
UNIVERSITI SAINS MALAYSIA

First Semester Examination
2013/2014 Academic Session

December 2013/January 2014

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. Though less rapidly than in other developed countries, the population of the United States of America (US) is aging. Below is the stem-and-leaf of the percents of population aged 65 and over in the 50 states, according to 2010 census. Over the years, it has been known that Alaska has the lowest percent of older populations while Florida has the highest. The leaves are tenths of a percent.

5 7
6
7
8 5
9 6 7 9
10 6
11 0 2 2 3 3 6 7 7
12 0 0 1 1 1 1 3 4 4 5 7 8 9
13 0 0 0 1 2 2 3 3 3 4 5 5 6 8
14 0 3 4 5 7 9
15 3 6
16
17 6

- (a) Find the mean and the three quartiles for the percentage of US population aged 65 and over.
- (b) Construct a boxplot for the percentage of US population aged 65 and over. What can you say about the percentage of older population in Alaska and Florida?

[20 marks]

1. *Walaupun kurang pantas berbanding di negara membangun yang lain, populasi Amerika Syarikat sedang mengalami penuaan. Di bawah adalah carta dahan-dan-daun bagi peratusan populasi berumur 65 tahun ke atas di 50 buah negeri, berdasarkan banci 2010. Selama ini, adalah sedia maklum bahawa Alaska mempunyai peratusan terendah populasi tua sementara Florida mempunyai peratusan tertinggi. Daun ialah persepuhluh daripada peratus.*

5 / 7
6 /
7 /
8 / 5
9 / 6 7 9
10 / 6
11 / 0 2 2 3 3 6 7 7
12 / 0 0 1 1 1 1 3 4 4 5 7 8 9
13 / 0 0 0 1 2 2 3 3 3 4 5 5 6 8
14 / 0 3 4 5 7 9
15 / 3 6
16 /
17 / 6

- (a) *Cari mean dan tiga kuartil bagi peratusan populasi Amerika Syarikat yang berumur 65 tahun ke atas.*

- (b) *Bina plot kotak untuk peratusan populasi Amerika Syarikat yang berumur 65 tahun ke atas. Apakah yang boleh diperkatakan tentang peratusan populasi tua di Alaska dan Florida?*
[20 markah]
2. (a) In the ATP World Final Championship, the top eight tennis players are to be divided into two groups of four players. While the top two players leads the groups, the third and fourth ranked players cannot be in the same group. In addition, the fifth and sixth ranked players as well as the seventh and eighth ranked players cannot be in the same group. How many ways the two groups can be formed?

(b) Jason is attending his high school reunion, the first time since leaving the school 10 years ago. Sixty percent of the attendees are female. Seventy-five percent are female or married. If twenty percent of the attendees are married, what percentage of the attendees is both female and married?
[20 marks]
2. (a) *Di Kejohanan Akhir Dunia ATP, lapan pemain tenis teratas dibahagikan kepada dua kumpulan yang mengandungi empat pemain. Sementara dua pemain teratas mengetuai dua kumpulan, pemain menduduki tempat ketiga dan keempat tidak boleh berada dalam kumpulan yang sama. Tambahan pula, pemain yang menduduki tempat kelima dan keenam serta pemain menduduki tempat ketujuh dan kelapan tidak boleh berada dalam kumpulan yang sama. Berapa carakah dua kumpulan tersebut boleh dibentuk?*

(b) *Jason menghadiri perjumpaan sekolah menengah beliau, kali pertama sejak meninggalkan sekolah 10 tahun yang lalu. Enam puluh peratus daripada yang hadir adalah perempuan. Tujuh puluh lima peratus adalah perempuan atau telah berkahwin. Jika dua puluh peratus daripada yang hadir telah berkahwin, apakah peratusan yang hadir adalah perempuan dan telah berkahwin?*
[20 markah]
3. (a) (i) The probability that heavy rain is in the vicinity of Penang International Airport (PIA) in a particular day in December is 0.70, The probability that heavy rain is in the vicinity of PIA and a plane lands on time is only 0.56. Find the probability that a plane lands on time when heavy rain is in the vicinity.
(ii) Sixty percent of students in a particular class are female. While nineteen percent of the male students possess the latest Samsung Galaxy S4 mobile phone, only three and a half percent of the female students possess the same brand of phone. If a randomly selected student is found to possess a Galaxy S4, find the probability that the student is a male.

- (b) "How many smart phones are there in your household?" was one of the questions in a household survey questionnaire sent to 1000 addresses in Sabah. The collected data resulted in the following distribution.

No. of smart phone	0	1	2	3	4
Percentage	1.9	37.5	35.8	18.3	6.5

Let X is a random variable for the number of smart phones per household in Sabah. Find the mean and standard deviation of X .

[30 marks]

3. (a) (i) Kebarangkalian bahawa hujan lebat di sekitar Lapangan Terbang Antarabangsa Pulau Pinang (LTAPP) pada hari tertentu pada bulan Disember adalah 0.70. Kebarangkalian bahawa hujan lebat disekitar LTAPP dan kapal terbang mendarat tepat pada waktunya hanyalah 0.56. Cari kebarangkalian bahawa kapal terbang mendarat tepat pada waktunya apabila hujan lebat adalah di sekitar lapangan terbang.
(ii) Enam puluh peratus daripada para pelajar dalam kelas tertentu adalah perempuan. Sementara sembilan belas peratus daripada pelajar lelaki memiliki telefon pintar Samsung Galaxy S4 yang terkini, hanya tiga setengah peratus daripada pelajar perempuan memiliki jenama telefon yang sama. Jika seorang pelajar yang dipilih secara rawak didapati memiliki Galaxy S4, cari kebarangkalian bahawa pelajar tersebut adalah seorang lelaki.
- (b) "Berapa banyak telefon pintar terdapat dalam isirumah anda?" adalah satu daripada soalan dalam kajian soal selidik isirumah yang telah dihantar ke 1000 alamat di Sabah. Data yang dikumpul menghasilkan taburan berikut.

Bil. telefon pintar	0	1	2	3	4
Peratusan	1.9	37.5	35.8	18.3	6.5

Andaikan X adalah pembolehubah rawak bagi bilangan telefon pintar setiap isirumah. Cari min dan sisihan piawai bagi X .

[30 markah]

4. (a) Like most companies, Pos Malaysia states that providing high quality services to customers is the top priority of the management. Despite all efforts, past services show that 20% of the documents sent through the Pos Laju service do not arrive at their destination within 48 hours.
(i) Suppose that a particular school in USM sends 10 documents through the Pos Laju service on a certain working day. Find the probability that at most two of the 10 documents will not arrive with 48 hours.
(ii) If a sample of 50 documents sent through Pos Laju in a given week is randomly selected, what is the probability that between 7 and 9 documents are not delivered within 48hours?

- (b) An average of 4 accidents occur per week along the Penang Bridge.
- (i) Find the probability that at least 2 accidents occur on the bridge in any given week.
- (ii) Find the probability that between 30 and 35 accidents occur in any given 8 weeks period.
- (c) A box of 12 cupcakes are packed randomly with two flavours, chocolate chips and blueberry. However, all cupcakes have the same colour of wrapper. If four cupcakes are randomly selected from a box with seven blueberry cupcakes and checked for quality control, find the probability that at most three of them are chocolate chips cupcakes.

[30 marks]

4. (a) Seperti kebanyakan syarikat, Pos Malaysia menyatakan bahawa memberikan servis berkualiti tinggi kepada pelanggan adalah keutamaan bagi pihak pengurusan. Disebalik segala usaha, servis terdahulu menunjukkan bahawa 20% daripada dokumen yang dihantar melalui servis Pos Laju tidak sampai di destinasi dalam jangkamasa 48 jam.
- (i) Andaikan bahawa sebuah pusat pengajian tertentu di USM menghantar 10 dokumen melalui servis Pos Laju dalam suatu hari bekerja tertentu. Cari kebarangkalian bahawa paling banyak dua daripada 10 dokumen tersebut tidak akan sampai dalam jangkamasa 48 jam.
- (ii) Sekiranya satu sampel 50 dokumen yang telah dihantar melalui Pos Laju dalam satu minggu tertentu dipilih secara rawak, apakah kebarangkalian bahawa antara 7 dan 9 dokumen tidak dihantar dalam jangkamasa 48 jam?
- (b) Secara purata 4 kemalangan berlaku setiap minggu disepanjang Jambatan Pulau Pinang.
- (i) Cari kebarangkalian bahawa sekurang-kurangnya 2 kemalangan berlaku di atas jambatan pada sesuatu minggu.
- (ii) Cari kebarangkalian bahawa di antara 30 hingga 35 kemalangan berlaku dalam jangkamasa 8 minggu tertentu.
- (c) Sekotak 12 kek cawan dipek secara rawak dengan dua rasa, cip choclat dan bluberi. Walau bagaimanapun, semua kek cawan mempunyai warna kertas bungkus yang sama. Sekiranya empat kek cawan dipilih secara rawak daripada satu kotak dengan tujuh kek cawan bluberi dan diperiksa untuk kawalan kualiti, cari kebarangkalian bahawa paling banyak tiga daripadanya adalah kek cawan bercip choclat.

[30 markah]

5. (a) The times that university students spend on revision per week have a distribution that is skewed to the right with a mean of 9 hours and standard deviation of 3 hours. Find the probability that the mean time spent revising per week by a randomly selected 35 students is at least 10 hours.
- (b) Hanis, who is running for student representative believes she is favoured by 63% of students at her school. If 300 students are randomly selected, find 99% confidence interval, the percentage of students who will be voting for her?

[20 marks]

5. (a) *Masa yang digunakan oleh pelajar universiti untuk mengulangkaji setiap minggu mempunyai taburan yang condong ke kanan dengan min 9 jam dan sisihan piawai 3 jam. Cari kebarangkalian bahawa min masa yang digunakan untuk mengulangkaji oleh 35 orang pelajar yang dipilih secara rawak adalah sekurang-kurangnya 10 jam.*

(b) *Hanis, yang akan bertanding untuk wakil pelajar percaya beliau disukai oleh 65% daripada para pelajar di pusat pengajian. Sekiranya 300 orang pelajar dipilih secara rawak, cari 99% selang keyakinan, peratusan pelajar yang akan mengundi beliau?*

[20 markah]

6. (a) The Ministry of Health claims that the average age of first-time mothers is now normally distributed with mean of 28 years old. A random sample of 20 first-time mothers is found to have a mean age of 26.5 years with standard deviation of 2 years. Construct a 95% confidence interval for the mean age of first-time mothers and use that to test the claim made by the ministry.

(b) The Ministry of Tourism and Culture claims that more than 75% of foreign tourists make a return visit to Malaysia less than 5 years after their first visit. A random sample of 2000 tourists was surveyed on arrival at the KLIA and 1700 of them stated that they are returning less than 5 years since their first visit. Find the p-value for the hypothesis being tested. What can be said about the claim made by the ministry?

[20 marks]

6. (a) *Kementerian Kesihatan mendakwa bahawa purata umur mereka yang menjadi ibu pertama kali sekarang ini tertabur normal dengan min 28 tahun. Suatu sampel rawak seramai 20 orang yang kali pertama menjadi ibu didapati mempunyai purata umur 26.5 tahun dengan sisihan piawai 2 tahun. Bina 95% selang keyakinan min umur bagi mereka yang kali pertama menjadi ibu dan gunakannya untuk menguji dakwaan yang dibuat oleh kementerian.*

(b) *Kementerian Pelancongan dan Kebudayaan mendakwa bahawa lebih 75% daripada pelancong asing melakukan lawatan semula ke Malaysia kurang 5 tahun selepas lawatan pertama mereka. Suatu sampel rawak seramai 2000 orang pelancong telah di soal selidik semasa mendarat di KLIA dan 1700 daripada mereka menyatakan bahawa mereka kembali kurang daripada 5 selepas lawatan pertama. Cari nilai-p bagi hipotesis yang diujji. Apakah yang boleh diperkatakan tentang dakwaan yang dibuat oleh kementerian?*

[20 markah]

7. The Ministry of Education claimed that while variation in the number of hours spent by secondary school students playing computer games is the same across all states in Malaysia, the average hours spent on computer games is higher in more developed states. A random sample of 22 secondary school students in Johor and 16 students in Kedah shows the following number of hours spent on computer games.

Johor:	37	27	36	31	30	30	29
	39	30	28	22	35	28	24
	40	26	23	30	26	27	32

Kedah:	20	23	24	28	21	22	24
	24	14	20	14	30	18	25

- (a) Using the p-value approach, does the sample provide enough evidence to reject the claim that the variation in the number of hours spent on playing computer games by secondary school students is the same in more developed state such as Johor and in less developed state such as Kedah?
- (b) Test the claim that the average number of hours spent on playing computer games by secondary school students in more developed states such as Johor is higher than that in less developed states such as Kedah. Based on the sample, what can you say about the amount of evidence to support the claim?

[30 marks]

7. Kementerian Pendidikan telah mendakwa bahawa sementara variasi dalam bilangan jam yang dihabiskan oleh pelajar sekolah menengah bermain permainan komputer adalah sama di kesemua negeri di Malaysia, purata jam yang dihabiskan bagi bermain permainan komputer adalah lebih tinggi di negeri yang lebih membangun di Semenanjung Malaysia. Suatu sampel rawak seramai 22 orang pelajar sekolah menengah di Johor dan 16 orang pelajar di Kedah menunjukkan bilangan jam yang dihabiskan bermain permainan komputer seperti berikut.

Johor:	37	27	36	31	30	30	29
	39	30	28	22	35	28	24
	40	26	23	30	26	27	32

Kedah:	20	23	24	28	21	22	24
	24	14	20	14	30	18	25

- (a) Menggunakan pendekatan nilai-p, adakah sampel memberi bukti yang cukup untuk menolak dakwaan bahawa variasi dalam bilangan jam yang dihabiskan bermain permainan komputer oleh pelajar sekolah menengah adalah sama di negeri yang lebih membangun seperti Johor dan di negeri yang kurang membangun seperti Kedah?
- (b) Uji dakwaan bahawa purata bilangan jam yang dihabiskan bermain permainan komputer oleh pelajar sekolah menengah di negeri yang lebih membangun seperti Johor adalah lebih tinggi berbanding dengan purata bilangan jam oleh pelajar di negeri yang kurang membangun seperti Kedah. Berdasarkan sampel, apakah yang boleh diperkatakan terhadap jumlah bukti untuk menyokong dakwaan?

[30 markah]

8. (a) A study is conducted to determine whether age of radio listener is related to the preferred singer of Hari Raya song. A sample of radio listeners yields the following data. Conduct the appropriate hypothesis testing and give your conclusion.

Age	S. M. Salim	Syarifah Aini	Sudirman Arshad	Siti Nurhaliza
20 - 34	8	9	14	29
35 - 49	16	29	39	13
50 - 64	38	22	21	17

- (b) KFC is creating a new marinating mix for its chicken rice meal. Prior to replacing the current flavour, the company would like to conduct a test to ensure that customers will like the new flavour better. The marketing manager has called in 15 regular customers and each customer has been given a set of the current chicken rice meal and asked to rate its overall taste on a scale of 1 to 20 with the value close to 20 indicates the customer have much liking to the current flavour. The same 15 customers were given a sample of the new set of chicken rice meal with the new marinating mix and again were asked to rate its taste on a scale of 1 to 20.

Customer	Current Flavour	New Flavour
Arianti	12	14
Amalina	16	8
Huzaifah	2	6
Jie Hui	4	18
Shukur	12	20
Hafiza	16	16
Shaheera	5	14
Shu Yee	16	6
Sutha	10	19
Wissely	10	18
Mei Shiang	13	16
Harith	2	18
Chun Kit	13	4
Wazir	14	7
Sakinah	4	16

The results from the testing are reported in the table above. Is it reasonable to conclude that the new marinating flavor is preferred?

[30 marks]

8. (a) *Suatu kajian dijalankan untuk menentukan samada umur pendengar radio adalah berhubung dengan penyanyi lagu Hari Raya yang disukai. Suatu sampel pendengar radio menghasilkan data seperti berikut. Jalankan ujian hipotesis yang sesuai dan berikan kesimpulan anda.*

Umur	S. M. Salim	Syarifah Aini	Sudirman Arshad	Siti Nurhaliza
20 – 34	8	9	14	29
35 – 49	16	29	39	13
50 – 64	38	22	21	17

- (b) KFC sedang mencipta suatu campuran memerap yang baru untuk juadah nasi ayam. Sebelum menggantikan perasa semasa, syarikat ingin menjalankan suatu ujian untuk memastikan bahawa pelanggan akan lebih menyukai perasa yang baru. Pengurus pemasaran telah memanggil 15 orang pelanggan tetap dan setiap pelanggan telah diberi satu set juadah nasi ayam semasa dan telah diminta untuk menilai rasa keseluruhan pada skala 1 hingga 20 dengan nilai yang hampir dengan 20 menandakan pelanggan mempunyai citarasa yang tinggi terhadap perasa semasa. 15 pelanggan yang sama telah diberikan satu set juadah nasi ayam dengan perasa memerap yang baru dan sekali lagi diminta untuk menilai rasanya pada skala 1 hingga 20.

<i>Customer</i>	<i>Perasa semasa</i>	<i>Perasa baru</i>
<i>Arianti</i>	12	14
<i>Amalina</i>	16	8
<i>Huzaifah</i>	2	6
<i>Jie Hui</i>	14	18
<i>Shukur</i>	12	20
<i>Hafiza</i>	16	16
<i>Shaheera</i>	5	14
<i>Shu Yee</i>	16	6
<i>Sutha</i>	10	19
<i>Wissely</i>	10	18
<i>Mei Shiang</i>	13	16
<i>Harith</i>	12	18
<i>Chun Kit</i>	13	4
<i>Wazir</i>	14	7
<i>Sakinah</i>	4	16

Keputusan daripada ujian yang dijalankan dilaporkan dalam jadual di atas. Adakah berpatutan untuk disimpulkan bahawa perasa yang baru adalah lebih disukai?

[30 markah]

APPENDIX

Confidence Interval

$\bar{X} \pm Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$ $\bar{d} \pm t_{\frac{\alpha}{2}} \frac{s_d}{\sqrt{n_d}}$	$\bar{X} \pm t_{\frac{\alpha}{2}} \frac{s}{\sqrt{n}}$ $b \pm t_{\frac{\alpha}{2}} s_b$	$(\bar{X} - \bar{Y}) \pm t_{\alpha/2} S_p \sqrt{\frac{1}{n_x} + \frac{1}{n_y}}$
$\hat{p} \pm Z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$ $(\hat{p}_x - \hat{p}_y) \pm z_{\alpha/2} \cdot \sqrt{\frac{\hat{p}_x(1-\hat{p}_x)}{n_x} + \frac{\hat{p}_y(1-\hat{p}_y)}{n_y}}$ $(\bar{X} - \bar{Y}) \pm Z_{\alpha/2} \cdot \sqrt{\frac{\sigma_x^2}{n_x} + \frac{\sigma_y^2}{n_y}}$ $(\bar{X} - \bar{Y}) \pm t_{\alpha/2} \cdot \sqrt{\frac{s_x^2}{n_x} + \frac{s_y^2}{n_y}}$	$\left(\frac{(n-1)s^2}{\chi_{\frac{\alpha}{2}, n-1}^2}, \frac{(n-1)s^2}{\chi_{1-\frac{\alpha}{2}, n-1}^2} \right)$ $\left(\frac{s}{Z_{\frac{\alpha}{2}}}, \frac{s}{Z_{\frac{\alpha}{2}}} \right)$ $\left(\frac{s_1^2}{S_2^2} F_{1-\frac{\alpha}{2}, (v_2, v_1)}, \frac{s_1^2}{S_2^2} F_{\frac{\alpha}{2}, (v_2, v_1)} \right)$	

Test Statistic

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

**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

- 000 O 000 -