UNIVERSITI SAINS MALAYSIA

First Semester Examination
Academic Session 2007/2008

October/November 2007

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

Duration : 3 hours
[Masa : 3 jam]

Please check that this examination paper consists of <u>NINE</u> pages of printed material before you begin the examination.

[Sila pastikan bahawa kertas peperiksaan ini mengandungi <u>SEMBILAN</u> muka surat yang bercetak sebelum anda memulakan peperiksaan ini.]

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

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

1. (a) Events M and N are mutually exclusive and $P(M) = 0.3$, $P(N) = 0.4$. Find

   (i) $P(M \text{ or } N)$
   (ii) $P(M \text{ and } N)$
   (iii) $P(M \mid N)$
   Are events M and N independent? Explain.

   (b) Two flower seeds are randomly selected from a package containing five seeds of red flowers and three seeds of white flowers.

   (i)   What is the probability that both seeds will result in red flowers?
   (ii)   What is the probability that one of each color is selected?

   [10 marks]

2. An airline's records show that its flights between two cities arrive on the average 5.5 minutes late with a standard deviation of 1.5 minutes.

   (i)   At least what percentage of the flights between the two cities arrive anywhere between 2.5 minutes late and 8.5 minutes late? Use Chebyshev's theorem.
   (ii)   About what percentage of the flights between the two cities arrive anywhere between 4 minutes late and 10 minutes late? Use empirical rule.

   [10 marks]

3. It is known from experience that 2% of the calls received by a switchboard are wrong numbers. Use the Poisson approximation to the binomial distribution to determine the probability that among 300 calls received by a switchboard,

   (i)   four will be wrong numbers.
   (ii)   more than three will be wrong numbers.

   [8 marks]

4. In a cannery, assembly lines X, Y and Z account for 35%, 40% and 20%, respectively, of the total output. If 0.5% of the cans from assembly line X are improperly sealed, while the corresponding percentages for assembly lines Y and Z are 0.6% and 1.2%, what is the probability that

   (i)   a can is improperly sealed?
   (ii)   an improperly sealed can came from assembly line Y?

   [12 marks]

5. According to the Malaysian Automobile Association, nearly 50% of Malaysian drivers are females. Assume a random sample of 60 drivers is to be selected for a survey.

1. (a) Peristiwa M dan N adalah saling eksklusif dan P(M) = 0.3, P(N) = 0.4. Dapatkan
    (i)   P(M atau N)
    (ii)   P(M dan N)
    (iii) P(M | N)
    Adakah peristiwa M dan N tak bersandar. Jelaskan.

   (b) Dua biji benih bunga dipilih secara rawak daripada sebuah pek yang mengandungi lima biji benih bunga merah dan tiga biji benih bunga putih.
    (i)   Apakah kebarangkalian kedua-dua biji benih menghasilkan bunga merah?
    (ii)   Apakah kebarangkalian satu biji benih bunga merah dan satu biji benih bunga putih dipilih?

[10 markah]

2. Rekod-rekod sebuah sistem penerbangan menunjukkan bahawa penerbangan di antara dua buah bandaraya tiba lewat pada purata 5.5 minit dengan sisihan piawai 1.5 minit.
    (i)   Sekurang-kurangnya berapa peratus daripada penerbangan di antara dua buah bandaraya tersebut tiba 2.5 minit hingga 8.5 minit lewat? Gunakan theorem Chebyshev.
    (ii)   Sekurang-kurangnya berapa peratus daripada penerbangan di antara dua buah bandaraya tersebut tiba 4 minit hingga 10 minit lewat? Gunakan petua empirical.

[10 markah]

3. Berdasarkan pengalaman lepas diketahui bahawa 2% daripada panggilan yang diterima oleh sebuah 'switchboard' adalah salah nombor. Gunakan penghampiran Poisson untuk taburan binomial bagi menentukan kebarangkalian bahawa daripada 300 panggilan yang diterima oleh sebuah 'switchboard'
    (i)   empat adalah salah nombor.
    (ii)   lebih daripada tiga adalah salah nombor.

[8 markah]

4. Di sebuah kilang pengetinan makanan, barisan pemasangan X, Y dan Z masing-masing menghasilkan 35%, 40% dan 20% daripada jumlah pengeluaran. Jika 0.5% daripada tin-tin yang dikeluarkan oleh barisan pemasangan X tidak ditutup dengan sempurna, manakala peratusan bagi Y dan Z masing-masing adalah 0.6% dan 1.2%, apakah kebarangkalian bahawa
    (i)   suatu tin yang dipilih tidak ditutup dengan sempurna?
    (ii)   suatu tin yang tidak ditutup dengan sempurna dihasilkan oleh barisan pemasangan Y ?

[12 markah]

5. Menurut Persatuan Automobile Malaysia, hampir 50% daripada pemandu Malaysia adalah wanita. Katakan suatu sampel rawak yang mengandungi 60 orang pemandu akan ambil.

(i)   What is the probability that no more than half (30) of the drivers will be female?

(ii)   What is the probability that at least three fourth (45) of the drivers will be female?

[10 marks]

6. (a) Among all the income-tax forms filed in a certain year, the mean tax paid was RM 2000 and the standard deviation was RM 500. In addition, for 10% of the forms, the tax paid was greater than RM 3000. A random sample of 625 tax forms is drawn

      (i)   What is the probability that the average tax paid on the sampled forms is greater than RM 1980?

      (ii)   What is the probability that more than 60 of the sampled forms have a tax of greater than RM 3000?

  (b) Resistance measurements were made on a sample of 81 wires of a certain type. The sample mean resistance was 17.3 $m\Omega$ and the variance was 1.44 $m\Omega$.

      (i)   Construct a 98% confidence interval for the mean resistant of this type of wire.

      (ii)   What is the confidence level for the interval (17.1, 17.5)?

      (iii)   How many wires must be sampled so that a 98% confidence interval will specify the mean to within $\pm 0.1\, m\Omega$?

[12 marks]

7. A new production process is being contemplated for the manufacture of stainless steel bearings. Measurements of the diameters of random samples of bearings from the old process X and the new process Y produced the following data:

Old process, X

16.3	15.9	15.8	16.2	16.1	16.0
15.7	15.8	15.9	16.1	16.3	16.1
15.8	15.7	15.8	15.7		

New process, Y

15.9	16.2	16.0	15.8	16.1	16.1
15.8	16.0	16.2	15.9	15.7	16.2
15.8	15.8	16.2	16.3		

$$\sum_i x_i = 255.2, \quad \sum_i x_i^2 = 4071.1, \quad \sum_i y_i = 256, \quad \sum_i y_i^2 = 4096.54$$

(i)   Can you conclude at the 5% level that one process yields a different mean size bearing than the other?

(ii)   Can you conclude at the 5% level that the variance of the new process is lower than the old process?

[15 marks]

...5/-

   (i)   *Apakah kebarangkalian tidak lebih daripada setengah (30) daripada pemandu adalah wanita?*

   (ii)   *Apakah kebarangkalian sekurang-kurangnya tiga perempat (45) daripada pemandu adalah wanita?*

                                                        [10 markah]

6. *(a)* *Di antara borang-borang cukai pendapatan bagi tahun tertentu, purata cukai terbayar adalah RM 2000 dan sisihan piawai adalah RM 500. Tambahan pula, didapati 10% daripada borang-borang tersebut, cukat terbayar melebihi RM 3000. Suatu samplel rawak 625 borang telah diambil.*

     (i)   *Apakah kebarangkalian bahawa purata cukai terbayar bagi sampel borang tersebut melebihi RM 1980?*

     (ii)   *Apakah kebarangkalian bahawa lebih daripada 60 sampel borang yang mempunyai cukai terbayar melebihi RM 3000?*

 (b) *Pengukuran rintangan telah dilakukan pada 81 kabel dari jenis tertentu. Didapati min rintangan sampel kabel tersebut adalah 17.3 mΩ dengan varians 1.44 mΩ.*

     (i)   *Bina suatu selang keyakinan 98% bagi min rintangan kabel tersebut.*

     (ii)   *Apakah aras keyakinan bagi selang (17.1, 17.5)?*

     (iii)   *Berapakah saiz sampel jika pada aras keyakinan 98%, min rintangan dalam lingkungan ± 0.1 mΩ ?*

                                                        [12 markah]

7.   *Suatu proses pengeluaran yang baru telah difikirkan bagi pengeluaran keluli bering. Suatu sampel rawak ukuran diameter bering telah diambil dari proses lama X dan proses baru Y. Data yang diperolehi adalah seperti berikut:*

Proses lama, X

16.3	*15.9*	*15.8*	*16.2*	*16.1*	*16.0*
15.7	*15.8*	*15.9*	*16.1*	*16.3*	*16.1*
15.8	*15.7*	*15.8*	*15.7*		

Proses baru, Y

15.9	*16.2*	*16.0*	*15.8*	*16.1*	*16.1*
15.8	*16.0*	*16.2*	*15.9*	*15.7*	*16.2*
15.8	*15.8*	*16.2*	*16.3*		

$$\sum_i x_i = 255.2, \quad \sum_i x_i^2 = 4071.1, \quad \sum_i y_i = 256, \quad \sum_i y_i^2 = 4096.54$$

     (i)   *Adakah kita dapat membuat kesimpulan pada aras 5% bahawa terdapat perbezaan min saiz diameter bearing antara kedua-dua proses pengeluaran?*

     (ii)   *Adakah kita dapat membuat kesimpulan pada aras 5% bahawa varians dari proses terbaru lebih rendah daripada varians proses lama?*

                                                        [15 markah]

8. A chemical manufacturing company is planning to locate a hazardous waste disposal site near a city of 20,000 residents. They conducted a poll on two hundred adults (110 women and 90 men) who are residents of the city randomly to know their opinion. Sixty percent of those polled appose the site, 32% are in favor and 8% are undecided. Of those who oppose the site, 65% are women; of those in favor, 62.5% are men.

(i)   Construct a contingency table to summarize the information given.
(ii) Using the 2.5% significance level, can you conclude that opinions on the disposal site are dependent on gender?

[10 marks]

9. A manufacturer is interested in the output voltage of a power supply used in a computer. Output voltage is assumed to be normally distributed, with variance 0.0625 V, and the manufacture wishes to test $H_o : \mu = 5\ V$ against $H_o : \mu \neq 5\ V$, using $n=8$.

(i)   If the acceptance region is $4.85 \leq \overline{X} \leq 5.15$, find the value of α.
(ii)   Find the power of the test for detecting a true mean output voltage of 5.1 V.
(iii) Suppose that the manufacture wants the type I error probability for the test to be 0.05. What is the acceptance region?

[6 marks]

10. A researcher at USM suspects that the grades of applied statistics majors tend to be lower in the second semester than in the first semester. He randomly selects 10 applied statistics majors and records their grade point averages (GPAs) for the first and the second semester. The data obtained are shown in the table. Using the Wilcoxon signed-rank test at 5% level of significance, can you conclude that the median GPA of all applied statistics majors tends to be lower in the second semester than in the first semester?

Student	1	2	3	4	5	6	7	8	9	10
First GPA	3.20	3.56	3.05	3.78	4.00	2.85	3.33	2.67	3.00	3.67
Second GPA	3.15	3.40	2.88	3.67	4.00	3.00	3.30	3.05	2.95	3.50

[7 marks]

8. Sebuah syarikat kimia bercadang membina kawasan pembuangan sisa kimia berbahaya berdekatan sebuah bandaraya dengan 20,000 penghuni. Syarikat tersebut telah menjalankan pungutan suara ke atas 200 penghuni (110 wanita dan 90 lelaki) secara rawak untuk mendapatkan pandangan mereka. Didapati enam puluh peratus yang mengundi membantah pembinaan kawasan pembuangan tersebut, 32% bersetuju dan 8% tidak pasti. Manakala dari kalangan yang membantah, 65% adalah wanita dan dari kalangan yang bersetuju pula 62.5% adalah lelaki.

(i) Bina jadual kontingensi untuk meringkaskan maklumat yang telah diberi.

(ii) Pada aras keertian 2.5%, dapatkah disimpul bahawa pandangan terhadap pembinaan kawasan pembuangan tersebut bersandar kepada jantina pengundi?

[10 markah]

9. Seorang pengeluar berminat mengetahui voltan yang dihasilkan oleh pembekal kuasa sebuah komputer. Voltan yang dihasilkan dianggap bertaburan normal dengan varians 0.0625 V dan pengeluar juga ingin menguji $H_o : \mu = 5$ V terhadap $H_o : \mu \neq 5$ V menggunakan $n=18$.

(i) Jika kawasan penerimaan $4.85 \leq \bar{X} \leq 5.15$, dapatkan nilai α.

(ii) Dapatkan kuasa ujian jika min sebenar pengeluarn voltan adalah 5.1 V.

(iii) Jika pengeluar menetapkan ralat jenis I sebagai 0.05, apakah lokasi kawasan penerimaan?

[6 markah]

10. Seorang penyelidik di USM mensyaki bahawa gred pelajar major Statistik Gunaan rendah pada semester kedua berbanding semester pertama. Beliau telah memilih secara rawak 10 pelajar major Statistik Gunaan dan mengambil rekod purata nilai gred (PNG) pada semester pertama dan semester kedua. Data-data yang didapati diberikan dalam jadual. Berdasarkan ujian Pangkat dan Tanda Wilcoxon pada aras keertian 5%, dapatkah disimpulkan bahawa median PNG semua pelajar major Statistik Gunaan adalah rendah pada semester kedua berbanding semester pertama?

Pelajar	1	2	3	4	5	6	7	8	9	10
PNG pertama	3.20	3.56	3.05	3.78	4.00	2.85	3.33	2.67	3.00	3.67
PNG kedua	3.15	3.40	2.88	3.67	4.00	3.00	3.30	3.05	2.95	3.50

[7 markah]

21

APPENDIX

Confidence Interval

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

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

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

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

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

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

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

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

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

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

Test Statistic

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Table A5   Table of Critical T Values for Wilcoxon's Signed-Ranks
and Matched-Pairs Signed-Ranks Test

	One-tailed level of significance					One-tailed level of significance			
	.05	.025	.01	.005		.05	.025	.01	.005
	Two-tailed level of significance					Two-tailed level of significance			
	.10	.05	.02	.01		.10	.05	.02	.01
n					n				
5	0	–	–	–	28	130	116	101	91
6	2	0	–	–	29	140	126	110	100
7	3	2	0	–	30	151	137	120	109
8	5	3	1	0	31	163	147	130	118
9	8	5	3	1	32	175	159	140	128
10	10	8	5	3	33	187	170	151	138
11	13	10	7	5	34	200	182	162	148
12	17	13	9	7	35	213	195	173	159
13	21	17	12	9	36	227	208	185	171
14	25	21	15	12	37	241	221	198	182
15	30	25	19	15	38	256	235	211	194
16	35	29	23	19	39	271	249	224	207
17	41	34	27	23	40	286	264	238	220
18	47	40	32	27	41	302	279	252	233
19	53	46	37	32	42	319	294	266	247
20	60	52	43	37	43	336	310	281	261
21	67	58	49	42	44	353	327	296	276
22	75	65	55	48	45	371	343	312	291
23	83	73	62	54	46	389	361	328	307
24	91	81	69	61	47	407	378	345	322
25	100	89	76	68	48	426	396	362	339
26	110	98	84	75	49	446	415	379	355
27	119	107	92	83	50	466	434	397	373

- ooo O ooo -