Volatility Estimation via First Exit Times

Chris Merrill* Euan Sinclair**

This Draft: 12/15/2014

Abstract

Estimation of volatility is important for many financial applications. The most common methods are based on a time series analysis, either sampling prices at regular intervals or using extreme values from within the time intervals. We examine an alternative, where we use the exit times from a price corridor. This method converges more quickly to the true volatility and is a natural fit for high frequency traders who continuously monitor price series.

JEL Classification Numbers: C00, C53, G00, G1, G10, E47

Keywords: Volatility, Risk, High frequency trading

* Chris Merrill is in the mathematics department of the University of Chicago.

**Euan Sinclair is risk manager at Bluefin Trading. Email: esinclair@bluefintrading.com

Introduction

Estimation of volatility is a central problem in finance and has been studied extensively over the past several decades. When using historical underlying data to estimate volatility it is most common to form a time series then study either the closing prices of each period or a selection of points from within each time period such as the opening price, high price, low price and closing price. A good survey of these time series methods is Poon, 2005.

The simplest example is the close-to-close estimator. At fixed time intervals (monthly, daily, hourly etc) we take the closing price, S_t and form a series of logarithmic returns

$$r_t = log(S_t/S_{t-1}) \tag{1}$$

These are a series of random returns sampled at fixed times. This sampling scheme is illustrated in Figure One.

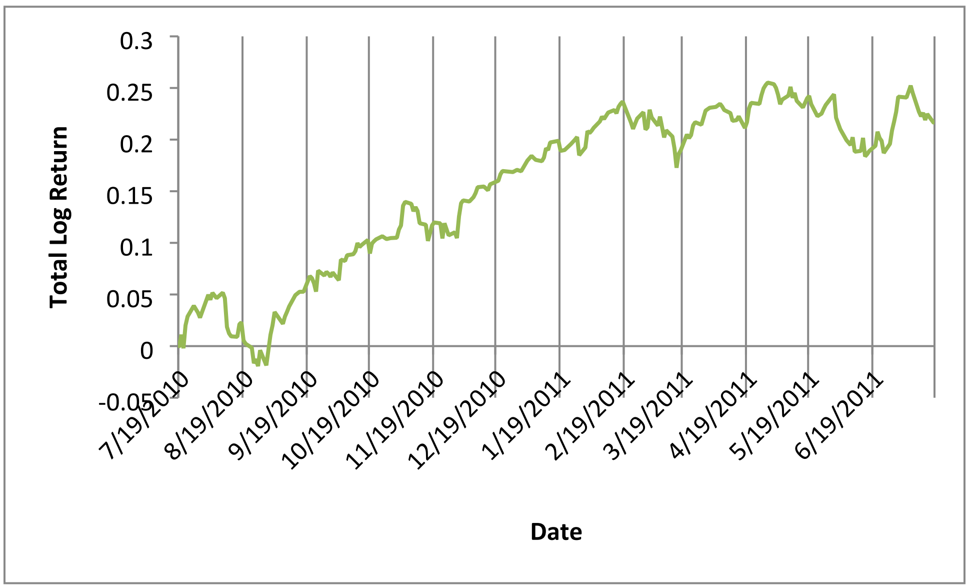

Figure One: The logarithmic returns for SPY sampled once a month from July 2010 until July 2011.

In this example we have a sequence of monthly logarithmic returns given by 0.005, 0.062, 0.022, 0.030, 0.040, 0.029, 0.048, -0.047, 0.027, 0.025, -0.048, 0.022.

Volatility is then calculated as

$$\sigma = \sqrt{\frac{1}{N-1} \sum_{i=1}^{N} (r_t - \tilde{r})^2} \tag{2}$$

This method has many strengths. It is simple conceptually and computationally. It has well understood biases and sampling properties. It is the industry standard, implemented in every analysis software package. However, it is a rather inefficient way to use the available data. Even if we slice the data in finer and finer increments most data is not used at all.

In this note we will turn the estimation problem on its side. Instead of asking, "how far did the price move?" we will ask, "how fast did the price move?"

The First Exit Time Method

This alternative viewpoint was first studied by Cho and Frees, 1988.

Define a logarithmic price corridor by forming a double barrier, Δ up and Δ down from the initial spot price. When the barrier is touched, we note the exit time, $\tau 1$, and reset the barrier around the current price. This generates a sequence of exit times $\tau 1$, $\tau 2$,… τn from which we want to estimate volatility.

The choice of Δ here is analogous to the choice of the fixed time interval in the close-to-close method, but our series now consists of a random sequence of times when the price has moved by a fixed amount. This contrast is illustrated in Figure Two, which also shows why it is apt to think of the first exit time method as "turning close to close on its side".

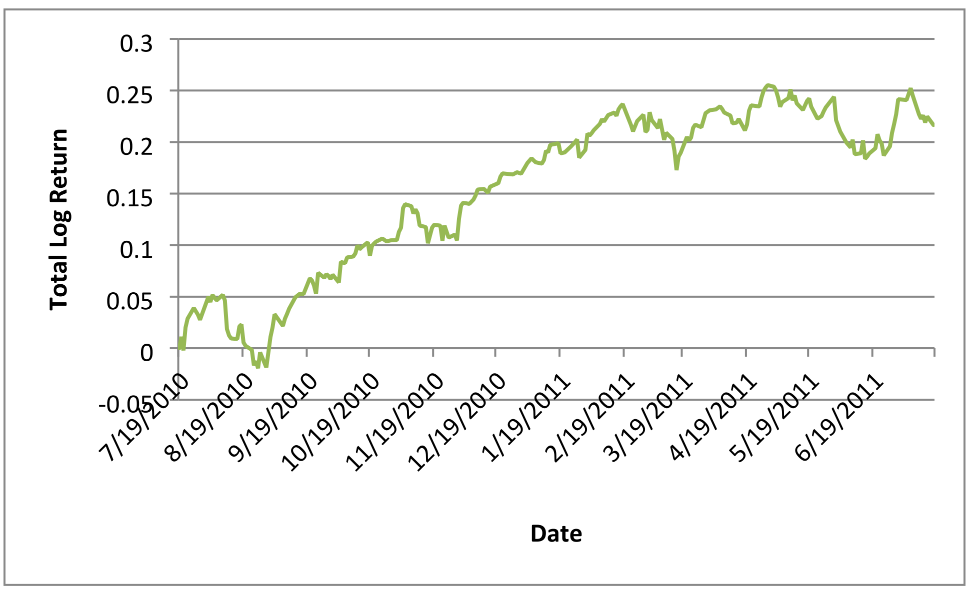

Figure Two: The logarithmic returns for SPY sampled each time it crosses a corridor of 0.05.

In this example we have a sequence of crossing times (in days) of 13, 14, 17, 24, 39, 32, 65, 27, 23.

We note that the data need not necessarily consist of trade prices. The data could also include quote data, especially if the security is continuously quoted but infrequently traded.

Assuming Brownian motion for the log-price process, and making the standard assumption that the drift rate is negligible for typical τ, we can derive the elegant result:

$$E[\tau] = \frac{\Delta^2}{\sigma^2} \quad (3)$$

(For details of the derivation refer to Borodin and Salminen, 2002.)

This can be inverted to obtain the following estimate for the volatility of the security.

$$\sigma = \frac{\Delta}{\sqrt{E[\tau]}} \quad (4)$$

where $E[\tau]$ is the population mean, which we don't know; we can only make estimates based on the observed data.

So we observe n first exit times and calculate a sample mean $\bar{\tau}$. We could naively substitute this sample mean for $E[\tau]$ in (4), but that will lead to a biased estimate for σ unless n is very large. The true value of σ will be slightly lower than the naïve estimate due to Jensen's inequality. Because $\bar{\tau}$ is a random variable with mean $E[\tau]$, and $1/\sqrt{\bullet}$ is convex in its argument:

$$\sigma = \Delta/\sqrt{E[\tau]} < E\left[\Delta/\sqrt{\bar{\tau}}\right] \quad (5)$$

Therefore, we need to know something about the sampling distribution of τ to calculate a Jensen-corrected volatility estimate.

From Borodin and Salminen:

$$Var(\tau) = \frac{2\Delta^2}{3\sigma^2} = \frac{2}{3}E[\tau]^2 \quad (6)$$

If $\bar{\tau}$ is the mean of the sample of the n first exit times we have, by the Central Limit Theorem,

$$E[\bar{\tau}] = E[\tau] \quad (7)$$

$$Var(\bar{\tau}) = \frac{Var(\tau)}{n} \quad (8)$$

Our goal is to derive an unbiased estimator of the population variance .To do this we define a new random variable $\delta\tau$ by

$$\delta\tau = \bar{\tau} - E[\tau] \quad (9)$$

This is a random variable with zero mean and a variance, $Var(\tau)/n$.
Now define a function, f.

$$f(\tau) = \frac{\Delta}{\sqrt{\tau}} \tag{10}$$

We can solve for the true population volatility, σ, by expanding our biased sample volatility, $E[f(\bar{\tau})]$, to second order to capture the variance term:

$$E[f(\bar{\tau})] = E[f(E[\tau] + \delta\tau)]$$
$$= E\left[f(E[\tau]) + f'(E[\tau])\delta\tau + \frac{1}{2}f''(E[\tau])\delta\tau^2 + \dots\right]$$
$$= f(E[\tau]) + \frac{1}{2}f''(E[\tau])E[\delta\tau]^2$$
$$= \sigma + \frac{1}{2}\frac{3}{4}\frac{\sigma}{E[\tau]^2}Var(\bar{\tau})$$
$$= \sigma\left(1 + \frac{1}{4n}\right) \tag{11}$$

So for a single observation we need to correct our observed volatility by multiplying by 4/5. This is a substantial correction but the bias decreases quickly as *n* increases. This convergence is illustrated in Table One.

n	Bias Factor
1	0.8
5	0.952
10	0.976
50	0.995

Table One: The bias correction factor for the sample volatility as a function of the number of observations.

Efficiency

It is difficult to consistently compare the barrier estimator to the close-to-close estimator in terms of efficiency. The barrier estimator generally uses a different amount of information; indeed it is a natural fit for online estimation which can be viewed as an almost infinitely large information set. However, most traders are now sitting in front of computers with this data streaming past their eyes anyway. Why not use all of it?

To get a visual comparison of the relative convergence speeds we simulated 100,000 20 day paths of a 0.30 volatility stock, with observation barriers set at 0.01. The dispersion of our measured estimates of volatility from the barrier method is shown in Figure Three. The standard deviation of the estimates is 0.028.

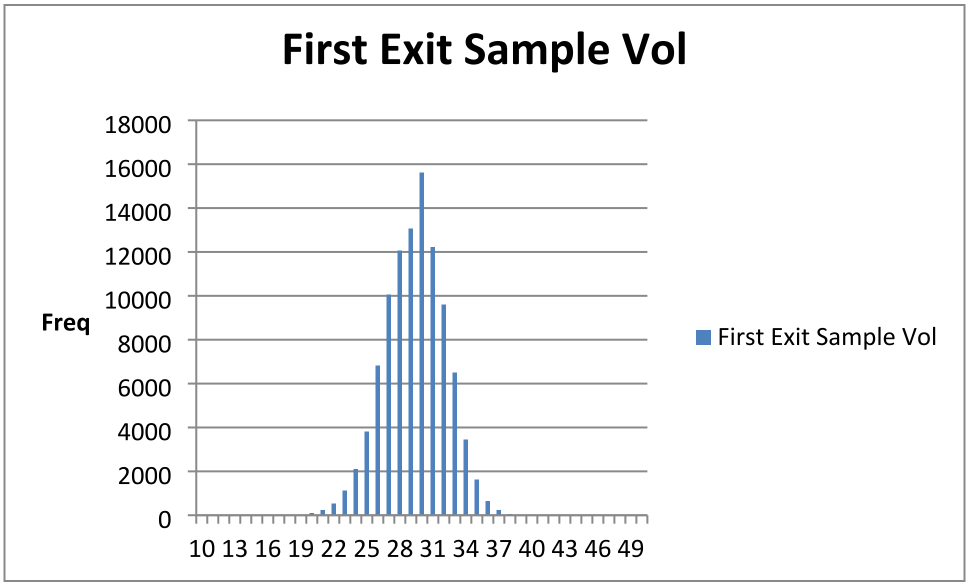

Figure Three: Volatility estimates using the barrier method.

We can invert equation 4 to obtain an estimate of $\tau = \left(\frac{0.01}{0.3}\right)^2 \approx 0.28\ days$. So next we do another simulation, sampling the stock every 0.28 of a day. This means both methods will be using the same amount of data on average. The dispersion of our measured estimates of volatility from the close-to-close method is shown in Figure Four. The standard deviation of the estimates is 0.048.

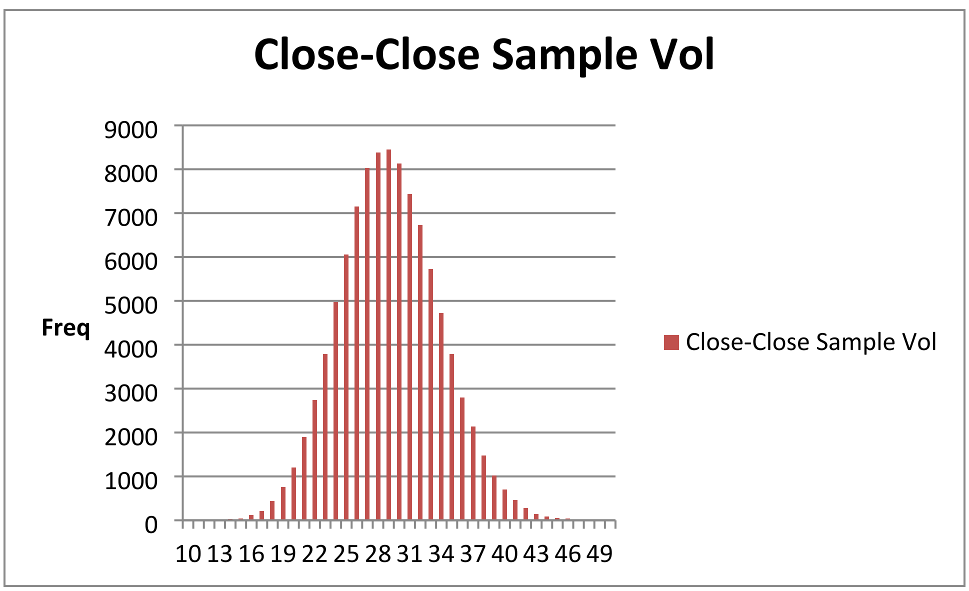

Figure Four: Volatility estimates using the close-to-close method.

An Example

Assume we are observing a $50 stock and set our barrier at 1% (technically we should set the barrier in log terms rather than percentages but the difference here is very small and traders generally develop intuition by thinking in percentage terms). This gives up and down barriers of $50.50 and $49.50 respectively. We choose to define the first exit time as the earliest time that the stock trades at or beyond either barrier. This doesn't have to be our choice. We could opt to observe the midpoint of the bid/ask spread or even wait until the stock is $50.50 bid or $49.50 offered to get a conservative buy-side estimate of volatility (i.e. biased downward).

Imagine that the stock trades at $49.50 halfway through the trading day. Here $\bar{\tau} = \frac{1}{2}\frac{1}{252}$, so our estimate of volatility is $\frac{\Delta}{\sqrt{\bar{\tau}}} = \frac{0.01}{\sqrt{1/504}} \approx 0.2245$. To un-bias this single observation we multiply by 0.8 to obtain $\sigma \approx 0.1796$.

Between barrier hits we can extrapolate the next t based on the current trading level relative to the nearest barrier. This might be useful for generating intraday volatility estimates that smoothly vary with the current stock price.

For instance, in the above example we would reset the barriers to $49.00 (if we were strictly following the 1% rule the barrier would be at $49.005) and $50.00 after the stock hit $49.50. Let's say that with some time left in the trading day we see the stock trading at $49.75. We already have $\tau_1 = \frac{1}{2}\frac{1}{252}$ locked in and we now want to estimate what τ_2 will be. Assume that 1/3 of a day has elapsed since the first hit. Consistent with our assumption that the stock follows a constant volatility Brownian motion, we can say $E[\tau_2] = 1/3 + \left(log(49.75/50.00)/\Delta\right)^2 \approx 1/3 + 1/4 = 7/12$ .The stock has moved half way towards its new upper barrier, hence the $(1/2)^2$ term. So we can expect to wait another 1/4 of a day to hit the upper barrier (obviously longer to hit the lower barrier).

With $\tau_1 = \frac{1}{2}\frac{1}{252}$ and τ_2 estimated at $\frac{7}{12}\frac{1}{252}$ we have $\bar{\tau} = \frac{13}{24}\frac{1}{252}$ and $\sigma \approx 0.2157$ for n=2. If we apply the bias correction, we get $\sigma \approx 0.1917$. Note that the τ_2 estimate is comprised of two terms: one absolutely certain (how long we have waited thus far) and one uncertain (how much longer we expect to wait). The uncertainty in the 2nd term decreases as the stock price approaches one of the barriers.

Conclusion

The first exit time method for estimating volatility is a natural fit for online volatility estimation. It can be easily adapted to take into account the stock's bid-ask spread and discrete trading. It also converges to the true volatility more quickly than the traditional close-to-close estimator.

This volatility estimation scheme has a tangible interpretation for dynamic hedgers. It has long been known that to simply hedge on the basis of elapsed time is sub-optimal (see for example, Hodges and Neuberger, 1989). To achieve an optimal balance between transaction costs and risk a trader will hedge every time the underlying "moves enough". This is our Δ parameter. Thus the trader simply does not care about moves smaller than the Δ threshold. The estimation method developed here is closer in spirit to this hedging scheme whereas the standard close-to-close estimator is more consistent with a time based hedging scheme.

References

Borodin, A., and P. Salminen, 2002, "Handbook of Brownian Motion - Facts and Formulae (Probability and its Applications)", Birkhauser-Verlag, Basel.

Cho, D and E. Frees, 1988, "Estimating the Volatility of Discrete Stock Prices," The Journal of Finance, Volume 43, 451-466.

Hodges, S., and A. Neuberger, 1989, "Optimal Replication of Contingent Claims Under Transaction Costs", Review of Futures Markets, Volume 8, 222-239.

Poon, S., 2005, "A Practical Guide to Forecasting Financial Market Volatility," Wiley, London.