SB2.1 Foundations of Statistical Inference

Sheet 2 — MT22

Section A

1. The number of phone calls a man receives in a week follows a Poisson distribution with mean θ. At the start of week 1, the man's beliefs about the value of θ is reprsented by the gamma distribution
$$\pi(\theta) = \frac{1}{54}\theta^2 e^{-\theta/3},\ \theta > 0.$$
In the 4 weeks following the start of week 1, the man received 3, 7, 6, and 10 phone calls, respectively. Determine the posterior distribution of θ and the predictive distribution for the number of calls that he will receive in week 5.

2. In a model $X \sim f(x, \theta)$, suppose T is complete sufficient statistic for θ. Show that if a minimal sufficient statistic S for θ exists, then T is also minimal sufficient.

George Deligiannidis: deligian@stats.ox.ac.uk

Section B

3. Let $X = (X_1, \ldots, X_n)$ be a random sample from a density $f_X(x;\theta)$ belonging to a parametric family $\mathcal{F}$. Let $T = t(X)$ be a function of X and denote the density of T by $f_T(t;\theta)$. Assuming statistical regularity, let $i_X(\theta)$ to be the Fisher information about θ in X. Finally, let $i_{X|t}(\theta)$ denote the Fisher information conditional on $T = t$, i.e.

$$i_{X|t}(\theta) := -\int f_{X|t}(x \mid t;\theta)\mathrm{d}x \left[-\frac{\partial^2}{\partial\theta^2}\log f_{X|t}(x \mid t;\theta)\right],$$

and let

$$i_{X|T}(\theta) = \int i_{X|t}(\theta) f_T(t;\theta)dt$$

(a) Show that

$$i_X(\theta) = i_{X|T}(\theta) + i_T(\theta)$$

(b) Show that

$$i_X(\theta) \geq i_T(\theta),$$

with equality for all θ if and only if $T = t(X)$ is sufficient for θ.

Hint: Use the factorization theorem for the density

$$f_X(x;\theta) = f_{X|T}(x \mid t;\theta) f_T(t;\theta).$$

(c) Hence, or otherwise, determine the Fisher information about θ in the first r order statistics

$$X_{(1)} < X_{(2)} < \cdots < X_{(r)}$$

of a sample of size n from the density

$$f(x;\theta) = \theta\exp(-\theta x),\ x > 0$$

4. In order to measure the intensity (on a logarithmic scale so that it can take negative values), θ, of a source of radiation in a noisy environment, a measurement X_1 is taken without the source present and a second, independent measurement X_2 is taken with the source present. It is known that X_1 is $N(\mu, 1)$ and X_2 is $N(\mu + \theta, 1)$, where μ is the mean noise level. The prior distribution for μ is $N(\mu_0, 1)$ while the prior for θ is constant (and thus improper).

(a) Write down the joint posterior distribution of μ and θ up to a constant of proportionality.

(b) Hence obtain the posterior marginal distribution of θ.

(c) The usual estimate of θ is $x_2 - x_1$; explain why $\frac{1}{2}(2x_2 - x_1 - \mu_0)$ might be better.

George Deligiannidis: deligian@stats.ox.ac.uk

5. A non-negative function $f(\theta)$ of the parameter is said to be *proper* if its integral is finite

$$\int f(\theta)d\theta < \infty$$

and can thus be normalized to be a probability distribution. Otherwise it is said to be *improper*.

Let $X_1, \ldots, X_n$ be a random sample from a normal distribution with unknown mean μ and unknown variance λ^{-1}, with the improper prior

$$\pi_0(\mu, \lambda) = \lambda^{-1},\ \lambda > 0, -\infty < \mu < \infty.$$

(a) Find the joint posterior density $\pi(\mu, \lambda|x)$ of (μ, λ).

(b) Find the marginal posterior density $\pi(\mu|x)$ of μ.

(c) Show that the joint posterior density in (a) is proper.

George Deligiannidis: deligian@stats.ox.ac.uk

6. *[This question is from a past paper and has three parts]* Let $E(a,b)$ be the distribution of the shifted exponential with density

$$\frac{1}{b}e^{-(x-a)/b}, \qquad x > a$$

where $a \in \mathbb{R}, b > 0$ are parameters. Let $X_1, \dots, X_n$ be a random sample from the distribution $E(a,b)$.

(a) In this part suppose that a is fixed and known.

(i) Let T be a statistic from the random sample. Write down the definition of *sufficiency, minimality and completeness* of T for a parameter θ.

(ii) Show that this is a one parameter exponential family. Without calculation, give a sufficient and complete statistic for b.

(iii) Find the minimal variance unbiased estimator (MVUE) $\hat{b}$ of b. (hint: remember that if $G \sim \Gamma(u,v)$ then $E(G) = u/v$) and that a sum of independent exponential variables with same means is Gamma distributed)

(iv) Compute the Fisher information for b and compare to the variance of $\hat{b}$. Does $\hat{b}$ attains the Cramer-Rao lower bound? How could you have inferred this directly from the log-likelihood ℓ? (Hint: if $G \sim \Gamma(u,v)$ then $Var(G) = u/v^2$).

(b) Now we suppose that it is b which is known and a that we wish to estimate.

(i) Can we still say that the distribution $E(a,b)$ with b known is an exponential family? Briefly justify your answer.

(ii) Find the distribution of $X_{(1)} = \min_{i=1,\dots,n} X_i$.

(iii) Show that $X_{(1)}$ is sufficient and complete.

(iv) Hence deduce the MVUE of a.

(c) Suppose now that both a and b are unknown. Define $T_1(X) = X_{(1)}$ and $T_2(X) = \sum_{i=1}^{n}(X_i - X_{(1)})$.

(i) Fix $i \in \{1, \dots, n\}$. What is the conditional distribution of $X_i - X_{(1)}$ given $X_{(1)}$.

(ii) Show that $T_2(X)$ has a Gamma distribution and give its parameters.

(iii) Show that (T_1, T_2) is a sufficient statistic for (a,b). Assuming that is also complete, obtain an MVUE for (a,b) (Hint: you can assume without proof that T_1 and T_2 are independent).

[It is an interesting (but non-trivial) problem to show that (T_1, T_2) is complete. In particular, it involves some measurability issues. You are not required to do so.]

George Deligiannidis: deligian@stats.ox.ac.uk

Section C

7. From 2015.

   (a) Let $X = (X_1, \dots, X_n)$ be a sample from a continuous distribution with probability density function $f(x;\theta)$ where θ is an unknown parameter. Let $L(\theta; X)$ be the likelihood function and $\ell(\theta) = \log L(\theta; X)$

   (i) Prove that $\mathbb{E}[\frac{\partial \ell}{\partial \theta}] = 0$.

   (ii) If we define $I_\theta = -\mathbb{E}[\frac{\partial^2 \ell}{\partial \theta^2}]$ show that $I_\theta = \mathbb{E}[(\frac{\partial \ell}{\partial \theta})^2] = \mathrm{Var}(\frac{\partial \ell}{\partial \theta})$.

   (iii) Show that the variance of an unbiased estimator of θ, denoted $\hat{\theta}(X)$, satisfies the Cramér-Rao inequality

$$\mathrm{Var}\big[\hat{\theta}(X)\big] \geq I_\theta^{-1},$$

   and provide a brief statement about why regularity conditions are needed for this result to hold.

   (iv) Show that there exists an unbiased estimator $\widehat{\theta}$ which attains the Cramér-Rao lower bound (under regularity conditions) if and only if

$$\frac{\partial \ell}{\partial \theta} = I_\theta(\widehat{\theta} - \theta)$$

   (v) If $\hat{g}(X)$ is an unbiased estimator of the function $g(\theta)$ derive a lower bound for the variance of the estimator.

   (b) Suppose that

$$f(x;\theta) = \frac{\theta^3 x(x+1)}{\theta + 2} e^{-\theta x}, \quad x \geq 0, \quad \theta > 0.$$

   Find an unbiased estimator of $(\theta^2 - 6)/[\theta(\theta+2)]$ whose variance attains the Cramér-Rao lower bound.

George Deligiannidis: deligian@stats.ox.ac.uk