Chapter 7

Itô's Lemma

1 Itô's Lemma

Itô's lemma is the most important result (and tool) in doing stochastic analysis. Roughly speaking, it relates the small change in a function of a random variable to the small change in the random variable itself. In the following, our heuristic approach to Itô's lemma is based on the Taylor series expansion.

Recall that we assume that the asset price $S(t)$ follows the geometric Brownian motion:

$$dS = \mu S dt + \sigma S dX_t, \tag{1}$$

where again we will use X_t, $X(t)$, or X interchangeably. Using Proposition 6.8, we have shown in Example 6.6 that

$$(dS)^2 = \sigma^2 S^2 dt. \tag{2}$$

For any sufficiently smooth function $f(S,t)$, by the Taylor series expansion, we have

$$df = \frac{\partial f}{\partial S} dS + \frac{\partial f}{\partial t} dt + \frac{1}{2}\frac{\partial^2 f}{\partial S^2}(dS)^2 + o(dt, dS),$$

where $o(dt, dS)$ represents higher order terms than dt and dS (i.e. the term converges to 0 faster than dt and dS, as dt and dS converge to 0). Using (1) and (2), we find a new expression for df:

$$df = (\mu S \frac{\partial f}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 f}{\partial S^2} + \frac{\partial f}{\partial t})dt + \sigma S \frac{\partial f}{\partial S} dX_t. \tag{3}$$

This is a special case of the Itô's lemma.

Theorem 1. (Itô's Lemma) *Let $\{Y_t; t \geq 0\}$ be an Itô process:*

$$dY_t = a(Y_t, t)dt + b(Y_t, t)dX_t, \tag{4}$$

where both $a(Y_t,t)$ and $b(Y_t,t)$ are sufficiently smooth functions. If $f(Y_t,t)$ is a sufficiently smooth function of Y_t and t, then $f(Y_t,t)$ will also be an Itô process satisfying

$$df = \left[a(Y_t,t)\frac{\partial f}{\partial Y_t} + \frac{1}{2}b(Y_t,t)^2\frac{\partial^2 f}{\partial Y_t^2} + \frac{\partial f}{\partial t}\right] dt + b(Y_t,t)\frac{\partial f}{\partial Y_t} dX_t. \tag{5}$$

Clearly (3) is a particular case of (5) with $Y_t = S(t)$. Let us consider more examples before proving the theorem.

Example 1. Let $Y_t = S_t$ in (4). We will simply write S_t as S. Thus we have $a(Y_t, t) = \mu S$ and $b(Y_t, t) = \sigma S$ in (4). Define $f(S, t) = \ln S$ in (5). Because

$$\frac{\partial f}{\partial S} = \frac{1}{S}, \qquad \frac{\partial^2 f}{\partial S^2} = -\frac{1}{S^2}, \qquad \frac{\partial f}{\partial t} = 0,$$

it follows from (5) that

$$d(\ln S) = (\mu - \frac{\sigma^2}{2})dt + \sigma dX_t. \tag{6}$$

It indicates that if $S(t)$ is the geometric Brownian motion given by (1), then $\ln S(t)$ follows a generalized Wiener process with constant drift parameter $(\mu - \sigma^2/2)$ and constant variance rate σ^2. Therefore, the change in $\ln S(t)$ between time 0 and some future time t is normally distributed with mean $(\mu - \sigma^2/2)t$ and variance $\sigma^2 t$. □

Integrating both side of (6) from time 0 to t, and let $S_0 = S(0)$, then we have

$$S(t) = S_0 e^{(\mu - \frac{\sigma^2}{2})t + \sigma X(t)}. \tag{7}$$

This is the solution to the stochastic differential equation (6). In (7), if we neglect the Brownian motion $X(t)$, then $S(t)$ is increasing exponentially at a rate of $(\mu - \sigma^2/2)$, or equivalently, geometrically with a common factor $e^{(\mu - \frac{\sigma^2}{2})}$. This is why $S(t)$ is called a *geometric* Brownian motion.

It is important to note that Itô calculus is completely different from the ordinary calculus we learned before when there is no stochastic process involved. To illustrate that, we first note that since $S(t)$ in (7) is derived from the process $S(t)$ that satisfies the stochastic differential equation (1), $S(t)$ given in (7) is also a solution to (1).

This may be surprising to someone unfamiliar to Itô calculus. For one may think that from (1), we get

$$\frac{dS}{S} = \mu dt + \sigma dX_t. \tag{8}$$

Hence integrating both side from 0 to t, we should get the solution

$$S(t) = S_0 e^{\mu t + \sigma X(t)},$$

which is different from (7). The paradox comes from the fact that if $S(t)$ is a stochastic process satisfying (1), then

$$\mu dt + \sigma dX_t = \frac{dS}{S} \neq d(\ln S) = (\mu - \frac{\sigma^2}{2})dt + \sigma dX_t. \tag{9}$$

We will explain this discrepancy later in §4. Right now, let us illustrate with a simple example why in general $d \ln Y_t \neq dY_t / Y_t$ for stochastic processes Y_t.

Example 2. Let $\{X_t, t \geq 0\}$ be a standard Wiener process. Let us compute $d[\ln(X_t)]$. First in (4), we set $Y_t = X_t$ and hence $a(Y_t, t) = 0$ and $b(Y_t, t) = 1$. In (5), we put $Y_t = X_t$, and $f(Y_t, t) = \ln(X_t)$, then we have

$$d \ln(X_t) = -\frac{1}{2X_t^2}dt + \frac{dX_t}{X_t} \neq \frac{dX_t}{X_t}.$$

This equation shows that $Z_t \equiv \ln X_t$ is the solution to the stochastic PDE

$$dZ_t = e^{-Z_t} dX_t - \frac{1}{2} e^{-2Z_t} dt. \quad \square$$

Thus we should be careful when we differentiate functions with stochastic processes as independent variables.

Example 3. Let $\{X_t, t \geq 0\}$ be a standard Wiener process. Let us compute $d(X_t)^2$. Again we set $Y_t = X_t$ in (4), and hence $a(Y_t, t) = 0$ and $b(Y_t, t) = 1$. In (5), we put $Y_t = X_t$, and $f(Y_t, t) = X_t^2$, then we have

$$d[f(X_t)] = d(X_t)^2 = dt + 2X_t dX_t. \tag{10}$$

In other words, $Z_t(X_t, t) = (X_t)^2$ solves the stochastic differential equation

$$\begin{cases} dZ_t(X_t, t) = dt + 2[Z_t(X_t, t)]^{1/2} dX_t, \\ \qquad\qquad Z_t = 0. \end{cases}$$

Integrating both sides of (10) from 0 to t and noting that for standard Wiener process $X_0 = 0$, we get

$$X_t^2 = t + 2\int_0^t X_s dX_s.$$

Thus

$$\int_0^t X_s dX_s = \frac{1}{2}(X_t^2 - t),$$

whereas the ordinary Riemann integral yields

$$\int_0^t W(s) dW(s) = \frac{1}{2} W(t)^2,$$

if $W(0) = 0$. This is just another example to show how different Itô integrals are from the ordinary Riemann integrals. $\square$

Example 4. Let Y_t be the Itô process given by

$$dY_t = \theta(t) dX_t - \frac{1}{2}\theta^2(t) dt.$$

Thus $a = -\frac{1}{2}\theta^2(t)$ and $b = \theta(t)$ in (4). Now, let $f(Y_t, t) = e^{Y_t} \equiv Z_t$. Observe that

$$\frac{\partial f}{\partial Y_t} = e^{Y_t} = \frac{\partial^2 f}{\partial Y_t^2} \quad \text{and} \quad \frac{\partial f}{\partial t} = 0.$$

Hence by (5),

$$dZ_t = \theta(t) e^{Y_t} dX_t = \theta(t) Z_t dX_t. \tag{11}$$

In other words, $Z_t = \exp\left\{\int_0^t \theta(s) dX_s - \frac{1}{2}\int_0^t \theta^2(s) ds\right\}$ solves the stochastic differential equation (11). $\square$

2 Proof of Itô's Lemma

In this section, we show how Itô's lemma can be regarded as a natural extension of Taylor's expansion. Consider a sufficiently smooth function f of two variables y and t. If dy and dt are the small changes in y and t and df is the resulting small change in f, then the Taylor series expansion of df to the second order is:

$$df = \frac{\partial f}{\partial y}dy + \frac{\partial f}{\partial t}dt + \frac{1}{2}\frac{\partial^2 f}{\partial y^2}dy^2 + \frac{\partial^2 f}{\partial y \partial t}dydt + \frac{1}{2}\frac{\partial^2 f}{\partial t^2}dt^2 + o(dy^2, dt^2). \tag{12}$$

We now extend (12) to cover functions with variables following Itô processes. Suppose that a variable Y_t is an Itô process, i.e.,

$$dY_t = a(Y_t, t)dt + b(Y_t, t)dX_t, \tag{13}$$

where X_t is a standard Wiener process. Let f be some function of Y_t and of time t. By analogy with (12), we can write

$$df = \frac{\partial f}{\partial Y_t}dY_t + \frac{\partial f}{\partial t}dt + \frac{1}{2}\frac{\partial^2 f}{\partial Y_t^2}dY_t^2 + \frac{\partial^2 f}{\partial Y_t \partial t}dY_t dt + \frac{1}{2}\frac{\partial^2 f}{\partial t^2}dt^2 + o(dY_t^2, dt^2). \tag{14}$$

Note that by (13),

$$dY_t^2 = a(Y_t, t)^2 dt^2 + 2a(Y_t, t)b(Y_t, t)dtdX_t + b(Y_t, t)^2 dX_t^2.$$

By Proposition 6.8, $dX_t^2 = dt$, hence $dY_t^2 = b(Y_t, t)^2 dt + o(dt)$. Putting this back to (14), we have

$$df = \frac{\partial f}{\partial Y_t}dY_t + \frac{\partial f}{\partial t}dt + \frac{1}{2}b(Y_t, t)^2 \frac{\partial^2 f}{\partial Y_t^2}dt + o(dY_t, dt). \tag{15}$$

Substituting dY_t from (13) into (15) and neglecting the higher order terms, we get the Itô lemma:

$$df = \left(a(Y_t, t)\frac{\partial f}{\partial Y} + \frac{\partial f}{\partial t} + \frac{1}{2}b(Y_t, t)^2 \frac{\partial^2 f}{\partial Y^2} \right) dt + b(Y_t, t)\frac{\partial f}{\partial Y}dX_t. \quad \square$$

3 Log-Normal Property of Geometric Brownian Motions

For an asset price $S(t)$ which follows the geometric Brownian motion (1), we know that $S(t)$ is given by (7). Modern mathematical economists usually prefer *geometric Brownian motion* over Brownian motion as a model for prices of assets, say shares of stock, that are traded in a perfect market. By (7), such prices are nonnegative and exhibit random fluctuations about a long-term exponential decay or growth curve. Both of these properties are possessed by geometric Brownian motion, but not by Brownian motion itself.

Note that if we take the logarithm on both side of (7), we have

$$\ln S(t) = \ln S_0 + (\mu - \frac{\sigma^2}{2})t + \sigma X(t),$$

where $S_0 = S(0)$. Since $X(t)$ is a Wiener process,

$$\ln S(t) \sim \mathcal{N}\left(\ln S_0 + (\mu - \frac{\sigma^2}{2})t, \sigma^2 t \right). \tag{16}$$

Thus the logarithm of $S(t)$ follows a normal distribution. Hence the distribution of $S(t)$ is also called *log-normal*. Equivalently, (16) can be written as

$$\ln\left(\frac{S_t}{S_0}\right) \sim \mathcal{N}\left((\mu - \frac{\sigma^2}{2})t, \sigma^2 t\right). \tag{17}$$

Note that S_t can take any value between zero and infinity. Recall in (6.16), we have

$$\frac{dS}{S} \sim \mathcal{N}(\mu dt, \sigma^2 dt). \tag{18}$$

This also indicates again that $dS/S \neq d(\ln S)$. Note that (17) holds for all t while (18) holds for small dt only.

Example 5. Consider a stock with an initial price of \$40, an expected return of 16% per annum, and a volatility of 20% per annum. Let us compute its price distribution in six months' time. Take $t = 0.5$ year. Clearly, $\mu = 0.16$ and $\sigma = 0.20$. By (16)

$$\ln S(0.5) \sim \mathcal{N}\left(\ln 40 + (0.16 - \frac{0.2^2}{2}) \times \frac{1}{2}, 0.2^2 \times \frac{1}{2}\right) \sim \mathcal{N}(3.759, 0.141^2).$$

Recall from (6.8) that there is a 95% probability that a standard normal random variable has a value within 1.96 standard deviations of its mean. Hence, within 95% confidence,

$$3.759 - 1.96 \times 0.141 < \ln S(0.5) < 3.759 + 1.96 \times 0.141.$$

This can be written

$$e^{3.759-1.96\times 0.141} < S(0.5) < e^{3.759+1.96\times 0.141},$$

or $32.55 < S(0.5) < 56.56$. Thus, there is a 95% probability that the stock price in six months will lie between \$32.55 and \$56.56. $\square$

From (17), we have

Theorem 2. *The probability density function of S is given by*

$$p_S(s) = \frac{1}{\sigma s\sqrt{2\pi t}} e^{-[\ln(s/S_0)-(\mu-\frac{1}{2}\sigma^2)t]^2/2\sigma^2 t}, \quad 0 < s < \infty. \tag{19}$$

Proof. From (16) and (6.5), the probability density function of $\ln S$ is given by

$$p_{\ln S}(x) = \frac{1}{\sigma\sqrt{2\pi t}} e^{-[x-\ln(S_0)-(\mu-\frac{1}{2}\sigma^2)t]^2/2\sigma^2 t}, \quad -\infty < x < \infty.$$

Note that by definition,

$$\int_{-\infty}^{w} p_{\ln S}(x)dx = \text{Prob}\{\ln S \leq w\} = \text{Prob}\{S \leq e^w\} = \int_{-\infty}^{e^w} p_S(x)dx.$$

Differentiate both sides with respect to w, we get

$$p_{\ln S}(w) = p_S(e^w)e^w.$$

Letting $w = \ln s$, we get (19). $\square$

The probability density function of S_t is useful, as for example, we can compute

$$\text{Prob}\{S_t \ge E\} = \int_E^\infty p_S(s)ds = \int_E^\infty \frac{1}{\sigma s\sqrt{2\pi t}} e^{-[\ln(s/S_0)-(\mu-\frac{1}{2}\sigma^2)t]^2/2\sigma^2 t} ds.$$

It will be useful later when we want to evaluate the price of an option by risk neutrality method, see §8.6. Using Theorem 2, one can also evaluate the expected value and the variance of S_t by direct integration.

Corollary 3. *We have*

$$\mathbb{E}[S_t] = \int_{-\infty}^{\infty} s p_S(s) ds = S_0 e^{\mu t}, \tag{20}$$

$$\text{Var}[S_t] = S_0^2 e^{2\mu t}[e^{\sigma^2 t} - 1]. \tag{21}$$

Proof. Note that $X_t \sim \mathcal{N}(0,t)$, then $X_t = \sqrt{t}Z$. By (6.4),

$$\mathbb{E}[S_t] = \mathbb{E}\left[S_0 e^{(\mu-\frac{\sigma^2}{2})t+\sigma X_t}\right] = S_0 e^{(\mu-\frac{\sigma^2}{2})t}\mathbb{E}[e^{\sigma\sqrt{t}Z}] = S_0 e^{(\mu-\frac{\sigma^2}{2})t} e^{\frac{\sigma^2}{2}t} = S_0 e^{\mu t}.$$

Similarly,

$$\begin{aligned}\mathbb{E}[S_t^2] &= \mathbb{E}\left[S_0^2 e^{(2\mu-\sigma^2)t+2\sigma X_t}\right] = S_0^2 e^{(2\mu-\sigma^2)t}\mathbb{E}[e^{2\sigma\sqrt{t}Z}] \\ &= S_0^2 e^{(2\mu-\sigma^2)t} e^{2\sigma^2 t} = S_0^2 e^{(2\mu+\sigma^2)t}.\end{aligned}$$

Hence

$$\text{Var}(S_t) = \mathbb{E}[S_t^2] - \mathbb{E}[S_t]^2 = S_0^2 e^{2\mu t}(e^{\sigma^2 t} - 1). \quad \square$$

Example 6. Consider a stock where the current price is \$20, the expected return is 20% per annum, and the volatility is 40% per annum. The expected stock price in one year and the variance of the stock price in one year are given by, respectively,

$$\begin{aligned}\mathbb{E}[S_1] &= 20e^{0.2\times 1} = 24.43, \\ \text{Var}[S_1] &= 400e^{2\times 0.2\times 1}(e^{0.4^2\times 1} - 1) = 103.54.\end{aligned}$$

The standard deviation of the stock price in one year is $\sqrt{103.54} = 10.18$. $\square$

4 Expected Returns of Geometric Brownian Motions

From (9), we know that

$$\mu dt = \mathbb{E}\left[\frac{dS}{S}\right] > \mathbb{E}[d(\ln S)] = \left(\mu - \frac{\sigma^2}{2}\right) dt. \tag{22}$$

The difference comes from measuring the returns differently. It will give different rates of return, and is closely related to an issue in the reporting of mutual fund returns. We illustrate that with an example first.

Example 7. A mutual fund, starting at \$100 five years ago, performed as follows in the last five years:

$$115, 138, 179.4, 143.52, 179.4.$$

Then the returns per annum measured using annual compounding will be:

$$15\%, \quad 20\%, \quad 30\%, \quad -20\%, \quad 25\%$$

i.e. $[S(1) - S(0)]/S(0) = 0.15$, $[S(2) - S(1)]/S(2) = 0.20$, etc. The arithmetic mean of the returns, calculated by taking the sum of the returns and dividing by 5, is 14%. However, an investor would actually earn less than 14% per annum by leaving the money invested in the fund for five years. In fact, for the \$100 he invested, at the end of the five years, he gets only \$179.40. If we measure that using continuous compound interest, that will be 11.69% as

$$100 \times e^{0.1169 \times 5} = 179.40. \tag{23}$$

Note that the average return 11.69% can also be obtained by computing the returns annually by compounding continuously and they are:

$$14\%, \quad 18.2\%, \quad 26.2\%, \quad -22.3\%, \quad 22.3\%$$

where $S(1) = S(0)e^{0.14}$ (i.e. $\ln S(1) - \ln S(0) = 0.14$), $S(2) = S(1)e^{0.182}$, etc. Then the average return 11.69% is obtained by averaging these five returns. The fact that $14\% > 11.69\%$ is just a validation of (22).

So what average return should the fund manager report? It is tempting for the manager to make a statement such as: "The average of the returns per year that we have realized in the last five years is 14%". Although true, this is misleading. It is much less misleading to say "The average return realized by someone who invested with us for the last five years is 11.69% per year." In some jurisdictions regulatory standards require fund managers to report returns the second way. $\square$

In the following, to distinguish the two different ways of computing average return, we will call the one computed by compounding discretely (like the 14% in Example 7) the "short-term" return rate, while the one computed by compounding continuously (like the 11.69% in Example 7) the "long-term return" rate. We call the latter long-term because it can be computed using the beginning price and the ending price only, (after 5 years in Example (7), see (23)). Thus (22) states that the long-term return rate $(\mu - \sigma^2/2)$ is always less than the short-term return rate μ. In the following, we verify that from their original definitions and show how exactly the term $\sigma^2/2$ comes in. We start with the long-term return rate.

Proposition 4. *Assume that the asset price follows the geometric Brownian motion (1). Let η be the continuously compounded rate of return per annum realized by the stock between times t and $t + \Delta t$. Then*

$$\mathbb{E}[\eta] = \mu - \frac{\sigma^2}{2}. \tag{24}$$

Proof. By definition, $S(t + \Delta t) = S(t)e^{\eta \Delta t}$. Thus,

$$\eta = \frac{1}{\Delta t} \ln \frac{S(t + \Delta t)}{S(t)}.$$

By (17), we obtain

$$\eta \sim \mathcal{N}\left(\mu - \frac{\sigma^2}{2}, \frac{\sigma^2}{\Delta t}\right). \tag{25}$$

Thus η, the continuously compounded rate of return per annum for the stock is normally distributed with mean $(\mu - \sigma^2/2)$ and standard deviation $\sigma/\sqrt{\Delta t}$. Hence we get (24). $\square$

We see that the rate of return on the stock, when measured in the continuously compounded way, will tend to $\mu - \sigma^2/2$ as Δt increases (because the variance decreases with Δt). That is, if we measure the return rate in a longer time frame, we are more certain that it will be closer to $\mu - \sigma^2/2$.

Example 8. Consider a stock with an expected return of 17% per annum and a volatility of 20% per annum. Then, by (25), the probability distribution for the actual rate of return (continuously compounded) realized over three years is normal with mean

$$0.17 - \frac{0.2^2}{2} = 0.15,$$

or 15% per annum and standard deviation

$$\frac{0.2}{\sqrt{3}} = 0.1155$$

or 11.55% per annum. Because there is a 95% chance that a normal random variable will lie within 1.96 standard deviations of its mean, we can be 95% confident that the actual return realized over three years will be between -7.6% $(= 0.15 - 1.96 \times 0.1155)$ and $+37.6\%$ $(= 0.15 + 1.96 \times 0.1155)$ per annum. $\square$

Next we compute the "short-term" expected return rate of the stock.

Proposition 5. *Assume that the asset price follows the geometric Brownian motion (1). Let λ be the rate of return of the stock over the short time interval $(t, t + \Delta t)$, i.e.*

$$\frac{S(t + \Delta t) - S(t)}{S(t)} = \lambda \Delta t. \tag{26}$$

Then

$$\mathbb{E}[\lambda] = \mu. \tag{27}$$

Proof. By (18), we know that $\mathbb{E}(dS/S) = \mu dt$, and hence $\mathbb{E}[\lambda]$ should be equal to μ. Let us prove it more formally so we see how the $\sigma^2/2$ term drops out. By (7), the left-hand side of (26) can be expressed as

$$\frac{S(t + \Delta t) - S(t)}{S(t)} = \frac{S(t + \Delta t)}{S(t)} - 1 = e^{(\mu - \frac{\sigma^2}{2})\Delta t + \sigma(X_{t+\Delta t} - X_t)} - 1.$$

Expanding the exponential term around 0 to the second order and noting from (6.15) that

$$X_{t+\Delta t} - X_t = dX_{\Delta t} = \varepsilon\sqrt{\Delta t},$$

where $\varepsilon \sim \mathcal{N}(0, 1)$, we get

$$\begin{aligned}\frac{S(t + \Delta t) - S(t)}{S(t)} &= \{1 + (\mu - \frac{\sigma^2}{2})\Delta t + \sigma\varepsilon\sqrt{\Delta t} + \frac{1}{2}\sigma^2\varepsilon^2\Delta t + o(\Delta t)\} - 1 \\ &= \sigma\varepsilon\sqrt{\Delta t} + \mu\Delta t - \frac{\sigma^2}{2}\Delta t + \frac{\sigma^2}{2}\varepsilon^2\Delta t + o(\Delta t).\end{aligned}$$

Taking the expectation and noting that $\varepsilon \sim \mathcal{N}(0,1)$, we get

$$\mathbb{E}\left[\frac{S(t+\Delta t)-S(t)}{S(t)}\right]=\mu\Delta t.$$

Comparing this with (26), we get (27), i.e. μ is the expected return on the stock for the short time period Δt. $\square$

Thus we have shown that the expected continuously compounded return of a stock following geometric Brownian motion is $(\mu-\sigma^2/2)$, which is less than μ, the expected return on the stock computed discretely. In fact this is true for any asset models. (Indeed, in Example 7, we did not assume any model on the fund prices.) The phenomenon is well known: the geometric mean of a set of numbers (not all the same) is always less than the arithmetic mean. To understand this further, consider an asset not necessarily follows a log-normal distribution. Define μ to be the expected return on the asset per annum in a period Δt, and η be the expected return on the asset per annum with a compounding frequency of Δt. An important general result is that $\mu > \eta$. As $\Delta t \to 0$, η becomes the expected continuously compounded return. For geometric Brownian motions, this expected continuously compounded return is $(\mu-\sigma^2/2)$. A simple proof of this general result for any process is given as follows.

Proposition 6. *Let μ_i be the return rate of the stock $S(t)$ from time t_i to t_{i+1}, and η_i be the return rate of the stock computed by continuously compounded return over the same period. If $S(t_{i+1}) \neq S(t_i)$, then $\mu_i > \eta_i$. In particular, the average of μ_i over a number of time intervals will always be greater than the average of η_i over the same set of intervals, unless $S(t_i)$ are all the same in the intervals.*

Proof. By definition,

$$\mu_i \cdot (t_{i+1}-t_i)=\frac{S(t_{i+1})-S(t_i)}{S(t_i)}=\frac{S(t_{i+1})}{S(t_i)}-1 \geq -1.$$

By assumption $\mu_i \cdot (t_{i+1}-t_i) \neq 0$. Also by definition

$$S(t_{i+1})=S(t_i)e^{\eta_i(t_{i+1}-t_i)}.$$

Hence

$$\eta_i \cdot (t_{i+1}-t_i)=\ln\left(\frac{S(t_{i+1})}{S(t_i)}\right)=\ln[1+\mu_i(t_{i+1}-t_i)] < \mu_i \cdot (t_{i+1}-t_i),$$

as $\ln(1+x) < x$ for all nonzero $x \geq -1$. (This last fact you can establish by showing that $x-\ln(1+x)$ is a decreasing function for $x \leq 0$ and an increasing function for $x \geq 0$ and its minimum is at $x=0$.) Hence $\eta_i < \mu_i$. $\square$

The main conclusion of this section is that it will be ambiguous to talk about the return rate of an asset. In the following, the *return rate of an asset* will always be referring to μ.

5 Calibrating Geometric Brownian Motions

None of the analysis that we have presented so far is of much use unless we can estimate the parameters in our model. In the next chapter, we will see that only the volatility parameter σ appears in the value of an option and the drift μ does not. But how do we obtain σ, the volatility, of the asset $S(t)$. One way is to estimate it from historic data.

From Definition 6.5, the volatility σ of a stock can be defined as *the standard deviation of the return* provided by the stock in one year when the return is expressed using continuous compounding. Equation (16) shows that σ is also the standard deviation of the natural logarithm of the stock price at the end of one year. Thus a simple approach for estimating σ from past data is as follows.

Suppose we have the values of asset price S at $n+1$ equal time-steps Δt. Call these values $S_0, S_1, \ldots, S_n$ in chronological order with S_0 being the first value. That is,

$$\begin{aligned} n+1 &: \text{Number of observations} \\ S_i &: \text{Stock price at the end of } i\text{th interval} \\ \Delta t &: \text{Length of time interval in years} \end{aligned}$$

We first form the data series

$$U_i = \ln(S_{i+1}) - \ln(S_i) = \ln\left(\frac{S_{i+1}}{S_i}\right), \quad i = 0, 1, \ldots n.$$

Since

$$\frac{S_{i+1} - S_i}{S_i} = \frac{S_{i+1}}{S_i} - 1 = e^{(\mu - \frac{\sigma^2}{2})\Delta t + \sigma dX_i} - 1,$$

where dX_i are independent, the successive ratios

$$\frac{S_1}{S_0}, \frac{S_2}{S_1}, \ldots, \frac{S_n}{S_{n-1}}$$

are independent random variables too.

Next we find the mean $\overline{U}$ and variance $\widehat{\sigma}^2$ of the data series $U_1, U_2, \ldots, U_n$:

$$\overline{U} = \frac{1}{n}\sum_{i=1}^{n} U_i \quad \text{and} \quad \widehat{\sigma}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(U_i - \overline{U})^2.$$

These statistics are estimates for the theoretical mean and variance of the population U_i. In fact, by (17),

$$U_i \sim \mathcal{N}\left((\mu - \frac{\sigma^2}{2})\Delta t, \sigma^2 \Delta t\right).$$

Hence the theoretical mean and variance of U_i are $(\mu - \sigma^2/2)\Delta t$ and $\sigma^2 \Delta t$ respectively. Thus we can solve the equations

$$\begin{cases} \overline{U} &= (\mu - \sigma^2/2)\Delta t \\ \widehat{\sigma}^2 &= \sigma^2 \Delta t \end{cases}$$

for μ and σ. The algebra is easy and the answers are

$$\mu = \frac{\overline{U} + \widehat{\sigma}^2/2}{\Delta t} \qquad \text{and} \qquad \sigma = \frac{\widehat{\sigma}}{\sqrt{\Delta t}}.$$