BSc and MSci EXAMINATIONS (MATHEMATICS)

May-June 2015

This paper is also taken for the relevant examination for the Associateship of the Royal College of Science.

M3S1/M4S1

Statistical Theory I

Date: Monday, 11th May 2015 Time: 2 pm – 4 pm

Solutions

seen ⇓

1. (a) (i) The conditional distribution of X given T does not depend on θ. 2

(ii) T is sufficient and for any other sufficient statistic S, T is a function of S. 2

(iii) The only function h which satisfies $\mathrm{E}[h(T)] = 0\ \forall\theta$ is $h(T) = 0$ (almost surely). 2

(b) (i) $\ell(\theta) = n\log(\theta) - \theta\sum_{i=1}^{n} x_i$. meth seen ⇓

$U_{\bullet}(\theta) = \ell'(\theta) = \frac{n}{\theta} - \sum_{i=1}^{n} X_i = n(\mu(\theta) - \bar{X})$. By inspection $\bar{X}$ is the CRUE of $\mu(\theta) = \frac{1}{\theta}$. 3

$\mathrm{var}(\bar{X}) = \frac{1}{n\theta^2} = \frac{\mu^2}{n}$ 1

(ii) Take

$$g(\mu) = \int \left(n\mathrm{var}(\bar{X})\right)^{-\frac{1}{2}}\, \mathrm{d}\mu = \int \frac{1}{\mu}\, \mathrm{d}\mu = \log(\mu)\ (+C).$$

3

(iii) Applying Bayes theorem the posterior is

$$\begin{aligned}\pi(\theta|\mathrm{x}) &\propto \pi(\theta)L(\theta)\\ &\propto \theta^{\alpha-1}\mathrm{e}^{-\beta\theta}\theta^{n}\mathrm{e}^{-\theta\sum_{i=1}^{n} x_i}\\ &= \theta^{\alpha+n-1}\mathrm{e}^{-\theta\left(\beta+\sum_{i=1}^{n} x_i\right)}\\ &\propto f_{\mathrm{Gamma}(\alpha+n,\beta+n\bar{x})}(\theta),\end{aligned}$$

which is again a Gamma distribution. 4

(iv) $\frac{\mathrm{d}\log\pi(\theta|\mathrm{x})}{\mathrm{d}\theta} = \frac{\alpha+n-1}{\theta} - (\beta + \sum_{i=1}^{n} x_i)$.

Solving $\frac{\mathrm{d}\log\pi(\theta|\mathrm{x})}{\mathrm{d}\theta} = 0$ gives us the posterior mode $\frac{\alpha+n-1}{\beta+\sum_{i=1}^{n} X_i}$ 3

2. (a) (i) $\alpha = \sup_{\theta \in \Theta_0} \mathrm{P}_\theta(X \in R)$. seen ⇓
$\beta(\theta) = \mathrm{P}_\theta(X \in R)$. $\boxed{2}$

(ii) A test is unbiased if $\beta(\theta_0) \leq \beta(\theta_1)$ for all $\theta_0 \in \Theta_0$, $\theta_1 \in \Theta_1$. Alternatively $\beta(\theta_1) \geq \alpha \; \forall \theta_1 \in \Theta_1$ $\boxed{1}$

(iii) Let $R(\theta_0)$ denote the critical region for a test of size α of $H_0 : \theta = \theta_0$ against $H_1 : \theta \neq \theta_0$. Let $\Psi = \{\theta_0 : X \notin R(\theta_0)\}$. Then Ψ is a $100(1-\alpha)\%$ confidence interval since $\mathrm{P}_\theta(\theta \in \Psi) = \mathrm{P}_\theta(X \notin R(\theta)) = 1 - \alpha$. $\boxed{3}$

(b) (i) $\ell(\theta) = -\frac{n}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^n (Y_i - \theta x_i)^2$.
$U_\bullet(\theta) = \frac{1}{\sigma^2}\sum_{i=1}^n x_i(Y_i - \theta x_i) = \frac{\sum_{i=1}^n x_i^2}{\sigma^2}\left(\frac{\sum_{i=1}^n x_i Y_i}{\sum_{i=1}^n x_i^2} - \theta\right)$.
Hence by inspection $\frac{\sum_{i=1}^n x_i Y_i}{\sum_{i=1}^n x_i^2}$ is the CRUE of θ. $\boxed{3}$
The CRLB is $\frac{\sigma^2}{\sum_{i=1}^n x_i^2}$. $\boxed{1}$

(ii) $\mathrm{E}[\bar{Y}] = \frac{\sum_{i=1}^n x_i}{n}\theta$. Hence an unbiased estimator is $T = \frac{n\bar{Y}}{\sum_{i=1}^n x_i}$. meth seen ⇓ $\boxed{1}$

$\mathrm{var}\left(\frac{n\bar{Y}}{\sum_{i=1}^n x_i}\right) = \frac{n\sigma^2}{\left(\sum_{i=1}^n x_i\right)^2}$. Hence
$\mathrm{Efficiency}(T) = \frac{CRLB}{\mathrm{var}(T)} = \frac{\left(\sum_{i=1}^n x_i\right)^2}{n\sum_{i=1}^n x_i^2}$. $\boxed{2}$

(c) (i) $L(\theta) = (1-\theta)\theta^{x-1}\frac{\mathrm{e}^{-\theta}\theta^y}{y!} = \frac{1}{y!}(1-\theta)\mathrm{e}^{-\theta}\theta^{x+y-1}$.
$t = x + y$ is minimal sufficient. $\boxed{2}$

(ii) Let $0 < \theta_0 < \theta_1 < 1$. The likelihood ratio for $H_0 : \theta = \theta_0$ against $H_1 : \theta = \theta_1$ is

$$\lambda = \frac{(1-\theta_1)\mathrm{e}^{-\theta_1}}{(1-\theta_0)\mathrm{e}^{-\theta_0}}\left(\frac{\theta_1}{\theta_0}\right)^{t-1}.$$

$\frac{\theta_1}{\theta_0} > 1$ hence this is an increasing function of the sufficient statistic t and the monotone likelihood ratio criterion is satisfied. $\boxed{4}$

(iii) Yes. The likelihood ratio test is uniformly most powerful because the monotone likelihood ratio criterion is satisfied. $\boxed{1}$

3. (a) $S = \sum_{i=1}^{n} X_i$ is complete and sufficient because it is the natural statistic τ of the 1-parameter exponential family followed by $X_1, \ldots, X_n$. meth seen $\Downarrow$ 2

(b) Let $t(x)$ be an unbiased estimate of θ^2. Then

$$\begin{aligned}
&\sum_{x=0}^{\infty} t(x)\theta(1-\theta)^x = \theta^2 \\
\iff &\sum_{x=0}^{\infty} t(x)(1-\theta)^x = \theta \\
\iff &t(0) + t(1)(1-\theta) = \theta \\
\iff &t(0) = 1,\, t(1) = -1,\, \&\, t(x) = 0 \text{ otherwise.}
\end{aligned}$$

Hence an unbiased estimator of θ^2 is given by $T = \mathbb{1}_{X_1=0} - \mathbb{1}_{X_1=1}$. 4

(c) $\ell(\theta) = n\log(\theta) + S\log(1-\theta)$.
$U_\bullet(\theta) = \ell'(\theta) = \frac{n}{\theta} - \frac{S}{1-\theta} = \frac{n}{\theta-1}\left(\bar{X} + 1 - \frac{1}{\theta}\right)$. 2
$I_\bullet(\theta) = \mathrm{E}\left[-\frac{\mathrm{d}}{\mathrm{d}\theta}U_\bullet(\theta)\right] = \mathrm{E}\left[\frac{n}{\theta^2} + \frac{S}{(1-\theta)^2}\right] = \frac{n}{\theta^2} + \frac{\frac{n}{\theta}-n}{(1-\theta)^2} = \frac{n}{\theta^2(1-\theta)}$. 2

(d) $U_\bullet(\theta)$ cannot be written in the form $\frac{1}{c(\theta)}(T-\theta)$ for any statistic T.
Alternatively, note that $\theta \pm \frac{U_\bullet(\theta)}{I_\bullet(\theta)}$ is not a statistic. 2

(e) The Rao-Blackwell estimate is given by

$$\begin{aligned}
\mathrm{E}[\mathbb{1}_{X_1=0}|S=s] &= \mathrm{P}(X_1=0|S=s) \\
&= \frac{\mathrm{P}(S=s|X_1=0)\mathrm{P}(X_1=0)}{\mathrm{P}(S=s)} \\
&= \frac{\binom{n-1+s-1}{s}\theta^{n-1}(1-\theta)^s \cdot \theta(1-\theta)^0}{\binom{n+s-1}{s}\theta^n(1-\theta)^s} \\
&= \frac{(n+s-2)!s!(n-1)!}{(n+s-1)!s!(n-2)!} \\
&= \frac{n-1}{s+n-1}.
\end{aligned}$$

6

(f) Yes. The improved estimator in (f) is unbiased, and it is a function of the complete sufficient statistic S. Hence by the Lehman-Scheffé theorem it is the MVUE. 2

4. (a) sim. seen ⇓

$$L(\theta_X,\theta_Y)=\prod_{i=1}^{n}\left(\frac{1}{\theta_X}\mathbb{1}_{X_i\le\theta_X}\frac{1}{\theta_Y}\mathbb{1}_{Y_i\le\theta_Y}\right)$$
$$=\theta_X^{-n}\theta_Y^{-n}\mathbb{1}_{X_{(n)}\le\theta_X}\mathbb{1}_{Y_{(n)}\le\theta_Y}$$
$$=g\left((\theta_X,\theta_Y),(X_{(n)},Y_{(n)})\right).$$

Hence $(X_{(n)},Y_{(n)})$ is sufficient by the Neyman factorisation theorem. $\boxed{2}$

(b) Both H_0 and H_1 are composite. $\boxed{1}$

(c) (i) unseen ⇓

$$\lambda=\frac{L(X_{(n)},Y_{(n)})}{L(T,T)}=T^{2n}X_{(n)}^{-n}Y_{(n)}^{-n}.$$

Hence $\Lambda=4n\log(T)-2n\log(X_{(n)})-2n\log(Y_{(n)})$. $\boxed{3}$

Assuming H_0 is true, then $\theta=\theta_X=\theta_Y$ is a scale parameter, and

$$\Lambda(\theta X_1,\ldots,\theta Y_n)=4n\log(\theta T)-2n\log(\theta X_{(n)})-2n\log(\theta Y_{(n)})$$
$$=\Lambda(X_1,\ldots,Y_n)-4n\log(\theta)-2n\log(\theta)-2n\log(\theta)$$
$$=\Lambda(X_1,\ldots,Y_n).$$

Hence Λ is ancillary for θ. $\boxed{2}$

(ii) T is a complete sufficient statistic and Λ is ancillary. Hence by Basu's theorem T is independent of Λ. $\boxed{2}$

Using the same approach as in (c)(i), we may write
$\Lambda=4n\log(\theta^{-1}T)-2n\log(\theta^{-1}X_{(n)})-2n\log(\theta^{-1}Y_{(n)})$,
which can be rearranged as
$-2n\log(\theta^{-1}X_{(n)})-2n\log(\theta^{-1}Y_{(n)})=\Lambda-4n\log(\theta^{-1}T)$.
(Alternatively note that we may assume $\theta=1$ without affecting the the distribution of Λ because it is ancillary.)
From facts 4 and 3, $-2n\log(\theta^{-1}X_{(n)})\sim\chi^2_2$ independently of $-2n\log(\theta^{-1}Y_{(n)})\sim\chi^2_2$. Hence by fact 1 the LHS is χ^2_4.
From facts 5 and 3 $-4n\log(\theta^{-1}T)\sim\chi^2_2$, and we showed that it is independent of Λ. Hence $\chi^2_4=\Lambda+\chi^2_2$. Now $\Lambda\sim\chi^2_2$ follows from fact 2. $\boxed{4}$

(iii) The critical region is $\Lambda>z$, where z is chosen such that $\mathrm{P}(\Lambda>z|H_0)=1-F_{\chi^2_2}(z)=\alpha$. Using the hint this simplifies to $e^{-\frac{1}{2}z}=\alpha$ and the solution is $z=-2\log(\alpha)$.
Equivalently, the critical region can be written as
$R=\{(X_1,\ldots,X_n,Y_1,\ldots,Y_n):\Lambda(X_1,\ldots,X_n,Y_1,\ldots,Y_n)>-2\log(\alpha)\}$. $\boxed{2}$

(d) Under H_1 the parameter space is 2-dimensional and under H_0 it is 1-dimensional. Based on Wilks' Theorem one would expect $\Lambda\xrightarrow{d}\chi^2_1$, where the degrees of freedom is the difference between the two dimensionalities. However this is contraddicted by (c)(ii) which states that $\Lambda\sim\chi^2_2$. Wilks' Theorem does not apply here because one of the regularity conditions requires that the range of the samples does not depend on the parameter(s). $\boxed{4}$