CONFIDENTIAL

UNIVERSITI TUN HUSSEIN ONN MALAYSIA

FINAL EXAMINATION
SEMESTER II
SESI 2014/2015

COURSE NAME	:	ENGINEERING STATISTICS
COURSE CODE	:	BDA 24103
PROGRAMME	:	4BDD
EXAMINATION DATE	:	JUNE 2015 / JULY 2015
DURATION	:	3 HOURS
INSTRUCTION	:	PLEASE ANSWER **FIVE (5)** QUESTIONS FROM **SIX (6)** QUESTIONS PROVIDED.

THIS PAPER CONSISTS OF **TWELVE (12)** PAGES

CONFIDENTIAL

Q1 (a) A random variable assumes that any of several different values as a result of some random event or experiment. There are two types of random variable which is discrete random variable and continuous random variable. By using appropriate example, describe the difference between these two random variables.

(4 marks)

(b) Specifications for the thickness of aluminium sheets to be made into cans between 8 and 11 thousandths of an inch. Let X is the thickness of an aluminium sheet and the probability density function of X is given by:

$$f(x) = \begin{cases} \dfrac{x}{54}, & 6 < x < 12 \\ 0, & \text{otherwise} \end{cases}$$

(i) Calculate the probability of the thickness of aluminium at least 9.5 thousandths of an inch.

(ii) Calculate the probability of the thickness of aluminium which not more than 6.5 thousandths of an inch.

(iii) Find the mean of the thickness an aluminium sheet.

(iv) Find the standard deviation of the thickness of an aluminium sheet.

(16 marks)

Q2 The average number of traffic accidents along Km 156 - 255 of PLAS highway is two per week. Assume that the number of accidents follows a Poisson distribution with μ=2

(a) Write the probability distribution notation for this problem.

(2 marks)

(b) Find the probability of no accident along km 156 - 255 of PLAS highway during a 1- week period.

(7 marks)

(c) Find the probability of at most three accidents on the same section of highway during a 2-week period.

(8 marks)

(d) Will the probability of no accident along km 300 – 399 of the highway be the same? State your reasons.

(3 marks)

Q3 (a) The total number of students at a large university is more than 25,000. The student height data follows the normal distribution with mean height of 1.625 meter and standard deviation of 0.075 meter. A random sample of 36 students is selected. Compute the probability that the average height is greater than 1.65 meter.

(8 marks)

(b) The average operating life of light bulbs produced by Company A is 8100 hours and a standard deviation is 400 hours. Company B produces similar light bulbs with a mean operating life of 8200 hours and a standard deviation of 480 hours. It is known that the populations are normally distributed. You are given 60 samples from Company A and 80 samples from Company B. Analyse the situation and compute the probability that sample light bulbs from Company A will have a mean operating life of at most 50 hours more than the sample mean operating life from Company B.

(12 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 a standard deviation of 0.085 mg/l. Construct 95% confidence interval for the population mean.

(8 marks)

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

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

(10 marks)

(ii) Explain the meaning of your answer.

(2 marks)

Q5 (a) A researcher decided to test the claim that the average age of lifeguards in Perhentian island is different than 33 years. He selects a sample of fourteen guards and finds the mean of the sample to be 32.1 years, with sample standard deviation of 2 years. By using significance testing, provide any evidence to support the claim by using 0.05 significance level.

(8 marks)

(b) An experiment was performed to compare the abrasive wear of two different laminated materials. Twelve pieces of material 1 were tested by exposing each piece to a machine measuring wear. Ten pieces of material 2 were similarly tested. In each case, the depth of wear was observed. The samples of material 1 gave an average (coded) wear of 85 units with a sample standard deviation of 4, while the samples of material 2 gave an average of 81 and a sample standard deviation of 5. Assume that the populations to be approximately normal with equal variance. If we conclude at the 0.05 level of significance that the abrasive wear of material 1 exceeds that of material 2 by more than 2 units;

(i) State the null and alternative hypotheses

(2 marks)

(ii) Find the standard deviation of this population

(3 marks)

(iii) Determine the critical region of this problem

(2 marks)

(iv) Test to coclude that the abrasive wear of material 1 that of material 2 by more than 2 units

(5 marks)

Q6 (a) A linear regression is a statistical technique used to find the relationships between variables for the purpose of predicting future values. By using scatter plot, one must be able to draw the line of best fit. Describe the meaning of best fit in the context of linear regression.

(2 marks)

(b) The following data (Table Q6) pertain to the chlorine residual in a swimming pool at various times after it has been treated with chemicals.

Table Q6: The chlorine residual in a swimming pool

No. of hours, x	2	4	6	8	10	12
Chlorine residual (parts per million), y	1.8	1.5	1.4	1.1	1.1	0.9

(i) Draw a scatter plot for the variables.

(4 marks)

(ii) Determine the regression line using the least squares method and interpret the result.

(10 marks)

(iii) Estimate the chlorine residual in a swimming pool when the various times after it has been treated with chemicals is 13 hours.

(4 marks)

END OF QUESTIONS

Appendix I

FINAL EXAMINATION

SEMESTER / SESSION : SEMESTER II /2014/2015
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 II /2014/2015 PROGRAM : 2BDD
COURSE : ENGINEERING STATISTICS 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	$P(X=x)=\frac{n!}{x!\cdot(n-x)!}\cdot p^x\cdot q^{n-x}={}^nC_x\cdot p^x\cdot q^{n-x}$
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	$\sigma^2=npq$

Appendix I

FINAL EXAMINATION

SEMESTER / SESSION : SEMESTER II /2014/2015 PROGRAM : 2BDD
COURSE : ENGINEERING STATISTICS COURSE CODE: BDA24103

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

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

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

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

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

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

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

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

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

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

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

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

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

Appendix I

FINAL EXAMINATION

SEMESTER / SESSION : SEMESTER II /2014/2015
COURSE : ENGINEERING STATISTICS

PROGRAM : 2BDD
COURSE CODE: BDA24103

T distribution case

$$(\bar{x}_1 - \bar{x}_2) \pm t_{\frac{\alpha}{2},v}\left(\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}\right); \qquad v = n_1 + n_2 - 2$$

For case; $n_1 \neq n_2,\ \sigma_1^2 \neq \sigma_2^2$

$$v = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_{12} - 1}}$$

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

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

Appendix I

FINAL EXAMINATION

SEMESTER / SESSION : SEMESTER II /2014/2015
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)	$n_1 = n_2 < 30$	$T_{Test} = \dfrac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{1}{n}\left(s_1^2 + s_2^2\right)}}$ $v = 2(n-1)$
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 II /2014/2015
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

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

(ii) Intercept

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

Appendix I

FINAL EXAMINATION

SEMESTER / SESSION : SEMESTER I /2014/2015
COURSE : ENGINEERING STATISTICS

PROGRAM : 2BDD
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.

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