Universiti Tun Hussein Onn Malaysia

UNIVERSITI TUN HUSSEIN ONN MALAYSIA

FINAL EXAMINATION
SEMESTER I
SESSION 2022/2023

COURSE NAME	:	STATISTICS
COURSE CODE	:	DAC 21302
PROGRAMME CODE	:	DAA
EXAMINATION DATE	:	FEBRUARY 2023
DURATION	:	2 HOURS AND 30 MINUTES
INSTRUCTIONS	:	1. ANSWER ALL QUESTIONS
		2. THIS FINAL EXAMINATION IS CONDUCTED VIA **CLOSED BOOK**
		3. STUDENTS ARE **PROHIBITED** TO CONSULT THEIR OWN MATERIAL OR ANY EXTERNAL RESOURCES DURING THE EXAMINATION CONDUCTED VIA CLOSED BOOK

THIS QUESTION PAPER CONSISTS OF **SIX (6)** PAGES

Q1 **Table Q1** contains data on blackouts in a new building constructed by your company in minutes for 50 days.

Table Q1

Class limit	f
46 – 57	7
58 – 69	3
70 – 81	5
82 – 93	14
94 – 105	12
106 – 117	7
118 – 129	2

(a) If x is the midpoint, construct a table that contains lower boundary, cumulative frequency, midpoint, x^2, $f_i x_i$, $f_i x_i^2$, $\sum f$, $\sum f_i x_i$, and $\sum f_i x_i^2$.

(8 marks)

(b) Find the mean, median, mode, variance and standard deviation.

(12 marks)

Q2 (a) The owner of the café on the ground floor of UTHM Pagoh surveyed 50 customers if they like tea, coffee, or chocolate drinks. 31 customers like chocolate drinks, 34 like tea, and 21 like coffee and tea. Seven like tea and chocolate but not coffee, while two of them like coffee and chocolate but not tea. 17 of them like all three drinks. However, seven of them like coffee only.

(i) Display all the information on a Venn diagram.

(3 marks)

(ii) Calculate the probability of customers who liked only chocolate drink.

(1 mark)

(iii) Two of the respondents were chosen at random. Determine the probability that both of them like coffee.

(3 marks)

(iv) Find the probability that the customer like tea given that they did not like coffee.

(3 marks)

(b) A continuous random variable X, has a probability density function given by the following:

$$f(x) = \begin{cases} cxe^{x} & ; \quad 0 \le x \le 1 \\ 0 & ; \quad \text{otherwise} \end{cases}.$$

(i) Determine the value of c.

(4 marks)

(ii) Solve for $P(0.45 \leq X \leq 0.8)$.

(3 marks)

(iii) Evaluate $E(2X+3)$, given that $E(X) = 0.7187$.

(3 marks)

Q3 (a) A factory is producing cookware that are in circular shape. A sample of cookware is taken, and the diameters are 5, 6, 7, 8.5, 10 and 15 centimeters. Estimate a 99% confidence interval for the mean diameter of cookware, assuming an approximate normal distribution.

(8 marks)

(b) Two independents sampling stations were chosen for an investigation of index chemical acids pollution in rivers of Malaysia. **Table Q3 (b)** shows the data, recorded in months, represent the monthly samples collected at different stations.

Table Q3 (b)

First Station	Second Station
$n_1 = 36$	$n_2 = 31$
$\overline{x_1} = 73.44$	$\overline{x_2} = 96.41$
$s_1^2 = 0.201$	$s_2^2 = 0.594$

Find a 90% confidence interval for the difference between the population means for the two stations. Assume that the population are approximately normal distributed.

(12 marks)

Q4 (a) A bus company advertised a mean time of 120 minutes for a trip between two cities. A consumer group had reason to believe that the mean time was more than 120 minutes. A sample of 40 trips showed a mean $\bar{x} = 125$ minutes and a standard deviation $s = 8.5$ minutes. At the 5% level of significance, test the consumer group's belief.

(10 marks)

(b) In a mathematics competition in secondary school, the mean score of 45 boys was 79 with a standard deviation of 8, while the mean score of 55 girls was 72 with standard deviation 7. Test the hypothesis testing at 1% level significance that the boys are performed better than the girls.

(10 marks)

Q5 You are studying employee salary movements by years of experience. **Table Q5** shows the employees data in your company.

Table Q5

Years of experience, x	1.5	1.1	2	2.2	1.3	3.2	5	4.5
Monthly salary, y	3750	3680	3900	3945	3700	4190	5590	4645

(a) Calculate S_{xx}, S_{yy} and S_{xy}.

(9 marks)

(b) Determine and interpret the sample correlation coefficient, r.

(3 marks)

(c) Evaluate $\hat{\beta}_1$ and $\hat{\beta}_0$.

(4 marks)

(d) Find the estimated regression line, $\hat{y}$.

(2 marks)

(e) Solve for the value of $\hat{y}$ if $x = 8$.

(2 marks)

– END OF QUESTIONS –

FINAL EXAMINATION

SEMESTER / SESSION : SEM I 2022/2023
COURSE NAME : STATISTICS

PROGRAMME CODE : DAA
COURSE CODE : DAC21302

Formula

$$S_{xy}=\sum x_i y_i-\frac{\sum x_i \sum y_i}{n},\ S_{xx}=\sum x_i^{\ 2}-\frac{\left(\sum x_i\right)^2}{n}, S_{yy}=\sum y_i^{\ 2}-\frac{\left(\sum y_i\right)^2}{n},$$

$$r=\frac{S_{xy}}{\sqrt{S_{xx}S_{yy}}},\quad \hat{\beta}_1=\frac{S_{xy}}{S_{xx}},\ \hat{\beta}_0=\bar{y}-\hat{\beta}_1\bar{x},$$

$$Z=\frac{\bar{x}-\mu}{\sigma/\sqrt{n}}, Z=\frac{\bar{x}-\mu}{s/\sqrt{n}},\ T=\frac{\bar{x}-\mu}{s/\sqrt{n}}$$

$$Z=\frac{(\bar{X}_1-\bar{X}_2)-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}, Z=\frac{(\bar{X}_1-\bar{X}_2)-(\mu_1-\mu_2)}{\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}}$$

$$\bar{x}=\frac{\sum f_i x_i}{\sum f_i},\ M=L_M+C\times\left(\frac{n/2-F}{f_m}\right),\ M_0=L+C\times\left(\frac{d_b}{d_b+d_a}\right)$$

$$s^2=\frac{1}{\Sigma f-1}\left[\sum_{i=1}^{n} f_i x_i^{\ 2}-\frac{(\Sigma f_i x_i)^2}{\Sigma f}\right]$$

$$\sum_{i=-\infty}^{\infty}p(x_i)=1,\ E(X)=\sum_{\forall x}xp(x),\ \int_{-\infty}^{\infty}f(x)\ dx=1,\ E(X)=\int_{-\infty}^{\infty}xp(x)\ dx,\ Var(X)=E(X^2)-[E(X)]^2,$$

$$P(x)=\binom{n}{x}\cdot p^x\cdot(1-p)^{n-x}\quad x=0,1,\ldots,n,\ P(X=r)=\frac{e^{-\mu}\cdot\mu^r}{r!}\quad r=0,1,\ldots,\infty,$$

$X\sim N(\mu,\sigma^2)$, $Z\sim N(0,1)$ and $Z=\dfrac{X-\mu}{\sigma}$

FINAL EXAMINATION

SEMESTER / SESSION : SEM I 2022/2023
COURSE NAME : STATISTICS

PROGRAMME CODE : DAA
COURSE CODE : DAC21302

$$\overline{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right), \quad Z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1), \quad \overline{X}_1 - \overline{X}_2 \sim N\left(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right),$$

$$\overline{x} - z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right) < \mu < \overline{x} + z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right)$$

$$\overline{x} - z_{\alpha/2}\left(\frac{s}{\sqrt{n}}\right) < \mu < \overline{x} + z_{\alpha/2}\left(\frac{s}{\sqrt{n}}\right)$$

$$\overline{x} - t_{\alpha/2,\,v}\left(\frac{s}{\sqrt{n}}\right) < \mu < \overline{x} + t_{\alpha/2,\,v}\left(\frac{s}{\sqrt{n}}\right), \; v = n - 1.$$

$$(\overline{x}_1 - \overline{x}_2) - z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} < \mu_1 - \mu_2 < (\overline{x}_1 - \overline{x}_2) + z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}},$$

$$(\overline{x}_1 - \overline{x}_2) - z_{\alpha/2}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} < \mu_1 - \mu_2 < (\overline{x}_1 - \overline{x}_2) + z_{\alpha/2}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}},$$

$$(\overline{x}_1 - \overline{x}_2) - t_{\alpha/2,\,v}\, S_P\sqrt{\frac{1}{n_1} + \frac{1}{n_2}} < \mu_1 - \mu_2 < (\overline{x}_1 - \overline{x}_2) + t_{\alpha/2,\,v}\, S_P\sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \text{ where}$$

$$S_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} \text{ and } v = n_1 + n_2 - 2,$$

$$(\overline{x}_1 - \overline{x}_2) - t_{\alpha/2,\,v}\sqrt{\frac{1}{n}\left(s_1^2 + s_2^2\right)} < \mu_1 - \mu_2 < (\overline{x}_1 - \overline{x}_2) + t_{\alpha/2,\,v}\sqrt{\frac{1}{n}\left(s_1^2 + s_2^2\right)} \text{ and } v = 2(n-1),$$

$$(\overline{x}_1 - \overline{x}_2) - t_{\alpha/2,\,v}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} < \mu_1 - \mu_2 < (\overline{x}_1 - \overline{x}_2) + t_{\alpha/2,\,v}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} \text{ and } v = \frac{\left(s_1^2/n_1 + s_2^2/n_2\right)^2}{\frac{\left(s_1^2/n_1\right)^2}{n_1 - 1} + \frac{\left(s_2^2/n_2\right)^2}{n_2 - 1}}$$