

Exercise

Exercise

Suppose $X_1, X_2, \dots \stackrel{\text{iid}}{\sim} p(x|\theta)$ given θ . Suppose the posterior predictive $p(x_{n+1}|x_{1:n})$ is too complicated to compute analytically, however, you can easily sample from the posterior $p(\theta|x_{1:n})$, and you can also easily compute the c.d.f.

$$F(c|\theta) = \mathbb{P}(X_i \leq c \mid \theta).$$

How would you construct a Monte Carlo approximation of

$$\mathbb{P}(X_{n+1} \leq c \mid x_{1:n})$$

for some given value of c ?

Solution

Approach 1 (Better)

$$\begin{aligned}\mathbb{P}(X_{n+1} \leq c \mid x_{1:n}) &= \int \mathbb{P}(X_{n+1} \leq c \mid \theta, x_{1:n}) p(\theta \mid x_{1:n}) d\theta \\ &= \int \mathbb{P}(X_{n+1} \leq c \mid \theta) p(\theta \mid x_{1:n}) d\theta \\ &\approx \frac{1}{N} \sum_{i=1}^N \mathbb{P}(X_{n+1} \leq c \mid \theta_i) \\ &= \frac{1}{N} \sum_{i=1}^N F(c \mid \theta_i)\end{aligned}$$

where $\theta_1, \dots, \theta_N \stackrel{\text{iid}}{\sim} p(\theta \mid x_{1:n})$. If you didn't understand the first line, note that by the law of iterated expectations, for any random variables X and Y ,

$$\mathbb{P}(X \leq c) = \mathbb{E}(\mathbf{1}(X \leq c)) = \mathbb{E}(\mathbb{E}(\mathbf{1}(X \leq c) \mid Y)) = \mathbb{E}(\mathbb{P}(X \leq c \mid Y)).$$

This can also be derived by writing out the integrals or sums for each of these expectations (and you should definitely do this if you didn't fully understand).

Approach 2

Another approach is to sample $(\theta_i, X_{n+1}^{(i)}) \sim p(\theta, x_{n+1} \mid x_{1:n})$ i.i.d. for $i = 1, \dots, N$, and use the approximation

$$\mathbb{P}(X_{n+1} \leq c \mid x_{1:n}) \approx \frac{1}{N} \sum_{i=1}^N \mathbf{1}(X_{n+1}^{(i)} \leq c).$$

These samples can be obtained by drawing $\theta_i \sim p(\theta \mid x_{1:n})$ and $X_{n+1}^{(i)} \sim p(x \mid \theta)$ independently for $i = 1, \dots, N$. However, this is probably not as good of an approximation as Approach 1, especially when $\mathbb{P}(X_{n+1} \leq c \mid x_{1:n})$ is very close to 0 or 1. The general principle is that if there is any part of the computation that can be done analytically, doing so will improve the approximation. (This is usually true, but not always.)