
UNIVERSITI SAINS MALAYSIA

First Semester Examination
2010/2011 Academic Session

November 2010

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

Duration : 3 hours
[Masa : 3 jam]

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

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

Instructions: Answer **all eight** [8] questions.

Arahan: Jawab **semua lapan** [8] soalan.]

In the event of any discrepancies, the English version shall be used.

[Sekiranya terdapat sebarang percanggahan pada soalan peperiksaan, versi Bahasa Inggeris hendaklah diguna pakai].

1. The following table shows the frequency distribution for the number of times of 80 students, randomly selected from those who stay at RST accommodation in USM main campus, uses the commuter service per week.

No. of times use commuter	Frequency
0 – 4	8
5 – 8	13
9 – 12	20
13 – 16	25
17 – 20	14

- (a) Estimates the mean and median for the number of times students staying at RST uses the commuter service to lecture per week.
- (b) (i) Find the percentile rank for number of times that equal to 15.
(ii) Calculate the 85th percentiles and give a brief interpretation of the value.

[20 marks]

2. (a) A class is given a list of 20 problems from which 10 will be part of the upcoming exam. If a particular student knows how to solve 15 of the problems, find the probability that the student will be able to answer:
- (i) Exactly 8 questions on the exam
(ii) At most 2 questions on the exam
- (b) A car dealer has kept records of the customers who visited his showroom. 30% of the customers who visited his showroom were women. Furthermore, his records show that 24% of the women who visited his showroom ended up buying a car while only 14% of the men ended up buying a car. If, while away for a lunch break, a customer came to the showroom and talk to the assistant but did not buy a car, what is the probability that the customer was a man?

[20 marks]

1. Jadual berikut menunjukkan taburan bagi kekerapan 80 orang pelajar, yang dipilih secara rawak daripada mereka yang tinggal di tempat penginapan RST di kampus utama USM, menggunakan perkhidmatan komuter setiap minggu.

Bilangan menggunakan komuter	Kekerapan
0 – 4	8
5 – 8	13
9 – 12	20
13 – 16	25
17 – 20	14

- (a) Anggarkan purata dan median bagi bilangan perkhidmatan komuter yang digunakan oleh pelajar yang tinggal di RTS
- (b) (i) Cari pangkat persentil bagi bilangan kali sama dengan 15.
(ii) Hitung persentil ke-85 dan beri interpretasi ringkas bagi nilai tersebut.
[20 markah]
2. (a) Satu kelas telah diberi senarai 20 masalah yang mana 10 akan menjadi sebahagian daripada peperiksaan yang akan diadakan. Jika seorang pelajar tahu menyelesaikan 15 daripada masalah tersebut, dapatkan kebarangkalian bahawa pelajar tersebut boleh menjawab:
- (i) Hanya 8 soalan dalam peperiksaan
(ii) Paling banyak 2 soalan dalam peperiksaan
- (b) Seorang pengedar kereta telah menyimpan rekod para pelanggan yang datang melawat bilik persembahan beliau. 30% daripada pelanggan yang datang melawat bilik pameran adalah wanita. Tambahan pula, rekod beliau menunjukkan bahawa 24% daripada wanita yang datang melawat akhirnya membeli sebuah kereta sementara hanya 14% daripada lelaki yang melawat akhirnya membeli sebuah kereta. Jika, sementara keluar untuk makan tengahari, seorang pelanggan datang ke bilik pameran dan berbual dengan seorang pembantu tetapi tidak membeli sebuah kereta, apakah kebarangkalian bahawa pelanggan tersebut adalah seorang lelaki?

[20 markah]

3. (a) Matt Cronin, a tennis analyst, has a rate of 0.125 for predicting the wrong winner of a tennis match. In a Grand Slam tournament there are 120 matches prior to quarter-final (64 matches in the 1st round, 32 matches in the 2nd round, 16 matches in the 3rd round and 8 matches in the 4th round). Calculate the probability that he predicted the winner:
- (i) correctly in more than 12 matches in the 3rd round
 - (ii) wrongly in between 12 and 24 matches in the 1st round.
- (b) Failure of a traffic light after a heavy thunderstorm at the Sunny Point intersection near USM occurs at random, at an average of twice per week. Find the probability that more than 75 failures will occur in a year. (Assume that a year consists of 52 week period).
- [20 marks]
4. (a) Assuming a 40-hour per week working period, the number of hours a USM lecturer spends on supervision is normally distributed with a mean of 16 hours and a standard deviation of 5 hours.
- (i) What is the probability that a randomly selected lecturer spends between 9 and 12 hours per week on research supervision?
 - (ii) If a sample of 35 lecturers is randomly selected from five schools under the Faculty of Science, what is the probability that the mean hours spend on research supervision is between 18 and 24 hours per week?
- (b) As part of a promotion for a new type of potato chips, free samples are offered to shoppers in a local supermarket. The probability that a shopper will buy a packet of potato chips after tasting the free sample is 0.20. Suppose, \hat{p} is the proportion of the next n shoppers buying a packet of potato chips after tasting a free sample. Assuming different shoppers can be regarded as independent trials, how large should n be so that the standard error of \hat{p} is no more than 0.05?

[20 marks]

3. (a) *Matt Cronin, seorang penganalisa tenis, mempunyai kadar 0.125 bagi meramal pemenang yang salah untuk sesuatu perlawanan tenis. Dalam suatu pertandingan Grand Slam terdapat 120 perlawanan sebelum ke peringkat suku akhir (64 perlawanan dalam pusingan pertama, 32 perlawanan dalam pusingan kedua, 16 perlawanan dalam pusingan ketiga dan 8 perlawanan dalam pusingan ke-empat). Hitung kebarangkalian bahawa ramalan beliau terhadap pemenang:*
- (i) *adalah betul bagi lebih daripada 12 perlawanan dalam pusingan ketiga*
 - (ii) *adalah salah bagi antara 12 dan 24 perlawanan dalam pusingan pertama*
- (b) *Kerosakan lampu isyarat selepas hujan ribut yang lebat di persimpangan Sunny Point berhampiran USM berlaku secara rawak, dengan purata dua kali seminggu. Dapatkan kebarangkalian bahawa lebih daripada 75 kerosakan berlaku dalam setahun. (Andaikan bahawa setahun mengandungi tempoh masa 52 minggu)*

[20 markah]

4. (a) *Mengandaikan tempoh bekerja 40-jam seminggu, bilangan jam seseorang pensyarah USM memberi penyeliaan adalah tertabur secara normal dengan min 16 jam dan sisihan piawai 5 jam.*
- (i) *Apakah kebarangkalian bahawa seorang pensyarah yang dipilih secara rawak memberi penyeliaan penyelidikan di antara 9 dan 12 jam seminggu?*
 - (ii) *Sekiranya suatu sampel 35 orang pensyarah dipilih secara rawak dari lima pusat pengajian di bawah Fakulti Sains, apakah kebarangkalian bahawa min jam yang diberi bagi penyeliaan penyelidikan adalah di antara 18 dan 24 jam seminggu?*
- (b) *Sebahagian daripada promosi bagi suatu jenis baru keropok kentang, sampel percuma diberikan kepada para pengunjung di suatu pasaraya tempatan. Kebarangkalian bahawa seorang pengunjung akan membeli satu paket keropok kentang selepas merasa sampel percuma adalah 0.20. Andaikan, \hat{p} adalah pecahan bagi n pengunjung seterusnya yang membeli satu paket keropok kentang selepas merasa sampel percuma. Dengan andaian bahawa setiap pengunjung boleh dianggapkan sebagai percubaan tak bersandar, berapa besarkah n diperlukan supaya ralat piawai bagi \hat{p} tidak lebih daripada 0.05?*

[20 markah]

5. (a) A clothing manufacturing company produces women's blouses with 4 buttons. A big-buyer notices that only 60% of the blouses he bought have full buttons, 30% of the blouses have 1 button missing while 9% have 2 buttons missing. The good news is that the buyer has never found blouses with all the buttons missing. If you are buying a new blouse from the manufacturer, find the expected number of buttons missing on the blouse.
- (b) A certain car company claims that at most 8% of its new car have a manufacturing defect. A quality control inspector randomly selects 300 new cars and finds that 33 of them have a defect. Find the 99% confidence interval for the true proportion of newly manufactured car from the company having a defect. Based on the confidence interval, should the 8% claim be rejected at 1% significance level?

[20 marks]

6. It is claimed that products promoted through a magazine would be more attractive if the advertisements are more readable. Market researchers use the number of sentences per advertisement, X as a measure of readability for magazine advertisements, whereby the advertisements are more attractive when having lesser number of sentences. The number of sentences found in a randomly selected sample of 12 advertisements in travel magazine and 18 advertisements in sports magazines are listed below.

Travel:	16	9	14	11	17	12
	39	18	13	12	25	9
	$\sum X = 195$			$\sum X^2 = 3951$		

Sports:	17	6	11	17	18	20	6	14	7
	11	12	5	18	6	4	13	11	12
	$\sum X = 208$				$\sum X^2 = 2840$				

- (a) Does the data provide enough evidence that the population of travel and sports advertisements have equal variance of the number of sentences? Solve using the critical region approach and at 10% significance level.
- (b) Construct a 90% confidence interval for the difference between the mean number of sentences in travel and sports magazines.
- (c) Using the p-value approach at 10% significance level, test the claim that more sports products are sold through magazines compared to those of traveling products as the advertisements are shorter and therefore more readable and attractive.

[40 marks]

5. (a) Suatu syarikat mengeluarkan pakaian menghasilkan blaus wanita dengan 4 butang. Seorang pembeli pukul memerhatikan bahawa hanya 60% daripada blaus yang dibeli mempunyai butang penuh, 30% daripada blaus kehilangan 1 butang sementara 9% hilang dua butang. Walau bagaimanapun pembeli tersebut tidak pernah menjumpai blaus dengan semua butang hilang. Sekiranya anda membeli blaus baru daripada pengeluar tersebut, cari jangkakan bilangan butang yang hilang pada blaus tersebut.
- (b) Suatu syarikat kereta mendakwa bahawa paling banyak 8% daripada kereta baru yang dikeluarkan mempunyai kecacatan pembuatan. Seorang pemeriksa kawalan kualiti memilih secara rawak 300 kereta baru dan mendapati 33 daripadanya mempunyai kecacatan. Dapatkan selang keyakinan 99% bagi nilai pecahan sebenar kereta yang baru dikeluarkan oleh syarikat tersebut mempunyai kecacatan. Berdasarkan selang keyakinan tersebut, patutkah dakwaan 8% ditolak pada aras kesignifikanan 1%?
- [20 markah]
6. Adalah didakwa bahawa hasil keluaran yang dipromosikan melalui majalah akan lebih menarik perhatian sekiranya iklan terbabit adalah lebih mudah dibaca. Penyelidik pemasaran menggunakan bilangan perkataan bagi setiap iklan, X sebagai ukuran tahap pembacaan bagi iklan majalah, yang mana iklan adalah lebih menarik apabila kurang mempunyai bilangan ayat. Bilangan ayat yang dijumpai pada suatu sampel rawak 12 iklan di dalam majalah pelancongan and 18 iklan di dalam majalah sukan adalah disenaraikan di bawah.

$$\begin{array}{l} \text{Pelancongan:} \\ 16 \quad 9 \quad 14 \quad 11 \quad 17 \quad 12 \\ 39 \quad 18 \quad 13 \quad 12 \quad 25 \quad 9 \\ \Sigma X = 195 \qquad \qquad \qquad \Sigma X^2 = 3951 \end{array}$$

$$\begin{array}{l} \text{Sukan:} \\ 17 \quad 6 \quad 11 \quad 17 \quad 18 \quad 20 \quad 6 \quad 14 \quad 7 \\ 11 \quad 12 \quad 5 \quad 18 \quad 6 \quad 4 \quad 13 \quad 11 \quad 12 \\ \Sigma X = 208 \qquad \qquad \qquad \Sigma X^2 = 2840 \end{array}$$

- (a) Adakah data memberi bukti yang cukup bahawa populasi iklan perlancongan dan sukan mempunyai varians bilangan ayat yang sama? Selesaikan menggunakan pendekatan rantau penolakan dan pada aras kesignifikanan 10%.
- (b) Bina suatu selang keyakinan 90% untuk perbezaan di antara purata bilangan ayat di dalam majalah perlancongan dan sukan.
- (c) Menggunakan pendekatan nilai-p pada aras kesignifikanan 10%, uji dakwaan bahawa lebih hasil keluaran barangan sukan dijual melalui majalah berbanding dengan hasil barangan perlancongan disebabkan iklannya adalah lebih pendek dan oleh itu lebih mudah dibaca dan menarik.

[40 markah]

7. (a) The manager of a clothing store claims that during a 3-day Merdeka sale, the number of visitors on each of the last two days is twice as large as the number of customers on the first day. Featured below are numbers of visitors at the store during the recent 3-day Merdeka sale. Do these data support the claim made by the store manager?

Day	1	2	3
Frequency	197	326	435

- (b) A factory dietician wants to know if the tendency to consume coffee is the same throughout the day. The following data represents the beverage purchases by a random sample of customers at Intel Penang's cafeteria.

Type of beverage	Early morning	Late morning	Early afternoon	Late afternoon
Coffee	9	6	12	8
Others	51	54	48	52

Using $\alpha = 0.05$, what conclusion can be made by the dietician?

[30 marks]

8. A set of 12 identical twins is given a psychological test to determine whether the first born of the twins tends to be more aggressive than the second born. Each twin is scored according to aggressiveness with a higher score indicates greater level of aggressiveness.

Twin-index	1 st born	2 nd born
1	86	88
2	71	77
3	77	76
4	68	64
5	91	96
6	72	72
7	77	65
8	91	90
9	70	65
10	71	80
11	88	81
12	87	72

Assuming the population distribution is not normal, test at 5% significance level if there is evidence that older twin tends to be more aggressive than a younger twin.

[30 marks]

7. (a) *Pengurus sebuah stor pakaian mendakwa bahawa semasa 3-hari Jualan Merdeka, bilangan pengunjung pada setiap daripada dua hari terakhir adalah dua kali lebih ramai daripada bilangan pengunjung pada hari pertama. Paparan di bawah adalah bilangan pengunjung di stor tersebut semasa 3-hari Jualan Merdeka yang lepas. Adakah data ini menyokong dakwaan oleh pengurus stor tersebut?*

<i>Hari</i>	<i>1</i>	<i>2</i>	<i>3</i>
<i>Kekerapan</i>	<i>197</i>	<i>326</i>	<i>435</i>

- (b) *Seorang pakar pemakanan disebuah kilang ingin tahu sekiranya kecenderongan meminum kopi adalah sama di sepanjang hari. Data berikut mewakili pembelian minuman oleh sampel rawak pelanggan di kafeteria Intel Penang*

<i>Jenis minuman</i>	<i>Awal pagi</i>	<i>Lewat pagi</i>	<i>Awal petang</i>	<i>Lewat petang</i>
<i>Kopi</i>	<i>9</i>	<i>6</i>	<i>12</i>	<i>8</i>
<i>Lain-lain</i>	<i>52</i>	<i>51</i>	<i>44</i>	<i>54</i>

Menggunakan $\alpha = 0.05$, apakah kesimpulan yang boleh dibuat oleh pakar pemakanan itu?

[30 markah]

8. *Suatu set 12 orang kembar yang sama telah diberi ujian psikologi untuk menentukan samada kembar kelahiran pertama lebih agresif daripada kelahiran kedua. Setiap kembar diberi markah berdasarkan tahap agresif dengan markah yang lebih tinggi menunjukkan tahap agresif yang lebih besar.*

<i>Index kembar</i>	<i>1st born</i>	<i>2nd born</i>
<i>1</i>	<i>86</i>	<i>88</i>
<i>2</i>	<i>71</i>	<i>77</i>
<i>3</i>	<i>77</i>	<i>76</i>
<i>4</i>	<i>68</i>	<i>64</i>
<i>5</i>	<i>91</i>	<i>96</i>
<i>6</i>	<i>72</i>	<i>72</i>
<i>7</i>	<i>77</i>	<i>65</i>
<i>8</i>	<i>91</i>	<i>90</i>
<i>9</i>	<i>70</i>	<i>65</i>
<i>10</i>	<i>71</i>	<i>80</i>
<i>11</i>	<i>88</i>	<i>81</i>
<i>12</i>	<i>87</i>	<i>72</i>

Andaikan taburan populasi adalah tidak normal, uji pada aras kesignifikanan 5% sekiranya terdapat bukti bahawa kembar yang lebih tua adalah lebih agresif berbanding kembar yang lebih muda.

[30 markah]

APPENDIX

Confidence Interval

$\bar{X} \pm Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$ $\bar{X} \pm t_{\frac{\alpha}{2}} \frac{s}{\sqrt{n}}$ $\bar{d} \pm t_{\frac{\alpha}{2}} \frac{s_d}{\sqrt{n_d}}$ $b \pm t_{\frac{\alpha}{2}} s_b$	$(\bar{X} - \bar{Y}) \pm t_{\alpha/2} S_p \sqrt{\frac{1}{n_x} + \frac{1}{n_y}}$
$\hat{p} \pm Z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$ $(\hat{p}_x - \hat{p}_y) \pm z_{\alpha/2} \sqrt{\frac{\hat{p}_x(1-\hat{p}_x)}{n_x} + \frac{\hat{p}_y(1-\hat{p}_y)}{n_y}}$ $(\bar{X} - \bar{Y}) \pm Z_{\alpha/2} \sqrt{\frac{\sigma_x^2}{n_x} + \frac{\sigma_y^2}{n_y}}$ $(\bar{X} - \bar{Y}) \pm t_{\alpha/2} \sqrt{\frac{s_x^2}{n_x} + \frac{s_y^2}{n_y}}$	$\left(\frac{(n-1)s^2}{\chi_{\frac{\alpha}{2}, n-1}^2}, \frac{(n-1)s^2}{\chi_{1-\frac{\alpha}{2}, n-1}^2} \right)$ $\left(\frac{s}{Z_{\frac{\alpha}{2}}}, \frac{s}{Z_{\frac{\alpha}{2}}} \right)$ $\left(1 + \frac{\frac{\alpha}{2}}{\sqrt{2n}}, 1 - \frac{\frac{\alpha}{2}}{\sqrt{2n}} \right)$ $\left(\frac{s_1^2}{s_2^2} F_{1-\frac{\alpha}{2}, (v_2, v_1)}, \frac{s_1^2}{s_2^2} F_{\frac{\alpha}{2}, (v_2, v_1)} \right)$

Test Statistic

$Z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}}$ $T = \frac{\bar{X} - \mu}{s / \sqrt{n}}$ $T = \frac{\bar{d} - \mu_d}{s_d / \sqrt{n_d}}$ $T = \frac{b - \beta_1}{s_b}$ $T = r \sqrt{\frac{n-2}{1-r^2}}$ $\chi^2 = \frac{(n-1)s^2}{\sigma^2}$	$Z = \frac{s - \sigma}{\sigma / \sqrt{2n}}$ $Z = \frac{(\bar{X} - \bar{Y}) - (\mu_x - \mu_y)}{\sqrt{\frac{\sigma_x^2}{n_x} + \frac{\sigma_y^2}{n_y}}}$ $T = \frac{(\bar{X} - \bar{Y}) - (\mu_x - \mu_y)}{\sqrt{S_p^2 \left(\frac{1}{n_x} + \frac{1}{n_y} \right)}}$ $S_p^2 = \frac{(n_x - 1)s_x^2 + (n_y - 1)s_y^2}{n_x + n_y - 2}$ $F = \frac{s_x^2}{s_y^2}$	$Z = \frac{(\hat{p}_x - \hat{p}_y) - (p_x - p_y)}{\sqrt{\hat{p}(1-\hat{p}) \left(\frac{1}{n_x} + \frac{1}{n_y} \right)}}$ $T = \frac{(\bar{X} - \bar{Y}) - (\mu_x - \mu_y)}{\sqrt{\frac{s_x^2}{n_x} + \frac{s_y^2}{n_y}}}$ $dk = \frac{\left(\frac{s_x^2}{n_x} + \frac{s_y^2}{n_y} \right)^2}{\frac{\left(\frac{s_x^2}{n_x} \right)^2}{n_x - 1} + \frac{\left(\frac{s_y^2}{n_y} \right)^2}{n_y - 1}}$ $\chi^2 = \sum \frac{(O - E)^2}{E}, \quad E = np$
---	--	---

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

<i>n</i>	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	
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