3 Estimating Volatility

In order to apply the Black-Scholes pricing formula, we need five parameters – initial stock price S_0, volatility of the stock σ^2, riskless interest rate r, maturity of the option T and strike price K. Note that four of the five parameters are known and the value of σ^2 needs to be estimated. One possibility is to use historic data. In this Section we will detail how to estimate σ^2 using daily closing prices and also opening and closing prices. We will also consider implied volatility.

3.1 Using Closing Data

Suppose that $Y_1, \ldots, Y_n$ are independent random variables having a common probability distribution with mean μ_0 and variance σ_0^2. For unknown μ_0, the usual estimator of σ_0^2 is the sample variance,

$$\hat{\sigma}_0^2 = \frac{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}{n-1} = \frac{\sum_{i=1}^{n} Y_i^2 - n\bar{Y}^2}{n-1}.$$

Note that the estimator is a random variable (and the capitalisation of the Y_i reflects this). Recall that the above is an unbiased estimator of σ_0^2 since

$$E(\hat{\sigma}_0^2) = \sigma_0^2.$$

Furthermore, if the Y_i are Normally distributed, it can be shown that

$$Var(\hat{\sigma}_0^2) = \frac{2\sigma_0^4}{n-1}.$$

In this case, $\hat{\sigma}_0^2$ is a consistent estimator of σ_0^2. Recall that for consistency, we need asymptotic unbiasedness and the variance of the estimator to tend to zero as the sample size increases.

Proof of unbiasedness

We can show that $E(\hat{\sigma}_0^2) = \sigma_0^2$ as follows. We have that

$$E(\hat{\sigma}_0^2) = \frac{1}{n-1}\left(\sum_{i=1}^{n} E(Y_i^2) - nE(\bar{Y}^2)\right).$$

Now,

$$Var(Y_i) = \sigma_0^2 = E(Y_i^2) - [E(Y_i)]^2.$$

Hence, using $E(Y_i) = \mu_0$ and re-arranging the above gives

$$E(Y_i^2) = \sigma_0^2 + \mu_0^2.$$

In a very similar manner, we obtain

$$E(\bar{Y}^2) = \frac{\sigma_0^2}{n} + \mu_0^2.$$

Finally,

$$
\begin{aligned}
E(\hat{\sigma}_0^2) &= \frac{1}{n-1}\left(n \times (\sigma_0^2 + \mu_0^2) - n \times (\sigma_0^2/n + \mu_0^2)\right) \\
&= \frac{1}{n-1}(n-1) \times \sigma_0^2 \\
&= \sigma_0^2
\end{aligned}
$$

as required.

Application to p.a. volatility estimation

We apply this procedure to estimate the volatility of a stock price as follows. Suppose the *closing price* of a stock is observed at fixed intervals of time (e.g. every day) and is assumed to follow a G.B.M. with volatility σ^2 (per annum). Define:

$n+1$: Number of observations

C_i: Closing price at the end of interval i $(i = 0, \ldots, n)$

Δt: Length of time interval in years

and let

$$Y_i = \log\left(\frac{C_i}{C_{i-1}}\right).$$

Now, under the assumption of G.B.M. Y_i follows a Normal distribution with variance $\sigma^2 \Delta t$. In this context, the sample variance would give an unbiased estimator of $\sigma^2 \Delta t$. Therefore we construct the estimator of σ^2 as

$$\hat{\sigma}^2 = \frac{1}{\Delta t}\frac{\sum_{i=1}^n (Y_i - \bar{Y})^2}{n-1} = \frac{1}{\Delta t}\frac{\sum_{i=1}^n Y_i^2 - n\bar{Y}^2}{n-1}.$$

Comments

- Choosing a value n is not as straightforward as would appear. Naturally, more data will lead to more accuracy. However, it may not be reasonable to assume that the volatility of a particular stock is the same now as, say ten years ago. We will come back to this point later on.

- The procedure is typically used with daily data. An important point is whether volatility should be measured per calendar year or trading year. Evidence suggests that volatility is largely caused by trading itself and therefore σ^2 is typically measured per trading year. Since there are (approximately) 252 trading days in a year (so that 1 day $= 1/252$ years), the estimator of σ^2 based on $n+1$ days of data is

$$\hat{\sigma}^2 = 252 \frac{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}{n-1} = 252 \frac{\sum_{i=1}^{n} Y_i^2 - n\bar{Y}^2}{n-1} \, .$$

Example 3.1

Consider the following closing day Gas prices (in dollars) observed on 21 consecutive days. Assuming that Gas price follows a G.B.M., we estimate the volatility σ^2 (per annum) by

Day (i)	Closing Price (C_i)	Day (i)	Closing Price (C_i)
0	52.75	11	59.48
1	53.43	12	58.12
2	54.51	13	57.40
3	53.77	14	56.38
4	53.90	15	57.60
5	53.66	16	57.25
6	54.54	17	57.44
7	54.92	18	56.07
8	55.00	19	56.21
9	56.88	20	57.76
10	57.80		

Table 2: Commodity Prices (dollars)

computing

$$\sum_{i=1}^{20} y_i^2 = \sum_{i=1}^{20} \left[\log\left(\frac{C_i}{C_{i-1}} \right) \right]^2 = 0.00615$$

and

$$\bar{y} = \frac{1}{20} \sum_{i=1}^{20} \log\left(\frac{C_i}{C_{i-1}} \right) = 0.00454 \, .$$

Hence we obtain the *estimate*

$$\hat{\sigma}^2 = 252 \frac{0.00615 - 20 \times 0.00454^2}{19} = 0.0761 \,.$$

Example 3.2

Here, we will consider daily closing values of the Standard and Poors 500 (S&P500) index from January 03, 2012 until October 25, 2022. The data set can be downloaded from the course page (click *Download SandP500.RData* and double click the file once it's downloaded). The data will then be loaded into R (RStudio) where a new vector called 'sp' will be loaded into the R environment. The following commands can be used to plot:

```
plot(ts(sp,start=2012,deltat=1/252),xlab="Time (years)",ylab="Price (dollars)")
```

We can plot the *daily returns*, $\log(C_i/C_{i-1})$ with

```
len = length(sp)
ret = log(sp[2:len]/sp[1:(len-1)])
plot(ts(ret,start=2012,deltat=1/252),xlab="Time (years)",ylab="Returns")
```
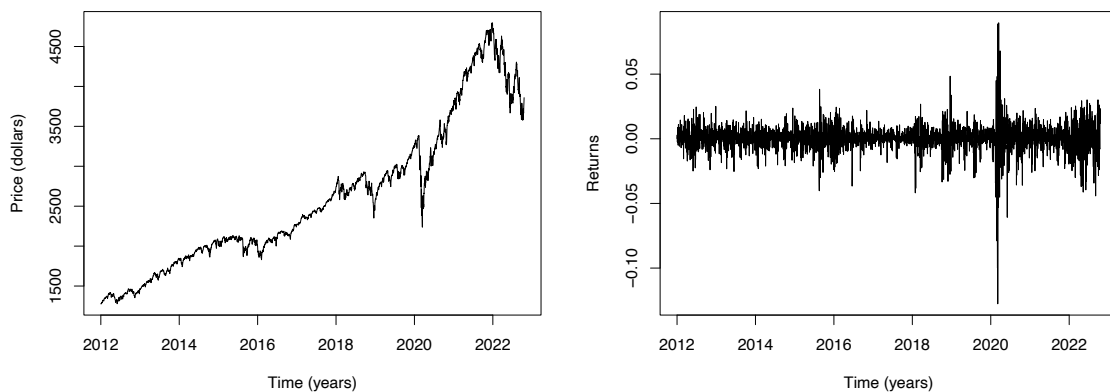

Figure 8: Daily closing prices and returns of the S&P500 index

Figure 8 shows daily closing prices and returns. Assuming that the S&P500 index data follows a G.B.M. we can estimate the volatility per annum, using the full set of observations, with the command,

```
var(ret)*252
```

which yields $\hat{\sigma}^2 = 0.0294$. Note that in computing this estimate we have assumed that the volatility is constant over the period January 03, 2012 – October 25, 2022. Figure 9 shows estimates of square root annual volatility, σ, obtained by taking the S&P500 data for each year in turn, and applying the formula (and square-rooting). On inspection of this plot, it appears that our assumption of a constant volatility over the entire period may be an unreasonable one.
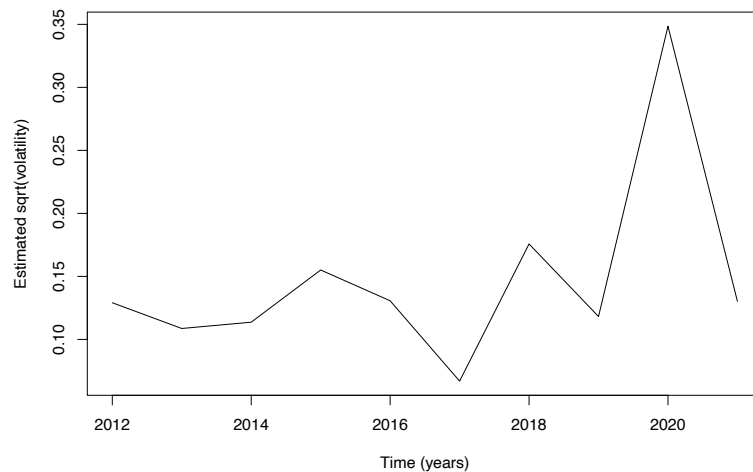

Figure 9: Estimated square root volatility against year, using the S&P500 data.

3.2 Using Opening and Closing Data (not examinable)

Let C_i denote the (closing) price of a particular stock at the end of trading day i. Assuming that stock price follows a geometric Brownian motion, $\log(C_i/C_{i-1})$ follows a Normal distribution with variance $\sigma^2/252$. Now let O_i denote the *opening price* at the start of trading day i and write

$$
\begin{aligned}
\log\left(\frac{C_i}{C_{i-1}}\right) &= \log\left(\frac{C_i}{O_i}\frac{O_i}{C_{i-1}}\right) \\
&= \log\left(\frac{C_i}{O_i}\right) + \log\left(\frac{O_i}{C_{i-1}}\right).
\end{aligned}
$$

Now, assuming that the ratio price change during a trading day is independent of the change that happened when the market was closed (i.e. C_i/O_i and O_i/C_{i-1} are independent) it follows that

$$
\begin{aligned}
\mathrm{Var}\left(\log(C_i/C_{i-1})\right) &= \mathrm{Var}\left(\log(C_i/O_i)\right) + \mathrm{Var}\left(\log(O_i/C_{i-1})\right) \\
&= \mathrm{Var}\left(C_i^* - O_i^*\right) + \mathrm{Var}\left(O_i^* - C_{i-1}^*\right)
\end{aligned}
$$

where $C_i^* = \log(C_i)$ and $O_i^* = \log(O_i)$. If we observe closing prices on days $0, 1 \ldots, n$ and opening prices on days $1, 2, \ldots n$ and assume that $C_i^* - O_i^*$ and $O_i^* - C_{i-1}^*$ both have a mean of approximately 0, then we may estimate σ^2 by

$$\hat{\sigma}^2 = 252 \frac{\sum_{i=1}^{n} \left[(C_i^* - O_i^*)^2 + (O_i^* - C_{i-1}^*)^2 \right]}{n - 1} . \tag{12}$$

Example 3.3

Consider the following opening and closing day prices (in dollars) of the S&P500 observed on 11 consecutive days.

Day (i)	Open (O_i)	Close (C_i)
0		241.13
1	241.24	242.49
2	242.71	245.73
3	245.73	248.13
4	248.18	244.35
5	244.12	241.99
6	241.13	244.06
7	244.56	250.84
8	250.67	252.04
9	252.12	254.70
10	254.70	251.79

Table 3: S&P500 opening and closing prices (dollars)

We estimate the per annum volatility σ^2 by computing

$$n + 1 = 11, \qquad n = 10,$$

so that

$$\sum_{i=1}^{10} (\log(C_i) - \log(O_i))^2 = \sum_{i=1}^{10} (C_i^* - O_i^*)^2 = 0.00165$$

and

$$\sum_{i=1}^{10} (\log(O_i) - \log(C_{i-1}))^2 = \sum_{i=1}^{10} (O_i^* - C_{i-1}^*)^2 = 1.94 \times 10^{-5}$$

Hence our estimate is

$$\hat{\sigma}^2 = \frac{252}{9} \sum_{i=1}^{10} \left[(C_i^* - O_i^*)^2 + (O_i^* - C_{i-1}^*)^2 \right] = 0.047$$

3.3 Implied Volatility

We have seen how to estimate σ from historic data. In practice, traders usually work with *implied volatilities*. These are the volatilities implied by option prices observed in the market.

Definition 3.1

The implied volatility of an option is the value of the volatility (of the underlying stock process) which when input into an option pricing model gives a value equal to the current/quoted price of the option.

Example 3.4

Suppose that stock prices follow a risk neutral G.B.M. (with parameters r and σ^2). Now suppose that the fair price of one ECO is 1.875 when $S_0 = 21$, $K = 20$, $r = 0.1$ and $T = 0.25$. The implied volatility is the value of σ that, when substituted into the Black-Scholes formula gives $C_0 = 1.875$.

Unfortunately, it is not possible to analytically express σ as a function of S_0, K, r, T using the Black-Scholes formula. An iterative search procedure can be used instead. For example, start with a value $\sigma = 0.2$ and calculate a value for C_0. The following R function should be useful:

```
bs=function(s0=21,K=20,r=0.1,T=0.25,sig)
{
 omega=(r*T+sig*sig*T/2-log(K/s0))/(sig*sqrt(T))
 #Return BS price
 s0*pnorm(omega,0,1)-K*exp(-r*T)*pnorm(omega-sig*sqrt(T))
}
```

Then, `bs(sig=0.2)` gives $C_0 = 1.76$, which is too low. Since the fair price is increasing in σ (proof of which is beyond the scope of the course), we could try $\sigma = 0.3$. We perform the following in R:

```
bs(sig=0.3)
[1] 2.101014
> bs(sig=0.25)
[1] 1.926831
> bs(sig=0.21)
[1] 1.795785
> bs(sig=0.22)
```

```
[1] 1.827633
> bs(sig=0.23)
[1] 1.860137
> bs(sig=0.235)
[1] 1.876611
> bs(sig=0.234)
[1] 1.873305
```

which suggests an implied volatility of 0.234-0.235. More powerful iterative search procedures, such as the Newton-Raphson method could be used.