Homework 3

PkGu

10/18/2021

P[66] 12

(1). Location parameters family, with Jeffrey prior $\pi(\theta) \propto 1$.

(2). Scale parameters family, with Jeffrey prior $\pi(\beta) \propto 1/\beta$.

(3). Either if one of $\{\mu, \sigma\}$ fixed, with Jeffrey prior $\pi(\mu, \sigma) \propto 1/\sigma$.

(4). Scale parameters family, with Jeffrey prior $\pi(x_0) \propto 1/x_0$.

P[66] 13

(1). By computing

```
Sum[
  -D[Log[Exp[-w]*w^i/i!],{w,2}]*Exp[-w]*w^i/i!,
{i,0,Infinity}]
```

the Fisher information for Poisson distribution $P(\lambda)$ equals to

$$I(\lambda) = 1/\lambda$$

And the Jeffrey prior distribution is

$$\pi(\lambda) \propto \lambda^{-\frac{1}{2}}$$

(2). By computing

```
Sum[
  -D[Log[Binomial[n+r-1,r-1]* w^r*(1-w)^n],{w,2}]*Binomial[n+r-1,r-1]*w^r*(1-w)^n,
{n,0,Infinity}]
```

the Fisher information for distribution $NBN(r, \theta)$ equals to

$$I(\theta) = \frac{r}{(1-\theta)\cdot\theta^2}$$

And the Jeffrey prior distribution is

$$\pi(\theta) \propto \frac{1}{\theta\sqrt{1-\theta}}$$

(3). By computing

```
Integrate[
  -D[Log[Exp[-x/w]/w],{w,2}]*Exp[-x/w]/w,
{x,0,Infinity}]
```

the Fisher information for distribution $Exp(\frac{1}{\lambda})$ equals to

$$I(\lambda) = \frac{1}{\lambda^2}$$

And the Jeffrey prior distribution is

$$\pi(\lambda) \propto \frac{1}{\lambda}$$

(4). By computing

```
Integrate[
  -D[Log[w^a*Exp[-w*x]*x^(a-1)/Gamma[a]],{w,2}]*w^a*Exp[-w*x]*x^(a-1)/Gamma[a],
{x,0,Infinity}]
```

the Fisher information for distribution $\Gamma(\alpha, \lambda)$ where α is already known equals to

$$I(\lambda) = \frac{\alpha}{\lambda^2}$$

And the Jeffrey prior distribution is

$$\pi(\lambda) \propto \frac{1}{\lambda}$$

(5). The likelihood function of the multinomial model is

$$p((x_1, ..., x_{k-1})|(p_1, ..., p_{k-1})) = A_{n,\mathbf{x}} \cdot (\prod_1^{k-1} {p_i}^{x_i}) \cdot (1 - \sum_1^{k-1} p_i)^{n - \sum_1^{k-1} x_i}$$

the log likelihood

$$log\ p(\mathbf{x}|\mathbf{p}) = C(\mathbf{x}) + \sum_1^{k-1} x_i log(p_i) + (n - \sum_1^{k-1} x_i) log(1 - \sum_1^{k-1} p_i)$$

and for $i \neq j$,

$$-D_{ij}(l_{\mathbf{x}}\mathbf{p}) = \frac{n - \sum_1^{k-1} x_i}{(1 - \sum_1^{k-1} p_i)^2}$$

whereas for i,

$$-D_{ii}(l_{\mathbf{x}}\mathbf{p}) = \frac{n - \sum_1^{k-1} x_i}{(1 - \sum_1^{k-1} p_i)^2} + \frac{x_i}{{p_i}^2}$$

By taking expectation over $\mathbf{X}_{k-1}$, we have for $i \neq j$,

$$
\begin{aligned}
&\mathbb{E}_{\mathbf{p}}\left[\frac{n-\sum_1^{k-1} x_i}{(1-\sum_1^{k-1} p_i)^2}\right] \\
&= \sum_{x_1+\ldots+x_{k-1}\leq n} \left(A_{n,\mathbf{x}} \cdot (\prod_1^{k-1} {p_i}^{x_i}) \cdot (1-\sum_1^{k-1} p_i)^{n-\sum_1^{k-1} x_i} \cdot \frac{n-\sum_1^{k-1} x_i}{(1-\sum_1^{k-1} p_i)^2} \right) \\
&= \sum_{x_1+\ldots+x_{k-1}\leq n-1} \left(A_{n-1,\mathbf{x}} \cdot (\prod_1^{k-1} {p_i}^{x_i}) \cdot (1-\sum_1^{k-1} p_i)^{n-1-\sum_1^{k-1} x_i} \cdot \frac{n}{1-\sum_1^{k-1} p_i} \right) \\
&= \frac{n}{1-\sum_1^{k-1} p_i}
\end{aligned}
$$

$i.e.$ for $i \neq j$,

$$
I_{ij}(\mathbf{p}) = \frac{n}{1-\sum_1^{k-1} p_i}
$$

and similarly, for each i,

$$
I_{ii}(\mathbf{p}) = \frac{n}{1-\sum_1^{k-1} p_i} + \frac{n}{p_i}
$$

Hence the Jeffrey prior distribution is

$$
\pi(\mathbf{p}) \propto \sqrt{|det(I(\mathbf{p}))|} \propto \prod_1^{k-1} {p_i}^{-\frac{1}{2}}
$$

P[66] 20

For prior distribution

$$
\pi(\theta) \propto e^{k_1 a(\theta)+k_2 c(\theta)}
$$

with an exponential family of single parameter

$$
p(x|\theta) = e^{a(\theta)b(x)+c(\theta)+d(x)}
$$

The posterior is

$$
\pi(\theta|x) \propto e^{a(\theta)(b(x)+k_1)+c(\theta)(1+k_2)+d(x)}
$$

i.e. for some k_3,

$$
\pi(\theta|x) = e^{a(\theta)(b(x)+k_1)+c(\theta)(1+k_2)+(d(x)+k_3)}
$$

P[67] 24

To maximize

$$
\begin{aligned}
KL(\pi;\pi_0) &= \int_a^b log\frac{d\Pi_0}{d\Pi}\ d\Pi \\
&= \int_a^b -log(\theta f(\theta))f(\theta)\ d\theta \\
&= (\int_a^z + \int_z^b) - log(\theta f(\theta))f(\theta)\ d\theta
\end{aligned}
$$

given that $F(z)=\frac{1}{2}$.

If F is optimized in (a,z), by variation inside $\mathcal{G}_1 = \{g| \int_a^z g\ d\theta = 0\}$

$$
\begin{aligned}
&\frac{\partial}{\partial\epsilon}\bigg|_{\epsilon=0} \int_a^z -log(\theta(f+\epsilon g))\cdot(f+\epsilon g)\ d\theta \\
&= -\int_a^z [log(\theta)+1+log(f)]\cdot g\ d\theta \\
&= 0, \quad \forall g\in\mathcal{G}_1
\end{aligned}
$$

Hence $(log(\theta)+1+log(f))$ is constant, *i.e.* $f=\frac{A_1}{\theta}$. Recall that $F(z)=\frac{1}{2}$, we have

$$
f(\theta) = \frac{1}{\theta}(2ln\frac{z}{a})^{-1}, \quad \theta\in(a,z)
$$

Similarly,

$$
f(\theta) = \frac{1}{\theta}(2ln\frac{b}{z})^{-1}, \quad \theta\in(z,b)
$$

P[67] 25

Since the reference prior is independent with the choice of coordinates, we have for coordinate $\{\mu_1,\mu_2\}$ and $\{\phi_1,\phi_2\}$,

$$
\begin{aligned}
\pi(\phi_1,\phi_2) &= \pi(\mu_1,\mu_2)\cdot|det(T_{\mu\to\phi})|^{-1} \\
&\propto \pi(\mu_1,\mu_2)\cdot\frac{\mu_1}{\mu_2}
\end{aligned}
$$

And then consider the reference prior $\pi(\mu_1,\mu_2)$, which is equivalent to the Jeffrey prior. (Bernardo,1979)

$$
\pi(\mu_1,\mu_2)\propto\frac{1}{\mu_1\mu_2}
$$

Hence

$$
\pi(\phi_1,\phi_2)\propto\frac{1}{{\mu_2}^2} = (\phi_1\phi_2)^{-1}
$$