UNIVERSITI TUN HUSSEIN ONN MALAYSIA

FINAL EXAMINATION
(ONLINE)
SEMESTER II
SESSION 2019/2020

COURSE NAME	:	STATISTICS
COURSE CODE	:	BIT 11603
PROGRAMME CODE	:	BIT
EXAMINATION DATE	:	JULY 2020
DURATION	:	2 HOURS 30 MINUTES
INSTRUCTION	:	1. ANSWER **ALL** QUESTIONS. 2. PLEASE MAKE SURE TO CLICK "SAVE ANSWER" BUTTON FOR SUBJECTIVE QUESTIONS. OBJECTIVE QUESTIONS ARE SAVED AUTOMATICALLY.

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

Q1 (a) During peak hours, the arrival of customers at a fast food restaurant is random with an average rate of 90 customers per hour.

(i) Calculate the probability that exactly two customers will arrive at the restaurant within a specified one-minute period.

(4 marks)

(ii) Calculate the probability that at most four customers will arrive within a specified two-minute period.

(5 marks)

(b) An automobile battery manufacturer claims that its midgrade battery has a mean life of 50 months with a standard deviation of 6 months. Suppose the distribution of battery lives of this particular brand is approximately normal.

(i) On the assumption that the manufacturer's claims are true, find the probability that a randomly selected battery of this type will last less than 48 months.

(5 marks)

(ii) On the same assumption, compute the probability that the mean of a random sample of 36 such batteries will be less than 48 months.

(6 marks)

Q2 (a) The mean height of 15-year-old boys is 175 *cm* and the variance is 64 *cm*. For girls, the mean is 165 *cm* and the variance is 64. If 8 boys and 8 girls were sampled, what is the probability that the mean height of the sample of boys would be at least 6 *cm* higher than the mean height of the girls sample?

(4 marks)

(b) A thread manufacturer test a sample of eight lengths of a certain type of thread made of blended materials and obtains a mean strength of 8.2 *lb* with standard deviation 0.06 *lb*. Assuming tensile strengths are normally distributed, construct a 90% confidence interval for the mean tensile strength of this thread.

(6 marks)

(c) A school decided that the number of students attending their high school was nearly unmanageable so it was split into two districts, with District 1 students going to the old high school and District 2 students going to a newly constructed building. A group of parents became concerned with how the two districts were constructed relative to income levels. A study was thus conducted to determine whether persons in suburban District 1 have a different mean income from those in District 2. A random sample of 20 homeowners was taken in District 1 and only 19 observations were taken from District 2. The data produced sample means and variances as shown in Table Q2. Use these data to construct a 95% confidence interval for $(\mu_1 - \mu_2)$.

TABLE Q2: Income Data

	District 1	District 2
Sample Size	20	20
Sample Mean	18.27	16.78
Sample Variance	8.74	6.58

(10 marks)

Q3 (a) A test of the hypothesis that eating supplement Z makes one stronger where the reading will be greater than 100 is done. A random sample of 12 persons take supplement Z for two year and then are given a fitness test. Here are the results:

116	111	101	120	99	94	106	115	107	101	110	92

(i) Write the null and alternative hypothesis for the given scenario.

(2 marks)

(ii) Test the hypothesis.

(5 marks)

(iii) Summarize your conclusion.

(2 marks)

(b) **Table Q3** relates to a study where infants listened to three types of Music in Utero and their advancement to crawling was then observed and categorized as either early, on time, or late. The study was done to determine if there was a statistically significant relation between music in Utero and time of advancement to crawling for the infants.

TABLE Q3: Infants listened to three types of Utero music and their advancement to crawling

	Advancement to crawling			Sample sizes
Music In Utero	Early	On Time	Late	N
Mozart (Piano Sonata)	50.8%	30.2%	19.0%	63
Philip Glass (minimalist music)	40.0%	38.3%	21.7%	60
White Noise and Silence	17.9%	21.1%	61.0%	44

(i) Determine the count of infant for each category of music based on the percentage.

(3 marks)

(ii) Conduct an appropriate hypothesis test to indicate an association between music in Utero and time of advancement to crawling in infants using a 0.05 significance level

(8 marks)

Q4 The following data represent the years of experience, X and salary, Y (in thousand dollars) of a random sample of professional engineers.

X	13	17	9	18	16	18	13	16
Y	21.6	25.8	15.9	48.3	38.2	56.4	28.4	43.3

(a) Plot a scatter diagram.

(2 marks: C2)

(b) Calculate the sample mean, the sample variances and the sample covariance.

(6 marks: C3)

(c) Compute the coefficient of correlation.

(5 marks: C3)

(d) Is it sufficient evidence to indicate that there is linear correlation between the years of experience and salary?

(2 marks: C4)

(e) Write the equation for estimated regression line.

(5 marks: C3)

- END OF QUESTIONS -

FINAL EXAMINATION	
SEMESTER/ SESSION: SEM II / 2019/2020	PROGRAMME CODE: BIT
COURSE NAME: STATISTICS	COURSE CODE: BIT 11603

Formula

Special Probability Distributions :

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

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

Sampling Distributions

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

Estimations :

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

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

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

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

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

where Pooled estimate of variance, $S_p^2=\dfrac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_3-2}$ with $v=n_1+n_2-2$,

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

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

FINAL EXAMINATION	
SEMESTER/ SESSION: SEM II / 2019/2020	PROGRAMME CODE: BIT
COURSE NAME: STATISTICS	COURSE CODE: BIT 11603

Formula

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

Hypothesis Testing :

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

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

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

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

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

Simple Linear Regressions :

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

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