Variance reduction techniques

Arabin Kumar Dey

Assistant Professor

Department of Mathematics
Indian Institute of Technology Guwahati

Talk at Indian Institute of Technology Guwahati

6-th january, 2012

Simulation efficiency

- Consider the usual problem of estimating $\theta = E[h(X)]$, where X is a random variable from distribution with density say f.
- The standard Monte Carlo simulation algorithm is to:
  - generate $X_1, X_2, \cdots, X_n$
  - estimate θ using $\hat{\theta}_n = \frac{1}{n}\sum_{j=1}^{n} Y_j$
    where $Y_j = h(X_j)$.
  - approximate a $100(1-\alpha)\%$ confidence interval given by

    $$\hat{\theta}_n - z_{\alpha/2}\frac{S_n}{\sqrt{n}}, \hat{\theta}_n + z_{\alpha/2}\frac{S_n}{\sqrt{n}}$$

    where S_n^2 is the usual estimate of the variance of Y based on the simulated values $Y_1, \cdots, Y_n$.
- We see from the confidence interval that one way to measure then the quality of the estimator is through the variance.
- This can also be justified because the mean square error is

$$MSE = E[(\bar{Y}-\theta)^2] = Var(\bar{Y}) = \frac{Var(Y)}{n}$$

Some steps you can take to improve simulation efficiency

There are a few recommended steps you can take to improve the efficiency of your simulation, without sacrificing much of the accuracy :

- Develop a good simulation routine.
  - avoid loops if at all possible.
  - work with vectors, strings of numbers, or matrices
- Develop a routine to minimize storing unnecessary values.
  - e.g. because U and $1 - U$ have the same distribution, maybe avoid the additional step of storing U and then computing $1 - U$.
- Write your routine to minimize the time of execution.
- Reduce the variability of the simulation output used to estimate the parameter of interest,θ.
  - These are called *Variance Reduction Techniques*.

Some variance reduction techniques

Some variance reduction techniques in simulation :

- Use of antithetic variables
- Use of control variates
- Variance reduction by conditioning

Use of antithetic variables

- This method involves generating estimates with negative correlation and then adding these estimates to obtain the final estimate.
- Consider two (simulated) estimates of θ: say, $\bar{Y}_1$ and $\bar{Y}_2$. Define

$$\bar{Y} = \frac{1}{2}(\bar{Y}_1 + \bar{Y}_2).$$

- Then, assuming both estimates have the sample number of simulations and they have same variance, we have

$$\begin{aligned} Var(\bar{Y}) &= \frac{1}{4}[Var(\bar{Y}_1) + Var(\bar{Y}_2) + 2\rho\sqrt{Var(\bar{Y}_1)Var(\bar{Y}_2)}] \\ &= \frac{1}{2}Var(\bar{Y}_1)(1+\rho) = \frac{Var(\bar{Y}_1)}{2n}(1+\rho), \end{aligned}$$

where ρ denotes the correlation between $\bar{Y}_1$ and $\bar{Y}_2$.
- Thus, clearly, if ρ is negative, you can gain a variance reduction.

One possible approach

- One possible approach to arrange for negatively correlated estimates is as follows :
  - Suppose $\bar{Y}_1$ is based on random numbers $U_1, U_2, \cdots, U_m$, say $\bar{Y}_1 = g(U_1, U_2, \cdots, U_m)$.
  - Then $\bar{Y}_2$ can be another estimate based on random numbers $1 - U_1, 1 - U_2, \cdots, 1 - U_m$, so that $\bar{Y}_2 = g(1 - U_1, 1 - U_2, \cdots, 1 - U_m)$.
- Both U and $1 - U$ are uniformly distributed and are clearly negatively correlated.
- It can be shown that $\bar{Y}_1$ and $\bar{Y}_2$ will be negatively correlated, and hence can obtain a variance reduction, so long as g is monotone (either increasing or decreasing).

Example 8d

- Suppose we are interested in estimating

$$\theta = E(e^U) = \int_0^1 e^U dU,$$

where U is a U(0,1) random variable.

- Show that the use of antithetic variables U and $1-U$ indeed the reduces the variance.
- To be discussed in lecture.

The use of control variates

- Consider the simple problem of estimating $\theta = E(X)$, where X is drawn from a simulation.
- Suppose there is another random variable Y with expectation $E(Y) = \mu_y$.Then for any constant c, the quantity

$$W = X + c(Y - \mu_y)$$

  is also an unbiased estimator of θ.
- Consider its variance:

$$\begin{aligned} Var(X + c(Y - mu_y)) &= Var(X + cY) \\ &= Var(X) + c^2 Var(Y) + 2cCov(X, Y). \end{aligned}$$

- It can be shown that this variance is minimized when c is equal to

$$\hat{c} = -\frac{Cov(X, Y)}{Var(Y)}$$

Variance of the new estimator

- The variance of the new estimator is:

$$Var(X + \hat{c}(Y - \mu_y)) = Var(X) - \frac{[Cov(X,Y)]^2}{Var(X)}.$$

- Y is called the control variate for the simulation estimator X.
- We can re-express this by dividing both sides by $Var(X)$:

$$\frac{Var(X + \hat{c}(Y - \mu_y))}{Var(X)} = 1 - [Corr(X,Y)]^2,$$

  where

$$Corr(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}},$$

  is correlation between X and Y.
- The variance is therefore reduced by $100[Corr(X,Y)]^2$percent.

–continued

$$\begin{aligned} Var(X+\hat{c}(Y-\mu_y)) &= Var(X+\hat{c}Y) \\ &= Var(X)+\hat{c}^2 Var(Y)+2\hat{c}Cov(X,Y) \\ &= Var(X)-\frac{[Cov(X,Y)]^2}{Var(Y)} \end{aligned} \tag{1}$$

$$\begin{aligned} \frac{Var(X)-\frac{[Cov(X,Y)]^2}{Var(Y)}}{Var(X)} &= 1-[\frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}}]^2 \\ &= 1-\rho_{x,y}^2 \end{aligned} \tag{2}$$

The controlled estimator

- The controlled estimator is therefore

$$\bar{X} + \hat{c}(\bar{Y} - \mu_y),$$

and its variance is given by

$$\bar{X} + \hat{c}(\bar{Y} - \mu_y) = \frac{1}{n}[Var(X) - \frac{[Cov(X,Y)]^2}{Var(Y)}].$$

Example 8h

- Consider the case of estimating (follow-up of Example 8d):

$$\theta = E(e^U) = \int_0^1 e^U dU,$$

- Use a random number U as a control variate and estimate the reduction in variance using the controlled estimator.
- To be discussed in lecture.

Interpreting the control variate approach

- One can indeed interpret the control variate estimator as a combination of estimators of the parameter θ.
- Consider two values X and W with mean $E(X) = E(W) = \theta$. clearly the estimator

$$\alpha X + (1-\alpha)W.$$

  is an unbiased estimator.
- We can choose α that would minimize the variance, giving it in some sense the better estimator. One can show (in class-details!) that the α minimizing the variance is

$$\hat{\alpha} = \frac{Var(W) - Cov(X, W)}{Var(X) + Var(W) - 2Cov(X, W)}.$$

- if $E(Y) = \mu_y$ s therefore known, we can combine the two unbiased estimators X and $X + Y - \mu_y$ without

$$(1-c)X + c(X + Y - \mu_y) = X + c(Y - \mu_y).$$

  giving exactly the control variate estimator.

The antithetic variate as a special case

- if $E(X) = \theta$ with $X = h(U_1, \cdots, U_n)$, then $E(W) = \theta$ where $W = h(1 - U_1, \cdots, 1 - U_n)$.
- Thus, we can then combine these two estimators in the form

$$\alpha X + (1 - \alpha) W.$$

- But note that both X and W have the same distribution, so that their variances $Var(X) = Var(W)$ are equal.
- When the variance are equal, the optimal value of α is clearly $\hat{\alpha} = 1/2$. This leads us to the estimator $\frac{1}{2}(X + W)$, the antithetic variable estimator.

Variance reduction by conditioning

- Recall the law of iterated expectations :

$$E(X) = E[E(X|Y)] = \theta.$$

- This implies that the estimator $E(X|Y)$ is also an unbiased estimator.
- Now, recall the conditional variance formula :

$$Var(X) = E[Var(X|Y)] + Var[E(X|Y)].$$

- Clearly, both terms on the right are non-negative, so that we have

$$Var(X) \geq Var[E(X|Y)].$$

- This implies that the estimator, by conditioning, produces a more superior variance.

Example 8I

- Suppose that Y is an exponential random variable with mean 1.
- Suppose further that conditional on $Y = y$, the random variable X is Normal with mean y and variance 4.
- You are interested in estimating $\theta = P(X > 1)$. Suggest ways of simulating to improve the variance efficiency of estimating θ.
- To be discussed in lecture.
- To estimate $\theta = P(X > 1)$ note that $\theta = E[I(X > 1)]$ where

$$I(X > 1) = \begin{cases} 1, & \text{if } X > 1 \\ 0, & \text{if } X \leq 1 \end{cases}$$