An additional note on importance sampling

March 13, 2002

This note is an integration to the illustration of variance reduction by importance sampling on page 230 of the book. For additional details see, e.g., [1].

Consider the problem of estimating

$$\theta = \mathrm{E}_f[h(\mathbf{X})],$$

where $\mathbf{X}$ is a random vector with joint density f (the notation E_f is used to emphasize this). Assume for simplicity that $h(\mathbf{x}) \geq 0$. Consider another density g whose support is equal to that of f (i.e., $g(\mathbf{z}) = 0$ only if $f(\mathbf{z}) = 0$). We may write:

$$\mathrm{E}_f[h(\mathbf{X})] = \int h(\mathbf{x}) f(\mathbf{x})\, d\mathbf{x} = \int \frac{h(\mathbf{x}) f(\mathbf{x})}{g(\mathbf{x})} g(\mathbf{x})\, d\mathbf{x} = \int h^*(\mathbf{x}) g(\mathbf{x})\, d\mathbf{x} = \mathrm{E}_g[h^*(\mathbf{X})],$$

where $h^*(\mathbf{X}) = h(\mathbf{x}) f(\mathbf{x}) / g(\mathbf{x})$. Note that the condition on the support of g is needed in order to avoid any trouble with the case $g(\mathbf{x}) = 0$ in the definition of h^*; we may think of integrating only on the support.

The two estimators have the same expectation, but what about the variance? Using the well-known properties of the variance:

$$\begin{aligned}
\mathrm{Var}_f[h(\mathbf{X})] &= \int h^2(\mathbf{x}) f(\mathbf{x})\, d\mathbf{x} - \theta^2 \\
\mathrm{Var}_g[h^*(\mathbf{X})] &= \int h^2(\mathbf{x}) \frac{f(\mathbf{x})}{g(\mathbf{x})} f(\mathbf{x})\, d\mathbf{x} - \theta^2.
\end{aligned}$$

From the second equation, it is easy to see that the choice

$$g(\mathbf{x}) = \frac{h(\mathbf{x}) f(\mathbf{x})}{\theta}$$

leads to the ideal condition $\mathrm{Var}_g[h^*(\mathbf{X})] \equiv 0$. Of course this is an ideal condition, but example 4.8 on page 231 shows how to take advantage of this information. Note also that the condition $h(\mathbf{x}) \geq 0$ is needed in order to ensure that this is a density; see, e.g., [2, p. 122] to see how to deal with a generic function h.

In general, the reduction in variance is:

$$\Delta\mathrm{Var} = \mathrm{Var}_f[h(\mathbf{X})] - \mathrm{Var}_g[h^*(\mathbf{X})] = \int h^2(\mathbf{x}) \left[1 - \frac{f(\mathbf{x})}{g(\mathbf{x})} \right] f(\mathbf{x})\, d\mathbf{x}.$$

From this expression we see that, in order to ensure that the reduction is positive, we should select a new density g such that:

$$\begin{cases} g(\mathbf{x}) < f(\mathbf{x}) & \text{when the term } h^2(\mathbf{x}) f(\mathbf{x}) \text{ is large,} \\ g(\mathbf{x}) > f(\mathbf{x}) & \text{when the term } h^2(\mathbf{x}) f(\mathbf{x}) \text{ is small.} \end{cases}$$

The name "importance sampling" derives from this observation.

[1]This supplement should be used in conjunction with the book: P. Brandimarte, *Numerical Methods in Finance: a MATLAB-Based Introduction*, Wiley, 2001. Please refer to the web page (`www.polito.it/~brandimarte`) for further updates and supplements. Any comment is welcome. My e-mail address is: `brandimarte@polito.it`.

References

[1] G.S. Fishman. *Monte Carlo: Concepts, Algorithms, and Applications.* Springer-Verlag, Berlin, 1996.

[2] R.Y. Rubinstein. *Simulation and the Monte Carlo Method.* Wiley, Chichester, 1981.