UNIVERSITI SAINS MALAYSIA

Second Semester Examination
Academic Session 2007/2008

April 2008

MAA 161 – Statistics for Science Students
[Statistik untuk Pelajar Sains]

Duration : 3 hours
[Masa : 3 jam]

Please check that this examination paper consists of EIGHT pages of printed material before you begin the examination.

[Sila pastikan bahawa kertas peperiksaan ini mengandungi LAPAN muka surat yang bercetak sebelum anda memulakan peperiksaan ini.]

Instructions: Answer **all ten** [10] questions.

[Arahan: *Jawab* ***semua sepuluh*** *[10] soalan.]*

1. The sizes (in carats) of 16 diamonds are as follows:

0.42	0.63	0.81	1.01	0.50	0.59	0.66	0.48
0.70	0.99	0.83	0.50	0.51	0.75	0.66	1.38

(i) Find the median of this data.
(ii) Compute the first and third quartiles.
(iii) Draw a box-and-whisker plot to represent the data. Is the reading 1.38 an outlier?

[10 marks]

2. A light fixture holds two light bulbs. Bulb A is of type whose lifetime is normally distributed with mean 800 hours and standard deviation 100 hours. Bulb B has a lifetime that is normally distributed with mean 900 hours and standard deviation 150 hours. Assume that the lifetimes of the bulbs are independent.
(i) What is the probability that bulb B lasts longer than bulb A?
(ii) What is the probability that bulb B lasts more than 200 hours longer than bulb A?

[8 marks]

3. Tahir and Ahmad each fire one shot at a target. Tahir has probability 0.5 of hitting the target, and Ahmad has probability 0.3. The shots are independent.
(i) Find the probability that the target is hit.
(ii) Find the probability that the target is hit by exactly one shot.
(iii) Given that the target was hit by exactly one shot, find the probability that Tahir hit the target.

[10 marks]

4. A company employs 800 men under the age of 55. Suppose that 30% carry a marker on the male chromosome that indicates an increased risk for high blood pressure.
(i) If 10 men in the company are tested for the marker in this chromosome, what is the probability that exactly 1 man has the marker?
(ii) If 10 men in the company are tested for the marker in this chromosome, what is the probability that more than 1 man has the marker?

[8 marks]

5. A rare disease exists in which only 1 in 500 are affected. A test for the disease exists but of course it is not infallible. A correct positive result (patient actually has the disease) occurs 95% of the time while a false positive result (patient does not have the disease) occurs 1% of the time. If a randomly selected individual is tested and the result is positive, what is the probability that the individual has the disease?

[6 marks]

1. *Saiz (dalam karat) bagi 16 butir belian diberikan seperti berikut:*

0.42	*0.63*	*0.81*	*1.01*	*0.50*	*0.59*	*0.66*	*0.48*
0.70	*0.99*	*0.83*	*0.50*	*0.51*	*0.75*	*0.66*	*1.38*

(i) Dapatkan median bagi data.
(ii) Hitungkan kuartils pertama dan ketiga.
(iii) Lukiskan plot kotak dan misai bagi menerangkan data. Adakah bacaan 1.38 suatu nilai yang terkeluar?

[10 markah]

2. *Suatu soket dapat dimuatkan dengan dua mentol iaitu mentol A dan mentol B. Mentol A adalah jenis yang mempunyai jangka hayat bertaburan secara normal dengan min 800 jam dan sisihan piawai 100 jam. Manakala, mentol B mempunyai jangka hayat bertaburan secara normal dengan min 900 jam dan sisihan piawai 150 jam. Andaikan bahawa jangka hayat kedua-dua mentol adalah tak bersandar.*
(i) Apakah kebarangkalian mentol B bertahan lebih lama dari mentol A?
(ii) Apakah kebarangkalian mentol B bertahan 200 jam lebih lama daripada mentol A?

[8 markah]

3. *Tahir dan Ahmad menembak seorang sekali pada suatu sasaran. Tahir mempunyai kebarangkalian 0.5 mengenai sasaran dan Ahmad mempunyai kebarangkalian 0.3. Setiap tembakan adalah tak bersandar.*
(i) Dapatkan kebarangkalian mengenai sasaran.
(ii) Dapatkan kebarangkalian bahawa sasaran dikenai dengan satu tembakan.
(iii) Sekiranya sasaran telah dikenai dengan satu tembakan, dapatkan kebarangkalian bahawa Tahir yang mengenai sasaran itu.

[10 markah]

4. *Suatu syarikat menggaji 800 orang lelaki berumur 55 tahun ke bawah. Andaikan 30% daripada mereka mempunyai petanda dalam kromosom yang menunjukkan risiko tinggi berpenyakit tekanan darah tinggi.*
(i) Jika 10 orang lelaki dalam syarikat diuji bagi kehadiran petanda dalam kromosom, apakah kebarangkalian bahawa tepat seorang mempunyai petanda?
(ii) Jika 10 orang lelaki dalam syarikat diuji bagi kehadiran petanda dalam kromosom, apakah kebarangkalian bahawa lebih daripada seorang mempunyai petanda.

[8 markah]

5. *Suatu penyakit yang jarang wujud di mana hanya 1 dalam 500 orang akan dijangkiti. Ujian untuk mengesan penyakit ini pula adalah tidak sempurna. Suatu keputusan positif yang benar (pesakit sebenarnya mempunyai penyakit ini) berlaku 95% pada sesuatu masa manakala suatu keputusan positif yang palsu (pesakit tidak mempunyai penyakit ini) berlaku 1% pada sesuatu masa. Jika seorang individu diuji secara rawak dan keputusan adalah positif, apakah kebarangkalian bahawa individu itu mempunyai penyakit ini?*

[6 markah]

...4/-

6. Seventy percent of adults favour some kind of government control on the prices of medicines. Assume that this percentage is true for the current population of all adults.

(i) Find the probability that the proportion of adults in a random sample of 400 who favour some kind of government control on the prices of medicines is less than 0.65 and between 0.73 and 0.76.

(ii) What is the probability that the proportion of adults in a random sample of 400 who favour some kind of government control is within 0.06 of the population proportion?

[12 marks]

7. The following data show the number of defects in 100,000 lines of code in a particular type of software program made in the U.S. and Japan. Is there enough evidence to claim that there is a significance difference between the programs of the two countries? Assume the two population variances are equal. Use $\alpha = 0.05$.

U.S, X

48	39	42	52	40	48	52	52
54	48	52	55	43	46	48	52

Japan, Y

50	48	42	40	43	48	50	46
38	38	36	40	40	48	48	45

$$\sum_i x_i = 771, \quad \sum_i x_i^2 = 37527, \quad \sum_i y_i = 700, \quad \sum_i y_i^2 = 30954$$

[14 marks]

8. The following data represent the number of hours that two different types of scientific pocket calculator operate before a recharge is required.

Calculator A	5.5	5.6	6.3	4.6	5.3	5.0	6.2	5.8	5.1
Calculator B	3.8	4.8	4.3	4.2	4.0	4.9	4.5	5.2	4.5

Use the appropriate nonparametric method to determine if calculator A operates longer than calculator B on a full battery recharge. Use $\alpha = 0.01$.

[12 marks]

6. *Tujuh puluh peratus orang dewasa memihak kawalan harga ubat-ubatan oleh pihak kerajaan. Andaikan bahawa peratusan ini adalah benar untuk penduduk semasa orang dewasa keseluruhan.*

(i) Dapatkan kebarangkalian bahawa kadar orang dewasa dalam satu sampel rawak 400 yang memihak kawalan harga ubat-ubatan oleh pihak kerajaan adalah kurang daripada 0.65 dan antara 0.73 dan 0.76

(ii) Apakah kebarangkalian bahawa kadar orang dewasa dalam satu sampel rawak 400 yang memihak kawalan kerajaan adalah dalam lingkungan 0.06 dari kadar populasi?

[12 markah]

7. *Data-data berikut menunjukkan bilangan kesalahan dalam 100,000 baris kod dalam suatu aturcara perisian tertentu yang dibuat di U.S dan di Jepun. Adakah terdapat bukti yang cukup untuk mendakwa bahawa terdapat perbezaan bererti antara aturcara dari kedua-dua negara? Andaikan kedua-dua varians populasi adalah sama? Guna* $\alpha = 0.05$.

U.S, X

48	39	42	52	40	48	52	52
54	48	52	55	43	46	48	52

Jepun, Y

50	48	42	40	43	48	50	46
38	38	36	40	40	48	48	45

$$\sum_i x_i = 771, \quad \sum_i x_i^2 = 37527, \quad \sum_i y_i = 700, \quad \sum_i y_i^2 = 30954$$

[14 markah]

8. *Data-data berikut mewakili bilangan jam operasi bagi dua jenis kalkulator saintifik yang berlainan sebelum perlu dicas semula.*

Kalkulator A	5.5	5.6	6.3	4.6	5.3	5.0	6.2	5.8	5.1
Kalkulator B	3.8	4.8	4.3	4.2	4.0	4.9	4.5	5.2	4.5

Dengan menggunakan kaedah tak berparameter yang sesuai tentukan sama ada kalkulator A beroperasi lebih lama daripada kalkulator B apabila bateri telah dicas sepenuhnya. Guna $\alpha = 0.01$.

[12 markah]

9. The Malaysian Department of Statistics compiles data on the current employment status of "displaced workers" who have lost or left a job for reasons given in the table below. The table gives the number of workers (in thousand) who lost jobs but are now employed, unemployed, or not in the labor force.

Reason for Job Loss	*Employed*	*Unemployed*	*Not in the Labor Force*
Plant or company closed down or moved	2532	340	429
Insufficient work	1892	379	258
Position or shift abolished	1635	212	276

Suppose that these data were obtained from a random sample of displaced worker. Using the 1% significance level, can you conclude that the current employment status and reason for job loss are related?

[12 marks]

10. Let μ be the mean hourly wage of carpenters in Penang. A random sample of carpenters from Penang yielded a 95% confidence interval for μ of RM11.45 to RM14.35. Assume that this sample included more than 30 carpenters.

(i) Find the value of $\bar{x}$ for this sample.

(ii) Find the 98% confidence interval for μ based on this sample.

[8 marks]

9. *Jabatan Statistik Malaysia sedang menyusun data status pekerjaan semasa "pekerja-pekerja yang diganti" iaitu sama ada yang telah hilang atau berhenti dari suatu pekerjaan untuk sebab-sebab diberi dalam jadual di bawah. Jadual menunjukkan bilangan pekerja (dalam ribu) yang hilang pekerjaan namun kini sedang bekerja, menganggur atau tidak mahu bekerja semula.*

Sebab Kehilangan Pekerjaan	***Bekerja***	***Menganggur***	***Tidak Mabu Bekerja Semula***
Kilang atau syarikat tutup atau berpindah	*2532*	*340*	*429*
Kerja berkurangan	*1892*	*379*	*258*
Posisi atau pertukaran dimansuhkan	*1635*	*212*	*276*

Jika data-data telah didapati daripada sampel rawak "pekerja yang diganti". Pada aras keertian 1%, bolehkah dibuat kesimpulan bahawa terdapat hubungan antara status pekerjaan terkini dan sebab kehilangan pekerjaan?

[12 markah]

10. *Andaikan μ adalah min upah sejam seorang tukang kayu di Pulau Pinang. Suatu sampel rawak tukang kayu dari Pulau Pinang menghasilkan selang keyakinan 95% untuk μ antara RM 11.45 dan RM 14.35. Anggapkan sampel tukang kayu yang diambil lebih daripada 30.*

(i) Dapatkan nilai $\bar{x}$ untuk sampel ini.

(ii) Dapatkan selang keyakinan 98% untuk μ bagi sampel ini.

[8 markah]

APPENDIX

Confidence Interval

$$\bar{X} \pm Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \qquad \bar{X} \pm t_{\frac{\alpha}{2}} \frac{s}{\sqrt{n}}$$

$$\bar{d} \pm t_{\frac{\alpha}{2}} \frac{s_d}{\sqrt{n_d}} \qquad b \pm t_{\frac{\alpha}{2}} s_b$$

$$\hat{p} \pm Z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

$$(\hat{p}_x - \hat{p}_y) \pm z_{\alpha/2} \cdot \sqrt{\frac{\hat{p}_x(1-\hat{p}_x)}{n_x} + \frac{\hat{p}_y(1-\hat{p}_y)}{n_y}}$$

$$(\bar{X} - \bar{Y}) \pm Z_{\alpha/2} \cdot \sqrt{\frac{\sigma_x^{\,2}}{n_x} + \frac{\sigma_y^{\,2}}{n_y}}$$

$$(\bar{X} - \bar{Y}) \pm t_{\alpha/2} \cdot \sqrt{\frac{s_x^{\,2}}{n_x} + \frac{s_y^{\,2}}{n_y}}$$

$$(\bar{X} - \bar{Y}) \pm t_{\alpha/2}\ S_p \sqrt{\frac{1}{n_x} + \frac{1}{n_y}}$$

$$\left(\frac{(n-1)s^2}{\chi^2_{\frac{\alpha}{2},n-1}}, \frac{(n-1)s^2}{\chi^2_{1-\frac{\alpha}{2},n-1}} \right)$$

$$\left(\frac{s}{1+\frac{Z_{\frac{\alpha}{2}}}{\sqrt{2n}}}, \frac{s}{1-\frac{Z_{\frac{\alpha}{2}}}{\sqrt{2n}}} \right)$$

$$\left(\frac{s_1^2}{s_2^2} F_{1-\frac{\alpha}{2},(v_2,v_1)}, \frac{s_1^2}{s_2^2} F_{\frac{\alpha}{2},(v_2,v_1)} \right)$$

Test Statistic

$$Z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}}$$

$$T = \frac{\bar{X} - \mu}{s / \sqrt{n}}$$

$$T = \frac{\bar{d} - \mu_d}{s_d / \sqrt{n_d}}$$

$$T = \frac{b - \beta_1}{s_b}$$

$$T = r\sqrt{\frac{n-2}{1-r^2}}$$

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2}$$

$$U_1 = n_1 n_2 + \frac{(n_2)(n_2+1)}{2} - R_2$$

$$U_2 = n_1 n_2 + \frac{(n_1)(n_1+1)}{2} - R_1$$

$$Z = \frac{s - \sigma}{\sigma / \sqrt{2n}}$$

$$Z = \frac{(\bar{X} - \bar{Y}) - (\mu_x - \mu_y)}{\sqrt{\frac{\sigma_x^{\,2}}{n_x} + \frac{\sigma_y^{\,2}}{n_y}}}$$

$$T = \frac{(\bar{X} - \bar{Y}) - (\mu_x - \mu_y)}{\sqrt{S_p^{\,2}\left(\frac{1}{n_x} + \frac{1}{n_y}\right)}}$$

$$S_p^{\,2} = \frac{(n_x - 1)s_x^{\,2} + (n_y - 1)s_y^{\,2}}{n_x + n_y - 2}$$

$$F = \frac{s_x^2}{s_y^2}$$

$$Z = \frac{(\hat{p}_x - \hat{p}_y) - (p_x - p_y)}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_x} + \frac{1}{n_y}\right)}}$$

$$T = \frac{(\bar{X} - \bar{Y}) - (\mu_x - \mu_y)}{\sqrt{\frac{s_x^2}{n_x} + \frac{s_y^2}{n_y}}}$$

$$dk = \frac{\left(\frac{s_x^2}{n_x} + \frac{s_y^2}{n_y}\right)^2}{\frac{\left(\frac{s_x^2}{n_x}\right)^2}{n_x - 1} + \frac{\left(\frac{s_y^2}{n_y}\right)^2}{n_y - 1}}$$

$$\chi^2 = \sum \frac{(O-E)^2}{E}, \quad E = np$$

-ooo00ooo-