Series 3

1. **Jeffreys prior**

   a. Consider the normal model $X \sim \mathcal{N}(0, \sigma^{-2})$ where σ^{-2} is the precision (inverse of the variance). Show that Jeffreys prior for σ^{-2} is proportional to σ^2.

   b. Consider the normal model $X \sim \mathcal{N}(\mu, \sigma)$ with both the mean and the standard deviation unknown. Show that Jeffreys prior for (μ, σ) is proportional to σ^{-2}.

   c. Consider the vector normal means example we saw in the lecture: X_i $1 \leq i \leq 2n$ are independent and normally distributed with variance σ^2 and means $\mathbb{E}(X_{2k-1}) = \mathbb{E}(X_{2k}) = \mu_k$. Show that Jeffreys prior for this model is $\pi(\mu_1, \ldots, \mu_n, \sigma^2) \propto \sigma^{-n-2}$.

   d. Consider the model $X \sim \text{binomial}(n, \theta)$. Show that Jeffreys prior for θ is the Beta($\frac{1}{2}, \frac{1}{2}$) distribution.

(Partial) Solution

d. We have

$$
\begin{aligned}
f(x|\theta) &= \binom{n}{x} \theta^x (1-\theta)^{n-x} \\
\log f(x|\theta) &= \log\left(\binom{n}{x}\right) + x \log(\theta) + (n-x)\log(1-\theta) \\
\frac{\partial^2}{\partial \theta^2} \log f(x|\theta) &= -\frac{x}{\theta^2} - \frac{n-x}{(1-\theta)^2} \\
I(\theta) &= \frac{n}{(1-\theta)\theta}
\end{aligned}
$$

2. **Jeffreys prior and variance stabilizing transformations**

   Find a transform of θ, $\tau = g(\theta)$, such that the Fisher information $I_\tau(\tau)$ is constant for:

   a. The Poisson distribution, $X \sim \text{Poisson}(\theta)$.

   b. The gamma distribution, $X \sim \text{Gamma}(\gamma, \theta)$, with known shape parameter γ.

   c. The binomial distribution, $X \sim \text{Binomial}(n, \theta)$.

   Remark: In frequentist statistics, such a transformation is called a variance stabilizing transformation since for the transformed parameter, the asymptotic variance of the maximum likelihood estimator does not depend on the paramter.

Solution

The following holds true

$$I_\tau(\tau) = I_\theta(g^{-1}(\tau))(g^{-1})'(\tau)^2 = I_\theta(\theta)g'(\theta)^{-2}.$$

This means that if we choose

$$g'(\theta) \propto I_\theta^{0.5}$$

then I_τ is constant. This is achieved if we use

$$g(\theta) \propto \int^\theta I_\theta(u)^{0.5} du.$$

a. We have

$$\begin{aligned} f(x|\theta) &= \frac{\theta^x}{x!} e^{-\theta} \\ \log f(x|\theta) &= -\theta + x \log \theta - \log x! \\ \frac{\partial^2}{\partial \theta^2} \log f(x|\theta) &= -\frac{x}{\theta^2} \\ I(\theta) &= \frac{1}{\theta} \end{aligned}$$

I.e., we choose

$$\begin{aligned} g(\theta) &\propto \int^\theta I_\theta(u)^{0.5} du \\ &\propto \theta^{0.5}. \end{aligned}$$

b. We have

$$\begin{aligned} f(x|\theta) &= \frac{\theta^\gamma}{\Gamma(\gamma)} x^{\gamma-1} e^{-\theta x} \\ \log f(x|\theta) &= \gamma \log(\theta) - \log(\Gamma(\gamma)) + (\gamma - 1)x - \theta x \\ \frac{\partial^2}{\partial \theta^2} \log f(x|\theta) &= -\frac{\gamma}{\theta^2} \\ I(\theta) &= \frac{\gamma}{\theta^2} \end{aligned}$$

I.e., we choose

$$\begin{aligned} g(\theta) &\propto \int^\theta I_\theta(u)^{0.5} du \\ &\propto \log(\theta). \end{aligned}$$

c. We know that

$$I(\theta) = \frac{n}{(1-\theta)\theta}.$$

This means that $I_\theta(u)^{0.5}$ is the density of the Beta($\frac{1}{2}, \frac{1}{2}$) distribution. I.e., we choose for $g(\theta)$ the cummulative distribution function of the Beta($\frac{1}{2}, \frac{1}{2}$) distribution.

3. **Refference prior**

In the lecture, we have seen that the reference prior defined using an asymptotic argument leads to the Jeffreys prior, which is equivariant under reparametrizations.

Show directly that the reference prior is also equivariant under reparametrization when defining it as the maximizer of the mutual information

$$I(X,\theta) = \int_X f(x) \int_\Theta \pi(\theta \mid x) \log \frac{\pi(\theta \mid x)}{\pi(\theta)} d\theta dx.$$

Solution

If we reparametrize $\tau = g(\theta)$, where g is invertible and differentiable, and apply the change-of-variables formula to the prior, we otain

$$\pi(\tau) = \pi(g^{-1}(\tau))|\det Dg^{-1}(\tau)|$$

and also

$$\pi(\tau|x) = \pi(g^{-1}(\tau)|x)|\det Dg^{-1}(\tau)|.$$

Since in the mutual information the Jacobian terms in the numerator and denominator of the logarithm cancel, we obtain

$$I(X,\tau) = \int_X f(x) \int_\Theta \pi(g^{-1}(\tau) \mid x)|\det Dg^{-1}(\tau)| \log \frac{\pi(g^{-1}(\tau) \mid x)}{\pi(g^{-1}(\tau))} d\theta dx,$$

which is exactly what we obtain when we apply a change-of-variables from θ to τ in the inner integral in the mutual information.

This means that first determining the reference prior and then changing the variable from θ to τ is equivalent to first changing the variable and then determining the reference prior. I.e., the reference prior is equivariant under reparametrizations.

4. **Expert priors**[1]

[1] Based on exercise 3 in Section 6.8 of Held and Sabanes Bove (2014).

Suppose that the heights of male students are normally distributed with mean 180 and unknown variance σ^2. We believe that σ^2 is in the range $[22, 41]$ with approximately 95% probability. Thus, we assign an inverse-gamma distribution IG$(\alpha = 38, \beta = 1110)$ as prior distribution for σ^2.

a. Verify with R that the above choice of parameters for the inverse-gamma distribution leads to a prior probability of approximately 95 % that $\sigma^2 \in [22, 41]$.
b. Derive and plot the posterior density of σ^2 when using the following data:
183, 173, 181, 170, 176, 180, 187, 176, 171, 190, 184, 173, 176, 179,181, 186.
c. Compute the posterior density of the standard deviation σ.

Hint:

- The R package `invgamma` implements the density and cummulative distribution function of the inverse-gamma distribution.

Solution

a. We can easily verify that $P(\sigma^2 \in [22, 41]) \approx 0.95$.

```
library(invgamma)
pinvgamma(41,shape=38,rate=1100)-pinvgamma(22,shape=38,rate=1100)

## [1] 0.9417763

##Note that the 'rate' parameter in the package corresponds to
## the parameter that is usually called 'scale' parameter.
```

b. For the normal likelihood, one can easily show that if μ is fixed and known and σ^2 has an inverse-gamma prior IG(α, β), the posterior for σ^2 is again an inverse-gamma distribution

$$IG\left(\alpha + \frac{n}{2}, \beta + \frac{1}{2}\sum_{i=1}^{n}(x_i - \mu)\right).$$

In our example, we obtain the following values.

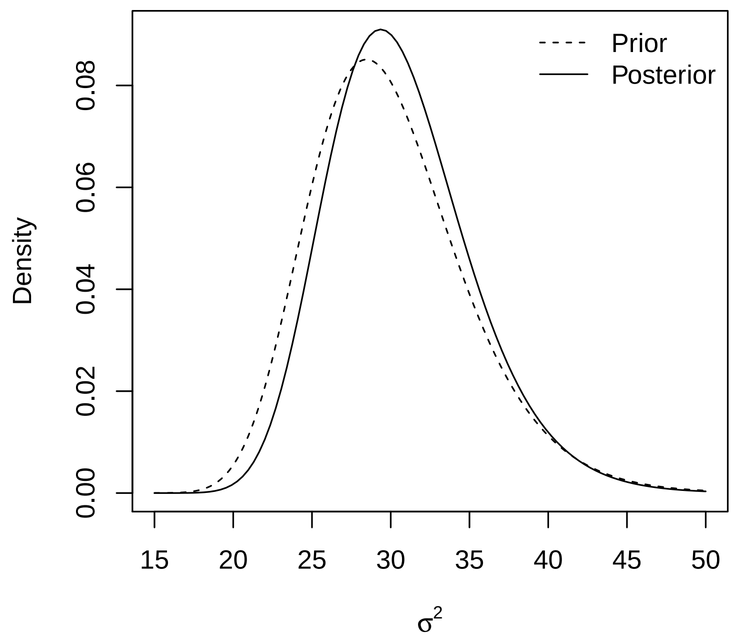
```
alpha=38
beta=1110
mu=180
heights=c(183,173,181,170,176,180,187,176,
          171,190,184,173,176,179,181,186)
n=length(heights)
(alpha_post=alpha+n/2)
```

```
## [1] 46

(beta_post=beta+sum((heights-mu)^2)/2)

## [1] 1380

##Next, we plot the posterior and the prior distribution
curve(dinvgamma(x, shape=alpha_post, rate=beta_post), from=15, to=50,
      xlab=expression(sigma^2), ylab="Density", col=1, lty=1)
curve(dinvgamma(x, shape=alpha, rate=beta), from=15, to=50,
      n=200, add=T, col=1, lty=2)
legend("topright",c("Prior","Posterior"), bg="white",
       lty=c(2,1), col=1, bty="n")
```



c. Using the change-of-variables formula, we obtain the following posterior density for σ

$$\frac{2\beta^\alpha}{\Gamma(\alpha)}\sigma^{-2\alpha-1}e^{-\frac{\beta}{\sigma^2}}.$$