Optimization and Simulation

Variance reduction

Michel Bierlaire

Transport and Mobility Laboratory
School of Architecture, Civil and Environmental Engineering
Ecole Polytechnique Fédérale de Lausanne

Outline

Example

Use simulation to compute

$$I = \int_0^1 e^x \, dx$$

We know the solution: $e - 1 = 1.7183$

Simulation: consider draws two by two

- Let $r_1, \ldots, r_R$ be independent draws from $U(0,1)$.
- Let $s_1, \ldots, s_R$ be independent draws from $U(0,1)$.

$$I \approx \frac{1}{2R} \left(\sum_{i=1}^{R} e^{r_i} + \sum_{i=1}^{R} e^{s_i} \right) = \frac{1}{R} \sum_{i=1}^{R} \frac{e^{r_i} + e^{s_i}}{2}$$

- Use $R = 10'000$ (that is, a total of 20'000 draws)
- Mean over R draws from $(e^{r_i} + e^{s_i})/2$: 1.720,variance: 0.123

Example

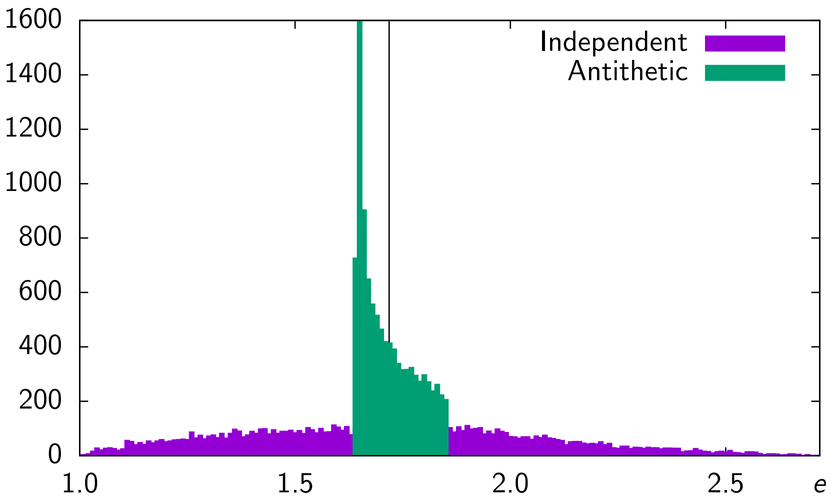
Now, use half the number of draws

- Idea: if $X \sim U(0,1)$, then $(1-X) \sim U(0,1)$
- Let $r_1, \ldots, r_R$ be independent draws from $U(0,1)$.

$$I \approx \frac{1}{R} \sum_{i=1}^{R} \frac{e^{r_i} + e^{1-r_i}}{2}$$

- Use $R = 10'000$
- Mean over R draws of $(e^{r_i} + e^{1-r_i})/2$: 1.7183,variance: 0.00388
- Compared to: mean of $(e^{r_i} + e^{s_i})/2$: 1.720,variance: 0.123

Example


Independent
Antithetic
1600
1400
1200
1000
800
600
400
200
0
1.0
1.5
2.0
2.5
e

Antithetic draws

- Let X_1 and X_2 i.i.d r.v. with mean θ
- Then

$$\operatorname{Var}\left(\frac{X_1 + X_2}{2}\right) = \frac{1}{4}\left(\operatorname{Var}(X_1) + \operatorname{Var}(X_2) + 2\operatorname{Cov}(X_1, X_2)\right)$$

- If X_1 and X_2 are independent, then $\operatorname{Cov}(X_1, X_2) = 0$.
- If X_1 and X_2 are negatively correlated, then $\operatorname{Cov}(X_1, X_2) < 0$, and the variance is reduced.

Back to the example

Independent draws

- $X_1 = e^U$, $X_2 = e^U$

$$
\begin{aligned}
\mathrm{Var}(X_1) = \mathrm{Var}(X_2) &= \mathrm{E}[e^{2U}] - \mathrm{E}[e^U]^2 \\
&= \int_0^1 e^{2x}\,dx - (e-1)^2 \\
&= \frac{e^2-1}{2} - (e-1)^2 \\
&= 0.2420
\end{aligned}
$$

$$
\mathrm{Cov}(X_1, X_2) = 0
$$

$$
\mathrm{Var}\left(\frac{X_1+X_2}{2}\right) = \frac{1}{4}\,(0.2420 + 0.2420)) = 0.1210
$$

Back to the example

Antithetic draws

- $X_1 = e^U$, $X_2 = e^{1-U}$

$$\mathrm{Var}(X_1) = \mathrm{Var}(X_2) = 0.2420$$

$$\begin{array}{rcrcl} \mathrm{Cov}(X_1, X_2) & = & \mathrm{E}[e^U e^{1-U}] & - & E[e^U]E[e^{1-U}] \\ & = & e & - & (e-1)(e-1) \\ & = & -0.2342 & & \end{array}$$

$$\mathrm{Var}\left(\frac{X_1 + X_2}{2}\right) = \frac{1}{4}\left(0.2420 + 0.2420 - 2\ 0.2342)\right) = 0.0039$$

Antithetic draws: generalization

- Suppose that
$$X_1 = h(U_1, \ldots, U_m)$$
where $U_1, \ldots U_m$ are i.i.d. $U(0, 1)$.
- Define
$$X_2 = h(1 - U_1, \ldots, 1 - U_m)$$
- X_2 has the same distribution as X_1
- If h is monotonic in each of its coordinates, then X_1 and X_2 are negatively correlated.
- If h is not monotonic, there is no guarantee that the variance will be reduced.

Another example

$$I = \int_0^1 \left(x - \frac{1}{2}\right)^2 dx$$

- Antithetic draws:

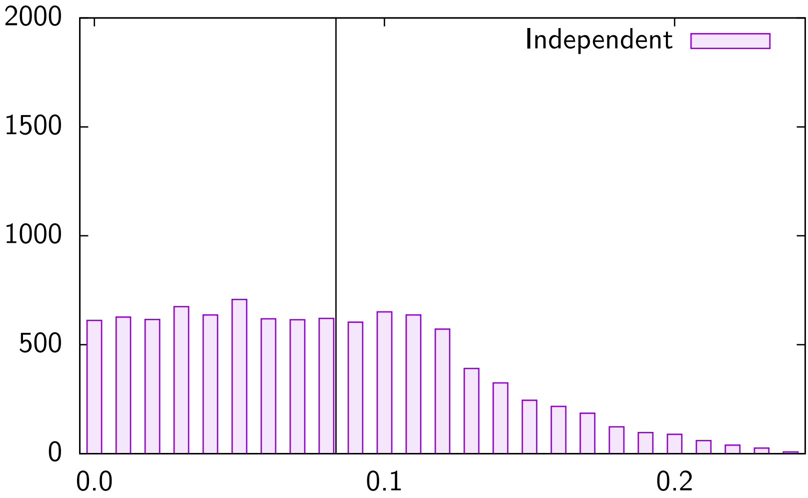
$$X_1 = \left(U - \frac{1}{2}\right)^2, \; X_2 = \left((1-U) - \frac{1}{2}\right)^2$$

- The covariance is positive:

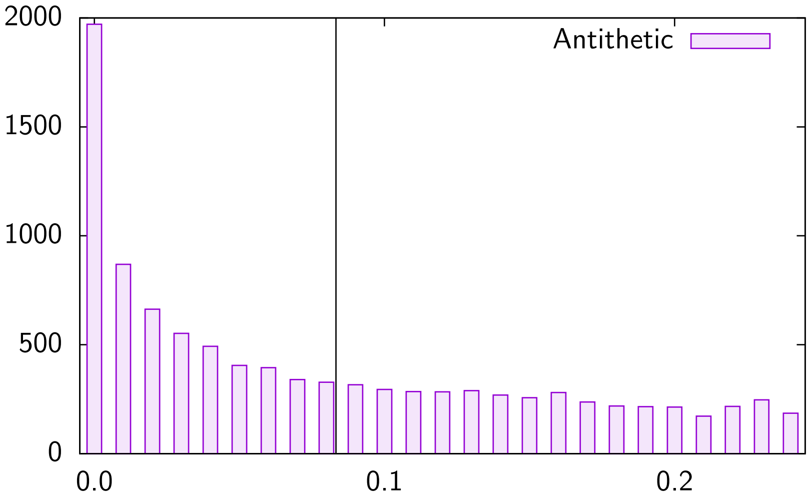
$$\mathrm{Cov}(X_1, X_2) = \frac{1}{180} > 0.$$

- The variance will therefore be (slightly) increased!

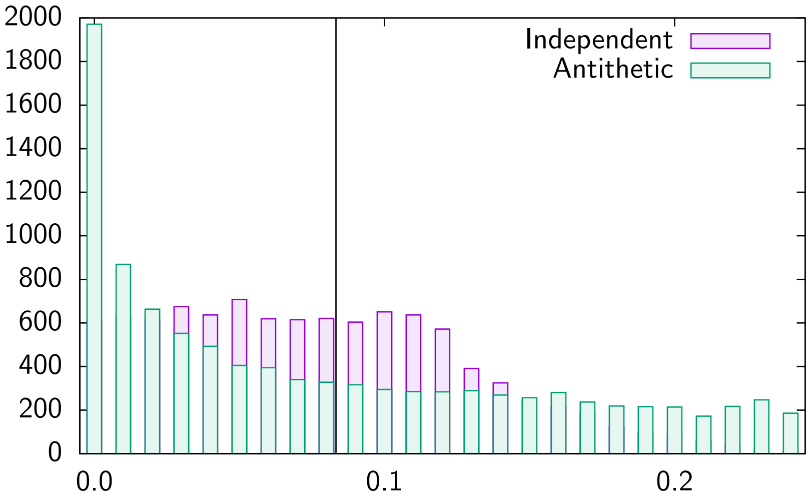
Another example



Another example



Another example



Outline

Control variates

- We use simulation to estimate $\theta = \mathrm{E}[X]$, where X is an output of the simulation
- Let Y be another output of the simulation, such that we know $\mathrm{E}[Y] = \mu$
- We consider the quantity:

$$Z = X + c(Y - \mu).$$

- By construction, $\mathrm{E}[Z] = \mathrm{E}[X]$
- Its variance is

$$\mathrm{Var}(Z) = \mathrm{Var}(X + cY) = \mathrm{Var}(X) + c^2 \,\mathrm{Var}(Y) + 2c\,\mathrm{Cov}(X, Y)$$

- Find c such that $\mathrm{Var}(Z)$ is minimum

Control variates

- First derivative:

$$2c\,\mathrm{Var}(Y) + 2\,\mathrm{Cov}(X, Y)$$

- Zero if

$$c^* = -\frac{\mathrm{Cov}(X, Y)}{\mathrm{Var}(Y)}$$

- Second derivative:

$$2\,\mathrm{Var}(Y) > 0$$

- We use

$$Z^* = X - \frac{\mathrm{Cov}(X, Y)}{\mathrm{Var}(Y)}(Y - \mu).$$

- Its variance

$$\mathrm{Var}(Z^*) = \mathrm{Var}(X) - \frac{\mathrm{Cov}(X, Y)^2}{\mathrm{Var}(Y)} \leq \mathrm{Var}(X)$$

Control variates

In practice...

- $\mathrm{Cov}(X, Y)$ and $\mathrm{Var}(Y)$ are usually not known.
- We can use their sample estimates:

$$\widehat{\mathrm{Cov}}(X, Y) = \frac{1}{n-1} \sum_{r=1}^{R} (X_r - \bar{X})(Y_r - \bar{Y})$$

and

$$\widehat{\mathrm{Var}}(Y) = \frac{1}{n-1} \sum_{r=1}^{R} (Y_r - \bar{Y})^2.$$

Control variates

In practice...

- Alternatively, use linear regression

$$X = aY + b + \varepsilon$$

  where $\varepsilon \sim N(0, \sigma^2)$.

- The least square estimators of a and b are

$$\hat{a} = \frac{\widehat{\mathrm{Cov}}(X,Y)}{\widehat{\mathrm{Var}}(Y)} = \frac{\sum_{r=1}^{R}(X_r - \bar{X})(Y_r - \bar{Y})}{\sum_{r=1}^{R}(Y_r - \bar{Y})^2}$$

$$\hat{b} = \bar{X} - \hat{a}\bar{Y}.$$

- Therefore

$$c^* = -\hat{a}$$

Control variates

- Moreover,

$$
\begin{aligned}
\hat{b} + \hat{a}\mu &= \bar{X} - \hat{a}\bar{Y} + \hat{a}\mu \\
&= \bar{X} - \hat{a}(\bar{Y} - \mu) \\
&= \bar{X} + c^*(\bar{Y} - \mu) \\
&= \widehat{\theta}
\end{aligned}
$$

- Therefore, the control variate estimate $\widehat{\theta}$ of θ is obtained by the estimated linear model, evaluated at μ.

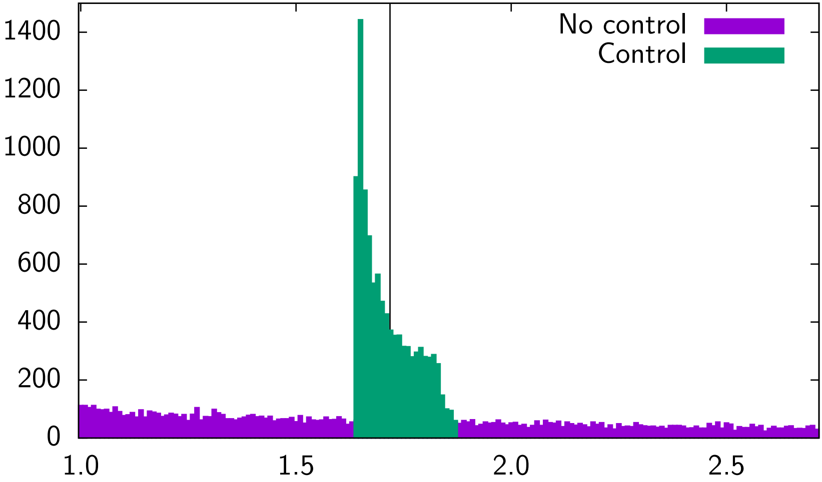
Back to the example

- Use simulation to compute $I = \int_0^1 e^x \, dx$
- $X = e^U$
- $Y = U$, $\mathrm{E}[Y] = 1/2$, $\mathrm{Var}(Y) = 1/12$
- $\mathrm{Cov}(X, Y) = (3 - e)/2 \approx 0.14$
- Therefore, the best c is

$$c^* = -\frac{\mathrm{Cov}(X, Y)}{\mathrm{Var}(Y)} = -6(3 - e) \approx -1.69$$

- Test with $R = 10'000$
- Result of the regression: $\hat{a} = 1.6893$, $\hat{b} = 0.8734$
- Estimate: $\hat{b} + \hat{a}/2 = 1.7180$, Variance: 0.003847 (compared to 0.24)

Back to the example

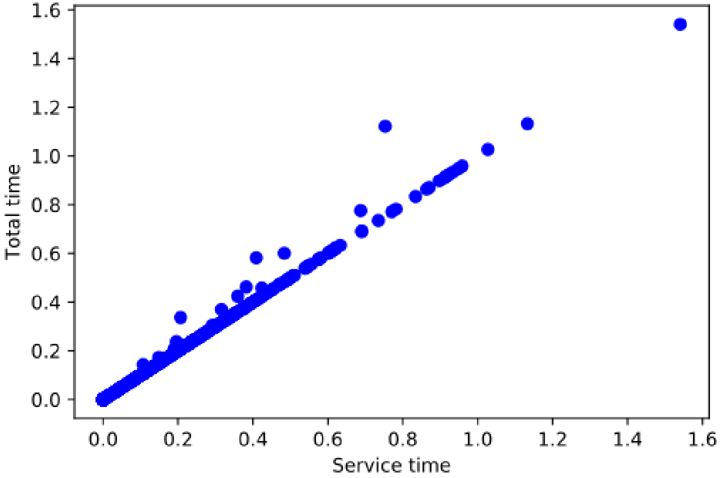


Satellite simulation

Variables

- X: average time spent by the customers in the bar.
- Y: average service time.

Satellite simulation



Satellite simulation

True value of $E[Y]$

- When customers are served, the average service time $\mu = 0.2$ is known.
- There are customers 63.2% of the days. Indeed, some days, the first customer arrives after the closure.
- Therefore,

$$E[Y] = 0.632\mu = 0.1264$$

Important

Do not use simulated values to calculate this quantity.

Satellite simulation

Scenario

- Mean service time: 0.2
- Mean inter-arrival time: 1
- Closure: 10.0
- $R = 1000$

Regular estimate

- $\mathrm{E}[\widehat{X}] = 0.133$
- $\mathrm{Var}[\widehat{X}] = 0.0336$

Control variates

- $\mathrm{E}[\widehat{X}] = 0.128$
- $\mathrm{Var}[\widehat{X}] = 0.000216$

Satellite simulation

Scenario

- Mean service time: 0.2
- Mean inter-arrival time: 1
- Closure: 10.0
- $R = 100000$

Regular estimate

- $\mathrm{E}[\widehat{X}] = 0.127$
- $\mathrm{Var}[\widehat{X}] = 0.0294$

Control variates

- $\mathrm{E}[\widehat{X}] = 0.128$
- $\mathrm{Var}[\widehat{X}] = 0.000261$

Comments

- The true value μ of the mean of the control variable Y must be available.
- Using the sample mean does **not** work.
- The higher the correlation between X and Y, the better.

Outline

Variance reductions techniques

Other techniques

- Conditioning
- Stratified sampling
- Importance sampling
- Draw recycling

In general

Correlation helps!