
UNIVERSITI SAINS MALAYSIA

First Semester Examination
2009/2010 Academic Session

November 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 THIRTEEN pages of printed material before you begin the examination.

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

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

Arahan: Jawab **semua sepuluh** [10] 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. (a) In a survey, 20 people were asked how many magazines they had purchased during the previous year and the results are shown below.

6	15	13	36	25	18	12	18	5	30
24	7	10	22	33	24	19	14	12	9

Find the 5-number summary and construct a boxplot for the data set. What can you say about the distribution of the number of magazine purchased?

[15 marks]

2. (a) A concert pianist plans to perform five classical music pieces during a recital. He has a choice of seven pieces by Beethoven, five pieces by Mozart and four pieces by Tchaikovsky. If he decides to perform two pieces each by Beethoven and Mozart, how many different possible combinations can he perform?
- (b) A study of consumer smoking habit includes 161 people in the 21-30 age category (50 of whom smoke), 144 people in the 31-40 age category (35 of whom smoke) and 80 people in the 41-50 age category (28 of whom smoke). If one person is randomly selected from this sample, find the probability of getting someone who is age 31-40 or smokes.

[20 marks]

3. A large shop owner purchases alarm clocks from two companies. From companies A, 350 alarm clocks are purchased and 2% are found to be defective. From company B, 1050 alarm clocks are purchased and 4% are found to be defective. Given that a randomly selected alarm clock is found to be defective, find the probability that it was purchased from company B.

[15 marks]

1. (a) Dalam suatu kaji selidik, 20 orang telah ditanya berapa banyak majalah telah mereka beli dalam tempoh satu tahun yang lalu dan keputusan ditunjukkan di bawah.

6	15	13	36	25	18	12	18	5	30
24	7	10	22	33	24	19	14	12	9

Cari ringkasan 5-nombor dan bina plot kotak bagi data set tersebut. Apa yang boleh anda katakan tentang taburan bagi nombor majalah yang telah dibeli?

[15 markah]

2. (a) Seorang pemain piano merancang untuk mempersembahkan lima buah muzik klasik semasa suatu konsert resital. Dia mempunyai tujuh pilihan oleh ciptaan Beethoven, lima pilihan oleh Mozart dan empat pilihan oleh Tchaikovsky. Sekiranya dia mengambil keputusan untuk mempersembahkan dua muzik oleh Beethoven dan Mozart, berapa banyak kombinasi yang mungkin boleh dipersembahkan?

- (b) Suatu kajian bagi tabiat merokok pengguna melibatkan 161 orang dalam kategori umur 21-30 tahun (50 daripada mereka merokok), 144 orang dalam kategori umur 31-40 tahun (35 daripada mereka merokok) dan 80 orang dalam kategori 41-50 tahun (28 daripada mereka merokok). Sekiranya seorang dipilih secara rawak daripada sampel ini, cari kebarangkalian untuk mendapat seorang dalam kategori umur 31-40 tahun atau merokok.

[20 markah]

3. Seorang pemilik kedai besar membeli jam penggera dari dua syarikat. Dari syarikat A, 350 jam penggera telah dibeli and 2% didapati rosak. Dari syarikat B, 1050 jam penggera telah dibeli dan 4% didapati rosak. Diberikan bahawa sebuah jam penggera yang dipilih secara rawak telah didapati rosak, cari kebarangkalian bahawa ia telah dibeli dari syarikat B.

[15 markah]

4. (a) (i) In a certain college, 23% of the social science students belong to ethnic minorities. If 10 students are selected at random from the social science students, what is the probability that not more than four belong to an ethnic minority?
- (ii) At an archaeological excavation site, workers find an average of 3 prehistoric artifacts per 10 litres of sediment. If a student is sent to dig up 50 litres of sediment, find the probability that he or she will find at least five artifacts.
- (b) The table below describes the probability distribution for the number of rotten star fruit in a packet of four, two days after it was put on the shelf. Find the mean and standard deviation for the distribution shown.

x	0	1	2	3	4
$P(x)$	0.2646	0.4116	0.2401	0.0756	0.0081

[20 marks]

5. (a) A machine is adjusted to fill 180ml bottles of fresh kiwi juice with mean 160ml and standard deviation 8ml. It was discovered that 7% of the bottles have volume of more than 170ml. Assume the amount of fill follows a normal distribution.
- (i) What is the probability that a randomly selected kiwi juice bottle contains more than 164ml?
- (ii) If 50 bottles of kiwi juice are randomly selected, find the probability that their mean volume is less than 155ml.
- (b) The Registrar at a local university would like to estimate, within three minutes, the mean travel time to work for all lecturers who live within twenty kilometers of the university. He knows that the standard deviation for all travel times is 12.5 minutes. If he wants to achieve a 95% confidence level, how many lecturers should he sample?

[20 marks]

4. (a) (i) Dalam suatu kolej, 23% daripada pelajar sains kemasyarakatan tergolong dalam puak etnik minoriti. Sekiranya 10 pelajar dipilih secara rawak daripada para pelajar sains kemasyarakatan, apakah kebarangkalian bahawa tidak lebih daripada empat adalah tergolong dalam suatu puak etnik minoriti?
- (ii) Di suatu tapak penggalian arkeologi, pekerja mendapat secara purata 3 artifak pra-sejarah bagi setiap 10 liter mendakan. Sekiranya seorang pelajar dihantar untuk menggali 50 liter mendakan, cari kebarangkalian bahawa beliau akan mendapat sekurang-kurangnya lima artifak.
- (b) Jadual di bawah menerangkan taburan kebarangkalian bagi bilangan biji belimbing yang rosak dalam satu paket mengandungi empat biji belimbing, selepas dua hari ia diletakkan di atas rak. Cari min dan sisihan piawai bagi taburan yang ditunjuk.

x	0	1	2	3	4
$P(x)$	0.2646	0.4116	0.2401	0.0756	0.0081

[20 markah]

5. (a) Sebuah mesin diselaraskan untuk mengisi botol-botol 180ml bagi jus kiwi segar dengan purata 160ml dan sisihan piawai 8ml. Adalah telah dijumpai bahawa 7% daripada botol-botol tersebut mempunyai isipadu lebih daripada 170ml. Andaikan amaun yang diisi mempunyai taburan normal.
- (i) Apakah kebarangkalian bahawa satu botol jus kiwi yang dipilih secara rawak mempunyai lebih daripada 164ml?
- (ii) Sekiranya 50 botol jus kiwi dipilih secara rawak, cari kebarangkalian bahawa purata isipadu adalah kurang daripada 155ml.
- (b) Pendaftar di sebuah universiti tempatan ingin menganggar, dalam lingkungan tiga minit, purata waktu perjalanan ke tempat kerja bagi semua pensyarah yang tinggal dalam lingkungan dua puluh kilometer dari universiti. Dia mengetahui bahawa sisihan piawai untuk semua waktu perjalanan adalah 12.5 minit. Sekiranya dia ingin mencapai 95% aras keyakinan, berapa ramai pensyarah yang patut diambil sebagai sampel?

[20 markah]

6. (a) Of 350 randomly interviewed medical students, only 30 said that they are willing to work in rural community. Find a 99% confidence interval for the true proportion of all medical students who are willing to work in a rural community.
- (b) The following data represent the number of minutes the Air Asia flights from Kuala Lumpur arriving later than originally scheduled at Penang's Bayan Lepas airport. The data are in order of arrival during a particular 5-weekday period.

6	1	3	9	10	10	2	5	5	6
12	3	7	8	9	4	5	8	11	14

- (i) Find the 95% confidence interval for the variance of minutes of late arrival.
- (ii) If 4% of the flights arrive more than 10 minutes late, what is the probability that of 200 flights, more than 10 flights arrive more than 10 minutes late.

[20 marks]

7. An economist wants to see if the mean income during the first year of work for university graduates who took statistics course exceeds that of those who did not. The mean first-year income of students who have not taken statistics is assumed to be RM26500 and that the standard deviation of the distribution of first-year incomes is assumed to be RM5250. The economist surveys 25 recent graduates who has taken a statistics course during university years and finds that the average first year income is RM29800.

Calculate the observed p-value. Using 1% significance level, is there enough evidence to conclude that the mean income for the first year of work for university graduates who took statistics course exceed that of those who did not.

[15 marks]

6. (a) *Daripada 350 pelajar perubatan yang ditemuduga secara rawak, hanya 30 orang menyatakan kesanggupan mereka untuk bekerja di kawasan komuniti luar bandar. Cari 99% selang keyakinan bagi peratusan sebenar pelajar perubatan yang sanggup bekerja di kawasan luar bandar.*
- (b) *Data berikut mewakili bilangan minit penerbangan Air Asia dari Kuala Lumpur yang tiba lewat daripada yang dijadualkan di lapangan terbang Bayan Lepas Pulau Pinang. Data adalah mengikut turutan ketibaan bagi tempoh 5-hari bekerja yang tertentu.*

6	1	3	9	10	10	2	5	5	6
12	3	7	8	9	4	5	8	11	14

- (i) *Cari 95% selang keyakinan bagi varians bilangan minit ketibaan lewat.*
- (ii) *Sekiranya 4% daripada bilangan penerbangan tiba lewat lebih daripada 10 minit, apakah kebarangkalian bahawa daripada 200 penerbangan, lebih daripada 10 penerbangan tiba lewat lebih daripada 10 minit.*

[20 markah]

7. *Seorang pakar ekonomi ingin melihat sekiranya purata pendapatan semasa tahun pertama bekerja bagi graduan universiti yang mengambil kursus statistik melebihi pendapatan mereka yang tidak mengambil kursus tersebut. Purata pendapatan tahun pertama bagi para pelajar yang tidak mengambil statistik diandaikan sebanyak RM26500 dan bahawa sisihan piawai bagi taburan untuk pendapatan tahun pertama diandaikan sebanyak RM5250. Pakar ekonomi tersebut menjalankan kaji selidik terhadap 25 graduan yang telah mengambil kursus statistik semasa di universiti dan mendapati bahawa purata pendapatan tahun pertama adalah RM29800.*

Hitung nilai-p yang terhasil. Menggunakan aras kesignifikanan 1%, adakah terdapat bukti yang cukup untuk menyimpulkan bahawa purata pendapatan semasa tahun pertama bekerja bagi graduan universiti yang mengambil kursus statistik melebihi nilai bagi mereka yang tidak mengambil kursus tersebut.

[15 markah]

8. The management of Malaysian Philharmonic Orchestra believes that although music by Mozart appeals to older audiences as compared to music by Shostakovich, music by Mozart attracts wider age of audiences. To test the two claims, 30 audiences were randomly selected (based on randomly chosen seating number) at the end of two concerts, one concert features music by Mozart and another concert features music by Shostakovich. The age of the selected audiences are recorded below.

	Mozart				Shostakovich			
	48	42	50	49	42	41	35	33
	44	59	32	43	38	48	40	45
	52	45	47	53	37	38	40	53
		25	51	58		39	41	44

(given that $\sum_{\text{mozart}} X_i^2 = 33596$ and $\sum_{\text{shostakovich}} X_i^2 = 25492$)

- (a) (i) State the null and alternative hypotheses for testing the related 2-sample variances.
(ii) State the null and alternative hypotheses for testing the related 2-sample means.
- (b) (i) At 5% significance level, is there enough evidence to support the claim that music by Mozart attracts wider age of audiences as compared to music by Shostakovich? Use rejection region approach.
(ii) At 5% significance level, is there enough evidence to support the claim that music by Mozart appeals to older audiences as compared to music by Shostakovich? Use p-value approach.

[30 marks]

9. (a) According to Mars Inc, 30% of the M&M's candies are brown and the rest are coloured (yellow, red, blue, orange and green) with equal ratios. A randomly selected 500gram packet of M&M is found to contain the following number of brown and coloured candies.

Brown	Yellow	Red	Blue	Orange	Green
95	45	40	33	36	51

Conduct a test at 10% significance level to determine whether the observed counts contradict the manufacturer's claim.

[15 marks]

8. Pihak pengurusan Orkestra Pilharmonik Malaysia percaya bahawa walaupun muzik oleh Mozart mempunyai daya tarikan pada pendengar yang lebih tua berbanding muzik oleh Shostakovich, muzik oleh Mozart menambat hati pendengar daripada jurang umur yang lebih besar. Untuk menguji dua dakwaan tersebut, 30 pendengar telah dipilih secara rawak (berdasarkan nombor tempat duduk yang dipilih secara rawak) setelah dua konsert berakhir, satu konsert menampilkan muzik oleh Mozart dan satu konsert lagi menampilkan muzik oleh Shostakovich. Umur para pendengar yang terpilih dicatatkan seperti di bawah.

	Mozart				Shostakovich		
48	42	50	49	42	41	35	33
44	59	32	43	38	48	40	45
52	45	47	53	37	38	40	53
25	51	58		39	41	44	
(diberi bahawa $\sum_{\text{mozart}} X_i^2 = 33596$)				dan $\sum_{\text{shostakovich}} X_i^2 = 25492$)			

- (a) (i) Nyatakan hipotesis nol dan hipotesis alternatif bagi menguji varians 2-sampel yang berkaitan.
- (ii) Nyatakan hipotesis nol dan hipotesis alternatif bagi menguji purata 2-sampel yang berkaitan.
- (b) (i) Pada aras kesignifikanan 5%, adakah terdapat bukti yang cukup untuk menyokong dakwaan bahawa musik oleh Mozart menambat hati pendengar daripada jurang umur yang lebih besar berbanding musik oleh Shostakovich? Gunakan kaedah rantau penolakan.
- (ii) Pada aras kesignifikanan 5%, adakah terdapat bukti yang cukup untuk menyokong dakwaan bahawa muzik oleh Mozart mempunyai daya tarikan pada pendengar yang lebih tua berbanding muzik oleh Shostakovich? Gunakan kaedah nilai-p.

[30 markah]

9. (a) Mengikut Mars Inc, 30% daripada gula-gula M&M's adalah berwarna coklat dan selebihnya bercampur warna (kuning, merah, biru, oren dan hijau) dengan nisbah yang sama. Satu paket 500gram M&M yang dipilih secara rawak dijumpai mengandungi bilangan gula-gula berwarna coklat dan lain-lain warna seperti berikut.

Coklat	Kuning	Merah	Biru	Oren	Hijau
95	45	40	33	36	51

Lakukan ujian pada aras kesignifikanan 10% untuk menentukan samada bilangan yang diperhati bercanggah dengan dakwaan oleh pengeluar tersebut.

[15 markah]

10. (a) Ten secretaries were selected at random from among the secretaries of a large university. The typing speed (number of words per 2-minute) was recorded for each secretary using a certain computer keyboard. The following results were obtained.

Secretary	No. of words
Khatijah	74
Zaiton	80
Dahlan	68
Norhaliza	74
Ramlah	86
Sarah	75
Anuar	78
Noraniza	69
Daniel	76
Jaclyn	77

Assume that the typing speeds are not normally distributed. Perform the Wilcoxon signed-rank test to determine if these data provide enough evidence at the 5% significance level to infer that the typing speed is less than 80 words per 2-minits.

- (b) A Northern Region manager of McDonalds fast food restaurant strongly feels that the time required for drive-through services at Penang Island's Greenlane branch is shorter than that at Penang's Seberang Jaya branch. The manager randomly observed eight cars using the drive-through services at both branches and records the time for completing the services. The following results were obtained.

Greenlane	115	126	49	56	98	76	78	95
Seberang Jaya	120	67	89	97	124	68	72	144

Using a suitable non-parametric test and at 5% significance level, is there sufficient evidence to support the claim made by the regional manager?

[30 marks]

10. (a) Sepuluh kerani telah dipilih secara rawak daripada para kerani di suatu universiti. Kepantasan menaip (bilangan perkataan bagi 2-minit) menggunakan papan kekunci komputer yang tertentu telah dicatat untuk setiap kerani. Keputusan berikut telah diperolehi.

Kerani	Bilangan perkataan
Khatijah	74
Zaiton	80
Dahlan	68
Norhaliza	74
Ramlah	86
Sarah	75
Anuar	78
Noraniza	69
Daniel	76
Jaclyn	77

Andaikan bahawa kelajuan menaip adalah tidak tertabur normal. Lakukan ujian tanda-berpangkat Wilcoxon untuk menentukan sekiranya data ini memberikan bukti yang cukup pada aras kesignifikanan 5% untuk menyimpulkan bahawa kepantasan menaip adalah kurang daripada 80 perkataan bagi 2-minit.

- (b) Seorang pengarah Zon Utara restoran makanan segera McDonalds amat merasakan bahawa masa yang perlu untuk perkhidmatan pandu-lalu di cawangan Greenlane Pulau Pinang adalah lebih pendek daripada masa yang diperlukan di cawangan Seberang Jaya. Pengurus tersebut memilih secara rawak lapan kereta yang menggunakan perkhidmatan pandu-lalu di kedua-dua cawangan dan mencatatkan masa untuk menyempurnakan perkhidmatan tersebut. Keputusan berikut telah diperolehi.

Greenlane	115	126	49	56	98	76	78	95
Seberang Jaya	120	67	89	97	124	68	72	144

Menggunakan ujian tidak-berparameter yang sesuai dan pada aras kesignifikanan 5%, adakah terdapat bukti yang cukup untuk menyokong dakwaan yang dibuat oleh pengurus zon tersebut?

[30 markah]

APPENDIX

Confidence Interval

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

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

n	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