Homework 6 Solution

Math 315, Fall 2019

***BSM* Chapter 2 exercises 16**

(a) $f(Y|\lambda) = \lambda e^{-\lambda Y}$, $\lambda > 0$, $Y \geq 0$

Target: Derive Jeffreys' prior

$\ell(\lambda) = \log(\lambda) - \lambda Y$

$\ell'(\lambda) = \frac{1}{\lambda} - Y$

$\ell''(\lambda) = -\frac{1}{\lambda^2}$

So $I(\lambda) = -E(\ell''(\lambda)) = \frac{1}{\lambda^2}$

Thus, Jeffreys' prior is $\pi(\lambda) \propto \frac{1}{\lambda}$

(b) $\int_0^\infty \frac{1}{\lambda} d\lambda = \infty$; thus Jeffreys' prior is <u>not</u> valid.

(c)

$p(\lambda|Y) \propto e^{-\lambda Y}$

Since $\int_0^\infty e^{-\lambda Y} = \frac{1}{Y}$, the posterior is proper for $Y > 0$.

1. *Jeffreys' prior.*

In class we showed that the Jeffreys' prior for the Binomial(n, θ) model is Beta(1/2, 1/2). In this problem you will explore why a Jeffreys' prior is said to be transformation invariant.

(a)

Suppose you reparameterize the binomial distribution with $\gamma = \log[\theta/(1-\theta)]$, so that

$$f(y|\gamma) = \binom{n}{y} e^{\gamma y}(1 + e^{\gamma})^{-n}.$$

Derive the Jeffreys' prior distribution for γ under this model.

(a) $\gamma = \log[\theta/(1-\theta)]$

$f(y|\gamma) = \binom{n}{y} e^{\gamma y}(1+e^{\gamma})^{-n}$

Target: Derive Jeffreys' prior

$l(\gamma) = \log\binom{n}{y} + \gamma y - n\log(1+e^{\gamma})$

$l'(\gamma) = y - \frac{ne^{\gamma}}{1+e^{\gamma}} = y - ne^{\gamma}(1+e^{\gamma})^{-1}$

$$\begin{aligned} l''(\gamma) &= -ne^{\gamma}(1+e^{\gamma})^{-1} - ne^{\gamma}(-1)(1+e^{\gamma})^{-2}e^{\gamma} \\ &= \frac{-ne^{\gamma}}{1+e^{\gamma}} + \frac{ne^{2\gamma}}{(1+e^{\gamma})^2} \\ &= \frac{-ne^{\gamma} - ne^{2\gamma}}{(1+e^{\gamma})^2} + \frac{ne^{2\gamma}}{(1+e^{\gamma})^2} = \frac{-ne^{\gamma}}{(1+e^{\gamma})^2} \end{aligned}$$

So $I(\gamma) = -E(l''(\gamma))$

$= \frac{ne^{\gamma}}{(1+e^{\gamma})^2}$

Thus the Jeffreys' prior is

$$\pi(\gamma) \propto \sqrt{\frac{ne^{\gamma}}{(1+e^{\gamma})^2}} \propto \sqrt{\frac{e^{\gamma}}{(1+e^{\gamma})^2}}$$

*Note: there's no need to find a matching distributional form here

(b)

Take the Beta(1/2, 1/2) prior distribution we derived in class and apply the change of variables formula to obtain the induced prior density on γ. Does this agree with your answer to part (a)? (If you have forgotten the change of variables formula, see equation 2.36 for a reminder.)

(b) Target: Show Jeffreys' prior for θ results in Jeffreys' prior for γ

Recall: $\theta \sim \text{Beta}(\frac{1}{2}, \frac{1}{2})$, so $\pi_\theta(\theta) \propto \theta^{-1/2}(1-\theta)^{-1/2}$

Applying the chang of variables formula we have

$$\pi_\gamma(\gamma) = \pi_\theta\left(\frac{e^\gamma}{1+e^\gamma}\right)\left|\frac{d}{d\gamma}\frac{e^\gamma}{1+e^\gamma}\right|$$

$$\propto \left(\frac{e^\gamma}{1+e^\gamma}\right)^{1/2}\left(1-\frac{e^\gamma}{1+e^\gamma}\right)^{1/2}\left|\frac{-e^{2\gamma}}{(1+e^\gamma)^2}+\frac{e^\gamma}{1+e^\gamma}\right|$$

$$= \left[\frac{e^\gamma}{1+e^\gamma}\left(\frac{1}{1+e^\gamma}\right)\right]^{-1/2}\left|\frac{e^\gamma}{(1+e^\gamma)^2}\right|$$

$$= \sqrt{\frac{e^\gamma}{1+e^\gamma}}$$

So, if we start with Jeffreys' prior on θ and apply the monotone trasformation, $\gamma = \log[\theta/(1-\theta)]$, the induced prior on γ is Jeffreys' prior on γ.

2. *Multiparameter model.*

```r
sleep <- c(9.0, 8.5, 7.0, 8.5, 6.0, 12.5, 6.0, 9.0, 8.5, 7.5,
           8.0, 6.0, 9.0, 8.0, 7.0, 10.0, 9.0, 7.5, 5.0, 6.5)
```

(a)

Use a Normal(μ, σ^2) sampling distribution to model the data, and assign the noninformative prior $\pi(\mu, \sigma^2) = 1/\sigma^2$. Describe how to draw a random sample from the posterior distribution for (μ, σ^2), and then draw a random sample of size 10,000 from this distribution.

There are two approaches that you could describe (so far): conditional sampling, or the grid approximation. I describe each below, but only one was necessary.

Conditional sampling

```r
# Calculate necessary summary statistics
n <- length(sleep)
s2 <- var(sleep)
ybar <- mean(sleep)

# Sample from sigma | data
sigma2s <- 1 / rgamma(1e4, (n-1)/2, (n-1) * s2 / 2)

# Sample from mu | sigma, data
mus <- rnorm(1e4, ybar, sqrt(sigma2s / n))
```

Grid approximation

I'll present the version that uses the for loop, but the vectorized log likelihood approach is a good idea!

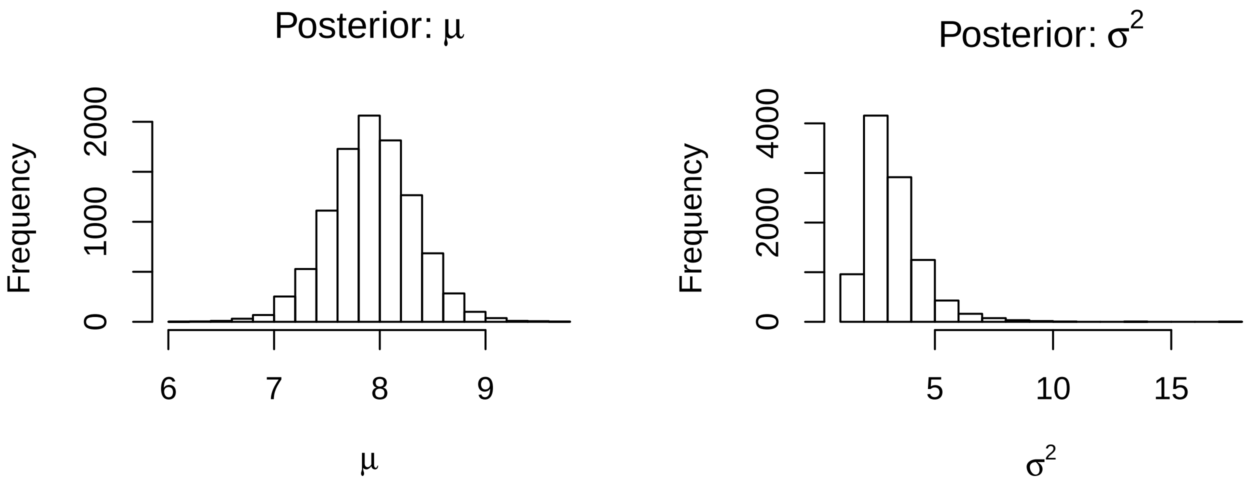
```r
# Create grid over the coordinate plane
param_grid <- expand.grid(
  mu = seq(from = 4, to = 14, length.out = 1000),
  sigma = seq(from = 0.1, to = 4, length.out = 1000)
)

# Calculate joint log prior for each point on grid
logprior <- log(1 / param_grid$sigma^2)

# Calculate log likelihood for each point on grid
ll <- numeric(length = nrow(param_grid))
for(i in 1:nrow(param_grid)) {
  ll[i] <- sum(dnorm(sleep, mean = param_grid[i, "mu"],
                     sd = param_grid[i, "sigma"], log = TRUE))
}

# Calculate log posterior, then exponentiate
logposterior <- logprior + ll
unstd_posterior <- exp(logposterior - max(logposterior)) # numeric stability
posterior <- unstd_posterior / sum(unstd_posterior)

# Sample from the joint posterior
posterior_draws <- dplyr::sample_n(
  param_grid, size = 1e4, replace = TRUE, weight = posterior
)
```

(b)

Plot the marginal posterior distributions for μ and σ^2, and calculate the posterior mean and 90% credible interval.

Either histograms or density plots are fine here. Remember that your answer may differ slightly due to Monte Carlo error.

```
par(mfrow = c(1, 2))
hist(mus, xlab = expression(mu), main = bquote("Posterior:" ~ mu))
hist(sigma2s, xlab = expression(sigma^2), main = bquote("Posterior:" ~ sigma^2))
```




```
mean(mus)
```

```
## [1] 7.924369
```

```
quantile(mus, probs = c(0.05, 0.95))
```

```
##       5%      95%
## 7.273026 8.574436
```

```
mean(sigma2s)
```

```
## [1] 3.220864
```

```
quantile(sigma2s, probs = c(0.05, 0.95))
```

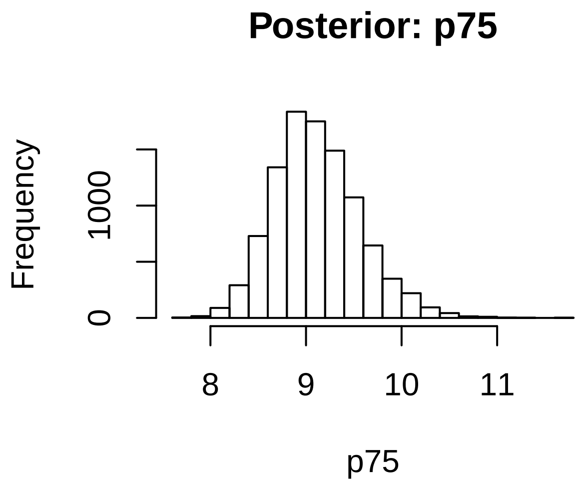
```
##       5%      95%
## 1.810541 5.424451
```

(c)

Plot the posterior distribution for the upper 75th quantile, $p_{75} = \mu + 0.674\sigma$ and calculate its mean and a 90% credible interval.

Either a histogram or density plot is fine here. Remember that your answer may differ slightly due to Monte Carlo error.

```
p75 <- mus + 0.674 * sqrt(sigma2s)
hist(p75, xlab = "p75", main = "Posterior: p75")
```




```r
mean(p75)
```

```
## [1] 9.116121
```

```r
quantile(p75, probs = c(0.05, 0.95))
```

```
##       5%      95%
## 8.441508 9.922587
```