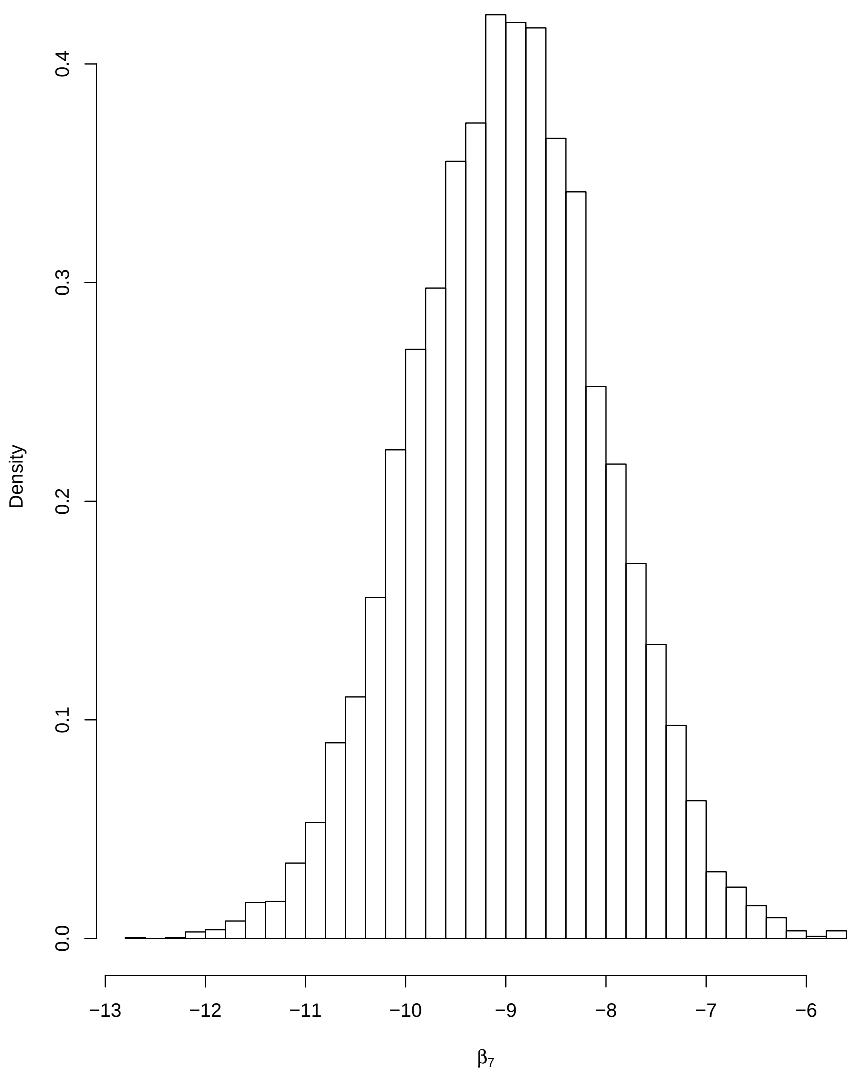
Introduction to Bayesian Econometrics, 2nd Edition
Answers to Exercises, Chapter 8

1. The R code in `AnswerEx8_1.r` uses *MCMCpack* to run the regression. We use the same prior on the intercept, smoking, and the variance as in exercise 7.2. For the other variables, we took a mean of 0 and a variance of 10. The Bayes factor for the model that includes smoking relative to the one that does not appears in the output. It is large enough to suggest that smoking should be included in the model. The figure below of the coefficient of the smoking dummy variable shows that it is less than zero with a very high probability.

Effect of smoking on birthweight, Gaussian errors: $\pi(\beta_7|y)$



2. Under the first specification, we have

$$\pi(\beta, \sigma^2|y) \propto (\sigma^2)^{-n/2} \left\{ \prod_i \left[1 + \frac{(y_i - x_i'\beta)^2}{\sigma^2} \right] \right\}^{-(\nu+1)/2}$$
$$\times \exp\left[-\frac{1}{2}(\beta - \beta_0)' B_0^{-1} (\beta - \beta_0) \right]$$
$$\times \left(\frac{1}{\sigma^2} \right)^{\alpha_0/2+1} \exp\left[-\frac{\delta_0}{2\sigma^2} \right].$$

You can see that there is no way to combine the βs in the first and second line, or the σ^2s in the first and third lines. Replacing the normal prior for β by an expression proportional to

$$\left[1 + \frac{1}{\nu_0}(\beta - \beta_0)' \Sigma_0^{-1} (\beta - \beta_0) \right]^{-(\nu_0+K)/2}$$

does not improve the situation.

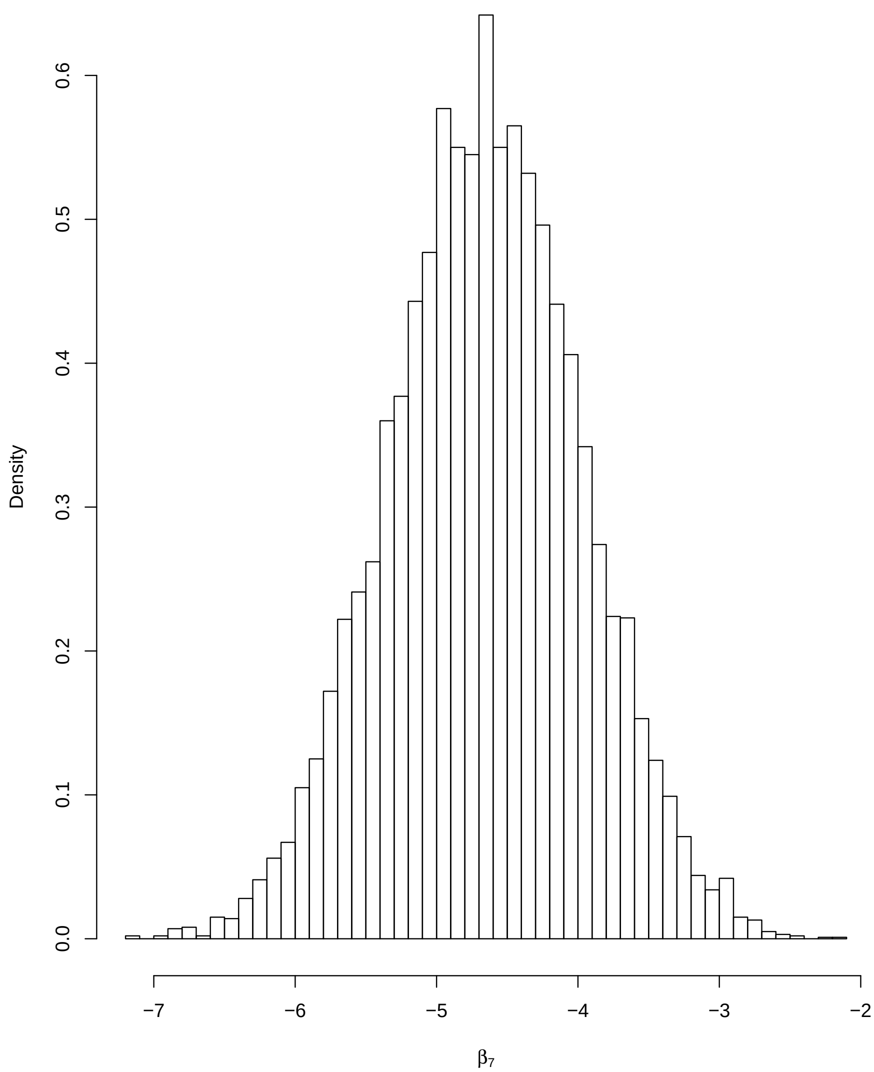
3. If you have a lot of time, you can run a program that I wrote, `AnswerEx8_3.r`. (I hope someone can write a program that runs faster.) It does the sampling with Student-t errors and provides a summary. As in exercise 8.1, the Bayes Factor favors the model with the smoking variable; also, the Student-t error model is favored by the Bayes factor over the Gaussian error model. The base e values for the marginal likelihood are as follows:

Normal errors with smoking variable	-4989.427
Normal errors without smoking variable	-5011.220
t errors with smoking variable	-4989.422
t errors without smoking variable	-5011.213

Note that the mean value of β_7 is -4.656, suggesting a smaller effect of smoking on birthweight than the normal error model, for which the mean is about -8.9.

The figure below summarizes the distribution of the coefficient of smoking.

Effect of smoking on birthweight, t errors: $\pi(\beta_7|y)$



4. See the R code `AnswerEx8_4.r`. It computes the mean value of β_U for nine combinations of prior mean (0, 0.1, 0.2) and prior variance (.001, .0036, .0050). See the following table:

	0.001	0.0036	0.005
0	0.051	0.096	0.105
0.1	0.116	0.129	0.132
0.2	0.18	0.163	0.159

Table 1: Mean of β_U

5. An algorithm that does this is coded in the file `egRegression.R` as the function `egRegStudent()`. It is specified by using the conditional distributions for $[\beta|y,\lambda,\sigma^2]$, $[\sigma^2|y,\beta,\lambda]$, and $[\lambda_i|y,\beta,\sigma^2]$, $i = 1,\dots,n$, that appear in the text. In algorithm form:

   (a) Choose starting values for $\lambda^{(1)}$ and $\sigma^{2(1)}$. Draw $\beta^{(1)}$ from

$$\beta|y,\lambda^{(1)},\sigma^{2(1)} \sim N_K(\bar{\beta}^{(1)}, B_1^{(1)}),$$

   where

$$\begin{aligned} B_1^{(1)} &= \left[\sigma^{-2(1)}X'\Lambda^{(1)}X + B_0^{-1}\right]^{-1}, \\ \bar{\beta}^{(1)} &= B_1^{(1)}\left[\sigma^{-2(1)}X'\Lambda^{(1)}y + B_0^{-1}\beta_0\right]. \end{aligned}$$

   (b) At the gth iteration, draw

$$\begin{aligned} \beta^{(g)}|y,\lambda^{(g-1)},\sigma^{2(g-1)} &\sim N_K(\bar{\beta}^{(g)}, B_1^{(g)}), \\ \sigma^{2(g)}|y,\beta^{(g)},\lambda^{(g-1)} &\sim \mathrm{IG}(\alpha_1/2, \delta_1^{(g)}/2), \\ \lambda_i^{(g)}|y,\beta^{(g)},\sigma^{2(g)} &\sim \mathrm{Ga}(\nu_1/2, \nu_{2i}^{(g)}/2), \quad i = 1,\dots,n, \end{aligned}$$

where

$$
\begin{aligned}
B_1^{(g)} &= \left[\sigma^{-2(g-1)} X' \Lambda^{(g-1)} X + B_0^{-1}\right]^{-1}, \\
\bar{\beta}^{(g)} &= B_1^{(g)} \left[\sigma^{-2(g-1)} X' \Lambda^{(g-1)} y + B_0^{-1}\beta_0\right], \\
\alpha_1^{(g)} &= \alpha_0 + n, \\
\delta_1^{(g)} &= \delta_0 + (y - X\beta^{(g)})' \Lambda^{(g-1)} (y - X\beta^{(g)}), \\
\nu_1^{(g)} &= \nu + 1, \\
\nu_{2i}^{(g)} &= \nu + \sigma^{-2(g)} (y_i - x_i' \beta^{(g)})^2.
\end{aligned}
$$

(c) Go to (b) until $g = B + G$, where B is the burn-in sample and G is the desired sample size.

6. Use the specification of section 8.2.1. At the first iteration, choose values for $\sigma^{2(1)}$, $\lambda^{(1)}$, and $y^{*(1)}$ to draw $\beta^{(1)}$ from its full conditional distribution.

   At the gth step,

   (a) Draw $\beta^{(g)}$ from $N_K(\bar{\beta}^{(g)}, B_1^{(g)})$, where

$$
\begin{aligned}
B_1^{(g)} &= \left[\sigma^{-2(g-1)} X' \Lambda^{(g-1)} X + B_0^{-1}\right]^{-1}, \\
\bar{\beta}^{(g)} &= B_1^{(g)} \left[\sigma^{-2(g-1)} X' \Lambda^{(g-1)} y^{*(g-1)} + B_0^{-1}\beta_0\right].
\end{aligned}
$$

   (b) Draw $\sigma^{2(g)}$ from $\mathrm{IG}(\alpha_1/2, \delta_1^{(g)}/2)$, where

$$
\begin{aligned}
\alpha_1 &= \alpha_0 + n, \\
\delta_1^{(g)} &= \delta_0 + (y^{*(g-1)} - X\beta^{(g)})' \Lambda^{(g-1)} (y^{*(g-1)} - X\beta^{(g)}).
\end{aligned}
$$

(c) Draw $\lambda_i^{(g)}$ from $\text{Ga}(\nu/2, \nu_i^{(g)}/2$, where

$$\nu_i^{(g)} = \nu + (y_i - x_i'\beta^{(g)})'(y_i - x_i'\beta^{(g)}).$$

(d) Draw $y_i^{*(g)}$ for $i \in C$ from $\text{TN}_{(-\infty,0]}(x_i'\beta^{(g)}, \lambda^{-1(g)}\sigma^{2(g)})$.

(e) Repeat steps (a)–(d) until the desired sample is obtaned.

7. For the Chib method applied to the binary probit model, we have

$$\hat{m}(y) = \frac{f(y|\beta^*)\,\pi(\beta^*)}{\hat{\pi}(\beta^*|y)},$$

where β^* is the the mean of the $\beta^{(g)}$. The likelihood function is computed directly from

$$f(y|\beta^*) = \prod_i \left[\Phi(x_i'\beta^*)\right]^{y_i} \left[1 - \Phi(x_i'\beta^*)\right]^{1-y_i}.$$

$\pi(\beta^*)$ is computed from its prior, $N_K(\beta^*|\beta_0, B_0)$. Finally, we can approximate the denominator by

$$\hat{\pi}(\beta^*|y) = \frac{1}{G}\sum_i N_K(\beta^*|\bar{\beta}^{(g)}, B_1),$$

where

$$B_1 = \left[X'X + B_0^{-1}\right]^{-1},$$
$$\bar{\beta}^{(g)} = B_1\left[X'y^{*(g)} + B_0^{-1}\beta_0\right].$$

8. (a) Choose $\beta^{(1)}$ and $\lambda^{(1)}$ and then draw $y^{*(1)}$ from its full conditional distribution.

(b) At the gth iteration, draw

$$y_i^{*(g)} \sim \begin{cases} \mathrm{TN}_{(-\infty,0]}(x_i'\beta^{(g-1)}, \lambda_i^{-1(g-1)}), & \text{if } y_i = 0, \\ \mathrm{TN}_{(0,\infty)}(x_i'\beta^{(g-1)}, \lambda^{-1(g-1)}), & \text{if } y_i = 1, \end{cases} \qquad i = 1, \ldots, n,$$

$$\beta^{(g)} \sim N_K(\bar{\beta}^{(g)}, B_1^{(g)}),$$

$$\lambda_i^{(g)} \sim \mathrm{Ga}\left(\frac{\nu+1}{2}, \frac{\nu_i^{(g)}}{2}\right),$$

where

$$\begin{aligned} B_1^{(g)} &= \left[X'\Lambda^{(g-1)}X + B_0^{-1}\right]^{-1}, \\ \bar{\beta}^{(g)} &= B_1^{(g)}\left[X'\Lambda^{(g-1)}y^{*(g)} + B_0^{-1}\beta_0\right], \\ \nu_i^{(g)} &= \nu + (y_i^{*(g)} - x_i'\beta^{(g)})^2. \end{aligned}$$

9. Since $\pi(\beta|y) \propto f(y|\beta)\pi(\beta)$, you should recognize the likelihood function

$$f(y|\beta) \propto \prod \left(\frac{\exp[x_i'\beta]}{1+\exp[x_i'\beta]}\right)^{y_i} \left(\frac{1}{1+\exp[x_i'\beta]}\right)^{1-y_i}$$

and

$$\pi(\beta) \propto \exp\left[-\frac{1}{2}(\beta-\beta_0)'B_0^{-1}(\beta-\beta_0)\right].$$

Given the y_i^*, which would be drawn from a truncated logistic distribution, we have

$$\pi(\beta|y^*) \propto \prod_i \frac{\exp(y_i^* - x_i'\beta)}{\left(1+\exp(y_i^* - x_i'\beta)\right)^2} \exp\left[-\frac{1}{2}(\beta-\beta_0)'B_0^{-1}(\beta-\beta_0)\right].$$

Clearly, the right-hand side does not simplify to a standard distribution for β.

10. We need to find the full conditional distributions of y_{ij}^*, γ_j, and x_i to verify the algorithm. Start with the likelihood function written in terms of the y_{ij}^*, where

y_{ij} can take on the values 0, 1, or NA:

$$p(y_{ij}|\gamma_j, x_i) =$$
$$1(y_{ij} = 0)\, 1(y^*_{ij} \leq 0) + 1(y_{ij} = 1)\, 1(y^*_{ij} > 0) + 1(y_{ij} = \text{NA}) 1(-\infty < y^*_{ij} < \infty).$$

It follows from Bayes theorem that

$$\pi(\gamma, x, y^*|y) \propto$$
$$\prod_{i,j} \left[1(y_{ij} = 0)\, 1(y^*_{ij} \leq 0) + 1(y_{ij} = 1)\, 1(y^*_{ij} > 0) + 1(y_{ij} = \text{NA}) 1(-\infty < y^*_{ij} < \infty)\right]$$
$$\times \quad \pi(y^*_{ij}|\gamma_j, x_i)\, \pi(\gamma_j)\, \pi(x_i).$$

Since $y^*_{ij} \sim N(x_i^{*'} \gamma_j, 1)$, the full conditional distribution of y^*_{ij} in step 2 of algorithm 8.5 follows immediately.

For γ_j, from the definitions of $y^*_{\cdot j}$, X^*, and $\epsilon_{\cdot j}$, we have

$$y^*_{\cdot j} = X^* \gamma_j + \epsilon_{\cdot j}.$$

Accordingly, given the prior distribution of γ,

$$\pi(\gamma_j|y^*_{\cdot j}, X^*) \propto \exp\left[-\frac{1}{2}(y^*_{\cdot j} - X^*\gamma_j)'(y^*_{\cdot j} - X^*\gamma_j)\right] \exp\left[-\frac{1}{2}(\gamma_j - t_{0j})' T_{0j}^{-1} (\gamma_j - t_{0j})\right].$$

Standard computations now justify the full conditional distribution for γ_j of algorithm 8.5.

As noted in the text, we have

$$w_{i\cdot} = Bx_i + \epsilon_{i\cdot}.$$

Therefore, for the x_i, from its prior and from the terms in which it appears in

$\pi(\gamma, x, y^*|y)$, we have

$$\pi(x_i|y^*_{i\cdot}, \gamma) \propto \exp\left[-\frac{1}{2}(w_{i\cdot} - x_i B)'(w_{i\cdot} - x_i B)\right] \exp\left[-\frac{1}{2}(x_i - v_{0i})' V_{0i}^{-1}(x_i - v_{0i})\right].$$

Completing the square in x_i justifies the last line of step 2 in algorithm 8.5.

11. By Bayes theorem, we have

$$\pi(\beta|y) \propto \prod_i \exp\left[-\exp[x_i'\beta]\right] [\exp[x_i'\beta]^{y_i} \exp\left[-\frac{1}{2}(\beta - \beta_0)' B_0^{-1}(\beta - \beta_0)\right].$$

It should be clear that an MH step is needed to sample for β. Chib, Greenberg, and Winkelmann (1998) suggest a tailored MH proposal density in the form of t distribution with mean at the mode $(\hat{\beta})$ of $\pi(\beta|y)$ and scale matrix (H_β) proportional to the negative inverse of the Hessian of $\log \pi(\beta|y)$. In an attempt to obtain proposal values that are relatively distant from the current point, they suggest reflecting the current value around $\hat{\beta}$ and then adding a drawing from a multivariate t distribution with mean 0 and scale matrix proportional to V_β.

Another possible proposal density is a random walk around the current value of β.