KOLEJ UNIVERSITI TEKNOLOGI TUN HUSSEIN ONN

PEPERIKSAAN AKHIR
SEMESTER I
SESI 2006/2007

NAMA MATA PELAJARAN	:	STATISTIK
KOD MATA PELAJARAN	:	DSM 2932
KURSUS	:	3 DET/3 DEE/3 DEX/3 DDT
TARIKH PEPERIKSAAN	:	NOVEMBER 2006
JANGKA MASA	:	2 JAM 30 MINIT
ARAHAN	:	JAWAB **LIMA(5)** SOALAN DARIPADA **ENAM(6)** SOALAN.

KERTAS SOALAN INI MENGANDUNGI 6 MUKA SURAT

Q1 (a) The following is cumulative distribution function for **discrete** random variable, X.

$$F(x)=\begin{cases}0, & x<0\\ \dfrac{1}{291}\left(2x^4+x\right), & 0\le x<5\\ 1, & x\ge 5.\end{cases}$$

Find

(i) $P(X=3)$,

(ii) $P(X>3)$,

(iii) $P(1\le X\le 2)$.

(10 marks)

(b) Let random variable X have the continuous probability density function defined

$$f(x)=\begin{cases}0.5x-1, & 2<x<4\\ 0, & otherwise.\end{cases}$$

Find the expected value and variance of X.

(10 marks)

Q2 (a) The health department has a problem with dengue cases at Taman Perling Utama. In a week, the department receives an average of 12 reported cases. If this problem still occurs, find the probability that,

(i) exactly 5 reported cases received within 1 week,

(ii) less than 22 reported cases received within 2 weeks.

(7 marks)

(b) In a driving license test, the score of the candidates were approximately normal distributed with a mean, μ=200 points and standard deviation, σ=50 points. Find the probability of the candidates who received the following scores;

(i) less than 180 points,

(ii) between 170 and 180 points.

(8 marks)

c) An IT company in Batu Pahat had found out that the probability of cartridges failed to be sold is 10%. By consider the distribution as binomial, 200 cartridges were selected at random by the customers at the show room. Find the probability that less than 30 cartridges failed to be sold.

(5 marks)

Q3 (a) Given X the random variable be the height of building with mean of 89.5m and variance of 19.6m. If 28 building are selected randomly from this population, calculate

(i) the mean and variance of the sampling distribution of $\overline{X}$,

(ii) the probability of sample means falling less than 87m.

(6 marks)

(b) Azfar Corporation manufactures light bulbs where the life of bulb is approximately normal distributed. The CEO claims that an average Azfar light bulb lasts 300 days, with a standard deviation of 50 days. If the CEO's claim was true, what is the probability that

(i) 15 randomly selected bulbs would have an average life of not more than 290 days?

(ii) 20 randomly selected bulbs would have an average life between 280 to 290 days?

(14 marks)

Q4 (a) To investigate the air pollution index at KUiTTHO, two independent sampling stations A and B were built. The 10 monthly samples collected from station A showed that the mean index is 3.22 with the standard deviation is 0.7, while the 8 monthly samples collected from station B showed that the mean index is 2.56 with the standard deviation is 0.4. Find the 95% confidence interval for the difference between two populations means index for both stations. Assume that the populations are approximately normal distributed with equal variances.

(10 marks)

(b) The data below shows the weight (in gram) of 10 samples of instant noodles that are run to test the new noodles-cutting machine.

80.0 81.2 80.0 79.8 80.9 79.7 81.0 80.0 78.0 80.5

Find the 90% confidence interval for the variance of the weight of the instant noodles within the samples. Assume that the population is normally distributed.

(10 marks)

Q5 (a) A large manufacturing firm is being charged with discrimination in its hiring practices.

(i) What hypothesis is being tested if a jury commits a type I error by finding the firm guilty ?

(ii) What hypothesis is being tested if a jury commits a type II error by finding the firm guilty ?

(2 marks)

(b) A researcher measures the self-esteem scores of a sample of statistics students, reasoning that their frustration with this course may lower their self-esteem to that of the typical college student. He takes a sample of 9 students and finds them to have a mean self-esteem score of 35.1. Given that the population mean is 55 and population standard deviation is 11.35, test your hypothesis at the 0.05 significance level.

(9 marks)

(c) Two different lighting techniques are compared by measuring the intensity of light at selected locations in areas lighted by the two methods. If 15 measurements in the first area had a standard deviation of 2.7 foot-candles and 21 measurements in the second area had a standard deviation of 4.2 foot-candles, can it be conclude that the lighting in the second area is less uniform ? Use a 0.01 level of significance.

(9 marks)

Q6 The data is about the profit margin of the one mini market in Parit Raja after they make their sales within 12 months last year. Assume that the profit margin has a linear relationship with the sales margin.

Table Q6: Sales margin versus profit margin

Sales (*X*)	Profit (*Y*)
8500	5200
7800	5000
7100	3000
9400	5100
10200	7000
11800	6300
11900	8000
15000	10000
14100	9000
7000	3500
13000	7300
12500	6400

(a) Plot the scatter diagram and give your explanation.
(4 marks)

(b) Fit the linear regression equation of the data.
(5 marks)

(c) Estimate the profit margin if the sales is RM11400.
(2 marks)

(d) Test the slope whether it is equal to 0.7 at 5% level of significance.
(9 marks)

STATISTICAL FORMULA

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

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

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

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

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

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

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

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

$$T = \frac{\overline{X} - \mu}{S/\sqrt{n}} \sim t_{\alpha, n-1} \qquad F = \frac{\sigma_2^2 S_1^2}{\sigma_1^2 S_2^2} \sim f_{\alpha, n_1-1, n_2-1}$$

$$\chi^2 = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{\alpha, n-1}$$

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

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

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

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

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

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