



UNIVERSITI KUALA LUMPUR
INSTITUTE OF MEDICAL SCIENCE TECHNOLOGY

FINAL EXAMINATION
MARCH 2025 SEMESTER

COURSE CODE : HGB 30103
COURSE NAME : BIOSTATISTICS
PROGRAMME NAME : BACHELOR OF ENVIRONMENTAL HEALTH (Hons)
BACHELOR OF BIOMEDICAL SCIENCE (HONOURS)
BACHELOR OF OCCUPATIONAL SAFETY & HEALTH
(HONOURS)
DATE : 26 JUNE 2025
TIME : 9:00 AM – 12:00 PM
DURATION : 3 HOURS



INSTRUCTIONS TO CANDIDATES

1. Please CAREFULLY read the instructions given in the question paper.
2. This question paper has information printed on both sides of the paper.
3. This question paper consists of TWO sections.
4. Answer ALL questions for Section A.
5. Section B consist of four questions. Answer THREE (3) questions only.
6. Please write your answers on the answer booklet provided.
7. Answer all questions in English language ONLY.
8. Please answer MCQ/EMQ questions using OMR sheet.
9. Refer to the attached Formula/ Appendices.

THERE ARE 15 PAGES OF QUESTIONS, EXCLUDING THIS PAGE.

SECTION A (Total: 40 marks)

INSTRUCTION: Answer ALL questions.
Please use the answer booklet provided.

Question 1

Answer all the following questions.

- a. According to a report from the Ministry of Health Malaysia as of the 17th April 2020, 56.5% of Covid-19 positive patients have recovered. If 20 patients are randomly selected throughout Malaysia, what is the probability that 95% of the positive patients have recovered?

(5 marks)

- b. A manufacturing plant uses toluene diisocyanate, a known trigger for asthma. Workers have reported an increase in asthma diagnoses recently. Medical records of current workers from morning shift reveal that 42.5% of workers were diagnosed with asthma, while only 28.3% of employees from evening shift had such diagnoses. To investigate a potential link, a safety and health officer randomly selected 300 workers and 250 workers from the entire morning and evening shift, respectively.

What is the probability that the difference between the sample proportions of asthmatic workers in the two groups ($\hat{p}_1 - \hat{p}_2$) will be less than 5%?

Refer below – Table 1: Prevalence of asthma among workers from morning and evening shift

(5 marks)

Table 1: Prevalence of asthma among workers from morning and evening shift

Shift	<i>n</i>	%
Morning	300	42.5
Evening	250	28.3

Question 2

Many neurological symptoms, including cognitive impairment, memory loss, and personality changes, have been reported among paddy farmers. All paddy farmers claimed that they had been using organophosphate pesticides for a long period (chronic exposure). The prevalence of neurological symptoms can be found in the table below.

Refer below – Table 2: The prevalence of neurological symptoms among paddy farmers

Table 2: The prevalence of neurological symptoms among paddy farmers

Chronic exposure to pesticides	Neurological symptoms		Total
	Yes	No	
Yes	35	85	120
No	17	155	172
Total	52	240	292

- Compute the relative risk of neurological symptoms associated with the chronic exposure to pesticides. Explain your result.
(3 marks)
- Calculate the attributable risk. Explain your result.
(4 marks)
- Compute the attributable risk percentage in the exposed group. Explain your result.
(3 marks)

Question 3

A group of environmental officers measured the concentration of sulfur dioxide (SO₂) (mg/m³) released from incinerators used to manage clinical wastes. The measurements were taken from the same incinerators in February and July. Engineering measures were taken as corrective actions to reduce the release of SO₂ from the incinerators. The collected data were analysed using SPSS and are illustrated in the figure below.

Refer below – Figure 1: A.

		Statistics			
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	February	257.73	30	49.040	8.954
	July	174.87	30	15.558	2.840

		Correlations		
		N	Correlation	Sig.
Pair 1	February & July	30	-.371	.043

		Test							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
					Lower	Upper			
Pair 1	February - July	82.867	56.687	10.349	61.700	104.034	8.007	29	.000

Figure 1: A

- Assuming the data is normally distributed, determine the test used to analyse the above data and state the justification of using the selected test. (2 mark)
- State the null hypothesis for this analysis. (1 mark)
- Construct a reporting table to summarize the findings and interpret the results. (5 mark)
- Determine the decision-making process and state the conclusion based on the findings. (2 marks)

Question 4

A researcher has conducted a cross-sectional study to evaluate the association between the concentration of nitrogen dioxide (NO₂) emitted from industrial plants and lung function among residents in Rome, Italy. Spirometry was used to assess forced vital capacity (FVC), which served as the lung function indicator. The level of NO₂ exposure was measured in each respondent's house. The collected data were analysed using SPSS and are illustrated in the figures below.

Refer below – Figure 2: B.

Descriptive Statistics

	Mean	Std. Deviation	N
FVC (Liters)	4.151	.2748	35
NO2 Level (ppb)	26.63	6.353	35

Variables Entered/Removed^a

Model	Variables Entered	Variables Removed	Method
1	NO2 Level (ppb) ^b	.	Enter

a. Dependent Variable: FVC (Liters)

b. All requested variables entered.

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.990 ^a	.979	.978	.0403

a. Predictors: (Constant), NO2 Level (ppb)

ANOVA^a

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	2.514	1	2.514	1547.198	.000 ^b
	Residual	██████████	33	.002		
	Total	2.567	34			

a. Dependent Variable: FVC (Liters)

b. Predictors: (Constant) ██████████

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	t	Sig.
1	(Constant)	5.291	.030		177.758	.000
	NO2 Level (ppb)	-.043	.001	-.990	-39.334	.000

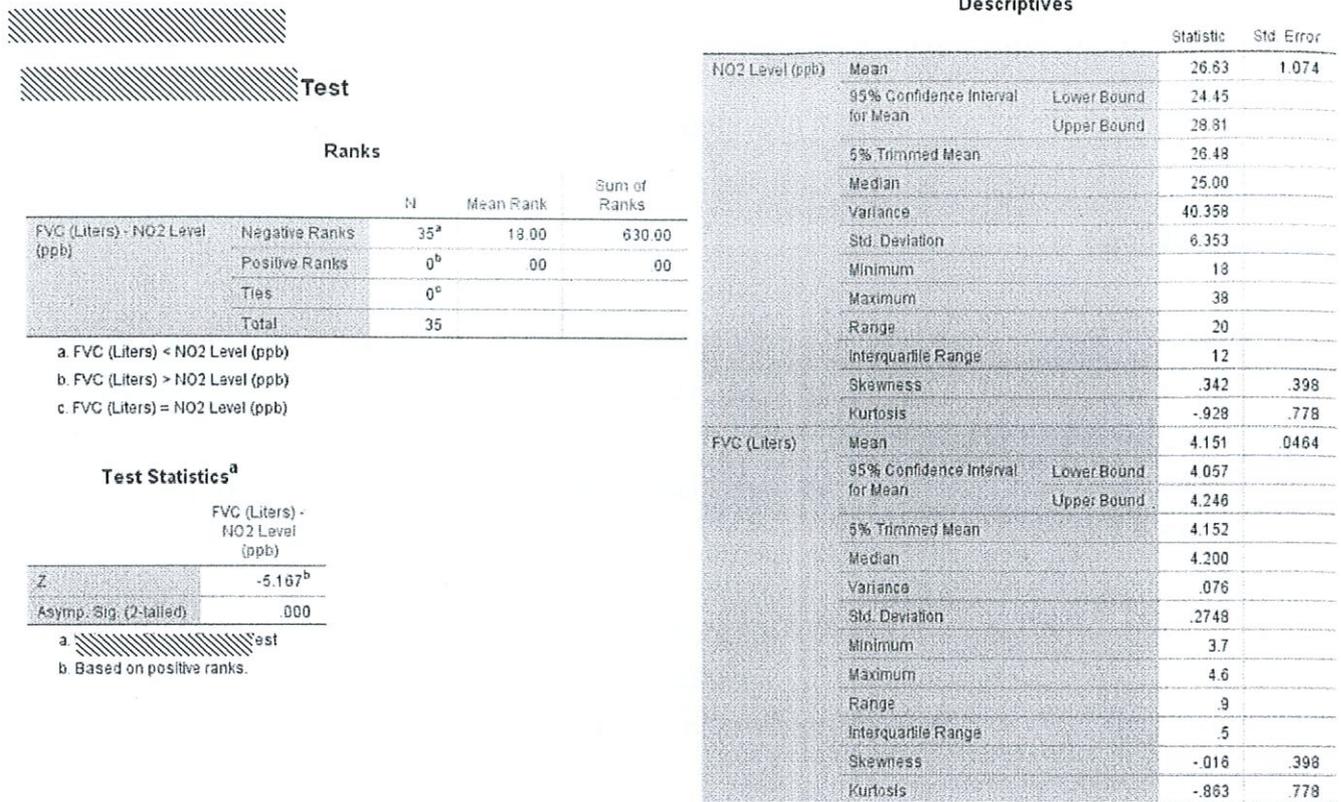


Figure 2: B

- Assuming the data is not normally distributed. From the SPSS output above, select the correct analysis test used to analyse the above situation. Provide justification of the test selection and state the null hypothesis for the analysis. (3 marks)
- Identify the dependent and independent variables used in this analysis. (1 mark)
- Using the SPSS Output, identify the relevant findings and construct a reporting table to summarize the findings. Provide clear interpretation of the results. (4 marks)
- Determine if there is a statistically significant association between the concentration of nitrogen dioxide (NO₂) emitted from industrial plants and lung function among residents in Rome, Italy. Justify your determination. (2 marks)

SECTION B (Total: 60 marks)

Answer only **THREE (3)** questions.

Please use the answer booklet provided.

Question 1

An FYP student conducted a survey to determine the cholesterol level (mg/dL) among school teachers with hypertension in two Malaysian states (Selangor and Johor). The results are summarized in Table below.

Refer below – Table 3: Summary of cholesterol levels among school teachers with hypertension in two states in Malaysia

Table 3: Summary of cholesterol levels among school teachers with hypertension in two states in Malaysia

State	Number of screened teachers	Mean	Standard deviation
Selangor	200	174.6	30.5
Johor	150	204.1	28.7

- What is the probability that the difference of mean cholesterol levels of school teachers, $\bar{x}_1 - \bar{x}_2$ between this two universities is more than 40.0 (mg/dL)?
(6 marks)
- Construct 95% Confident Interval for the difference between population means.
(7 marks)
- Construct 90% Confident Interval for the difference between population means.
(7 marks)

Question 2

A research team investigated the prevalence of anemia among pregnant women in Selangor. They collected blood samples from 400 women attending antenatal clinics and measured their hemoglobin levels (g/dL). The analysis revealed an average blood hemoglobin concentration of 12.8 g/dL with a standard deviation of 1.5 g/dL.

- a) Calculate the percentage of pregnant women with hemoglobin level of 10 g/dL and lower.
(4 marks)

- b) Calculate the percentage of pregnant women with hemoglobin levels 13 to 17 g/dL.
(9 marks)

- c) Determine the central limits for 90 percent of hemoglobin levels present.
(7 marks)

Question 3

Answer all the following questions.

- a) A screening for anemia among blood donors was conducted in the Blood Bank. Anemia refers to a condition where red blood cell count or hemoglobin concentration is below the normal range. The survey aimed to determine the association between type of occupation (professional, managerial & technical work) and anemia status among blood donors. The collected data were analysed using SPSS and are presented in figures below.

Refer below – Figure 3: C & Figure 4: D

Case Processing Summary

	Valid		Cases Missing		Total	
	N	Percent	N	Percent	N	Percent
Occupation * Anemia_Status	75	100.0%	0	0.0%	75	100.0%

Occupation * Anemia_Status

			Anemia_Status		Total
			No	Yes	
Occupation	Professional	Count	18	6	24
		% within Occupation	75.0%	25.0%	100.0%
	Managerial	Count	7	13	20
		% within Occupation	35.0%	65.0%	100.0%
	Technical work	Count	12	19	31
		% within Occupation	38.7%	61.3%	100.0%
Total		Count	37	38	75
		% within Occupation	49.3%	50.7%	100.0%

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	9.369 ^a	2	.009
Likelihood Ratio	9.688	2	.008
Linear-by-Linear Association	6.491	1	.011
N of Valid Cases	75		

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 9.87.

Figure 3: C

			Occupation	Anemia_Status	
⇒	Spearman's rho	Occupation	Correlation Coefficient	1.000	.289*
			Sig. (2-tailed)	.	.012
			N	75	75
	Anemia_Status	Correlation Coefficient	.289*	1.000	
			Sig. (2-tailed)	.012	.
			N	75	75

*. Correlation is significant at the 0.05 level (2-tailed).

Figure 4: D

- i. Assuming the data is not normally distributed. From the SPSS Outputs above, select the correct analysis test used to analyse the above situation. Provide justification of the test selection and state the null hypothesis for the analysis.
(3 marks)
- ii. Identify the dependent and independent variables used in this data analysis.
(1 marks)
- iii. Using the SPSS Output, identify the relevant findings and construct a reporting table to summarize the findings. Provide clear interpretation of the results.
(4 marks)
- iv. Determine if there is a statistically significant association between the different type of occupation and anemia status among blood donors. Justify your determination.
(2 marks)

b) A small survey was conducted by the occupational safety officer to determine the relationship of musculoskeletal disease disorder (MSD) and lifting heavy materials among 14 store workers in Factory CD. All 14 workers were sent for physical examination by the physician to confirm their health issues. The collected data were analysed using SPSS and are presented in the figures below.

Refer below – Figure 5: E

Case Processing Summary

	Valid		Cases Missing		Total	
	N	Percent	N	Percent	N	Percent
Lifting heavy materials * Musculoskeletal disease disorder (MSD)	14	100.0%	0	0.0%	14	100.0%

Lifting heavy materials * Musculoskeletal disease disorder (MSD)

			Musculoskeletal disease disorder (MSD)		Total
			No	Yes	
Lifting heavy materials	Everyday	Count	1	10	11
		% within Lifting heavy materials	9.1%	90.9%	100.0%
	Sometimes	Count	3	0	3
		% within Lifting heavy materials	100.0%	0.0%	100.0%
Total		Count	4	10	14
		% within Lifting heavy materials	28.6%	71.4%	100.0%

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	9.545 ^a	1	.002		
Continuity Correction ^b	5.611	1	.018		
Likelihood Ratio	10.050	1	.002		
Fisher's Exact Test				.011	.011
Linear-by-Linear Association	8.864	1	.003		
N of Valid Cases	14				

a. 3 cells (75.0%) have expected count less than 5. The minimum expected count is .86.

b. Computed only for a 2x2 table

Figure 5: E

- i. Identify the correct analysis used to analyse the above situation. Provide justification of the test selection and state the null hypothesis for the analysis.
(3 marks)
- ii. Identify the dependent and independent variables used in this analysis.
(1 mark)
- v. Using the SPSS Output, identify the relevant findings and construct a reporting table to summarize the findings. Provide clear interpretation of the results.
(4 marks)
- iii. Determine if there is a statistically significant relationship of musculoskeletal disease disorder (MSD) and lifting heavy items among 14 store workers in Factory CD. Justify your determination.
(2 marks)

Question 4

a) A comparative study was conducted to determine the fungal counts (CFU) levels in hospital wards situated on Level 3 and Level 4. Measurements were taken for 7 days using aerosol biosamplers in both wards. The collected data were analysed using SPSS and are presented in figure below.

Refer below – Figure 6: F

Group Statistics

	Level	N	Mean	Std. Deviation	Std. Error Mean
Fungi Count (CFU)	Level 3	20	35.15	6.418	1.435
	Level 4	24	50.93	16.499	3.366

Test

		Levene's Test for Equality of Variances		t-test for Equality of Means			95% Confidence Interval of the Difference			
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Fungi Count (CFU)	Equal variances assumed	24.597	.000	-5.533	42	.000	-21.683	3.919	-29.592	-13.775
	Equal variances not assumed			-5.926	30.887	.000	-21.683	3.659	-29.147	-14.220

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.649 ^a	.422	.408	12.943

a. Predictors: (Constant), Level

ANOVA^a

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	5129.094	1	5129.094	30.618	.000 ^b
	Residual	7035.883	42	167.521		
	Total	12164.977	43			

a. Dependent Variable: Fungi Count (CFU)

b. Predictors: (Constant), Level

ANOVA^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	13.467	6.363		2.116	.040
	Level	21.683	3.919	.649	5.533	.000

a. Dependent Variable: Fungi Count (CFU)

Figure 6: F

- i. Assuming the data is normally distributed. From the SPSS Outputs above, select the correct analysis test used to analyse the above situation. Provide justification of the test selection and state the null hypothesis for the analysis.
(3 marks)
- ii. Using the SPSS Output, identify the relevant findings and construct a reporting table to summarize the findings. Provide clear interpretation of the results.
(5 marks)
- iii. Determine if there is a statistically significant difference in fungal counts (CFU) between wards on Level 3 and Level 4. Justify your determination.
(2 marks)

b) An environmental consultant company was appointed to determine the levels of lead (Pb) in the well water in the Gua Musang area, Kelantan. During the investigation, it was confirmed that this well water was being used for consumption by the villagers. The investigation also involved the measurement of Pb in the villagers' hair samples to assess the health effects of consuming the contaminated water. The collected data were analysed using SPSS and are presented in figures below.

Refer below – Figure 7: G

			Lead (ppb) in Water	Lead (µg/g) in Nails
Lead (ppb) In Water	Correlation Coefficient		1.000	.844**
	Sig. (2-tailed)		.	.000
	N		50	50
Lead (µg/g) in Nails	Correlation Coefficient		.844**	1.000
	Sig. (2-tailed)		.000	.
	N		50	50

Ranks

		N	Mean Rank	Sum of Ranks
Lead (µg/g) in Nails - Lead (ppb) in Water	Negative Ranks	50 ^a	25.50	1275.00
	Positive Ranks	0 ^b	.00	.00
	Ties	0 ^c		
	Total	50		

a. Lead (µg/g) in Nails < Lead (ppb) in Water

b. Lead (µg/g) in Nails > Lead (ppb) in Water

c. Lead (µg/g) in Nails = Lead (ppb) in Water

Test Statistics^a

	Lead (µg/g) in Nails - Lead (ppb) in Water
Z	-6.154 ^b
Asymp. Sig. (2-tailed)	.000

Descriptives			Statistic	Std. Error
Lead (ppb) in Water	Mean		25.50	2.062
	95% Confidence Interval for Mean	Lower Bound	21.36	
		Upper Bound	29.64	
	5% Trimmed Mean		25.50	
	Median		25.50	
	Variance		212.500	
	Std. Deviation		14.577	
	Minimum		1	
	Maximum		50	
	Range		49	
	Interquartile Range		26	
	Skewness		.000	.337
	Kurtosis		-1.200	.662
	Lead (µg/g) in Nails	Mean		4.854
95% Confidence Interval for Mean		Lower Bound	4.142	
		Upper Bound	5.566	
5% Trimmed Mean			4.813	
Median			4.700	
Variance			6.269	
Std. Deviation			2.5038	
Minimum			.5	
Maximum			10.1	
Range			9.6	
Interquartile Range			3.5	
Skewness			.155	.337
Kurtosis			-.640	.662

Figure 7: G

- i. Assuming the data is not normally distributed. From the SPSS output, select the correct analysis used to analyse the above situation. Provide justification of the test selection and state the null hypothesis for the analysis. (3 marks)
- ii. Identify the dependent and independent variables used in this analysis. (1 mark)
- iii. Using the SPSS Output, identify the relevant findings and construct a reporting table to summarize the findings. Provide clear interpretation of the results. (4 marks)
- iv. Determine if there are a statistical effects of lead (Pb) exposure from consumed well water on the development of neurological effects among residents living in Gua Musang area, Kelantan. Justify your determination. (2 marks)

END OF EXAMINATION PAPER

