Final Examination
2017/2018 Academic Session

May/June 2018

JIM414 – Statistical Inference
[Pentaabiran Statistik]

Duration: 3 hour
[Masa: 3 jam]

Please ensure that this examination paper contains **NINE** 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 **SEMBILAN** 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 distribution with $E(X_i) = \mu$ and Var$(X_i) = \sigma^2, 0 < \sigma^2 < \infty$. Define $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$, $S_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2$ and $G_n(x)$ be the cumulative distribution function of $\sqrt{n}(\bar{X} - \mu)/\sigma$. Prove or disprove the following.

   (a). Suppose the distribution is discrete. $\bar{X}_n$ converges in probability to μ.

   (25 marks)

   (b). Suppose the distribution is Normal $(0, \sigma^2)$. S_n^2 converges in probability to σ^2.

   (25 marks)

   (c). Regardless of the type of distribution, $G_n(x)$ has a limiting standard normal distribution.

   (50 marks)

2. (a). Let $X_1, X_2, \ldots, X_n$ be a random sample from a Poisson (λ) distribution. Find the sufficient statistic for λ.

   (25 marks)

   (b). Let $X_1, X_2, \ldots, X_n$ be a random sample from a gamma (k, β) distribution with k is fixed. Define $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$.

   (i). Evaluate $\hat{\beta} = \frac{1}{k}\bar{X}_n$ based upon unbiasedness, consistency and efficiency.

   (ii). Show that $\hat{\beta}$ is a minimum variance unbiased estimator for β.

   (75 marks)

3. $X_1, X_2, \ldots, X_n$ is a random sample from a uniform $(0,\theta)$ distribution. Let $Y = \text{Max}\{X_1, X_2, \ldots, X_n\}$. Consider two candidates for interval estimators of $\theta : [aY, bY], 1 \le a < b$ and $[Y+c, Y+d], 0 \le c < d$, where a, b, c, d are constants.

   (a). Find the confidence coefficient of both interval estimators.

   (30 marks)

   (b). What is the main difference between the two confidence coefficients?

   (40 marks)

   (c). What happens to the two confidence coefficients when θ approaches ∞?

   (30 marks)

4. (a). $X_1, X_2, \ldots, X_n$ is a random sample from a Normal $(\theta,1)$ distribution. Given that the maximum likelihood estimator of θ is $\bar{X}$, the sample mean, construct a likelihood ratio test for $H_0 : \theta = \theta_0$ versus $H_1 : \theta \ne \theta_0$.

   (30 marks)

   (b). $X_1, X_2, \ldots, X_{30}$ is a random sample from a Normal $(\theta, 5.76)$ distribution. Let $\bar{X}$, the sample mean, be a statistic in the following test of hypothesis: $H_0 : \theta = 25$ versus $H_1 : \theta > 25$.

   (i). Evaluate the following decision rules: reject H_0 if and only if $\bar{X} > k$, for k = 25.25, 25.718, 26.50. Which would you prefer? Why?

   (ii). Calculate Type II errors when the true mean $\theta = 25.75$ and 26.8 based upon your preferred decision rule in (i).

   (iii). Graph the power curve of this test.

   (70 marks)

5. (a). Let $X_1, X_2, \ldots, X_n$ be a random sample from a negative binomial (r, p) distribution. Define $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$. Approximate $P\left(\bar{X} \le 11\right)$ in terms of r, p and n.

(25 marks)

(b). Let $X_1, X_2, \ldots, X_n$ be a random sample from a uniform $(0,1)$ distribution. Identify the distribution of the k^{th} order statistic, $X_{(k)}$ from this sample.

(25 marks)

(c). What is a pivot? Is it also an ancillary statistic? Give an example to support your answers.

(25 marks)

(d). Let $X_1, X_2, \ldots, X_n$ be a random sample from a Bernoulli (p) distribution. Consider testing $H_0 : p \le p_0$ versus $H_1 : p > p_0$, where $0 < p_0 < 1$. The maximum likelihood estimator of p is given by $\hat{p} = \frac{1}{n}\sum_{i=1}^{n} X_i$. Construct the approximate test for p.

(25 marks)

…5/-

1. Andaikan $X_1, X_2, \ldots, X_n$ sebagai suatu sampel rawak daripada taburan yang mempunyai $E(X_i) = \mu$ dan Var $(X_i) = \sigma^2, 0 < \sigma^2 < \infty$. Takrifkan $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$, $S_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \bar{X}_n\right)^2$ dan $G_n(x)$ sebagai fungsi taburan longgokan bagi $\sqrt{n}\left(\bar{X} - \mu\right)/\sigma$. Buktikan atau sangkalkan pernyataan-pernyataan berikut.

   (a). Andaikan taburan tersebut adalah diskrit. $\bar{X}_n$ menumpu secara kebarangkalian pada μ.

   (25 markah)

   (b). Andaikan taburan tersebut adalah Normal $\left(0, \sigma^2\right)$. S_n^2 menumpu secara kebarangkalian pada σ^2.

   (25 markah)

   (c). Tanpa menghiraukan jenis taburan populasi, taburan penghad bagi $G_n(x)$ ialah taburan normal piawai.

   (50 markah)

2. (a). Andaikan $X_1, X_2, \ldots, X_n$ sebagai suatu sampel rawak daripada taburan Poisson (λ). Cari statistik cukup bagi λ.

   (25 markah)

   (b). Andaikan $X_1, X_2, \ldots, X_n$ sebagai suatu sampel rawak daripada taburan gamma (k, β) dengan k ditetapkan. Takrifkan $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$.

   (i). Nilaikan $\hat{\beta} = \frac{1}{k}\bar{X}_n$ berdasarkan kesaksamaan, konsistensi dan kecekapan.

   (ii). Tunjukkan bahawa $\hat{\beta}$ adalah suatu penganggar saksama bervarians minimum bagi β.

   (75 markah)

…6/-

3. $X_1, X_2, \ldots, X_n$ adalah suatu sampel rawak daripada taburan seragam $(0, \theta)$. Andaikan $Y = \text{Maks}\{X_1, X_2, \ldots, X_n\}$. Pertimbangkan dua calon penganggar selang untuk $\theta : [aY, bY], 1 \le a < b$ dan $[Y + c, Y + d], 0 \le c < d,$ di mana a, b, c, d adalah pemalar.

   (a). Cari pekali keyakinan untuk kedua-dua penganggar selang.

   (30 markah)

   (b). Apakah perbezaan utama di antara kedua-dua pekali keyakinan tersebut?

   (40 markah)

   (c). Apakah yang berlaku kepada kedua-dua pekali keyakinan tersebut apabila θ menuju ∞?

   (30 markah)

4. (a). $X_1, X_2, \ldots, X_n$ adalah suatu sampel rawak daripada suatu taburan Normal $(\theta, 1)$. Diberikan maklumat bahawa penganggar kebolehjadian maksimum bagi θ ialah min sampel, $\bar{X}$, bina suatu ujian nisbah kebolehjadian untuk menguji $H_0 : \theta = \theta_0$ lawan $H_1 : \theta \neq \theta_0$.

   (30 markah)

   (b). $X_1, X_2, \ldots, X_{30}$ adalah suatu sampel rawak daripada taburan Normal $(\theta, 5.76)$. Andaikan min sampel, $\bar{X}$, sebagai statistik di dalam ujian berikut: $H_0 : \theta = 25$ lawan $H_1 : \theta > 25$.

   (i). Nilaikan peraturan-peraturan keputusan berikut: tolak H_0 jika dan hanya jika $\bar{X} > k$, untuk k = 25.25, 25.718, 26.50. Yang mana satu anda utamakan? Kenapa?

   (ii). Kira ralat-ralat Jenis II apabila min sebenar $\theta = 25.75$ dan 26.8 berdasarkan peraturan keputusan yang anda utamakan di dalam (i).

   (iii) Lakarkan graf keluk kuasa ujian ini.

   (70 markah)

5. (a). Andaikan $X_1, X_2, \ldots, X_n$ sebagai suatu sampel rawak daripada taburan binomial negatif (r, p). Takrifkan $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$. Hampirkan $P\left(\bar{X} \le 11\right)$ di dalam sebutan r, p dan n.

(25 markah)

(b). Andaikan $X_1, X_2, \ldots, X_n$ sebagai suatu sampel rawak daripada taburan seragam $(0,1)$. Camkan taburan statistik tertib ke-k, $X_{(k)}$ daripada sampel ini.

(25 markah)

(c). Apakah suatu pangsi? Adakah suatu pangsi itu suatu statistik tambah? Beri suatu contoh untuk menyokong kedua-dua jawapan anda.

(25 markah)

(d). Andaikan $X_1, X_2, \ldots, X_n$ sebagai suatu sampel rawak daripada taburan Bernoulli (p). Pertimbangkan ujian $H_0 : p \le p_0$ lawan $H_1 : p > p_0$, di mana $0 < p_0 < 1$. Penganggar kebolehjadian maksimum bagi p diberikan oleh $\hat{p} = \frac{1}{n}\sum_{i=1}^{n} X_i$. Bina ujian hampiran bagi p.

(25 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. $I\left(\theta\right) = E\left[\left(\frac{\partial \log f\left(X;\theta\right)}{\partial\theta}\right)^2\right] = -E\left(\frac{\partial^2 \log f\left(X;\theta\right)}{\partial\theta^2}\right)$

4. $Var\left(Y\right) \geq \frac{\left[k'\left(\theta\right)\right]^2}{nI\left(\theta\right)}$

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

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

7. Let $Y_1 < Y_2 < \ldots < Y_n$.

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

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

9. $f\left(x\right) = \binom{n}{x} p^x\left(1-p\right)^{n-x}$, $x = 0, 1, 2, \ldots, n$, $0 < p < 1$. $E(X) = np$, $\text{Var}(X) = np(1-p)$.

   $m(t) = (1 - p + pe^t)^n$.

10. $f(x) = p(1-p)^x$, $x = 0, 1, 2, \ldots$, $0 < p < 1$. $E\left(X\right) = \frac{1-p}{p}$, $\text{Var}\left(X\right) = \frac{1-p}{p^2}$.

    $m\left(t\right) = \frac{p}{1-\left(1-p\right)e^t}$.

11. $f\left(x\right) = \binom{r+x-1}{x} p^r\left(1-p\right)^x$, $x = 0, 1, 2, \ldots$, $0 < p < 1$. $E\left(X\right) = \frac{r\left(1-p\right)}{p}$,

    $\text{Var}\left(X\right) = \frac{r\left(1-p\right)}{p^2}$. $m\left(t\right) = \left(\frac{p}{1-\left(1-p\right)e^t}\right)^r$, $t < -\log(1-p)$.

12. $f\left(x\right) = \frac{e^{-\lambda}\lambda^x}{x!}$, $x = 0, 1, 2, \ldots, \lambda \geq 0$. $E\left(X\right) = \text{Var}(X) = \lambda$. $m(t) = e^{\lambda\left(e^t - 1\right)}$.

13. $f(x) = \frac{1}{b-a}, a < x < b.\ E(X) = \frac{a+b}{2},\ \text{Var}(X) = \frac{(b-a)^2}{12}.\ m(t) = \frac{e^{bt} - e^{at}}{(b-a)t}.$

14. $f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right], \sigma > 0.\ E(X) = \mu, \text{Var}(X) = \sigma^2.$

$m(t) = \exp\left[\mu t + \tfrac{1}{2}\sigma^2 t\right].$

15. $f(x) = \frac{1}{B(\alpha,\beta)} x^{\alpha-1}(1-x)^{\beta-1},\ 0 \le x \le 1,\ \alpha > 0,\ \beta > 0.\ B(\alpha,\beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}.$

$E[X] = \frac{\alpha}{\alpha+\beta}, \text{Var}(X) = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}.$

$m(t) = 1 + \sum_{k=1}^{\infty}\left(\prod_{r=0}^{k-1}\frac{\alpha+r}{\alpha+\beta+r}\right)\frac{t^k}{k!}.$

16. $f(x) = \frac{1}{2^{\frac{k}{2}}\Gamma\left(\frac{k}{2}\right)} x^{\frac{k}{2}-1} e^{-\frac{x}{2}}, x \ge 0.\ E(X) = k,\ \text{Var}(X) = 2k.\ m_X(t) = \left(\frac{1}{1-2t}\right)^{\frac{k}{2}}, t < \frac{1}{2}.$

17. $f(x) = \frac{1}{\beta} e^{-x/\beta}, x \ge 0, \beta > 0.\ E(X) = \beta,\ \text{Var}(X) = \beta^2.\ m(t) = \frac{1}{1-\beta t}, t < \frac{1}{\beta}.$

18. $f(x) = \frac{1}{\Gamma(\alpha)\beta} x^{\alpha-1} e^{-x/\beta}, x \ge 0, \alpha > 0, \beta > 0.\ E(X) = \alpha\beta,\ \text{Var}(X) = \alpha\beta^2.$

$m(t) = \left(\frac{1}{1-\beta t}\right)^{\alpha}, t < \frac{1}{\beta}.$

- oooOooo -