UNIVERSITI TUN HUSSEIN ONN MALAYSIA

FINAL EXAMINATION
SEMESTER I
SESI 2016/2017

COURSE NAME	:	ENGINEERING STATISTICS
COURSE CODE	:	BDA 24103
PROGRAMME	:	2 BDD
EXAMINATION DATE	:	DECEMBER 2016 / JANUARY 2017
DURATION	:	3 HOURS
INSTRUCTION	:	PLEASE ANSWER **FIVE (5)** QUESTIONS FROM **SIX (6)** QUESTIONS PROVIDED.

THIS PAPER CONSISTS OF **FOURTEEN (14)** PAGES

Q1 (a) Choose the right answer

i. Which of these is NOT continuous data

A. A person's height each year
B. The volume of water in a swimming pool each day
C. Cars produced in a factory each day
D. A person's weight on each birthday

(1 mark)

ii. Which of these is **NOT** discrete data

A. Height of a sunflower as measured each day
B. Number of students absent from school each day
C. Number of widgets sold each day
D. The number of people who drive through a red light each hour during rush hour

(1 mark)

iii. Which of the following is a discrete random variable?

I. Average height of a randomly selected group of boys.
II. Number of sweepstakes winners from a City.
III. Number of general elections in the 20th century.

A. I only
B. II only
C. III only
D. I and II
E. II and III

(2 marks)

iv. The number of adults living in homes on a randomly selected city block is described by the following probability distribution.

Number of adults, x	1	2	3	4 or more
Probability, P(x)	0.25	0.50	0.15	???

What is the probability that 4 or more adults reside at a randomly selected home?

A. 0.10
B. 0.15
C. 0.25
D. 0.50
E. 0.90

(2 marks)

v. In a recent little softball league game, each player went to bat 4 times. The number of hits made by each player is described by the following probability distribution.

Number of hits, x	0	1	2	3	4
Probability, P(x)	0.10	0.20	0.30	0.25	0.15

What is the mean of the probability distribution?

A. 1.00
B. 1.75
C. 2.00
D. 2.25
E. None of the above

(2 marks)

(b) Let X be a continuous random variable with the following probability distribution function

$$fX(x)=\begin{cases} ce^{-x} & x \geq 0 \\ 0 & otherwise \end{cases}$$

where c is a positive constant.

i. Find c. (2 marks)
ii. Find the cumulative distribution function of X (2 marks)
iii. Find P(1<X<3). (3 marks)

(c) Let X be a continuous random variable with PDF

$$fX(x)=\begin{cases} \frac{3}{x^4} & x \geq 1 \\ 0 & otherwise \end{cases}$$

Find the mean and variance of X. (5 marks)

Q2 (a) The number of defective components produced by a certain process in one day has a Poisson distribution with a mean of 20. Each defective component has the probability 0.60 of being repairable.

(i) Find the probability that exactly 15 are defective product.

(2 marks)

(ii) Given that exactly 15 defective components are produced, find the probability that exactly 10 of them are repairable.

(2 marks)

(b) The bolts manufactured for a certain application, 90% met the length specification and could be used immediately, 6% were too long and could be used after being cut, and 4% were too short and must be scrapped.

(i) Find the probability that a randomly selected bolt can be used (either immediately or after being cut).

(2 marks)

(ii) Find the probability that fewer than 9 out of a sample of 10 bolts can be used (either immediately or after being cut).

(2 marks)

(c) A fiber-spinning process currently produces a fiber whose strength is normally distributed with a mean of 75 N/m^2. The minimum acceptable strength is 65 N/m^2.

(i) Ten percent of the fiber produced by the current method fails to meet the minimum specification. What is the standard deviation of fiber strengths in the current process?

(4 marks)

(ii) If the mean remains at 75 N/m2, what must the standard deviation be so that only 1% of the fiber fail to meet the specification?

(4 marks)

(iii) If the standard deviation is 5 N/m2, to what value must the mean be set so that only 1% of the fiber fail to meet the specification?

(4 marks)

Q3 A certain type of thread is manufactured with a mean tensile strength of 78.3 kg and a standard deviation of 5.6 kg. Assuming that the strength of this type of thread is distributed approximately normal, find;

(a) The probability that the mean strength of a random sample of 10 such thread falls between 77 kg and 78 kg.

(5 marks)

(b) The probability that the mean strength is greater than 79 kg.

(5 marks)

(c) The probability that the mean strength is less than 76 kg.

(5 marks)

(d) The value of $\bar{x}$ to the right of which 15% of the means computed from random samples of size 10 would fall.

(5 marks)

Q4 (a) A study was conducted to monitor the quality of water from a pretreatment plant. 80 samples were taken during the study and the average of suspended solids was found to be 2.15 mg/l with standard deviation of 0.085 mg/l. Construct 95% confidence interval for the population mean.

(8 marks)

(b) A production engineer collected diameter data of copper rods produced by two different rolling machines. The population diameter data on both machines follow normal distribution with standard deviation of 0.45 mm for machine A and 0.38 mm for machine B. 40 random samples were taken from machine A yielding the average rod diameter of 32.5 mm. 50 random samples were taken from machine B yielding 31.9 mm average rod diameter.

(i) Calculate 90% confidence interval for the difference between means rod diameters from machine A and machine B.

(10 marks)

(ii) Explain your answer.

(2 marks)

Q5 Table 1 shows a fitness test among primary school female students, 8 girls were timed to see how long (in seconds) they could hold their breath in a relaxed position and after jumping for two minutes. The fitness instructor thought that the jumping would not affect their times, Test her hypothesis by using a 0.01 level of significance.

Table 1: A fitness test data

Relaxed mode (seconds)	After Jumping (seconds)
26	21
47	40
30	28
22	21
23	25
45	43
37	35
29	32

(20 marks)

Q6 A manager of PROCAR Enterprise conducted a study to determine whether there is a relationship between the number of sales people on duty and the number of cars sold in a week. The data of the study are shown in the Table 2 below.

Table 2: Data of the relationship study

Week	Number of Sales People	Number of Cars Sold
1	6	76
2	6	68
3	5	58
4	3	32
5	4	54
6	3	28

(a) Estimate the regression line by using the least square method. Interpret the result.
(10 marks)

(b) Estimate the number of car sold when the number of sales people is 9. Interpret the result.
(2 marks)

(c) Test the slope whether it is greater than four at 5% level of significance.
(8 marks)

END OF QUESTIONS

Appendix I

FINAL EXAMINATION

SEMESTER / SESSION : SEMESTER I /2016/2017
COURSE : ENGINEERING STATISTICS

PROGRAM : 2BDD
COURSE CODE: BDA24103

EQUATIONS

P(X=x) = P(x) ;

- $0 \le P(x) \le 1$
- $\sum_{i=1}^{n} P(X) = 1$
- $P(x_i) = P(X=x_i)$

$$F(x) = P(X \le x) = \sum_{-\infty}^{n} P(X = x)$$

$$\mu = E(X) = \sum_{all\ xi} X_i P(X_i)$$

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

$$E(X^2) = \sum_{all\ xi} X_i^2 . P(X_i)$$

(a)	$P(X \ge k)$ = from table
(b)	$P(X < k) = 1 - P(X \ge k)$
(c)	$P(X \le k) = 1 - P(X \ge k+1)$
(d)	$P(X > k) = P(X \ge k+1)$
(e)	$P(X = k) = P(X \ge k) - P(X \ge k+1)$
(f)	$P(k \le X \le l) = P(X \ge k) - P(X \ge l+1)$
(g)	$P(k < X < l) = P(X \ge k+1) - P(X \ge l)$
(h)	$P(k \le X < l) = P(X \ge k) - P(X \ge l)$
(i)	$P(k < X \le l) = P(X \ge k+1) - P(X \ge l+1)$

Cumulative distribution function

- $P(X \le r) = F(r)$
- $P(X > r) = 1 - F(r)$
- $P(X < r) = P(X \le r-1) = F(r-1)$
- $P(X = r) = F(r) - F(r-1)$
- $P(r < X \le s) = F(s) - F(r)$
- $P(r \le X \le s) = F(s) - F(r) + f(r)$
- $P(r \le X < s) = F(s) - F(r) + f(r) - f(s)$
- $P(r < X < s) = F(s) - F(r) - f(s)$

$$\sigma = Std(X) = \sqrt{Var(X)}$$

$$E(aX + b) = aE(x) + b$$

$$Var(aX + b) = a^2 Var(x)$$

PDF of Cont. random variable

- $f(x) \ge 1$
- $\int_{-\infty}^{\infty} f(x)dx = 1$
- $P(a < x < b) = P(a \le x < b) = P(a < x \le b) = P(a \le x \le b) = \int_a^b f(x)dx$

$$F(x) = P(X \le x) = \int_{-\infty}^{x} f(x)dx$$

$$\mu = E(X) = \int_{-\infty}^{\infty} x.f(x)dx$$

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

$$E(X^2) = \int_{-\infty}^{\infty} x^2.f(x)dx$$

$$\sigma = Sd(X) = \sqrt{Var(X)}$$

Appendix I

FINAL EXAMINATION

SEMESTER / SESSION : SEMESTER I /2016/2017
COURSE : ENGINEERING STATISTICS

PROGRAM : 2BDD
COURSE CODE: BDA24103

Binomial Distribution	Normal Distribution
(i) $P(X = a)$	$P(a - 0.5 < X < a + 0.5)$
(ii) $P(X \geq a)$	$P(X > a - 0.5)$
(iii) $P(X > a)$	$P(X > a + 0.5)$
(iv) $P(X \leq a)$	$P(X < a + 0.5)$
(v) $P(X < a)$	$P(X < a - 0.5)$
(vi) $P(a \leq X \leq b)$	$P(a - 0.5 < X < b + 0.5)$
(vii) $P(a < X < b)$	$P(a + 0.5 < X < b - 0.5)$

For all cases, $\mu = E(x) = np$, $\sigma = \sqrt{npq}$, $np \geq 5$, and $nq \geq 5$.

Binomial Distribution	
Formula	$P(X = x) = \frac{n!}{x! \cdot (n-x)!} \cdot p^x \cdot q^{n-x} = {}^n C_x \cdot p^x \cdot q^{n-x}$
Mean	$\mu = np$
Variance	$\sigma^2 = npq$

Poisson Distribution	
Formula	$P(X = x) = \frac{e^{-\mu} \cdot \mu^x}{x!}$, $x = 0, 1, 2, \ldots, \infty$
Mean	$\mu = \mu$
Variance	$\sigma^2 = \mu$

Normal Distribution	
Formula	$P\left(Z = \frac{x - \mu}{\sigma}\right)$

Poisson Approximation to the Binomial Distribution	
Condition	Use if $n \geq 30$ and $p \leq 0.1$
Mean	$\mu = np$

Normal Approximation to the Binomial Distribution	
Condition	Use if n is large and $np \geq 5$ and $nq \geq 5$
Mean	$\mu = np$
Variance	$\sigma^2 = npq$

Appendix I

FINAL EXAMINATION

SEMESTER / SESSION : SEMESTER I /2016/2017 PROGRAM : 2BDD
COURSE : ENGINEERING STATISTICS COURSE CODE: BDA24103

Sampling error of single mean : $e = \left| \bar{x} - \mu \right|$.

$P\left(\bar{x} > r\right) = P\left(Z > \dfrac{r - \mu_{\bar{x}}}{\sigma_{\bar{x}}}\right)$.

Population mean, $\mu = \dfrac{\sum x}{N}$.

Sample mean, is $\bar{x} = \dfrac{\sum x}{n}$.

$\sigma_{\bar{x}_1 - \bar{x}_2} = \sqrt{\dfrac{\sigma^2_1}{n_1} + \dfrac{\sigma^2_2}{n_2}}$.

Z-value for sampling distribution of $\bar{x}$ is $Z = \dfrac{\bar{x} - \mu}{\sigma / \sqrt{n}}$.

$P\left(\bar{x}_1 - \bar{x}_2 > r\right) = P\left(Z > \dfrac{r - \mu_{\bar{x}_1 - \bar{x}_1}}{\sigma_{\bar{x}_1 - \bar{x}_2}}\right)$.

$\sigma_{\bar{x}} = \sigma / \sqrt{n}$.

Maximum error : $E = Z_{\alpha/2}\left(\dfrac{\sigma}{\sqrt{n}}\right)$, Sample size : $n = \left(\dfrac{Z_{\alpha/2}(\sigma)}{E}\right)^2$

(i) σ is known : $\left(\bar{x} - z_{\alpha/2}(\sigma / \sqrt{n}) < \mu < \bar{x} + z_{\alpha/2}(\sigma / \sqrt{n})\right)$

(ii) σ is unknown : $\left(\bar{x} - z_{\alpha/2}(s / \sqrt{n}) < \mu < \bar{x} + z_{\alpha/2}(s / \sqrt{n})\right)$

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

***Z* distribution case**

(i) σ is known : $(\bar{x}_1 - \bar{x}_2) \pm z_{\alpha/2}\left(\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}\right)$

(ii) σ is unknown : $(\bar{x}_1 - \bar{x}_2) \pm z_{\alpha/2}\left(\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}\right)$

Appendix I

FINAL EXAMINATION

SEMESTER / SESSION : SEMESTER I /2016/2017 PROGRAM : 2BDD
COURSE : ENGINEERING STATISTICS COURSE CODE: BDA24103

t distribution case

(i) $n_1 = n_2,\ \sigma_1^2 = \sigma_2^2\ :\ (\bar{x}_1 - \bar{x}_2) \pm t_{\alpha/2,v}\left(\sqrt{\frac{1}{n}\left(s_1^{\ 2} + s_2^{\ 2}\right)}\right)$; $v = 2n - 2$

(ii) $n_1 = n_2,\ \sigma_1^2 = \sigma_2^2\ :\ (\bar{x}_1 - \bar{x}_2) \pm t_{\alpha/2,v} S_p\left(\sqrt{\frac{2}{n}}\right)$; $v = 2n - 2$

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

(iii) $n_1 \neq n_2,\ \sigma_1^2 = \sigma_2^2\ :\ (\bar{x}_1 - \bar{x}_2) \pm t_{\alpha/2,v} S_p\left(\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}\right)$; $v = n_1 + n_2 - 2$

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

(iv) $n_1 \neq n_2,\ \sigma_1^2 \neq \sigma_2^2\ :\ (\bar{x}_1 - \bar{x}_2) \pm t_{\alpha/2,v}\left(\sqrt{\frac{s_1^{\ 2}}{n_1} + \frac{s_2^{\ 2}}{n_2}}\right)$, $v = \dfrac{\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)^2}{\frac{\left(\frac{S_1^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{S_2^2}{n_2}\right)^2}{n_2 - 1}}$

$$\frac{(n-1)s^2}{\chi^2_{\alpha/2,v}} \leq \sigma^2 \leq \frac{(n-1)s^2}{\chi^2_{1-\alpha/2,v}} \ ;\ v = n - 1$$

TERBUKA

$$\frac{s_1^2}{s_2^2}\frac{1}{f_{\alpha/2,v1,v2}} \leq \frac{\sigma_1^2}{\sigma_2^2} \leq \frac{s_1^2}{s_2^2} f_{\alpha/2,v2,v1} \ ;\ v_1 = n_1 - 1 \text{ and } v_2 = n_2 - 1$$

Appendix I

FINAL EXAMINATION

SEMESTER / SESSION : SEMESTER I /2016/2017
COURSE : ENGINEERING STATISTICS

PROGRAM : 2BDD
COURSE CODE: BDA24103

Case	Variances	Samples size	Statistical Test
A	Known	$n_1, n_2 \geq 30$	$Z_{Test} = \dfrac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}$
B	Known	$n_1, n_2 < 30$	$Z_{Test} = \dfrac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}$
C	Unknown	$n_1, n_2 \geq 30$	$Z_{Test} = \dfrac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}}$
D	Unknown (Equal)	$n_1, n_2 < 30$	$T_{Test} = \dfrac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{S_p \cdot \sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}}$ $v = n_1 + n_2 - 2$
E	Unknown (Not equal)	$n_1 = n_2 < 30$	$T_{Test} = \dfrac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{1}{n}(s_1^2 + s_2^2)}}$ $v = 2(n-1)$
F	Unknown (Not equal)	$n_1, n_2 < 30$	$T_{Test} = \dfrac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}}$ $v = \dfrac{\left(\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}\right)^2}{\dfrac{\left(\dfrac{s_1^2}{n_1}\right)^2}{n_1 - 1} + \dfrac{\left(\dfrac{s_2^2}{n_2}\right)^2}{n_2 - 1}}$

Appendix I

FINAL EXAMINATION

SEMESTER / SESSION : SEMESTER I /2016/2017
COURSE : ENGINEERING STATISTICS

PROGRAM : 2BDD
COURSE CODE: BDA24103

(i) Least Squares Method

The model : $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$

$\hat{\beta}_1 = \dfrac{Sxy}{Sxx}$ (slope) and $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1\bar{x}$, ($y$-intercept) where

$$Sxy = \sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) = \sum_{i=1}^{n} x_i y_i - \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} y_i\right),$$

$$Sxx = \sum_{i=1}^{n}(x_i - \bar{x})^2 = \sum_{i=1}^{n} x_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)^2,$$

$$Syy = \sum_{i=1}^{n}(y_i - \bar{y})^2 = \sum_{i=1}^{n} y_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} y_i\right)^2$$

and n = sample size

Inference of Regression Coefficients

(i) Slope

$$SSE = Syy - \hat{\beta}_1 S_{xy}\ , \quad MSE = \frac{SSE}{n-2}\ , \qquad T_{test} = \frac{\hat{\beta}_1 - \beta_C}{\sqrt{MSE/S_{xx}}}$$

(ii) Intercept

$$T_{test} = \frac{\hat{\beta}_0 - \beta_C}{\sqrt{MSE(1/n + \bar{x}^2/Sxx)}}$$

TERBUKA

FINAL EXAMINATION

SEMESTER / SESSION : SEMESTER I /2016/2017 PROGRAM : 2BDD
COURSE : ENGINEERING STATISTICS COURSE CODE: BDA24103

3. Confidence Intervals of the Regression Line

(i) Slope, β_1

$$\hat{\beta}_1 - t_{\alpha/2,v}\sqrt{MSE/Sxx} < \beta_1 < \hat{\beta}_1 + t_{\alpha/2,v}\sqrt{MSE/Sxx},$$

where $v = n$-2

(ii) Intercept, β_0

$$\hat{\beta}_0 - t_{\alpha/2,v}\sqrt{MSE\left(\frac{1}{n}+\frac{\bar{x}^2}{Sxx}\right)} < \beta_0 < \hat{\beta}_0 + t_{\alpha/2,v}\sqrt{MSE\left(\frac{1}{n}+\frac{\bar{x}^2}{Sxx}\right)},$$

where $v = n$-2

4. Coefficient of Determination, r^2.

$$r^2 = \frac{Syy - SSE}{Syy} = 1 - \frac{SSE}{Syy}$$

5. Coefficient of Pearson Correlation, r.

$$r = \frac{Sxy}{\sqrt{Sxx \cdot Syy}}$$