Standard Errors with Antithetic Sampling

Nicholas R. Kirchner

July 18, 2013

1 Introduction

There was a discussion as to whether to divide by N (the number of pairs generated) or $2N$ (the total number of values) when computing the standard error of a quantity estimated through antithetic variance reduction. From a theoretical perspective, the correct answer differs from both of these.

In fact, if the two values x_i and x_i' generated via antithetic sampling are negatively correlated, this method turns out to be even more powerful than dividing by $2N$ from the perspective of reducing standard error.

The reason for the reduced standard error is that the covariance between the two values should be added before division by $2N$, and therefore a negative covariance results in a lowered standard error.

From a computational perspective, it is not necessary to compute any covariances. Rather, it is easiest to simply make a list of the quantities $\frac{x_i+x_i'}{2}$ and compute the mean and standard error of that list.

2 Properties of Variance

If X and Y are random variables and a a constant, we have the following properties:

$$\mathrm{Var}(aX) = a^2\mathrm{Var}(X)$$

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

In particular, when X and Y are independent we have $\mathrm{Cov}(X,Y) = 0$, so

$$\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$$

3 The Usual Case

First, back to basics. Our usual mean estimator is

$$\hat{\mu} = \frac{1}{N}\sum_{i=1}^{N} x_i$$

where N is the number of data points, and the x_i are the data points drawn from random variables X_i. If we presume that the X_i are I.I.D. (independent and identically distributed), we can calculate

$$\begin{aligned}\mathrm{Var}(\hat{\mu}) &= \mathrm{Var}\left(\frac{1}{N}\sum_{i=1}^{N} X_i\right)\\ &= \frac{1}{N^2}\mathrm{Var}\left(\sum_{i=1}^{N} X_i\right)\end{aligned}$$

Here's where independence comes in: we needn't worry about covariance.

$$= \frac{1}{N^2}\sum_{i=1}^{N}\mathrm{Var}(X_i)$$

$$= \frac{1}{N}\mathrm{Var}(X_i)$$

The last step is justified by noting that the X_i are identically distributed, so that all the $\mathrm{Var}(X_i)$ are equal. Taking a square root yields the standard error.

4 Correlated Pairs

4.1 Finding the Standard Error

When we estimate price using antithetic variance reduction, we are generating estimates of the price (or delta, or gamma, etc.) in pairs. Let x_i denote the price obtained from the ith path actually generated, and x_i' denote the price obtained from the corresponding anti-correlated path. X_i will denote the random variable from which x_i is drawn, and X_i' will denote the random variable from which x_i' is drawn.

We have the following properties:

1. X_i and X_i' are not independent.

2. Each pair X_i,X_i' is independent of any other pair X_j, X_j'.

3. All of the random variables here have the same distribution.

Our estimator for the mean is

$$\hat{\mu} = \frac{1}{2N}\sum_{i=1}^{N} x_i + x_i'$$

where N is the number of pairs of values generated (so that the total number of values considered here is $2N$). This remains an unbiased estimator. Its variance is given by

$$\mathrm{Var}(\hat{\mu}) = \mathrm{Var}\left(\frac{1}{2N}\sum_{i=1}^{N} X_i + X_i'\right)$$

$$= \mathrm{Var}\left(\frac{1}{N}\sum_{i=1}^{N}\frac{X_i + X_i'}{2}\right)$$

$$= \frac{1}{N^2}\mathrm{Var}\left(\sum_{i=1}^{N}\frac{X_i + X_i'}{2}\right)$$

Now, the random variables $X_i + X_i'$ are I.I.D., which gives us

$$= \frac{1}{N^2}\sum_{i=1}^{N}\mathrm{Var}\left(\frac{X_i + X_i'}{2}\right)$$

$$= \frac{1}{N}\mathrm{Var}\left(\frac{X_i + X_i'}{2}\right)$$

(As before, all the items in the sum are the same.) The square root of this is the standard error. This implies that the standard error for antithetic estimation should be calculated like so:

1. Make a list of the quantities $\frac{x_i + x_i'}{2}$ (no need to keep track of x_i or x_i')
2. Calculate the average of that list, which gives you the mean
3. Calculate the variance of that list and divide by its length
4. Take a square root of the previous step to get the standard error.

4.2 Analysis of Result

Let us now analyze the effectiveness of this method.

$$\mathrm{Var}(\hat{\mu}) = \frac{1}{N}\mathrm{Var}\left(\frac{X_i + X_i'}{2}\right)$$

$$= \frac{1}{4N}\mathrm{Var}(X_i + X_i')$$

X_i and X_i' are not independent. Therefore,

$$\mathrm{Var}(\hat{\mu}) = \frac{1}{4N}\left(\mathrm{Var}(X_i) + \mathrm{Var}(X_i') + 2\mathrm{Cov}(X_i, X_i')\right)$$

X_i and X_i' are identically distributed, so their variances are equal. Thus,

$$\mathrm{Var}(\hat{\mu}) = \frac{1}{4N}\left(2\mathrm{Var}(X_i) + 2\mathrm{Cov}(X_i, X_i')\right)$$

$$= \frac{1}{2N}\left(\mathrm{Var}(X_i) + \mathrm{Cov}(X_i, X_i')\right)$$

(And we get the standard error by taking the square root of this.)

Now, we can be confident that our covariance is negative. If one member of our pair is high, we would strongly suspect that the other is low. That makes antithetic variance reduction powerful. Not only do we get to divide the variance by $2N$, we also get to subtract a bit off it first. Antithetic variance reduction as presented in class on N pairs of prices actually results in a lower standard error than straight random sampling on $2N$ data points!

For example, if X_i and X_i' are prices generated with the inputs $S = 50$, $K = 50$, $r = 0.05$, $\sigma = 0.5$ and $T = 1$, it happens that $\text{Cov}(X_i, X_i') \approx -\frac{1}{4}\text{Var}(X_i)$. (Note: this covariance will vary depending on the inputs). If we generate N pairs, we will get

$$\text{Var}(\hat{\mu}) \approx \frac{3}{8}\frac{\text{Var}(X_i)}{N}$$

$$\text{Var}(\hat{\mu}_{antithetic}) \approx \frac{3}{8}\text{Var}(\hat{\mu}_{normal})$$

For these particular inputs, we would therefore need only 3/8 as many random numbers generated in the antithetic case compared to the usual case in order to match standard errors. To put it another way, if we generate 1,000,000 pairs via the antithetic method, we'd need roughly 2,700,000 simulations by the usual method to get the same standard error.

5 Appendix: The European Option Price Distribution

When it comes to our particular case, we generate our two prices by first generating two numbers from a standard normal distribution, with the second being the negative of the first. Let's notate these as ϵ_i and ϵ_i', with $\epsilon_i' = -\epsilon_i$. We then compute the prices X_i and X_i' based on ϵ_i and ϵ_i'.

The correlation coefficient between two random variables X and Y is defined as

$$\rho_{XY} = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}}$$

A famous theorem (Cauchy-Schwarz-Bunyakovsky) asserts that this quantity is always between -1 and 1. Further, it will be equal to ± 1 if and only if the $Y = cX$ for some constant $c \neq 0$. (This is to say that whenever we simultaneously sample a value from X and a value from Y, the ratio of the two samples will always be c.)

If X and X' are identically distributed then their variances are equal, so that

$$\rho_{XX'} = \frac{\text{Cov}(X, X')}{\text{Var}(X)}$$

We therefore have

$$\rho_{\epsilon_i \epsilon_i'} = -1$$

However, for the inputs $S = 50$, $K = 50$, $r = 0.05$, $\sigma = 0.5$ and $T = 1$, I empirically found

$$\rho_{X_i X_i'} \approx -\frac{1}{4}$$

One might be a bit confused by this. After all, if we know ϵ_i, we can say definitively and with no effort what ϵ_i' is. We get X_i and X_i' with these values. In this sense, X_i and X_i' are completely dependent, just as ϵ_i and ϵ_i' are. Why, then, are the respective correlation coefficients different?

The answer to the question is that correlation is not measuring dependence in general. It's measuring linearity, which is only one specific kind of dependence.

To illustrate the point, we can sample ϵ_i and ϵ_i' from a standard normal as shown in figure 1. If we make many such samples and plot the ordered pairs $(\epsilon_i, \epsilon_i')$, we obtain the graph shown in figure 2. The relation $\epsilon_i' = -\epsilon_i$ gives us a perfectly straight line.

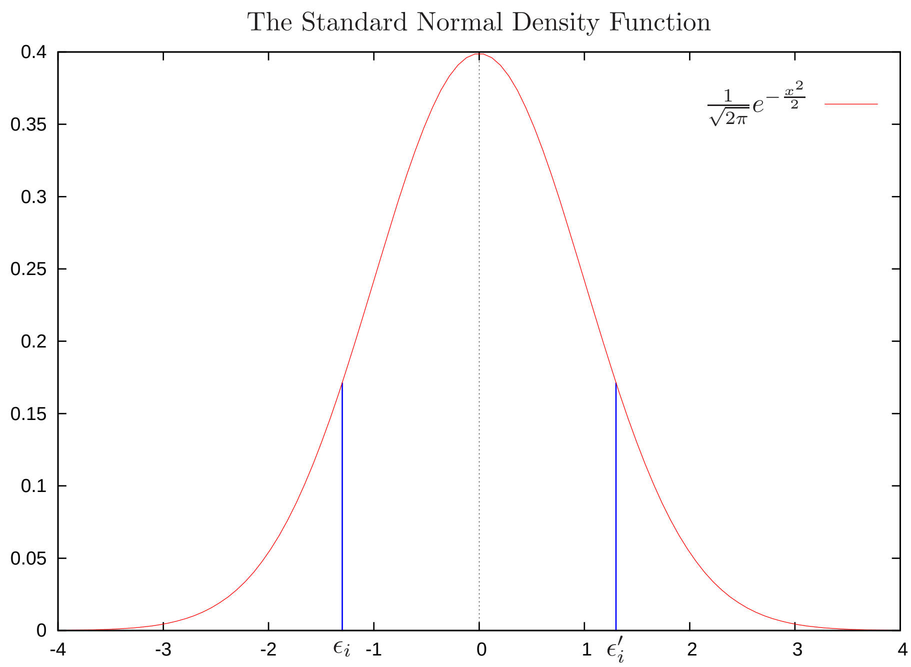

Figure 1: Antithetic samples from the standard normal

Now, we use ϵ_i and ϵ_i' to obtain option values X_i and X_i' for an option with $S = 50$, $K = 20$ (note the change), $r = 0.05$, $\sigma = 0.5$ and $T = 1$. The X_i and X_i' are shown in figure 3.

If we do many such samples and plot the ordered pairs (X_i, X_i'), we obtain the plot in figure 4. Note that while there is clearly much dependence between X_i and X_i', it is not all linear dependence.

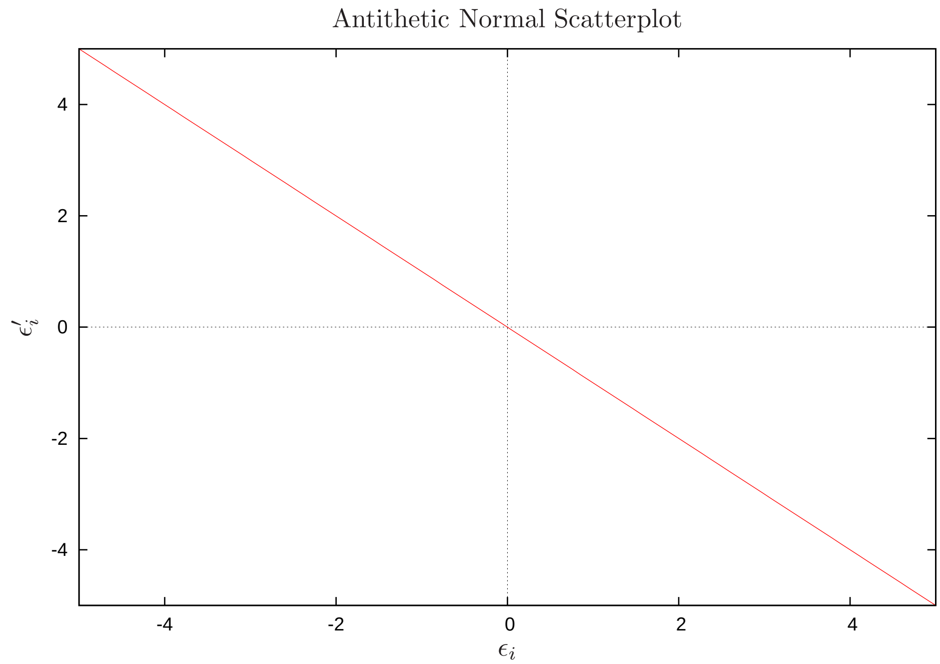

Figure 2: Relation between antithetic samples from the standard normal

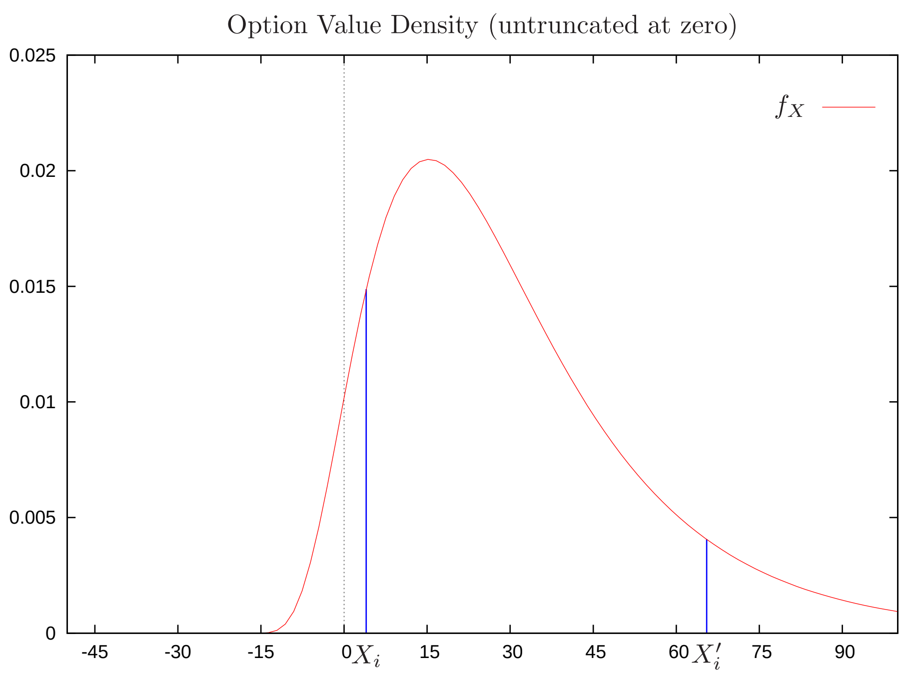

Figure 3: Distribution of possible option values (K = 20). Note that about 4.62% of the area is to the left of $x = 0$.

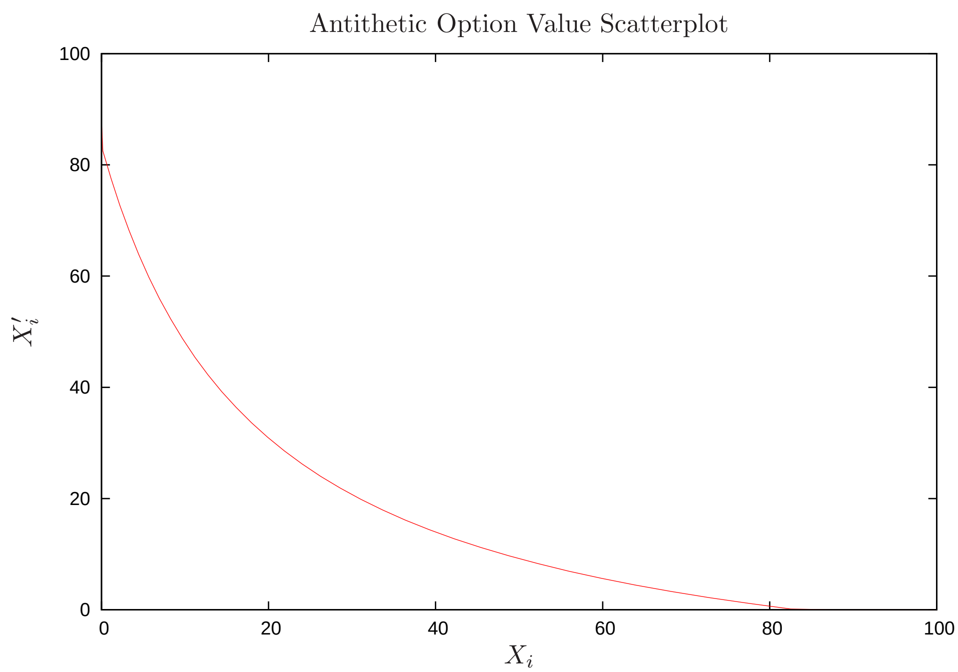

Figure 4: Relation between antithetic samples for the option price