UNIVERSITI TUN HUSSEIN ONN MALAYSIA

FINAL EXAMINATION
SEMESTER I
SESSION 2009/2010

SUBJECT	:	STATISTICS
CODE	:	BSM 1413
COURSE	:	1 BIT / 2 BIT
DATE	:	NOVEMBER 2009
DURATION	:	3 HOURS
INSTRUCTION	:	ANSWER ALL QUESTIONS IN **PART A** AND **FOUR (4)** QUESTIONS IN **PART B**

THIS EXAMINATION PAPER CONSISTS OF 7 PAGES

PART A

Q1 The following data in **Table Q1** is the sales record of heating fuel (thousands of litres) by a distributor over ten weeks along with the average temperature (^oC) for each week.

Table Q1 : Sales of Heating Fuel

Week	1	2	3	4	5	6	7	8	9	10
Heating fuel sales	26	17	7	12	30	40	20	15	10	5
Average temperature	4	10	14	12	4	5	8	11	13	15

(a) Sketch a scatter plot for the above data.

(2 marks)

(b) What is the degree of correlation between fuel sales and temperature? Interpret the results.

(3 marks)

(c) Compute the best fitting regression equation between fuel sales and temperature, assuming that fuel sales are dependent on temperature.

(7 marks)

(d) Test the alternative hypothesis of $\beta_1 \neq 2$ at the 0.05 level of significance.

(8 marks)

Q2 (a) A time and motion study is conducted to test whether the mean length of time required to perform a certain task is the same for the employees on the day shift and the employees on the night shift. The data given are as follow, $n_1 = 10, \bar{x}_1 = 26, s_1^2 = 64, \quad n_2 = 8, \bar{x}_2 = 29, s_2^2 = 50$. Test the claim that the mean length of time required to perform a certain task is the same for the employees on the day shift and the employees on the night shift. Use 5% level of significance with the assumption that the populations are approximately normal and the population variances are equal.

(8 marks)

(b) Two methods of filling cereal box produce the same package fill weight on the average. The second method is faster, but some people suspect that this method produces greater variances in fill amount than the first method does. A random sample of 25 packages filled by method 1 is obtained, and the observed sample variance is $s_1^2 = 0.4$. A random sample of 25 packages filled by method 2 yields a sample variance of $s_2^2 = 0.6$.

(i) Some people claim that the variance for method 1 is more than 0.57. At $\alpha = 0.05$, is there enough evidence to reject the claim?

(ii) Is there enough evidence to support the claim that the variation for method 1 is less than method 2? Use $\alpha = 0.05$ level of significance.

(12 marks)

PART B

Q3 A study of the mass (to nearest 0.1 kg) for 45 newborn babies is shown in the following **Table Q3.**

Table Q3 : Mass of 45 Newborn Babies

3.1	3.4	2.3	2.5	3.2	3.6	2.7	3.8	3.6	3.4	2.2	2.3	3.4	3.3	3.8
2.6	2.4	2.7	2.3	3.1	3.0	3.2	3.3	3.4	2.5	3.2	3.4	3.6	3.3	3.1
2.1	3.2	3.4	3.5	3.4	2.4	2.7	2.9	3.0	3.1	3.8	3.5	2.2	3.2	2.9

(a) Complete the table given below.

Mass (kg)	f	X	fx	lower boundary	x^2	fx^2
2.0 – 2.2						
2.3 – 2.5						
2.6 – 2.8						
2.9 – 3.1						
3.2 – 3.4						
3.5 – 3.7						
3.8 – 4.0						

(3 marks)

(b) Find the
(i) mean.
(ii) median.
(iii) mode.
(iv) standard deviation.

(12 marks)

Q4 (a) The probability distribution function for the random variables X is given as below:

$$f(x) = \begin{cases} \dfrac{(7-x)}{k}, & x = 0,2,4,6 \\ 0, & \text{otherwise} \end{cases}$$

Find the
(i) value of k.
(ii) values of $E(X)$ and $E(X^2)$.
(iii) value of $V(4X + 1)$.

(9 marks)

(b) The length of time X in seconds has the probability density function as below :

$$f(x) = \begin{cases} 5(1-x)^4, & 0 < x < 1 \\ 0, & \text{otherwise} \end{cases}$$

(i) Find $P\left(\dfrac{1}{2} < X < \dfrac{3}{4}\right)$.

(ii) Find the cumulative distribution, $F(x)$.

(6 marks)

Q5 (a) Assuming that the probability of giving birth to a baby boy is 0.15, find the probability that in a family of five children

(i) at least two are boys.

(ii) exactly two are girls.

(5 marks)

(b) The number of customers who enter a service station to buy gas in any five minute period has the Poisson distribution with mean 0.3. Find the probability that during

(i) a five minute period, three customers enter the station.

(ii) an hour, at least three customers enter the station.

(5 marks)

(c) In a large class, suppose your instructor tells you that you need to obtain a good mark in order the top 10% of your class to get an A in a particular exam. From past experience she is able to estimate that the mean and standard deviation on this exam will be 72 and 13, respectively, by assuming that the marks will be approximately normal distributed. What will be the minimum mark needed to obtain an A?

(5 marks)

Q6 (a) Give four ways to select sample from population.

(2 marks)

(b) The customers who rent compact cars from the Weera Car Rental Agency travel 400 miles on the average. The population standard deviation is $\sigma = 70$. A random sample of 64 customers is checked during a certain time period. Find the probability that the mean customers travel

(i) exceeds 420 miles.

(ii) fewer than 390 miles.

(iii) between 380 to 410 miles.

(7 marks)

(c) The mean annual income of union carpenters is RM 1000 higher than the mean annual income of nonunion carpenters. For each population, the standard deviation is RM 3000.

(i) Random samples of 100 union members and 200 nonunion members are taken. What is the probability that the mean annual income for union members at least RM1050 more than nonunion members?

(ii) What is the probability that the sample of size 76 union members is at least RM1020 more than the sample of size 86 nonunion members?

(6 marks)

Q7 (a) The data in **Table Q7 (a)** below show the typing rates (words per minute) achieved by each of ten secretaries working in two branches of an insurance company which used two different types of word processing programs.

Table Q7 (a) : Typing Rates

Branch 1	64	55	58	55	52	68	72	60	72	49
Branch 2	57	64	55	64	60	66	75	82	48	64

Construct a 90% confidence interval for the difference between the two branches by assuming that the variances are unknown but equal.

(9 marks)

(b) In a study of computer science majors, the data in **Table Q7(b)** were obtained on two groups, those who left their profession within a few months after graduation (leaved) and those who remained in their profession after they graduated (stayed).

Table Q7 (b) : Summary on Profession

Leaved	Stayed
$\bar{x}_1 = 3.16$	$\bar{x}_2 = 3.28$
$s_1 = 0.52$	$s_2 = 0.46$
$n_1 = 13$	$n_2 = 11$

Compute a 98% confidence interval for the ratio population variances, $\dfrac{\sigma^2_{\text{leaved}}}{\sigma^2_{\text{stayed}}}$.

(6 marks)

FINAL EXAMINATION

SEMESTER / SESSION: SEM I / 2009/2010 | COURSE : 2 BIT / 1 BIT

SUBJECT : STATISTICS | CODE : BSM 1413

Formulae

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

$$r=\frac{S_{xy}}{\sqrt{S_{xx}S_{yy}}},\quad \hat{\beta}_1=\frac{S_{xy}}{S_{xx}},\quad \hat{\beta}_0=\bar{y}-\hat{\beta}_1\bar{x},$$

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

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

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

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

$$\chi^2=\frac{(n-1)s^2}{\sigma_0^2},\quad F=\frac{S_1^2}{S_2^2}$$

$$\bar{x}=\frac{\sum_{i=1}^{n} f_i x_i}{\sum_{i=1}^{n} f_i},\ M=L_M+c\left(\frac{n/2-F}{f_m}\right),\ M_0=L+c\left(\frac{d_B}{d_B+d_A}\right),\ s^2=\frac{1}{\Sigma f-1}\left[\sum_{i=1}^{n} f_i x_i^2-\frac{(\Sigma f_i x_i)^2}{\Sigma f}\right]$$

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

FINAL EXAMINATION

SEMESTER / SESSION: SEM I / 2009/2010 COURSE : 2 BIT / 1 BIT

SUBJECT : STATISTICS CODE : BSM 1413

Formulae

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

$$r = \frac{S_{xy}}{\sqrt{S_{xx} S_{yy}}},\quad \hat{\beta}_1 = \frac{S_{xy}}{S_{xx}},\quad \hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x},$$

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

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

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

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

$$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2},\quad F = \frac{S_1^2}{S_2^2}$$

$$\bar{x} = \frac{\sum_{i=1}^{n} f_i x_i}{\sum_{i=1}^{n} f_i},\ M = L_M + c\left(\frac{n/2 - F}{f_m}\right),\ M_0 = L + c\left(\frac{d_B}{d_B + d_A}\right),\ s^2 = \frac{1}{\Sigma f - 1}\left[\sum_{i=1}^{n} f_i x_i^2 - \frac{(\Sigma f_i x_i)^2}{\Sigma f}\right]$$

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

FINAL EXAMINATION

SEMESTER / SESSION: SEM I / 2009/2010

COURSE : 2 BIT / 1 BIT

SUBJECT : STATISTICS

CODE : BSM 1413

$P(x) = \binom{n}{x} \cdot p^x \cdot (1-p)^{n-x} \quad x = 0, 1, \ldots, n$, $P(X = r) = \frac{e^{-\mu} \cdot \mu^r}{r!} \quad r = 0, 1, \ldots, \infty$,

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

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

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

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

$$(\bar{x}_1 - \bar{x}_2) - t_{\alpha/2, v}\, S_P \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} < \mu_1 - \mu_2 < (\bar{x}_1 - \bar{x}_2) + t_{\alpha/2, v}\, S_P \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \text{ where}$$

$$S_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} \text{ and } v = n_1 + n_2 - 2,$$

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

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

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

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