UTHM
Universiti Tun Hussein Onn Malaysia

UNIVERSITI TUN HUSSEIN ONN MALAYSIA

FINAL EXAMINATION
SEMESTER II
SESSION 2021/2022

COURSE NAME	:	STATISTICS
COURSE CODE	:	BIT 11603
PROGRAMME CODE	:	BIT
EXAMINATION DATE	:	JULY 2022
DURATION	:	3 HOURS
INSTRUCTION	:	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

Q1 (a) Classify each random variable below as either discrete or continuous.

(i) The number of boys in a randomly selected three-child family.
(1 mark)

(ii) The temperature of a cup of tea served at a restaurant.
(1 mark)

(iii) The number of no-shows for every reservation made with a commercial airline.
(1 mark)

(iv) The number of vehicles owned by a randomly selected household.
(1 mark)

(v) The average amount spent on electricity each February by a randomly selected household in a certain state.
(1 mark)

(b) Random variable X has a probability function define as below.

$$f(x)=\begin{cases} 0.2 & ; -1<x\le 0 \\ 0.2+cx & ; 0<x\le 1 \\ 0 & ; \text{elsewhere} \end{cases}$$

Calculate

(i) the value of c.
(3 marks)

(ii) the cumulative distribution function, $F(x)$.
(6 marks)

(iii) $P(-0.5<X<0.5)$
(3 marks)

(c) Ten percent of computer chips produces are defective. Compute the probability that in a random sample of 400 computer chips produced

(i) at most 30 chips will be defective.
(3 marks)

(ii) not less than 20 chips will be defective.
(2 marks)

(iii) between 35 and 45 chips will be defective.

(3 marks)

Q2 A group of nut packets has a mean weight of 5.02 grams and a standard deviation of 0.30 grams. A random sample of 100 nut packets are chosen from the group. Compute

(a) the probability that an average weight of nut packets chosen is between 4.96 and 5.00 grams.

(6 marks)

(b) the probability that an average weight of nut packets chosen is more than 5.10 grams.

(4 marks)

Q3 (a) As part of determining what to charge for its policies, an insurance company need to estimate the average cost of bodily injury claims when the policyholder is at fault. They would like to be 95% confident that the estimate is within ± RM200 and the standard deviation for bodily injury claims is around RM8000. Estimate the sample size should be used?

(4 marks)

(b) Two hundred randomly selected adults were timed as they filled out a particular form. The times required had mean 12.8 minutes with standard deviation 1.7 minutes. Construct a 90% confidence interval for the mean time taken for all adults to fill out this form.

(6 marks)

(c) A company claims that its medicine, Brand *Z*, provides faster relief from pain than another company's medicine, Brand *Y*. A researcher tested both brands of medicine on two groups randomly selected patients. The results of the test are given in **Table 1**. Assume that the two populations are normally distributed with equal standard deviations. Construct a 99% confidence interval for the difference between the mean relief times for the two brands of medicines.

Table 1: Medicine Testing Data

Brand	Sample size	Mean of relief times (Minutes)	Standard deviation of relief times (Minutes)
Z	25	44	11
Y	22	49	9

(10 marks)

Q4 (a) An Ice cream company claimed that its product contain on average 500 calories per pint.

(i) Test the claim if 24 pint containers were analyzed giving the mean is 507 calories and a standard deviation of 21 calories at 1% level of significance.
(6 marks)

(ii) Test the claim if 42 pint containers were analyzed giving the mean is 509 calories and a standard deviation of 18 calories at 1% level of significance.
(6 marks)

(b) The manufacturer of a certain brand of light bulbs claims that the variance of the lives of bulbs is 4200 square hours. A consumer agency took a random sample of 25 bulbs and tested them. The variance of the lives of bulbs was found to be 5200 square hours. Assume that the lives of all bulbs are approximately normally distributed. Test at the 5% significance level whether the variance of such bulbs is different from 4200 square hours.

(13 marks)

Q5 A company has recorded data on the weekly sales for its product and the unit price of the competitor's product. The data resulting from a random sample of 8 weeks as follows. The company is interested in knowing how their sales might be impacted by the competitor's price.

Week	Price	Sales
1	0.33	20
2	0.25	14
3	0.44	22
4	0.4	21
5	0.35	16
6	0.39	19
7	0.29	15
8	0.37	17

(a) Draw a scatter plot diagram of the data and explain if the graph indicates whether simple linear regression is appropriate for these data.

(3 marks)

(b) In conducting a simple linear regression, state what is the dependent variable and independent variable.

(2 marks)

(c) Construct the linear regression equation of the data.

(9 marks)

(d) Compute the R Square value. Interpret the R square value in relation to the linear regression model from **Q5(c)**.

(4 marks)

(e) Is it sufficient evidence to indicate that there is linear correlation between the competitor's price and the company sales?

(2 marks)

-END OF QUESTIONS –

FINAL EXAMINATION	
SEMESTER/ SESSION: SEM II / 20212022	**PROGRAMME CODE: BIT**
COURSE NAME: STATISTICS	**COURSE CODE: BIT 11603**

Formulae

Special Probability Distributions :

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

$$X\sim P_0(\mu),\ Z=\frac{X-\mu}{\sigma},\ Z\sim N(0,1),\ X\sim N(\mu,\sigma^2).$$

Sampling Distributions :

$$\overline{X}\sim N(\mu,\sigma^2/n),\ Z=\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}\sim N(0,1),\ T=\frac{\overline{x}-\mu}{s/\sqrt{n}},\ \overline{X}_1-X_2\sim N\left(\mu_1-\mu_2,\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}\right).$$

Estimations :

$$n=\left(\frac{Z_{\alpha/2}\cdot\sigma}{E}\right)^2,\quad \overline{x}\pm z_{\alpha/2}(\sigma/\sqrt{n}),\quad \overline{x}\pm z_{\alpha/2}(s/\sqrt{n}),\quad \overline{x}\pm t_{\alpha/2,v}\left(\frac{s}{\sqrt{n}}\right)$$

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

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

$$(\bar{x}_1-\bar{x}_2)-t_{\alpha/2,v}\cdot S_p\sqrt{\frac{2}{n}}<\mu_1-\mu_2<(\bar{x}_1-\bar{x}_2)+t_{\alpha/2,v}\cdot S_p\sqrt{\frac{2}{n}};v=2n-2$$

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

where Pooled estimate of variance, $S_p^2=\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}$ with $v=n_1+n_2-2$,

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

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

FINAL EXAMINATION	
SEMESTER/ SESSION: SEM II / 20212022	**PROGRAMME CODE: BIT**
COURSE NAME: STATISTICS	**COURSE CODE: BIT 11603**

Formulae

$$\frac{(n-1)\cdot s^2}{\chi^2_{\alpha/2,v}} < \sigma^2 < \frac{(n-1)\cdot s^2}{\chi^2_{1-\alpha/2,v}} \text{ with } v = n-1,$$

Hypothesis Testing :

$$Z_{Test} = \frac{\bar{x}-\mu}{\sigma/\sqrt{n}},\quad Z_{Test} = \frac{\bar{x}-\mu}{s/\sqrt{n}},\quad T_{Test} = \frac{\bar{x}-\mu}{s/\sqrt{n}}$$

$$Z = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}},\quad T = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{S_p \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \text{ with } v = n_1 + n_2 - 2,$$

$$Z = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}},\quad T = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{1}{n}\left(s_1^2 + s_2^2\right)}},\quad T = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \text{ with}$$

$$v = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2 - 1}}.\ ;\quad S_p^2 = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1 + n_2 - 2}\ ;\quad \chi^2 = \frac{(n-1)s^2}{\sigma^2}$$

$$F = \frac{s_1^2}{s_2^2},\quad \text{with } \frac{1}{f_{\alpha/2}(v_2, v_1)} \quad \text{and} \quad f_{\alpha/2}(v_1, v_2)$$

Simple Linear Regressions :

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

$$\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}},\quad \hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x},\quad \hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x,\quad r = \frac{S_{xy}}{\sqrt{S_{xx} \cdot S_{yy}}},\quad SSE = S_{yy} - \hat{\beta}_1 S_{xy},\quad MSE = \frac{SSE}{n-2}.$$