Introduction to Bayesian Econometrics, Second Edition
Answers to Exercises, Chapter 7

1. (a) To derive the conditional posterior distribution of h, we treat μ as a constant and drop the parts that do not have h. Thus, dropping the second $\exp[\cdot]$ of $\pi(\mu, h|y)$, we have

$$\begin{aligned}\pi(h|\mu, y) &\propto h^{n/2} \exp\left[-\frac{h}{2}\sum(y_i-\mu)^2\right] h^{\alpha_0/2-1}\exp\left[-h\frac{\delta_0}{2}\right] \\ &= h^{(\alpha_0+n)/2-1}\exp\left[-h\frac{\delta_0+\sum(y_i-\mu)^2}{2}\right].\end{aligned}$$

   (b) To derive the conditional posterior distribution of μ, treat h as a constant and drop the parts that do not have μ. Thus,

$$\begin{aligned}\pi(\mu|h, y) &\propto \exp\left[-\frac{h}{2}\sum(y_i-\mu)^2\right]\exp\left[-\frac{h_0}{2}(\mu-\mu_0)^2\right] \\ &\propto \exp\left[-\frac{hn+h_0}{2}\left(\mu^2-2\mu\left(\frac{h_0\mu_0+hn\bar{y}}{h_0+hn}\right)\right)\right] \\ &\propto \exp\left[-\frac{hn+h_0}{2}\left(\mu-\frac{h_0\mu_0+hn\bar{y}}{h_0+hn}\right)^2\right].\end{aligned}$$

2. The R code necessary to process these data with the *MCMCpack* function `MCMCregress()` is given in `AnswerEx7_2.r`. It will generate a summary of the posterior distribution, trace graphs, and graphs of the marginal posterior distributions of the intercept, the coefficient of smoking, and the error variance.

3. The prior distributions are

$$\begin{aligned}\pi(\theta_1) &= \frac{\beta_{10}^{\alpha_{10}}}{\Gamma(\alpha_{10})}\,\theta_1^{\alpha_{10}-1}e^{-\beta_{10}\theta_1}, \\ \pi(\theta_2) &= \frac{\beta_{20}^{\alpha_{20}}}{\Gamma(\alpha_{20})}\,\theta_2^{\alpha_{20}-1}e^{-\beta_{20}\theta_2}, \\ \pi(k=j) &= 1/n,\end{aligned}$$

and the likelihood function is

$$f(y|\theta_1, \theta_2, k) = \frac{1}{y!} \prod_{i=1}^{k} e^{-\theta_1} \theta_1^{y_i} \prod_{i=k+1}^{n} e^{-\theta_2} \theta_2^{y_i}.$$

Thus,

$$\pi(\theta_1, \theta_2, k|y) \propto \quad \theta_1^{\alpha_{10}-1} e^{-\beta_{10}\theta_1} \theta_2^{\alpha_{20}-1} e^{-\beta_{20}\theta_2} \prod_{i=1}^{k} e^{-\theta_1} \theta_1^{y_i} \prod_{i=k+1}^{n} e^{-\theta_2} \theta_2^{y_i},$$

and

$$\begin{aligned}
\pi(\theta_1|\theta_2, y, k) &\propto \theta_1^{\alpha_{10}-1} e^{-\beta_{10}\theta_1} \prod_{i=1}^{k} e^{-\theta_1} \theta_1^{y_i} \\
&= G(\alpha_{10} + \sum_{1}^{k} y_i, \beta_{10} + k), \\
\pi(\theta_2|\theta_1, y, k) &\propto \theta_2^{\alpha_{20}-1} e^{-\beta_{20}\theta_2} \prod_{i=k+1}^{n} e^{-\theta_2} \theta_2^{y_i} \\
&= G(\alpha_{20} + \sum_{k+1}^{n} y_i, \beta_{20} + n - k), \\
\pi(k|\theta_1, \theta_2, y) &\propto e^{-k\theta_1} \theta_1^{\sum_1^k y_i} e^{-(n-k)\theta_2} \theta_2^{\sum_{k+1}^n y_i} \\
&\propto e^{-k(\theta_1-\theta_2)} \theta_1^{\sum_1^k y_i} \theta_2^{\sum_1^n y_i - \sum_1^k y_i} \\
&\propto e^{-k(\theta_1-\theta_2)} \left(\frac{\theta_1}{\theta_2}\right)^{\sum_1^k y_i}.
\end{aligned}$$

Since we know that $\sum_k \pi(k|\theta_1, \theta_2, y) = 1$, we have

$$\pi(k|\theta_1, \theta_2, y) = \frac{e^{-k(\theta_1-\theta_2)} \left(\frac{\theta_1}{\theta_2}\right)^{\sum_1^k y_i}}{\sum_k e^{-k(\theta_1-\theta_2)} \left(\frac{\theta_1}{\theta_2}\right)^{\sum_1^k y_i}}.$$

The R program `AnswerEx7_3.r` contains code that estimates models with zero and one change point; it utilizes the `MCMCpoisson()` and `MCMCpoissonChange()`

functions of *MCMCpack*.

The output of `AnswerEx7_3.r` provides a summary of the coefficients and graphs of the coefficients and the probabilities of the period in which a change is likely to have occurred. The logarithms of the marginal likelihoods strongly favor the model with one change (-81.917) to the model with no change (-89.737). The summary and the graphs of the θs show that the mean is about 3.270 before the change point and 1.032 after. The graphs of the probabilities indicate that the change point is likely to have occurred within a few years of 1891 (the 40th observation), which is consistent with the new union's beginning in 1889. (You can move between the two graphs by using the Page Up and Page Down keys.)

4. Since the proposal density is a symmetric random-walk kernel, $q(y) = q(x)$, we have

$$\begin{aligned}
\alpha(x,y) &= \frac{f(y)/q(y)}{f(x)/q(x)} \\
&= \frac{f(y)}{f(x)} \\
&= \frac{y^2(1-y)^3}{x^2(1-x)^3}
\end{aligned}$$

5.

$$\begin{aligned}
\alpha(x,y) &= \min\left\{\frac{f(y)/q(y)}{f(x)/q(x)}, 1\right\} \\
&= \min\left\{e^{-|y|+|x|}, 1\right\}.
\end{aligned}$$

An example of R code for this problem is in `AnswerEx7_5.r`. Running the program will produce summaries, acceptances rates, histograms of the approximate density overlaid with the true Laplace distribution, and autocorrelations of the sampled values.

The plots show that mixing can be slow either due to a very small variance of the random walk process, $\sigma = .05$, despite an extremely high acceptance rate (0.9814) or due to a large value, $\sigma = 100$, with an extremely low acceptance rate (0.0199). The results with a σ of 1 or 2 seem reasonable, both in terms of acceptance rates and the autocorrelations.

6. An example of R code for this problem is in `AnswerEx7_6.r`.

   For the AR approach, first obtain c by computing the mode of the $\mathrm{Beta}(5, 5)$ distribution. In this case, since the distribution is symmetric around 1/2, the mode occurs at 1/2. (More generally, by differentiation, you can show that the mode occurs at $(\alpha - 1)/(\alpha + \beta - 2)$.) R tells us that $\mathrm{Beta}(1/2|5, 5) = 2.460938$. To be on the safe side, we use $c = 2.460939$. An algorithm may be found in the R code.

   For the MH approach,
$$\alpha(x, y) = \frac{\mathrm{Beta}(y|5, 5)}{\mathrm{Beta}(x|5, 5)}$$
   because $q(x, y) = q(y, x) = q(y) = 1(0 \leq y \leq 1)$.

   Your output will depend on the seed for the random number generator. With `set.seed(123)`, the mean was estimated well with both algorithms, an acceptance rate for the AR of 0.409, which is close to $1/c = 0.406$, and an MH acceptance rate of 0.505. The graphs appear in Figure 1.

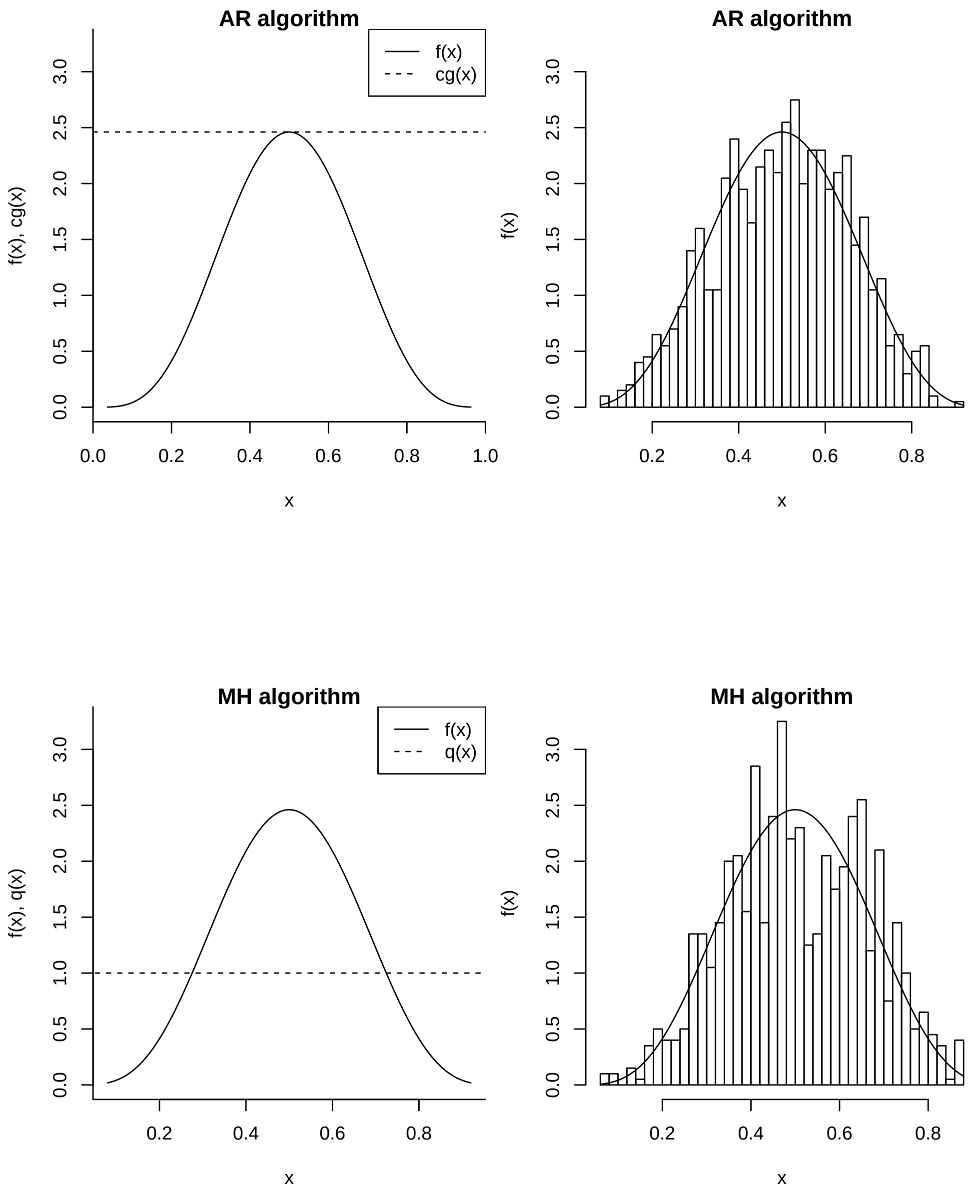


Figure 1: Graphs for exercise 7.6

7. Since

$$f(y|\beta) = \left(\frac{1}{\sqrt{2\pi}}\right)^n \exp\left[-\frac{1}{2}\sum_{i=1}^{n}(y_i - \beta x_i)^2\right],$$
$$\pi(\beta) = \frac{1}{\Gamma(2)}\beta e^{-\beta}\, 1(\beta > 0),$$

we have

$$\pi(\beta|y) \propto \beta \exp\left[-\beta\right] \exp\left[-\frac{1}{2}\sum_{i=1}^{n}(y_i - \beta x_i)^2\right].$$

By taking the log,

$$\log \pi(\beta|y) = \text{constant} + \log\beta - \beta - \frac{1}{2}\sum_{i=1}^{n}(y_i - \beta x_i)^2,$$

we have

$$\frac{\partial \log \pi(\beta|y)}{\partial \beta} = \frac{1}{\beta} - 1 + \sum_{i=1}^{n}(y_i - \beta x_i)x_i = 0, \tag{1}$$
$$\frac{\partial^2 \log \pi(\beta|y)}{\partial \beta^2} = -\frac{1}{\beta^2} - \sum x_i^2. \tag{2}$$

Accordingly, from (1)

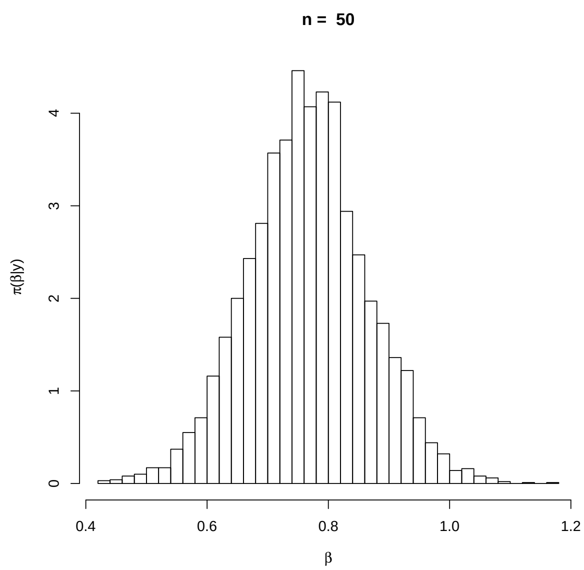
$$\hat{\beta} = \frac{(\sum x_i y_i - 1) + \sqrt{(\sum x_i y_i - 1)^2 + 4\sum x_i^2}}{2\sum x_i^2},$$

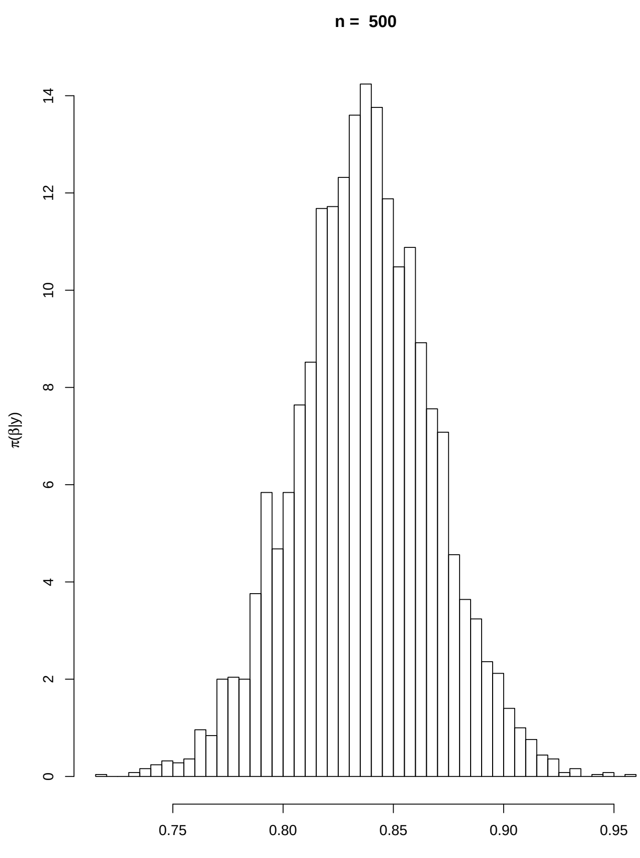
and from (2)

$$\left[-\frac{\partial^2 \ln \pi(\beta|y)}{\partial \beta^2}\right]^{-1}_{\beta=\hat{\beta}} = \left[\frac{1}{\hat{\beta}^2} + \sum x_i^2\right]^{-1}.$$

R code for this problem is in `AnswerEx7_7n50.r` for $n = 50$ and `AnswerEx7_7n500.r` for $n = 500$.

As the following plots show, the posterior density has a smaller variance when the size of the data is larger.


n = 50
π(β|y)
0
1
2
3
4
0.4
0.6
0.8
1.0
1.2
β


n = 500
π(β|y)
0
2
4
6
8
10
12
14
0.75
0.80
0.85
0.90
0.95
β

8. Since

$$\begin{aligned}
f(y|\beta,\sigma^2) &= \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^n (y_i-\beta x_i)^2\right], \\
\pi(\beta) &= \frac{1}{\Gamma(2)}\beta e^{-\beta}, \\
\pi(\sigma^2) &= \frac{(3/2)^{5/2}}{\Gamma(5/2)}\frac{1}{\sigma^7}e^{-3/2\sigma^2},
\end{aligned}$$

therefore

$$\pi(\beta|y) \propto \beta\exp\left[-\beta\right]\exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^n (y_i-\beta x_i)^2\right],$$

and

$$\begin{aligned}
\pi(\sigma^2|\beta,y) &\propto \sigma^{-7}e^{-3/2\sigma^2}\sigma^{-n}\exp\left[-\frac{1}{2\sigma^2}\sum(y_i-\beta x_i)^2\right] \\
&= \frac{1}{(\sigma^2)^{n/2+5/2+1}}\exp\left[-\frac{1}{\sigma^2}\frac{3+\sum(y_i-\beta x_i)^2}{2}\right] \\
&\propto IG\left(\frac{n+5}{2}, \frac{3+\sum(y_i-\beta x_i)^2)}{2}\right).
\end{aligned}$$

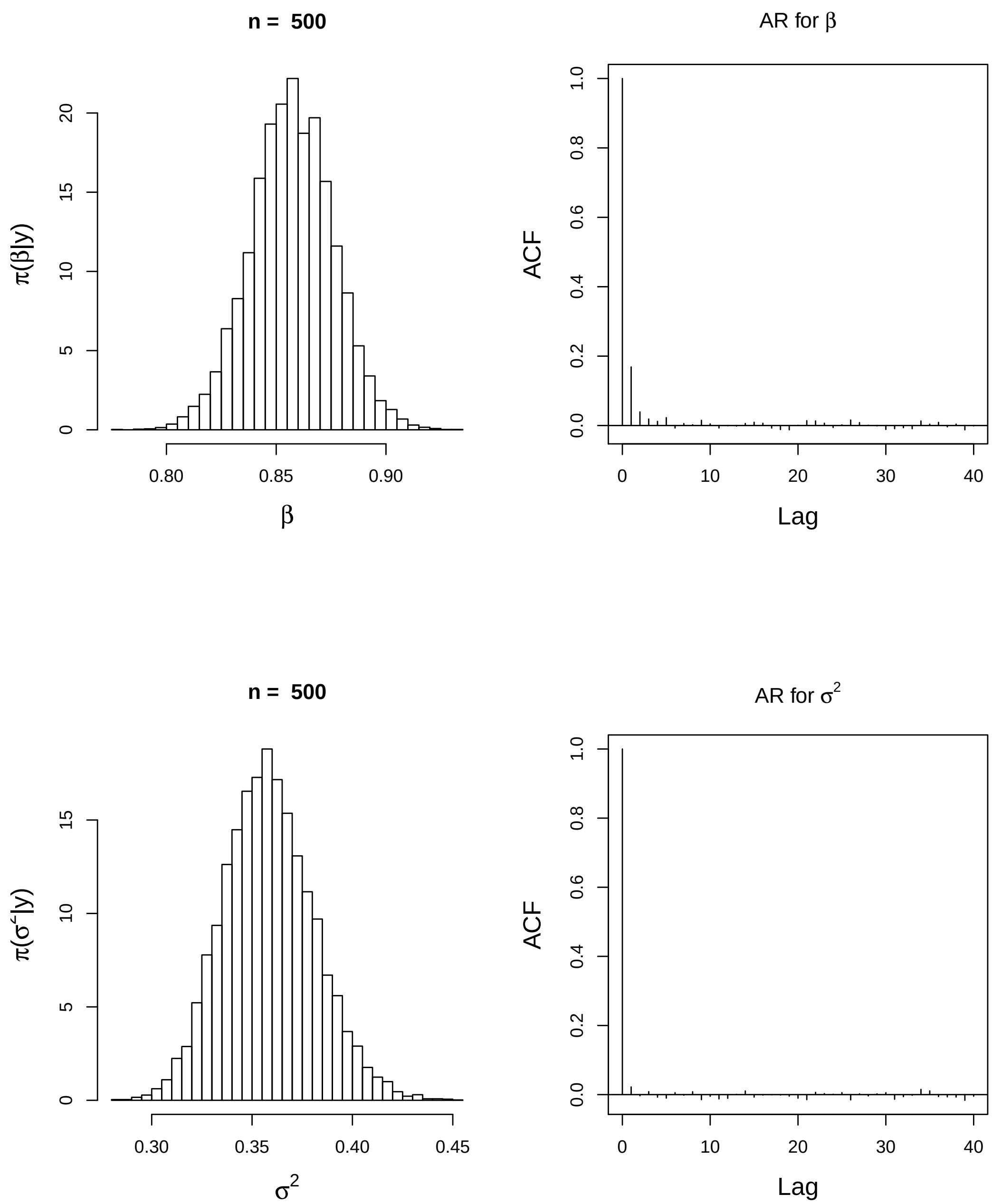
Note that

$$\hat{\beta} = \frac{(\frac{1}{\sigma^2}\sum x_i y_i - 1) + \sqrt{(\frac{1}{\sigma^2}\sum x_i y_i - 1)^2 + 4\frac{1}{\sigma^2}\sum x_i^2}}{2\frac{1}{\sigma^2}\sum x_i^2}$$

and

$$\left[-\frac{\partial^2 \ln \pi(\beta|y)}{\partial\beta^2}\right]^{-1}_{\beta=\hat{\beta}} = \left[\frac{1}{\hat{\beta}^2} + \frac{1}{\sigma^2}\sum x_i^2\right]^{-1}.$$

R code for this problem is in `AnswerEx7_8n500.r` for $n = 500$, $G = 10,000$, and 1,000 burn in values. With the seed set in the program, the true value of $\beta = 0.892$ and the true value of $\sigma^2 = 0.372$. The figure below shows the posterior distributions for β and σ^2 and the autocorrelation functions for the output. Running the program displays various summaries. You can experiment with different n and modify the scale factor.



n = 500
π(β|y)
β
AR for β
ACF
Lag
n = 500
π(σ²|y)
σ²
AR for σ²
ACF
Lag