CONFIDENTIAL

UNIVERSITI TUN HUSSEIN ONN MALAYSIA

FINAL EXAMINATION
SEMESTER II
SESSION 2022/2023

COURSE NAME	:	STATISTICS
COURSE CODE	:	DAE 23602
PROGRAMME CODE	:	DAE
EXAMINATION DATE	:	JULY / AUGUST 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 **SEVEN (7)** PAGES

CONFIDENTIAL

Q1 A manager wishes to study the relationship between the cost oil expenditure (Y) and car miles (X) for their taxi company over 22 time periods. The data is shown in **Table Q1**.

(a) Find S_{xx}, S_{yy}, S_{xy}, $\widehat{\beta}_0$ and $\widehat{\beta}_1$ (up to 3 decimal places) to fit the regression line, $\hat{y}$.

(15 marks)

(b) Calculate the coefficient of determination, R^2.

(4 marks)

(c) Write a conclusion based on your answer in **Q1(a)** and **Q1(b)**.

(1 mark)

Q2 (a) Farid has two bags that contain green and blue balls. The first bag contains 4 green balls and 3 blue balls, and the second bag contains 3 green balls and 5 blue balls. Farid took one ball from the first bag and placed it unseen in the second bag. Then Farid asks his daughter to draw a ball from the second bag.

(i) Construct a tree diagram based on the above situation.

(4 marks)

(ii) Determine the probability that a ball now drawn from the second bag is blue.

(3 marks)

(iii) By multiplicative rule, find the probability of picking the same colour balls from the first bag and the second bag?

(3 marks)

(b) The shelf life for packets of biscuits in days is a random variable having the distribution function of the following:

$$f(x)=\begin{cases}\dfrac{5000}{(x+50)^3} & ; \quad x>0\\ 0 & ; \quad \text{elsewhere}\end{cases}.$$

Find the probability that the biscuits will have a shelf life of:

(i) at least 50 days.

(5 marks)

(ii) anywhere from 20 to 30 days.

(5 marks)

Q3 (a) Given a test that is normally distributed with a mean of 100 and a standard deviation of 12. Calculate the probability that a sample of:

(i) 25 scores will have a mean greater than 106.

(5 marks)

(ii) 16 scores will have a mean between 95 and 105.

(5 marks)

(b) The mean GPA for students in School A is 3.0 and the mean GPA for students in School B is 2.8. The standard deviation in both schools is 0.25. The GPAs of both schools are normally distributed. If 9 students are randomly sampled from each school, what is the probability that the sample mean for:

(i) School A will exceed that of School B by 0.5 or more.

(5 marks)

(ii) School B will be greater than the sample mean for School A.

(5 marks)

Q4 (a) A random sample of 10 cocoa bars of a certain brand has, on average, 215 calories with a sample standard deviation of 19 calories. Assume that the distribution of the calories is approximately normal.

(i) Calculate the point estimate of μ.

(1 mark)

(ii) Construct a 95% confidence interval for the true mean calorie content of this brand of cocoa bar.

(9 marks)

(b) Two kinds of thread are being compared for strength. Thirty-five pieces of each type of thread are tested under similar conditions. Brand *A* had an average tensile strength of 89.5 kilograms with a standard deviation of 6.7 kilograms, while brand *B* had an average tensile strength of 87.3 kilograms with a standard deviation of 6.2 kilograms. Construct a 98% confidence interval for the difference of the population means.

(10 marks)

Q5 (a) Each packet of snacks must weigh 100g. Omar randomly selected 50 packets and found that the mean weight is 104g and the standard deviation is 3.5g. Assume the population is distributed approximately normal. Test at a 5% significance level whether the mean weight per packet is more than 100g.

(10 marks)

(b) The mean lifetime of 35 batteries produced by Company Pagoh is 50 hours and the mean lifetime of 38 batteries produced by Company Muar is 48 hours. If the population standard deviation of all batteries produced by Company Pagoh is 3.5 hours and the population standard deviation of all batteries produced Company Muar is 3.8 hours, test at 1% significance level that the mean lifetime of batteries produced by Company Pagoh is better than the mean lifetime of Company Muar. Assume the data was taken from a normal distribution.

(10 marks)

– END OF QUESTIONS –

FINAL EXAMINATION

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

PROGRAMME CODE : DAE
COURSE CODE : DAE23602

Table Q1

Period	Cost	Miles	Period	Cost	Miles
1	2.139	3.147	12	2.027	3.141
2	2.126	3.16	13	1.985	2.928
3	2.153	3.197	14	1.956	3.063
4	2.153	3.173	15	2.004	3.096
5	2.154	3.292	16	2.001	3.096
6	2.282	3.561	17	2.015	3.158
7	2.456	4.013	18	2.132	3.338
8	2.599	4.244	19	2.195	3.492
9	2.509	4.159	20	2.437	4.019
10	2.345	3.776	21	2.623	4.394
11	2.059	3.232	22	2.523	4.251

FINAL EXAMINATION

SEMESTER / SESSION : SEM II 2022/2023
COURSE NAME : STATISTICS
PROGRAMME CODE : DAE
COURSE CODE : DAE23602

Formula

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

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

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

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

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

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

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

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

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

FINAL EXAMINATION

SEMESTER / SESSION : SEM II 2022/2023
COURSE NAME : STATISTICS
PROGRAMME CODE : DAE
COURSE CODE : DAE23602

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

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

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

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

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

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

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

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

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

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