7.5 The Sampling Importance Resampling (SIR) Algorithm

Suppose that we are interested in simulating values from a random vector \( \mathbf{X} \) with p.m.f. \( f(\mathbf{x}) = C_1 f_0(\mathbf{x}) \), known up to a normalising constant, using simulated values from a Markov chain with limiting probabilities given by the p.m.f. \( g(\mathbf{x}) = C_2 g_0(\mathbf{x}) \), also known up to a normalising constant.

The SIR method starts by generating \( m \) successive states \( y_1, \ldots, y_m \), of a Markov chain with limiting p.m.f. \( g(\mathbf{x}) \). Next we define the weights \( w_i \) as

\[
w_i = \frac{f_0(y_i)}{g_0(y_i)}, \quad i = 1, 2, \ldots, m,
\]

and generate a random vector \( \mathbf{X} \) from the p.m.f.

\[
p_j = P(\mathbf{X} = y_j) = \frac{w_j}{\sum_{i=1}^{m} w_i}, \quad j = 1, 2, \ldots, m.
\]

It can be shown that the latter distribution for \( \mathbf{X} \) converges to the target distribution \( f(\mathbf{x}) \) as \( m \to +\infty \).

7.5.1 Application to Bayesian computation

The SIR method is as if tailored for Bayesian statistics. Here is the idea: suppose we observe random vectors \( \mathbf{X}_1, \ldots, \mathbf{X}_n \) with distribution \( f(\mathbf{x}|\theta) \), and let \( \pi(\theta) \) denote a prior distribution we entertain for \( \theta \in \Theta \), that is known. What we are interested of course, if the posterior distribution of \( \theta | \mathbf{X} \), given by

\[
\pi(\theta|\mathbf{x}_1, \ldots, \mathbf{x}_n) = \frac{f(\mathbf{x}_1, \ldots, \mathbf{x}_n|\theta)\pi(\theta)}{\int f(\mathbf{x}_1, \ldots, \mathbf{x}_n|\theta)\pi(\theta) d\theta} \propto f(\mathbf{x}_1, \ldots, \mathbf{x}_n|\theta)\pi(\theta).
\]

We can now use the SIR algorithm by generating vectors \( \theta_1, \ldots, \theta_m \) from the prior \( \pi(\theta) \). More precisely, let

\[
f_0(\theta) = f(\mathbf{x}_1, \ldots, \mathbf{x}_n|\theta)\pi(\theta),
\]

\[
g(\theta) = g_0(\theta) = \pi(\theta), \quad \text{and using weights}
\]

\[
w(\theta) = f(\mathbf{x}_1, \ldots, \mathbf{x}_n|\theta),
\]

and then easily apply the method.

Obtaining means of functions of the parameters is now trivial, since

\[
E(h(\theta)|\mathbf{X}) = \sum_{j=1}^{m} p_j h(\theta_j),
\]

where

\[
p_j = \frac{\int f(\mathbf{x}_1, \ldots, \mathbf{x}_n|\theta_j)\pi(\theta)}{\sum_{j=1}^{m} \int f(\mathbf{x}_1, \ldots, \mathbf{x}_n|\theta_j)}.
\]