M3/M4
S1

Correction to Q 4(ii)

Imperial College
London

BSc, MSci and MSc EXAMINATIONS (MATHEMATICS)

May – June 2012

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

Statistical Theory I

Date: Monday, 21 May 2012. Time: 2.00pm. Time allowed: 2 hours.

This paper has FOUR questions.

Candidates should use ONE main answer book.

Supplementary books may only be used after the main book is full.

Statistical tables will not be provided.

- DO NOT OPEN THIS PAPER UNTIL THE INVIGILATOR TELLS YOU TO.
- Answer all the questions. Each question carries equal weight.
- Credit will be given for all questions attempted, but extra credit will be given for complete or nearly complete answers.
- Affix one of the labels provided to each answer book that you use, but DO NOT USE THE LABEL WITH YOUR NAME ON IT.
- Calculators may not be used.

1. (i) For a parametric family of probability density functions $\{f(x \mid \theta)\}$, define

   (a) *total efficient score* $U_\bullet(\theta)$,

   (b) *total Fisher information* $I_\bullet(\theta)$.

   (ii) Let $X_1, X_2, \ldots, X_n$ be independent identically distributed random variables having the probability density function

$$f(x \mid \theta) \;=\; c(\theta)\, x^\theta (1-x)^{1-\theta} \qquad\qquad (0 \le x \le 1),$$

   where, for unknown parameter θ $(0 < \theta < 1)$, $c(\theta)$ makes the probability density function integrate to 1.

   Let $\xi(\theta) \;=\; \dfrac{\mathrm{d}}{\mathrm{d}\theta}\, \ln c(\theta)$.

   (a) Find $U_\bullet(\theta)$.
   From $U_\bullet(\theta)$ identify the unbiased estimate $\widehat{\xi}$ of $\xi(\theta)$.
   Find $I_\bullet(\theta)$.

   (b) Find $U_\bullet(\xi)$.
   Obtain the variance of $\widehat{\xi}$ in terms of $\xi(\theta)$.

   (c) Explain why there is no unbiased estimator $\widehat{c}$ of $c(\theta)$ having a variance which is the Cramér-Rao lower bound.

2. (i) (a) Explain what is meant by *complete* and *sufficient* when describing a *complete sufficient statistic.*

   (b) Explain briefly why these properties are important in statistical theory.

   (ii) Let $X_1, X_2, \ldots, X_n$ be independent identically distributed random variables from the delayed exponential distribution having the probability density function

$$f(x \mid \theta) \;=\; \begin{cases} e^{\theta - x} & (x > \theta), \\ 0 & \text{(otherwise)}, \end{cases}$$

   where θ is an unknown parameter.

   (a) Why does this probability density function not belong to the Exponential Family?

   (b) Show that $T \;=\; \min(X_1, X_2, \ldots, X_n)$ is sufficient for θ.

   (c) Find the distribution of T.

   (d) Show that the distribution of T is complete.

   (e) Obtain the unique minimum variance unbiased estimator $\widehat{\theta}$ for θ.

   (f) Obtain the variance of $\widehat{\theta}$.

3. (i) What is meant by an *unbiased test of size* α $(0 < \alpha < 1)$?

(ii) Let $x_1, x_2, \ldots, x_n$ be a random sample from *Normal* $(0, \theta^2)$, where parameter θ $(\theta > 0)$ is unknown.

(a) Show that there is a uniformly most powerful test of $H_0 : \theta = \theta_0$ against $H_1 : \theta < \theta_0$, where θ_0 is specified, and find the explicit form of the test.

(b) Find the power function of the test in (a), expressing it in terms of a standard distribution.

(c) Show that the test in (a) is biased for $H_1 : \theta \neq \theta_0$.

Give your reasoning throughout.

4. (i) (a) Given data $\boldsymbol{x} \in \mathbb{X}$ from a known probability distribution having unknown parameter $\theta \in \Theta$, define a $100(1-\alpha)\%$ confidence set $\Psi(\boldsymbol{x})$ for θ.

(b) Illustrate (a) by considering a size α test for each $\theta_0 \in \Theta$ of the hypothesis $H_0 : \theta = \theta_0$ against $H_1 : \theta \neq \theta_0$, and the relationship between the acceptance set $\overline{R}(\theta_0)$ of values of $\boldsymbol{x}$ and the confidence set $\Psi(\boldsymbol{x})$.

(c) What is meant by a *best confidence set*?

(ii) Let $x_1, x_2, \ldots, x_n$ be a random sample from the distribution having the probability density function

$$f(x \mid \theta) \;=\; \theta\, x^{\theta - 1} \qquad (0 < \cancel{\theta} < 1),$$ x

where $\theta > 0$. The prior probability density function for θ, $\pi(\theta)$, is *Exponential*(λ), where λ $(\lambda > 0)$ is known.

(a) Find the posterior probability density function, and identify the distribution.

(b) Find the posterior mean (Bayes mean).

(c) Find the posterior mode (Bayes MLE).

You may wish to use that Gamma (ν, ξ) *has expectation* ν/ξ, *and variance* ν/ξ^2.
Give your reasoning throughout.

M3S1/M4S1 EXAMINATION SOLUTIONS 2011–12

Course: M3S1

Question: 1

Marks & seen/unseen

Parts

(i) Total efficient score $U_{\cdot}(\theta) = \sum_j U_j(\theta) = \sum_j \frac{\partial}{\partial\theta} \ln f_{X_j|\theta}(X_j|\theta)$ — bookwork 1.

Total Fisher information

$$I_{\cdot}(\theta) = \sum_j I_j(\theta) = \sum_j E_{X_j|\theta}\left\{-\frac{\partial U_j(\theta)}{\partial\theta}\right\} \text{ or } \sum_j E_{X_j|\theta}\{U_j(\theta)^2\}$$ 1.

— rest unseen

(ii)(a) $\ln f(x|\theta) = \ln c(\theta) + \theta \ln x + (1-\theta)\ln(1-x)$

$\frac{d}{d\theta}\ln f(x|\theta) = \xi(\theta) - \ln\left(\frac{1-x}{x}\right)$

so $U(\theta) = \xi(\theta) - \ln\left(\frac{1-X}{X}\right)$

$$U_{\cdot}(\theta) = n\Big\{\xi(\theta) - \underbrace{\frac{1}{n}\sum_j \ln\left(\frac{1-X_j}{X_j}\right)}_{=\hat{\xi}}\Big\} = -n(\hat{\xi} - \xi) \quad \text{— (*)}$$

From (*), $\hat{\xi}$ is unbiased for $\xi(\theta)$ (since $E\{U_{\cdot}(\theta)\} = 0$) 6.

$\frac{d^2}{d\theta^2}\ln f(x|\theta) = \frac{d\xi}{d\theta} = \xi'$ say

$I(\theta) = E\left(-\frac{d^2}{d\theta^2}\ln f_{X|\theta}(X|\theta)\right) = -\xi'$ so $I_{\cdot}(\theta) = -n\xi'$ — (**) 5.

(b) $U_{\cdot}(\xi) = \frac{1}{\xi'}U_{\cdot}(\theta)$ & $I_{\cdot}(\xi) = \frac{1}{\xi'^2}I_{\cdot}(\theta) = -\frac{n}{\xi'}$ from (**)

$\text{var}(\hat{\xi}) = \frac{1}{I_{\cdot}(\xi)} = -\frac{\xi'}{n}$ 5.

[Note: From (*) $I_{\cdot}(\theta) = E\{U_{\cdot}(\theta)^2\} = n^2\,\text{var}(\hat{\xi})$]

(c) $U_{\cdot}(\theta) = n\{\ln c(\theta) - \hat{\xi}(\underset{\sim}{X})\}$

so $U_{\cdot}(c) = \frac{n}{\frac{dc(\theta)}{d\theta}}\{\ln c - \hat{\xi}(\underset{\sim}{X})\}$

which cannot be written after non-linear transformation in the form $\kappa(c)\{c - \hat{c}(\underset{\sim}{X})\}$ necessary for $\text{var}(\hat{c})$ to be the CRLB 2.

Setter's initials — Checker's initials

M3S1/M4S1 EXAMINATION SOLUTIONS 2011–12 | Course M3S1

Question 2 | Marks & seen/unseen

Parts

(i)(a) • A statistic $t = t(\underset{\sim}{x})$ is <u>sufficient</u> for θ if $f_{Z|T,\theta}(z|t,\theta)$ does not depend on θ for any statistic $z = z(\underset{\sim}{x})$. bookwork $1\frac{1}{2}$

• A family of distributions $\{f_{T|\theta}\}$ of a statistic t is <u>complete</u> iff the only unbiased estimator of 0 that is a function of t is the statistic that is 0 with probability 1. $1\frac{1}{2}$

• If $t = t(\underset{\sim}{x})$ is <u>sufficient</u> and its distribution is <u>complete</u>, it is a <u>complete sufficient statistic</u>.

(b) For a complete sufficient statistic $t = t(\underset{\sim}{x})$, any function of $t(\underset{\sim}{X})$ is a minimum variance unbiased estimator (MVUE) of its expectation. 1.

rest unseen

(ii)(a) The range (support) of $f_{X|\theta}$ depends on θ, so $f_{X|\theta}$ cannot be an Exponential Family. 2.

(b)
$$f_{\underset{\sim}{X}|\theta}(\underset{\sim}{x}|\theta) = \prod_1^n e^{(\theta - x_i)} \cdot H(x_{min} > \theta)$$
$$= e^{-\Sigma x_i} e^{n\theta} H(x_{min} > \theta)$$
$$= h(\underset{\sim}{x})\, g(x_{min}, \theta)$$
so $t = x_{min}$ is sufficient for θ by Neyman Factorisation. 3.

(c)
$$P_T(X_{min} > t|\theta) = P(\text{each } X_i > t) = \prod_i e^{-(t-\theta)}$$
$$= e^{-n(t-\theta)} \qquad (t > \theta)$$
$$f_{T|\theta}(t|\theta) = n e^{-n(t-\theta)} \qquad (t > \theta)$$
3.

(d) To show that $E\{h(T)\} = 0 \Rightarrow h(t) \equiv 0$ w.p. 1:
$$\int_\theta^\infty h(t) n e^{n(\theta - t)} dt = e^{n\theta} \int_\theta^\infty h(t) n e^{-nt} dt$$
$ne^{n\theta} > 0$ so to show $\int_\theta^\infty h(t) e^{-nt} dt = 0 \Rightarrow h(t) \equiv 0$

Differentiate wrt θ, $-h(\theta) e^{-n\theta} = 0$ whatever θ $\Rightarrow h(\theta) = 0$
since $-e^{-n\theta} < 0$ i.e. $\neq 0$ 3.

Setter's initials | Checker's initials | Page number 2

Course: M3S1

Question: 2 ctd

Marks & seen/unseen

Parts

(ii)(e) Let $Z = n(T-\theta)$, $z = n(t-\theta)$

$$P(Z > n(t-\theta)) = e^{-n(t-\theta)}$$

i.e. $P(Z > z) = e^{-z}$

so Z is Exponential (1) with $E(Z)=1$, $var(Z)=1$

Then $E\{n(T-\theta)\} = 1$ so $E(T) = \frac{1}{n} + \theta$

So $W = T - \frac{1}{n} = X_{min} - \frac{1}{n}$ is unbiased for θ, and is a function only of the sufficient statistic so $X_{min} - \frac{1}{n}$ is UMVU for θ (by Lehmann-Scheffé) — unseen, 4.

(f) $var\{n(T-\theta)\} = 1$ so $n^2 var(T) = 1$

so $var(T) = \frac{1}{n^2}$ — 1.

Setter's initials

Checker's initials

Parts

(i) A test of size α is <u>unbiased</u> if its power function $\beta(\theta) = P(\underset{\sim}{X} \in R \mid \theta)$, where R is the rejection set, satisfies — bookwork

$$\beta(\theta) \begin{cases} \leq \alpha & \theta \in \Theta_0 \quad (\text{size } \alpha) \\ \geq \alpha & \theta \in \overline{\Theta}_0 . \end{cases}$$

2.

(ii)(a) Consider $H_0: \theta = \theta_0$ v. $H_1: \theta = \theta_1 < \theta_0 \quad (\theta_1 > 0)$ — ~~test unseen~~ similar seen in exercise

These are 2 simple hypotheses so the most powerful test by the Neyman-Pearson Lemma is the likelihood ratio test, to reject H_0 if the likelihood ratio $\lambda(\underset{\sim}{x}) > k$ for some k to be determined. — 1.

Here

$$\lambda(\underset{\sim}{x}) = \frac{\left(\frac{1}{\sqrt{2\pi}\,\theta_1}\right)^n e^{-\frac{1}{2}\frac{1}{\theta_1^2}\sum x_i^2}}{\left(\frac{1}{\sqrt{2\pi}\,\theta_0}\right)^n e^{-\frac{1}{2}\frac{1}{\theta_0^2}\sum x_i^2}} = \left(\frac{\theta_0}{\theta_1}\right)^n e^{-\frac{1}{2}\left(\frac{1}{\theta_1^2} - \frac{1}{\theta_0^2}\right)t}$$

where $t = \sum x_i^2$ is sufficient for θ. — unseen

We reject H_0 if $t < c_\alpha$ since $\theta_1 < \theta_0$ so $\frac{1}{\theta_1^2} > \frac{1}{\theta_0^2}$ — 3.

where c_α is s.t.

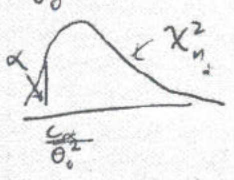
$$\alpha = P(T < c_\alpha \mid \theta = \theta_0) = P\left(Z = \frac{T}{\theta_0^2} < \frac{c_\alpha}{\theta_0^2} \,\middle|\, \theta = \theta_0\right)$$

$$= F_Z\left(\frac{c_\alpha}{\theta_0^2}\right) \quad \text{so} \quad c_\alpha = \theta_0^2 F_Z^{-1}(\alpha) \quad \text{i.e. on } \{\underset{\sim}{x} \mid t(\underset{\sim}{x}) < c_\alpha\}$$ — 3.

Since this test does not depend on the value of θ_1 in $(0, \theta_0)$, it is uniformly most powerful (UMP) for all θ in $(0, \theta_0)$ i.e. for $H_1: \theta < \theta_0$. — 2.

Under H_0,

$\frac{X_i}{\theta_0}$ is $N(0,1)$ so $Z = \sum_i \frac{X_i^2}{\theta_0^2}$ is χ^2_n — 2.



Parts

(ii)(b) For θ, $\frac{X_i}{\theta}$ is $N(0,1)$ & $Z = \sum_i \frac{X_i^2}{\theta^2}$ is χ^2_n — unseen

The power function

$$\beta(\theta) = P(T < c_\alpha | \theta)$$
$$= P\left(Z < \frac{c_\alpha}{\theta^2} \middle| Z \text{ is } \chi^2_n\right)$$
$$= F_Z\left(\frac{c_\alpha}{\theta^2}\right) \quad \text{where } c_\alpha = \theta_0^2 F_Z^{-1}(\alpha)$$

4.

(c) $\beta(\theta) \downarrow$ as $\theta \uparrow$

$\beta(\theta_0) = \alpha$

$$\beta(\theta) \begin{cases} > \alpha & (\theta < \theta_0) \\ = \alpha & (\theta = \theta_0) \\ < \alpha & (\theta > \theta_0) \end{cases}$$

The test is thus biased for testing

$H_0: \theta = \theta_0$ v. $H_1: \theta \neq \theta_0$

3.

Parts

(i)(a) A $100(1-\alpha)\%$ confidence set $\Psi(\underset{\sim}{x})$ for θ is a random set (an interval or union of intervals for one-dimensional parameter θ) that contains the true, fixed unknown θ with probability $1-\alpha$. — bookwork — 2.

(b) Let $\Psi(\underset{\sim}{x})$ be the values in Θ containing those for which an acceptance set is $\underset{\sim}{x} \in \bar{R}(\theta)$, then

$$P(\theta \in \Psi(\underset{\sim}{X}) \mid \theta) = P(\underset{\sim}{X} \in \bar{R}(\theta) \mid \theta) = 1-\alpha.$$ — 2.

(c) A confidence set is <u>best</u> if the probability of its containing a value of θ other than the true one is a small as possible. — 1.

unseen

(ii)(a)

$$\pi(\theta \mid \underset{\sim}{x}) \propto f(\underset{\sim}{x} \mid \theta)\pi(\theta) = \theta^n \left(\prod_i x_i\right)^{\theta-1} \cdot \lambda e^{-\lambda\theta}$$

$$\propto \theta^n z^\theta e^{-\lambda\theta} \quad \text{where } z = \prod_i x_i$$

$$= \theta^n e^{\theta \ln z - \lambda\theta} = \theta^n e^{-\theta t}$$

$$\text{where } t = \lambda - \ln z = \lambda - \sum_i \ln x_i$$ — 6.

$$\propto \frac{t^{n+1}}{\Gamma(n+1)} \theta^{(n+1)-1} e^{-\theta t} \quad \text{i.e. Gamma}(n+1, t)$$ — 2.

(b) The posterior mean is the expectation of Gamma$(n+1, t)$

so $$E(\Theta \mid t) = \frac{n+1}{t} = \frac{n+1}{\lambda - \sum \ln x_i}$$ — 3.

(c) $$\ln \pi(\theta \mid \underset{\sim}{x}) = \text{const.} + n \ln\theta - \theta t$$

$$\frac{d}{d\theta} \ln \pi(\theta \mid \underset{\sim}{x}) = \frac{n}{\theta} - t$$

so $\pi(\theta \mid \underset{\sim}{x})$ is a maximum when $\theta = \frac{n}{t}$

i.e. the posterior mode is at $\theta = \frac{n}{t}$ — 4.