1. For model $Y_i = g(x_i; \beta) + \epsilon_i$, we studied two methods for estimating parameters $\beta$ and doing statistical inference. In the parametric regression, the distribution of $\epsilon_i$ is specified, i.e. $\epsilon_i \sim (0, \sigma^2)$ with unknown variance $\sigma^2$. In the bootstrap approach, we did not assume any distribution form for $\epsilon_i$. In both parametric and nonparametric bootstrap approach, we assumed $\beta$ to be fixed unknown scalar values to be estimated.

In the Bayesian framework, parameters are assumed to be random and described by a probability distribution called the prior distribution $\pi(\beta)$. This is formulated before the data are observed based on a certain prior belief.

Let $f(Y|\beta)$ be the sampling distribution of a sample $Y_i$. The posterior distribution of $\beta$ is given by Bayes formula

$$
\pi(\beta|Y) = \frac{f(Y|\beta)\pi(\beta)}{\int f(Y|\beta)\pi(\beta) \, d\beta}.
$$

Note that

$$
\pi(\beta|Y) \propto f(Y|\beta)\pi(\beta) = \pi(\beta)\prod_{i=1}^n f(y_i|\beta).
$$

The product term is the likelihood. The Bayes estimate $\hat{\beta}$ would be the mean of the posterior distribution, i.e.

$$
\hat{\beta} = \mathbb{E}(\beta|Y) = \frac{\int \beta f(Y|\beta)\pi(\beta) \, d\beta}{\int f(Y|\beta)\pi(\beta) \, d\beta}.
$$

In general one need to compute $\mathbb{E}(h(\beta)|Y)$ for Bayesian inference but this may be hard and Monte-Carlo integration based on MCMC might be one answer.

2. Consider a simple linear model

$$
Y_i = \beta_0 + \epsilon_i, \quad i = 1, \ldots, n
$$

where $\epsilon_i \sim N(0, \sigma^2)$ with known $\sigma^2$. This can be written as $Y_i|\beta_0 \sim N(\beta_0, \sigma^2)$. Let the prior be $N(\mu, \tau^2)$. Then the posterior is

$$
\pi(\beta_0|y) \propto \exp\left(-\frac{(\beta_0 - \mu)^2}{2\tau}\right) \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^n (\beta_0 - y_i)^2\right).
$$

We can find, for instance, the Bayes estimate $\hat{\beta}$ analytically but let us estimate via MCMC. Let $y_i$ be grip from strength data. Based on the random-walk Metropolis algorithm, we simulate the posterior $\pi(\beta_0|Y)$ based on the symmetric propositional density $q(y|x) \sim N(x, 1)$.

3. Homework 5. Fit model $Y_i = \beta_0 + \beta_1 x_i + \epsilon_i$ using the 2D Gibbs sampler from lecture 23. You need to simulate 2D Markov chain $(\beta_0^t, \beta_1^t)$ whose invariant distribution is the posterior $\pi(\beta_0, \beta_1|Y)$. All these examples are somewhat boring use of MCMC. Use it for complicated nonlinear model fitting when it is hard to compute $\mathbb{E}(h(\beta)|Y)$.

4. Lecture 25-27 will deal with Monte-Carlo optimization and the EM algorithm. Please read the textbook.