UNIVERSITI SAINS MALAYSIA

Final Examination
2016/2017 Academic Session

May/June 2017

JIM 414 – Statistical Inference
[Pentaabiran Statistik]

Duration: 3 hour
[Masa: 3 jam]

Please ensure that this examination paper contains **ELEVEN** printed pages before you begin the examination.

Answer **ALL** questions. You may answer either in Bahasa Malaysia or in English.

Read the instructions carefully before answering.

Each question is worth 100 marks.

In the event of any discrepancies, the English version shall be used.

*Sila pastikan bahawa kertas peperiksaan ini mengandungi **SEBELAS** muka surat yang bercetak sebelum anda memulakan peperiksaan.*

*Jawab **SEMUA** soalan. Anda dibenarkan menjawab sama ada dalam Bahasa Malaysia atau Bahasa Inggeris.*

Baca arahan dengan teliti sebelum anda menjawab soalan.

Setiap soalan diperuntukkan 100 markah.

Sekiranya terdapat sebarang percanggahan pada soalan peperiksaan, versi Bahasa Inggeris hendaklah digunapakai.

1. Let $X_1, X_2, \ldots, X_n$ be a random sample from a Bernoulli(θ) distribution, $0 < \theta < 1$.

(a) Find the maximum likelihood estimator of θ.

(20 marks)

(b) Find the sufficient statistic of θ.

(20 marks)

(c) Find the minimum variance unbiased estimator of θ.

(20 marks)

(d) Let $\delta = \theta(1-\theta)$. Find the minimum variance unbiased estimator of δ.

(20 marks)

(e) Show that (d) is a consistent estimator of δ.

(20 marks)

2. (a) Determine whether the following statistics are minimally sufficient or not. Support your determination with an argument.

(i) $X \sim$ Normal(0, σ^2). Is $T = X$ minimally sufficient?

(ii) $X \sim$ Normal(0, σ^2). Is $T = X^2$ minimally sufficient?

(iii) X_1, X_2, X_3 is a random sample from a Bernoulli(θ) distribution. Is $T = \sum_{i=1}^{3} X_i$ minimally sufficient?

(iv) X_1, X_2, X_3 is a random sample from a Bernoulli(θ) distribution. Given that the values of T are assigned as follows:

$$T = \begin{cases} 0, & (X_1, X_2, X_3) = (0,0,0) \\ 1, & (X_1, X_2, X_3) = (0,0,1) \\ 1, & (X_1, X_2, X_3) = (0,1,0) \\ 1, & (X_1, X_2, X_3) = (1,0,0) \\ 73, & (X_1, X_2, X_3) = (0,1,1) \\ 73, & (X_1, X_2, X_3) = (1,0,1) \\ 91, & (X_1, X_2, X_3) = (1,1,0) \\ 103, & (X_1, X_2, X_3) = (1,1,1). \end{cases}$$

Is T minimally sufficient?

(v) $X_1, X_2, \ldots, X_n$ is a random sample from a Uniform$\left(\theta-1, \theta+1\right)$ distribution. Let Y_1 = min($X_1, X_2, \ldots, X_n$) and Y_n = max($X_1, X_2, \ldots, X_n$). Is $T = \frac{Y_1 + Y_n}{2}$ minimally sufficient?

(75 marks)

(b) The following data are the ordered realizations of a random sample of size 15 on a random sample *X*.

56 70 89 94 96 101 102 102 102 105 106 108 110 113 116

Let $\xi_{1/2}$ denote the median of $F(x)$. Obtain the 88% confidence interval for $\xi_{1/2}$.

(25 marks)

3. (a) $X_1, X_2, \ldots, X_n$ is a random sample from a Normal(μ, σ^2) distribution. Given

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \text{ and } S^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2.$$

(i) Find the distribution of the pivot $T = \dfrac{\bar{X} - \mu}{S/\sqrt{n}}$.

(ii) Is T an ancillary statistic? State the reason for your answer.

(iii) Obtain the $100(1-\alpha)\%$ confidence interval for μ.

(30 marks)

(b) Let X_1 and X_2 be a random sample of size 2 taken from an exponential distribution with mean θ. Reject H_0: $\theta = 2$ and accept H_1: $\theta = 1$ if the observed values of X_1 and X_2 resulted in

$$\frac{f\left(x_1;\theta=2\right) f\left(x_2;\theta=2\right)}{f\left(x_1;\theta=1\right) f\left(x_2;\theta=1\right)} \le \frac{1}{2}.$$

Find the significance level of the test and power of the test when H_0 is false.

(40 marks)

(c) $X_1, X_2, \ldots, X_{10}$ is a random sample from a Poisson(λ) distribution. A critical region for testing H_0: $\lambda = 0.1$ against H_1: $\lambda > 0.1$ is given by $Y = \sum_{i=1}^{10} X_i \geq 3$.

(i) Find the significance level of this test.

(ii) Suppose the critical region is now $Y = \sum_{i=1}^{10} X_i \geq 4$. What is the significance level now?

(iii) A significance level of 0.05 can be achieved if we state the rejection rule as: Reject H_0 if $\sum_{i=1}^{10} X_i \geq 4$ or if $\sum_{i=1}^{10} X_i = 3$ and W = 1, where W has a Bernoulli(p) distribution. Find $P(W = 1) = p$.

(30 marks)

4. (a) $X_1, X_2, \ldots, X_n$ is a random sample from an Exponential(θ) distribution. We want to test H_0: $\theta = \theta_0$ versus H_1: $\theta \neq \theta_0$.

(i) Show that $\frac{2}{\theta_0} \sum_{i=1}^{n} X_i \sim \chi^2_{2n}$.

(ii) Construct likelihood ratio test of size $\alpha = 0.05$, when $\theta_0 = 1$, $n = 10$.

(30 marks)

(b) $X_1, X_2, \ldots, X_n$ is a random sample from a Normal(μ, σ^2) dsitribution. Let $\boldsymbol{\theta} = \begin{bmatrix} \mu \\ \sigma^2 \end{bmatrix}$.

(i) Obtain the maximum likelihood estimator, $\widehat{\boldsymbol{\theta}}$.

(ii) Determine whether the elements of the maximum likelihood estimator vector $\widehat{\boldsymbol{\theta}}$ are consistent and/or unbiased.

(30 marks)

(c) The recent United States presidential race had 4 candidates. Prior to election day, a poll was conducted. Assume that those polled are selected independently of one another and that each has selected one candidate. Let (X_1, X_2) be the random vector of 0's and 1's where X_i = 1 when the i^{th} candidate was selected, i =1, 2 representing the top 2 candidates. Let p_i be the probability of selecting candidate i =1, 2.

(i) What distribution is appropriate to model this event?

(ii) Construct the likelihood ratio test statistic, Λ for H_0: $p_1 = p_2$ versus H_1: $p_1 \neq p_2$.

(iii) Suppose 200 people were polled, with 92 selecting candidate X_1, 86 selecting candidate X_2 while 22 selecting either one of the remaining candidates. Run the test in (ii) using the rejection criterion $2\log \Lambda^{-1} > \chi^2_{0.05;1}$.

(40 marks)

5. (a) Let X be the binomial random variable with probability of success p and the number of trials n = 5. We want to test H_0: $p = \frac{1}{2}$ versus H_1: $p = \frac{3}{4}$.

(i) Fill in the following table:

x	$f(x;\ p = \frac{1}{2})$	$f(x;\ p = \frac{3}{4})$	$f(x:\ p = \frac{1}{2})/f(x;\ p = \frac{3}{4})$
0			
1			
2			
3			
4			
5			

(ii) Using X, determine the best critical region for the test at $\alpha = \frac{6}{32}$. What is its power?

(50 marks)

(b) X_1, X_2 is a random sample from an Exponential(θ) distribution. Construct a uniformly most powerful test for H_0: $\theta = 2$ versus H_1: $\theta > 2$ at $\alpha = 0.05$.

(20 marks)

(c) $X_1, \ldots, X_n$ is a random sample from a population with density function

$$f\left(x;\gamma,\theta\right) = \frac{1}{\theta}e^{-(x-\gamma)/\theta},\ \ x > \gamma.$$

Construct the likelihood ratio test of H_0: $\gamma = \gamma_0$ versus H_1: $\gamma \neq \gamma_0$ using the F sampling distribution.

(30 marks)

1. Andaikan $X_1, X_2, \ldots, X_n$ sebagai suatu sampel rawak daripada suatu taburan Bernoulli(θ), $0 < \theta < 1$.

   (a) Cari penganggar kebolehjadian maksimum bagi θ.
   (20 markah)

   (b) Cari statistik cukup bagi θ.
   (20 markah)

   (c) Cari penganggar saksama bervarians minimum bagi θ.
   (20 markah)

   (d) Andaikan $\delta = \theta(1-\theta)$. Cari penganggar saksama bervarians minimum bagi δ.
   (20 markah)

   (e) Tunjukkan bahawa (d) adalah penganggar konsisten bagi δ.
   (20 markah)

2. (a) Tentukan sama ada statistik-statistik berikut cukup secara minimum. Sokong penentuan anda dengan hujah.

   (i) X ~ Normal(0, σ^2). Adakah $T = X$ cukup secara minimum?

   (ii) X ~ Normal(0, σ^2). Adakah $T = X^2$ cukup secara minimum?

   (iii) X_1, X_2, X_3 adalah suatu sampel rawak daripada suatu taburan Bernoulli(θ). Adakah $T = \sum_{i=1}^{3} X_i$ cukup secara minimum?

   (iv) X_1, X_2, X_3 adalah suatu sampel rawak daripada suatu taburan Bernoulli(θ). Diberikan nilai-nilai T yang diagihkan secara berikut:

$$T = \begin{cases} 0, & (X_1, X_2, X_3) = (0,0,0) \\ 1, & (X_1, X_2, X_3) = (0,0,1) \\ 1, & (X_1, X_2, X_3) = (0,1,0) \\ 1, & (X_1, X_2, X_3) = (1,0,0) \\ 73, & (X_1, X_2, X_3) = (0,1,1) \\ 73, & (X_1, X_2, X_3) = (1,0,1) \\ 91, & (X_1, X_2, X_3) = (1,1,0) \\ 103, & (X_1, X_2, X_3) = (1,1,1). \end{cases}$$

   Adakah T cukup secara minimum?

(v) $X_1, X_2, \ldots, X_n$ adalah suatu sampel rawak daripada suatu taburan Seragam $(\theta - 1, \theta + 1)$. Andaikan Y_1 = min($X_1, X_2, \ldots, X_n$) dan Y_n = maks($X_1, X_2, \ldots, X_n$). Adakah $T = \dfrac{Y_1 + Y_n}{2}$ cukup secara minimum?

(75 markah)

(b) Data berikut adalah nilai-nilai mengikut tertib daripada suatu sampel rawak X bersaiz 15.

56 70 89 94 96 101 102 102 102 105 106 108 110 113 116

Andaikan $\xi_{1/2}$ sebagai median bagi $F(x)$. Dapatkan selang keyakinan 88% bagi $\xi_{1/2}$.

(25 markah)

3. (a) $X_1, X_2, \ldots, X_n$ adalah suatu sampel rawak daripada taburan Normal(μ, σ^2).

Diberikan $\bar{X} = \dfrac{1}{n}\sum_{i=1}^{n} X_i$ dan $S^2 = \dfrac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2$.

(i) Cari taburan bagi pangsi $T = \dfrac{\bar{X} - \mu}{S/\sqrt{n}}$.

(ii) Adakah T suatu statistik sampingan? Nyatakan sebab anda menjawab sedemikian.

(iii) Dapatkan selang keyakinan $100(1 - \alpha)$% bagi μ.

(30 markah)

(b) Andaikan X_1 dan X_2 sebagai suatu sampel rawak bersaiz 2 diambil daripada taburan eksponen yang mempunyai min θ. Tolak H_0: $\theta = 2$ dan terima H_1: $\theta = 1$ jika nilai-nilai X_1 dan X_2 yang dicerap menyebabkan

$$\frac{f\left(x_1;\theta=2\right) f\left(x_2;\theta=2\right)}{f\left(x_1;\theta=1\right) f\left(x_2;\theta=1\right)} \le \frac{1}{2}.$$

Cari aras keertian ujian dan kuasa ujian apabila H_0 didapati palsu.

(40 markah)

(c) $X_1, X_2, \ldots, X_{10}$ adalah suatu sampel rawak daripada taburan Poisson(λ). Rantau genting untuk menguji H_0: $\lambda = 0.1$ lawan H_1: $\lambda > 0.1$ diberikan oleh $Y = \sum_{i=1}^{10} X_i \geq 3$.

(i) Cari aras keertian ujian ini.

(ii) Andaikan rantau genting sekarang ialah $Y = \sum_{i=1}^{10} X_i \geq 4$. Apakah aras keertian ujian ini sekarang?

(iii) Aras keertian 0.05 boleh dicapai jika kita nyatakan petua penolakan sebagai: Tolak H_0 jika $\sum_{i=1}^{10} X_i \geq 4$ atau jika $\sum_{i=1}^{10} X_i = 3$ dan $W = 1$, di mana W tertabur secara Bernoulli(p). Cari $P(W = 1) = p$.

(30 markah)

4. (a) $X_1, X_2, \ldots, X_n$ adalah suatu sampel rawak daripada suatu taburan Eksponen(θ). Kita hendak menguji H_0: $\theta = \theta_0$ lawan H_1: $\theta \neq \theta_0$.

(i) Tunjukkan bahawa $\frac{2}{\theta_0}\sum_{i=1}^{n} X_i \sim \chi^2_{2n}$.

(ii) Bina ujian nisbah kebolehjadian bersaiz $\alpha = 0.05$, apabila $\theta_0 = 1$, $n = 10$.

(30 markah)

(b) $X_1, X_2, \ldots, X_n$ adalah suatu sampel rawak daripada suatu taburan Normal(μ, σ^2). Biar $\boldsymbol{\theta} = \begin{bmatrix} \mu \\ \sigma^2 \end{bmatrix}$.

(i) Dapatkan pengangar kebolehjadian maksimum, $\widehat{\boldsymbol{\theta}}$.

(ii) Tentukan sama ada unsur-unsur vektor penganggar kebolehjadian maksimum konsisten dan/atau saksama.

(30 markah)

(c) Terdapat 4 orang calon di dalam pilihan raya presiden Amerika Syarikat yang lepas. Sebelum hari pilihan raya, suatu cabutan suara dijalankan. Andaikan mereka yang mengambil bahagian di dalam cabutan suara tersebut dipilih secara tak bersandar dan setiap orang telah memilih calon mereka. Andaikan (X_1, X_2) sebagai vektor rawak yang mengandungi 0 dan 1 di mana $X_i = 1$ apabila calon ke-i dipilih, i =1, 2 mewakili 2 calon teratas. Andaikan p_i sebagai kebarangkalian memilih calon i =1, 2.

(i) Apakah taburan berpatutan yang boleh dijadikan model peristiwa ini?

(ii) Bina statistik ujian nisbah kebolehjadian, Λ buat H_0: $p_1 = p_2$ lawan H_1: $p_1 \neq p_2$.

(iii) Sekiranya 200 orang mengambil bahagian di dalam cabutan suara, dan 92 orang memilih calon X_1, 86 orang memilih calon X_2 manakala 22 orang memilih salah satu daripada calon-calon yang tertinggal. Jalankan ujian di dalam (ii) dengan menggunakan kriteria penolakan 2 log $\Lambda^{-1} > \chi^2_{0.05;1}$.

(40 markah)

5. (a) Andaikan X ialah pembolehubah rawak binomial dengan kebarangkalian kejayaan p dan bilangan percubaan n = 5.
Kita hendak menguji H_0: $p = \frac{1}{2}$ lawan H_1: $p = \frac{3}{4}$.

(i) Lengkapkan jadual berikut:

x	$f(x;\ p=\frac{1}{2})$	$f(x;\ p=\frac{3}{4})$	$f(x:\ p=\frac{1}{2})/f(x;\ p=\frac{3}{4})$
0			
1			
2			
3			
4			
5			

(ii) Dengan menggunakan X tentukan rantau genting terbaik bagi ujian tersebut pada $\alpha = \frac{6}{32}$. Berapakah kuasa ujian di tahap α ini?

(50 markah)

(b) X_1, X_2 adalah suatu sampel rawak daripada suatu taburan Eksponen(θ). Bina suatu ujian paling berkuasa bagi H_0: $\theta = 2$ lawan H_1: $\theta > 2$ apabila $\alpha = 0.05$.

(20 markah)

(c) $X_1,\ldots,X_n$ adalah suatu sampel rawak daripada suatu populasi yang mempunyai fungsi ketumpatan

$$f\left(x;\gamma,\theta\right)=\frac{1}{\theta}e^{-(x-\gamma)/\theta} \text{ , } x > \gamma.$$

Bina ujian nisbah kebolehjadian untuk H_0: $\gamma = \gamma_0$ lawan H_1: $\gamma \neq \gamma_0$ dengan menggunakan taburan pensampelan F.

(30 markah)

Formulas

1. $\lim_{n\to\infty} P\left(\left|X_n - c\right| < \varepsilon\right) = 1$, for any $\varepsilon > 0$.

2. $\prod_{i=1}^{n} f\left(x_i;\theta\right) = k_1\left[u_1\left(x_1, x_2, \ldots, x_n\right);\theta\right] k_2\left(x_1, x_2, \ldots, x_n\right)$

3. Let $Y_1 < Y_2 < \ldots < Y_n$. $g(y_1, y_2, \ldots, y_n) = n! f(y_1) f(y_2) \ldots f(y_n)$, $y_1 < y_2 < \ldots < y_n$.

4. Let $Y_1 < Y_2 < \ldots < Y_n$. $g_k\left(y_k\right) = \dfrac{n!}{(k-1)!(n-k)!}\left[F\left(y_k\right)\right]^{k-1}\left[1 - F\left(y_k\right)\right]^{n-k} f\left(y_k\right)$

5. Let $Y_1 < Y_2 < \ldots < Y_n$.
$$g_{ij}\left(y_i, y_j\right) = \frac{n!}{(i-1)!(j-i-1)!(n-j)!}\left[F\left(y_i\right)\right]^{i-1}\left[F\left(y_j\right) - F\left(y_i\right)\right]^{(j-i-1)}$$
$$\times\left[1 - F\left(y_j\right)\right]^{n-j} f\left(y_i\right) f\left(y_j\right), y_i < y_j.$$

6. Let $Y_1 < Y_2 < \ldots < Y_n$. Take $c_{a/2}$ to be the a/2th quantile of a Binomial(n, 1/2) distribution. Then $P\left(Y_{c_{\alpha/2}+1} < \xi_{1/2} < Y_{n-c_{\alpha/2}}\right) = 1 - \alpha$

7. $f(x) = p^x(1-p)^{1-x}$, $x = 0, 1$, $0 < p < 1$. $E(X) = p$, $\mathrm{Var}(X) = p(1-p)$. $m(t) = 1 - p + pe^t$.

8. $f\left(x\right) = \binom{n}{x} p^x\left(1-p\right)^{n-x}$, $x = 0, 1, 2, \ldots, n$, $0 < p < 1$. $E(X) = np$, $\mathrm{Var}(X) = np(1-p)$.
$m(t) = (1 - p + pe^t)^n$.

9. $f(x) = p\,(1-p)^x$, $x = 0, 1, 2, \ldots$, $0 < p < 1$. $E\left(X\right) = \dfrac{1-p}{p}$, $\mathrm{Var}\left(X\right) = \dfrac{1-p}{p^2}$.
$m\left(t\right) = \dfrac{p}{1-\left(1-p\right)e^t}$.

10. $f\left(x\right) = \dfrac{1}{b-a}, a < x < b$. $E\left(X\right) = \dfrac{a+b}{2}$, $\mathrm{Var}\left(X\right) = \dfrac{\left(b-a\right)^2}{12}$. $m\left(t\right) = \dfrac{e^{bt} - e^{at}}{\left(b-a\right)t}$.

11. $f\left(x\right) = \dfrac{1}{\sigma\sqrt{2\pi}}\exp\left[-\dfrac{\left(x-\mu\right)^2}{2\sigma^2}\right], \sigma > 0$. $E(X) = \mu$, $\mathrm{Var}(X) = \sigma^2$.
$m\left(t\right) = \exp\left[\mu t + \tfrac{1}{2}\sigma^2 t\right]$.

12. $f\left(x\right) = \dfrac{1}{\theta}e^{-\frac{1}{\theta}x}, x \geq 0, \theta > 0$. $E(X) = \theta$, $\mathrm{Var}\left(X\right) = \theta^2$. $m\left(t\right) = \dfrac{1}{1-\theta t}, t < \dfrac{1}{\theta}$.

13. $f\left(x_1, x_2, \ldots, x_k\right) = \dfrac{n!}{x_1!x_2!\ldots x_k!} p_1^{x_1} p_2^{x_2} \ldots p_k^{x_k}, 0 \leq p_i \leq 1, x_i = 1, 2, \ldots n, i = 1, 2, \ldots k. \sum_{i=1}^{k} p_i = 1,$
$\sum_{i=1}^{k} x_i = n$. $E\left(X_i\right) = np_i$, $\mathrm{Var}\left(X_i\right) = np_i\left(1-p_i\right)$, $\mathrm{Cov}\left(X_i, X_j\right) = -np_i p_j$, $i \neq j$.
$m\left(t_1, t_2, \ldots, t_n\right) = \left(\sum_{i=1}^{k} p_i e^{t_i}\right)^n$.

- oooOooo -