Statistics 582, Problem Set 6, Solutions

Wellner; 2/20/98

1. Suppose that $X \sim P_\theta$ for $\theta \in \Theta \subset R^k$ has well-defined Fisher information matrix $I(\theta)$ for θ. The *Jeffreys prior* distribution Λ_J has density $\lambda_J(\theta) = \det(I(\theta))^{1/2}$ with respect to Lebesgue measure on Θ. Note that Λ_J may not be a finite measure, and even if Λ_J is a finite measure, it may not have total mass 1. If a prior distribution is a finite measure, then call it a *proper prior distribution*, and correspondingly if it is not a finite measure, call it an *improper prior distribution*. If the resulting posterior distribution is a finite measure, call it a *proper posterior distribution*, and (by convention) normalize it to have total mass 1.

   A. Suppose that $X \sim \text{Bernoulli}(\theta)$. Find the Jeffrey's prior density λ_J for θ. Is Λ_J a finite measure? If it is finite, what is $\Lambda_J((0,1))$? Find the corresponding posterior distribution of Θ starting with the Jeffrey's prior.

   Solution: When $X \sim \text{Bernoulli}(\theta)$, the Information for θ is $I(\theta) = \{\theta(1-\theta)\}^{-1}$, so the Jeffrey's prior for θ is has density

$$\lambda_J(\theta) = \frac{1}{\sqrt{\theta(1-\theta)}}.$$

   This density is proportional to the Beta(1/2, 1/2) density

$$\lambda(\theta) = \frac{\Gamma(1)}{\Gamma(1/2)\Gamma(1/2)} \theta^{1/2-1}(1-\theta)^{1/2-1} 1_{(0,1)}(\theta)$$

   (which is also known as the "arcsin" distribution because the corresponding distribution function is $\Lambda(\theta) = (2/\pi)\arcsin(\sqrt{\theta})$; it arises naturally as the limiting distribution of the proportion of time a random walk stays positive). Thus $\lambda_J((0,1)) = \Gamma(1/2)^2 = \pi$. The corresponding posterior distribution is proportional to

$$\theta^{x-1/2}(1-\theta)^{1-x-1/2} = \theta^{(x+1/2)-1}(1-\theta)^{(3/2-x)-1};$$

   i.e. the posterior density is Beta$(x+1/2, 3/2-x)$ if $X = x$ is observed.

   B. Suppose that $X \sim \text{Geometric}(\theta)$, i.e. the number of trials until the first success in i.i.d. Bernoulli trials with probability θ of success for each trial – recall Chapter 1, section 1. Find the Jeffrey's prior density λ_J for θ. Is Λ_J a finite measure? If it is finite, what is $\Lambda_J((0,1))$? Find the corresponding posterior distribution of Θ starting with the Jeffrey's prior. If we observe $X_1, \ldots, X_n$

i.i.d. Geometric(θ), so that $\sum X_i \sim$ Negative Binomial(n, θ) is the posterior distribution "proper" for some n?

Solution: When $X \sim$ Geometric(θ) with $P_\theta(X = x) = (1-\theta)^{x-1}\theta$ for $x = 1, 2, \ldots$, the score function is

$$\dot{l}_\theta(x) = \frac{1}{1-\theta}\left(\frac{1}{\theta} - X\right)$$

and the information for θ is

$$I(\theta) = E_\theta(\dot{l}_\theta(X))^2 = -E_\theta \ddot{l}_{\theta\theta}(X) = \frac{1}{\theta^2(1-\theta)}.$$

Hence the Jeffrey's prior for θ has density

$$\lambda_J(\theta) = \frac{1}{\theta\sqrt{(1-\theta)}}.$$

Here we have

$$\int_0^1 \lambda_J(\theta)d\theta == \int_0^1 \frac{1}{\theta\sqrt{(1-\theta)}}d\theta = \infty$$

because the integral diverges at 0. Hence this density does not correspond to a finite measure. It corresponds in some sense to a Beta($0, 1/2$) density; but recall that the Beta(α, β) densities are defined for $\alpha, \beta > 0$. In spite of this the posterior density is proportional to

$$\theta^0(1-\theta)^{x-1-1/2} = (1-\theta)^{x-1/2-1}$$

which corresponds to Beta($1, x - 1/2$) if $X = x$ is observed. This is completely proper for any $x \in \{1, 2, \ldots\}$ since $x - 1/2 > 0$.

C. Suppose that $X \sim$ Weibull(θ) with $\theta = (\alpha, \beta) \in (0, \infty) \times (0, \infty)$ as in chapters 3 and 4. Find the Jeffrey's prior density λ_J for θ. Is Λ_J a finite measure? If it is finite, what is $\Lambda_J((0, \infty)^2)$? Find the corresponding posterior distribution of Θ starting with the Jeffrey's prior.

Solution: From Example 3.2.7 the information matrix for the Weibull density is

$$I(\theta) = \begin{pmatrix} \frac{\beta^2}{\alpha^2} & \frac{a}{\alpha} \\ \frac{a}{\alpha} & \frac{b^2}{\beta^2} \end{pmatrix}$$

where $a = -(1-\gamma)$ and $b^2 = \pi^2/6 + (1-\gamma)^2$. Therefore the Jeffrey's prior is given by

$$\lambda_J(\theta) = \det(I(\theta))^{1/2} = \frac{\pi/\sqrt{6}}{\alpha}.$$

This is not the density of a finite measure:

$$\int_0^\infty \int_0^\infty \lambda_J(\alpha, \beta) d\alpha d\beta = \frac{\pi}{\sqrt{6}} \int_0^\infty \int_0^\infty \frac{1}{\alpha} d\alpha d\beta = \infty \, .$$

The posterior density is proportional to

$$\begin{aligned} p_\theta(x)\lambda_J(\theta) &= \frac{\beta}{\alpha}\left(\frac{x}{\alpha}\right)^{\beta-1} \exp(-(x/\alpha)^\beta)\frac{\pi/\sqrt{6}}{\alpha} 1_{(0,\infty)}(\alpha)1_{(0,\infty)}(\beta) \\ &= \frac{\pi}{\sqrt{6}}\frac{\beta}{\alpha^2}\left(\frac{x}{\alpha}\right)^{\beta-1} \exp(-(x/\alpha)^\beta) 1_{(0,\infty)}(\alpha)1_{(0,\infty)}(\beta) \, . \end{aligned}$$

When we try to compute the normalizing constant to form a density (namely the marginal density $p(x)$), we find, however, that (using the change of variables $u = \alpha^{-\beta}$ in the second line so that $du = -\beta\alpha^{-\beta-1}d\alpha$)

$$\begin{aligned} p(x) &= \frac{\pi}{\sqrt{6}} \int_0^\infty \int_0^\infty \frac{\beta}{\alpha^2}\left(\frac{x}{\alpha}\right)^{\beta-1} \exp(-(x/\alpha)^\beta) d\alpha d\beta \\ &= \frac{\pi}{\sqrt{6}} \int_0^\infty \frac{x^{\beta-1}}{x^\beta} \int_0^\infty x^\beta \exp(-x^\beta u) du d\beta \\ &= \frac{\pi}{\sqrt{6}} \int_0^\infty \frac{1}{x} d\beta = \infty \, . \end{aligned}$$

Hence the posterior (for one observation) is not proper for any x.

For n observations we compute

$$p_\theta(\underline{x})\lambda_J(\theta) = \frac{\beta^n}{\alpha^n}\frac{(\prod x_i)^{\beta-1}}{\alpha^{n(\beta-1)}} \exp(-(\sum x_i^\beta/\alpha^\beta)\frac{\pi/\sqrt{6}}{\alpha} 1_{(0,\infty)}(\alpha)1_{(0,\infty)}(\beta) \, .$$

Therefore, using the same change of variables as before, $u = \alpha^{-\beta}$,

$$\begin{aligned} p(\underline{x}) &= \frac{\pi}{\sqrt{6}} \int_0^\infty \beta^n (\prod x_i)^{\beta-1} \int_0^\infty \frac{1}{\alpha^{n+1}} \frac{1}{\alpha^{n(\beta-1)}} \exp(-\sum x_i/\alpha^\beta) d\alpha d\beta \\ &= \frac{\pi}{\sqrt{6}} \int_0^\infty \beta^{n-1} (\prod x_i)^{\beta-1} \int_0^\infty u^{n-1} \exp(-\sum x_i^\beta u) du d\beta \\ &= \frac{\pi}{\sqrt{6}} \Gamma(n) \int_0^\infty \beta^{n-1} \frac{(\prod x_i)^{\beta-1}}{(\sum x_i^\beta)^n} d\beta \end{aligned}$$

This converges or diverges depending on the x_i's: when $x_i = c$ for all i it is easily seen to diverge; when $x_i = ci$ for $i = 1, \ldots, n$, it converges. I suspect is converges when $0 < x_{(1)} < \ldots < x_{(n)} < \infty$, but I do not yet have a proof.

2. **Optional bonus problem:** A. Suppose that $X \sim F$, and let $m = F^{-1}(1/2)$, $\mu = E(X)$, $\sigma^2 = Var(X)$, and we assume that $E(X^2) < \infty$. Show that

$$|m - \mu| \le \sqrt{2\sigma^2}\,.$$

Hint: use Chebychev's inequality.

Solution: Chebychev's inequality says that

$$P(|X-\mu| > t) \le \frac{Var(X)}{t^2} \equiv \frac{\sigma^2}{t^2}$$

and this implies

$$P(|X-\mu| > \sqrt{2}\sigma) \le \frac{1}{2}\,.$$

or, taking complements, that

$$P(\mu - \sqrt{2}\sigma \le X \le \mu + \sqrt{2}\sigma) > \frac{1}{2}\,.$$

This implies that

$$P(X \le \mu + \sqrt{2}\sigma) > \frac{1}{2} \qquad \text{and} \qquad P(X \ge \mu - \sqrt{2}\sigma) > \frac{1}{2}\,,$$

and these in turn imply that

$$\mu - \sqrt{2}\sigma \le m \le \mu + \sqrt{2}\sigma\,.$$

B. Let m and μ be the median and mean of F as in A, and let M be the *mode* of F, assuming that it is well-defined: i.e. we suppose that F has density f which is strictly increasing to the left of M and strictly decreasing to the right of M. Show that $\mu \le m \le M$ if there is an x_0 such that

$$f(m+x) - f(m-x) \begin{cases} \ge 0 & \text{for} \quad 0 \le x < x_0 \\ \le 0 & \text{for} \quad x_0 < x < \infty \end{cases} \,. \tag{0.1}$$

If the inequalities in (0.1) are reversed, then $M \le m \le \mu$. Hint: show that

$$m - \mu = \int_0^\infty \{F(m-x) + F(m+x) - 1\}dx\,.$$

C. Examine (0.1) and the conclusion for the distributions Gamma$(r,1)$ with $r = 1, 2$.