UNIVERSITI TUN HUSSEIN ONN MALAYSIA

FINAL EXAMINATION
SEMESTER II
SESI 2018/2019

COURSE NAME	:	ENGINEERING STATISTICS
COURSE CODE	:	BDA 24103
PROGRAMME	:	BDD
EXAMINATION DATE	:	JUNE/JULY 2019
DURATION	:	3 HOURS
INSTRUCTION	:	**SECTION A:** ANSWER **ALL** QUESTIONS. **SECTION B:** ANSWER **THREE (3)** FROM FOUR (4) QUESTIONS.

THIS PAPER CONSISTS OF **ELEVEN (11)** PAGES

SECTION A

Instruction: Please answer **ALL questions** in this section.

Q1 (a) State the null and alternative hypothesis for the following cases:

(i) A hypothesis test will be used to potentially provide evidence that the population mean is more than 20.

(2 marks)

(ii) A hypothesis test will be used to potentially provide evidence that the population mean is not equal to 9.

(2 marks)

(b) The weights of salmon grown at a commercial hatchery are normally distributed with a standard deviation of 0.5 kg. The hatchery claims that the mean weight of this year's crop is at least 1.5 kg. Suppose a random sample of 16 fish yielded an average weight of 3.2 kg. Is this strong enough evidence to reject the hatchery's claims at the

(i) 5 percent level of significance and (3 marks)

(ii) 1 percent level of significance? (3 marks)

(c) The production of large electrical transformers and capacitators in a semiconductor plant located in Tanjung Agas requires the use of lead (Pb), which are extremely hazardous when released into the environment. Two methods have been suggested to monitor the levels of lead in fish pond, located near to the plant. It is believed that each method will result in a normal random variable that depends on the method. Test the hypothesis at the $\alpha = .10$ level of significance that both methods have the same variance, if a given fish is checked 8 times by each method as shown in Table Q1 (data recorded is in parts per million).

Table Q1

Method 1	6.2	5.8	5.7	6.3	5.9	6.1	6.2	5.7
Method 2	6.3	5.7	5.9	6.4	5.8	6.2	6.3	5.5

(10 marks)

Q2 The table shows the total fertilizer weight and harvested black paper permonth recoreded from 10 different plant.

Fertilizer (g/month)	1.8	1.9	1.8	2.4	5.3	3.1	5.5	5.1	8.3	13.7
Black paper (g/month)	21	26	25	34	53	25	41	45	54	94

a) Develop a model relating fertilizer weight (in gram) and harvested blackpaper weight (in gram).

(6 marks)

b) Draw a graph and interperatyue the value of $\hat{\beta}$ obtained from the derived model.

(4 marks)

c) Is the harvested black paper weight related to the consumed fertilizer weight? Conduct an appropriate testing test using $\alpha = 0.01$.

(6 marks)

d) If possible, predict the minimum fertilizer weight need to be consumed in a month to consistently produce 42 gram/month of black paper.

(4 marks)

SECTION B

Instruction: Please answer **THREE (3) questions** from FOUR (4) questions provided in this section.

Q3 (a) Let the X is a continuous random variable with the next probability density function

$$f(y) = \begin{cases} c(2x^3 + 5) & , 0 \leq x \leq 2 \\ 0 & , \text{otherwise} \end{cases}$$

(i) Evaluate c

(2 marks)

(ii) Find P(0 <X ≤1)

(2 marks)

(b) Given that the distribution function for a random variables Y is as

$$F(y)=\begin{cases} 0 & , \quad y \le 0 \\ \dfrac{y}{8} & , \quad 0 < y < 2 \\ \dfrac{y^2}{16} & , \quad 2 \le y \le 4 \\ 1 & , \quad y > 4 \end{cases}$$

(i) Find the probability distribution of y.

(4 marks)

(ii) Find the probability of y is larger or equal to 1 but equal or less than 3

(6 marks)

(iii) Find the probability of y is larger than or equal to 1.5

(6 marks)

Q4 A Maths quiz consists of 10 questions were given to a group of students. Each question has 4 choices. Grade A is given to the students managed to get 70% marks and above. According to the compiled data, 45% of students managed to get grade A.

(a) What is the probability that exactly 3 of a random sample of 5 the students were grade A?

(6 marks)

(b) What is the probability that not more than 7 of a random sample of those 15 students were grade A?

(6 marks)

(c) Among the candidates, one student has not studied and decided to guess the answers to every questions. Find the probability that the candidate will get grade A in the quiz.

(8 marks)

Q5 **(a)** Drilling tool produced from a production line were measured in quality control department. The tool tip mean diameter was 5.02 mm and the standard deviation was 0.3 mm. A random sample of 100 tools were selected from the same production line. Find the :

i. Probability that the tool diameter will fall in between 4.96 and 5.0 mm.

(4 marks)

ii. Probability that the tool diametr will be more than 5.10 mm

(4 marks)

(b) An experiment was conducted to evaluate the performance of two drilling tools with different design by measuring maximum drilling time before it reach the allowable wear limit. It was recoreded that drilling tool A maximum drilling time is distributed normally with mean of 6 hour and standard deviation of 1 hour, while tool B maximum drilling time is normally distributed with mean of 5 hour and standard deviantion of 0.5 hour. A random sample of 40 drilling tool was used teste dfrom each design.

i. Calculate the probability that the difference in the sample means is larger than 0.5 hour

(6 marks)

ii. What is the probability that the sample mean of maximum drilling hour recorded by drilling tool A is larger than B.

(6 marks)

Q6 (a) From a zinc measurement sample in 36 different location, the obtained average zinc concentration recovered is 2.6 g/ml. Construct the 99% confidence interval for the mean zinc concentration. Assume that the population standard deviation is 0.3 g/ml

(6 marks)

(b) Two independent sampling station were chosen for an investigation of index acid mine pollution in rivers Malaysia. For 12 monthly samples collected at the first station, the pollution index has a mean, 3.11 and a standard deviation 0.771, while 10 monthly sample collected at the second station had a mean 2.04 and a standard deviation 0.448. Assume that the population are approximately normally distributed with equal variance. Find the 90% confidence interval for the difference between the population means for the two stations.

(14 marks)

END OF QUESTIONS

FINAL EXAMINATION

SEMESTER / SESSION : SEMESTER II /2018/2019
COURSE : ENGINEERING STATISTICS

PROGRAMME : BDD
COURSE CODE: BDA 24103

EQUATIONS

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

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

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

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

Note :

- $E(aX + b) = a\,E(x) + b.$
- $Var(aX + b) = a^2\,Var(x)$

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

$$F(x) = P(X \le x) = \int_{-\infty}^{x} f(x)\,dx \text{ for } -\infty < x < \infty.$$

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

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

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

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

(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)	$P(k < X \le l) = P(X \ge k+1) - P(X \ge l-1)$

FINAL EXAMINATION

SEMESTER / SESSION : SEMESTER II /2018/2019
COURSE : ENGINEERING STATISTICS

PROGRAMME : BDD
COURSE CODE: BDA 24103

EQUATIONS

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

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

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

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

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

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

$\bar{x}\sim N\left(\mu_{\bar{x}},\sigma^2_{\bar{x}}\right)$

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

$\mu_{\bar{x}_1-\bar{x}_2}=\mu_1-\mu_2$

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

$\bar{x}\sim N\left(\mu_{\bar{x}_1-\bar{x}_2},\sigma^2_{\bar{x}_1-\bar{x}_2}\right)$

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

FINAL EXAMINATION

SEMESTER / SESSION : SEMESTER II /2018/2019 | PROGRAMME : BDD
COURSE : ENGINEERING STATISTICS | COURSE CODE: BDA 24103

Confidence Interval for Single Mean

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

(a) $n \geq 30$ **or** σ **known**

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

(b) $n < 30$ **and** σ **unknown**

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

Confidence Interval for a Difference Between Two Means

(a) ***Z*** **distribution case**

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

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

(b) ***t*** **distribution case**

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

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

TERBUKA

FINAL EXAMINATION

SEMESTER / SESSION : SEMESTER II /2018/2019
COURSE : ENGINEERING STATISTICS

PROGRAMME : BDD
COURSE CODE: BDA 24103

Confidence Interval for Single Population Variance

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

Confidence Interval for Ratio of Two Population Variances

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

Case	Variances	Samples size	Statistical Test
A	Known	$n_1, n_2 \ge 30$	$Z_{Test} = \dfrac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}$
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	$n_1, n_2 \ge 30$	$Z_{Test} = \dfrac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}}$
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)	$n_1, n_2 < 30$	$T_{Test} = \dfrac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}}$ $v = \dfrac{\left(\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}\right)^2}{\dfrac{\left(\dfrac{s_1^2}{n_1}\right)^2}{n_1 - 1} + \dfrac{\left(\dfrac{s_2^2}{n_2}\right)^2}{n_2 - 1}}$

FINAL EXAMINATION

SEMESTER / SESSION : SEMESTER II /2018/2019 PROGRAMME : BDD
COURSE : ENGINEERING STATISTICS COURSE CODE: BDA 24103

Simple Linear Regression Model

(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

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

Coefficient of Determination, r^2.

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

Coefficient of Pearson Correlation, r.

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

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