KOLEJ UNIVERSITI TEKNOLOGI TUN HUSSEIN ONN

PEPERIKSAAN AKHIR
SEMESTER I
SESI 2004/2005

NAMA MATA PELAJARAN	:	STATISTIK
KOD MATA PELAJARAN	:	BSM 2742/ BSM 2622
KURSUS	:	2BTM / 3BTA / 3BTE / 4BTM
TARIKH PEPERIKSAAN	:	OKTOBER 2004
JANGKA MASA	:	3 JAM
ARAHAN	:	JAWAB **SEMUA SOALAN** DARI BAHAGIAN A DAN **TIGA** (3) SOALAN SAHAJA DARI BAHAGIAN B.

KERTAS SOALAN INI MENGANDUNGI **6** MUKA SURAT

PART A

Q1 (a) *Simple Linear Regression* is a statistical technique for investing and modeling the linear relationship between variables. The model is represented by a linear equation $y = \beta_0 + \beta_1 x + \varepsilon$.

(i) Write down the formula for the regression coefficients $\hat{\beta}_0$ and $\hat{\beta}_1$ so that the regression equation can be obtained.

(7 marks)

(ii) What is the role of $\hat{\beta}_1$ in the equation? (1 mark)

(b) The manager of a car plant wishes to investigate how the plant's electricity usage upon the plant's production. The **SPSS** output for this problem is shown in **Figure Q1**, which provides the plant's production and electrical usage for each month of the previous years.

Refer to the given **SPSS** output and answer the following questions.

(i) From the graph in **Figure Q1**, identify whether equation is linear or quadratic. Give your reason.

(2 marks)

(ii) State that the coefficient of determination R^2 and the correlation coefficient r.

(4 marks)

(iii) Give the value of $\hat{\beta}_0$ and $\hat{\beta}_1$.

(4 marks)

(iv) Estimate the regression line.

(3 marks)

(v) If a production level of RM5.5 million worth of cars is planned for next month, estimate for the electrical usage.

(4 marks)

SPSS OUTPUT:

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.896(a)	.802	.782	.31051

a Predictors: (Constant), electricity usage (kWh)

ANOVA(b)

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	3.908	1	3.908	40.533	.000(a)
	Residual	.964	10	.096		
	Total	4.872	11			

a Predictors: (Constant), electricity usage (kWh)
b Dependent Variable: production

Coefficients(a)

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	.309	.724		.427	.679
	electricity usage (kWh)	1.608	.253	.896	6.367	.000

a Dependent Variable: production

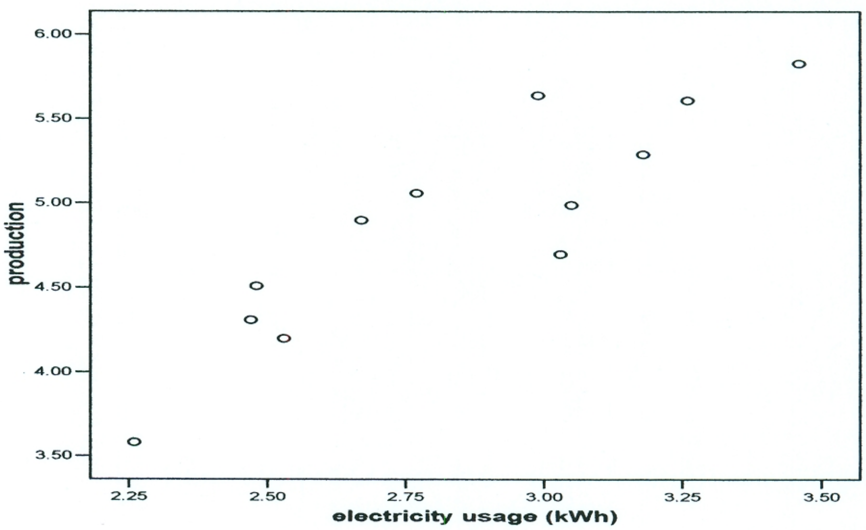

Figure Q1

PART B

Q2 (a) In an automobile manufacturing, there are 0.004 flaws per square meter of plastic panel that used in the interior of automobiles. Assume an automobile interior contains 10 square meters of plastic panel. If 10 cars are sold to a rental company, what is the probability that at most one car has that surface flaws?

(7 marks)

(b) The probability that your call to a service line is answered in less than 30 seconds is 0.08. Assume that your calls are independent. If you call 30 times, what is the probability that less than 9 of your calls are answered within 30 seconds? Use the appropriate approximation.

(9 marks)

(c) The number of taxis on average that arrive at a hotel in each period of 10 minutes is 6. Find the probability that at least two taxis will arrive in a given 15-minute period.

(9 marks)

Q3 (a) The distribution of height of a certain breed of terrier dogs has a mean height of 72 cm and standard deviation of 10 cm, whereas the distribution of heights of a certain breed of poodles has a mean height of 28 cm and standard deviation of 5 cm. Assuming that the sample means can be measured to any degree of accuracy, find the probability that the sample mean for a random sample of heights of 64 terriers exceeds the sample mean for a random sample of heights of 100 poodles by at most 44.2 cm.

(12 marks)

(b) A researcher found that 7% of all football helmets could potentially cause injury to the wearer. Find the probability that a randomly selected sample of 220 football helmets could cause injury to at least 15 players.

(13 marks)

Q4 (a) In estimating the mean IQ score for the population of statistics professors, we are 95% confidence that the sample mean $\bar{x}$ is within 2 IQ points of the population mean μ. Use $\sigma = 15$ and $n = 217$, show that the difference between $\bar{x}$ and μ is 2.

(5 marks)

(b) Assume that a random sample is selected to estimate the mean IQ score of statistics professors. The sample with size of $n = 315$ give the mean $\bar{x} = 103$ and the standard deviation $s = 17$. Construct a 99% confidence interval for the mean IQ score of statistics professors.

(10 marks)

(c) A comparison of the mean IQ scores between 2 groups of statistics professors has been studied. The results are shown below.

Group I	**Group II**
$\bar{x}_1 = 98$	$\bar{x}_2 = 102$
$s_1 = 16.5$	$s_2 = 14$
$n_1 = 13$	$n_2 = 11$

Given that the population variances of the Group I and Group II statistics professors are σ_1^2 and σ_2^2, respectively. Construct a 95% confidence interval for the ratio of σ_1^2/σ_2^2.

(10 marks)

Q5 (a) A study of seat-belt use involved children who were hospitalized caused by motor vehicle crashes has been done. For a group of 290 children who were not wearing seat belts, the numbers of days spent in intensive care units (ICUs) have a mean of 1.39 and a standard deviation of 3.06. For a group of 123 children who were wearing seat belts, the numbers of days in ICU have a mean of 0.83 and a standard deviation of 1.77. Using a significance level $\alpha = 0.02$, is there sufficient evidence to support the claim that the population of children not wearing seat belts has a higher mean number of days spent in ICU?

(Hint: Test the hypothesis: H_0: $\mu_1 - \mu_2 = 0$ versus H_1: $\mu_1 - \mu_2 > 0$)

(9 marks)

(b) Researchers studied crashes of general aviation (noncommercial and nonmilitary) airplanes and found that pilots died in 437 of 8411 crash landings. Use a significance level $\alpha = 0.06$ to test the claim that pilots die is not equal to 5% of such crashes.

(Hint: Test the hypothesis: H_0: $p = 0.05$ versus H_1: $p \neq 0.05$)

(8 marks)

(c) A pharmaceutical company uses a machine to pour cold medicine into a random sample of 25 bottles with the standard deviation of the weights 0.12mg. A machine company claims that a new machine can fills bottles with a lower variation. It gives a standard deviation of the weight 0.19mg for a random sample of 25 bottles. At the significance level $\alpha = 0.05$, test the claim made by the machine company.

(Hint: Test the hypothesis: H_0: $\frac{\sigma_1^2}{\sigma_2^2} = 1$ versus H_1: $\frac{\sigma_1^2}{\sigma_2^2} > 1$)

(8 marks)

STATISTICS FORMULAS

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

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

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

$$p(x) = \frac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}},\ x = 0, 1, 2, \ldots, k$$

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

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

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

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

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

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

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

$$\hat{P} \sim N(p, \frac{pq}{n}) \qquad Z = \frac{\hat{P}-p}{\sqrt{pq/n}} \sim N(0, 1)$$

$$\hat{P}_1 - \hat{P}_2 \sim N(p_1 - p_2,\ \hat{p}\hat{q}\ [1/n_1 + 1/n_2]) \qquad Z = \frac{(\hat{P}_1 - \hat{P}_2) - (p_1 - p_2)}{\sqrt{\hat{p}\hat{q}[1/n_1 + 1/n_2]}} \sim N(0, 1)$$

$$\hat{p} = \frac{x_1 + x_2}{n_1 + n_2}$$