UNIVERSITI TUN HUSSEIN ONN MALAYSIA

FINAL EXAMINATION
SEMESTER I
SESSION 2010/2011

COURSE	:	ENGINEERING STATISTICS
CODE	:	BSM 2922
PROGRAMME	:	2 BFF, 3BFF, 4 BFF, 1 BDD, 2 BDD, 3 BDD, 4 BDD, 2 BEE, 3 BEE, 4 BEE
DATE	:	NOVEMBER/DECEMBER 2010
DURATION	:	2 HOURS 30 MINUTES
INSTRUCTION	:	ANSWER **ALL** QUESTIONS.

THIS EXAMINATION PAPER CONSISTS OF 6 PAGES

Q1 Continuous random variables X has probability density function with

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

(a) Show that the 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)

Q2 (a) It has been reported that 72% of federal government employees use e-mail. If a sample of 110 federal government employees is selected, use appropriate approximation to find the probability that

(i) not more than 90 federal government employees use e-mail.

(ii) at least 75 federal government employees use e-mail.

(7 marks)

(b) The lifetime of a wristwatch is normally distributed with the mean of lifetime is 35 months and a standard deviation of 5 month. Find the probability that the lifetime of a wristwatch is

(i) between 38 and 40 months.

(ii) less than 33 months.

(6 marks)

(c) Suppose the mean tensile strength of Type A steel is 105 ksi and the standard deviation of tensile strength is 10 ksi. For Type B steel, suppose the mean tensile strength and standard deviation of tensile strength are 100 ksi and 8 ksi, respectively. Sample of 40 specimens of Type A and 35 specimens of Type B are selected randomly. If the distribution are approximate normally distributed, what is the probability that the mean tensile strength of Type A is more than Type B? (7 marks)

Q3 A study of two kinds of photocopying equipment shows that 9 failures of equipment A took on the average 8.69 minutes to repair with a variance of 0.30 minutes, whereas 13 failures of the equipment B took on average 8.51 minutes to repair with a variance of 0.44 minutes. Assuming that the data constitute independent random samples from normal populations with unequal variance.

(a) Construct a 95% confidence interval for the difference between the true average time takes to repair failures of two kinds of photocopy.

(8 marks)

(b) Construct a 95% confidence interval for the variance for the equipment A.

(6 marks)

(c) Construct a 95% confidence interval for the ratio of the variances of the two populations sampled, σ_A^2/σ_B^2.

(6 marks)

Q4 (a) The data is about the average of mileage record by two type of engine in certain company. The sample size of engine Type I is 18 with sample mean 114. Meanwhile the sample size of engine Type II is 14 with sample mean 123. If the sample standard deviations of both engines are 1.6 and 1.7 respectively, test the hypothesis the average of mileage engine Type I is lower than the average of mileage engine Type II (use $\alpha = 0.01$). Assuming that the data are normally distributed with unknown and equal population variances.

(9 marks)

(b) Two chemical companies can supply a raw material. The concentration of a particular element in this material is important. The mean concentrations for both raw materials supply by both suppliers are the same, but we suspect that the variability in concentration may differ between the two companies. The standard deviation of concentration in a random sample of 11 batches produced by company 1 is 4.7 g/l, while for company 2, a random sample of 13 batches yields standard deviation of 5.8g/l. Is there sufficient evidence to conclude that the two population variances differ? Use $\alpha = 0.05$.

(11 marks)

Q5 It is assumed that achievement test scores should be correlated with student's classroom performance. One would expect that students who consistently perform well in the classroom (tests, quizzes, etc.) would also perform well on a standardized achievement test (0 - 100 with 100 indicating high achievement). A lecturer decides to examine this hypothesis. At the end of the academic year, she computes a correlation between the students' achievement test scores and the overall g.p.a. for each student computed over the entire year. The data for her class are given in **Table Q5**.

Table Q5: The data of students' achievement and overall g.p.a over entire year.

Students' Achievement	G.P.A.
72	3.2
95	3.6
80	2.7
70	2.8
65	2.5
40	1.9
54	1.8
67	2.4
53	1.9
77	3.0
87	3.3
90	3.7

(a) Find the value of Σx_i, Σy_i, Σx_i^2, Σy_i^2 and $\Sigma x_i y_i$.

(4 marks)

(b) Using the least square method, find the regression line that approximates the G.P.A score.

(5 marks)

(c) Calculate the sample correlation coefficient and interpret the result.

(3 marks)

(d) Calculate the value of SSE and MSE.

(2 marks)

(e) Test the null hypothesis $\beta_1 = 2.0$ against the alternative hypothesis $\beta_1 > 2.0$ at the 0.05 level of significance.

(6 marks)

FINAL EXAMINATION

SEMESTER / SESSION: SEM I / 2010/2011

COURSE: 2 BFF, 3BFF, 4 BFF, 1 BDD, 2 BDD, 3 BDD, 4 BDD, 2 BEE, 3 BEE, 4 BEE

SUBJECT : ENGINEERING STATISTICS

CODE: BSM 2922

Formulae

Random variables:

$$\sum_{i=-\infty}^{\infty} P(x_i) = 1, \quad E(X) = \sum_{\forall x} x \cdot P(x), \quad E(X^2) = \sum_{\forall x} x^2 \cdot P(x), \quad \int_{-\infty}^{\infty} f(x)\,dx = 1,$$

$$E(X) = \int_{-\infty}^{\infty} x \cdot P(x)\,dx, \quad E(X^2) = \int_{-\infty}^{\infty} x^2 \cdot P(x)\,dx, \quad Var(X) = E(X^2) - [E(X)]^2.$$

Special Probability Distributions :

$$P(x=r) = {}^nC_r \cdot p^r \cdot q^{n-r},\ r = 0, 1, \dots, n,\ X \sim B(n, p),\ P(X=r) = \frac{e^{-\mu} \cdot \mu^r}{r!},\ r = 0, 1, \dots, \infty,$$

$$X \sim P_0(\mu),\ Z = \frac{X-\mu}{\sigma},\ Z \sim N(0,1),\ X \sim N(\mu, \sigma^2).$$

Sampling Distributions :

$$\bar{X} \sim N(\mu, \sigma^2/n),\ Z = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1),\ T = \frac{\bar{x}-\mu}{s/\sqrt{n}},\ \bar{X}_1 - \bar{X}_2 \sim N\left(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right).$$

Estimations :

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

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

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

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

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 $v = n_1 + n_2 - 2$,

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

$$\left(\bar{x}_1 - \bar{x}_2\right) - t_{\alpha/2,v}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} < \mu_1 - \mu_2 < \left(\bar{x}_1 - \bar{x}_2\right) + t_{\alpha/2,v}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} \text{ with } v = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1-1} + \frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2-1}},$$

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

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

Hypothesis Testings :

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

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

$$v = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2 - 1}}.\ ;\ S_p^2 = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1 + n_2 - 2}\ ;\ \chi^2 = \frac{(n-1)s^2}{\sigma^2}$$

$$\frac{s_1^2}{s_2^2}\cdot\frac{1}{f_{\alpha/2}(v_2,v_1)} < \frac{\sigma_1^2}{\sigma_2^2} < \frac{s_1^2}{s_2^2}\cdot f_{\alpha/2}(v_1,v_2)$$

Simple Linear Regressions :

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

$$\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}},\ \hat{\beta}_0 = \bar{y} - \hat{\beta}_1\bar{x},\ \hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x,\ r = \frac{S_{xy}}{\sqrt{S_{xx}\cdot S_{yy}}},\ SSE = S_{yy} - \hat{\beta}_1 S_{xy},\ MSE = \frac{SSE}{n-2},$$

$$T = \frac{\hat{\beta}_1 - \beta_1^*}{\sqrt{\frac{MSE}{S_{xx}}}} \sim t_{n-2},\ T = \frac{\hat{\beta}_0 - \beta_0^*}{\sqrt{MSE\left(\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}\right)}} \sim t_{n-2}.$$