

STAT 513 fa 2020 hw 8

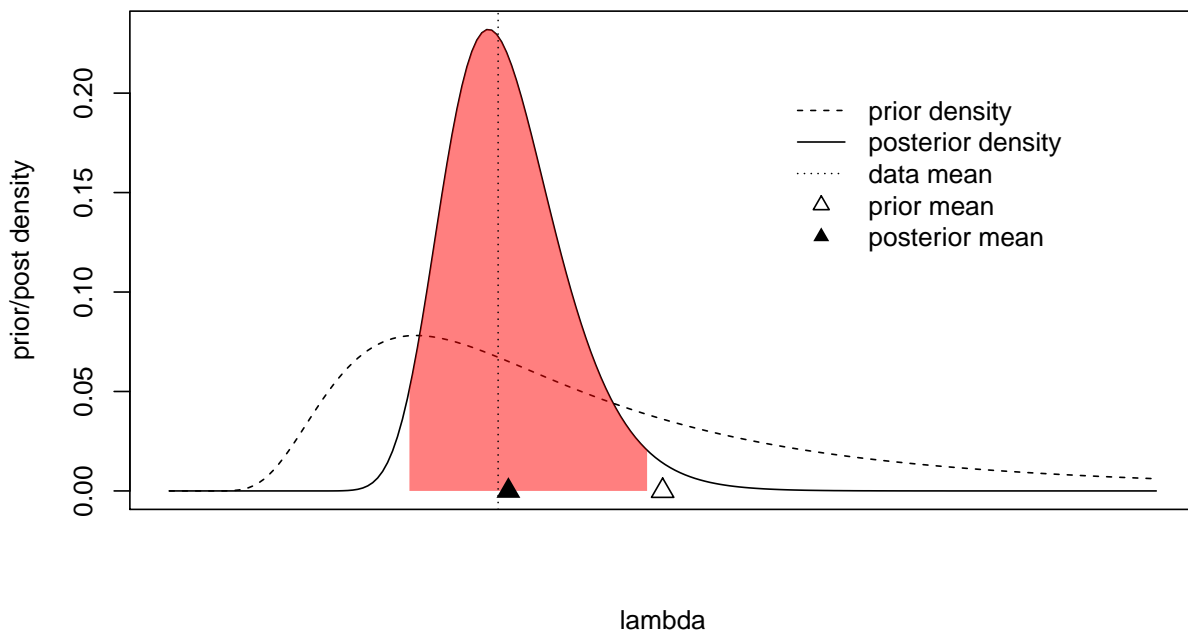
1. Suppose you are to observe the random variables X_1, \dots, X_n coming from the hierarchical model

$$\begin{aligned} X_1, \dots, X_n | \lambda &\stackrel{\text{ind}}{\sim} \text{Exponential}(\lambda) \\ \lambda &\sim \text{Inverse-Gamma}(\alpha, \beta), \end{aligned}$$

where the pdf of the $\text{Inverse-Gamma}(\alpha, \beta)$ distribution is given by

$$f(\lambda; \alpha, \beta) = \frac{1}{\Gamma(\alpha)\beta^\alpha} \left(\frac{1}{\lambda}\right)^{\alpha+1} \exp\left(-\frac{1}{\lambda\beta}\right) \mathbf{1}(\lambda > 0).$$

- (a) Find an expression for the prior (unconditional) mean of λ .
- (b) Write down the joint pdf of the data $X_1, \dots, X_n | \lambda$ conditional on λ .
- (c) Find the posterior distribution of $\lambda | X_1, \dots, X_n$.
- (d) Give an expression for the posterior mean $\mathbb{E}[\lambda | X_1, \dots, X_n]$ in terms of α , β , n , and the data mean \bar{X}_n .
- (e) Refer to the plot below, which shows the prior and posterior densities of λ under $\beta = 1/30$ and $\alpha = 3$ for an observed value of $\bar{X}_n = 10$ based on a sample of size $n = 30$. The shaded region represents the 95% equal-tails Bayesian credible interval for $\lambda | X_1, \dots, X_n$.



- i. Give the prior mean of λ .

- ii. Give the posterior mean of $\lambda|X_1, \dots, X_n$.
- iii. Find the endpoints of the 95% equal-tails credible interval for λ , which is depicted in the plot above. *Hint: Use the fact*

$$\lambda \sim \text{Inverse-Gamma}(\alpha, \beta) \iff 1/\lambda \sim \text{Gamma}(\alpha, \beta)$$

to find a credible interval for $1/\lambda$, and then transform the interval so that it contains λ with probability 0.95, conditional on the data.

2. Suppose we are interested in estimating a success probability p based on the number of successes in n Bernoulli trials, and suppose that we use the following Bayesian setup:

$$Y|p \sim \text{Binomial}(n, p)$$

$$p \sim \text{Uniform}(0, 1).$$

- (a) What beliefs about p does the prior distribution reflect?
- (b) Find the posterior distribution of $p|Y$.
- (c) Give an expression for the posterior mean of $p|Y$.
- (d) Suppose $n = 10$ and $Y = 3$ is observed. Give the equal-tails 95% credible interval for $p|Y$.
- (e) i. Run a simulation in R to compare the coverage of the 95% frequentist Wald-type confidence interval for p given by

$$\hat{p} \pm z_{0.025} \sqrt{\hat{p}(1 - \hat{p})/n},$$

where $\hat{p} = Y/n$ and the coverage of the 95% Bayesian credible interval: Generate 1000 values of p from the $\text{Uniform}(0, 1)$ distribution, and for each of these, generate a value of Y from the $\text{Binomial}(n, p)$ distribution. Then compute the upper and lower bounds of the frequentist confidence interval and of the Bayesian credible interval and record whether the true value of p was captured in the interval or not. Then compute the proportion of times each interval captured the true value of p . Use $n = 10$. Show your R code and give the coverage results.

- ii. Interpret your simulation; which interval performed better? What happens when you use $n = 100$?

- (f) Suppose you wish to test the hypotheses

$$H_0: p \leq 1/2 \text{ versus } H_1: p > 1/2.$$

Let $n = 10$ and suppose you observe $Y = 6$.

- i. Compute the prior odds in favor of H_0 over H_1 and interpret the number.
- ii. Compute the posterior odds in favor of H_0 over H_1 and interpret the number.
- iii. Compute the Bayes factor associated with observing $Y = 6$.
- iv. In your opinion, is the evidence against H_0 strong enough to reject it?

3. Let \mathbf{Y} be an $n \times 1$ vector of random variables and \mathbf{X} be an $n \times p$ matrix of real numbers such that

$$\begin{aligned}\mathbf{Y}|\boldsymbol{\beta} &\sim \text{Normal}(\mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{I}_n) \\ \boldsymbol{\beta} &\sim \text{Normal}(\mathbf{0}, \sigma^2\kappa_0^{-1}\mathbf{I}_p),\end{aligned}$$

where $\boldsymbol{\beta}$ is a $p \times 1$ vector of random variables and σ^2 and κ_0 are known, fixed constants. This is the multiple linear regression model with Normally distributed errors and with the regression coefficients regarded as random variables. It can be shown that the posterior distribution of $\boldsymbol{\beta}|\mathbf{Y}$ is given by

$$\boldsymbol{\beta}|\mathbf{Y} \sim \text{Normal}\left((\mathbf{X}^T\mathbf{X} + \kappa_0\mathbf{I}_p)^{-1}\mathbf{X}^T\mathbf{Y}, \sigma^2(\mathbf{X}^T\mathbf{X} + \kappa_0\mathbf{I}_p)^{-1}\right).$$

- (a) Suggest a Bayesian estimator for $\boldsymbol{\beta}$ based on the posterior distribution of $\boldsymbol{\beta}|\mathbf{Y}$.
- (b) Install the R package `hdi` and read in the data set `riboflavin` which is included in that package. Enter `?riboflavin` into the console to learn more about the data; the data contain for $n = 71$ samples the values of a response variable measuring riboflavin production and the values of 4088 predictor variables, which are gene expression levels. Store the responses in a vector \mathbf{Y} and store in a matrix \mathbf{X} the values of the first 100 predictors, centering both \mathbf{Y} and the columns of \mathbf{X} to have mean 0 and scaling them to have standard deviation 1. All of this is achieved with the following code:

```
library(hdi)
data(riboflavin)
Y <- scale(riboflavin[,1])
X.allcovariates <- scale(riboflavin[,-1])
X <- X.allcovariates[,1:100]
```

Try to compute the least-squares estimator of $\boldsymbol{\beta}$ based on the data in the matrix \mathbf{X} and the vector \mathbf{Y} . Explain why you get an error.

- (c) Set $\kappa_0 = 1$ and compute the Bayesian estimator you suggested in part (a). Give the first 5 entries of the estimated value of $\boldsymbol{\beta}$. *Hint: You can make the $p \times p$ identity matrix \mathbf{I}_p in R with the command `diag(p)`.*
- (d) Set $\kappa_0 = 10$ and do the same thing.
- (e) Describe how the estimates changed.
- (f) For what value of κ_0 would your Bayesian estimator be the same as the least-squares estimator?
- (g) What happens if $\kappa_0 \rightarrow \infty$? *Hint: You can try very large values of κ_0 and see what happens.*
- (h) What beliefs about $\boldsymbol{\beta}$ are reflected by choosing κ_0 to be very large?