اونيۏرسيتي مليسيا ڤهڠ السلطان عبد الله
UNIVERSITI MALAYSIA PAHANG
AL-SULTAN ABDULLAH

CENTRE FOR MATHEMATICAL SCIENCES
EXAMINER ANSWER SCRIPT (FINAL EXAM)
SEMESTER: I SESSION: 2023/2024

QUESTION 1	Marks	Remark
STEP 1: Formulate the hypothesis $H_0 : \sigma^2 \le 15$ $H_1 : \sigma^2 > 15$ (Claims) **STEP 2:** Calculate the test statistics $\chi^2_{test} = \dfrac{(n-1)s^2}{\sigma_0^2}$ $\chi^2_{test} = \dfrac{(13-1)(6.1442)^2}{15} = 30.2010$ **STEP 3:** Determine the critical value		

$\chi^2_{0.01,12} = 26.2170$ **STEP 4:** Make a decision Since $\left(\chi^2_{test} = 30.2010\right) > \left(\chi^2_{0.01,12} = 26.2170\right)$, reject H_0. **STEP 5:** Draws a conclusion At $\alpha = 0.01$, there is sufficient evidence to support that the population variance for Indian Mackerel caught on rainy days is exceeded than 15.		
TOTAL	**8 Marks**	

QUESTION 2		Marks	Remarks
i)	Let **1** symbolized normal air, while **2** symbolized enriched air. **STEP 1:** Formulate the hypothesis $H_0 : \sigma_1^2 = \sigma_2^2$ (Claims) $H_1 : \sigma_1^2 \neq \sigma_2^2$ **STEP 2:** Calculate the test statistics $f_{test} = \dfrac{s_1^2}{s_2^2}$ $f_{test} = \dfrac{0.9143}{2.5915} = 0.3528$ **STEP 3:** Determine the critical value $f_{0.01,11,7} = 6.5382$; $f_{0.99,11,7} = \dfrac{1}{f_{0.01,7,11}} = \dfrac{1}{4.8861} = 0.2047$ **STEP 4:** Make a decision Since $\left(f_{0.99,11,7} = 0.2047\right) < \left(f_{test} = 0.3528\right) < \left(f_{0.01,11,7} = 6.5382\right)$, failed to reject H_0.		

	STEP 5: Draws a conclusion At $\alpha = 0.02$, there is sufficient evidence to support that no statistical significance of the difference in variability of the population plants' weights.		
ii)	**STEP 1:** Formulate the hypothesis $H_0 : \mu_1 \geq \mu_2$ $H_1 : \mu_1 < \mu_2$ (Claims) **STEP 2:** Determine the P-value P-value = 0.0583 **STEP 3:** Make a decision Since (P-value = 0.0583) > $(\alpha = 0.02)$, failed to reject H_0. **STEP 4:** Draws a conclusion At $\alpha = 0.02$, there is no sufficient evidence to support that a CO_2-enriched atmosphere increases the average of plant growth.		
	TOTAL	**13 Marks**	

	QUESTION 3	Marks	Remark
i)	Plant growth / height of each plant		
ii)	$X = 56.0257/32 = 1.7508$ $Y = 40.3542/1.7508 = 23.0490$		
iii)	**Water Frequency effect:** H_0: There is no effect of water frequency on the plant growth H_1: There is an effect of water frequency on the plant growth $p - value = 0.9760$ or $f_{test} = 0.0009$ Since $(p - value = 0.9760) > (\alpha = 0.05)$ or $(f_{test} = 0.0009) < (Fcrit = 4.1491)$, the decision is do not reject H_0: At $\alpha = 0.05$, there is no effect of water frequency on the plant growth **Exposure to sunlight effect:** H_0: There is no effect of exposure to sunlight on the plant growth		

H_1: There is an effect of exposure to sunlight on the plant growth $p - value = 0.0000$ or $f_{test} = 23.0490$ Since $(p - value = 0.0000) < (\alpha = 0.05)$ or $(f_{test} = 23.0490) > (Fcrit = 2.9011)$, the decision is reject H_0: At $\alpha = 0.05$, there is an effect of exposure to sunlight on the plant growth		
TOTAL	**15 Marks**	

QUESTION 4	**Marks**	**Remark**
i) Independent variable: temperature Dependent variable: life time of the device ii) There is a negative linear correlation between the variables. iii) $\hat{\beta}_1 = \frac{-31880}{6000} = -5.3133$ $\hat{y} = 453.5556 - 5.3133x$ iv) The life time of the device decrease by 5 hours for each increasing in temperature. v) $H_0: \beta_1 = 0$ (no linear relationship) $H_1: \beta_1 \neq 0$ (has linear relationship) $S_{yy} = 489857 - \frac{1691^2}{9} = 172136.8889$ $MS_{Res} = \frac{172136.8889 - (-5.3133)(-31880)}{7} = 392.6978$ $t_{test} = \frac{-5.3133}{\sqrt{392.6978\left(\frac{1}{6000}\right)}} = -20.7688$ $t_{0.01,7} = 2.9980$ Since $(t_{test} = -20.7688) < (-t_{0.01,7} = -2.9980)$, the decision is reject H_0 At $\alpha = 0.02$, there is a linear relationship between the range of temperature and the life time of the device		
TOTAL	**16 Marks**	

QUESTION 5		Marks	Remark
(i)	Simple linear regression involves only one independent variable while multiple linear regression involves more than one independent variables.		
(ii)	Salary		
(iii)	H_0: $\beta_1 = \beta_2 = 0$ H_1: $\beta_j \neq 0$ for at least one $j = 1,2$ $p = 0.0000$ $(p = 0.0000) < 0.025 \rightarrow$ reject H_0 Thus, at least one of the independent variables has a significant relationship with the dependent variable at 2.5% significance level.		
(iv)	0.8147 81.47% of the variation in salaries can be predicted by the number of years of experience of the programmer and aptitude test score.		
(v)	<u>For β_1:</u> $p = 0.0000$ $(p = 0.0000) < 0.025 \rightarrow$ reject H_0 Thus, x_1 is a significant predictor.		

For β_2:

$p = 0.0048$

$(p = 0.0048) < 0.025 \rightarrow$ reject H_0

Thus, x_2 is a significant predictor.

(vi) $\beta_0 = 3.1739$

The estimated salary is RM3173.90 when experience is 0 year and aptitude test score also 0.

$\beta_2 = 0.2509$

When number of years of experience of the programmer is held constant, the estimated salary increases by RM250.90 for every aptitude test score.

(vii)

Predictor	***P*-value**	***r***	***r*²**	**Adjusted *r*²**	**Regression model**
x_1	0.0000	0.8553	0.7316	0.7167	$\hat{y} = 22.8111 + 1.6200\, x_1$
x_2	0.0063	0.5887	0.3466	0.3103	$\hat{y} = -4.7097 + 0.4344\, x_2$
x_1, x_2	0.0000	0.9133	0.8342	0.8147	$\hat{y} = 3.1739 + 1.4039\, x_1 + 0.2509\, x_2$

(viii) The best model: $\hat{y} = 3.1739 + 1.4039x_1 + 0.2509\, x_2$

Because the model has (p-value $= 0.0000) < (\alpha = 0.05)$ with high adjusted $r^2 = 0.8147$.

(ix) $\hat{y} = 3.1739 + 1.4039x_1 + 0.2509\, x_2$

$40 = 3.1739 + 1.4039x_1 + 0.2509(90)$

$40 = 1.4039x_1 + 25.7549$

$1.4039x_1 = 14.2451$

$x_1 = 10.1468$ years ≈ 10 years		
TOTAL	**23 Marks**	

QUESTION 6	**Marks**	**Remark**
i) A + B = 100-50 A+ B = 50 A = 25, B=25 ii) H_o: The proportions of number of motorcycle accidents do not differ significantly from those with the stated ratio H_1: The proportions of number of motorcycle accidents differ significantly from those with the stated ratio		

O_i	P_i	E_i	$\frac{(O_i - E_i)^2}{E_i}$
30	0.4	0.4(100) = 40	2.5
25	0.3	0.3(100) = 30	0.8333
25	0.2	0.2(100) = 20	1.25
20	0.1	0.1(100) = 10	10
			$\chi^2_{test} = 14.5833$

Since $(\chi^2_{test} = 14.5833) > (\chi^2_{0.005,3} = 12.8382)$, then we reject H_o At α = 0.005, we can conclude that the proportions of number of motorcycle accidents differ significantly from those with the stated ratio. iii) Not a good fit. .		
TOTAL	**12 Marks**	

QUESTION 7	**Marks**	**Remark**

a)

Contingency table	**Linear Regression**
It cannot be used for prediction	It can be used for prediction
Determine the relationship between two categorical variables	Determine the strength of a relationship between dependent and independent variables
It treats neither variable as independent nor dependent	It treats one variable as dependent

b)

W = 90 – 10 – 30
= 50

ii)
H_o: The sport preference and age of students are independent
H_1: The sport preference and age of students are dependent

O_{ij}	$E_{ij} = \frac{n_i \times n_j}{n_{..}}$	$\frac{(O_{ij}-E_{ij})^2}{E_{ij}}$
$O_{11} = 40$	21.875	15.0179
$O_{12} = 10$	26.25	10.0595
$O_{13} = 20$	21.875	0.1607

$O_{21} = 10$	28.125	11.6806
$O_{22} = 50$	33.75	7.8241
$O_{23} = 30$	28.125	0.1250
		$\chi^2_{test} = 44.8677$

Since $(\chi^2_{test} = 44.8677) > (\chi^2_{0.01,2} = 9.2103)$, then we reject H_o

At α = 0.01, we can conclude that the sport preference and age of students are dependent.

TOTAL	**13 Marks**	