UNIVERSITI SAINS MALAYSIA

First Semester Examination
2018/2019 Academic Session

December 2018/January 2019

MAA161 - Statistics For Science Students
(Statistik Untuk Pelajar Sains)

Duration : 3 hours
[Masa : 3 jam]

Please check that this examination paper consists of NINE (9) pages of printed material before you begin the examination.

[Sila pastikan bahawa kertas peperiksaan ini mengandungi SEMBILAN (9) muka surat yang bercetak sebelum anda memulakan peperiksaan ini.]

Instructions: Answer **NINE** (9) questions.

*[**Arahan:** Jawab **SEMBILAN** (9) 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.]

Questions 1

The frequency distribution shows the blood glucose levels (in milligrams per deciliter) for 50 patients at a medical centre.

Blood Glucose Level	Frequency, f_i
60 - 64	2
65 - 69	1
70 - 74	5
75 - 79	12
80 - 84	18
85 - 89	6
90 - 94	5
95 - 99	1

Given that $\sum_{i=1}^{k} x_i f_i = 4030$ and $\sum_{i=1}^{k} x_i^2 f_i = 327,420$

(a) Calculate the mean of the blood glucose level and its standard deviation.

(b) The patients with glucose level exceeding 87 mg/dl will be admitted in the hospital. What is the percentage of the patients who will be admitted in the hospital?

(c) Use Chebyshev's Theorem to obtain an interval of the blood glucose levels of at least 75% of the patients at a medical facility.

[10 marks]

Soalan 1

Taburan frekuensi menunjukkan tahap glukosa darah (dalam milligram per desiliter) bagi 50 pesakit di sebuah pusat perubatan.

Tahap Glukosa Darah	*Frekuensi,* f_i
60 - 64	*2*
65 - 69	*1*
70 - 74	*5*
75 - 79	*12*
80 - 84	*18*
85 - 89	*6*
90 - 94	*5*
95 - 99	*1*

Diberikan bahawa $\sum_{i=1}^{k} x_i f_i = 4030$ *dan* $\sum_{i=1}^{k} x_i^2 f_i = 327,420$

(a) Kira min bagi tahap glukosa darah dan sisihan piawainya.

(b) Pesakit dengan tahap glukosa melebihi 87 mg/dl akan dimasukkan ke hospital. Berapakah peratusan pesakit yang akan dimasukkan ke hospital?

(c) Guna Teorem Chebyshev bagi mendapatkan selang bagi tahap glukosa darah untuk sekurang-kurangnya 75% pesakit di pusat perubatan tersebut.

[10 markah]

Question 2

(a) The probability that Henry will like a new movie is 0.70 and the probability that Jean, his girlfriend, will like it is 0.60. If the probability is 0.28 that he will like it and she will dislike it, what is the probability that he will like it given that she is not going to like it?

(b) If A and B are two nonempty and independent events, show that $\overline{A}$ and $\overline{B}$ are also independent.

[10 marks]

Soalan 2

(a) Kebarangkalian Henry menyukai wayang gambar baru adalah 0.70 dan kebarangkalian teman wanitanya, Jean menyukai wayang gambar tersebut adalah 0.60. Sekiranya kebarangkalian Henry suka dan teman wanitanya tidak suka wayang gambar tersebut adalah 0.28, apakah kebarangkalian bahawa Henry suka diberikan Jean tidak menyukai wayang gambar itu?

(b) Sekiranya A dan B adalah dua peristiwa bukan sifar dan tidak bersandar, tunjukkan $\overline{A}$ dan $\overline{B}$ juga adalah tidak bersandar.

[10 markah]

Question 3

The probability density function of a continuous variable *X* is given by

$$f(x)=\begin{cases}\dfrac{x+1}{6}, & 1<x<3\\ 0, & \text{otherwise}\end{cases}$$

(a) Find E(*X*) and Var(*X*).

(b) The median for a continuous probability distribution is a value *m* that divides the distribution into two equal areas, i.e.

$$\int_{-\infty}^{m} f(x)\,dx=\int_{m}^{\infty} f(x)\,dx=0.5$$

Find the median of *f(x)*.

[15 marks]

Soalan 3

Fungsi ketumpatan kebarangkalian bagi pemboleh ubah selanjar X dinyatakan sebagai

$$f(x)=\begin{cases}\dfrac{x+1}{6}, & 1<x<3\\ 0, & \textit{selainnya}\end{cases}$$

(a) Cari E(X) dan Var(X).

(b) Median bagi taburan kebarangkalian selanjar ialah nilai m yang membahagikan taburan kepada dua kawasan yang sama iaitu

$$\int_{-\infty}^{m} f(x)\,dx=\int_{m}^{\infty} f(x)\,dx=0.5$$

Dapatkan median bagi f(x).

[15 markah]

Question 4

In a reality program show, it has been studied that Mawi's winning votes are according to a Poisson distribution with 5 votes per minute. Meanwhile, Siti's winning votes are following the Poisson distribution with 4.5 votes per minute.

(a) Calculate the probability that Mawi wins 5 votes or more within three minutes.

(b) Within 1 hour of the program, calculate the probability that Siti wins more than 300 votes. [Hint: use normal approximation to solve the question]

(c) After two hours running of the reality program, using the normal approximation and assuming that Mawi's votes and Siti's votes are mutually independent, find the probability that total Siti's votes is higher than Mawi's.

[10 marks]

Soalan 4

Dalam suatu program realiti, kajian menunjukkan bahawa undian kemenangan Mawi adalah berdasarkan taburan Poisson dengan 5 undian per minit. Manakala Siti meraih undian berdasarkan taburan Poisson dengan 4.5 undian per minit.

(a) Kira kebarangkalian bahawa Mawi meraih 5 atau lebih undian dalam tempoh tiga minit.

(b) Dalam tempoh 1 jam, kira kebarangkalian Siti meraih undian lebih daripada 300 undian. [Petunjuk: gunakan penghampiran normal untuk menyelesaikan soalan ini].

(c) *Selepas dua jam program tersebut berlangsung, dengan menggunakan penghampiran normal serta andaian bahawa undian Mawi dan Siti adalah tidak bersandar, cari kebarangkalian keseluruhan undian Siti adalah lebih tinggi daripada keseluruhan undian Mawi.*

[10 markah]

Question 5

A random variable *Y* has a normal distribution with mean μ and variance σ^2. Given that 10% of the values of *Y* exceed 17.24 and that 25% of the values of *Y* are less than 14.37. Find the values of μ and σ^2.

[10 marks]

Soalan 5

Pemboleh ubah rawak Y mempunyai taburan normal dengan min μ dan varians σ^2. Diberikan bahawa 10% daripada nilai Y melebihi 17.24 dan 25% daripada nilai Y kurang daripada 14.37. Dapatkan nilai μ dan σ^2.

[10 markah]

Question 6

A university is subdivided into faculties of Arts, Engineering, Humanities, Law and Science. The numbers of students in these faculties are 1300, 800, 1100, 500 and 1400 respectively. A questionnaire concerning library usage is sent to all the students. The first 300 replies opened contained 101 from Arts students, 30 from Engineering students and 69, 17 and 83 from students of the other three faculties respectively. Using a 0.1% significance level, test whether the replies appear to be providing an unbiased representation of the students in the university.

[10 marks]

Soalan 6

Sebuah universiti dibahagikan kepada fakulti Sastera, Kejuruteraan, Kemanusiaan, Undang-undang dan Sains. Bilangan pelajar bagi fakulti-fakulti tersebut adalah 1300, 800, 1100, 500 dan 1400 masing-masing. Soal selidik berkaitan penggunaan perpustakaan diedarkan kepada semua pelajar. 300 jawapan pertama yang diterima adalah daripada 101 pelajar Sastera, 30 pelajar Kejuruteraan, dan masing-masing 69, 17 dan 83 pelajar dari fakulti-fakulti lain. Menggunakan aras keertian 0.1%, uji sama ada jawapan yang diberikan adalah perwakilan saksama dengan bidang pelajar di universiti.

[10 markah]

Question 7

Ten students sat for a Mathematics examination which comprises two different papers. Their scores out of 100 are given in the table below:

Student	1	2	3	4	5	6	7	8	9	10
Paper 1	31	43	50	52	62	40	44	60	55	27
Paper 2	35	58	43	68	92	59	25	55	50	18

(a) Find a 95% confidence interval for the mean of the differences between the scores for Paper 1 and Paper 2.

(b) At the 5% significance level, test whether the mean scores of Paper 1 and Paper 2 are different.

[15 marks]

Soalan 7

Sepuluh pelajar menduduki peperiksaan Matematik yang terdiri daripada dua kertas berbeza. Markah mereka daripada 100 markah penuh diberikan dalam jadual di bawah:

Pelajar	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
Kertas 1	*31*	*43*	*50*	*52*	*62*	*40*	*44*	*60*	*55*	*27*
Kertas 2	*35*	*58*	*43*	*68*	*92*	*59*	*25*	*55*	*50*	*18*

(a) Dapatkan selang keyakinan 95% bagi perbezaan min skor Kertas 1 dan Kertas 2.

(b) Pada aras keertian 5%, uji sama ada min skor Kertas 1 dan Kertas 2 adalah berbeza.

[15 markah]

Question 8

(a) Briefly describe Type I and Type II errors.

(b) In a poll of college athletes in 2003, 76% of the athletes declared that their favourite brand of shoe was Brand X. The manufacturer of Brand X is interested to know whether this is still the case. Therefore, 120 athletes are polled, and 73.3% declared that they prefer Brand X. The manufacturer then performs a hypothesis test at the 10% level of significance. What is the manufacturer's conclusion?

[10 marks]

Soalan 8

(a) Huraikan secara ringkas ralat jenis I dan ralat jenis II.

(b) Dalam undian atlet kolej pada 2003, 76% daripada atlet menyatakan jenama kasut kegemaran mereka adalah Jenama X. Pengeluar Jenama X berminat untuk mengetahui sama ada kegemaran mereka masih sama. Oleh yang demikian, seramai 120 atlet telah diminta mengundi dan 73.3% mengisytiharkan bahawa mereka gemar Jenama X. Pengeluar kemudiannya menjalankan ujian hipotesis pada aras keertian 10%. Apakah kesimpulan pengeluar?

[10 markah]

Question 9

Jenifer has decided to change the girls physical education at Alexander High School to include a volunteer weight reduction and fitness program for the high school seniors who want to reduce their body weight. The 15 girls who signed up for the program weighed in at the onset and again one month later. The data of the 15 girls involved in this program are given in the table below:

Weight Before	Weight After	Weight Before	Weight After	Weight Before	Weight After
146	140	135	136	151	150
159	151	178	170	132	131
139	132	146	141	149	141
175	165	153	150	129	129
152	143	166	168	139	133

Using the 5% level of significance, can we conclude that the new physical education program helped reduced the median body weight of high school senior females?

[10 marks]

Soalan 9

Jenifer membuat keputusan untuk mengubah pendidikan jasmani pelajar perempuan di Sekolah Menengah Alexander bagi memasukkan sukarelawati dalam program kecergasan dan penurunan berat badan dalam kalangan pelajar senior yang ingin menurunkan berat badan mereka. Sebanyak 15 pelajar perempuan menyertai program ini dan berat badan mereka diambil pada masa program bermula dan selepas sebulan kemudian. Data 15 pelajar perempuan yang terlibat dalam program ini diberikan dalam jadual di bawah:

Berat Sebelum	*Berat Selepas*	*Berat Sebelum*	*Berat Selepas*	*Berat Sebelum*	*Berat Selepas*
146	*140*	*135*	*136*	*151*	*150*
159	*151*	*178*	*170*	*132*	*131*
139	*132*	*146*	*141*	*149*	*141*
175	*165*	*153*	*150*	*129*	*129*
152	*143*	*166*	*168*	*139*	*133*

Menggunakan aras keertian 5%, bolehkah disimpulkan bahawa program pendidikan jasmani baru membantu dalam menurunkan median berat badan pelajar senior perempuan sekolah menengah?

[10 markah]

FORMULA

$$\bar{x} = \frac{\sum xf}{\sum f}$$

$$s^2 = \frac{\sum (x^2 f) - \frac{(\sum xf)^2}{\sum f}}{\sum f - 1}$$

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

$$\hat{p} = \frac{X + Y}{n_x + n_y}$$

$$m = b + \frac{c\left(\frac{n}{2} - l\right)}{f_m}$$

Confidence Intervals:

$$\overline{X} \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

$$\overline{X} \pm t_{\alpha/2}\frac{s}{\sqrt{n}}$$

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

$$\frac{(n-1)s^2}{\chi^2_{\alpha/2}} \quad to \quad \frac{(n-1)s^2}{\chi^2_{1-\alpha/2}}$$

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

$$\left(\overline{X} - \overline{Y}\right) \pm z_{\alpha/2}\sqrt{\frac{s_x^2}{n_x} - \frac{s_y^2}{n_y}}$$

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

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

Test Statistics:

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

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

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

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

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

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

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

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

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

$$F = \frac{s_1^2}{s_2^2}$$

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

Nonparametric Statistics:

Wilcoxon Signed-rank: $W = \sum R^+$, $W = \sum R^-$

$$Z = \frac{T - \mu_W}{\sigma_W} , \quad \mu_W = \frac{n(n+1)}{4}, \ \sigma_W = \sqrt{\frac{n(n+1)(2n+1)}{24}}$$

Wilcoxon Rank Sum Test: $U = R - \frac{n(n+1)}{2}$

$$Z = \frac{T - \mu_T}{\sigma_T} , \quad \mu_T = \frac{n_1(n_1 + n_2 + 1)}{2}, \ \sigma_T = \sqrt{\frac{n_1 n_2 (n+1)(n_1 + n_2 + 1)}{12}}$$

- ooo000ooo -