
UNIVERSITI SAINS MALAYSIA

Second Semester Examination
2015/2016 Academic Session

June 2016

MGM 561 – Statistical Methods for Research
[Kaedah Statistik untuk Penyelidikan]

Duration : 3 hours
[Masa : 3 jam]

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

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

Instructions: Answer **all six** [6] questions.

Arahan: Jawab **semua enam** [6] 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. Given the data as below:-

| | | | | | | | | | |
|----|---|----|---|----|---|----|---|----|---|
| 11 | 7 | 7 | 7 | 5 | 2 | 11 | 3 | 5 | 5 |
| 5 | 3 | 4 | 4 | 10 | 6 | 3 | 4 | 12 | 9 |
| 9 | 2 | 7 | 4 | 6 | 4 | 5 | 1 | 3 | 5 |
| 5 | 8 | 2 | 6 | 6 | 7 | 4 | 4 | 4 | 4 |
| 6 | 2 | 13 | 6 | 5 | 7 | 6 | 4 | 7 | 5 |

- (a) Draw the stem-and-leaf plot. [3 marks]
- (b) Find the mean, median, first and third quartiles. [7 marks]
- (c) Draw the boxplot and explain the shape of the data. [3 marks]
- (d) Based on these plots, determine which of the estimates, is the best in explaining the measure of location of this data. [2 marks]

1. Diberi data seperti di bawah:-

| | | | | | | | | | |
|----|---|----|---|----|---|----|---|----|---|
| 11 | 7 | 7 | 7 | 5 | 2 | 11 | 3 | 5 | 5 |
| 5 | 3 | 4 | 4 | 10 | 6 | 3 | 4 | 12 | 9 |
| 9 | 2 | 7 | 4 | 6 | 4 | 5 | 1 | 3 | 5 |
| 5 | 8 | 2 | 6 | 6 | 7 | 4 | 4 | 4 | 4 |
| 6 | 2 | 13 | 6 | 5 | 7 | 6 | 4 | 7 | 5 |

- (a) Lukiskan plot batang-dan-daun. [3 markah]
- (b) Cari min, median, kuartil pertama dan ketiga. [7 markah]
- (c) Lukiskan plot-kotak dan terangkan bentuk data tersebut. [3 markah]
- (d) Berdasarkan plot-plot tersebut, tentukan yang manakah penganggar terbagus, dalam menerangkan ukuran lokasi bagi data ini. [2 markah]

Questions number 2 and 3 are based on Table 1

Table 1: Athletes, physical endurance scores

| Girls, x | | | Boys, y | | | |
|------------|-------|-------|-----------|-------|-------|-------|
| -2.65 | 0.60 | -2.19 | 0.16 | -1.04 | 0.82 | 0.96 |
| -1.71 | -3.04 | 0.99 | -1.98 | -0.57 | -0.12 | -0.11 |
| -1.66 | 0.92 | 0.02 | 1.23 | 0.81 | 1.59 | |
| -1.55 | 0.57 | -0.13 | -1.09 | -0.51 | 0.81 | |
| 3.49 | | | 1.03 | 0.76 | 1.58 | |

$$\sum x = -6.34, \sum x^2 = 43.8512, \sum y = 4.33, \sum y^2 = 17.9093$$

2. In a study of the performance of athletes, one of the factors influencing here is physical endurance. A study has been conducted and the factor score of the athletes' physical endurance has been computed and shown in Table 1:
- (a) At 5% significance level, test whether the boy athletes have the variance equals to 1.
[9 marks]
- (b) At 5% significance level, test whether the scores of girls and boys have the common variance.
[8 marks]
- (c) Based on answer in (b), find the 90% confidence interval of the difference of the mean scores between girls and boys.
[5 marks]
- (d) Thus, at 5% significance level, test whether the mean scores of the physical endurance of the boys is greater than the girls.
[3 marks]

Soalan nombor 2 dan 3 adalah berdasarkan Jadual 1

Jadual 1: Skor ketahanan fizikal atlit

| <i>Perempuan, x</i> | | | <i>Lelaki, y</i> | | | |
|---------------------|-------|-------|------------------|-------|-------|-------|
| -2.65 | 0.60 | -2.19 | 0.16 | -1.04 | 0.82 | 0.96 |
| -1.71 | -3.04 | 0.99 | -1.98 | -0.57 | -0.12 | -0.11 |
| -1.66 | 0.92 | 0.02 | 1.23 | 0.81 | 1.59 | |
| -1.55 | 0.57 | -0.13 | -1.09 | -0.51 | 0.81 | |
| 3.49 | | | 1.03 | 0.76 | 1.58 | |

$$\sum x = -6.34, \sum x^2 = 43.8512, \sum y = 4.33, \sum y^2 = 17.9093$$

2. Dalam suatu kajian terhadap pencapaian atlit, satu daripada faktor yang mempengaruhinya adalah ketahanan fizikal. Satu kajian telah dilaksanakan dan skor faktor bagi ketahanan fizikal atlit telah dikira dan ditunjukkan dalam Jadual 1.

(a) Pada aras keertian 5%, uji sama ada atlit lelaki mempunyai varians bersamaan 1.

[9 markah]

(b) Pada aras keertian 5%, uji sama ada skor bagi atlit perempuan dan lelaki mempunyai varians sepunya.

[8 markah]

(c) Berdasarkan jawapan (b), cari selang keyakinan 90% bagi perbezaan skor min antara atlit perempuan dan lelaki.

[5 markah]

(d) Seterusnya, pada aras keertian 5%, uji sama ada min bagi skor ketahanan fizikal bagi atlit lelaki adalah lebih besar daripada perempuan.

[3 markah]

3. Using data in Table 1, with the assumption of common variance between the two groups,
- (a) Estimate the pooled variance of the data from these two groups.
[2 marks]
- (b) Find the mean square of Treatment.
[4 marks]
- (c) Using one way Analysis of Variance in testing whether the mean of the physical endurance score of the boys is equal to the girls,
- (ii) Determine the F-test statistic, and its distribution.
- (iii) At 5% significance level, determine whether the null hypothesis above is being rejected, or not. State the conclusion.
[6 marks]

3. *Dengan menggunakan data yang diberikan di dalam Jadual 1, dengan andaian varians sepunya antara dua kumpulan itu,*
- (a) *Anggar varians terkumpul bagi kedua-dua kumpulan ini.*
[2 markah]
- (b) *Cari min kuasa dua bagi Rawatan.*
[4 markah]
- (c) *Dengan menggunakan analisis varians sehala dalam menguji samada min skor ketahanan fizikal bagi lelaki adalah sama dengan perempuan,*
- (i) *Tentukan ujian statistik-F, serta taburannya.*
- (ii) *Pada aras keertian 5%, tentukan sama ada hipotesis di atas ditolak atau tidak. Nyatakan kesimpulannya.*
[6 markah]

4. A study has been constructed to observe the efficiency of the new method of teaching science subject of a small class. All students of the class are required to take the test before the method of teaching being conducted, and the test after being implemented. The marks are given in Table 2,

Table 2: Marks for Science Subject

| Student | Before | After |
|----------------|---------------|--------------|
| S1 | 85 | 85 |
| S2 | 53 | 42 |
| S3 | 67 | 59 |
| S4 | 78 | 86 |
| S5 | 62 | 51 |
| S6 | 67 | 56 |
| S7 | 89 | 93 |
| S8 | 62 | 68 |
| S9 | 86 | 89 |
| S10 | 75 | 72 |
| S11 | 71 | 72 |
| S12 | 48 | 45 |
| S13 | 69 | 75 |
| S14 | 70 | 82 |
| S15 | 67 | 50 |

- (a) Find the 90% confidence interval of the difference scores between before and after implementation.

[9 marks]

- (b) Thus, at 5% significance level, test whether the implementation efficiently increase the students' marks.

[3 marks]

4. Suatu kajian telah dijalankan untuk menilai keberkesanan suatu kaedah baru pengajaran mata pelajaran sains bagi suatu kelas kecil. Kesemua pelajar di dalam kelas tersebut dikehendaki menduduki ujian sebelum kaedah pengajaran dijalankan, dan ujian selepas ia dilaksanakan. Markah diberikan seperti di Jadual 2,

Jadual 2: Markah bagi mata pelajaran Sains

| <i>Pelajar</i> | <i>Sebelum</i> | <i>Selepas</i> |
|----------------|----------------|----------------|
| <i>S1</i> | 85 | 85 |
| <i>S2</i> | 53 | 42 |
| <i>S3</i> | 67 | 59 |
| <i>S4</i> | 78 | 86 |
| <i>S5</i> | 62 | 51 |
| <i>S6</i> | 67 | 56 |
| <i>S7</i> | 89 | 93 |
| <i>S8</i> | 62 | 68 |
| <i>S9</i> | 86 | 89 |
| <i>S10</i> | 75 | 72 |
| <i>S11</i> | 71 | 72 |
| <i>S12</i> | 48 | 45 |
| <i>S13</i> | 69 | 75 |
| <i>S14</i> | 70 | 82 |
| <i>S15</i> | 67 | 50 |

- (a) Cari 90% selang keyakinan bagi perbezaan markah ujian antara sebelum dan selepas perlaksanaan dibuat.

[9 markah]

- (b) Justeru, pada aras keertian 5%, uji sama ada perlaksanaan tersebut berkesan menaikkan markah pelajar-pelajar tersebut.

[3 markah]

5. Based on the previous national tournament of athletic last year, the athletes of a state were attaining the number of podiums as follows,

| |
|--|
| 75% of the athletes in this state did not attain any podium for the last year tournament |
| 8% of the athletes only attained one podium |
| 7.5% of the athletes attained two podiums |
| 5% of the athletes achieved three podiums |
| 4.5% of the athletes achieved more than three podiums in the last year tournament |

In this year, 50 athletes, representing this state, are entering the national athletic tournament. The number of podiums achieved by the athletes is shown in Table 3,

Table 3: Number of athletes attaining the number of podiums

| Number of Podiums | Number of Athletes |
|-------------------|--------------------|
| 0 | 35 |
| 1 | 5 |
| 2 | 4 |
| 3 | 3 |
| 4 | 2 |
| >4 | 1 |

- (a) Calculate the expected number of athletes attaining the respective number of podiums, based on previous year of achievement.

[3 marks]

- (b) Construct a necessary hypothesis and the test to evaluate whether the achievement of the athletes in this year is indifferent with previous year (use 10% significance level).

[5 marks]

- (c) Based on previous year achievement, how many athletes attain more than one podium?

[2 marks]

5. Berdasarkan satu kejohanan olahraga peringkat kebangsaan pada tahun lepas, atlit-atalit bagi sesebuah negeri telah memperolehi bilangan podium seperti berikut,

| |
|---|
| 75% atlit tidak memperolehi bilangan podium bagi siri kejohanan yang lepas |
| 8% atlit hanya memperolehi satu podium sahaja |
| 7.5% atlit memperolehi dua podium |
| 5% atlit mencapai tiga podium |
| 4.5% atlit memperolehi lebih daripada tiga podium bagi siri kejohanan musim lalu. |

Pada tahun ini, seramai 50 atlit dari negeri tersebut menyertai kejohanan olahraga peringkat kebangsaan. Bilangan podium yang dicapai oleh atlit-atalit tersebut ditunjukkan seperti di Jadual 3,

Jadual 3: bilangan atlit yang memperolehi bilangan podium

| <i>Bilangan Podium</i> | <i>Bilangan Atlit</i> |
|------------------------|-----------------------|
| 0 | 35 |
| 1 | 5 |
| 2 | 4 |
| 3 | 3 |
| 4 | 2 |
| >4 | 1 |

- (a) Kira jangkaan bilangan atlit yang memperolehi bilangan podium masing-masing, berdasarkan pencapaian tahun lalu.

[3 markah]

- (b) Bina hipotesis yang sesuai berserta ujiannya untuk menilai sama ada pencapaian atlit pada tahun ini tiada bezanya dengan tahun lepas. (gunakan aras keertian 10%).

[5 markah]

- (c) Berdasarkan pencapaian tahun lalu, berapa ramai atlit memperolehi lebih daripada satu podium?

[2 markah]

6. A sample of seven households from a small city consisting of the information on their incomes and food expenditures, are collected. The information (in RM '000) is shown in Table 4

Table 4: Income and food expenditure

| Income (RM '000), x | Food Expenditure (RM '000), y |
|--------------------------|------------------------------------|
| 55 | 14 |
| 83 | 24 |
| 38 | 13 |
| 61 | 16 |
| 33 | 9 |
| 49 | 15 |
| 67 | 17 |

$$\sum x = 386, \sum x^2 = 23,058, \sum y = 108, \sum y^2 = 1,792, \sum xy = 6,403$$

- (a) Find the correlation between income and food expenditure.
- [6 marks]
- (b) Test, at 1% significance level, whether the household income and the household food expenditure are correlated.
- [4 marks]
- (c) Let the food expenditure be the respond variable.
- (i) Draw the scatter diagram between income and food expenditure.
 - (ii) Consider $y = a + bx$, where y is the food expenditure, and x is the household income. Find the estimates of a and b .
 - (iii) Find the 95% confidence interval of expected value of b . Then, test whether $b = 0$.
 - (iv) Find the expected value of food expenditure, if given the income to be RM 70,000.

[16 marks]

6. Suatu sampel bagi tujuh isi rumah dari sebuah bandar kecil, yang mengandungi maklumat berkenaan pendapatan mereka dan perbelanjaan makanan, dikumpulkan. Maklumat (dalam RM '000) seperti ditunjukkan dalam Jadual 4.

Jadual 4: Pendapatan dan perbelanjaan makanan

| Pendapatan (RM '000), x | Perbelanjaan makanan (RM '000), y |
|---|---|
| 55 | 14 |
| 83 | 24 |
| 38 | 13 |
| 61 | 16 |
| 33 | 9 |
| 49 | 15 |
| 67 | 17 |

$$\sum x = 386, \sum x^2 = 23,058, \sum y = 108, \sum y^2 = 1,792, \sum xy = 6,403$$

- (a) Cari korelasi antara pendapatan dan perbelanjaan makanan.

[6 markah]

- (b) Uji, pada aras keyakinan 1%, sama ada pendapatan isi rumah dan perbelanjaan makanan isi rumah adalah berkaitan.

[4 markah]

- (c) Biarkan perbelanjaan makanan sebagai pemboleh ubah respon.

- (i) Lukiskan gambarajah serakan antara pendapatan dan perbelanjaan makanan.
- (ii) Biarkan kaitan di atas diberikan sebagai $y = a + bx$, yang mana y adalah perbelanjaan makanan, dan x adalah pendapatan isi rumah. Cari dan anggarkan a dan b .
- (iii) Cari 95% selang keyakinan bagi nilai anggaran b . Kemudian, uji sama ada $b = 0$.
- (iv) Cari nilai jangkaan perbelanjaan makanan, jika diberi pendapatan adalah RM 70,000.

[16 markah]

APPENDIX/LAMPIRAN

$$1. w = \frac{(s_1^2/n_1 + s_2^2/n_2)^2}{\frac{(s_1^2/n_1)^2}{n_1-1} + \frac{(s_2^2/n_2)^2}{n_2-1}}$$

$$2. s_p^2 = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2 + \dots + (n_k-1)s_k^2}{(n_1+n_2+\dots+n_k-k)} = \frac{\sum_j^k (n_j-1)s_j^2}{\sum_j^k n_j - k}$$

$$3. MST = \frac{\sum_{j=1}^k n_j (\bar{y}_j - \bar{y})^2}{k-1}$$

$$4. S_{XX} = \sum (x - \bar{x})^2 = \sum x^2 - \frac{(\sum x)^2}{n}$$

$$5. S_{XY} = \sum (x - \bar{x})(y - \bar{y}) = \sum xy - \frac{(\sum x)(\sum y)}{n}$$

$$6. S_{Y^*X}^2 = \frac{S_{YY} - bS_{XY}}{n-2}$$

$$7. K = \sum_i \frac{(O_i - E_i)^2}{E_i}$$