STAT 675

Statistical Computing

***Antithetic variables* for Monte Carlo Integration.**

Suppose you have a target integral to calculate which has either been given with or transformed to a range of integration over the unit interval $0 < u < 1$. Denote this as

$$\theta = \int_0^1 g(u)\,du.$$

For simple Monte Carlo integration as presented in Textbook Sec. 6.2, we generate a set of $m > 0$ pseudo-random variates $U_i \sim$ i.i.d.U(0,1) ($i = 1, ..., m$) and from these calculate $G_i = g(U_i)$. Notice that the G_is are themselves i.i.d. so that $E[G_i] = \theta$ and $Var[G_i] = \sigma_G^2$ are both constant. (We could find the precise distribution of G_i via the transformation technique from mathematical statistics – see Casella & Berger, 2002, Sec. 2.1 – but this is not needed here.)

With these, find $\bar{G} = \sum_{i=1}^{m} G_i/m$. Notice that since the G_is are i.i.d., G_i, $E[\bar{G}] = \theta$ and

$$Var[\bar{G}] = \frac{\sigma_G^2}{m}.$$

Also, from the strong law of large numbers, $P[\lim_{m\to\infty} |\bar{G} - \theta| < \varepsilon] = 1$ for every $\varepsilon > 0$, thus on balance $\bar{G}$ represents a reasonable approximation for θ. An estimate of the variance of $\bar{G}$ is s_G^2/m, where s_G^2 is the usual sample variance of the pseudo-random sample of G_is; this latter quantity is found, e.g., in **R** via `var(G)`. An estimate for $Var[\bar{G}]$ is then `var(G)/m`.

By contrast, the method of Antithetic Variables for Monte Carlo integration presented in Textbook Sec. 6.4 takes advantage of the fact that if $U_i \sim$ i.i.d.U(0,1) then $Corr[U_i, 1-U_i] = -1$ (see Textbook Exercise 6.8). To wit, instead of generating m pseudo-random variates, generate only $\frac{m}{2}$ pseudo-random variates

$$U_j \sim \text{i.i.d.U}(0,1) \left(j = 1, ..., \tfrac{m}{2}\right)$$

and then calculate G_j and $G_j' = g(1-U_j)$. Notice that U_j and $1-U_j$ are identically distributed as U(0,1), and so G_j and G_j' are also identically distributed. Further, $E[G_j] = E[G_j'] = \theta$ while $Var[G_i] = Var[G_i'] = \sigma_G^2$.

From these, find the new, $\frac{m}{2}$, pseudo-random quantities

$$W_j = ½(G_j + G_j')$$

and calculate $\bar{W} = \sum_{i=1}^{m/2} W_i/(\frac{m}{2})$. Clearly $E[W_i] = ½E[G_j + G_j'] = ½(E[G_j] + E[G_j']) = ½(\theta + \theta) = 2\theta/2 = \theta$, so $\bar{W}$ represents a reasonable alternative for approximating θ.

The choice between $\bar{G}$ and $\bar{W}$ is informed by the following feature:

$$\begin{aligned} Var[W_i] &= ¼(Var[G_i] + Var[G_j'] + 2Cov[G_j, G_j']) \\ &= ¼(2Var[G_i] + 2Cov[G_j, G_j']) \qquad (G_j \text{ and } G_j' \text{ are identically distributed}) \\ &= ½(\sigma_G^2 + Cov[G_j, G_j']) \end{aligned}$$

Note that $Cov[G_j, G_j']$ is constant with respect to *j*; for simplicity, write this as $Cov[G_j, G_j'] = \gamma$. Then $Var[W_j] = ½(\sigma_G^2 + \gamma)$ in

$$Var[\bar{W}] = Var[W_j]/(\tfrac{m}{2}) = ½(\sigma_G^2 + \gamma)/(\tfrac{m}{2}) = \frac{\sigma_G^2 + \gamma}{m}.$$

But then $Var[\bar{W}] < Var[\bar{G}]$ whenever $\gamma = Cov[G_j, G_j'] < 0$. If g(u) is a monotone function this will be true, as G_j and G_j' are functions of the negatively correlated antithetic variables U_j and $1-U_j$; see Textbook Corollary 6.1. Thus use of $\bar{W}$ to approximate θ via antithetic variables is an approach for *reducing variance* in Monte Carlo integration.