Note Title 03-10-2020

# MA 323

## Lecture # 09

### Variance Reduction Techniques:

The goal of variance reduction techniques is to increase the efficiency of Monte Carlo simulation, by reducing the variance of simulation estimates. The two broad strategies applicable are: (1) Taking advantage of the tractable features of a model to adjust or correct simulation outputs

## (2) Reducing the variability in simulation inputs.

**(A) Control Variates:** The method of control variates is among the most effective and widely used approaches for improving efficiency of Monte Carlo simulation.

It exploits information about the errors in estimation of known quantities to reduce the error in estimation of an unknown quantity.

More specifically, we let $Y_1, Y_2, \ldots, Y_n$ be outputs from $n$ replications of a simulation.

Further, we suppose that $Y_1, Y_2, \ldots, Y_n$ are i.i.d variables.

The goal is to estimate $E[Y_i]$.

The natural choice for this is the sample mean (which is the unbiased estimator): $\overline{Y} = \dfrac{Y_1 + Y_2 + \cdots + Y_n}{n}$

In the next step, for each replication we calculate an output $X_i$ corresponding to $Y_i$.

Suppose that the expectation $E(X)$ of $X_i$ is known, and that the pairs $(X_i, Y_i)$ $i = 1, 2, .., n$ are i.i.d.

Let $(X, Y)$ be a generic pair random variable, with the same distribution as each $(X_i, Y_i)$. For any fixed $b$, we can calculate:

$$Y_i(b) = Y_i - b(X_i - E(X)) \quad (\text{i-th replication})$$

We compute the sample mean

$$\overline{Y}(b) = \overline{Y} - b\left(\overline{X} - E(X)\right) = \frac{1}{n}\sum_{i=1}^{n}\left(Y_i - b(X_i - E(X))\right). \qquad (1)$$

This is a control variate estimator.

The observed error $\overline{X} - E(X)$ serves as a control in estimating $E(Y)$.

Now, $E\left(\overline{Y}(b)\right) = E\left(\overline{Y} - b(\overline{X} - E(X))\right) = E\left(\overline{Y}\right) = E(Y).$

This shows that the control variate is an unbiased estimator of $E(Y)$. Further it is consistent, because with probability 1,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} Y_i(b) = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \left(Y_i - b(X_i - E(X_i))\right)$$

$$= E\left(Y - b(X - E(X))\right) = E(Y)$$

Now, each $Y_i(b)$ has variance

$$Var(Y_i(b)) = Var\left(Y_i - b(X_i - E(X))\right)$$

$$= \sigma_Y^2 - 2b\sigma_X \sigma_Y \rho_{XY} + b^2 \sigma_X^2$$

$$\equiv \sigma^2(b). \qquad \text{———— } (2)$$

Here $\sigma_X^2 = Var(X)$, $\sigma_Y^2 = Var(Y)$ and $\rho_{XY} = \dfrac{\sigma_{XY}}{\sigma_X \sigma_Y}$.

<u>Observe</u> : (1) The control variate estimator $\overline{Y}(b)$ has the variance $\dfrac{\sigma^2(b)}{n}$.

(2) The ordinary sample mean $\overline{Y}$ ( corresponding to $b=0$) has the variance $\dfrac{\sigma_Y^2}{n}$.

Consequently, the control variate estimator has smaller variance than the standard estimator PROVIDED

$$-2b\sigma_X\sigma_Y\rho_{XY} + b^2\sigma_X^2 < 0 \quad \text{i.e.,} \quad b^2\sigma_X < 2b\sigma_Y\rho_{XY}$$

The value of $b$ that minimizes $\sigma^2(b)$ is given by,

$$b^* = \frac{\sigma_Y}{\sigma_X}\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X^2} \quad \text{——— (3)}$$

Substituting this value of $b^*$ in (2), and simplifying, we obtain,

$$\frac{Var\left(\bar{Y} - b^*\left(\bar{X} - E(X)\right)\right)}{Var(\bar{Y})} = 1 - \rho_{XY}^2 \quad \text{——— (4)}$$

Observations: (i) With the optimal $b^*$, the effectiveness of a control variate, as measured by the variance reduction ratio, is ascertained by the magnitude of the correlation between the quantity of interest $Y$ and the control $X$.

(ii) If the computational cost per replication is roughly the same with and without a control variate, then (4) measures the computational speedup resulting from the use of a control.

Specifically, the number of replications of the $Y_i$ required to achieve the same variance as $n$ replications of the control estimator is $n/(1-\rho_{XY}^2)$.

(iii) The variance reduction factor $\frac{1}{1-\rho_{XY}^2}$ increases / decreases as $|\rho_{XY}|$ approaches towards 1 / decreases away from 1.

The problem with the control approach is that the determination of $b^*$ requires the values of $\sigma_Y$ and $\rho_{XY}$. However, in practice we can still benefit from the control variate approach using an estimate of $b^*$.

For example, one could replace the population parameters in 3 with their sample counterparts to obtain the estimate

$$\hat{b}_n = \frac{\sum_{i=1}^{n} (X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^{n} (X_i - \bar{X})^2} \quad \text{--- } 5$$

Dividing numerator and denominator by $n$ and upon application of strong law of large numbers, results in $\hat{b}_n \to b^*$, with probability 1. ( Note: Replacing $b^*$ with $\hat{b}_n$

introduces some bias).

<u>Note</u> : The expression in equation 5 is the slope of the least-square regression line through the points $(X_i, Y_i)$, $i = 1, 2, \ldots, n$.