UNIVERSITI SAINS MALAYSIA

Second Semester Examination
Academic Session 2013/2014

June 2014

BOI 109/4 – Biostatistics
[Biostatistik]

Duration: 3 hours
[Masa : 3 jam]

Please ensure that this examination paper contains <u>TWENTY SEVEN</u> printed pages before you begin the examination.

[Sila pastikan bahawa kertas peperiksaan ini mengandungi <u>DUA PULUH TUJUH </u>muka surat yang bercetak sebelum anda memulakan peperiksaan ini.]

Instructions: Answer **FIVE** (5) out of **SIX** (6) questions, in English or Bahasa Malaysia. Each question carries 20 marks.

*[**Arahan:** Jawab **LIMA** (5) daripada **ENAM** (6) soalan yang diberikan dalam Bahasa Inggeris atau Bahasa Malaysia. Tiap-tiap soalan bernilai 20 markah.]*

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]   Nine patients were asked to evaluate a new therapeutic pillow to treat sleeping disorder. Each patient was asked to judge the pillow based on two variables, firmness and comfort. The patients each gave a score from 1 to 10; higher scores indicate greater firmness or comfort. Assume the data are not normally distributed. The results are already converted into ranks:

[Sembilan pesakit diminta untuk menilai sejenis bantal terapeutik untuk merawat gangguan tidur. Setiap pesakit diminta untuk menilai bantal itu berdasarkan dua pembolehubah, kepejalan dan keselesaan. Setiap pesakit memberi markah dari 1-10, nilai yang tinggi menunjukkan keselesaan dan kepejalan yang paling baik. Andaikan data tidak bertaburan normal. Keputusan telah ditukarkan kepada pangkat:]

Patients [Pesakit]	Firmness (x) [Kepejalan(x)]	Comfort(y) [Keselesaan(y)]
1	2	7
2	4	6
3	8	1.5
4	7	4
5	6	5
6	1	1.5
7	5	9
8	9	8
9	3	3

Calculate the strength of relationship between these two variables and determine whether it is significant.

[Hitung kekuatan hubungkait antara dua pembolehubah ini dan tentukan sama ada ianya signifikan.]

(10 marks / 10 markah)

[b] A study hypothesized that intelligence (based on IQ scores) might be related to height (inches). Ten students were selected and their height and IQ scores were recorded. Sketch a simple scatter gram graph based on the data given. Assume the data is normally distributed, calculate the correlation coefficient (r) and determine whether it is significant.

[Satu kajian mengemukakan hipotesis bahawa kebijaksanaan (berdasarkan markah IQ) mungkin berkorelasi dengan ketinggian (inci). Sepuluh pelajar telah dipilih dan ketinggian serta markah IQ mereka direkodkan. Lakarkan graf gram berselerak ringkas berdasarkan data yang diberi. Andaikan data bertaburan normal, hitungkan nilai koefisien korelasi (r) dan tentukan sama ada ianya signifikan.]

Students [Pelajar]	Height (inches) [Ketinggian (inci)]	IQ score [Markah IQ]
1	60	103
2	62	100
3	63	98
4	65	95
5	65	110
6	67	108
7	68	104
8	70	110
9	70	97
10	71	100

(10 marks / 10 markah)

2. [a] Write short notes on:

[Tulis nota ringkas mengenai]

[i] Central Limit Theorem

[Central Limit Theorem]

[ii] Normal Distribution

[Taburan Normal]

[iii] Binomial Distribution

[Taburan Binomial]

[iv] Poisson Distribution

[Taburan Poisson]

(8 marks / 8 markah)

[b] A common symptom of food poisoning is an increase in body temperature causing fever. A sample of body temperature readings (Fahrenheit) were taken from 12 patients suffering from food poisoning. Caution: sample size is quite small. The results are:

[Satu simptom biasa bagi keracunan makanan ialah peningkatan suhu badan mengakibatkan demam. Sampel bacaan suhu badan (Fahrenheit) telah diambil dari 12 orang pesakit yang mengalami keracunan makanan. Perhatian: saiz sampel yang agak kecil. Keputusannya ialah:]

Body Temperature (Fahrenheit) [Suhu Badan(Fahrenheit)]
100.5
101.9
101.3
100.4
102.4
100.8
101.7
101.7
100.8
100.1
101.5
101.6

[i] Find the point estimation for the mean and the variance of the patient's body temperature.

[Cari titik anggaran bagi nilai min dan varians bagi suhu badan pesakit.]

[ii] Find the interval estimate at 95 % confidence interval (two tail test) for the patient's body temperature.

[Cari anggaran selang pada 95 % selang keyakinan (ujian dua hala) untuk suhu badan pesakit.]

[iii] Describe what would happen if we were to increase the sample size.

[Terangkan apa yang mungkin terjadi jika saiz sampel ditambah.]

[iv] A new patient was admitted to a hospital (Patient X) with a fever and body temperature of 102 °F. Can we conclude patient X suffered from food poisoning?

[Seorang pesakit baru (Pesakit X) telah dimasukkan ke hospital akibat demam dengan suhu badan 102 °F. Bolehkah kita membuat kesimpulan bahawa pesakit X mengalami keracunan makanan?]

(12 marks / 12 markah)

3. [a] In hypothesis testing, which of the following statements is always true?

    [Dalam ujian hipotesis, manakah pernyataan yang sentiasa benar?]

    [i] The P-value is greater than the significance level.

    [Nilai P melebihi nilai signifikan.]

    [ii] The P-value is computed from the significance level.

    [Nilai P dikira daripada nilai signifikan.]

    [iii] The P-value is the parameter in the null hypothesis.

    [Nilai P ialah parameter dalam hipotesis nol.]

    [iv] The P-value is a test statistic.

    [Nilai P ialah nilai kajian statistik.]

    [v] The P-value is a probability.

    [Nilai P ialah kebarangkalian.]

    (4 marks / 4 markah)

    [b] [i] Describe the two types of errors that can be made when you conduct a test of hypothesis.

        [Terangkan dua jenis ralat yang boleh berlaku apabila menjalankan ujian hipotesis.]

        [ii] What is the main differences between paired and independent samples?

        [Apakah perbezaan utama antara sampel berpasangan dan sampel tak bersandar?]

        [iii] Explain briefly the purpose of paired data.

        [Terangkan secara ringkas tujuan data berpasangan]

        [iv] How do nonparametric tests differ from parametric ones?

        [Bagaimana ujian tidak berparametrik berbeza dengan ujian parametrik?]

        (16 mark / 16 markah)

4. Suppose that you are interested in determining whether exposure to the organochlorine DDT, which has been used extensively as pesticide from 1940s until the 1970s, is associated with breast cancer in women. As part of a study that investigated this issue, blood was drawn from a sample of women diagnosed with breast cancer over a ten-year period and from a sample of healthy control persons matched to the cancer patients on age, menopausal status, and date of blood donation. Each woman's blood level of [o,p'-DDT]—(an important by-product of DDT in the human blood) was measured, and the difference in levels for each patient and her matched control calculated. A sample of 171 such differences has mean $\bar{d} = 2.7$ ng/ml and standard deviation $s_d = 15.9$ ng/ml.

[Andaikan kamu berminat untuk menentukan sama ada pendedahan kepada organoklorin , DDT, yang digunakan secara meluas sebagai pestisid dari 1940 hingga 1970 dikaitkan dengan kanser payudara wanita. Sebagai sebahagian dari kajian yang menyiasat isu ini, sampel darah wanita yang didiagnos dengan kanser payudara telah diambil dalam tempoh 10 tahun dan sampel dari kumpulan kawalan yang sihat berpadanan dengan pesakit kanser berdasarkan umur, status menopaus, dan tarikh darah diambil. Paras [o,p'-DDT] – (bahan sampingan DDT di dalam darah) dalam darah setiap wanita telah diukur, dan perbezaan paras bagi setiap pesakit dan kawalan berpadanan telah dikira. Sebanyak 171 sampel nilai perbezaan tersebut mempunyai min $\bar{d} = 2.7$ ng/ml dan sisihan piawai $s_d = 15.9$ ng/ml.]

[a] Test the null hypothesis that the mean blood levels of DDT are identical for women with breast cancer and for healthy control subjects.

[Lakukan ujian hipotesis nol bahawa nilai min darah DDT adalah sama bagi wanita berpenyakit kanser dengan subjek kawalan yang sihat.]

[b] What do you conclude?

[Apa kesimpulan anda?]

(20 marks / 20 marakah)

5. A study was carried out to determine whether an expectant mother's cigarette smoking has any effect on the bone mineral content of her otherwise healthy child. A sample of 77 newborns whose mothers smoked during pregnancy has mean bone mineral content $\overline{X}_1$ = 0.098 g/cm and standard deviation s_1 = 0.026 g/cm; a sample of 161 infants whose mothers did not smoke has mean $\overline{X}_2$ = 0.095 g/cm and standard deviation s_2 = 0.025 g/cm. Assume that these population variances are equal.

[Satu kajian telah dijalankan untuk menentukan sama ada tabiat merokok seorang ibu mengandung akan memberi kesan kepada kandungan mineral tulang bayi yang sepatutnya sihat. Sebanyak 77 sampel bayi baru lahir yang ibunya merokok sewaktu mengandung mempunyai nilai min kandungan mineral tulang $\overline{X}_1$ = 0.098 g/cm dan sisihan piawai s_1 = 0.026 g/cm; manakala sebanyak 161 sampel bayi yang ibunya tidak merokok mempunyai nilai min $\overline{X}_2$ = 0.095 g/cm dan sisihan piawai s_2 = 0.025 g/cm. Andaikan varians populasi adalah sama.]

[a] Are the two samples of data paired or independent?

[Adakah dua sampel data itu berpasangan atau tidak bersandar?]

[b] State the null and alternative hypotheses of the two-sided test.

[Nyatakan hipotesis nol dan hipotesis kajian dalam ujian dua hala.]

[c] Conduct the test at the α = 0.05 level of significance.

[Jalankan ujian pada α = 0.05 aras signifikan.]

[d] What do you conclude?

[Apakah kesimpulan anda?]

(20 marks / 20 marakah)

6. A botanist fertilized mangrove trees *Avicennia marina* with urea, triple phosphorus and guano fertilizer. After a few months, the growth rate (measured as tree height) was recorded. The table below shows the height (meter) of *Avicennia marina* trees after fertilization.

[Seorang penyelidik botani membaja pokok bakau Avicennia marina dengan urea, phosphorus tiga ganda dan guano. Selepas beberapa bulan, kadar tumbesaran Avicennia marina (diukur dari segi ketinggian pokok) telah direkodkan. Jadual berikut menunjukkan ketinggian (meter) Avicennia marina setelah pembajaan.]

Height (meter)		
[Ketinggian (meter)]		
Urea (Ur)	Triple phosphorus (TP)	Guano (G)
2.0	2.3	3.1
2.1	2.9	1.5
2.4	1.5	2.2
2.8	1.2	2.9
2.9	1.9	1.7
3.1	1.9	2.1
3.2	3.4	2.8
3.7	2.1	1.5
3.8	2.6	2.8
4.1	2.4	2.2

[a] Identify whether there were significant differences in height among the trees treated with the three fertilizers.

[Kenalpasti sama ada terdapat perbezaan signifikan antara ketinggian pokok yang telah dibaja dengan tiga jenis baja tersebut.]

(15 marks / 15 markah)

[b] This experiment took place on a homogeneous plot of coastal area. Sketch the diagram of the experimental design to include the assumption, type of experimental design used and allocation of the fertilizers used in this experiment.

[Eksperimen ini telah dilakukan di kawasan pesisir pantai yang sekata. Lukis gambar rajah reka bentuk eksperimen yang mempunyai andaian, jenis reka bentuk eksperimen dan peruntukan jenis baja yang telah digunakan untuk eksperimen ini.]

(5 marks / 5 markah)

- 10 -

- oooOooo -