UNIVERSITI SAINS MALAYSIA

Second Semester Examination
Academic Session 2008/2009

April/May 2009

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 materials 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. According to a survey on the nutritional value of a type of low fat cheese, it is found that on the average, a piece of cheese has 3.5 grams of fat with the standard deviation of 0.04 gram fat. What can be said about the percentage of cheese that the fat content is more than 3.62 grams if:

   (a) the form of the data distribution is not known?
   (b) the data distribution is symmetrical?

   [10 marks]

2. The staff of a college is categorized as academic staff, administrative staff or support staff. The following table shows the number of staff in every category and their sex.:

	Male(L)	Female (W)	Total
Academic Staff (A)	42	28	70
Administrative Staff (B)	7	13	20
Support Staff (C)	26	9	35
Total	75	50	125

   Find

   (a) $P(L \cup \overline{A})$
   (b) $P(W|A)$

   Assume that 90% of the academic staff, 80% of the administrative staff and 30% of the support staff are more than 40 years of age.

   (c) If a staff is chosen randomly, what is the probability that he or she is more than 40 years old?
   (d) If a staff chosen randomly is more than 40 years old, what is the probability that he or she is a support staff?

   [12 marks]

3. In a group of 14 statistics students, 10 have gone through industrial training. If 3 are chosen at random for a research work, find the probability that

   (a) all three students have gone through industrial training before
   (b) only 2 students have gone through industrial training before.

   [8 marks]

1. *Menurut kajian mengenai nilai nutrisi sejenis keju rendah lemak, didapati bahawa secara puratanya, sekeping keju mengandungi 3.5 gram lemak dengan sisihan piawai 0.04 gram lemak. Apakah yang boleh dinyatakan berkenaan peratusan keju bahawa kandungan lemaknya adalah lebih daripada 3.62 gram jika:*

   (a) bentuk taburan data tidak diketahui?
   (b) taburan data berbentuk simetri?

   [10 markah]

2. *Kakitangan sebuah kolej dikategorikan sebagai staf akademik, staf pentadbiran atau staf sokongan. Jadual yang berikut menunjukkan bilangan kakitangan dalam setiap kategori dan jantina mereka:*

	Lelaki(L)	*Wanita (W)*	*Jumlah*
Staf Akademik (A)	*42*	*28*	*70*
Staf Pentadbiran (B)	*7*	*13*	*20*
Staf Sokongan (C)	*26*	*9*	*35*
Jumlah	*75*	*50*	*125*

   Dapatkan

   (a) $P(L \cup \overline{A})$
   (b) $P(W|A)$

   Andaikan bahawa 90% daripada staf akademik, 80% daripada staf pentadbiran dan 30% daripada staf sokongan berusia lebih daripada 40 tahun

   (c) Jika seorang kakitangan dipilih secara rawak, apakah kebarangkalian bahawa dia berusia lebih daripada 40 tahun?
   (d) Jika seorang kakitangan yang dipilih secara rawak berusia lebih daripada 40 tahun, apakah kebarangkalian bahawa beliau adalah staf sokongan?

   [12 markah]

3. *Dalam kalangan 14 orang pelajar statistik, 10 orang pernah menjalani latihan industri. Jika 3 orang pelajar itu dipilih secara rawak untuk suatu kerja penyelidikan, cari kebarangkalian bahawa*

   (a) ketiga-tiga orang pelajar itu pernah menjalani latihan industri
   (b) hanya 2 orang pelajar itu yang pernah menjalani latihan industri.

   [8 markah]

4. A factory produces 50,000 pencils every day. It is found that 5% of the pencils produced by the factory is faulty. If 400 pencils from the factory are chosen randomly, what is the probability of obtaining

(a) at least 14 but not more than 25 faulty pencils?
(b) 33 or more faulty pencils?

[10 marks]

5. A random sample of size 30 is taken from a population which is uniformly distributed in the range $3 \leq X \leq 6$.

(a) Describe the distribution of the sample mean
(b) Draw a graph of the population distribution and the distribution of the sample mean
(c) Find the probability that the sample mean will be more than 4.7

[12 marks]

6. The following data are marks obtained by 10 students in a history test:

56 87 64 38 26 67 54 57 78 61

(a) Find the 95% confidence interval for μ.
(b) What is the maximum estimation error in section (a).

[8 marks]

7. Random samples $X_1, X_2, \ldots, X_{25}$ are taken from a normal population with mean μ and variance 6.5.

The hypothesis test is:

$H_0 : \mu = 50$
$H_1 : \mu < 50$

and the tester wants to have a probability of type I error of 1%.

(a) Find the rejection region in terms of $\overline{X}$.
(b) Calculate the type II error when $\mu = 45$

[10 marks]

4. *Suatu kilang mengeluarkan 50,000 pensel setiap hari. Didapati bahawa 5% daripada pensel yang dikeluarkan oleh kilang itu cacat. Jika suatu sampel 400 pensel dari kilang itu dipilih secara rawak, apakah kebarangkalian mendapat*

   (a) sekurang-kurangnya 14 tetapi tidak melebihi 25 pensel yang cacat ?
   (b) 33 atau lebih pensel cacat ?

   [10 markah]

5. *Satu sampel rawak bersaiz 30 diambil daripada populasi yang tertabur secara seragam dalam julat* $3 \leq X \leq 6$.

   (a) Huraikan taburan pensampelan bagi min sampel
   (b) Lukiskan secara bergraf taburan populasi dan taburan pensampelan bagi min sampel
   (c) Cari kebarangkalian bahawa minsampel akan melebihi 4.7

   [12 markah]

6. *Data-data berikut adalah markah yang diperoleh oleh 10 orang pelajar dalam satu ujian sejarah :*

   56 87 64 38 26 67 54 57 78 61

   (a) Cari selang keyakinan 95% bagi μ.
   (b) Apakah ralat maksimum penganggaran di bahagian (a).

   [8 markah]

7. *Sampel rawak* $X_1, X_2, \ldots, X_{25}$ *diambil dari suatu populasi normal dengan min* μ *dan varians 6.5.*

   Ujian hipotesis ialah :

   $H_0 : \mu = 50$
   $H_1 : \mu < 50$

   dan penguji inginkan kebarangkalian ralat jenis I sebanyak 1%.

   (a) Dapatkan rantau penolakan dalam sebutan $\overline{X}$.
   (b) Hitung ralat jenis II apabila nilai $\mu = 45$

   [10 markah]

8. A doctor claims that the diet he prepares can reduce body weight. The body weights of 7 women who use the diet are recorded before and after a period of 3 weeks.

Participant	1	2	3	4	5	6	7
Body weight before	129	133	136	152	141	138	120
Body weight after	130	126	128	148	129	135	120

Can the doctor's claim be accepted at the 5% level of significance?

[10 marks]

9. According to Mendelian genetic theory, a mixture of red flowering legumes and blue flowering legumes will produce three types of legumes with red, blue and purple flowering legumes in the ratio of 1: 1: 2. An experiment is done with 333 pairs of legumes and it is found to produce the following results: 84 red flowering legumes, 92 blue flowering legumes and 157 purple flowering legumes. At the 1% level of significance, test whether the data support the Mendelian theory.

[10 marks]

10. A researcher attempts to determine if a drug has an effect on a particular disease. Counts of individuals are given in the table with the diagnosis either positive or negative before treatment given in the columns (before), and the diagnosis after treatment in the rows (after). The test requires the same subject to be included in the before and after measurements.

		Before		
		positive	negative	total
After	positive	40	10	50
	negative	20	50	70
total		60	60	120

Test at 5% level of signifance as to whether the drug is effective or not.

[10 marks]

8. *Seorang doctor mendakwa bahawa diet yang disediakannya dapat mengurangkan berat badan. Berat badan 7 orang perempuan yang menggunakan diet ini dicatatkan sebelum dan selepas suatu tempoh 3 minggu.*

Peserta	*1*	*2*	*3*	*4*	*5*	*6*	*7*
Berat badan sebelum	*129*	*133*	*136*	*152*	*141*	*138*	*120*
Berat badan selepas	*130*	*126*	*128*	*148*	*129*	*135*	*120*

Bolehkah dakwaan doctor itu diterima pada aras keertian 5%.

[10 markah]

9. *Menurut teori genetic Mendelian, kacukan kekacang berbunga merah dengan kekacang berbunga biru akan menghasilkan tiga jenis baka kekacang yang berbunga merah, biru atau ungu, dalam nisbah 1: 1: 2. Suatu ujikaji dijalankan dengan 333 pasangan kekacang dan didapati hasil berikut: 84 baka berbunga merah, 92 baka berbunga biru dan 157 baka berbunga ungu. Pada aras keertian 1%, uji sama ada data ini menyokong teori Mendelian.*

[10 markah]

10. *Seorang penyelidik ingin menentukan sama ada sejenis ubat mempunyai kesan pada sesuatu penyakit. Sebilangan individu dalam jadual dengan diuji terhadap penyakit tersebut sama ada positif atau negative sebelum diberi ubat dalam ruang (sebelum) dan diuji selepas diberi ubat dalam baris (selepas). Ujian menghendaki individu yang sama dimasukkan dalam ujian sebelum dan selepas*

		Sebelum		
		positif	*negatif*	*jumlah*
Selepas	*positif*	*40*	*10*	*50*
	negatif	*20*	*50*	*70*
jumlah		*60*	*60*	*120*

Uji pada aras keertian 5% sama ada ubat tersebut berkesan atau tidak.

[10 markah]

APPENDIX

FORMULA
MAA 161 – Statistics For Science Students

Confidence Interval:

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

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

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

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

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

$$\left(\frac{s}{Z_{\frac{\alpha}{2}}}, \frac{s}{Z_{\frac{\alpha}{2}}} \right)$$

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

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

$$\left(\frac{s_1^2}{s_2^2} F_{1-\frac{\alpha}{2},(\nu_2,\nu_1)}, \frac{s_1^2}{s_2^2} F_{\frac{\alpha}{2},(\nu_2,\nu_1)} \right)$$

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

Test Statistic:

$$Z = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \qquad Z = \frac{s-\sigma}{\sigma/\sqrt{2n}} \qquad Z = \frac{(\hat{p}_x - \hat{p}_y) - (p_x - p_y)}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_x}+\frac{1}{n_y}\right)}}$$

$$T = \frac{\bar{X}-\mu}{s/\sqrt{n}} \qquad Z = \frac{(\bar{X}-\bar{Y}) - (\mu_x - \mu_y)}{\sqrt{\frac{{\sigma_x}^2}{n_x} + \frac{{\sigma_y}^2}{n_y}}} \qquad T = \frac{(\bar{X}-\bar{Y}) - (\mu_x - \mu_y)}{\sqrt{\frac{s_x^2}{n_x} + \frac{s_y^2}{n_y}}}$$

$$T = \frac{\bar{d}-\mu_d}{s_d/\sqrt{n_d}} \qquad T = \frac{(\bar{X}-\bar{Y}) - (\mu_x - \mu_y)}{\sqrt{{S_p}^2\left(\frac{1}{n_x}+\frac{1}{n_y}\right)}} \qquad dk = \frac{\left(\frac{s_x^2}{n_x}+\frac{s_y^2}{n_y}\right)^2}{\frac{\left(\frac{s_x^2}{n_x}\right)^2}{n_x-1} + \frac{\left(\frac{s_y^2}{n_y}\right)^2}{n_y-1}}$$

$$T = \frac{b-\beta_1}{s_b} \qquad {S_p}^2 = \frac{(n_x-1){s_x}^2 + (n_y-1){s_y}^2}{n_x+n_y-2}$$

$$T = r\sqrt{\frac{n-2}{1-r^2}} \qquad F = \frac{s_x^2}{s_y^2} \qquad \chi^2 = \sum \frac{(O-E)^2}{E}, \quad E = np$$

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