Imperial College
London

M4/5S1

BSc, MSci and MSc EXAMINATIONS (MATHEMATICS)

May-June 2018

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

Statistical Theory I

Date: Friday, 01 June 2018

Time: 2:00 PM - 4:30 PM

Time Allowed: 2.5 hours

This paper has 5 questions.

Candidates should use ONE main answer book.

Supplementary books may only be used after the relevant main book(s) are full.

All required additional material will be provided.

- DO NOT OPEN THIS PAPER UNTIL THE INVIGILATOR TELLS YOU TO.
- Affix one of the labels provided to each answer book that you use, but DO NOT USE THE LABEL WITH YOUR NAME ON IT.
- Each question carries equal weight.
- Calculators may not be used.

1. (a) Write down the definition for each of the following concepts:

   (i) Ancillary statistic.

   (ii) Minimal sufficient statistic.

   (iii) Complete statistic.

   (b) Let $X_1, \ldots, X_n$ be n i.i.d. observations from the following probability density function:

$$f_\theta(x) = \frac{1}{\theta} e^{-\frac{(x-\theta)}{\theta}}, \qquad x > \theta, \ \theta > 0,$$

   where we know that $E(X) = 2\theta$, $E(X^2) = 5\theta^2$, and $E(\min(X_1, \ldots, X_n)) = (1 + n^{-1})\theta$.

   (i) Find a minimal sufficient statistic for θ.

   (ii) Show whether the minimal sufficient statistic obtained in part (i) is complete or not.

   (iii) Obtain a method of moments (MM) estimator for θ.

   (iv) Find the asymptotic distribution of the MM estimator of θ in part (iii).

2. (a) Suppose that $X_1, \ldots, X_n$ are n i.i.d. observations from the following probability mass function:

$$f_{\theta_1,\theta_2}(x) = \begin{cases} \theta_1 & x = 1 \\ \frac{1-\theta_1}{\theta_2 - 1} & x = 2, \ldots, \theta_2 \end{cases} \qquad 0 \le \theta_1 \le 1, \quad \theta_2 \in \{1, 2, 3, \ldots\}.$$

   (i) Obtain the maximum likelihood estimates (MLEs) of parameters θ_1 and θ_2 when at least one observation is equal to 1 and at least one observation is greater than 1.
   (You do not need to check the regularity conditions or evaluate the second derivative of the likelihood.)

   (ii) Derive the asymptotic distribution of the MLE of θ_1 when θ_2 is known. For the asymptotic variance, it is adequate to just provide its general form without calculating it.

   (b) Suppose that $X_1, \ldots, X_n$ are n i.i.d. observations from the Bernoulli distribution with parameter θ where $0 < \theta < 1$.

   (i) Find the UMVUE of $\theta^2(1-\theta)$.

   (ii) Compute the Cramér-Rao lower bound for the variance of unbiased estimators of $\theta^2(1-\theta)$. Explain whether the variance of the UMVUE of $\theta^2(1-\theta)$ attains the lower bound or not.
   [*Hint: The Fisher information for one observation is* $\frac{1}{\theta(1-\theta)}$]

3. (a) Prove that a unique Bayes estimator is admissible.

(b) Suppose that $X_1, \ldots, X_n$ are n i.i.d. observations from $U(\theta - a, \theta + a)$ where a is a known positive constant, and further assume θ has as prior distribution the Exponential distribution with mean 1.

(i) Obtain the posterior distribution of θ.

(ii) Is the prior distribution a conjugate prior? Justify your answer.

(iii) Find the Bayesian point estimator of θ under the squared error loss function.
[*Hint: You may use integration by parts:* $\int u dv = uv - \int v du$]

(iv) Is the Bayes estimator obtained in part (iii) admissible or not?

4. (a) Let $X_1, \ldots, X_n$ be i.i.d. observations from the Exponential distribution with parameter λ where $\lambda > 0$.

(i) Find the uniformly most powerful (UMP) size α test for testing $H_0: \lambda = \lambda_0$ versus $H_1: \lambda < \lambda_0$ where λ_0 is a known positive value.
[*Hint: If $X \sim$ Gamma(a, b) then $cX \sim$ Gamma$(a, \frac{b}{c})$*]

(ii) Find a uniformly most accurate (UMA) $1 - \alpha$ confidence interval for λ.

(b) Now, let $X_1, \ldots, X_n$ be independent Exponential random variables with $E(X_i) = \frac{1}{\lambda t_i}$ where $t_1, \ldots, t_n$ are known positive constants and λ is an unknown parameter. We want to test

$$H_0: \lambda = 1 \quad \text{versus} \quad H_1: \lambda \neq 1.$$

(i) Derive the likelihood ratio test statistic.

(ii) For a given significance level α, perform the likelihood ratio test in part (i) using the asymptotic distribution of the likelihood ratio test statistic under H_0.

Mastery Question:

5. Suppose that $X_1, \ldots, X_n$ are i.i.d. observations from a uniform distribution on interval $(0, \theta)$ where θ is an unknown parameter.

   (a) Find the maximum likelihood estimator (MLE) $\hat{\theta}$ of θ.

   (b) Compute the bias of the MLE of θ. Does the MLE underestimate θ or overestimate it? Why?

   (c) Obtain the jackknife estimator of bias of $\hat{\theta}$.

   (d) Derive the jackknifed estimator of θ.

   (e) Which estimator has a smaller bias: the MLE of θ or the jackknifed estimator of θ? Explain why (without giving a proof).

DISCRETE DISTRIBUTIONS

	RANGE $\mathbb{X}$	PARAMETERS	MASS FUNCTION f_X	CDF F_X	$\mathrm{E}_{f_X}[X]$	$\mathrm{Var}_{f_X}[X]$	MGF M_X
$Bernoulli(\theta)$	$\{0,1\}$	$\theta \in (0,1)$	$\theta^x(1-\theta)^{1-x}$		θ	$\theta(1-\theta)$	$1-\theta+\theta e^t$
$Binomial(n,\theta)$	$\{0,1,\ldots,n\}$	$n \in \mathbb{Z}^+, \theta \in (0,1)$	$\binom{n}{x}\theta^x(1-\theta)^{n-x}$		$n\theta$	$n\theta(1-\theta)$	$(1-\theta+\theta e^t)^n$
$Poisson(\lambda)$	$\{0,1,2,\ldots\}$	$\lambda \in \mathbb{R}^+$	$\frac{e^{-\lambda}\lambda^x}{x!}$		λ	λ	$\exp\{\lambda(e^t-1)\}$
$Geometric(\theta)$	$\{1,2,\ldots\}$	$\theta \in (0,1)$	$(1-\theta)^{x-1}\theta$	$1-(1-\theta)^x$	$\frac{1}{\theta}$	$\frac{(1-\theta)}{\theta^2}$	$\frac{\theta e^t}{1-e^t(1-\theta)}$
$NegBinomial(n,\theta)$	$\{n,n+1,\ldots\}$	$n \in \mathbb{Z}^+, \theta \in (0,1)$	$\binom{x-1}{n-1}\theta^n(1-\theta)^{x-n}$		$\frac{n}{\theta}$	$\frac{n(1-\theta)}{\theta^2}$	$\left(\frac{\theta e^t}{1-e^t(1-\theta)}\right)^n$
or	$\{0,1,2,\ldots\}$	$n \in \mathbb{Z}^+, \theta \in (0,1)$	$\binom{n+x-1}{x}\theta^n(1-\theta)^x$		$\frac{n(1-\theta)}{\theta}$	$\frac{n(1-\theta)}{\theta^2}$	$\left(\frac{\theta}{1-e^t(1-\theta)}\right)^n$

For **CONTINUOUS** distributions (see over), define the **GAMMA FUNCTION**

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1}e^{-x}\,dx$$

and the LOCATION/SCALE transformation $Y = \mu + \sigma X$ gives

$$f_Y(y) = f_X\left(\frac{y-\mu}{\sigma}\right)\frac{1}{\sigma} \qquad F_Y(y) = F_X\left(\frac{y-\mu}{\sigma}\right) \qquad M_Y(t) = e^{\mu t}M_X(\sigma t) \qquad \mathrm{E}_{f_Y}[Y] = \mu + \sigma\mathrm{E}_{f_X}[X] \qquad \mathrm{Var}_{f_Y}[Y] = \sigma^2\mathrm{Var}_{f_X}[X]$$

CONTINUOUS DISTRIBUTIONS

	$\mathbb{X}$	PARAMS.	PDF f_X	CDF F_X	$\mathrm{E}_{f_X}[X]$	$\mathrm{Var}_{f_X}[X]$	MGF M_X
$Uniform(\alpha,\beta)$ (standard model $\alpha=0,\beta=1$)	(α,β)	$\alpha<\beta\in\mathbb{R}$	$\frac{1}{\beta-\alpha}$	$\frac{x-\alpha}{\beta-\alpha}$	$\frac{(\alpha+\beta)}{2}$	$\frac{(\beta-\alpha)^2}{12}$	$\frac{e^{\beta t}-e^{\alpha t}}{t(\beta-\alpha)}$
$Exponential(\lambda)$ (standard model $\lambda=1$)	$\mathbb{R}^+$	$\lambda\in\mathbb{R}^+$	$\lambda e^{-\lambda x}$	$1-e^{-\lambda x}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$	$\left(\frac{\lambda}{\lambda-t}\right)$
$Gamma(\alpha,\beta)$ (standard model $\beta=1$)	$\mathbb{R}^+$	$\alpha,\beta\in\mathbb{R}^+$	$\frac{\beta^\alpha}{\Gamma(\alpha)}x^{\alpha-1}e^{-\beta x}$		$\frac{\alpha}{\beta}$	$\frac{\alpha}{\beta^2}$	$\left(\frac{\beta}{\beta-t}\right)^\alpha$
$Weibull(\alpha,\beta)$ (standard model $\beta=1$)	$\mathbb{R}^+$	$\alpha,\beta\in\mathbb{R}^+$	$\alpha\beta x^{\alpha-1}e^{-\beta x^\alpha}$	$1-e^{-\beta x^\alpha}$	$\frac{\Gamma(1+1/\alpha)}{\beta^{1/\alpha}}$	$\frac{\Gamma(1+2/\alpha)-\Gamma(1+1/\alpha)^2}{\beta^{2/\alpha}}$	
$Normal(\mu,\sigma^2)$ (standard model $\mu=0,\sigma=1$)	$\mathbb{R}$	$\mu\in\mathbb{R},\sigma\in\mathbb{R}^+$	$\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}$		μ	σ^2	$e^{\{\mu t+\sigma^2 t^2/2\}}$
$Student(\nu)$	$\mathbb{R}$	$\nu\in\mathbb{R}^+$	$\frac{(\pi\nu)^{-\frac{1}{2}}\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)\left\{1+\frac{x^2}{\nu}\right\}^{(\nu+1)/2}}$		0 (if $\nu>1$)	$\frac{\nu}{\nu-2}$ (if $\nu>2$)	
$Pareto(\theta,\alpha)$	$\mathbb{R}^+$	$\theta,\alpha\in\mathbb{R}^+$	$\frac{\alpha\theta^\alpha}{(\theta+x)^{\alpha+1}}$	$1-\left(\frac{\theta}{\theta+x}\right)^\alpha$	$\frac{\theta}{\alpha-1}$ (if $\alpha>1$)	$\frac{\alpha\theta^2}{(\alpha-1)(\alpha-2)}$ (if $\alpha>2$)	
$Beta(\alpha,\beta)$	$(0,1)$	$\alpha,\beta\in\mathbb{R}^+$	$\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}x^{\alpha-1}(1-x)^{\beta-1}$		$\frac{\alpha}{\alpha+\beta}$	$\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$	

BSc and MSci EXAMINATIONS (MATHEMATICS)

May-June 2018

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

M3S1/M4S1

Statistical Theory I

Date: Friday, 1 June 2018 Time: 14.00 – 16.00

Solutions

1. (a) (i) **Ancillary statistic:** A statistic is ancillary if its distribution does not depend on any parameter θ. seen ⇓ 2

(ii) **Minimal sufficient statistic:** A minimal sufficient statistic is a sufficient statistic that represents the maximal reduction of the data that contains as much information about the unknown parameter as the data itself does. Equivalently, a minimal sufficient statistic is a statistic that is a function of any other sufficient statistic (i.e., if T is a minimal sufficient statistic, then for any other sufficient statistic T^*, there is a function h such that $T = h(T^*)$). seen ⇓ 2

(iii) **Complete statistic:** A statistic $T = T(X)$ is said to be complete for parameter θ if for any function g, if $E_\theta[g(T)] = 0$ for all $\theta \in \Theta$ then $P_\theta(g(T) = 0) = 1$ for all $\theta \in \Theta$. In simple words, T is a complete statistic if the only unbiased estimator of zero based on T is zero. seen ⇓ 2

meth seen ⇓

(b) (i) For any two sample points x and x^*, we can write

$$\frac{f_\theta(x)}{f_\theta(x^*)} = e^{-\frac{n}{\theta}(\bar{x}-\bar{x}^*)}\frac{I_{(\theta,\infty)}(x_{(1)})}{I_{(\theta,\infty)}(x^*_{(1)})},$$

and this is independent of θ if and only if $\bar{x} = \bar{x}^*$ and $x_{(1)} = x^*_{(1)}$. Therefore, $(\bar{X}, X_{(1)})$ or $(\sum_{i=1}^{n} X_i, X_{(1)})$ as a one-to-one function of it is a minimal sufficient statistic for θ. 4

(ii) Since $E(\bar{X}) = E(X_1) = 2\theta$, we can get unseen ⇓

$$E(\frac{1}{2}(1 + n^{-1})\bar{X} - X_{(1)}) = 0,$$

while $\frac{1}{2}(1 + n^{-1})\bar{X} - X_{(1)} \neq 0$. Therefore, a non-zero function of $(\bar{X}, X_{(1)})$ is an unbiased estimator of zero, and hence $(\bar{X}, X_{(1)})$ is not complete. 4

(iii) A method of moments estimator for θ can be obtained by equating the sample moments with the population moments. From the first moments, we have sim. seen ⇓

$$2\theta = E(X) \doteq \frac{1}{n}\sum_{i=1}^{n} X_i = \bar{X},$$

so an MM estimator for θ is $\widehat{\theta}_{MM} = \bar{X}/2$. 3

(iv) Since $E(X) = 2\theta$ and $\text{var}(X) = E(X^2) - E^2(X) = 5\theta^2 - 4\theta^2 = \theta^2$, by the central limit theorem we get meth seen ⇓

$$\sqrt{n}(\bar{X} - 2\theta) \xrightarrow{D} N(0, \theta^2).$$

Now using the delta method, by considering $g(x) = x/2$, we obtain that

$$\sqrt{n}(\widehat{\theta}_{MM} - \theta) \xrightarrow{D} N(0, \frac{\theta^2}{4}).$$

3

2. (a) (i) First note that the distribution function can be rewritten as sim. seen ⇓

$$f_{\theta_1,\theta_2}(x) = \theta_1^{I(x=1)} \Big(\frac{1-\theta_1}{\theta_2 - 1}\Big)^{I(x\neq 1)} I_{\{1,\dots,\theta_2\}}(x).$$

Then, the likelihood function for the n observations is given by

$$L(\theta_1,\theta_2) = \theta_1^{\sum_{i=1}^{n} I(x_i=1)} \Big(\frac{1-\theta_1}{\theta_2 - 1}\Big)^{\sum_{i=1}^{n} I(x_i\neq 1)} I_{\{1,\dots,\theta_2\}}(x_{(n)}).$$

For each value of θ_1, the likelihood is decreasing in θ_2, therefore $\widehat{\theta}_{2ML} = X_{(n)}$.

4

To obtain the MLE of θ_1, by replacing θ_2 with $x_{(n)}$ we can write the log-likelihood function as follows

$$l(\theta_1) = \ln(\theta_1)\sum_{i=1}^{n} I(x_i = 1) + \ln\Big(\frac{1-\theta_1}{x_{(n)}-1}\Big)\sum_{i=1}^{n} I(x_i \neq 1).$$

Now by taking the derivative w.r.t θ_1, we get the log-likelihood equation as follows

$$\frac{d}{d\theta_1} l(\theta_1) = \frac{\sum_{i=1}^{n} I(x_i = 1)}{\theta_1} - \frac{\sum_{i=1}^{n} I(x_i \neq 1)}{1-\theta_1} = 0.$$

Since $I(x_i \neq 1) = 1 - I(x_i = 1)$, we obtain that

$$\widehat{\theta}_{1ML} = \frac{\sum_{i=1}^{n} I(X_i = 1)}{n}.$$

4

meth seen ⇓

(ii) From a theorem in the lecture notes we know that the MLEs are asymptotically normally distributed, and in fact for this question we have

$$\sqrt{n}(\widehat{\theta}_{1ML} - \theta_1) \xrightarrow{D} N(0, \frac{1}{I(\theta_1)}),$$

where $I(\theta_1)$ is the Fisher information for one observation, which is given by

$$I_{X_1}(\theta_1) = E\Big(\frac{d}{d\theta_1} l_{X_1}(\theta_1)\Big)^2.$$

2

meth seen ⇓

(b) (i) An unbiased estimator of $\theta^2(1-\theta)$ is $I(X_1 = 1, X_2 = 1, X_3 = 0)$. Also, because Bernoulli distribution is a member of full rank exponential families, $T = \sum_{i=1}^{n} X_i$ is a complete and sufficient statistic for θ. Therefore, using the Lehmann-Scheffe Theorem, the UMVUE of $\theta^2(1-\theta)$ is as follows

$$
\begin{aligned}
E(I(X_1=1,X_2=1,X_3=0)|T=t) &= P(X_1=1,X_2=1,X_3=0|\sum_{i=1}^{n} X_i=t) \\
&= \frac{P(X_1=1,X_2=1,X_3=0,\sum_{i=1}^{n} X_i=t)}{P(\sum_{i=1}^{n} X_i=t)} \\
&= \frac{P(X_1=1,X_2=1,X_3=0,\sum_{i=4}^{n} X_i=t-2)}{P(\sum_{i=1}^{n} X_i=t)} \\
&= \frac{\theta^2(1-\theta)\begin{pmatrix} n-3 \\ t-2 \end{pmatrix}\theta^{t-2}(1-\theta)^{n-3-t+2}}{\begin{pmatrix} n \\ t \end{pmatrix}\theta^t(1-\theta)^{n-t}} \\
&= \frac{\begin{pmatrix} n-3 \\ t-2 \end{pmatrix}}{\begin{pmatrix} n \\ t \end{pmatrix}} = \frac{t(t-1)(n-t)}{n(n-1)(n-2)},
\end{aligned}
$$

and, hence, the UMVUE of $\theta^2(1-\theta)$ is

$$\frac{T(T-1)(n-T)}{n(n-1)(n-2)},$$

where $T=\sum_{i=1}^{n} X_i$. $\boxed{6}$

Note: another way to obtain the above UMVUE is to directly find a function of $T=\sum_{i=1}^{n} X_i$ that is unbiased for $\theta^2(1-\theta)$.

meth seen ⇓

(ii) The Cramer-Rao lower bound here is

$$\frac{\left(\frac{d}{d\theta}\left(\theta^2(1-\theta)\right)\right)^2}{I(\theta)} = \frac{\left(\theta(2-3\theta)\right)^2}{nI_{X_1}(\theta)} = \frac{\theta^2(2-3\theta)^2}{\frac{n}{\theta(1-\theta)}} = \frac{\theta^3(1-\theta)(2-3\theta)^2}{n}.$$

$\boxed{2}$

Only estimators of the form $\left\{a\sum_{i=1}^{n} X_i+b\right\}$ achieve the Cramer-Rao lower bound. So the variance of the UMVUE of $\theta^2(1-\theta)$ does not attain the lower bound. $\boxed{2}$

3. (a) Suppose that $\widehat{\theta}_{\text{Bayes}}$ is not admissible, then there must be an estimator $\widetilde{\theta}$ such that seen ⇓

$$R_\theta(\widetilde{\theta}) \leq R_\theta(\widehat{\theta}_{\text{Bayes}}) \qquad \forall \theta \in \Theta.$$

Hence

$$R_B(\widetilde{\theta}) = \int_\Theta R(\widetilde{\theta})\pi(\theta)d\theta \leq \int_\Theta R(\widehat{\theta}_{\text{Bayes}})\pi(\theta)d\theta = R_B(\widehat{\theta}_{\text{Bayes}}),$$

which contradicts the uniqueness of $\widehat{\theta}_{\text{Bayes}}$. Therefore, $\widehat{\theta}_{\text{Bayes}}$ must be admissible. 4

unseen ⇓

(b) (i) First note that we must have $\theta - a \leq x_{(1)} \leq x_{(n)} \leq \theta + a$ which gives us $x_{(n)} - a \leq \theta \leq x_{(1)} + a$.

Now, the posterior is given by

$$\begin{aligned}
\pi(\theta|x_1,\dots,x_n) &= \frac{f_\theta(x_1,\dots,x_n)\pi(\theta)}{\int f_\theta(x_1,\dots,x_n)\pi(\theta)d\theta} = \frac{(2a)^{-n}e^{-\theta}}{\int\limits_{x_{(n)}-a}^{x_{(1)}+a}(2a)^{-n}e^{-\theta}d\theta} \\
&= \frac{e^{-\theta}}{\int\limits_{x_{(n)}-a}^{x_{(1)}+a} e^{-\theta}d\theta} = \frac{e^{-\theta}}{e^{a-x_{(n)}} - e^{-a-x_{(1)}}}, \qquad x_{(n)} - a \leq \theta \leq x_{(1)} + a.
\end{aligned}$$

6

(ii) No, because the posterior distribution is not an Exponential distribution. In fact, the support of θ in the posterior depends on the data (the posterior is a truncated Exponential distribution). meth seen ⇓ 4

(iii) Under the squared error loss, we know that the Bayes estimator is the posterior mean, which can be directly obtained as follows: meth seen ⇓

$$\begin{aligned}
\widehat{\theta}_{\text{Bayes}} &= E(\theta|x_1,\dots,x_n) = \int\limits_{x_{(n)}-a}^{x_{(1)}+a} \frac{\theta e^{-\theta}}{e^{a-x_{(n)}} - e^{-a-x_{(1)}}}d\theta \\
&= \frac{1}{e^{a-x_{(n)}} - e^{-a-x_{(1)}}} \int\limits_{x_{(n)}-a}^{x_{(1)}+a} \theta e^{-\theta}d\theta \\
&= \frac{1}{e^{a-x_{(n)}} - e^{-a-x_{(1)}}}\Big[-\theta e^{-\theta} + \int e^{-\theta}d\theta\Big]_{x_{(n)}-a}^{x_{(1)}+a} \\
&= \frac{1}{e^{a-x_{(n)}} - e^{-a-x_{(1)}}}\Big[-(\theta+1)e^{-\theta}\Big]_{x_{(n)}-a}^{x_{(1)}+a} \\
&= \frac{(x_{(n)} - a + 1)e^{a-x_{(n)}} - (x_{(1)} + a + 1)e^{-a-x_{(1)}}}{e^{a-x_{(n)}} - e^{-a-x_{(1)}}}.
\end{aligned}$$

4

(iv) Yes, the Bayes estimator obtained in part (iii) is admissible because it is unique. meth seen ⇓ 2

4. (a) (i) The joint distribution can be written as $\boxed{\text{meth seen} \Downarrow}$

$$f_\lambda(x) = \lambda^n \exp(-\lambda \sum_{i=1}^{n} x_i)$$

which belongs to the exponential family of distributions and it has monotone likelihood ratio in $-\sum_{i=1}^{n} x_i$. Using the Karlin-Rubin Theorem, the UMP size α test rejects H_0 if $\sum_{i=1}^{n} X_i \geq k$ and k is calculated such that

$$\alpha = P_{H_0}(\sum_{i=1}^{n} X_i \geq k) = P_{\lambda_0}(2\lambda_0 \sum_{i=1}^{n} X_i \geq 2\lambda_0 k) = P_{\lambda_0}(\chi^2(2n) \geq 2\lambda_0 k).$$

Hence $k = \frac{\chi^2_\alpha(2n)}{2\lambda_0}$, where $\chi^2_\alpha(2n)$ satisfies $P(\chi^2(2n) \geq \chi^2_\alpha(2n)) = \alpha$. $\boxed{6}$

(ii) By inverting the UMP size α test we get a UMA $1-\alpha$ confidence interval as follows $\boxed{\text{unseen} \Downarrow}$

$$C(x) = \{\lambda : \ 0 \leq \lambda \leq \frac{\chi^2_\alpha(2n)}{2\sum_{i=1}^{n} X_i}\}.$$

$\boxed{4}$

$\boxed{\text{meth seen} \Downarrow}$

(b) (i) The log-likelihood under the whole parameter space is

$$l(\lambda) = n\ln(\lambda) + \sum_{i=1}^{n} \ln(t_i) - \lambda \sum_{i=1}^{n} t_i x_i,$$

which gives $\widehat{\lambda}_{ML} = \frac{n}{\sum_{i=1}^{n} t_i X_i}$ as the MLE of λ under the whole parameter space. And under H_0, the MLE of λ is $\widehat{\lambda} = 1$. Therefore, the likelihood ratio test statistic is as follows

$$\lambda(x) = \frac{\left(\sum_{i=1}^{n} t_i X_i\right)^n e^{-\sum_{i=1}^{n} t_i X_i}}{n^n e^{-n}}.$$

$\boxed{6}$

$\boxed{\text{meth seen} \Downarrow}$

(ii) Under H_0 and under regularity conditions, the asymptotic distribution of $-2\log(\lambda(x))$ is $\chi^2(1)$. Hence, the likelihood ratio level α test based on the asymptotic distribution rejects H_0 if $-2\log(\lambda(x)) \geq \chi^2_\alpha(1)$. $\boxed{4}$

meth seen ⇓

5. (a) The likelihood function is

$$l(\theta) = \frac{1}{\theta^n} I_{(0,\theta)}(x_{(n)}),$$

which is a decreasing function of θ. Therefore, the MLE of θ is $\widehat{\theta} = X_{(n)}$. 4

meth seen ⇓

(b) The bias of the MLE of θ is equal to

$$b_{\widehat{\theta}}(\theta) = E(X_{(n)}) - \theta = \frac{n}{n+1}\theta - \theta = -\frac{\theta}{n+1}.$$

Because the bias is always negative, the MLE of θ underestimates θ. 4

unseen ⇓

(c) First, it follows that $\widehat{\theta}_{-i} = X_{(n)}$ for $n-1$ values of i and $\widehat{\theta}_{-i} = X_{(n-1)}$ for the other value of i. Thus, we obtain that

$$\widehat{\theta}_{\bullet} = \frac{1}{n}\sum_{i=1}^{n}\widehat{\theta}_{-i} = \frac{n-1}{n}X_{(n)} + \frac{1}{n}X_{(n-1)}.$$

Then, the jackknife estimator of bias is

$$\widehat{b}(\widehat{\theta}) = (n-1)(\widehat{\theta}_{\bullet} - \widehat{\theta}) = (n-1)(\frac{n-1}{n}X_{(n)} + \frac{1}{n}X_{(n-1)} - X_{(n)}) = \frac{n-1}{n}(X_{(n-1)} - X_{(n)}).$$

4

unseen ⇓

(d) The jackknifed estimator of θ is as follows

$$\begin{aligned} \widehat{\theta}_{Jack} &= \widehat{\theta} - \widehat{b}(\widehat{\theta}) = n\widehat{\theta} - (n-1)\widehat{\theta}_{\bullet} \\ &= X_{(n)} + \frac{n-1}{n}(X_{(n)} - X_{(n-1)}). \end{aligned}$$

4

unseen ⇓

(e) The jackknifed estimator $\widehat{\theta}_{Jack}$ has a smaller bias than the MLE $\widehat{\theta}$ because $\widehat{\theta}_{Jack}$ is a bias-corrected version of $\widehat{\theta}$ by reducing the finite-sample bias of the MLE. 4