UNIVERSITI SAINS MALAYSIA

Peperiksaan Kursus Semasa Cuti Panjang
Sidang Akademik 2009/2010

Jun 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 EIGHT pages of printed material before you begin the examination.

*[Sila pastikan bahawa kertas peperiksaan ini mengandungi LAPAN 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. As part of a project for a statistics class, a student collects data on wristwatch accuracy. The following error times (in seconds) are obtained. (Positive values represent watches that are ahead of the correct time, and negative values represent watches that are behind the correct time.)

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 140 | -125 | 105 | -241 | -85 | 41 | 186 | -151 |
| 325 | 80 | 27 | 20 | 20 | 30 | -65 | |

(i) Find the mean, median and mode.
(ii) Find the variance, standard deviation and range.
(iii) Are error times from a population discrete or continuous?
(iv) What is the level of measurement of these error times?

[12 marks]

2. Based on the National Health Survey, the heights of women in Malaysia have a mean of 63.6 inches and a standard deviation of 2.5 inches,

(i) What can you conclude from Chebyshev's Theorem about the percentage of women with height between 58.6 inches and 68.6 inches?
(ii) If the heights of women have a bell-shaped distribution with similar mean and standard deviation, use the empirical rule to approximate the percentage of women with height between 56.1 inches and 71.1 inches?

[7 marks]

3. Pollution of rivers in Malaysia has been a problem for many years. Consider the following events:

A={A river is polluted}
B={A sample of water tested detects pollution}
C={Fishing permitted}

Assume $P\ A = 0.3$ , $P\ B|A = 0.75$, $P\ B|A' = 0.20$, $P\ C|A \cap B = 0.20$, $P\ C|A' \cap B = 0.15$, $P\ C|A \cap B' = 0.80$, and $P\ C|A' \cap B' = 0.90$. Find

(i) $P\ A \cap B \cap C$
(ii) $P\ B' \cap C$
(iii) $P\ C$
(iv) The probability that the river is polluted, given that fishing is permitted and the sample tested did not detect pollution.

[10 marks]

1. *Sebagai sebahagian daripada projek untuk kelas statistik, pelajar mengumpul data tentang ketepatan jam tangan. Berikut adalah ralat waktu (dalam saat) yang diperoleh. (Nilai-nilai positif merujuk kepada jam tangan yang melepasi waktu yang tepat dan nilai-nilai negatif merujuk kepada jam tangan yang lambat daripada waktu yang tepat)*

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| *140* | *-125* | *105* | *-241* | *-85* | *41* | *186* | *-151* |
| *325* | *80* | *27* | *20* | *20* | *30* | *-65* | |

*(i) Cari min, median dan mod.*
*(ii) Cari varians, sisihan piawai dan julat.*
*(iii) Adakah ralat waktu dari suatu populasi diskrit atau selanjar?*
*(iv) Apakah peringkat pengukuran ralat waktu ini?*

*[12 markah]*

2. *Berdasarkan suatu Soalselidik Kesihatan Nasional, ketinggian wanita di Malaysia mempunyai min 63.6 inci dan sisihan piawai 2.5 inci,*

*(i) Apa yang boleh anda simpulkan dari Teorem Chebyshev tentang peratusan wanita antara 58.6 inci dan 68.6 inci?*
*(ii) Jika ketinggian wanita mempunyai taburan berbentuk loceng dengan min dan sisihan piawai yang sama, gunakan peraturan empirik untuk menganggar peratusan ketinggian wanita antara 56.1 inci dan 71.1 inci?*

*[7 markah]*

3. *Pencemaran sungai di Malaysia telah menjadi masalah bertahun-tahun lamanya. Pertimbangkan peristiwa berikut:*

*A={Sungai tercemar}*
*B={ Suatu sampel air diuji dan didapati tercemar}*
*C={Perikanan dibenarkan}*

*Anggapkan* $P(A)=0.3$, $P(B|A)=0.75$, $P(B|A')=0.20$, $P(C|A\cap B)=0.20$, $P(C|A'\cap B)=0.15$, $P(C|A\cap B')=0.80$, *and* $P(C|A'\cap B')=0.90$. *Cari*

*(i)* $P(A\cap B\cap C)$
*(ii)* $P(B'\cap C)$
*(iii)* $P(C)$
*(iv) Kebarangkalian bahawa sungai tercemar, diberikan bahawa perikanan dibenarkan dan sampel yang diuji tidak mengesan pencemaran.*

*[10 markah]*

4. It is estimated that 20% of all schoolchildren in Malaysia are left-handed. Let us assume that this estimate is reasonably accurate. Consider a random sample of 13 schoolchildren taken from the population.

   (i) Compute the probability that fewer than 3 children in the sample will be left-handed.
   (ii) Compute the probability that between 6 and 9 children (inclusive) will be left-handed.
   (iii) Compute the probability that more than 5 children will be right-handed.

   [6 marks]

5. The number of messages sent to a computer bulletin board is a Poisson random variable with a mean of 5 messages per hour. What is the probability that

   (i) 5 messages are received in 1 hour?
   (ii) 10 messages are received in 1.5 hours?
   (iii) less than two messages are received in one-half hour?

   [8 marks]

6. The monthly sick-leave time of employees in a firm is normally distributed with a mean of 100 hours and a standard deviation of 20 hours.

   (i) What is the probability that the sick-leave time for next month will be between 50 and 80 hours?
   (ii) How much time should be allocated for sick leave if the allocated amount should be exceeded with a probability of only 10%?

   [9 marks]

7. (a) Suppose you want to use a paired difference experiment to estimate the difference between two population means correct to within 1.8 with a 95% confidence interval. If prior information suggests that the population variance of the differences is approximately equal to 28, how many pairs should be selected?

   (b) A 90% confidence interval for $p$ is (0.26, 0.54). How large was the sample used to construct this interval?

   [12 marks]

4. *Dianggarkan bahawa 20% daripada kanak-kanak sekolah di Malaysia adalah kidal. Anggapkan bahawa anggaran ini adalah agak tepat. Pertimbangkan sampel rawak 13 kanak-kanak sekolah yang diambil dari populasi ini.*

   *(i) Hitung kebarangkalian bahawa kurang daripada 3 kanak-kanak dalam sampel adalah kidal.*
   *(ii) Hitung kebarangkalian bahawa antara 6 dan 9 kanak-kanak (terangkum) adalah kidal.*
   *(iii) Hitung kebarangkalian bahawa lebih daripada 5 kanak-kanak adalah bukan kidal.*

   *[6 markah]*

5. *Bilangan pesanan yang dihantar ke sebuah komputer lembaga buletin adalah suatu pembolehubah rawak Poisson dengan min 5 pesanan setiap jam. Apakah kebarangkalian bahawa*

   *(i) 5 pesanan diterima dalam 1 jam?*
   *(ii) 10 pesanan diterima dalam 1.5 jam?*
   *(iii) kurang daripada dua pesanan diterima dalam satu setengah jam?*

   *[8 markah]*

6. *Waktu cuti sakit pekerja di sebuah syarikat dalam satu bulan biasanya bertaburan normal dengan min 100 jam dan sisihan piawai 20 jam.*

   *(i) Apakah kebarangkalian bahawa waktu cuti sakit untuk bulan seterusnya adalah antara 50 dan 80 jam?*
   *(ii) Berapa banyak waktu yang harus diperuntukkan untuk cuti sakit jika jumlah yang diperuntukkan harus dilampaui dengan kebarangkalian hanya 10%?*

   *[9 markah]*

7. *(a) Misalkan anda ingin menggunakan ujikaji perbezaan pasangan untuk menganggarkan perbezaan antara dua min penduduk betul dalam linkungan 1.8 dengan selang keyakinan 95%. Jika maklumat terdahulu menunjukkan bahawa varians populasi perbezaan adalah lebih kurang sama dengan 28, berapa banyak pasangan harus dipilih?*

   *(b) Suatu selang keyakinan 90% untuk p ialah (0.26, 0.54). Berapa besar sampel yang digunakan untuk membina selang ini?*

   *[12 markah]*

8. A study was conducted to investigate the effectiveness of hypnotism in reducing pain. The sensory measurements(in centimeters) on a pain scale, before and after are as follow:

| Subject | A | B | C | D | E | F | G | H |
|---|---|---|---|---|---|---|---|---|
| Before | 6.6 | 6.5 | 9.0 | 10.3 | 11.3 | 8.1 | 6.3 | 11.6 |
| After | 6.8 | 2.4 | 7.4 | 8.5 | 8.1 | 6.1 | 3.4 | 2.0 |

(i) Construct a 95% confidence interval for the mean differences.
(ii) Use a 0.05 significance level to test the claim that the sensory measurements are lower after hypnotism.
(iii) Does hypnotism appear to be effective in reducing pain?

[14 marks]

9. The following data shows the number of road accidents in Georgetown for the period of 100 weeks.

| Number of accidents per week | 0 | 1 | 2 | 3 | 4 | 5 | $\geq 6$ | Total |
|---|---|---|---|---|---|---|---|---|
| Frequency | 7 | 17 | 21 | 24 | 19 | 12 | 0 | 100 |

(i) Estimate the value for $\mu$, the mean number of accidents per week for the period of 100 weeks.
(ii) Test at $\alpha = 0.05$ whether the Poisson model is suitable for the random variable *X* which represents the number of accidents per week.

[12 marks]

10. A random sample of fifty students was asked what salary the university should be prepared to pay to attract the right individual to coach the football team. An independent random sample of fifty faculty members was also asked the same question. The 100 salary figures were then pooled and ranked in order (with rank 1 assigned to the lowest salary). The sum of ranks for faculty members and students were 2024 and 2204 respectively. Test at $\alpha = 0.05$ the claim that students would propose a higher salary to attract a football coach.

[10 marks]

8. *Suatu kajian telah dilakukan untuk mengkaji keberkesanan hipnotisme dalam mengurangkan rasa sakit. Nilai pengukuran deria (dalam sentimeter) pada skala rasa sakit sebelum dan selepas adalah seperti berikut:*

| *Subjek* | *A* | *B* | *C* | *D* | *E* | *F* | *G* | *H* |
|---|---|---|---|---|---|---|---|---|
| *Sebelum* | *6.6* | *6.5* | *9.0* | *10.3* | *11.3* | *8.1* | *6.3* | *11.6* |
| *Selepas* | *6.8* | *2.4* | *7.4* | *8.5* | *8.1* | *6.1* | *3.4* | *2.0* |

*(i) Bina selang keyakinan 95% untuk perbezaan min.*
*(ii) Gunakan aras keertian 0.05 untuk menguji dakwaan bahawa pengukuran deria adalah lebih rendah selepas hipnotisme.*
*(iii) Apakah hipnotisme lebih berkesan dalam mengurangkan rasa sakit?*

*[14 markah]*

9. *Data berikut menunjukkan bilangan kemalangan jalan raya di Georgetown untuk tempoh 100 minggu.*

| *Bil. Kemalangan seminggu* | *0* | *1* | *2* | *3* | *4* | *5* | *≥6* | *Jumlah* |
|---|---|---|---|---|---|---|---|---|
| *Kekerapan* | *7* | *17* | *21* | *24* | *19* | *12* | *0* | *100* |

*(i) Anggarkan nilai $\mu$, min bilangan kemalangan per minggu untuk tempoh 100 minggu tersebut.*
*(ii) Uji pada $\alpha = 0.05$ samada model Poisson adalah sesuai bagi pemboleh ubah rawak X yang mewakili bilangan kemalangan per minggu*

*[12 markah]*

10. *Suatu sampel rawak lima puluh orang pelajar ditanya berapa gaji universiti harus bersedia untuk membayar bagi menarik individu yang sesuai menjadi jurulatih pasukan bola sepak. Suatu sampel rawak tak bersandar lima puluh ahli fakulti juga diajukan pertanyaan yang sama. Seratus nilai-nilai gaji kemudian disatukan dan dipangkatkan dalam turutan (dengan pangkat 1 merujuk gaji terendah). Jumlah pangkat untuk ahli fakulti dan pelajar adalah 2024 dan 2204 masing-masing. Uji pada $\alpha = 0.05$ dakwaan bahawa pelajar akan mencadangkan gaji yang lebih tinggi untuk menarik jurulatih bola sepak.*

*[10 markah]*

## APPENDIX

$$Q_r = b' + \frac{c\left(\frac{rn}{4} - I'\right)}{f_{Q_r}}$$

$$\bar{x} \pm \frac{Z_{\alpha/2}\sigma}{\sqrt{n}}$$

$$\bar{x} \pm \frac{t_{n-1,\alpha/2}, s}{\sqrt{n}}$$

$$\bar{X} - \bar{Y} \pm Z_{\alpha/2}\sqrt{\frac{\sigma_x^2}{n_x} + \frac{\sigma_y^2}{n_y}}$$

$$\bar{X} - \bar{Y} \pm t_{\alpha/2,\ n_x+n_y-2}\, S_p \sqrt{\frac{1}{n_x} + \frac{1}{n_y}} \qquad S_p^2 = \frac{n_x - 1\ s_x^2 + n_y - 1\ s_y^2}{n_x + n_y - 2}$$

$$\hat{p}_x - \hat{p}_y \pm Z_{\alpha/2}\sqrt{\frac{\hat{p}_x\ 1-\hat{p}_x}{n_x} + \frac{\hat{p}_y\ 1-\hat{p}_y}{n_y}}$$

$$\left(\frac{s_x^2}{s_y^2}\frac{1}{f_{\alpha/2}\ \nu_x,\nu_y}, \frac{s_x^2}{s_y^2} f_{\alpha/2}\ \nu_y,\nu_x\right)$$

$$\left(\frac{n-1\ s^2}{\chi^2_{n-1,\alpha/2}}, \frac{n-1\ s^2}{\chi^2_{n-1,1-\alpha/2}}\right)$$

$$Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$$

$$T = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

$$Z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}}$$

$$\chi^2 = \frac{n-1\ s^2}{\sigma_0^2}$$

$$Z = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{s_x^2}{n_x} + \frac{s_y^2}{n_y}}}$$

$$T = \frac{\bar{X} - \bar{Y}}{S_p\sqrt{\frac{1}{n_x} + \frac{1}{n_y}}}$$

$$T = \frac{\bar{d} - \mu_d}{s_d/\sqrt{n}}$$

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

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

$$\chi^2 = \sum_{i=1}^{k} \frac{O_i - E_i\ ^2}{E_i}$$

$$z = \frac{W - \frac{n\ n+1}{4}}{\sqrt{\frac{n\ n+1\ \ 2n+1}{24}}}$$

$$U_1 = n_1 n_2 + \frac{n_2\ \ n_2+1}{2} - R_2$$

$$U_2 = n_1 n_2 + \frac{n_1\ \ n_1+1}{2} - R_1$$

$$z = \frac{U - \frac{n_1 n_2}{2}}{\sqrt{\frac{n_1 n_2\ \ n_1 + n_2 + 1}{12}}}$$

**- ooo O ooo -**