Universiti Tun Hussein Onn Malaysia

UNIVERSITI TUN HUSSEIN ONN MALAYSIA

FINAL EXAMINATION
SEMESTER I
SESSION 2014/2015

COURSE NAME	:	CIVIL ENGINEERING STATISTICS
COURSE CODE	:	BFC34303
PROGRAMME	:	3 BFF
EXAMINATION DATE	:	DECEMBER 2014 / JANUARY 2015
DURATION	:	3 HOURS
INSTRUCTION	:	ANSWER **ALL** QUESTIONS IN **SECTION A** AND **ONLY TWO** QUESTIONS IN **SECTION B**

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

SECTION A

Q1 (a) Define the Type I and Type II errors in terms of hypothesis testing.

(2 marks)

(b) The compressive strength of concrete is being tested by a civil engineer. Eight specimens and the following data are obtained (in spi).

11.0 10.7 9.4 7.8 11.3 9.1 10.2 10.5

Test the hypothesis whether the mean compressive strength of the concrete is less than 11 spi at 0.10 level of significant.

(10 marks)

(c) A manufacturer of video display departments is testing two microcircuit designs to determine whether they produce equivalent mean current flow. Development engineering departments has the following data.

Design 1 : $n_1 = 45$ $\bar{x}_1 = 5.2$ $s_1 = 2.4$

Design 2 : $n_2 = 40$ $\bar{x}_2 = 4.8$ $s_2 = 1.2$

By using $\alpha = 0.05$, determine whether the design 1 performs better than design 2 in mean current flow. Assume that both populations are normally distributed.

(8 marks)

Q2 Medical researchers have noted that adolescent females are much more likely to deliver low birth weight babies than adult females. Because low birth weight babies have higher mortality rates, there have been a number of studies examining the relationship between birth weight and mother's age for babies born by young mother. The summary of statistics elicited from the studies are shown below:

$$n = 10 \quad \sum x = 170 \quad \sum x^2 = 2910$$
$$\sum y = 30{,}041 \quad \sum y^2 = 91{,}785{,}351 \quad \sum xy = 515{,}600$$

(a) Find the equation of the regression line.

(5 marks)

(b) Find the average birth weight of babies born by 18 year-old mothers.

(2 marks)

(c) Find the correlation coefficient, r and coefficient of determination, r^2. Interpret the results for each coefficient.

(8 marks)

(d) Test the null hypothesis $\beta_1 = 100$ against the alternative hypothesis $\beta_1 > 100$ at the 0.05 level of significance.

(5 marks)

Q3 (a) Manufactured gas plants are used to produce gas for lighting, heating and feedstock for the chemical industry. This process creates wastes that include toxic hydrogen sulfide. The amount of sulfide (in meg/g) from three independent runs produced by a gas plant is collected at random. The output of one way ANOVA is shown in Table Q3(a).

Table Q3(a): ANOVA

	Df	Sum of Squares	Mean Square Variance	F
Between Groups	2	0.011	**B**	**D**
Within Groups	**A**	0.285	**C**	
Total	14	0.296		

Assume that $\alpha = 0.05$

(i) Find the value of **A**, **B**, **C** and **D**.

(2 marks)

(ii) Test the hypothesis, is there any a difference amount of sulfide from three independent runs produce by a gas plant.

(6 marks)

(b) An experiment was carried out in a laboratory to investigate the amount of dirt (in mg) removed by detergents. They were four brands of detergent, and three samples are selected at random from each brand. The data are shown in Table Q3(b). Test the hypothesis that there is no difference of the amount of dirt removed among the four brands at the 0.05 level of significance.

Table Q3(b): Amount of dirt removed from the detergent

Brand			
A	B	C	D
11	12	18	11
13	14	16	12
17	17	20	16

(12 marks)

SECTION B

Q4 Based on continuous random variables X with probability density function given below

$$f(x) = \begin{cases} qx, & 1 < x < 3 \\ 6q - xq, & 3 < x < 5 \\ 0, & otherwise \end{cases}$$

(a) Show that value of $q = 1/8$.

(3 marks)

(b) Find $P(x > 4)$.

(3 marks)

(c) Find $E(X)$.

(3 marks)

(d) Find $E(3+4X)$.

(3 marks)

(e) Find $Var(X)$.

(5 marks)

(f) Find $Var(3+4X)$.

(3 marks)

Q5 (a) Normal distribution is a continuous random variable. (True/False)

(1 mark)

(b) In a driving license test, the score of the candidates were approximately normal distributed with mean 20 point and standard deviation 50 points. Find the probability of the candidates who received scores less than 180 points.

(3 marks)

(c) About 4.4% of motor vehicles crashed are caused by defective tires. If highway safety study begins with the random selection of 750 cases of motor vehicles crashes, estimate the probability that exactly 35 of them were caused by defective tires by using normal approximation.

(7 marks)

(d) Suppose that number of oil tankers arriving each day at certain port city is a Poisson random variable and P(X=0) = 0.05.

(i) Determine the mean of oil tankers arrive.

(4 marks)

(ii) State the Poisson distribution based on (i).

(1 mark)

(iii) Compute the probability that more than 3 oil tankers arriving each day.

(2 marks)

(iv) Find the standard deviation.

(2 marks)

Q6 (a) In a studies of water pollution in construction area, a random sample in industrial area 1 size 18 is selected from normal population with a mean of 80 and a standard deviation of 8. A second random sample in industrial area B size 10 is selected from another normal population with mean 75 and standard deviation 5. Let $\bar{X}_1$ and $\bar{X}_2$ be the two sample means. Find

(i) $P(\bar{X}_1 > 80.4)$,

(3 marks)

(ii) $P(74.9 < \bar{X}_2 < 75.3)$,

(3 marks)

(iii) $P(|\bar{X}_1 - \bar{X}_2| > 2.3)$.

(4 marks)

(b) Quality control engineer is measuring of their finishing product in two different machine. The distribution of the finishing product is normally distributed with mean and standard deviation such as below.

Table 6(b): The distribution of the finishing product

	Finishing product in Machine I	Finishing product in Machine II
Sample mean	14.2	12.3
Sample standard deviation	2.1	1.9
Sample size	33	29

Find the probability

(i) the mean of finishing product in Machine I is greater than 15,

(3 marks)

(ii) the mean of finishing product in Machine II is between 11.9 to 13,

(3 marks)

(iii) the mean of finishing product in Machine I is greater than the mean of finishing product in Machine II by 1.8.

(4 marks)

- END OF QUESTION -

STATISTICAL FORMULAE

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

$E(X) = \sum_{\forall x} xp(x)$ $\quad$ $E(X) = \int_{-\infty}^{\infty} xf(x)\ dx$

$E(X^2) = \sum_{\forall x} x^2 p(x)$ $\quad$ $E(X^2) = \int_{-\infty}^{\infty} x^2 f(x)\ dx$

$p(x) = \binom{n}{x} p^x (1-p)^{n-x},\ x = 0, 1, 2, \ldots, n$ $\quad$ $p(x) = \frac{e^{-\mu}\mu^x}{x!},\ x = 0, 1, 2, \ldots$

$X \sim N(\mu, \sigma^2)$

$\overline{X} \sim N(\mu, \frac{\sigma^2}{n})$ $\quad$ $Z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \sim N(0, 1)$

$Z \sim N(0, 1)$ $\quad$ $Z = \frac{X - \mu}{\sigma}$

$\overline{X}_1 - \overline{X}_2 \sim N(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2})$ $\quad$ $Z = \frac{\overline{X}_1 - \overline{X}_2 - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0, 1)$

$T = \frac{\overline{X} - \mu}{S/\sqrt{n}} \sim t_{\alpha, n-1}$ $\quad$ $F = \frac{S_1^2}{S_2^2} \sim f_{\alpha, n_1 - 1, n_2 - 1}$

$\chi^2 = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{\alpha, n-1}$ $\quad$ $s^2 = \frac{1}{n-1}\left(\sum x^2 - \frac{(\sum x)^2}{n}\right)$

$T = \frac{\overline{X}_1 - \overline{X}_2 - (\mu_1 - \mu_2)}{S_P\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t_{\alpha,\ n_1 + n_2 - 2}$ $\quad$ $S_P^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$

$T = \frac{\overline{X}_1 - \overline{X}_2 - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \sim t_{\alpha, v}$ $\quad$ $v = \frac{(S_1^2/n_1 + S_2^2/n_2)^2}{\frac{(S_1^2/n_1)^2}{(n_1 - 1)} + \frac{(S_2^2/n_2)^2}{(n_2 - 1)}}$ or $v = 2(n-1)$

$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$ $\quad$ $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$ $\quad$ $\text{SSE} = S_{yy} - \hat{\beta}_1 S_{xy}$ $\quad$ $\text{MSE} = \frac{\text{SSE}}{n-2}$

$\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}}$

$s_{xx} = \sum_{i=1}^{n} x_i^2 - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n}$ $\quad$ $s_{yy} = \sum_{i=1}^{n} y_i^2 - \frac{\left(\sum_{i=1}^{n} y_i\right)^2}{n}$ $\quad$ $s_{xy} = \sum_{i=1}^{n} x_i y_i - \frac{\sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n}$

$\hat{\beta}_1 \sim N\left(\beta_1, \frac{\sigma^2}{S_{xx}}\right)$ $\quad$ $T = \frac{\hat{\beta}_1 - \beta_1^*}{\sqrt{\text{MSE}/S_{xx}}} \sim t_{\alpha, n-2}$ $\quad$ $r^2 = \frac{(S_{xy})^2}{S_{xx}S_{yy}}$

$\hat{\beta}_0 \sim N\left(\beta_0, \sigma^2\left(\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}\right)\right)$ $\quad$ $T = \frac{\hat{\beta}_0 - \beta_0^*}{\sqrt{\text{MSE}(1/n + \bar{x}^2/S_{xx})}} \sim t_{\alpha, n-2}$

$SSTR(\text{between group}) = \sum r_j (\bar{x}_j - \bar{\bar{x}})^2$ $\quad$ $SSE(\text{within group}) = \sum\sum (x_{jk} - \bar{x}_j)^2$

$SST(\text{total}) = \sum_{i=1}^{r}\sum_{j=1}^{c} (x_{ij} - \bar{\bar{x}})^2$ $\quad$ $F = \frac{MSTR}{MSE} \sim f_{\alpha,(df1, df2)}$