID. NO. : _____

LECTURER : _____

SECTION : _____

DO NOT OPEN THIS QUESTION PAPER UNTIL YOU ARE TOLD TO DO SO

This question paper consists of 10 printed pages including the front page

QUESTION 1

- a. A fitness researcher wants to study the relationship between the number of hours spent exercising per week and body fat percentage among adults. Data were collected from 15 adults on their weekly exercise hours and body fat percentage. The number of exercise hours ranges from 1 to 10 hours per week. Statistical software was used to analyse the data, and the results are shown in Tables 1 and 2 below.

Table 1

<i>Regression Statistics</i>	
Multiple R	0.875
R Square	X
Adjusted R Square	0.745
Standard Error	2.150
Observations	15

Table 2

	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>
Intercept	31.25	1.420	22.007	0.000
Exercise Hours	-1.850	0.320	-5.781	0.000

- i. State the dependent and independent variables. (2 marks)
- ii. Name a measure used to determine the strength of the relationship between the variables. State its value and interpret it. (4 marks)
- iii. Identify a suitable graph to display the relationship. (1 mark)
- iv. Compute the value of **X** in Table 1 and explain its meaning. (3 marks)
- v. State the y-intercept, **a** and slope coefficient, **b**. Hence, determine the regression equation. (4 marks)
- vi. Interpret the slope coefficient, **b** in (v). (2 marks)
- vii. Predict the body fat percentage of an individual who exercises 6 hours per week. Is the estimation reliable? Explain. (3 marks)

- b. A company wants to evaluate the quality of its customer service by analysing the number of customer complaints received each day. The company recorded the daily number of complaints over a period of 35 working days. The summary statistics of the data are presented in Table 3 below.

Table 3	
<i>Number of complaints</i>	
Mean	12.40
Standard Error	0.52
Median	12.00
Mode	11.00
Standard Deviation	3.08
Kurtosis	-0.18
Skewness	0.22
Minimum	6
Maximum	19
Count	35

- i. Name **TWO (2)** methods used to assess normality. (2 marks)
- ii. Based on Table 3 above, can we indicate that the distribution is normally distributed? Explain. (2 marks)
- iii. For a normal distribution, where are the mean, median and mode located relative to each other? (2 marks)

(Total: 25 marks)

QUESTION 2

- a. A smartphone manufacturer claims that the average battery life of its new model is 18 hours on a full charge. To verify this claim, a random sample of 25 smartphones was selected and tested under standard operating conditions. The battery life of each phone was recorded, and the results were analysed using statistical software. The summary statistics are presented in Table 4 below.

Table 4
One-Sample Statistics

Battery Life (hours)	n	Mean	Std. Deviation	Std. Error Mean
	25	17.20	2.50	Y

- i. Calculate Y. (2 marks)
- ii. State the null and alternative hypotheses. (2 marks)
- iii. Construct a 95% confidence interval for the true population mean battery life. (Given $t_{0.025,24} = \pm 2.064$). Explain the values obtained. (5 marks)
- iv. Based on the confidence interval in (iii), comment on the manufacturer's claim. (3 marks)

- b. A school is interested in determining whether two different teaching methods produce different mathematics achievement levels among students. A random sample of students was taught using either Method A or Method B, and their mathematics test scores were recorded. The results were analysed using an independent two-sample t-test assuming equal variances. The summary statistics are presented in Table 5 below.

Table 5
t-Test: Two-Sample Assuming Equal Variances

	<i>Method A</i>	<i>Method B</i>
Mean	78.5	73.0
Variance	20.25	18.00
Observations	15	15
Pooled Variance	19.13	
Hypothesized Mean Difference	0	
df	28	
t Stat	3.455	
P(T<=t) one-tail	0.001	
t Critical one-tail	1.701	
P(T<=t) two-tail	0.002	
t Critical two-tail	2.048	

- i. State the null and alternative hypotheses. (2 marks)
- ii. Using the critical method approach, determine whether there is a significant difference in the mean score between both methods at $\alpha = 0.01$. (3 marks)

- c. A nutritionist wants to evaluate the effectiveness of a 6-week diet programme in helping participants lose weight. The weight of 18 participants was measured before the programme began and after completing the programme. The results were analysed using a paired-sample t-test to determine whether there was a significant change in the mean weight of participants. The summary statistics are presented in Table 6 below.

Table 6
t-Test: Paired Two Sample for Means

	<i>Before</i>	<i>After</i>
Mean	82.4	79.8
Variance	24.0	20.0
Observations	18	18
Pearson Correlation	0.88	
Hypothesized Mean Difference	0	
df	17	
t Stat	<i>M</i>	
P(T<=t) one-tail	0.000	
t Critical one-tail	1.740	
P(T<=t) two-tail	0.000	
t Critical two-tail	2.110	

- i. Calculate the value of ***M***. (Given the standard deviation of the mean differences, $s_d = 1.80$).
(3 marks)
- ii. Using the p -value approach, determine whether there is a significant difference in mean weight of participants before and after the diet programme at $\alpha = 0.05$.
(3 marks)
- iii. Explain why a paired-sample t-test is more appropriate than an independent-sample t-test for this study.
(2 marks)

(Total : 25 marks)

QUESTION 3

- a. A bottled water company is evaluating the consistency of the volume filled by two automated filling machines, Machine A and Machine B. Random samples of bottles were collected from each machine, and the filling volumes were measured. An F-test was performed at the 10% significance level to determine whether there is a significant difference in the variability of the filling volumes. The results are shown in Table 7 below.

Table 7
F-Test Two-Sample for Variances

	<i>Machine A</i>	<i>Machine B</i>
Mean	500.8	501.2
Variance	2.25	1.44
Observations	20	20
df	19	19
F	<i>K</i>	
P(F<=f) one-tail	0.146	
F Critical one-tail	2.168	

- i. State the null and alternative hypotheses. (2 marks)
- ii. Calculate the value of *K*. (2 marks)
- iii. Using the critical-value approach, determine whether there is a significant difference in variability between Machine A and Machine B at $\alpha = 0.10$. (3 marks)

- b. A university is investigating the effectiveness of three online learning platforms on students' statistics achievement. Thirty students were randomly assigned equally to Platform A, Platform B, and Platform C. Their final examination scores were analysed using a one-way ANOVA at the 1% significance level. The results are presented in Table 8 and Table 9.

Table 8

ANOVA : Single Factor
SUMMARY

<i>Groups</i>	<i>Count</i>	<i>Sum</i>	<i>Average</i>	<i>Variance</i>
Platform A	10	720	72.0	25.0
Platform B	10	790	79.0	36.0
Platform C	10	755	75.5	49.0

Table 9

ANOVA

<i>Source of Variation</i>	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>P-value</i>	<i>F crit</i>
Between Groups	A	2	C	E	0.041	3.188
Within Groups	B	27	D			
Total	1310					

- Using the Sum of Squares Between formula, show that **A** = 245. Then, determine **B**, **C**, **D** and **E**.
(11 marks)
- State the null and alternative hypotheses for the ANOVA test.
(2 marks)
- Using the critical value approach at $\alpha = 0.01$, determine whether there is a significant difference in mean examination scores among the three learning platforms.
(3 marks)
- Identify the learning platform with the highest and lowest mean examination score.
(2 marks)

(Total: 25 marks)

QUESTION 4

- a. A shopping mall conducted a survey to identify customers' preferred promotion types. Customers were asked to select one of five promotion options: Discount Voucher, Cashback, Free Gift, Buy 1 Free 1, and Lucky Draw. The management wants to determine whether customer preferences are equally distributed among the five promotion types. The data were analysed using MS Excel at a significance level of $\alpha = 0.10$. The results are shown in Tables 10 and 11 below.

Table 10

Promotion Type	Observed N	Expected N	Residual
Discount Voucher	45	50	H
Cashback	I	50	14
Free Gift	38	50	-12
Buy 1 Free 1	56	J	6
Lucky Draw	47	50	-3

Table 11

Pearson Chi-Square	L
p-value	0.085
Critical value	7.779

- i. For a goodness-of-fit test with 5 categories, what is the degree of freedom (df)?
(1 mark)
- ii. Determine the values of **H**, **I** and **J**.
(3 marks)
- iii. Using the Pearson Chi-Square test formula, show that **L** = 8.20.
(3 marks)
- iv. State the null and alternative hypotheses for this study.
(2 marks)
- v. Based on the *p*-value in Table 11, decide whether there is a significant difference in customer promotion preferences among the five promotion types at the 10% significance level?
(3 marks)

- b. A researcher investigates whether education level is associated with consumers' preferred online shopping platform. A sample of 210 respondents was surveyed, and each respondent was classified according to their education level and preferred platform. The data were analysed using a Chi-Square Test of Independence at a significance level of $\alpha = 0.05$. The results are presented in Tables 12 and 13 below.

Table 12

			Shopping Platform			Total
			Shopee	Lazada	Tik Tok Shop	
Education Level	Diploma	Observed	30	25	15	70
		Expected	26.667	<i>P</i>	<i>Q</i>	70
	Degree	Observed	35	40	30	105
		Expected	<i>R</i>	42.500	22.500	105
	Postgraduate	Observed	15	20	0	35
		Expected	13.333	<i>S</i>	7.500	35
Total			80	85	45	210

Table 13

Pearson Chi-Square	14.191
p-value	0.00671
Critical value	5.991

- Compute the values of *P*, *Q*, *R* and *S*.
(8 marks)
- State the null and alternative hypotheses for this study.
(2 marks)
- Based on the critical value in Table 13, determine whether there a significant relationship between education level and preferred online shopping platform at $\alpha = 0.05$.
(3 marks)

(Total: 25 marks)

(TOTAL: 100 MARKS)

END OF QUESTION PAPER

APPENDIX 1(1)

Correlation and Regression

- i. Pearson's Product Moment Correlation Coefficient

$$r = \frac{n(\sum XY) - (\sum X)(\sum Y)}{\sqrt{[n(\sum X^2) - (\sum X)^2][n(\sum Y^2) - (\sum Y)^2]}}$$

- ii. The least-squares regression line of Y against X, $Y = a + bX$

$$b = \frac{n(\sum XY) - (\sum X)(\sum Y)}{n(\sum X^2) - (\sum X)^2}$$

$$a = \frac{(\sum y)}{n} - b \frac{(\sum x)}{n}$$

Confidence Intervals

- i. Mean μ , for large samples

$$\bar{x} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \quad \text{or} \quad \bar{x} \pm z_{\alpha/2} \frac{s}{\sqrt{n}}$$

- ii. Mean μ , for small samples

$$\bar{x} \pm t_{\alpha/2} \frac{s}{\sqrt{n}} \quad df = n - 1$$

- iii. Test statistic for two population mean, paired observation

$$t = \frac{\bar{d} - 0}{\frac{s_d}{\sqrt{n}}}$$

- iv. Difference in means, $\mu_1 - \mu_2$, for large and independent samples

$$(\bar{x}_1 - \bar{x}_2) \pm z_{\alpha/2} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \quad \text{or} \quad (\bar{x}_1 - \bar{x}_2) \pm z_{\alpha/2} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

- v. Difference in means, $\mu_1 - \mu_2$, for small and independent samples

$$(\bar{x}_1 - \bar{x}_2) \pm t_{\alpha/2} s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \quad df = n_1 + n_2 - 2$$

- vi. Difference in means, $\mu_1 - \mu_2 = \mu_d$ for paired samples

$$\bar{d} \pm t_{\alpha/2} \frac{s_d}{\sqrt{n}} \quad df = n - 1$$

APPENDIX 1(2)

Analysis of Variance (ANOVA)

- i. Degrees of freedom for the numerator = $k - 1$
- ii. Degrees of freedom for the denominator = $n - k$
- iii. Total sum of squares:

$$SST = \sum x^2 - \frac{(\sum x)^2}{n}$$

- iv. Between-samples sum of squares:

$$SSB = \left(\frac{T_1^2}{n_1} + \frac{T_2^2}{n_1} + \frac{T_3^2}{n_1} + \dots \right) - \frac{(\sum x)^2}{n}$$

- v. Within-samples sum of squares = $SST - SSB$

- vi. Mean Square between samples:

$$MSB = \frac{SSB}{(k - 1)}$$

- vii. Mean Square within samples:

$$MSW = \frac{SSW}{(n - k)}$$

- viii. Test statistic for a one-way ANOVA test:

$$F = \frac{MSB}{MSW}$$

Chi-square Test

- i. Test statistics for Chi-square test

$$\chi^2 = \sum \frac{(O - E)^2}{E}$$

- ii. $Expected\ value = \frac{row\ sum \times column\ sum}{grand\ total}$