STK 4021 - Applied Bayesian Analysis and Numerical Methods
Thordis L. Thorarinsdottir
Fall 2014

Problem Set 8

Problem 1 (rejection-acceptance sampling). This exercise provides the basics of the so-called rejection-acceptance sampling strategy which is discussed in Chapter 3.3.2 of the book. It is presented here in the general framework of random variables on certain sample spaces where to goal is to obtain samples from the density $f(x)$. In terms of the current course, such a sampling strategy can be particularly useful when sampling from the posterior density $p(\theta|\mathcal{D})$.

(a) Let Y come from some density $g(y)$, and assume that we choose to keep Y with probability $h(y)$; otherwise we throw it away and go to the next round. Show that an accepted Y then follows the density

$$f(x) = \frac{g(x)h(x)}{\int g(x)h(x)\mathrm{d}x}.$$

(b) Suppose we wish to draw a sample $\{x_1, x_2, \ldots\}$ from some density $f(x)$ from which we cannot sample directly. Assume that $f(x) \leq Mg(x)$ for all x and a constant $M > 0$, where it is easy to obtain samples from g. Show that the two-step algorithm that first draws Y from g, and then keeps this value with probability $f(y)/[Mg(y)]$ returns a sample from f. What is the frequency of rejected Y values, that is of "wasted efforts"?

(c) Let

$$f(x) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1}(1-x)^{b-1} \quad \text{for } 0 < x < 1,$$

that is the Beta distribution with parameters (a, b). Show that f is unlimited if a or b is smaller than 1, and finds its maximum value M_0 for the case $a \geq 1, b \geq 1$.

(d) Let $a = 1.33$ and $b = 1.67$. Draw 1000 samples from the Beta distribution with these parameters, using the rejection algorithm that startes with uniformly distributed variables. How many samples from the uniform distribution did you need in order to produce 1000 samples from the target distribution?

(e) Suppese in general terms that we wish to sample from a density of the form $f(x) = g(x)/I$, where g is nonnegative over a certain region and $I = \int g(x)\mathrm{d}x$. Assume (i) that we sample X from a (simpler) start-density $h(x)$, where $g(x) \leq Mh(x)$ for all x, for some M; and (ii) that we keep this candidate X with probability $g(x)/[Mh(x)]$. Verify that the probability density of the remaining Xs is really $f(x)$.

The importance of this variation on the rejection sampling recipe in (a) is that, here, we don't need to know the normalizing constant I.

(f) Set up a rejection sampling regime to obtain 100,000 samples (x_i, y_i) from the density

$$f(x,y) = g(x,y)/I, \quad g(x,y) = \exp\big(\sin(\sqrt{|xy|})\exp(|y|^{3/2}\big),$$

where I is the integral of g over $[0,1] \times [0,1]$. Make histograms of the two marginal distributions, and find means, standard deviations, and the correlation numerically.

(g) Consider the following recipe for creating samples from $\mathcal{N}(0,1)$: Sample X from $\mathcal{N}(0,2)$ and keep X with probability $\exp(-x^2/4)$. Verify that this recipe works and obtain 100,000 samples.

(h) The bivariate normal density with mean $(0,0)$ and variance $(1,1)$ is of the form

$$f(x,y) = \frac{1}{2\pi}\frac{1}{(1-\rho^2)^{1/2}}\exp\Big(-\frac{1}{2}\frac{1}{1-\rho^2}(x^2+y^2-2\rho xy)\Big).$$

Use rejection sampling to generate 10,000 pairs (x,y) from this bivariate normal distribution, for a couple of values of the correlation parameter ρ. Plot the pairs and check empirically that your algorithm works properly.

Solutions will be discussed in class on October 24.