smfsoln2(1617)

SMF SOLUTIONS 2. 2.2.2017

Q1. (i) Consider

$$Q(\lambda) := \int (\lambda f + g)^2 = \lambda^2 \int f^2 + 2\lambda \int fg + \int g^2.$$

This is always non-negative, so its discriminant is ≤ 0 (if it were > 0, there would be two distinct real roots with a sign change between them). So ("$b^2 - 4ac \leq 0$")

$$(\int fg)^2 \leq \int f^2 \int g^2.$$

Equality holds iff Q has a double root, λ_0 say. Then

$$Q(\lambda_0) = \int (\lambda_0 f + g)^2 = 0.$$

This forces

$$\lambda_0 f + g = 0$$

(a.e.), the required linear relation.
(ii) We apply this to $X - E[X]$, $Y - E[Y]$, we find equality in CR iff a linear relation of the form $a(X - E[X]) + b(Y - E[Y]) = 0$ holds. Taking $X = \ell'$ (recall $E[\ell'] = 0$), $Y = u$ (recall $EY = \theta$ as u is unbiased for θ), we find

$$u - \theta = a\ell'$$

for some a (so $b = \theta$).
(iii) So

$$\ell' = u/a(\theta) - \theta/a(\theta) : \qquad \ell = u \int d\theta/a(\theta) - \int \theta d\theta/a(\theta) + k(\mathbf{X}) :$$

$$L(\theta; \mathbf{X}) = \exp\{\alpha(\theta)u(\mathbf{X}) + \beta(\theta) + k(\mathbf{X})\}.$$

(iv) By Fisher-Neyman, this shows u is sufficient for θ; by Lehmann-Scheffé, this shows u is minimal sufficient for θ.

Q2.

$$\ell = -n \log 2 - \sum |x_i - \theta|.$$

To maximise this – i.e. minimise $\sum |x_i - \theta|$ – draw a graph. From this, the sum is minimised by $\theta = Med$, and increases linearly on either side of the sample median. So the MLE is $\hat{\mu} = Med$.

(ii) With one reading, as above, ℓ decreases with slope -1 to the right of Med, slope +1 to the left of Med. So $(\ell')^2 = 1$ (except at $\lambda = Med$, where the derivative is not defined – but we are going to integrate, and so can neglect null sets, let alone single points, so this does not matter). So $I = \int (\partial \log f / \partial \theta)^2 f = \int f = 1$, as f is a density. So the CR bound is $1/n$.

We are given that Med is asymptotically normal, and that its mean is $med = \theta$, so Med is asymptotically unbiased. By symmetry, the population median is $med = \theta$, where the density is $\frac{1}{2}$. So $4f(med)^2 = 1$, and the asymptotic variance of the sample median is $1/n$, the CR bound, so Med is also asymptotically efficient.

Q3. (i)

$$f(x;\mu) = \frac{1}{\pi(1 + (x-\mu)^2)}, \qquad \ell = \log f = c - \log[1 + (x-\mu)^2],$$

$$\ell' = \frac{2(x-\mu)}{1 + (x-\mu)^2}, \qquad \ell'(\mathbf{x};\theta) = 2\sum_1^n \frac{(x_i - \mu)}{1 + (x_i - \mu)^2}.$$

But (Q1) we have efficiency iff ℓ' factorises in the form $\ell'(\mathbf{x};\theta) = A(\theta)(u(\mathbf{x}) - \theta)$. The likelihood here does not factorise, so there is no efficient estimator.

(ii) The information per reading is

$$E[(\ell')^2] = \int (\partial f/\partial \mu)^2 f = \frac{4}{\pi}\int \frac{(x-\mu)^2}{[1 + (x-\mu)^2]^3} dx = \frac{4}{\pi}\int \frac{x^2}{[1+x^2]^3} dx = \frac{4}{\pi} I,$$

say. One can evaluate I by Complex Analysis ($f(z) := z^2/[1+z^2]^3$, round the contour Γ – semicircle in the upper half-plane on base $[-R, R]$; f has a triple pole inside Γ of residue $-i/16$, so $I = 2\pi i \ Res = \pi/8$), giving the information per reading as $\frac{1}{2}$. So the information in a sample of size n is $n/2$, and the MLE has asymptotic variance $2/n$. As in Q2, the sample median has asymptotic variance $\pi^2/4n$. So the asymptotic efficiency is their ratio, $8/\pi^2 \sim 81\%$. NHB