In regression, the priors on the regression coefficients are crucial.

Whether or not $\beta_j = 0$ defines whether $X_j$ is “important” in the regression.

For any $j$, a useful prior for $\beta_j$ is:
Spike-and-Slab Priors for Linear Models

- Here: $P(\beta_j = 0) = h_{0j}$ (prior probability that $X_j$ is not needed in the model)

- $P(\beta_j \neq 0) = 1 - h_{0j} = h_{1j}(f_j - (-f_j)) = 2f_j h_{1j}$ (where $[-f_j, f_j]$ contains all “reasonable” values for $\beta_j$)

- To include $X_j$ in the model with certainty, set $h_{0j} = 0$.

- To reflect more doubt that $X_j$ should be in the model, increase the ratio

  $$\frac{h_{0j}}{h_{1j}} = \frac{h_{0j}}{(1 - h_{0j})/2f_j} = 2f_j \frac{h_{0j}}{1 - h_{0j}}$$

- Recently, “nonparametric priors” have become popular, typically involving a mixture of Dirichlet processes.
The Monte Carlo method involves studying a distribution (e.g., a posterior) and its characteristics by generating many random observations having that distribution.

If $\theta^{(1)}, \ldots, \theta^{(S)} \overset{iid}{\sim} \pi(\theta|\mathbf{x})$, then the empirical distribution of $\{\theta^{(1)}, \ldots, \theta^{(S)}\}$ approximates the posterior, when $S$ is large.

By the law of large numbers,

$$\frac{1}{S} \sum_{s=1}^{S} g(\theta^{(s)}) \to E[g(\theta)|\mathbf{x}]$$

as $S \to \infty$. 
So as $S \to \infty$:

$$\bar{\theta} = \frac{1}{S} \sum_{s=1}^{S} \theta^{(s)} \to \text{posterior mean}$$

$$\frac{1}{S-1} \sum_{s=1}^{S} (\theta^{(s)} - \bar{\theta})^2 \to \text{posterior variance}$$

$$\frac{\#\{\theta^{(s)} \leq c\}}{S} \to P[\theta \leq c|\mathbf{x}]$$

$$\text{median}\{\theta^{(1)}, \ldots, \theta^{(S)}\} \to \text{posterior median}$$

(and similarly for any posterior quantile).
If the posterior is a “common” distribution, as in many conjugate analyses, we could draw samples from the posterior using R functions.

**Example 1:** (General Social Survey)

- **Sample 1:** # of children for women age 40+, no bachelor’s degree.
- **Sample 2:** # of children for women age 40+, bachelor’s degree or higher.
- Assume Poisson(θ₁) and Poisson(θ₂) models for the data.
- We use gamma(2,1) priors for θ₁ and for θ₂.
The Monte Carlo Method

- **Data:** \( n_1 = 111, \sum_i x_{i1} = 217 \)
- **Data:** \( n_2 = 44, \sum_i x_{i2} = 66 \)
- \( \Rightarrow \) Posterior for \( \theta_1 \) is gamma(219,112).
- \( \Rightarrow \) Posterior for \( \theta_2 \) is gamma(68, 45).
- Find \( P[\theta_1 > \theta_2 | x_1, x_2] \).
- Find posterior distribution of the ratio \( \frac{\theta_1}{\theta_2} \).
- See R example using Monte Carlo method on course web page.