PPS.

KOLEJ UNIVERSITI TEKNOLOGI TUN HUSSEIN ONN

PEPERIKSAAN AKHIR SEMESTER II SESI 2004/2005

NAMA MATA PELAJARAN : STATISTIK

KOD MATA PELAJARAN : DSM 2932

KURSUS : 2DDT

TARIKH PEPERIKSAAN : MAC 2005

JANGKA MASA : 2 ½ JAM

ARAHAN : JAWAB **SOALAN** DARI BAHAGIAN A DAN **EMPAT (3) SOALAN** DARI BAHAGIAN B.

KERTAS SOALAN INI MENGANDUNGI 6 MUKA SURAT

Q4 This is a sample data for the number of claims for an insurance company base on their customer ages with some values missing.

Age	f	x	fx	F	x^2	fx^2
33-37	2	35	70	2	1225	2450
38-42	3	……	120	5	1600	4800
43-47	3	45	135	8	2025	6075
……	7	50	350	……	2500	17500
53-57	5	55	275	20	3025	15125
58-62	7	60	……	27	3600	……
63-67	3	65	195	30	4225	12675
Total	……		1565			83825

(a) Copy the table and insert the missing values.

(6 marks)

(b) Draw a histogram from the sample data.

(3 marks)

(c) Calculate the median of the sample data.

(2 marks)

(d) Find mean and variance of the sample data.

(9 marks)

Q5 (a) Given S is a sample space of events of number with each number having equal probability of being selected.

$$S = \{2, 3, 5, 7, 11, 12, 13, 16, 17, 21, 23, 26, 33, 37, 41\}$$

Let A, B, C and D are events of the sample space, defined as

$A = \{7, 16, 21, 37, 41\}$

$B = \{2, 3, 21, 37\}$

$C = \{2, 11, 12, 23, 37\}$

$D = \{5, 17\}$

(i) Draw a Venn diagram showing all events.

(ii) Calculate the probability of $P(B \cap C)$, $P(A \cap B \cap C)'$, $P(A \mid C)$.

(10 marks)

(b) In the production of electrical components of type A, type B and type C, 40% of the production is from type A, 20% from type B and the others from type C. The probability of defective product for the type A, B and C are 0.04, 0.02 and 0.05 respectively.

(i) Draw a Venn diagram for the events above.

(ii) Find the probability that the components are non-defective.

(iii) Find the probability that the component is type B given that it is defective product.

(10 marks)

PART A

Q1 The heights of KUiTTHO students may be approximated by a normal distribution $N(\mu,\sigma^2)$. A random sample of 12 students are taken in order to estimate the value of μ and σ. The results of this sampling show a mean of 174.5 cm and a standard deviation 6.9 cm.

(a) What is the best statistical estimate for μ and σ. (4 marks)

(b) Calculate the value of $\hat{\mu}$ and $\hat{\sigma}$. (2 marks)

(c) Obtain a 98% for the two sided confidence interval on mean height, μ. (7 marks)

(d) Obtain a 99% for the two sided confidence interval on standard deviation of students height σ. (7 marks)

Q2 A manufacturer develops a new and cheaper brick whose mean compressive strength exceeds the standard μ_o = 2500 psi. To test the manufacturers' claim is right, a random sample of size 9 were taken and their compressive strengths are measured. The following results were $\bar{x}$ = 2600 psi and s = 75 psi. Assume the compressive strengths follow a normal distribution $N(\mu,\sigma^2)$.

(a) Write the manufacturers claim in null and alternative hypothesis. (4 marks)

(b) Name the most appropriate statistical test. (4 marks)

(c) Describe the rejection region at 5 percent level. (4 marks)

(d) What your conclusion on manufacturers claim. (8 marks)

PART B

Q3 The *Lembaga Kemajuan Ikan Malaysia* had made a survey on the price of 'ikan kembung' in a sample of size 16 at fish outlets near the Klang Valley. Their output prices are shown below.

Prices in RM per kilogram							
8.90	6.90	7.80	7.60	8.50	6.40	5.90	7.30
7.30	7.40	8.20	6.70	7.30	8.60	7.20	6.50

Find,

(a) the mode, median and mean of the prices of this species of fish. (6 marks)

(b) the deciles of D_3 and D_8 . (5 marks)

(c) the percentiles of P_{24} and P_{45}. (5 marks)

(d) the variance of the data. (4 marks)

Q6 A random variable X has a probability density function f as follows,

$$f(x)=\begin{cases} a-5+x & 5\le x\le 6 \\ 0 & others \end{cases}$$

(a) Show that $a=\dfrac{1}{2}$

(4 marks)

(b) Determine the expected value $E(X)$.

(4 marks)

(c) Obtain the cumulative distribution function, F.

(6 marks)

(d) Find the probability $F(5.2)$, $P(X>5.5)$, $P(5\le X\le 5.5)$

(6 marks)

Q7 (a) The average number of claims per hour made to the Takaful Nasional Sdn. Bhd. for damages or losses incurred in any accidents is 1.5. Using Poisson distribution, what is the probability that,

(i) in any given hour, exactly 2 claims will be made?

(ii) in 3 hours, 3 or more claims will be made?

(7 marks)

(b) When a customer places an order with Ahmad Supplies, a computerized accounting information system automatically checks if the customer has exceeded his or her credit limit. Past records indicate that the probability of customers exceeding their credit limit is 0.09. Assume that the number of customers who are detected as having exceeded their credit limits is distributed as a binomial random variable.

(i) Suppose that on a given day, 20 customers place the orders. Find the probability that no customer will exceed their credit limit.

(ii) If in a certain day, 100 customers place orders. Find the probability that more than ten customers will exceed their credit limit by using binomial distribution and appropriate approximation distribution.

(13 marks)

Q8 (a) The amounts of electric bills for all households in a city have normal distribution with a mean of RM 80.00 and a standard deviation of RM 25.00. Find the probability that the mean amount of electric bills for a random sample of 75 households selected at random from this city will be

(i) less than RM 76.00

(ii) between RM 72.00 and RM 77.00

(iii) greater than RM 85.00

(3 marks)

(b) The distribution X of heights of a certain breed of hill goats has a mean height of 72 centimeters and a standard deviation of 10 centimeters, whereas the distribution Y of heights of a certain breed of sheep has a means height of 68 centimeters with a standard deviation 5 centimeters. A random sample of height of 64 hill goats and a sample of 100 sheep were taken for surveying.

(i) What is the mean sampling distribution of hill type goat, $\overline{X}$

(ii) What is the different mean sampling distribution these two types goats, $\overline{X}-\overline{Y}$

(iii) find the probability that the sample mean for the hill type goats between 70.0 centimeter and 74.2 centimeters

(iv) Calculate the probability $P(\overline{X}-\overline{Y}\geq 8)$

(11 marks)

STATISTICAL FORMULA

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

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

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

$$p(x) = \frac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}},\ x = 0, 1, 2, \ldots, k \qquad X \sim N(\mu, \sigma^2) \qquad Z \sim N(0, 1) \qquad Z = \frac{X-\mu}{\sigma}$$

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

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

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

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

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

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

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

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

BAHAGIAN A

S1 Ketinggian pelajar KUiTTHO boleh diandaikan tertabur normal distribution $N(\mu,\sigma^2)$. Satu sampel rawak sebanyak12 pelajar diambil untuk tujuan mendapatkan nilai anggaran bagi μ dan σ. Hasil daripada pensampelan ini menunjukkan bahawa min ialah 174.5 cm dan sisihan piawainya 6.9 cm.

(a) Apakah taburan anggaran terbaik bagi μ dan σ.

(4 markah)

(b) Kira nilai bagi $\hat{\mu}$ dan $\hat{\sigma}$.

(2 markah)

(c) Dapatkan 98% selang keyakinan dua pihak bagi min tinggi, μ.

(7 markah)

(d) Dapatkan 99% selang keyakinan dua pihak bagi sisihan piawai σ.

(7 markah)

S2 Satu pengeluar kilang mencipta bata baru dengan kos yang rendah dengan min kekuatan kompresif melebihi kekuatan piawai μ_o = 2500 psi. Untuk menuji bahawa dakwaan kilang ini betul, satu sampel rawak bersaiz 9 diambil dan kekuatan kompresif disukat. Berikut adalah hasilnya iaitu $\bar{x}$ = 2600 psi dan s = 75. Anggap bahawa kekekuatan kompresif mengikuti taburan normal $N(\mu,\sigma^2)$

(a) Tulis dakwan pengeluar kilang dalam hipotesis nul dan alternative..

(4 markah)

(b) Namakan ujian statistik yang yang digunakan

(4 markah)

(c) Nyatakan rantau penolakan bila paras keertian ialah 5 %.

(4 markah)

(d) Apakah pendapat anda dengan dakwaan pengeluar lilang ini.

(8 markah)

BAHAGIAN B

S3 Lembaga Kemajuan Ikan Malaysia telah membuat tinjauan untuk harga ikan kembung dengan sampel bersaiz 16 di pasar-pasar sekitar Lembah Klang. Harga-harga ikan adalah seperti di bawah ini.

RM per kilogram							
8.90	6.90	7.80	7.60	8.50	6.40	5.90	7.30
7.30	7.40	8.20	6.70	7.30	8.60	7.20	6.50

Cari,

(a) mod, median dan min bagi harga ikan kembung ini.

(6 markah)

(b) desil bagi D_3.

(5 markah)

(c) persentil bagi P_{24}

(5 markah)

(d) varians untuk data ini

(4 markah)

S4 Jadual dibawah menunjukkan sampel data bagi bilangan tuntutan untuk sebuah syarikat insurans. Data ini menunjukkan umur pelanggan dengan beberapa nilai yang tertinggal.

Umur	f	x	fx	F	x^2	fx^2
33-37	2	35	70	2	1225	2450
38-42	3	......	120	5	1600	4800
43-47	3	45	135	8	2025	6075
......	7	50	350	......	2500	17500
53-57	5	55	275	20	3025	15125
58-62	7	60	......	27	3600	......
63-67	3	65	195	30	4225	12675
Total	......		1565			83825

(a) Salin semula jadual dan masukkan nilai-nilai yang tertinggal

(6 markah)

(b) Lukiskan sebuah histogram daripada sampel data ini.

(3 markah)

(c) Kirakan median untuk sampel data ini.

(2 markah)

(d) Cari min dan varians untuk sampel data ini.

(9 markah)

S5 (a) Diberi S ialah ruang sampel untuk peristiwa bagi nombor-nombor yang mempunyai peluang yang sama untuk dipilih.

$$S = \{2, 3, 5, 7, 11, 12, 13, 16, 17, 21, 23, 26, 33, 37, 41\}$$

A, B, C and D adalah peristiwa-peristiwa bagi ruang sampel seperti di bawah:

$A = \{7, 16, 21, 37, 41\}$
$B = \{2, 3, 21, 37\}$
$C = \{2, 11, 12, 23, 37\}$
$D = \{5, 17\}$

(i) Lukiskan gambarajah Venn kesemua peristiwa tadi.
(ii) Kirakan kebarangkalian untuk $P(B \cap C)$, $P(A \cap B \cap C)'$, $P(A \mid C)$

(10 markah)

(b) Terdapat tiga jenis komponen di dalam sebuah pengeluaran komponen elektrik, iaitu jenis A, jenis B dan jenis C. Empat puluh peratus daripada pengeluaran terdiri dari jenis A, dua puluh peratus dari jenis B dan yang selebihnya dari jenis C. Kebarangkalian untuk komponen yang rosak bagi jenis A, B dan C masing-masing ialah 0.04, 0.02 dan 0.05.

(i) Lukiskan gambarajah pokok untuk masalah di atas.
(ii) Cari kebarangkalian untuk komponen yang rosak.
(iii) Cari kebarangkalian yang komponen adalah jenis B jika diketahui bahawa komponen itu rosak.

(10 markah)

S6 Pembolehubah rawak X mempunyai fungsi ketumpatan kebarangkalian f seperti berikut

$$f(x) = \begin{cases} a - 5 + x & 5 \leq x \leq 6 \\ 0 & \textit{selainnya} \end{cases}$$

(a) Tunjukkan bahawa $a = \dfrac{1}{2}$

(4 markah)

(b) Tentukan nilai jangkaan bagi X, $E(X)$.

(4 markah)

(c) Dapatkan fungsi taburan longgokan, F.

(6 markah)

(d) Cari kebarangkalian $F(5.2)$, $P(X > 5.5)$, dan $P(5 \leq X \leq 5.5)$

(6 markah)

S7 (a) Purata bilangan tuntutan sejam yang dibuat ke Takaful Nasional Sdn. Bhd. untuk kerosakan atau kehilangan ialah 1.5. Menggunakan taburan Poisson, cari kebarangkalian bahawa,

(i) dalam sebarang satu jam, kurang dari 2 tuntutan akan dibuat?

(ii) dalam 3 jam, 3 atau lebih tuntutan akan dibuat?

(7 markah)

(b) Bila pelanggan membuat tempahan dengan Ahmad Supplies, sistem maklumat perakaunan berkomputer secara automatik akan memeriksa sama ada pelanggan telah melebihi had kredit beliau. Rekod yang lalu menunjukkan bahawa kebarangkalian pelanggan melebihi had kredit beliau ialah 0.09. Andaikan bahawa bilangan pelanggan yang dikesan membuat tempahan melebihi had kredit adalah tertabur secara pembolehubah rawak Binomial.

(i) Andaikan pada suatu hari tertentu terdapat 20 pelanggan membuat tempahan. Cari kebarangkalian bahawa tiada pelanggan melebihi had kredit.

(ii) Katalah pada suatu hari tertentu, terdapat 100 pelanggan telah membuat tempahan. Cari kebarangkalian bahawa lebih dari 10 pelanggan akan melebihi had kredit dengan menggunakan taburan binomial dan taburan penghampiran yang sesuai.

(13 markah)

S8 (a) Jumlah bil elektrik untuk semua penghuni di dalam sebuah bandar mempunyai taburan normal dengan min RM 80.00 dan sisihan piawai RM 25.00 Cari kebarangkalian min jumlah bil elektrik bagi 75 orang penghuni yang dipilih secara rawak supaya minnya

(i) kurang daripada RM 76.00

(ii) di antara RM 72.00 hingga RM 77.00

(iii) lebih besar daripada RM 85.00

(9 markah)

(b) Katakan X ialah taburan ketinggian untuk baka kambing gurun yang mempunyai min ketinggian 72 cm dengan sisihan piawai 10 cm, manakala pula Y ialah taburan ketinggian untuk baka kambing biri-biri yang mempunyai min ketinggian 68 cm dengan sisihan piawai 5 cm. Suatu sampel rawak sebanyak 64 ekor kambing gurun dan 100 ekor kambing biri-biri diambil untuk kajian.

(i) Apakah taburan min pensampelan untuk baka kambing gurun, $\overline{X}$?

(ii) Apakah taburan bagi perbezaan min pensampelan untuk kedua-dua jenis baka kambing ini, $\overline{X} - \overline{Y}$?

(iii) Cari kebarangkalian min sampel untuk baka kambing gurun berada diantara 70.0 cm dan 74.2 cm.

(iv) Kirakan kebarangkalian $P(\overline{X} - \overline{Y} \geq 8)$

(11 markah)

FORMULA STATISTIK

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

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

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

$$p(x) = \frac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}},\ x = 0, 1, 2, \ldots, k \qquad X \sim N(\mu, \sigma^2)$$

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

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

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

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

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

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

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

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

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