CONFIDENTIAL

UNIVERSITI TUN HUSSEIN ONN MALAYSIA

FINAL EXAMINATION
SEMESTER II
SESSION 2016/2017

COURSE NAME	:	STATISTICS
COURSE CODE	:	BIT 11603
PROGRAMME CODE	:	BIT
EXAMINATION DATE	:	JUNE 2017
DURATION	:	3 HOURS
INSTRUCTION	:	ANSWER **ALL** QUESTIONS

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

CONFIDENTIAL

Q1 (a) Let X be a continuous random variable with the function below.

$$f(x) = \begin{cases} 20(x^3 - x^4), & 0 < x < 1 \\ 0, & \text{otherwise} \end{cases}$$

Show that $f(x)$ satisfy with the probability density function.

(5 marks)

(b) Ahmad is very forgetful. Every time he logs in to his Maybank2u.com, he only has a 40% chance of remembering his password correctly. He is allowed 3 unsuccessful attempts on any one day and then the bank will not let him try again until the next day.

(i) Draw a fully labelled tree diagram to illustrate this situation.

(4 marks)

(ii) Let X be the number of unsuccessful attempts Ahmad make on any day that he tries to log in to his bank. Find the probability distribution of X.

(6 marks)

(iii) Calculate the expected value and variance of unsuccessful attempts made by Ahmad on any day that he tries to log in.

(5 marks)

Q2 (a) KelanaConvoyOnline.com is a website based travel agency in which the customers can watch videos of the European countries they plan to visit. The number of daily hits at its website is normally distributed with $\mu = 1{,}000$ and $\sigma = 240$. The website has a limited bandwidth, which is measured in terms of the number of hits the site can handle.

(i) Find the probability of getting fewer than 900 hits.

(5 marks)

(ii) Determine a bandwidth to handle 99% of the daily traffic for KelanaConvoyOnline.com.

(5 marks)

(b) A computer crashes once every 2 days on average. Determine the probability of:

(i) exactly 5 crashes in a week.

(4 marks)

(ii) at least 3 crashes in a week.

(6 marks)

Q3 (a) The lifetime of a particular type of laptop battery is normally distributed with mean of 1100 days and a standard deviation of 80 days.

(i) Find the sample mean and standard deviation for a random sample of 400 laptop batteries.

(5 marks)

(ii) Determine the probability that the average lifetime of these 400 laptop batteries is greater than 1105 days.

(5 marks)

(b) The average internet speed in Malaysia is ranked 74th place worldwide, far behind Singapore and Thailand, reported by content delivery network services firm, Akamai Technologies Inc. Akamai measured the internet speed from 10 unique IPv4 addresses shown in **Table 1**.

Table 1: Internet speed in Mbps

5.6	5.2	6.0	6.1	5.9
4.3	4.9	5.0	3.3	3.0

Construct a 98% confidence interval about the population mean.

(10 marks)

Q4 (a) Test the following hyphoteses.

(i) $H_0 : \mu = 33$, $H_1 : \mu \neq 33$, $n = 14$, $\bar{x} = 32.1$, $s = 2$, $\alpha = 0.05$.

(5 marks)

(ii) $H_0 : \mu = 70$, $H_1 : \mu > 70$, $n = 120$, $\bar{x} = 71.8$, $\sigma = 8.9$, $\alpha = 0.01$.

(5 marks)

(b) Screening of a computer system by a new antivirus software takes a random time with a standard deviation $\sigma = 12$ min. When this software was tried on 40 randomly selected computers, the average screening time was 27 minutes. Analyse if there is a sufficient evidence that the expected screening time is less than 30 min at 5% level of significance.

(10 marks)

Q5 As a performance analyst, you will receive data from database department. One of the task is to find a relationship between two technical metrics, number of order lines entered into the system and CPU utilization. This task is to ensure the system can be supported without any performance problem. **Table 2** shows 5 samples number of order line entries and CPU utilization.

Table 2: Number of order lines and CPU utilization

No. of order lines	CPU utilization (%)
164	27
131	32
120	21
119	20
20	2

(a) State the independent and dependent variables.

(2 marks)

(b) Based on the data in **Table 2**, construct a least-squares regression line between number of order lines and CPU utilization.

(10 marks)

(c) Predict the CPU utilization based on 150 order lines entered into the system.

(4 marks)

(d) Determine and interpret the Pearson correlation coefficient between the two variables.

(4 marks)

- END OF QUESTION –

FINAL EXAMINATION

SEMESTER/SESSION : SEM II / 2016/2017 PROGRAMME CODE : BIT
COURSE NAME : STATISTICS COURSE CODE : BIT 11603

Formula

Table 1 : Descriptive Statistics

$\mu = \frac{\sum_{i=1}^{n} x_i}{N}$	$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$
$\sigma^2 = \frac{\sum_{i=1}^{n}(x_i - \mu)^2}{N}$	$s^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}$
$M = L_m + C\left(\frac{\frac{n}{2} - F}{f_m}\right)$	$M_o = L + C\left(\frac{d_1}{d_1 + d_2}\right)$

Table 2: Random Variable

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

Table 3 : Probability Distribution

$P(X = x) = {}^{n}C_x \cdot p^x \cdot q^{n-x}, x = 0,1,2\ldots,n \;, X \sim B(n.p)$
$P(X = x) = \frac{e^{-\mu} \cdot \mu^x}{x!} \quad , \quad x = 0, 1, 2, \ldots, \infty \;, X \sim P_o(\mu)$
$X \sim N(\mu, \sigma^2),\; Z \sim N(0,1),\; Z = \frac{X - \mu}{\sigma}$
$\chi^2 = \frac{(n-1).s^2}{\sigma^2{}_0} \;, F = \frac{s_1^2}{s_2^2}$

FINAL EXAMINATION

SEMESTER/SESSION : SEM II / 2016/2017 PROGRAMME CODE : BIT
COURSE NAME : STATISTICS COURSE CODE : BIT 11603

Table 4 : Sampling distribution, estimation and hypothesis testing

$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$, $\bar{X}_1 - \bar{X}_2 \sim N\left(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right)$
$Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$, $Z = \frac{\bar{x} - \mu}{s/\sqrt{n}}$, $T = \frac{\bar{x} - \mu}{s/\sqrt{n}}$
$\bar{x} - z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right) < \mu < \bar{x} + z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right)$
$\bar{x} - z_{\alpha/2}\left(\frac{s}{\sqrt{n}}\right) < \mu < \bar{x} + z_{\alpha/2}\left(\frac{s}{\sqrt{n}}\right)$
$\bar{x} - t_{\alpha/2,\,v}\left(\frac{s}{\sqrt{n}}\right) < \mu < \bar{x} + t_{\alpha/2,\,v}\left(\frac{s}{\sqrt{n}}\right)$, $v = n-1$
$(\bar{x}_1 - \bar{x}_2) \pm z_{\alpha/2}\left(\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\right)$
$(\bar{x}_1 - \bar{x}_2) \pm z_{\alpha/2}\left(\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}\right)$
$(\bar{x}_1 - \bar{x}_2) \pm t_{\alpha/2,v}\left(\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}\right)$, $v = \frac{\left(s_1^2/n_1 + s_2^2/n_2\right)^2}{\frac{\left(s_1^2/n_1\right)^2}{n_1 - 1} + \frac{\left(s_2^2/n_2\right)^2}{n_2 - 1}}$
$(\bar{x}_1 - \bar{x}_2) \pm t_{\alpha/2,v}\left(\sqrt{\frac{S_p^2}{n} + \frac{S_p^2}{n}}\right)$
$S_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}, v = n_1 + n_2 - 2$
$\frac{(n-1)s^2}{\chi^2_{\alpha/2,v}} \le \sigma^2 \le \frac{(n-1)s^2}{\chi^2_{1-\alpha/2,v}}$, $v = n - 1$
$\frac{s_1^2}{s_2^2}\frac{1}{f_{\alpha/2,v1,v2}} \le \frac{\sigma_1^2}{\sigma_2^2} \le \frac{s_1^2}{s_2^2} f_{\alpha/2,v2,v1}$, $v_1 = n_1 - 1$, $v_2 = n_2 - 1$

FINAL EXAMINATION

SEMESTER/SESSION : SEM II / 2016/2017	PROGRAMME CODE : BIT
COURSE NAME : STATISTICS	COURSE CODE : BIT 11603

Table 5: Simple Linear Regression

$S_{xy}=\sum x_i y_i - \frac{\sum x_i \sum y_i}{n}$	$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$
$S_{xx}=\sum x_i^2 - \frac{\left(\sum x_i\right)^2}{n}$	$\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}}$
$S_{yy}=\sum y_i^2 - \frac{\left(\sum y_i\right)^2}{n}$	$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$
$r = \frac{S_{xy}}{\sqrt{S_{xx} S_{yy}}}$	$SSE = S_{yy} - \hat{\beta}_1 S_{xy}$
$T = \frac{\hat{\beta}_1 - \beta_1^*}{\sqrt{MSE/S_{xx}}} \sim t_{n-2}$	$MSE = \frac{SSE}{n-k}$
$T = \frac{\hat{\beta}_0 - \beta_o^*}{\sqrt{MSE\left(\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}\right)}} \sim t_{n-2}$	