Statistical Inference: Examination 2023

30 January 2023

Instructions: The time allotted for the examination is 180 minutes. You may answer in either English or French. No written material may be brought into the examination, but a simple calculator may be used if necessary. Full marks may be obtained with complete answers to four questions. The final mark will be based on the best four solutions.

First name:

Last name:

SCIPER number:

Exercise	Points	Indicative marks
1		/10 points
2		/10 points
3		/10 points
4		/10 points
5		/10 points
Total:		/40 points

Some formulae

Definition 1 *The moment-generating and cumulant-generating functions of a real-valued random variable X are*

$$M_X(t) = \mathrm{E}\left(e^{tX}\right), \quad K_X(t) = \log M_X(t), \quad t \in \mathcal{T},$$

where $\mathcal{T} = \{t \in \mathbb{R} : M_X(t) < \infty\}$.

Definition 2 *A normal (or Gaussian) random variable $X \sim \mathcal{N}(\mu, \sigma^2)$ has probability density function*

$$f(x; \mu, \sigma^2) = \frac{1}{\sigma}\phi\left(\frac{x-\mu}{\sigma}\right), \quad x \in \mathbb{R}, \quad \mu \in \mathbb{R}, \sigma^2 > 0,$$

where $\phi(u) = (2\pi)^{-1/2} e^{-u^2/2}$ for $u \in \mathbb{R}$, and we also define $\Phi(x) = \int_{-\infty}^{x} \phi(u)\,\mathrm{d}u$.

Definition 3 *A gamma random variable with shape parameter $\alpha > 0$ and rate parameter $\beta > 0$, $X \sim \mathrm{Gamma}(\alpha, \beta)$, has probability density function*

$$f(x; \alpha, \beta) = \begin{cases} \dfrac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}, & x \geq 0, \\ 0, & x < 0, \end{cases}$$

where $\Gamma(\alpha+1) = \alpha\Gamma(\alpha)$, $\Gamma(\alpha) = (\alpha-1)!$ when α is a positive integer, and $\Gamma(1/2) = \sqrt{\pi}$.

Definition 4 *An exponential random variable X with rate parameter β, $X \sim \exp(\beta)$, has the gamma distribution with $\alpha = 1$.*

Definition 5 *A chi-squared random variable V with ν degrees of freedom, $V \sim \chi^2_\nu$, has the gamma distribution with $\alpha = \nu/2$ and $\beta = 1/2$, and can be expressed as $V \stackrel{\mathrm{D}}{=} Z_1^2 + \cdots + Z_\nu^2$, where $Z_1, \ldots, Z_\nu \stackrel{\mathrm{iid}}{\sim} \mathcal{N}(0,1)$.*

Solution 1

(a) [3, seen] Definition 3. Examples are scattered throughout the course.

(b) [2, seen] Lemma 17.

(c) [5, unseen] The linear combination $\psi X_1 - X_2 \sim \mathcal{N}(\psi \times \lambda - \psi\lambda, \psi^2 + 1)$, so

$$\frac{\psi X_1 - X_2}{\sqrt{\psi^2 + 1}} \sim \mathcal{N}(0, 1)$$

is a pivot, and a $(1 - \alpha)$ confidence set for ψ is

$$\mathcal{I}_{1-\alpha} = \left\{\psi : |\psi X_1 - X_2|/\sqrt{\psi^2 + 1} \leq z_{1-\alpha/2}\right\} = \left\{\psi : \psi^2 X_1^2 - 2\psi X_1 X_2 + X_2^2 \leq \psi^2 q + q\right\},$$

where for simplicity we have written $q = z_{1-\alpha/2}^2$. This set can be expressed as

$$\mathcal{I}_{1-\alpha} = \left\{\psi : \psi^2(X_1^2 - q) - 2\psi X_1 X_2 + (X_2^2 - q) \leq 0\right\}.$$

There are now several possibilities (in principle, some with zero probability of occurring):

- if $X_1^2 - q > 0$, then $\mathcal{I}_{1-\alpha}$ could be empty (if the quadratic has no roots), a single point (if it has one root), or a finite interval (if it has two distinct roots);
- if $X_1^2 - q = 0$, then $\mathcal{I}_{1-\alpha}$ could be a half-line (if $2X_1X_2 \neq 0$) or empty (if $2X_1X_2 = 0$ and $X_2^2 - q > 0$) or the entire real line (if $2X_1X_2 = 0$ and $X_2^2 - q \leq 0$); and
- if $X_1^2 - q < 0$, then $\mathcal{I}_{1-\alpha}$ could be the entire real line (if the quadratic has zero or one roots), or two disjoint half-lines (if it has two distinct roots).

The question did not ask for the exact confidence intervals, just their possible forms, so something like this description would give full marks, if correct.

Solution 2

(a) [3, seen]

(b) [3, unseen] The joint density is

$$\prod_{j=1}^{n} (\alpha + x_j\beta) \exp\{-(\alpha + \beta x_j)y_j\} = \exp\left\{-\alpha \sum_j y_j - \beta \sum x_j y_j + \sum_j \log(\alpha + x_j\beta)\right).$$

The random quantities here are $\sum_j y_j$ and $\sum_j x_j y_j$, so these together form a sufficient statistic. To show that this is minimal, we take another dataset $y_1', \ldots, y_n'$ with the same x_js, and note that the ratio of the corresponding densities will be free of the parameters iff $(\sum_j y_j, \sum_j x_j y_j) = (\sum_j y_j', \sum_j x_j y_j')$.

(c) [4, unseen] The given uniform density is $(2\theta)^{-1} I(-\theta < x < \theta)$, so the likelihood function for a random sample from this density is

$$L(\theta) = \prod_{j=1}^{n} (2\theta)^{-1} I(-\theta < y_j < \theta) = (2\theta)^{-n} I(-\theta < y_{(1)} < y_{(n)} < \theta), \quad \theta > 0.$$

The likelihood is also the joint density of $y_1, \ldots, y_n$, so we can write

$$f(y; \theta) = \theta^{-n} I(-\theta < y_{(1)} < y_{(n)} < \theta) \times 2^{-n}, \quad -\theta < y_1, \ldots, y_n < \theta,$$

and the factorisation theorem implies that $(y_{(1)}, y_{(n)})$ is sufficient.

As $y_{(1)} < y_{(n)}$ by construction, we can write

$$I(-\theta < y_{(1)} < y_{(n)} < \theta) = I(-\theta < y_{(1)}, y_{(n)} < \theta) = I(\theta > -y_{(1)}, y_{(n)} < \theta) = I(t < \theta),$$

where $t = \max(y_{(n)}, -y_{(1)})$, so

$$L(\theta) = \theta^{-n} I(t < \theta) \times 2^{-n}, \quad \theta > 0.$$

To see that t is minimal sufficient we consider another dataset of size n with corresponding t' and note that the ratio of likelihoods is

$$\frac{(2\theta)^{-n} I(t' < \theta)}{(2\theta)^{-n} I(t < \theta)}, \quad \theta > 0,$$

which is independent of θ only if $t = t'$.

Solution 3

(a) [5, seen] Slides 55–56.

(b) [2, seen] Example 30.

(c) [3, unseen] In this case the likelihood ratio is

$$\frac{\theta_1^n e^{-\theta_1 s}}{\theta_0 e^{-\theta_0 s}} = \exp\{(\theta_0 - \theta_1)s + n \log(\theta_1/\theta_0)\},$$

where $s = \sum y_j$, so the critical region is of the form

$$\mathcal{Y}_1 = \{(y_1, \ldots, y_n) : \exp\{(\theta_0 - \theta_1)s + n \log(\theta_1/\theta_0)\} > t_\alpha\}.$$

As n and the parameters are known (the hypotheses are simple), this is equivalent to

$$\mathcal{Y}_1 = \{(y_1, \ldots, y_n) : (\theta_0 - \theta_1)s > t'_\alpha\} = \{(y_1, \ldots, y_n) : s > t''_\alpha\},$$

for some t'_α (and t''_α derived from t'_α), because division by the known positive quantity $\theta_0 - \theta_1$ does not change the direction of the inequality.

We want to choose t''_α so that

$$\alpha = \mathrm{P}_0(Y \in \mathcal{Y}_1) = \mathrm{P}_0\left(\sum Y_j > t''_\alpha\right),$$

and as $S \sim \mathrm{Gamma}(n, \theta_0)$ under H_0, we see that t''_α must be the $(1-\alpha)$ quantile of this distribution.

The corresponding calculations with $n = 1$ are also accepted, if correct.

Solution 4

(a) [3, seen] Classical results for the limiting normal distribution of the MLE. Slides 103–109.

(b) [3, unseen] The density function is

$$f(y; \alpha, \beta) = \frac{\alpha \beta^\alpha}{(\beta + y)^{\alpha+1}}, \quad y > 0,$$

so the log likelihood based on a random sample $y_1, \dots, y_n$ is

$$\ell(\alpha,\beta) = n\log\alpha + n\alpha\log\beta - (\alpha+1)\sum_{j=1}^n \log(\beta+y_j) = n\log(\alpha/\beta) - (\alpha+1)\sum_{j=1}^n \log(1+y_j/\beta),$$

which thus gives $S(\beta) = \sum_{j=1}^n \log(1+y_j/\beta)$. Differentiation with respect to α gives

$$\frac{\partial\ell(\alpha,\beta)}{\partial\beta} = \frac{n}{\alpha} - S(\beta), \quad \frac{\partial^2\ell(\alpha,\beta)}{\partial\beta^2} = -\frac{n}{\alpha^2},$$

so $\widehat{\alpha}_\beta = n/S(\beta)$ is the MLE of α for fixed β, and substitution of this into $\ell(\alpha,\beta)$ gives

$$\log(\widehat{\alpha}_\beta,\beta) = n\log\{n/\beta S(\beta)\} - \{n/S(\beta)+1\}S(\beta) \equiv -n\log S(\beta) - n\log\beta - S(\beta),$$

as required; the $\equiv$ means that constants have been dropped.

(c) [4, unseen] The left-hand graph shows strong association between the two parameters, and the right-hand one shows that the profile log likelihood is highly asymmetric. Hence the limiting normal distribution stated in (a) is likely to give poor inferences here. Basing a confidence interval for β on the profile log likelihood, i.e.,

$$\{\beta : 2\{\ell(\widehat{\alpha},\widehat{\beta}) - \ell(\widehat{\alpha}_\beta,\beta) \le \chi_1^2(1-\alpha)\},$$

will take into account the asymmetry and the lower bound on the range of β.

Solution 5

(a) [3, seen] Slides 175–179.

(b) [2, seen] Problem 2 of week 13.

(c) [2, unseen] The posterior density is proportional to

$$\theta^n\exp(-\theta s) \times b^a\theta^{a-1}\exp(-b\theta)/\Gamma(a) \propto \theta^{a+n-1}\exp\{-\theta(b+s)\}, \quad \theta > 0,$$

i.e., it must be the Gamma$(a+n, b+s)$ distribution, where $s = \sum y_j$.

(d) [3, unseen] This is a bit trickier: writing $y \equiv (y_1,\dots,y_n)$ for compactness, the posterior predictive density $f(z \mid y)$ equals

$$\begin{aligned}
\int_0^\infty f(z\mid\theta)\pi(\theta\mid y)\,\mathrm{d}\theta &= \int_0^\infty (x\theta)\exp(-x\theta z) \times \frac{\theta^{a+n-1}(b+s)^{a+n}}{\Gamma(a+n)}\exp\{-\theta(b+s)\}\,\mathrm{d}\theta \\
&= \frac{x(b+s)^{a+n}}{\Gamma(a+n)}\int_0^\infty \theta^{a+n+1-1}\exp\{-\theta(b+s+xz)\}\,\mathrm{d}\theta \\
&= \frac{(a+n)x(b+s)^{a+n}}{(b+s+xz)^{a+n+1}} \\
&= \frac{(a+n)}{c}(1+z/c)^{-(a+n+1)}, \quad z > 0,
\end{aligned}$$

where $c = (b+s)/x$. This is a Lomax density (see question 4).

——————— END OF THE EXAM PAPER ———————