

Computational Statistics

Antithetic variables Section 5.2

Junqi Tang

Based on the lecture notes developed by Prof. Jinglai Li

Motivation: balancing simulated scenarios

Suppose a simulation output increases with a random input.

- An unusually large input often produces an unusually large output.
- A paired small input can produce a compensating small output.
- The pair average may fluctuate less than two independent outputs.

A simple construction

Draw $U \sim U(0, 1)$ and evaluate both $f(U)$ and $f(1 - U)$. Each input is uniform, but the inputs are dependent.

Example: a model driven monotonically by a scalar uncertain load.

Pairs and marginal distributions

Let (Y, Y') have the same marginal distribution as $f(X)$, where $X \sim p$. Assume

$$\mathbb{E}[Y] = \mathbb{E}[Y'] = \mathcal{I}, \quad \text{Var}(Y) = \text{Var}(Y') = \sigma^2 < \infty.$$

If $\sigma^2 > 0$, write $\rho = \text{Corr}(Y, Y')$.

The notes' definition

The pair is antithetic when the marginals agree and $\text{Cov}(Y, Y') < 0$.

Equal marginal distributions do **not** imply independence. Negative covariance concerns the **outputs being averaged**.

Uniform reflection

If $U \sim U(0, 1)$, set $V = 1 - U$. For $0 \leq t \leq 1$,

$$\mathbb{P}(V \leq t) = \mathbb{P}(U \geq 1 - t) = t.$$

Thus $V \sim U(0, 1)$ as well.

$$Y = f(U), \quad Y' = f(1 - U).$$

- The outputs have the same marginal distribution.
- $\text{Corr}(U, 1 - U) = -1$.
- The correlation of $f(U)$ and $f(1 - U)$ depends on f .

For an interval $[a, b]$, reflect X to $a + b - X$.

Independent pairs

Generate independent pairs

$$(Y_1, Y'_1), \dots, (Y_n, Y'_n).$$

For uniform reflection, draw $U_1, \dots, U_n \stackrel{\text{iid}}{\sim} U(0, 1)$ and set

$$Y_i = f(U_i), \quad Y'_i = f(1 - U_i), \quad A_i = \frac{Y_i + Y'_i}{2}.$$

The independent observation unit

$A_1, \dots, A_n$ are independent and identically distributed. The $2n$ individual outputs generally are not independent.

n pairs use $2n$ function evaluations in the ordinary implementation.

The antithetic estimator

Define the antithetic Monte Carlo estimator by

$$\widehat{\mathcal{I}}_A = \frac{1}{n} \sum_{i=1}^n A_i = \frac{1}{2n} \sum_{i=1}^n (Y_i + Y'_i).$$

For a transformation T preserving p ,

$$X_i \stackrel{\text{iid}}{\sim} p, \quad Y_i = f(X_i), \quad Y'_i = f(T(X_i)).$$

In the notes' notation, $g(x) = f(T(x))$.

Contrast with Section 5.1

Here $g(X)$ has the same unknown mean as $f(X)$. No separately known control mean is required.

Unbiasedness

Assume $\mathbb{E}|Y| < \infty$ and $Y' \stackrel{d}{=} Y$. Then

$$\mathbb{E}[A_i] = \frac{1}{2}\{\mathbb{E}[Y_i] + \mathbb{E}[Y'_i]\} = \mathcal{I}.$$

Consequently,

$$\mathbb{E}[\widehat{\mathcal{I}}_A] = \frac{1}{n} \sum_i \mathbb{E}[A_i] = \mathcal{I}.$$

- Dependence within a pair does not invalidate linearity of expectation.
- Negative covariance is needed for the variance benefit, not for unbiasedness.
- Incorrect partner marginals can change the expectation.

Variance of a pair average

Assume equal marginal variance σ^2 and covariance $\rho\sigma^2$.

$$\begin{aligned}\text{Var}(A_i) &= \frac{1}{4} \text{Var}(Y_i + Y'_i) \\ &= \frac{1}{4} \{\sigma^2 + \sigma^2 + 2\rho\sigma^2\} \\ &= \frac{\sigma^2}{2}(1 + \rho).\end{aligned}$$

Since different pairs are independent,

$$\text{Var}(\hat{\mathcal{I}}_A) = \frac{\sigma^2}{2n}(1 + \rho).$$

The covariance term is retained within each pair and vanishes across independent pairs.

A fair comparison at equal evaluations

With $2n$ independent evaluations, ordinary MC has

$$\text{Var}(\hat{\mathcal{I}}_{\text{MC},2n}) = \frac{\sigma^2}{2n}.$$

Therefore

$$\frac{\text{Var}(\hat{\mathcal{I}}_A)}{\text{Var}(\hat{\mathcal{I}}_{\text{MC},2n})} = 1 + \rho.$$

- $\rho < 0$: antithetic sampling reduces variance.
- $\rho = 0$: no variance gain at equal evaluations.
- $\rho > 0$: antithetic sampling increases variance.

Compared with only n MC evaluations, the ratio is $(1 + \rho)/2$, but the evaluation budgets differ.

Correlation and computational cost

At equal evaluations, the variance fraction is $1 + \rho$ and the SD fraction is $\sqrt{1 + \rho}$.

ρ	Variance fraction	SD fraction
-0.5	0.5	0.707
-0.9	0.1	0.316
-1	0	0

If one ordinary output costs c and a pair costs c_A , then at a large fixed budget the variance ratio is approximately

$$\frac{c_A}{2c}(1 + \rho).$$

Include partner evaluation, random generation, and any reusable model calculations.

When uniform reflection helps

Suppose f is monotone on $[0, 1]$ and $\mathbb{E}[f(U)^2] < \infty$.

- If f increases, $f(u)$ increases while $f(1 - u)$ decreases.
- If f decreases, the directions reverse.

In either case,

$$\text{Cov}(f(U), f(1 - U)) \leq 0.$$

A sufficient condition

Monotonicity guarantees nonpositive covariance. Strict improvement requires strictly negative covariance.

For a nonmonotone function, inspect or analyse the output covariance.

Why opposite monotonicity gives nonpositive covariance

Let U' be an independent copy of U . For square-integrable $a(U)$, $b(U)$,

$$2 \operatorname{Cov}(a(U), b(U)) = \mathbb{E}[(a(U) - a(U'))(b(U) - b(U'))].$$

Expanding the right side gives two copies of $\mathbb{E}[ab]$ and two of $-\mathbb{E}[a]\mathbb{E}[b]$.

- If a increases and b decreases, the two differences have opposite signs.
- Their product is therefore nonpositive for every (U, U') .

Set $a(u) = f(u)$ and $b(u) = f(1 - u)$ to obtain the reflection result.

Square example: mean and variance

Let $U \sim U(0, 1)$, $Y = U^2$ and $Y' = (1 - U)^2$. The uniform density is $p(u) = 1$ on $[0, 1]$, so

$$\mathcal{I} = \mathbb{E}[Y] = \int_0^1 u^2 du = \left[\frac{u^3}{3} \right]_0^1 = \frac{1}{3},$$

$$\mathbb{E}[Y^2] = \mathbb{E}[U^4] = \int_0^1 u^4 du = \left[\frac{u^5}{5} \right]_0^1 = \frac{1}{5}.$$

Therefore

$$\text{Var}(Y) = \mathbb{E}[Y^2] - (\mathbb{E}[Y])^2 = \frac{1}{5} - \frac{1}{9} = \frac{9 - 5}{45} = \frac{4}{45}.$$

Because $1 - U$ is also uniform, Y' has the same distribution as Y :

$$\mathbb{E}[Y'] = \frac{1}{3}, \quad \text{Var}(Y') = \frac{4}{45}.$$

Square example: expectation of the product

Both outputs use the *same* U . Thus

$$YY' = U^2(1 - U)^2 = U^2 - 2U^3 + U^4.$$

Integrating term by term,

$$\begin{aligned}\mathbb{E}[YY'] &= \int_0^1 (u^2 - 2u^3 + u^4) du \\ &= \left[\frac{u^3}{3} - \frac{u^4}{2} + \frac{u^5}{5} \right]_0^1 \\ &= \frac{1}{3} - \frac{1}{2} + \frac{1}{5} = \frac{10 - 15 + 6}{30} = \boxed{\frac{1}{30}}.\end{aligned}$$

Dependence matters

We cannot replace $\mathbb{E}[YY']$ by $\mathbb{E}[Y]\mathbb{E}[Y'] = 1/9$. The paired outputs are dependent.

Square example: covariance and correlation

Using the moments just calculated,

$$\begin{aligned}\text{Cov}(Y, Y') &= \mathbb{E}[YY'] - \mathbb{E}[Y]\mathbb{E}[Y'] \\ &= \frac{1}{30} - \frac{1}{3} \frac{1}{3} = \frac{3}{90} - \frac{10}{90} = \boxed{-\frac{7}{90}}.\end{aligned}$$

Correlation divides covariance by the product of the standard deviations:

$$\begin{aligned}\rho &= \frac{\text{Cov}(Y, Y')}{\sqrt{\text{Var}(Y)}\sqrt{\text{Var}(Y')}} \\ &= \frac{-7/90}{\sqrt{4/45}\sqrt{4/45}} = \frac{-7/90}{4/45} = -\frac{7}{90} \frac{45}{4} = \boxed{-\frac{7}{8}}.\end{aligned}$$

The inputs have correlation -1 , but the squared outputs have correlation $-7/8$.

Exact variance reduction in the square example

The pair average is

$$A = \frac{U^2 + (1 - U)^2}{2} = U^2 - U + \frac{1}{2}.$$

Using $\rho = -7/8$,

$$\text{Var}(A) = \frac{4/45}{2} \left(1 - \frac{7}{8}\right) = \frac{1}{180}, \quad \text{Var}(\hat{\mathcal{I}}_A) = \frac{1}{180n}.$$

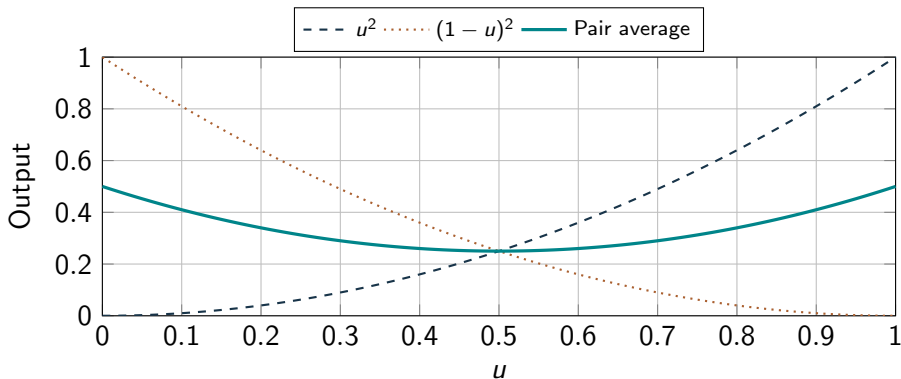
With $2n$ independent MC evaluations,

$$\text{Var}(\hat{\mathcal{I}}_{\text{MC},2n}) = \frac{4/45}{2n}.$$

At equal evaluations

Antithetic sampling reduces variance by a factor of 8 and standard deviation by a factor of $\sqrt{8} \approx 2.83$.

Visualising cancellation



The pair average ranges from 1/4 to 1/2, while each individual output ranges from 0 to 1. Curves are analytic functions, not simulated data.

What is Monte Carlo standard error?

MCSE = Monte Carlo standard error.

- A fresh set of random draws gives a different estimate $\hat{\mathcal{I}}$.
- MCSE describes how much that estimate varies across repeated runs of the same algorithm with the same sample size.

$$\text{MCSE}(\hat{\mathcal{I}}) = \sqrt{\text{Var}(\hat{\mathcal{I}})}.$$

Uncertainty about the computed average

MCSE measures simulation variability. It is not the realised error $|\hat{\mathcal{I}} - \mathcal{I}|$, which usually cannot be observed.

For an unbiased estimator, theoretical MCSE equals RMSE. Bias and model error need separate assessment.

Sample standard deviation versus MCSE

For independent, identically distributed outputs $Y_1, \dots, Y_N$, with finite variance σ_Y^2 ,

$$\hat{\mathcal{I}} = \bar{Y}, \quad \text{Var}(\bar{Y}) = \frac{\sigma_Y^2}{N}.$$

Estimate the output variance using

$$s_Y^2 = \frac{1}{N-1} \sum_{i=1}^N (Y_i - \bar{Y})^2, \quad \boxed{\widehat{\text{MCSE}} = \frac{s_Y}{\sqrt{N}}}.$$

Quantity	What varies?
s_Y	Individual outputs within a run
$s_Y/\sqrt{N}$	The estimated average across repeated runs

Neither definition requires Gaussian outputs.

Where does the number 1.96 come from?

Let $Z \sim N(0, 1)$ and write $\Phi(z) = \mathbb{P}(Z \leq z)$. For a central 95% interval, the remaining 5% is split equally:

$$\mathbb{P}(Z < -z) = 0.025, \quad \mathbb{P}(Z > z) = 0.025.$$

Hence the upper cutoff satisfies

$$\Phi(z) = 1 - 0.025 = 0.975,$$

$$z = \Phi^{-1}(0.975) = 1.959963 \dots \approx 1.96.$$

A standard normal probability statement

$$\mathbb{P}(-1.96 \leq Z \leq 1.96) \approx 0.95.$$

Find 0.975 in a cumulative normal table, or use `scipy.stats.norm.ppf(0.975)`.

From the normal quantile to a confidence interval

For iid outputs with finite, positive variance, the studentised CLT gives

$$\frac{\hat{\mathcal{I}} - \mathcal{I}}{s_Y/\sqrt{N}} \xrightarrow{d} N(0, 1).$$

Therefore, for sufficiently large N ,

$$\mathbb{P}\left(-1.96 \leq \frac{\hat{\mathcal{I}} - \mathcal{I}}{\widehat{\text{MCSE}}} \leq 1.96\right) \approx 0.95.$$

Rearrange the inequalities to obtain

$$\boxed{\hat{\mathcal{I}} - 1.96\widehat{\text{MCSE}} \leq \mathcal{I} \leq \hat{\mathcal{I}} + 1.96\widehat{\text{MCSE}}.}$$

In general, a two-sided $(1 - \alpha)$ interval uses $\Phi^{-1}(1 - \alpha/2)$. For normal iid outputs with unknown variance, a Student t_{N-1} quantile gives an exact interval. Here the normal interval is a large-sample approximation.

Reading an estimate and its MCSE

Illustration: $N = 1000$ independent outputs give $\hat{\mathcal{I}} = 0.340$ and $s_Y = 0.300$.

$$\widehat{\text{MCSE}} = \frac{0.300}{\sqrt{1000}} \approx 0.00949.$$

With enough observations for the central limit approximation,

$$\hat{\mathcal{I}} \pm 1.96 \widehat{\text{MCSE}} \approx 0.340 \pm 0.0186 = [0.3214, 0.3586].$$

An approximate 95% confidence interval

Across repeated runs, about 95% of intervals constructed this way would contain the fixed target, when the approximation is accurate.

Finite, positive variance is assumed. Strong skewness or rare events may require much larger samples. These numbers are illustrative.

Precision and the independent sampling unit

For iid averaging with the same output distribution,

$$\text{MCSE}_N = \frac{\sigma_Y}{\sqrt{N}}.$$

Desired MCSE	Required sample count
Half as large	$4N$
One tenth as large	$100N$

Why antithetic sampling needs care

The two outputs within a pair are dependent. Form each pair average $A_i = (Y_i + Y'_i)/2$ first. The n independent observations are $A_1, \dots, A_n$.

Reducing the variance of these independent units reduces MCSE.

MCSE uses the pair averages

Compute the n independent quantities $A_i = (Y_i + Y'_i)/2$.

$$\hat{\mathcal{I}}_A = \bar{A}, \quad s_A^2 = \frac{1}{n-1} \sum_{i=1}^n (A_i - \bar{A})^2.$$

$$\widehat{\text{MCSE}} = \frac{s_A}{\sqrt{n}}.$$

For large n under a finite, positive pair variance, an approximate 95% interval is

$$\bar{A} \pm 1.96 \frac{s_A}{\sqrt{n}}.$$

The number in the square root is the number of independent pairs.

Why pooling all outputs gives the wrong MCSE

The $2n$ individual outputs are dependent within pairs. Treating them as iid gives a variance estimate close to

$$\frac{\sigma^2}{2n},$$

while the true estimator variance is

$$\frac{\sigma^2}{2n}(1 + \rho).$$

- Negative correlation: the naive iid MCSE overstates the asymptotic error.
- Positive correlation: it understates the asymptotic error.
- Using pair averages handles both cases correctly.

Other distribution-preserving constructions

Symmetric normal input

If $Z \sim N(0, 1)$, then $-Z \sim N(0, 1)$. Evaluate $f(Z)$ and $f(-Z)$. For $X \sim N(\mu, \sigma^2)$, reflect to $2\mu - X$.

Inverse CDF construction

For a distribution with generalised inverse CDF F^{-1} ,

$$X = F^{-1}(U), \quad X' = F^{-1}(1 - U), \quad U \sim U(0, 1).$$

Both marginals have CDF F . For monotone scalar f , output covariance is nonpositive when second moments exist.

In code, guard against endpoint uniforms when inverse CDFs are infinite there.

A counterexample: reflection can waste evaluations

Let $U \sim U(0, 1)$ and $f(u) = (u - 1/2)^2$. Then

$$f(1 - u) = f(u).$$

The outputs are identical, so $\rho = 1$ and $A = f(U)$.

$$\text{Var}(\hat{\mathcal{I}}_A) = \frac{\sigma^2}{n}, \quad \text{Var}(\hat{\mathcal{I}}_{\text{MC}, 2n}) = \frac{\sigma^2}{2n}.$$

At equal evaluations

This reflection construction has twice the variance of ordinary MC. It is a candidate pairing, not an antithetic pair under the strict definition.

Likewise, $f(z) = z^2$ gives identical outputs at z and $-z$.

In-class exercise: perfect cancellation

Let $U \sim U(0, 1)$ and $f(u) = a + bu$ for fixed a, b .

1. Find the target $\mathcal{I} = \mathbb{E}[f(U)]$.
2. Compute $A = \{f(U) + f(1 - U)\}/2$.
3. What is $\text{Var}(\hat{\mathcal{I}}_A)$?

Solution:

$$\mathcal{I} = a + \frac{b}{2}, \quad A = \frac{a + bU + a + b(1 - U)}{2} = a + \frac{b}{2}.$$

Thus the estimator is exact for every draw. If $b \neq 0$, $\rho = -1$. If $b = 0$, both outputs are constant and correlation is undefined.

Implementation checklist

1. Check that each partner input has the correct distribution.
2. Check the covariance of the *outputs*, not only the inputs.
3. Generate independent base randomness across pairs.
4. Form one average per pair.
5. Report its sample mean and MCSE using n independent pairs.
6. Compare methods at equal evaluation counts or measured runtime.

Debugging

Use a benchmark with a known answer, repeat independent runs, and compare empirical error with the reported MCSE.

Lecture summary

For n independent pairs with equal marginal variance σ^2 ,

$$\hat{\mathcal{I}}_A = \frac{1}{n} \sum_i \frac{Y_i + Y'_i}{2}, \quad \text{Var}(\hat{\mathcal{I}}_A) = \frac{\sigma^2}{2n}(1 + \rho).$$

- Correct marginal expectations give unbiasedness.
- Negative output covariance reduces variance at equal evaluations.
- Uniform reflection works for monotone scalar integrands.
- Reflection can fail for nonmonotone integrands.
- MCSE is computed from the independent pair averages.