Chapter 1
Financial Models

1.1 Introduction

The movement of financial assets and products generally displays some type of expected return, even over a short period of time. This expected return trends at a predictable rate that may be positive indicating growth, negative indicating a decline, or zero. Additionally, there are random movements that are individually unpredictable; however, the general distribution of these fluctuations are predictable based on historical movements. The common approach to model randomness is to assume a single or multi-component Gaussian process. The generalized format to describe a time-dependent stochastic process is

$$dS_t = \alpha(S,t)dt + \sigma(S,t)dW_t ,$$

where the drift α and volatility σ are functions of time, t, and asset price, S, and W_t is a Wiener process. If the drift α(S,t) = μ and volatility σ are constants then the process

$$dS_t = \mu dt + \sigma dW_t ,$$

is known as *arithmetic Brownian motion*. This process by itself states that the stock price S will increase (or decrease) without bound at a rate that is not dependent on current stock price. Clearly this does not describe the typical behavior for an asset but modified versions of arithmetic motion are useful in finance and will be revisited later in the text.

1.2 Geometric Brownian Motion

A more appropriate description of a stock price process is that the movements in the stock are proportional to the value of the stock. A specific description is that the overall drift, $\alpha(S,t) = \mu S_t$, is the product of a expected return μ and the current asset price S_t. Adding a stochastic movement that is also proportional to the current price level gives

$$dS_t = \mu S dt + \sigma S dW_t ,$$

which is the well known *geometric Brownian motion* process. A crude discrete approximation of the stochastic differential equation for geometric Brownian motion given by

$$\frac{\Delta S_t}{S_t} = \mu \Delta t + \sigma \Delta W_t$$

is *only valid over short time intervals*. This form does highlight that the percentage change in the stock price $\frac{\Delta S}{S}$ over a short time interval is normally distributed with mean $\mu \Delta t$ and standard deviation $\sigma\sqrt{\Delta t}$, where μ is the drift and σ is the volatility. The shorthand for a normal distribution is

$$\frac{\Delta S}{S} \sim N\left(\mu \Delta t, \sigma \sqrt{\Delta t}\right).$$

The variance of this stochastic return is proportionally to the time interval, $\mathrm{var}\left(\frac{\Delta S}{S}\right) = \sigma^2 \Delta t$. [Hull 2006] One benefit of geometric Brownian motion is negative asset prices are not possible since any price change is proportional to the current price. Bankruptcy could drive an asset price down to but not past the natural absorbing barrier at zero. [Chance 1994]

The discrete approximation of the geometric Brownian motion stochastic equation is composed of a trend (or expectation) term $E\left(\frac{dS}{S}\right) = \mu dt$ and an uncertainty (of deviation) term. The uncertainty term is given by the Wiener increment

$$dW_t = \varepsilon_t \sqrt{dt},$$

with $E(dW_t) = 0$ where ε is the standard normal distribution. It turns out that the variance of dW_t is equal to the time interval dt. The variance of the Wiener increment was found by evaluating

$$Var(dW_t) = E\left[dW_t - \mu_{dW_t}\right]^2 = E(dW_t^2) - E(dW_t)^2 = E(dW_t^2).$$

The expected value of $(dW_t)^2$ is, by definition, $E\left[(dW_t)^2\right] = E\left[\left(\varepsilon_t \sqrt{dt}\right)^2\right]$. [Chance 2005] Pulling the dt factor out of the expectation gives

$$E(dW_t^2) = E(\varepsilon_t^2 dt) = \left(E(\varepsilon_t^2)\right)dt.$$

To evaluate the $E(\varepsilon_t^2)$ term requires the computational formula for the variance

$$Var(\varepsilon_t) = E[(\varepsilon_t - E(\varepsilon_t))^2] = E[(\varepsilon_t^2 - 2\varepsilon_t E(\varepsilon_t) + E(\varepsilon_t)^2]$$

$$Var(\varepsilon_t) = E(\varepsilon_t^2) - 2E(\varepsilon_t)E(\varepsilon_t) + E(\varepsilon_t)^2 = E(\varepsilon_t^2) - 2E(\varepsilon_t)^2 + E(\varepsilon_t)^2$$

$$Var(\varepsilon_t) = E(\varepsilon_t^2) - E(\varepsilon_t)^2$$

$$\longrightarrow E(\varepsilon_t^2) = Var(\varepsilon_t) - E(\varepsilon_t)^2.$$

The variance term of a standard normal variable is one, $Var(\varepsilon_t) = 1$. The expected value of a standard normal variable $E(\varepsilon_t)$ and the square of the expected value $E(\varepsilon_t)^2$ are zero. Therefore,

$$E(\varepsilon_t^2) = Var(\varepsilon_t) - E(\varepsilon_t)^2 = 1 - 0,$$

the expected value of a squared standard normal variable is one. The value for the expectation of the Wiener increment squared is given by

$$E(dW_t^2) = \left(E(\varepsilon_t^2)\right)dt = dt.$$

This important result states that the square of a Wiener process equals the time interval, $dt = dW_t^2$. In other words, the Weiner process is unpredictable but the square of the Weiner process is predictable.

The percentage price change of the stochastic representation of $\frac{dS_t}{S_t}$ is normally distributed since the stochastic differential equation, written as

$$\frac{dS_t}{S_t} = \mu dt + \sigma dW_t ,$$

is a linear transformation of the normally distributed variable dW_t. The *relative return* $\frac{S_0 + dS_t}{S_0} = 1 + \frac{dS_t}{S_0} = \frac{S_t}{S_0}$ over a time period T is the product of the intervening price changes as displayed by

$$\frac{S_T}{S_0} = \frac{S_1}{S_0}\frac{S_2}{S_1}\cdots\frac{S_{t-1}}{S_{t-2}}\frac{S_T}{S_{t-1}} ,$$

where each increment in relative return $\frac{S_i}{S_{i-1}}$ is capable of being further subdivided in time. A logarithm of the product converts the product series into a summation series as given by

$$\ln\left(\frac{S_T}{S_0}\right) = \ln\left(\frac{S_1}{S_0}\right) + \ln\left(\frac{S_2}{S_1}\right) + \cdots + \ln\left(\frac{S_{t-1}}{S_{t-2}}\right) + \ln\left(\frac{S_T}{S_{t-1}}\right) .$$

The central limit theorem states that the summation a large number of identically distributed and independent random variables, each with finite mean and variance, will be approximately normally distributed. [Rice 1995]

1.2.1 Log-normal stochastic differential equation

The insight of the previous paragraph provides a motivation to recast the geometric Brownian motion

$$dS = \mu S dt + \sigma S dW ,$$

as a log-normal diffusion stochastic differential equation. The alteration is accomplished by using the function

$$G = \ln S ,$$

and its derivatives

$$\frac{\partial G}{\partial S} = \frac{1}{S} \qquad \frac{\partial^2 G}{\partial S^2} = -\frac{1}{S^2} \qquad \frac{\partial G}{\partial t} = 0 .$$

Application of Ito's lemma gives

$$dG = \frac{\partial G}{\partial S}dS + \frac{1}{2}\frac{\partial^2 G}{\partial S^2}(dS)^2 + \frac{\partial G}{\partial t}dt + \frac{1}{2}\frac{\partial^2 G}{\partial t^2}\underbrace{(dt)^2}_{\substack{\text{infintisimally}\\ \text{small} \to 0}}$$

$$dG = \frac{1}{S}dS - \frac{1}{2}\frac{1}{S^2}(dS)^2 + 0$$

$$dG = \frac{1}{S}(\mu S dt + \sigma S dW) - \frac{1}{2}\frac{1}{S^2}(\mu S dt + \sigma S dW)^2$$

$$dG = (\mu dt + \sigma dW) - \frac{1}{2}\frac{1}{S^2}\left(\sigma^2 S^2 \underbrace{dW^2}_{dW^2 = dt} + \mu\sigma S^2 \underbrace{dt dW}_{\substack{\text{infintisimally}\\ \text{small} \to 0}} + \alpha^2 S^2 \underbrace{dt^2}_{\substack{\text{infintisimally}\\ \text{small} \to 0}}\right)$$

$$dG = (\mu dt + \sigma dW) - \frac{1}{2}\sigma^2 dt$$

$$dG = \Bigg(\underbrace{\mu - \frac{1}{2}\sigma^2}_{\eta}\Bigg)dt + \sigma dW$$

,

where we have introduced a *log-normal return drift factor*, η, which is the *continuously compounded return*. A solution to the log return stochastic differential equation is found by integrating

$$\int dG_u = \int \eta dt + \int \sigma dW_u$$

.

The deterministic drift term can be integrated similar to an ordinary differential equation. The σ coefficient is taken as a time-invariant constant, which greatly simplifies the stochastic integral to

$$\int_0^t \sigma dW_u = \sigma(W_t - W_0) = \sigma(W_t - 0)$$

,

where the random term is zero at the time origin by definition. The integral of $\int dG_u$ is G_t-G_0; therefore,

$$G_t - G_0 = \ln S_t - \ln S_0 = \ln\left(\frac{S_t}{S_0}\right) = \eta t + \sigma W_t$$

.

Hence the evolution of the logarithm of the price change follows a drift $\eta = \mu - \frac{1}{2}\sigma^2$ and has a Gaussian distribution with a mean at ηT and a variance σ^2T. Compactly this is expressed as

$$\ln S_t - \ln S_0 = \ln\left(\frac{S_t}{S_0}\right) \sim \phi\left[\eta t, \sigma\sqrt{t}\right]$$

$$\ln S_t \sim \phi\left[(\ln S_0 + \eta t), \sigma\sqrt{t}\right]$$

,

where $\phi(m,\sigma_{sd})$ denotes a normal distribution with mean m and standard deviation σ_{sd}. It follows that the *continuously compounded return* is found from

$$\eta = \frac{1}{t}\ln\left(\frac{S_t}{S_0}\right).$$

The continuously compounded return is normally distributed with a mean or expected value of $E(\eta) = \mu - \frac{\sigma^2}{2}$ and a standard deviation of $\frac{\sigma}{\sqrt{t}}$. [Hull 2006] Formally this is written as

$$\eta = \phi\left[\mu - \frac{\sigma^2}{2}, \frac{\sigma}{\sqrt{t}}\right],$$

which implies that likelihood of returns are more certain when examined over a longer time series.

Recalling that $S = e^G$ allows an expression for the evolution of S_t as

$$S_t = S_0 e^{\eta t + \sigma W_t}.$$

The variance of the stock price S_t has a lognormal distribution with a variance given by

$$\text{var}(S_t) = S_0^2 e^{2\alpha T}\left(e^{\sigma^2 T} - 1\right).$$

1.3 Expected Value, Variance and Moments of Lognormal Distribution

We will provide a brief proof of the lognormal distribution as discussed by Hull [2006]. The logarithm of the asset price, $G = \ln S$, has a normal distribution, $\phi(m,\sigma_{SD})$. The previous derivation showed that the mean of the distribution is dependent on logarithm of the starting stock price and a product of the continuously compounded rate and the time period of the analysis, $m = (\ln S_0 + \eta t)$; the standard deviation is a product of the annualized volatility and the square root of the time period of analysis, $\sigma_{SD} = \sigma\sqrt{t}$.

The related probability densities for G and ln(S) are

$$h(G) = \frac{1}{\sqrt{2\pi}\sigma_{SD}} e^{\left(\frac{-(G-m)^2}{2\sigma_{SD}^2}\right)} \qquad h(S) = \frac{1}{\sqrt{2\pi}\sigma_{SD} S} e^{\left(\frac{-(\ln(S)-m)^2}{2\sigma_{SD}^2}\right)}$$

The nth raw moment for a probability distribution h(S) is given by the integral

$$\mu'_n = E(S^n) = \int_0^\infty S^n h(S)\,dS,$$

where n signifies the nth moment. The nth moment after inserting the exponential of G, $e^G = S$, is

$$\mu_n' = E\left(\left(e^G\right)^n\right) = \int_{-\infty}^{\infty}\left(e^{nG}\right)\frac{1}{\sqrt{2\pi}\sigma_{SD}}e^{\left(\frac{-(G-m)^2}{2\sigma_{SD}^2}\right)}dG = \int_{-\infty}^{\infty}\left(e^{nG\frac{2\sigma_{SD}^2}{2\sigma_{SD}^2}}\right)\frac{1}{\sqrt{2\pi}\sigma_{SD}}e^{\left(\frac{-\left(G^2-Gm+m^2\right)^2}{2\sigma_{SD}^2}\right)}dG$$

$$\mu_n' = \int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}\sigma_{SD}}e^{\frac{2\sigma_{SD}^2nG-\left(G^2-Gm+m^2\right)^2}{2\sigma_{SD}^2}}dG = \int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}\sigma_{SD}}e^{\frac{2\sigma_{SD}^2nG-\left(G^2-Gm+m^2\right)^2}{2\sigma_{SD}^2}}dG$$

$$\mu_n' = \int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}\sigma_{SD}}e^{\frac{\left(G^2-m-\sigma_{SD}^2\right)^2}{2\sigma_{SD}^2}}e^{\frac{2mn\sigma_{SD}^2+n^2\sigma_{SD}^4}{2\sigma_{SD}^2}}dG = e^{mn+n^2\sigma_{SD}^2/2}\underbrace{\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}\sigma_{SD}}e^{\frac{\left(G^2-m-\sigma_{SD}^2\right)^2}{2\sigma_{SD}^2}}dG}_{\text{Integral of Normally Distributed Function}=1}$$

$$\mu_n' = e^{mn+n^2\sigma_{SD}^2/2}$$

Therefore the raw moments (taken about 0) are [Weisstein 2011]

$$\begin{aligned}
\mu_1' &= e^{m+\sigma_{SD}^2/2} = \exp ected \quad value \\
\mu_2' &= e^{2\left(m+\sigma_{SD}^2\right)} \\
\mu_3' &= e^{3m+9\sigma_{SD}^2/2} \\
\mu_4' &= e^{4\left(m+2\sigma_{SD}^2\right)}
\end{aligned}.$$

A little more mathematics with the first raw moment, $\mu_1' = e^{m+\sigma_{SD}^2/2}$, generates the expected value of the stock price. Substituting $m = \left(\ln S_0 + \eta t\right)$ and $\sigma_{SD} = \sigma\sqrt{t}$ yields

$$\begin{aligned}
\mu_1' = E(S_t) &= e^{\left(\ln S_0+\eta t\right)+\frac{1}{2}\sigma^2 t} = S_0 e^{\eta t+\frac{1}{2}\sigma^2 t} \\
E(S_t) &= S_0 e^{\left(\mu-\frac{1}{2}\sigma^2\right)t+\frac{1}{2}\sigma^2 t} \\
E(S_t) &= S_0 e^{\mu t}
\end{aligned}.$$

An alternate derivation of the expected value of lognormal stock price without invoking the first raw moment that may be more intuitive is presented below.

Log-Normal Distribution by Expectation

As discussed by Chance [2005], the solution to stochastic differential equation $\frac{dS_t}{S_t} = \mu dt + \sigma dW_t$ was developed as $S_t = S_0 e^{\eta t+\sigma W_t}$. Taking the expectation of the expression for the evolution of S_t gives

$$E\left[S_t\right] = E\left[S_0 e^{\eta t+\sigma W_t}\right],$$

which can be simplified by moving the constant factors out of the expectation to give

$$E\left[S_t\right] = S_0 e^{\eta t} E\left[e^{\sigma W_t}\right].$$

The Weiner process W_t follows a standard normal probability with a mean of zero and standard deviation of $\sqrt{t}$ as written by

$$f(W_t) = \frac{1}{\sqrt{2\pi t}} e^{-W_t^2 / 2t}$$
.

In general, the expected value of a random variable is the integral of the variable and its probability density function by $E(X) = \int_{-\infty}^{\infty} x f(x) dx$ or for g(X), an arbitrary function of X, the expected value is the integral of the inner product as given by $E[g(X)] = \int_{-\infty}^{\infty} g(x) f(x) dx$. Therefore, the expected value of the exponential of the Wiener process is

$$E\left[e^{\sigma W_t}\right] = \int_{-\infty}^{\infty} e^{\sigma W_t} f(W_t) dW_t = \int_{-\infty}^{\infty} e^{\sigma W_t} \frac{1}{\sqrt{2\pi t}} e^{-W_t^2 / 2t} dW_t$$
.

This expression can be placed into a more useful form by completing the square in the exponent by

$$E\left[e^{\sigma W_t}\right] = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi t}} e^{\frac{2t\sigma W_t - W_t^2}{2t}} dW_t = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi t}} e^{\frac{2t\sigma W_t - W_t^2 - \sigma^2 t}{2t} + \frac{\sigma^2 t}{2}} dW_t$$

$$E\left[e^{\sigma W_t}\right] = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi t}} e^{\frac{(W_t - \sigma t)^2}{2t} + \frac{\sigma^2 t}{2}} dW_t = e^{\frac{\sigma^2 t}{2}} \overbrace{\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi t}} e^{\frac{-1}{2}\left(\frac{W_t - \sigma t}{t}\right)^2} dW_t}^{\int \phi(\sigma t, \sqrt{t}) = 1}$$

$$E\left[e^{\sigma W_t}\right] = e^{\frac{\sigma^2 t}{2}}$$
.

The integral was eliminated by manipulating the expression into the form of a probability density function. The integral of a probability density function is intrinsically equal to one. Relying on the relation that the continuously compounded return η is normally distributed with a mean or expected value of $E(\eta) = \mu - \frac{\sigma^2}{2}$ allows the expected stock price to be written as

$$E(S_t) = S_0 e^{\left(\eta + \frac{1}{2}\sigma^2\right)t}$$

$$E(S_t) = S_0 e^{\mu t}$$
.

Moments and Variance of Log-Normal Distribution

Next, two related approaches are given to find the variance of S. The central moment μ_n taken about the expected value μ_1' is

$$\mu_n = \left\langle \left(S - \langle S \rangle\right)^n \right\rangle = \int \left(S - \mu_1'\right)^n h(S) dS$$
.

A more specific form of this equation is commonly used to express the variance of a process. The variance of S is given as

$$\mathrm{var}(S) = \mu_2 = \left\langle \left(S - \langle S \rangle\right)^2 \right\rangle$$
$$\mathrm{var}(S) = E(S^2) - [E(S)]^2 .$$

The term $E(S^2)$ is the raw moment $\mu_2' = e^{2\left(m+\sigma_{SD}^2\right)}$ and the expected value of a squared asset price is

$$[E(S)]^2 = (\mu_1')^2 = \left(e^{m+\sigma_{SD}^2/2}\right)^2 = e^{2\left(m+\sigma_{SD}^2/2\right)} .$$

Substituting these terms along with $m = \left(\ln S_0 + \eta t\right)$ and $\sigma_{SD} = \sigma\sqrt{t}$ yields variance of lognormal stock price

$$\mathrm{var}(S) = -(\mu_1')^2 + \mu_2' = e^{2\left(m+\sigma_{SD}^2\right)} - e^{\left(2m+\sigma_{SD}^2\right)}$$
$$\mathrm{var}(S) = e^{2m+\sigma_{SD}^2}\left(e^{\sigma_{SD}^2} - 1\right)$$
$$\mathrm{var}(S) = e^{2(\ln S_0+\eta t)+\sigma^2 t}\left(e^{\sigma^2 t} - 1\right)$$
$$\mathrm{var}(S_t) = S_0^2 e^{2\mu t}\left(e^{\sigma^2 t} - 1\right) .$$

Alternatively, using a binomial transform, not derived here, the central moments can be expressed in terms of the raw moments as given by

$$\mu_1 = 0$$
$$\mu_2 = -(\mu_1')^2 + \mu_2' = e^{2m+\sigma_{SD}^2}\left(e^{\sigma_{SD}^2} - 1\right)$$
$$\mu_3 = 2(\mu_1')^3 + 3\mu_1'\mu_2' + \mu_3' = e^{3m+3\sigma_{SD}^2/2}\left(e^{\sigma_{SD}^2} - 1\right)^2\left(e^{\sigma_{SD}^2} + 2\right)$$
$$\mu_4 = -3(\mu_1')^4 + 6(\mu_1')^2\mu_2' - 4\mu'\mu_3' + \mu_4' = e^{4m+2\sigma_{SD}^2}\left(e^{\sigma_{SD}^2} - 1\right)^2\left(e^{4\sigma_{SD}^2} + 2e^{3\sigma_{SD}^2} + 3e^{2\sigma_{SD}^2} - 3\right),$$

where the second central moment μ_2 yields the variance relative to the mean. [Papoulis 1984] Similarly the third and fourth central moments provide a construct for the skewness and kurtosis, respectively.

Log-Normal Distribution by Candidate Solution

Neftci [2000] provides an alternate approach to solve the stochastic differential equation

$$\int \frac{dS_u}{S_u} = \int \mu dt + \int \sigma dW_u .$$

Again, the Riemann and stochastic integration to the right side of the differential equation are solved to give

$$\int_0^t \frac{dS_u}{S_u} = \mu t + \sigma W_t .$$

Often in the financial literature a solution is not available so a candidate is proposed and back-checked in the original differential equation.

For example, the candidate given by

$$S_t = S_0 e^{\left\{(\mu - \frac{1}{2}\sigma^2)t + \sigma W_t\right\}},$$

is a strong solution in that the W_t is given exogenously and the error process is considered as another given in the equation. At this point Ito's lemma is employed to validate that the candidate solution satisfies the stochastic differential equation and the integral equation. The value S is a function of t and W with partial derivatives given by

$$\frac{\partial f}{\partial t} = \left(\mu - \frac{1}{2}\sigma^2\right)S_t \quad \frac{\partial f}{\partial z} = \sigma S_t \quad \frac{\partial^2 f}{\partial W^2} = \sigma^2 S_t.$$

Application of Ito's lemma

$$dS_u = \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial W}dW + \frac{1}{2}\frac{\partial^2 f}{\partial z^2}dW^2,$$

with $dW^2 = dt$ gives

$$dS_t = S_t\left[\left(\mu - \frac{1}{2}\sigma^2\right)dt + \sigma S_t dW_t + \frac{1}{2}\sigma^2 dt\right]$$
$$dS_t = S_t\left[\mu dt + \sigma dW\right],$$

where the original stochastic differential equation is recovered.

To provide some clarity to this section, the important equations for *geometric Brownian motion* are given in table 1-1.

$\frac{dS_t}{S_t} = \mu dt + \sigma dW_t$	Stochastic Stock Process for Geometric Brownian Motion
μ	Drift
$SD\left(\frac{dS_t}{S_t}\right) = \sigma\sqrt{\Delta t}$	Standard deviation of percentage change
$d(\ln S_t) = \left(\mu - \frac{\sigma^2}{2}\right)dt + \sigma dW_t$	Transformation to log price process via Ito's Lemma
$E(S_t) = S_0 e^{\mu t}$	Expected Value of S_t
$\eta = \mu - \sigma^2/2$ $\eta = \frac{1}{T}\ln\left(\frac{S_T}{S_0}\right)$	Continuously compounded return over a period of time length T
$u_i = \ln\left(\frac{S_i + D}{S_{i-1}}\right)$	Log return over time period T_i-T_{i-1}, e.g., daily return with potentially a dividend D
$\sigma = \frac{\sigma_{SD_\tau}}{\sqrt{\Delta t}}$	annualized volatility, σ, is the standard deviation of the asset's logarithmic returns in a year, e.g., σ_{SD} is standard

	deviation of daily logarithmic returns $\rightarrow$ 252 Trading Days / yr $\rightarrow$ $\Delta t = 1/252$
$\sigma_{SD} = \sigma\sqrt{t}$	Standard deviation over a time period t calculated from annualized volatility
$Var = (\sigma_{SD})^2 = \sigma^2 t$	Variance over a time period t calculated from annualized volatility
$\sigma_{SD} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(u_i - \overline{u})^2}$ $\sigma_{SD} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}u_i^2 - \frac{1}{n(n-1)}\left(\sum_{i=1}^{n}u_i\right)^2}$	General formula to calculate standard deviation from a data series, e.g., daily log return data. Matlab provides a built-in function std for this calculation.

Table 1-1. Summary of parameters for geometric Brownian motion.

The Black-Scholes option model is based on geometric Brownian motion,

$$dS_t = \mu S_t dt + \sigma S_t dz ,$$

and assumes a constant volatility, σ, or an effective volatility over the life of the contract, In the limiting case, a constant volatility is the square root of a constant variance, $\sigma = \sqrt{Var}$. For non-constant variance, the effective volatility is found as the square root of the mean of time-weighted variance or squared volatility over time,

$$\sigma = \sqrt{\frac{T_1 Var_1 + T_2 Var_2 + T_3 Var_3 + \cdots}{T_{total}}} = \sqrt{\frac{T_1\sigma^2 + T_2\sigma^2 + T_3\sigma^2 + \cdots}{T_{total}}} ,$$

where T_i is the time length of the ith period.

1.4 Fitting Geometric Brownian Motion

The code to simulate and analyze geometric Brownian motion is provided below as the function *GBM(S)*. The function only accepts one parameter S, which is the daily price of an asset assumed to follow geometric Brownian motion. Figure 1-1 is the graphical output of the function where the jagged line is the adjusted close price for Exxon Mobil stock. The continuously compounded return η is found by calculating the logarithm of the daily price change and then normalized for the daily time period. The daily standard deviation $\sigma_{SD} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(u_i - \overline{u})^2}$ is calculated by the Matlab *std* function. The approach in *GBMfit(S)* to calculate standard deviation is

$$\sigma_{SD} = std\left[\ln(S_t) - \ln(S_{t-1})\right] .$$

The annualized volatility is found as the daily standard deviation divided by the square-root of the time period measured in years, e.g., 1/252 for one day,

$$\sigma = \sigma_{SD} / \sqrt{\Delta t}\ .$$

Now the expected return per year can be determined from the mean of the continuously compounded return as given by

$$mean\left[\ln\left(\frac{S_t}{S_{t-1}}\right)\right] = mean[\ln(S_t) - \ln(S_{t-1})] = \eta = \mu - \sigma^2/2\ .$$

The expected return μ demanded by investors depends on the continuously compounded return η and the volatility risk of the stock, $\mu = \eta + \sigma^2/2$.

For the Exxon Mobil data from 2003 to mid-2010, a volatility σ = 0.2718 and an expected return μ = 0.096 are estimated. The solid central line of the expected value of the stock process $E(S_t) = S_0 e^{\mu t}$ is also called the mean future stock price.

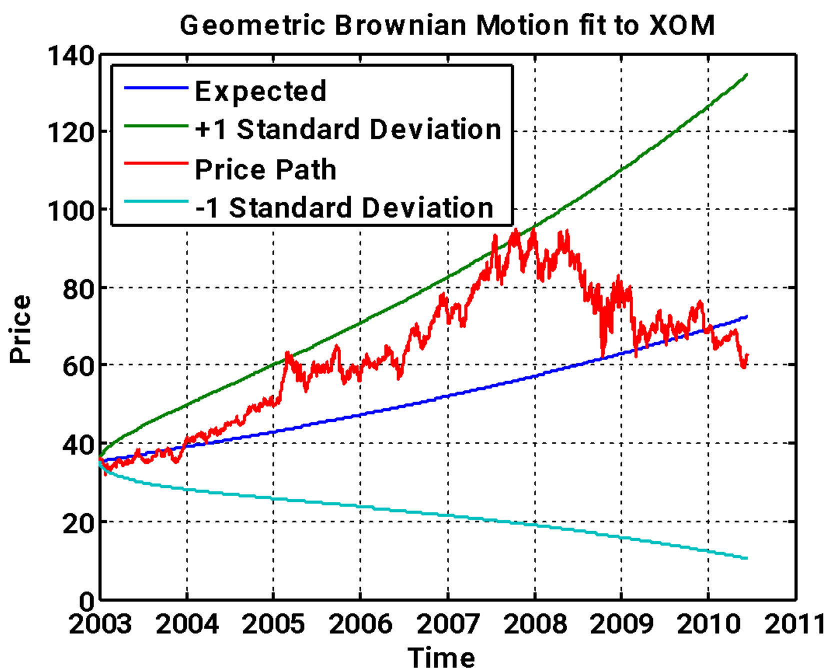

Figure 1-1. Price process of Exxon Mobil adjusted close stock price fitted to a geometric Brownian motion with volatility σ = 0.27 and return μ = 0.096.

1.5 Mean Price Simulation

When the function *GBMfit* is called without an argument the function will self-simulate a price process based on internally given parameters and then analyze this self-generated price process. The price path is formed by iteratively stepping through

$$S_{t+\Delta t} = S_t e^{\left\{(\mu - \frac{1}{2}\sigma^2)\Delta t + \sigma\sqrt{\Delta t}N(0,1)\right\}},$$

where N(0,1) is a random draw from a normal distribution. The function *GBMexpected* is used to generate several thousand simulated price paths. To speed up the execution, some Matlab vectorization and built-in functions are used. Specifically, the iterative step is accomplished via the built-in *cumprod* function, which calculates a cumulative product in time from an array of exponential drift and random movements. An alternate approach would be to use the built-in *cumsum* function for the logarithmic sum of the drift and random movements

Analyzing the price simulation formula shows that the medium stock price after one time step would be

$$S_{t+1}^{median} = S_t e^{\left(\mu - \frac{1}{2}\sigma^2\right)t},$$

that is, half the random movements in the asset price will fall above or below the medium value point. Contrast this to our earlier derivation of the expected asset price given as

$$S_t = S_0 e^{\mu t}.$$

What is different is that the geometric Brownian motion is not symmetric and one feature of the lognormal distribution is that the one-step mean future stock price is higher than the one-step median future stock price. For example, a symmetric random movement of +/- 0.1 is not symmetric in the exponential $e^{+/-0.1} = 0.905 / 1.105$.

Figure 1-2 clearly shows that the mean or expected price process will be higher as a consequence of the long tail in the stock price distribution towards higher asset prices. This repeated simulation is a Monte Carlo analysis in its most basic form.

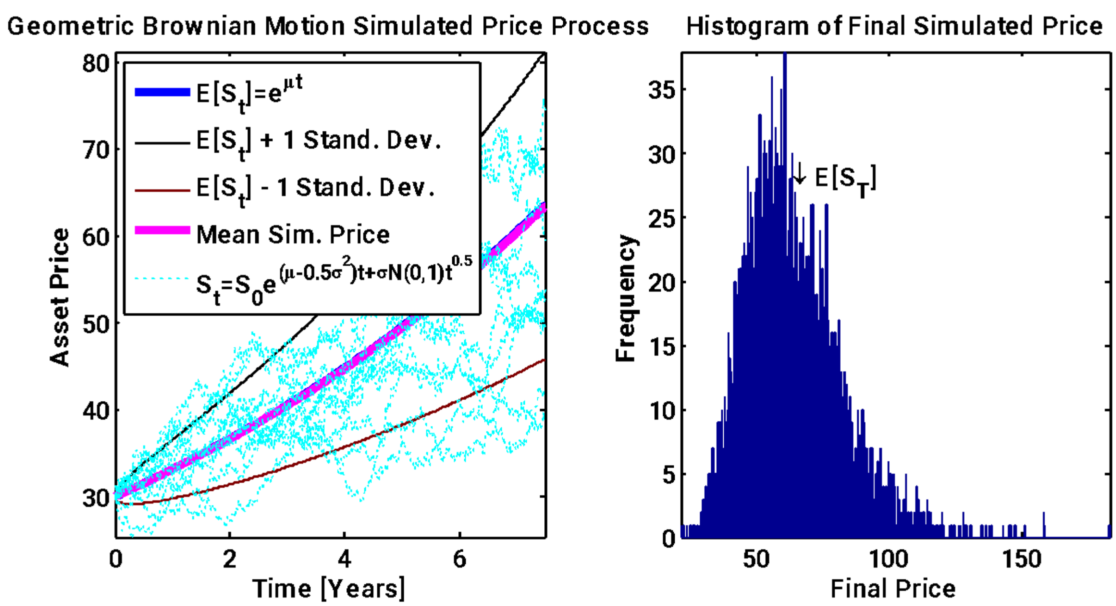

Figure 1-2. (left) Simulated geometric Brownian motion daily price process simulated. The mean of all the simulated price paths is exactly equal to the calculated expected value. (right) Examining the log-normal distribution of prices shows that the high-price tail shifts the mean or expected value above the peak in the distribution.

1.6 Mean Reversion Models

The first description of an ordinary mean reversion process was by Ornstein and Uhlenbeck. [1930] The Ornstein–Uhlenbeck process is the continuous-time analogue of the discrete-time AR(1) process. The behavior and economic principle of commodities, interest rates, and foreign exchange rates are well described by reversion to a mean. The microeconomic viewpoint is that the long term marginal production cost of a commodity, such as oil, determines the long run cost. [Dias 2004] Bessembinder et al. [1995] show that significant mean-reversion is observable for prices of agricultural and oil commodities.

An alternate viewpoint, which reaches the same conclusion is that a cartel will target a consistent price level. This target point may vary but the underlying profit level targets and political motivation tend not to change in the short term. [Laughton 1995] Pindyck and Rubinfeld [1991] examined over one hundred years of oil price data and found a slow mean reversion but a Dickey-Fuller unit root test rejected a simple random walk process. Baker et al. [1998] found that a mean reversion model was more consistent with the inter-relationship between oil price spot and futures data. Specifically, the spot price data is more volatile than the futures price data as is predicted by a mean-reversion model. For reference, a random walk model predicts equal volatility in futures and spot data. Additionally, Baker et al. [1998] showed that low spot prices tend to associate with futures prices increasing towards the long-run equilibrium, i.e., in contango; and a high spot price tends to associate with futures prices decreasing towards the long-run equilibrium, i.e., backwardation.

Unlike the geometric Brownain motion process, an arithmetic Brownian motion can have negative stochastic movements, which will result in a negative asset price. Cox [1985] developed a square root model that effectively prevents negative random movements below zero as given by

$$dr_t = \kappa(\theta - r_t)dt + \sigma\sqrt{r_t}dz \, .$$

A useful implementation for modeling commodity prices occurs by examining the logarithm of price, x = ln(S). In this form, a negative spot price is prevented as the negative logarithm of the spot price x maintains a positive spot price.

Several mean-reverting forms have been proposed in the literature to model commodity prices. One example is a geometric mean-reverting price process, which is also referred to as the Dixit and Pindyck [1994] model, as given by

$$dS = \lambda S(\mu - S)dt + \sigma S dz \, ,$$

where λ is the reversion rate and μ is the long-run mean price. If the price S is higher than the mean price μ then the negative $(\mu - S)$ factor pulls the price level down at a rate determined by the reversion rate λ. A large $(\mu - S)$ delta implies a faster rate of reversion.

Reversion due to a small $(\mu - S)$ delta may be difficult to differentiate from the stochastic variation generated by the σPdz term. This model displays log-normal diffusion similar to a non-mean-reverting geometric Brownian motion model; however, the variance increases with time up only until a stabilization level is reached.

1.7 Solving the Ornstein-Uhlenbeck process

Generally, the arithmetic Ornstein-Uhlenbeck process as given by

$$dx_t = \lambda(\mu - x_t)dt + \sigma dW_t$$

is pulled towards an equilibrium level μ at a rate λ and the σ is the volatility or average magnitude, per square-root time, of the random fluctuations that are modeled as Brownian motion. Integration of the deterministic term gives the expected value as

$$dx = \lambda(\mu - x)dt$$

$$u = (\mu - x) \quad du = -x$$

$$\int \frac{-du}{u} = \int \lambda dt$$

$$\ln|u| = -\lambda t + C$$

$$|u| = \pm e^{C} e^{-\eta t}$$

$$u_0 = Ae^{-\eta 0} \rightarrow A = u_0 = (m - x_0)$$

$$\mu - x = (m - x_0)e^{-\lambda t}$$

$$E[x(t)] = \mu + (x_0 - \mu)e^{-\lambda t}$$

$$E[x(t)] = x_0 e^{-\lambda t} + \mu(1 - e^{-\lambda t}) .$$

Thus, the expected value is approaching the long-term equilibrium price at a rate proportional to the present displacement from the equilibrium price.

Determining the stochastic integral of the arithmetic Ornstein-Uhlenbeck process requires a variation of parameters procedure to define a new function as

$$f(x_t, t) = x_t e^{\lambda t} ,$$

with the derivative found via Ito's lemma

$$df(x_t, t) = \lambda x_t e^{\lambda t} dt + e^{\lambda t} dx_t$$

$$df(x_t, t) = \lambda x_t e^{\lambda t} dt + e^{\lambda t} \lambda(\mu - x_t)dt + e^{\lambda t} \sigma dW$$

$$df(x_t, t) = e^{\lambda t} \lambda \mu_t dt + e^{\lambda t} \sigma dW .$$

Integration gives

$$f(x_t, t) = x_t e^{\lambda t} = x_t + \int_0^t e^{\lambda s} \lambda \mu ds + \int_0^t e^{\lambda s} \sigma dW_s$$

$$x_t = x_t e^{-\lambda t} + \mu(1 - e^{-\lambda t}) + \int_0^t e^{\lambda(s-t)} \sigma dW_s ,$$

where the mean is the first two terms as given by

$$E[x_t] = x_t e^{-\lambda t} + \mu(1 - e^{-\lambda t}) .$$

The variance is found from the integral of the stochastic process by

$$\mathrm{var}(x_t) = E[(x_t - E[x_t])^2]$$

$$\mathrm{var}(x_t) = \sigma^2 e^{-2\lambda t} E\left[\int_0^t e^{2\lambda s} \sigma dW_s\right] = \frac{\sigma^2}{2\lambda} e^{-2\lambda t}(e^{2\lambda t} - e^0)$$

$$\mathrm{var}(x_t) = \frac{\sigma^2}{2\lambda}\left(1 - e^{-2\lambda t}\right) .$$

Therefore the long-term (stationary) variance and standard deviation are

$$\mathrm{var}(x_t) = \frac{\sigma^2}{2\lambda} \quad SD(x_t) = \sqrt{\frac{\sigma^2}{2\lambda}}.$$

1.8 Simulating the Ornstein-Uhlenbeck process

Based on the previous derivation, a simulation is given as the sum of the mean and the stochastic fluctuations for an asset price as

$$S_t = S_0 e^{-\lambda t} + \mu(1 - e^{-\lambda t}) + \sqrt{\frac{\sigma^2}{2\lambda}\left(1 - e^{-2\lambda t}\right)} N(0,1),$$

where the time interval t can be arbitrarily large or small since this is an exact solution to the Ornstein-Uhlenbeck process. In this form, random movements are generated by multiplying the magnitude of the standard deviation with a random sampling from the standard normal distribution N(0,1).

The function *MRpath* consists of a simulation followed by a series of calibration approaches. Focusing first on the simulation, *MRpath* calculates a vector of the expected (mean) price path as well as vector of random movements. The summation of the expected and stochastic movements is displayed as the jagged line in figure 1-3. Starting from initial price of 5, the asset price is pulled towards an equilibrium price μ at a rate determined by the mean reversion parameter λ. The confidence interval is shown by the two lines displaced by plus or minus one standard deviation from the expected price path. An interesting effect of the Ornstein-Uhlenbeck process is that the variance of the process initially grows but then tends to a constant long-term variance as given by $\mathrm{var}_{long-term}(x_t) = \frac{\sigma^2}{2\lambda}$.

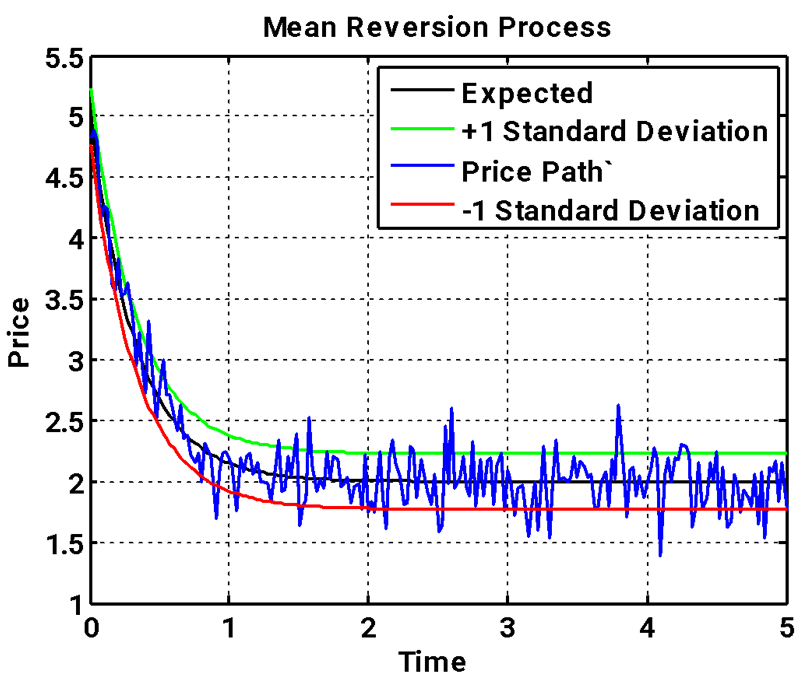

Figure 1-3. Ornstein-Uhlenbeck process initiated at a price of 5 displays a strong tendency to revert to the equilibrium price of 2.

1.9 Calibrating the Ornstein-Uhlenbeck process

There a many techniques to regress or fit data and here we will focus on two effective techniques, namely least squares fitting and maximum likelihood fitting. The function MRpath calibrates the simulated path with function calls to *WeightedLeastSquaresOU* and *MLweightedOU*. The former is a linear least squares fit that has been modified to put the solution of the Ornstein-Uhlenbeck process into a workable form.

1.10 Least Squares Fitting

The function *WeightedLeastSquaresOU* defaults when the StdDev parameter is null (or blank) to a linear unweighted least squares fit. This regression finds a best fit line through a set of data points. The linear least squares procedure fits a straight line, y=ax+b+ε, with slope m, intercept b, and an error term ε, to a set of data by minimizing the sum of squared error residuals. The squared error residuals meet the necessity for a continuous differential quantity in contrast to an absolute error residual which may not be a continuous differential quantity. One characteristic of squared error residuals is that a few outliers may have more influence than the majority of the data points. [Weisstein 2011]

The general least squares procedure finds a set of parameters that minimizes the squared vertical offsets of the data from the best fit line, plane, etc. The method is quite flexible as it can be applied to any linear combination of basis functions $X_k(x)$ including sines, cosines, or a polynomial as given by

$$y(x) = \sum_{k=1}^{k} a_k X_k(x) \; .$$

This approach is to fit a function with an arbitrarily large number of linear parameters a_i to minimize the squared deviation as given by

$$\varepsilon^2 = \sum \left[y_i - f\left(x_i, a_1, a_2, \ldots a_n\right)\right]^2 \; .$$

A minimum of a convex function exists when the first derivative is zero with respect to each dependent variable as given by

$$\frac{\partial\left(\varepsilon^2\right)}{\partial a_i} = 0 \; .$$

The simplest application and the form we are interested in is a linear fit to $f\left(a,b\right) = a + bx$. Deviations from this line are a function of the slope a and intercept b as given by

$$\varepsilon^2\left(a,b\right) = \sum_{i=1}^{n} \left[y_i - \left(a + bx_i\right)\right]^2 \; ,$$

with respective minima found at

$$\frac{\partial(\varepsilon^2)}{\partial a} = -2\sum_{i=1}^{n}\left[y_i - (a + bx_i)\right] = 0$$
$$\frac{\partial(\varepsilon^2)}{\partial b} = -2\sum_{i=1}^{n}\left[y_i - (a + bx_i)\right]x_i = 0$$
.

This leads to two coupled equations given by

$$\sum_{i=1}^{n} y_i = na + b\sum_{i=1}^{n} x_i$$
$$\sum_{i=1}^{n} y_i x_i = a\sum_{i=1}^{n} x_i + b\sum_{i=1}^{n} x_i^2$$
,

which can be solved by substitution. Equivalently the linear least squares can be expressed as matrices [Weisstein 2011] as

$$\begin{bmatrix} \sum_{i=1}^{n} y_i \\ \sum_{i=1}^{n} y_i x_i \end{bmatrix} = \begin{bmatrix} n & \sum_{i=1}^{n} x_i \\ \sum_{i=1}^{n} x_i & \sum_{i=1}^{n} x_i^2 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix}$$
,

where the parameters a and b can be solved by the Matlab backslash operator or by a standard matrix inversion as

$$\begin{bmatrix} a \\ b \end{bmatrix} = \frac{1}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2} \begin{bmatrix} \sum_{i=1}^{n} y_i \sum_{i=1}^{n} x_i^2 - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i y_i \\ n\sum_{i=1}^{n} x_i^2 - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i \end{bmatrix}$$
.

This provides a direct solution to linear unweighted least squares.

Applying this regression procedure to Ornstein-Uhlenbeck solution

$$S_i = S_{i-1}e^{-\lambda\delta_t} + \mu(1 - e^{-\lambda\delta_t}) + \sigma\sqrt{\frac{\left(1 - e^{-2\lambda\delta_t}\right)}{2\lambda}}N(0,1)$$
,

requires viewing this equation with the present observation S_i as the y-data and the previous observation S_{i-1} as the x-data for a time change δ_t as given by $S_i=a+bS_{i-1}+\varepsilon$. Following van den Berg [2007], the model variables are directly substituted as

$$slope = b = e^{-\lambda\delta_t}$$
$$intercept = a = \mu(1 - e^{-\lambda\delta_t})$$
$$sd = \sigma\sqrt{\frac{\left(1 - e^{-2\lambda\delta_t}\right)}{2\lambda}}$$
,

which can be inverted to give

$$\lambda = \ln b \big/ \delta_t$$

$$\mu = a \big/ (1 - e^{-\lambda \delta_t}) = a \big/ (1 - b)$$

$$\sigma = sd \sqrt{\frac{2\lambda}{(1 - e^{-2\lambda\delta_t})}} = sd \sqrt{\frac{-2\ln b}{\delta_t\left(1 - b^2\right)}} .$$

To add flexibility, we will derive a weighted least squares approach where the weight w_i is inversely proportional to the square of the standard deviation (or measurement error) of each data point x_i by $w_i = 1 \big/ \sigma_i^2$. The weighted general least squares merit function is

$$\varepsilon^2 = \sum_{i=1}^{n} \left[\frac{y_i - \sum_{k=1}^{k} a_k X_k(x)}{\sigma_i} \right]^2 .$$

The weights and measurement error are often unknown and simply set to unity to recover the unweighted least square formula. Here we present the weighted merit function fit to a line with an intercept a and a slope b as

$$\varepsilon^2(a,b) = \sum_{i=1}^{n} \left[\frac{y_i - (a + bx_i)}{\sigma_i} \right]^2 .$$

If the errors are normally distributed, i.e., a Gaussian distribution, then this approach will replicate the maximum likelihood estimate (MLE). [Press 1989] The maximum likelihood approach will be derived in a slightly different manner in the next section. The maximum likelihood approach can be applied to a known distribution of any type, e.g., Gaussian, exponential, etc. The least squares approach is powerful in that it provides a good estimate in most cases even if no information is available as to the size or distribution type of the measurement error.

The minima of a weighted convex function is expressed as

$$\frac{\partial\left(\varepsilon^2\right)}{\partial a} = -2 \sum_{i=1}^{n} \frac{\left[y_i - (a + bx_i)\right]}{\sigma_i^2} = 0$$

$$\frac{\partial\left(\varepsilon^2\right)}{\partial b} = -2 \sum_{i=1}^{n} \frac{x_i\left[y_i - (a + bx_i)\right]}{\sigma_i^2} = 0 .$$

This again gives two coupled equations. To simplify the algebra in the code, several convenient sums will be used as given by

$$S = \sum_{i=1}^{n} \frac{1}{\sigma_i^2} \qquad S_x = \sum_{i=1}^{n} \frac{x_i}{\sigma_i^2} = \sum_{i=1}^{n} \frac{S_{i-1}}{\sigma_i^2}$$

$$S_y = \sum_{i=1}^{n} \frac{y_i}{\sigma_i^2} = \sum_{i=1}^{n} \frac{S_i}{\sigma_i^2} \qquad S_{xx} = \sum_{i=1}^{n} \frac{x_i^2}{\sigma_i^2} = \sum_{i=1}^{n} \frac{S_{i-1}^2}{\sigma_i^2}$$

$$S_{yy} = \sum_{i=1}^{n} \frac{y_i^2}{\sigma_i^2} = \sum_{i=1}^{n} \frac{S_i^2}{\sigma_i^2} \qquad S_{xy} = \sum_{i=1}^{n} \frac{x_i y_i}{\sigma_i^2} = \sum_{i=1}^{n} \frac{S_i S_{i-1}}{\sigma_i^2} .$$

Substituting these sums into the two coupled equations gives

$$S_y = aS + bS_x$$

$$S_{xy} = aS_x + bS_{xx}.$$

The two coupled equations can be solved for the two unknowns, intercept a and slope b, as well as the standard deviation as

$$slope = b = \frac{SS_{xy} - S_x S_y}{SS_{xx} - S_x^2}$$

$$intercept = a = \frac{S_{xx} S_y - S_x S_{xy}}{SS_{xx} - S_x^2}$$

$$sd = \sqrt{\frac{SS_y - S_y^2 - b(SS_{xy} - S_x S_y)}{S(S-2)}}.$$

These unknowns allow direct calculation of the parameters in the Ornstein-Uhlenbeck model.

The function *WeightedLeastSquaresOU* is called with a vector of asset prices, a constant time delta parameter, and an optional vector of measurement errors. If the last parameter is absent or each parameter is equal to a single value then the function runs as an unweighted least squares fit. Otherwise, the weights correspond to the confidence or importance of the data.

In the function *WeightedLeastSquaresOU*, the earlier data is weighted to coincide with data that is far from the equilibrium level. This was done to improve the fit to λ, which determines the reversion to the mean. The price movements far from equilibrium are dominated by λ. Near equilibrium, the true reversion rate is obscured by noise (via sigma) in the data.

A graphical representation of weighted and unweighted least squares best fit lines are displayed in figure 1-4 along with the current vs. previous price data. The data near the equilibrium level of 2 are clustered at the bottom left hand corner of figure 1-4. The price points far from equilibrium are represented by moving up and right in figure 1-4.

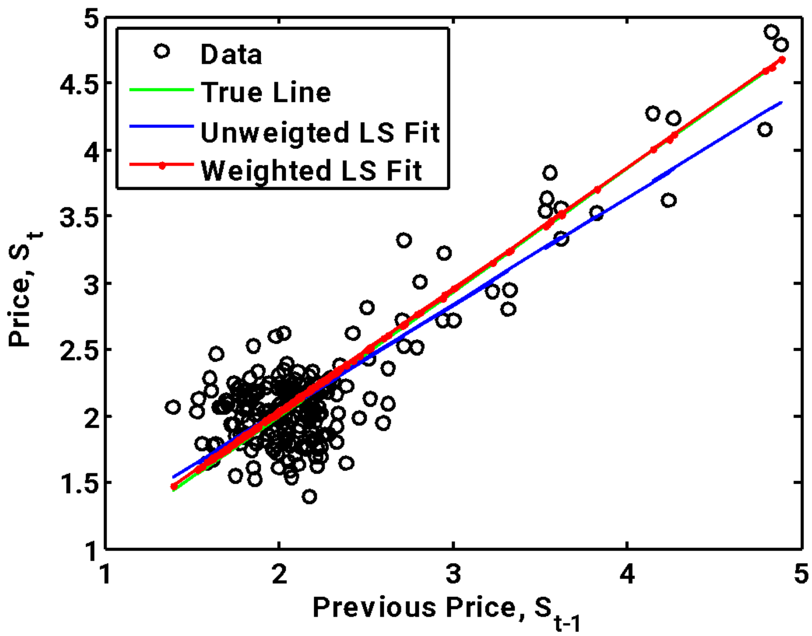

Figure 1-4. Graphical output of function *MRpath* depicting the unweighted and weighted least squares best fit line to the current price vs. previous price data.

The text output of *MRpath* gives numerical values for the underlying parameters of the various best fit lines to the S_t vs. S_{t-1} data:

	Mu	*lambda*	*sigma*	*Slope*	*Intercept*	*Standard Deviation*
True	*2.00*	*3.00*	*1.50*	*0.93*	*0.14*	*0.61*
Standard LS	*2.12*	*8.52*	*2.01*	*0.81*	*0.41*	*0.29*
Weight LS	*2.39*	*3.45*	*1.83*	*0.92*	*0.20*	*0.28*
Standard ML	*2.12*	*8.52*	*2.00*			
Weight ML	*2.39*	*3.45*	*1.83*			

The least squares procedures provide fairly good fits to the data. A large volatility and short time step δ_t (in contrast to the example by van den Berg [2007]) was selected to make the data more noisy and thus more challenging to fit. Nevertheless, in this limited example, the weighted approach does provide a better fit to the mean reversion parameter λ.

1.11 Maximum likelihood

The previous section sought to find the mean square error under the assumption that the expected variation of the observed data is best modeled as a Gaussian distribution. From another viewpoint, the minimization of the mean square error provides an estimate that maximizes the likelihood of the observed data. This approach can be generalized to find the maximum likelihood of any particular distribution chosen to fit the data. Usually, the distribution type, e.g., Gaussian, Bernouilli, Poisson, etc., is known but one or more of the parameters describing the distribution are not know. [Weisstein 2011] The conditional density function for n sequential data points x_i with a normal distribution for any mean μ and standard deviation σ is

$$f\left(x_1,\ldots,x_n \mid \mu,\sigma\right)=\prod \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x_i-\mu)^2}{2\sigma^2}} = \frac{2\pi^{(-n/2)}}{\sigma^n} e^{-\frac{\sum (x_i-\mu)^2}{2\sigma^2}}.$$

The logarithm of the likelihood is more convenient for estimation as it is expressed as a summation rather than a multiplication. The log-likelihood function *L* is given as

$$L=\sum_{i=1}^{n}\ln f=\frac{-1}{2}n\ln\left(2\pi\right)-n\ln\left(\sigma\right)-\frac{\sum\left(x_i-\mu\right)^2}{2\sigma^2}.$$

The maximum of a concave function is found when the first derivative with respect to the dependent parameters is zero,

$$\frac{\partial L}{\partial \mu}=0=\frac{\sum\left(x_i-\mu\right)}{\sigma^2}\rightarrow \sum\left(x_i-\mu\right)=0.$$

Rearranging this relation finds the mean μ that makes the function f the most likely,

$$\mu=\frac{\sum x_i}{n},$$

which clearly is the expression for an average. Similarly,

$$\frac{\partial L}{\partial \sigma}=\frac{n}{\sigma}+\frac{\sum\left(x_i-\mu\right)^2}{\sigma^3}=0,$$

recovers the usual formula for standard deviation,

$$\sigma=\sqrt{\frac{\sum\left(x_i-\mu\right)^2}{n}}.$$

Using our format above for the Ornstein-Uhlenbeck model,

$$S_i=S_{i-1}e^{-\lambda\delta_t}+\mu(1-e^{-\lambda\delta_t})+\sigma\sqrt{\frac{\left(1-e^{-2\lambda\delta_t}\right)}{2\lambda}}N(0,1),$$

the conditional probability density f_i of observation S_i given previous observation S_{i-1} after a time step δ_t is

$$f\left(S_i \mid S_{i-1},\mu,\hat{\sigma},\lambda\right)=\frac{1}{\sqrt{2\pi\hat{\sigma}^2}} e^{-\frac{\left(S_i-\overbrace{S_{i-1}e^{-\lambda\delta_t}-\mu(1-e^{-\lambda\delta_t})}^{mean}\right)^2}{2\hat{\sigma}^2}},$$

where

$$\hat{\sigma}=sd=\sigma\sqrt{\frac{\left(1-e^{-2\lambda\delta_t}\right)}{2\lambda}}.$$

was previously referred to as sd. Our initial derivation will assume $\hat{\sigma}$ is constant. Following van den Berg [2007], given n + 1 observations $\{S_0,\ldots,S_n\}$, the log-likelihood function is

$$L=\sum_{i=1}^{n}\ln f\left(S_i \mid S_{i-1},\mu,\hat{\sigma},\lambda\right)=-\frac{n}{2}\ln\left(2\pi\right)-n\ln\left(\hat{\sigma}\right)-\frac{1}{2\hat{\sigma}^2}\sum_{i=1}^{n}\left(S_i-S_{i-1}e^{-\lambda\delta_t}-\mu(1-e^{-\lambda\delta_t})\right)^2.$$

The parameters $\hat{\sigma}$ and μ that maximize the likelihood function are found by setting the partial derivatives of the log-likelihood function equal to zero by

$$\frac{\partial L(\mu,\hat{\sigma},\lambda)}{\partial \mu} = \frac{1}{\hat{\sigma}^2}\sum_{i=1}^{n}\left(S_i - S_{i-1}e^{-\lambda\delta_t} - \mu(1-e^{-\lambda\delta_t})\right) = 0$$

$$\rightarrow \mu = \frac{\sum_{i=1}^{n}\left(S_i - S_{i-1}e^{-\lambda\delta_t}\right)}{n(1-e^{-\lambda\delta_t})}$$

$$\frac{\partial L(\mu,\hat{\sigma},\lambda)}{\partial \hat{\sigma}} = \frac{n}{\hat{\sigma}} - \frac{1}{\hat{\sigma}^3}\sum_{i=1}^{n}\left(S_i - S_{i-1}e^{-\lambda\delta_t} + \mu(1-e^{-\lambda\delta_t})\right)^2 = 0$$

$$\rightarrow \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}\left(S_i - S_{i-1}e^{-\lambda\delta_t} + \mu(1-e^{-\lambda\delta_t})\right)^2 .$$

Finding optimal choice for the mean reversion parameter λ is simpler after a slight rearrangement of the log likelihood function to isolate the exponential functions by

$$L = -\frac{n}{2}\ln(2\pi) - n\ln(\hat{\sigma}) - \frac{1}{2\hat{\sigma}^2}\sum_{i=1}^{n}\left((S_i-\mu) - e^{-\lambda\delta_t}(S_{i-1}-\mu)\right)^2$$

$$L = -\frac{n}{2}\ln(2\pi) - n\ln(\hat{\sigma}) - \frac{1}{2\hat{\sigma}^2}\sum_{i=1}^{n}\left((S_i-\mu)^2 - 2e^{-\lambda\delta_t}(S_i-\mu)(S_{i-1}-\mu) + e^{-2\lambda\delta_t}(S_{i-1}-\mu)^2\right) .$$

The partial derivative of the log likelihood function with respect to λ is set to zero to find the optimal mean reversion parameter λ by

$$\frac{\partial L(\mu,\hat{\sigma},\lambda)}{\partial \lambda} = -\frac{\delta_t e^{-\lambda\delta_t}}{2\hat{\sigma}^2}\sum_{i=1}^{n}\left((S_i-\mu)(S_{i-1}-\mu) - e^{-\lambda\delta_t}(S_{i-1}-\mu)^2\right) = 0$$

$$\rightarrow \lambda = -\frac{1}{\delta_t}\ln\left(\frac{\sum_{i=1}^{n}\left((S_i-\mu)(S_{i-1}-\mu)\right)}{\sum_{i=1}^{n}\left((S_{i-1}-\mu)^2\right)}\right) .$$

In the previous section on least squares fitting, it was shown that individually weighting data points can improve the fit in some situations. It is thus beneficial to add a similar capability to our maximum likelihood approach. The basic idea is to re-derive the equations of this section with $\hat{\sigma}_i$ inside the summation or product. This approach takes some liberties with the underlying concept of the Gaussian distribution; however, this exercise is interesting when the numerical values of the weighted maximum likelihood and weighted least squares are compared. Briefly, the log-likelihood function is

$$L = \sum_{i=1}^{n}\ln f\left(S_i \mid S_{i-1},\mu,\hat{\sigma}_i,\lambda\right)$$

$$L = -\frac{n}{2}\ln(2\pi) - \sum_{i=1}^{n}\left[\ln(\hat{\sigma}_i)\right] - \sum_{i=1}^{n}\left[\frac{1}{2\hat{\sigma}_i^2}\left(S_i - S_{i-1}e^{-\lambda\delta_t} - \mu(1-e^{-\lambda\delta_t})\right)^2\right] ,$$

and the optimal parameters are

$$\mu = \frac{\sum_{i=1}^{n} \left(S_i - S_{i-1}e^{-\lambda\delta_t}\right) \Big/ \hat{\sigma}_i^2}{n(1-e^{-\lambda\delta_t})\sum_{i=1}^{n} 1 \Big/ \hat{\sigma}_i^2}$$

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}\left(S_i - S_{i-1}e^{-\lambda\delta_t} + \mu(1-e^{-\lambda\delta_t})\right)^2$$

$$\lambda = -\frac{1}{\delta_t}\ln\left(\frac{\sum_{i=1}^{n}\left(\frac{(S_i-\mu)(S_{i-1}-\mu)}{\hat{\sigma}_i^2}\right)}{\sum_{i=1}^{n}\left(\frac{(S_{i-1}-\mu)^2}{\hat{\sigma}_i^2}\right)}\right).$$

The equation just derived for $\hat{\sigma}$ is dependent on both μ and λ. Fortunately, the two coupled equations for μ and λ are only dependent on each other. Therefore, either μ or λ can be solved for λ or μ, respectively. Once μ and λ are know then $\hat{\sigma}$ can be solved directly. Again, to simplify the algebra in the code, several convenient sums will be used as given by

$$S = \sum_{i=1}^{n}\frac{1}{\sigma_i^2} \qquad S_x = \sum_{i=1}^{n}\frac{x_i}{\sigma_i^2} = \sum_{i=1}^{n}\frac{S_{i-1}}{\sigma_i^2}$$

$$S_y = \sum_{i=1}^{n}\frac{y_i}{\sigma_i^2} = \sum_{i=1}^{n}\frac{S_i}{\sigma_i^2} \qquad S_{xx} = \sum_{i=1}^{n}\frac{x_i^2}{\sigma_i^2} = \sum_{i=1}^{n}\frac{S_{i-1}^2}{\sigma_i^2}$$

$$S_{yy} = \sum_{i=1}^{n}\frac{y_i^2}{\sigma_i^2} = \sum_{i=1}^{n}\frac{S_i^2}{\sigma_i^2} \qquad S_{xy} = \sum_{i=1}^{n}\frac{x_i y_i}{\sigma_i^2} = \sum_{i=1}^{n}\frac{S_i S_{i-1}}{\sigma_i^2}.$$

Substituting these sums into the two coupled equations gives

$$\mu = \frac{S_y S_{xx} - S_x S_{xy}}{S(S_{xx} - S_{xy}) - (S_x^2 - S_x S_y)}$$

$$\lambda = -\frac{1}{\delta_t}\ln\left(\frac{S_{xy} - \mu S_x - \mu S_y + S\mu^2}{S_{xx} - 2\mu S_x + S\mu^2}\right),$$

The equation for the standard deviation is

$$\hat{\sigma}^2 = \left(S_{yy} - 2e^{-\lambda\delta_t}S_{xy} + e^{-2\lambda\delta_t}S_{xx} - 2\mu\left(1-e^{-\lambda\delta_t}\right)\left(S_y - e^{-\lambda\delta_t}S_x\right) + S\mu^2\left(1-e^{-\lambda\delta_t}\right)\right)\Big/S.$$

As discussed earlier, a least squares estimate of normally distributed errors will replicate the maximum likelihood estimate (MLE). This similarity was seen in the output of the function *MRpath* discussed previously. The log-likelihood form is quite flexible and will be used again later in this book in conjunction with the Kalman filer.

1.12 Summary

This chapter provided a derivation and application overview on the equations underlying geometric and arithmetic Brownian motion as well as the related mean reversion models. These models are readily applied to the pricing of financial derivatives and real options. A major vein of this book is the addition of jump processes or stochastic volatility to the drift-diffusion models of this chapter.

References

Baker, M.P., Mayfield, E.S., Parsons, J.E. (1998) Alternative Models of Uncertain Commodity Prices for Use with Modern Asset Pricing, Energy Journal **19**, 115

Bessembinder, H., Coughenour, J.F., Seguin, P.J., Smoller, M.M. (1995) Mean Reversion in Equilibrium Asset Prices: Evidence from the Futures Term Structure, Journal of Finance **50**, 361

Chance, D. (1994) The ABCs of Geometric Brownian Motion, Derivatives Quarterly **1**, 41

Chance, D. (2005) Mathematical Probability Theory and Finance: Connecting the Dots, Journal of Financial Education **31**, 1

Cox, J.C., Ingersoll, J.E., Ross, S.A. (1985) A Theory of the Term Structure of Interest Rates, Econometrica **53**, 385

Dias, M.A.G. (2004) Valuation of Exploration & Production Assets: An Overview of Real Options Models, Journal of Petroleum Science and Engineering **44**, 93

Dixit, A.K., Pindyck, R.S. (1994) Investment under Uncertainty, Princeton University Press

Hull, J. (2006) Options, Futures, and Other Derivatives, Prentice Hall

Laughton, D.G., Jacoby, H.D. (1995) The Effects of Reversion on Commodity Projects of Different Length, Real Options in Capital Investments: Models, Strategies, and Aplications, Trigeorgis, L. (ed.), Praeger Publisher185

Neftci, S.N. (2000) An Introduction to the Mathematics of Financial Derivatives, Academic Press Advanced Finance

Papoulis, A. (1984) Probability, Random Variables, and Stochastic Processes, McGraw-Hill

Pindyck, R. S., Rubinfeld, D. L. (1991) Econometric Models and Economic Forecasts, McGraw-Hill, Inc.

Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P. (1989) Numerical Recipes: The Art of Scientific Computing, Cambridge University Press

Rice, J. (1995) Mathematical Statistics and Data Analysis, Duxbury Press

Uhlenbeck, G.E., Ornstein, L.S. (1930) Theory of Brownian Motion Physical Review **36**

van den Berg, M.A. (2007) Calibrating the Ornstein-Uhlenbeck Model, White Paper, sitmo.com

Weisstein, E.W. Raw Moments, MathWorld, Wolfram Research, wolfram.com