CONFIDENTIAL

Universiti Tun Hussein Onn Malaysia

UNIVERSITI TUN HUSSEIN ONN MALAYSIA

FINAL EXAMINATION
SEMESTER I
SESSION 2019/2020

COURSE NAME	:	STATISTICS
COURSE CODE	:	DAS 20502
PROGRAMME CODE	:	DAM
EXAMINATION DATE	:	DECEMBER 2019 / JANUARY 2020
DURATION	:	2 HOURS 30 MINUTES
INSTRUCTION	:	ANSWER **ALL** QUESTIONS IN SECTION A AND **THREE (3)** QUESTIONS IN SECTION B

THIS QUESTION PAPER CONSISTS OF **EIGHT (8)** PAGES

CONFIDENTIAL

SECTION A

Q1 A manager in sales department of Audi's cars wants to study the relationship between number of experiences with the total sales in a year of 7 randomly selected salespersons in Johor Bahru showroom. The following data is obtained:

Table Q1

Number of experiences (years), x	Total sales in a year in million (RM), y
9	82
12	86
5	43
6	74
5	58
15	90
8	78

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

(9 marks)

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

(3 marks)

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

(4 marks)

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

(2 marks)

(e) Estimate value of $\hat{y}$ if $x = 7.5$.

(2 marks)

Q2 (a) It is claimed that an automobile is driven on the average more than 20,000 kilometers per year. To test this claim, a random sample of 100 automobile owners are asked to keep a record of the kilometers they travel. Determine if the random sample showed an average of 23,500 kilometers and a standard deviation of 3900 kilometers. Test the hypothesis at 1% level of significance.

(10 marks)

(b) The mean lifetime of 40 power banks produced by Company Pening is 65 hours and the mean lifetime of 55 power banks produced by Company Lalat is 60 hours. If the standard deviation of all power banks produced by Company Pening is 3.5 hours and the standard deviation of all power banks produced Company Lalat is 4 hours. Test at 5% significance level that the mean lifetime of power banks produced by Company Pening is better than the mean lifetime of Company Lalat. Assume the data was taken from a normal distribution.

(10 marks)

SECTION B

Q3 The money spend by 80 students of DAS20502 in a month is recorded as in **Table Q3.**

Table Q3

Class limit	f
100 – 199	1
200 – 299	8
300 – 399	15
400 – 499	21
500 – 599	18
600 – 699	12
700 – 799	5

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

(8 marks)

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

(12 marks)

Q4 (a) **Table Q4** is a two-way contingency table that gives the breakdown of the population of adults in a particular locale according to employment type and level of life insurance.

Table Q4

Employment type	Level of insurance		
	Low	Medium	High
Unskilled	0.07	0.19	0.00
Semi-skilled	0.04	0.28	0.08
Skilled	0.03	0.18	0.05
Professional	0.01	0.05	0.02

An adult is selected at random. Hence, calculate the probability:

(i) that the person has a high level of life insurance.

(1 mark)

(ii) that the person has a high level of insurance, given that he does not have a professional position.

(3 marks)

(iii) of selecting the person has a medium level of life insurance or have a semi-skilled position.

(2 marks)

(iv) Determine whether or not the events "has a high level of life insurance" and "has a professional position" are independent.

(2 marks)

(b) The queueing time, X minutes of a customer at a cash register of a supermarket has the probability density function,

$$f(x)=\begin{cases}\frac{3}{32}x(k-x) & ,0\le x\le k\\ 0 & ,\text{otherwise}\end{cases}.$$

(i) Show that the value of k is 4.

(5 marks)

(ii) Solve for $E(X)$ and $Var(X)$.

(7 marks)

Q5 (a) A lawyer commutes daily from his suburban home to his midtown office. The average time for a one-way trip is 24 minutes with a standard deviation of 3.8 minutes. Assume the distribution of trip times is normally distributed.

(i) Find the probability that a trip will take at least $\frac{1}{2}$ hour.

(3 marks)

(ii) If the office opens at 9.00 am and he leaves his house at 8.45 am daily, find the percentage of the time that he is late for work.

(3 marks)

(iii) If he leaves the house at 8.35 am and coffee is served at the office from 8.50 am until 9.00 am, find the probability that he misses coffee.

(3 marks)

(iv) Find the length of time above which we find the slowest 15% of the trips.

(3 marks)

(b) Statistics released by the National Highway Traffic Safety Administration and the National Safety Council show that on an average weekend night, 1 out of every 10 drivers on the road is drunk. If 400 drivers are randomly checked next Saturday night, find the probability that the number of drunk drivers will be:

(i) less than 32.

(2 marks)

(ii) more than 49.

(2 marks)

(iii) at least 35 but less than 47.

(4 marks)

Q6 (a) The number of cherries in a cherry puff has the probability distribution as in **Table Q6**. Analyse:

Table Q6

x	4	5	6	7
$P(X=x)$	0.2	0.4	0.3	0.1

(i) the mean μ and the variance σ^2 of X.

(5 marks)

(ii) the sample mean $\mu_{\overline{X}}$ and sample variance $\sigma^2_{\overline{x}}$ for random sample of 36 cherry puffs.

(3 marks)

(iii) the probability that the average number of cherries in 36 cherry puff will be less than 5.5.

(4 marks)

(b) The distribution of heights of a certain breed of terrier dogs has a mean height of 72 centimeters and a standard deviation of 10 centimeters, whereas the distribution of heights of a certain breed of poodles has a mean height 28 centimeters with a standard deviation of 5 centimeters. Assuming that the sample means can be measured to any degree of accuracy.

(i) Compute the probability that the sample mean for a random sample of heights of 64 terriers exceeds the sample mean for a random sample of heights of 100 poodles by at most 44.2 centimeters.

(5 marks)

(ii) Find the probability that the mean height of a random sample of 100 terriers falls between 70.5 centimeters and 72.5 centimeters.

(3 marks)

Q7 (a) Many cardiac patients wear implanted pacemakers to control their heartbeat. A plastic connector module mounts on the top of the pacemaker. Assuming a standard deviation of 0.0015 and an approximate normal distribution, construct a 95% confidence interval for the mean of all connector modules made by a certain manufacturing company. A random sample of 75 modules has an average of 0.310 inch.

(8 marks)

(b) A study was conducted to determine if a certain metal treatment has any effect on the amount of metal removed in a pickling operation. A random sample of 100 pieces was immersed in a bath for 24 hours without the treatment, yielding an average of 12.2 milimeters of metal removed and a sample standard deviation of 1.1 milimeters. A second sample of 200 pieces was exposed to the treatment followed by the 24-hour immersion in the bath, resulting in an average removal of 9.1 milimeter. Compute a 98% confidence interval estimate for the difference between the population means.

(12 marks)

- END OF QUESTIONS -

FINAL EXAMINATION

SEMESTER / SESSION: SEM I / 2019/2020 PROGRAMME CODE: DAM
COURSE NAME : STATISTICS COURSE CODE : DAS 20502

Formula

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

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

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

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

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

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

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

$$Var(X)=E(X^2)-[E(X)]^2,$$

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

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

FINAL EXAMINATION

SEMESTER / SESSION: SEM I / 2019/2020 PROGRAMME CODE: DAM
COURSE NAME : STATISTICS COURSE CODE : DAS 20502

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

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

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

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

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

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

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

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

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

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

$$v = \frac{\left(s_1^2/n_1 + s_2^2/n_2\right)^2}{\dfrac{\left(s_1^2/n_1\right)^2}{n_1 - 1} + \dfrac{\left(s_2^2/n_2\right)^2}{n_2 - 1}}$$