CONFIDENTIAL

UTHM
Universiti Tun Hussein Onn Malaysia

UNIVERSITI TUN HUSSEIN ONN MALAYSIA

FINAL EXAMINATION
SEMESTER I
SESSION 2013/2014

COURSE NAME	:	CIVIL ENGINEERING STATISTICS
COURSE CODE	:	BFC 34303
PROGRAMME	:	3BFF
EXAMINATION DATE	:	DECEMBER 2013/JANUARY 2014
DURATION	:	3 HOURS
INSTRUCTION	:	ANSWER ALL QUESTIONS

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

CONFIDENTIAL

Q1 (a) Explain about the data skewness if the value of mode is greater than mean and median. Sketch the distribution of data.

(2 marks)

(b) The lifetime of EAGLE tyre is normally distributed with mean 24000 km and standard deviation 4000 km.

(i) Find the probability that the lifetime of EAGLE tyre exceeds 27000 km.

(3 marks)

(ii) If 10% of EAGLE tyres have low lifetime, find the maximum distance it can achieve.

(3 marks)

(iii) A consumer organization is comparing the lifetime of EAGLE tyre and HAWK tyre. A random sample of 36 tyres from each type is selected. The lifetime of HAWK tyre is normally distributed with mean 25000 km and standard deviation 4500 km. What is the probability that the mean lifetime of Hawk tyre is greater than EAGLE tyre?

(6 marks)

(iv) Compute a 90% confidence interval on the difference of the means lifetime the tyre lasts.

(6 marks)

Q2 (a) Ahmad is testing a power suuply used in notebook computer. The complete table of observed frequencies is as follows:

Class Interval	Observed frequencies, O_i
$x < 4.948$	12
$4.948 \leq x < 4.986$	14
$4.986 \leq x < 5.014$	12
$5.014 \leq x < 5.040$	13
$x \geq 5.040$	14

Test the hypothesis whether the output voltage is adequately described by a normal distribution with mean 5.04V and standard deviation 0.08V at a significant level $\alpha = 0.05$?

(20 marks)

Q3 (a) Define a type I error.

(1 marks)

(b) List three different names of Type I error.

(3 marks)

(c) The following data represents the time taken by two machines in producing an electrical part.

Machine	Time (in miliseconds)						
1	86	102	98	109	92		
2	81	105	97	114	92	87	104

Assuming that the distributions of the times are approximately normal,

(i) can we conclude that at $\alpha = 0.05$, the average time taken by Machine 1 is greater than 95 miliseconds.

(8 marks)

(ii) can we conclude that there is a significant difference is variability of the times in producing an electrical part by Machine 1 and Machine 2 at $\alpha = 0.10$.

(8 marks)

Q4 (a) The following data represent the time, in minutes, that Raju has to wait during 12 visits to Dr. Samy's office before being seen by him.

17	15	20	20	32	28
26	25	25	35	24	12

Use the sign test at the 0.05 level of significance to test whether Dr Samy's claim that the median waiting time for Raju is not more than 20 minutes before being admitted to the examination room is significant.

(10 marks)

(b) A company is considering four brands of light bulbs to choose from the market. Before the company decides which light bulbs to buy, they want to investigate if the mean lifetimes of the four types of light bulbs are the same. The company's research department randomly selected a few bulbs of each brands and tested them. The following results are based on the number of hours (in thousands) that each of the bulbs lasted before being burned out.

Anova: Single Factor

SUMMARY

Groups	Count	Sum	Average	Variance
Brand A	8	187	23.375	7.982143
Brand B	8	164	20.5	5.142857
Brand C	8	197	24.625	5.410714
Brand D	8	209	26.125	8.410714

ANOVA

Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	136.5938	3	c	e	0.00143	2.946685
Within Groups	188.625	**b**	**d**			
Total	a	31				

(i) Find the value of **a**, **b**, **c**, **d** and **e**.

(5 marks)

(ii) At 5% significance level, test the null hypothesis that the mean lifetime of bulbs for each of these four brands is the same.

(5 marks)

Q5 The following data are a result of an investigation as to effect of reaction temperature x on percent conversion of a chemical process y.

Temperature $(^\circ C)$	% conversion
200	43
250	78
200	69
250	73
189.65	48
260.35	78
225	65
225	74
225	76
225	79
225	83
225	81

where $\sum X^2 = 7290000$, $\sum Y^2 = 717409$ and $\sum XY = 2286900$.

(a) Fit a straight line of these data by using the method of least squares.
(4 marks)

(b) What can you infer from the estimated value of the slope?
(1 mark)

(c) Estimate the percentage of conversion if the temperature is $300\,^{o}C$.
(2 marks)

(d) Do we have sufficient evidence to conclude that there exists a linear relationship between the temperature and percentage of conversion at 0.10 level of significance?
(8 marks)

(e) Obtain the Pearson correlation coefficient for the sample data. Comment your result.
(4 marks)

- END OF QUESTION –

FINAL EXAMINATION

SEMESTER / SESSION: SEM I / 2013/2014 COURSE: 3BFF

SUBJECT : CIVIL ENGINEERING STATISTICS CODE: BFC 34303

Formula

Special Probability Distributions :

$$P(x=r)={}^{n}C_r \cdot p^r \cdot q^{n-r},\ r=0,1,\ldots,n,\ X \sim B(n,p),$$

$$P(X=r)=\frac{e^{-\mu}\cdot \mu^r}{r\,!},\ r=0,1,\ldots,\infty,\ X\sim P_0(\mu),$$

$$Z=\frac{X-\mu}{\sigma},\ Z\sim N(0,1),\ X\sim N(\mu,\sigma^2).$$

Sampling Distributions :

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

$$\overline{X}_1-\overline{X}_2\sim N\left(\mu_1-\mu_2,\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}\right).$$

Estimations :

$$n=\left(\frac{Z_{\alpha/2}\cdot\sigma}{E}\right)^2,\ \overline{x}\pm z_{\alpha/2}(\sigma/\sqrt{n}),\ \overline{x}\pm z_{\alpha/2}(s/\sqrt{n}),\ \overline{x}\pm t_{\alpha/2,v}\left(\frac{s}{\sqrt{n}}\right)$$

$$\left(\bar{x}_1-\bar{x}_2\right)-Z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}<\mu_1-\mu_2<\left(\bar{x}_1-\bar{x}_2\right)+Z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}},$$

$$\left(\bar{x}_1-\bar{x}_2\right)-Z_{\alpha/2}\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}<\mu_1-\mu_2<\left(\bar{x}_1-\bar{x}_2\right)+Z_{\alpha/2}\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}},$$

$$\left(\bar{x}_1-\bar{x}_2\right)-t_{\alpha/2,v}\cdot S_p\sqrt{\frac{2}{n}}<\mu_1-\mu_2<\left(\bar{x}_1-\bar{x}_2\right)+t_{\alpha/2,v}\cdot S_p\sqrt{\frac{2}{n}};v=2n-2$$

$$\left(\bar{x}_1-\bar{x}_2\right)-t_{\alpha/2,v}\cdot S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}<\mu_1-\mu_2<\left(\bar{x}_1-\bar{x}_2\right)+t_{\alpha/2,v}\cdot S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$$

where Pooled estimate of variance, $S_p^2 = \frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}$ with $v = n_1+n_2-2$,

$$\left(\bar{x}_1-\bar{x}_2\right)-t_{\alpha/2,v}\sqrt{\frac{1}{n}\left(s_1^2+s_2^2\right)} < \mu_1-\mu_2 < \left(\bar{x}_1-\bar{x}_2\right)+t_{\alpha/2,v}\sqrt{\frac{1}{n}\left(s_1^2+s_2^2\right)} \text{ with } v = 2(n-1),$$

$$\left(\bar{x}_1-\bar{x}_2\right)-t_{\alpha/2,v}\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}} < \mu_1-\mu_2 < \left(\bar{x}_1-\bar{x}_2\right)+t_{\alpha/2,v}\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}$$

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

$$\frac{(n-1)\cdot s^2}{\chi^2_{\alpha/2,v}} < \sigma^2 < \frac{(n-1)\cdot s^2}{\chi^2_{1-\alpha/2,v}} \text{ with } v = n-1,$$

$$\frac{s_1^2}{s_2^2}\cdot\frac{1}{f_{\alpha/2}(v_1,v_2)} < \frac{\sigma_1^2}{\sigma_2^2} < \frac{s_1^2}{s_2^2}\cdot f_{\alpha/2}(v_2,v_1) \text{ with } v_1 = n_1-1 \text{ and } v_2 = n_2-1.$$

Hypothesis Testing :

$$Z_{Test} = \frac{\bar{x}-\mu}{\sigma/\sqrt{n}},\quad Z_{Test} = \frac{\bar{x}-\mu}{s/\sqrt{n}},\quad T_{Test} = \frac{\bar{x}-\mu}{s/\sqrt{n}}$$

$$Z = \frac{(\bar{X}_1-\bar{X}_2)-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}},\quad T = \frac{(\bar{X}_1-\bar{X}_2)-(\mu_1-\mu_2)}{S_p\cdot\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \text{ with } v = n_1+n_2-2,$$

$$Z = \frac{(\bar{X}_1-\bar{X}_2)-(\mu_1-\mu_2)}{\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}},\quad T = \frac{(\bar{X}_1-\bar{X}_2)-(\mu_1-\mu_2)}{\sqrt{\frac{1}{n}\left(s_1^2+s_2^2\right)}},$$

$$T = \frac{(\bar{X}_1-\bar{X}_2)-(\mu_1-\mu_2)}{\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}} \text{ with } v = \frac{\left(\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1-1}+\frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2-1}}\ ;\ S_p^2 = \frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}\ ;$$

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2}$$

$$F = \frac{s_1^2}{s_2^2},\quad \text{with } \frac{1}{f_{\alpha/2}(v_2,v_1)} \quad \text{and} \quad f_{\alpha/2}(v_1,v_2)$$

Simple Linear Regressions :

$$S_{xy} = \sum x_i y_i - \frac{\sum x_i \cdot \sum y_i}{n},$$

$$S_{xx} = \sum x_i^{\,2} - \frac{\left(\sum x_i\right)^2}{n},$$

$$S_{yy} = \sum y_i^{\,2} - \frac{\left(\sum y_i\right)^2}{n}, \qquad \bar{x} = \frac{\sum x}{n}, \qquad \bar{y} = \frac{\sum y}{n},$$

$$\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}}, \qquad \hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}, \qquad \hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x,$$

$$r = \frac{S_{xy}}{\sqrt{S_{xx} \cdot S_{yy}}}, \qquad SSE = S_{yy} - \hat{\beta}_1 S_{xy}, \qquad MSE = \frac{SSE}{n-2},$$

$$T = \frac{\hat{\beta}_1 - \beta_1^*}{\sqrt{\frac{MSE}{S_{xx}}}} \sim t_{n-2},$$

$$T = \frac{\hat{\beta}_0 - \beta_0^*}{\sqrt{MSE\left(\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}\right)}} \sim t_{n-2}.$$

Goodness-of-fit test

$$\chi^2 = \sum_{i=1}^{k} \frac{\left(O_i - E_i\right)^2}{E_i} \sim \chi^2_{\alpha,k-p-1}$$