Honors Theory of Probability
Date completed: Final, 2013-05-13

BLACK-SCHOLES-MERTON EQUATION AND RISK-NEUTRAL PRICING *

Yao XIAO   Xinqi ZOU   Yingxuan LI
NYU Shanghai

Abstract The Black-Scholes-Merton (BSM) model gives a theoretical estimate of the price of European-style options, taking into account the impact of time and other risk factors. We derive the Black-Scholes Merton equation and discuss two solutions [1] to it. One way is to directly solve the Black-Scholes-Merton partial differential equation. The other way is to approach by changing the probability measure and using risk-neutral pricing.

Key words Black-Scholes-Merton equation; Brownian motion; partial differential equation; change of measure; risk-neutral pricing.

1 Brownian Motion

1.1 Definition of Brownian Motion and the Filtration

Definition 1.1 Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. For each $\omega \in \Omega$, suppose there is a continuous function $W(t)$, $t \geq 0$ that satisfies $W(0) = 0$ and that depends on ω. Then $W(t)$, $t \geq 0$, is a *Brownian motion* if for all $0 = t_0 < t_1 < ... < t_m$ the increments

$$W(t_1) = W(t_1) - W(t_0),\ W(t_2) - W(t_1),\ \cdots,\ W(t_m) - W(t_{m-1}) \tag{1.1}$$

are independent and each of these increments is normally distributed with expected value 0 and variance $t_{i+1} - t_i$ for the ith increment.

Remark 1.2 One important property is that Brownian motion paths are nowhere differentiable. The intuition is that, since the increments in (1.1) are independent and normally distributed with expected value 0 and variance $t_{i+1} - t_i$ for the ith increment, the random variable

$$Z = \frac{W(t+h) - W(t)}{h}$$

has variance h^{-1}, which as $h \to 0$, becomes infinitely large.

In addition to the Brownian motion itself, we also need to denote the amount of information available at each time. We do it with filtration.

Definition 1.3 Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and define a Brownian motion $W(t)$, $t \geq 0$. A *filtration for the Brownian motion* is a collection of σ-algebras $\mathcal{F}(t)$, $t \geq 0$, satisfying:

(i) **(Information accumulates)** For $0 \leq s < t$, every set in $\mathcal{F}(s)$ is also in $\mathcal{F}(t)$.

(ii) **(Adaptivity)** For each $t \geq 0$, the Brownian motion $W(t)$ at time t is $\mathcal{F}(t)$-measurable.

*MATH-SHU 238 Honors Theory of Probability.

(iii) **(Independence of future increments)** For $0 \le t < u$, $W(u) - W(t)$ is independent of $\mathcal{F}(t)$.

Let $\Delta(t)$, $t \ge 0$ be a stochastic process. We say that $\Delta(t)$ is *adapted* to the filtration $\mathcal{F}(t)$ if for each $t \ge 0$ the random variable $\Delta(t)$ is $\mathcal{F}$-measurable.

Theorem 1.4 Brownian motion is a martingale.

Proof Let $0 \le s \le t$ be given. Then

$$\begin{aligned}\mathbb{E}(W(t)|\mathcal{F}(s)) &= \mathbb{E}((W(t) - W(s)) + W(s)|\mathcal{F}(s)) \\ &= \mathbb{E}(W(t) - W(s)|\mathcal{F}(s)) + \mathbb{E}(W(s)|\mathcal{F}(s)) = \mathbb{E}(W(t) - W(s)) + W(s) = W(s).\end{aligned}$$

□

1.2 Quadratic Variation

Definition 1.5 Let $f(t)$ be a function defined for $0 \le t \le T$. The *quadratic variation* of f up to time T is

$$[f, f](T) = \lim_{\|\Pi\| \to 0} \sum_{j=0}^{n-1} (f(t_{j+1}) - f(t_j))^2, \tag{1.2}$$

where $\Pi = \{t_0, t_1, \cdots, t_n\}$, $\|\Pi\| = \max_{0 \le k \le n-1} (t_{k+1} - t_k)$ and $0 = t_0 < t_1 < \cdots < t_n = T$.

Remark 1.6 Suppose the function f has a continuous derivative. Then

$$\sum_{j=0}^{n-1} (f(t_{j+1}) - f(t_j))^2 = \sum_{j=0}^{n-1} |f'(t_j^*)|^2 (t_{j+1} - t_j)^2 \le \|\Pi\| \cdot \sum_{j=0}^{n-1} |f'(t_j *)|^2 (t_{j+1} - t_j),$$

and thus

$$[f, f](T) \le \lim_{\|\Pi\| \to 0} \|\Pi\| \cdot \int_0^T |f'(t)|^2 dt = 0$$

as long as we assume that $\int_0^T |f'(t)|^2 dt < \infty$. However, Brownian motion is nowhere differentiable (Remark 1.2). In fact, it has bounded nonzero quadratic variation, as is stated in the following theorem.

Theorem 1.7 Let W be a Brownian motion. Then $[W, W](T) = T$ for $T \ge 0$ almost surely.

Proof Let

$$Y_{j+1} = \frac{W(t_{j+1}) - W(t_j)}{\sqrt{t_{j+1} - t_j}}.$$

We choose a large number n and take $t_j = jT/n$, $j = 0, 1, \cdots, n$. Then $t_{j+1} - t_j = T/n$ for all j and $(W(t_{j+1}) - W(t_j))^2 = TY_{j+1}^2/n$. The Law of Large Numbers then implies that

$$\sum_{j=0}^{n-1} \frac{Y_{j+1}^2}{n} \to \mathbb{E}(Y_{j+1}^2) \quad \text{as } n \to \infty,$$

where $Y_{j+1} \sim N(0, 1)$. Hence $\mathbb{E}(Y_{j+1}^2) = 1$, then we have

$$\sum_{j=0}^{n-1} (W(t_{j+1}) - W(t_j))^2 = \sum_{j=0}^{n-1} T \frac{Y_{j+1}^2}{n} \to T \quad \text{as } \|\Pi\| \to 0. \tag{1.3}$$

This completes the proof. □

Remark 1.8 We write informally

$$dW(t)dW(t) = dt \tag{1.4}$$

for the previous theorem, which indicates that Brownian motion accumulates quadratic variation at rate one per unit time. However, this does not mean $(W(t_{j+1})-W(t_j))^2 \approx t_{j+1}-t_j$. It is only when we sum up both sides and apply the Law of Large Numbers to cancel errors that we get the correct statement.

Remark 1.9 Let $\Pi = \{t_0, t_1, \cdots, t_n\}$ be a partition of $[0,T]$. Since

$$\lim_{\|\Pi\|\to 0}\left|\sum_{j=0}^{n-1}(W(t_{j+1})-W(t_j))(t_{j+1}-t_j)\right| \le \lim_{\|\Pi\|\to 0}\sum_{j=0}^{n-1}\max_{0\le k\le n-1}|W(t_{k+1})-W(t_k)|(t_{j+1}-t_j)$$

$$= \lim_{\|\Pi\|\to 0}\max_{0\le k\le n-1}|W(t_{k+1})-W(t_k)|\cdot T = 0,$$

$$\lim_{\|\Pi\|\to 0}\sum_{j=0}^{n-1}(t_{j+1}-t_j)^2 \le \lim_{\|\Pi\|\to 0}\sum_{j=0}^{n-1}\|\Pi\|(t_{j+1}-t_j) = \lim_{\|\Pi\|\to 0}\|\Pi\|\cdot T = 0,$$

we can conclude that

$$\lim_{\|\Pi\|\to 0}\sum_{j=0}^{n-1}(t_{j+1}-t_j)^2 = 0, \quad \lim_{\|\Pi\|\to 0}\sum_{j=0}^{n-1}(W(t_{j+1})-W(t_j))(t_{j+1}-t_j) = 0. \tag{1.5}$$

Then, just as we capture (1.3) by writing (1.4), we can capture (1.5) by writing informally

$$dW(t)dt = 0, \quad dtdt = 0. \tag{1.6}$$

Such informal notations as in (1.4) and (1.6) in differential form will be rapidly reused in the following sections to simplify computations and proofs.

2 Stochastic Calculus

If $g(t)$ is a differentiable function, then we can define

$$\int_0^T \Delta(t)dg(t) = \int_0^T \Delta(t)g'(t)dt,$$

where the right-hand side is an ordinary (Lebesgue) integral with respect to time. However, this will not work for an integral with respect to a Brownian motion, which is called an Itô's integral, since Brownian motion paths are nowhere differentiable. Our approach to make sense of an Itô's integral is that we first define the Itô's integral for simple processes and then define that for a general integrand by taking the limit of a sequence of simple processes.

2.1 Itô's Integral

Definition 2.1 Let $\Pi = \{t_0, t_1, \cdots, t_n\}$ be a partition of $[0,T]$; i.e., $0 = t_0 < t_1 < \cdots < t_n = T$. Suppose there is a sequence of *simple processes* $\{\Delta_n(t)\}$ such that it is constant in t on each subinterval $[t_j, t_{j+1})$, and that it converges to the continuously varying $\Delta(t)$ in $\mathcal{L}^2$ sense. Let $W(t)$ be a Brownian motion. The *Itô's integral of* each *simple process* $\Delta_n(t)$ with respect to $W(t)$ is defined by, for each $t_j \le t \le t_{k+1}$,

$$I_n(t) = \int_0^t \Delta_n(u)dW(u) = \sum_{j=0}^{n-1}\Delta_n(t_j)(W(t_{j+1})-W(t_j)) + \Delta_n(t_k)(W(t)-W(t_k)), \tag{2.1}$$

and the *Itô's integral* with respect to $W(t)$ is defined by

$$I(t) = \int_0^t \Delta(t)dW(t) = \lim_{n\to\infty}\int_0^t \Delta_n(u)dW(u).$$

Theorem 2.2 Let $W(t)$ be a Brownian motion. Let $T > 0$, and let $\Delta(t)$, $0 \le t \le T$, be an adapted stochastic process. Then, the Itô's integral $I(t) = \int_0^t \Delta(u)dW(u)$ satisfies:

(i) **(Continuity)** As a function of the upper limit of integration t, the paths of $I(t)$ are continuous.

(ii) **(Adaptivity)** For each t, $I(t)$ is $\mathcal{F}(t)$-measurable.

(iii) **(Linearity)** Suppose $I(t) = \int_0^t \Delta(u)dW(u)$, $J(t) = \int_0^t \Gamma(u)dW(u)$, and c is an arbitrary constant. Then $I(t) \pm J(t) = \int_0^t (\Delta(u) \pm \Gamma(u))dW(u)$ and $cI(t) = \int_0^t c\Delta(u)dW(u)$.

(iv) **(Martingale)** $I(t)$ is a martingale.

(v) **(Itô isometry)** $\mathbb{E}(I^2(t)) = \mathbb{E}\left(\int_0^t \Delta^2(u)du\right)$.

(vi) **(Quadratic variation)** $[I, I](t) = \int_0^t \Delta^2(u)du$.

Proof It suffices to prove the properties only for the Itô's integral of a simple process, since these properties can be inherited after simply taking limits in $\mathcal{L}^2$ sense. Properties (i)–(iii) are trivial, and we only prove properties (iv)–(vi) explicity.

(Martingale) Let $0 \le s \le t \le T$ be given. Assume that there exists partition points $t_l < t_k$ such that $s \in [t_l, t_{l+1})$, $t \in [t_k, t_{k+1})$. Let $D_j = W(t_{j+1}) - W(t_j)$ for $0 \le j \le k-1$ and $D_k = W(t) - W(t_k)$. Thus we split the sum as

$$I(t) = \sum_{j=0}^{l-1} \Delta(t_j)D_j + \Delta(t_l)D_l + \sum_{j=l+1}^{k-1} \Delta(t_j)D_j + \Delta(t_k)D_k.$$

We investigate the four summands respectively. For the first summand, since $t_l \le s$, for $0 \le j \le l-1$, $\Delta(t_j)D_j$ is $\mathcal{F}(s)$-measurable. Hence, $\mathbb{E}(\Delta(t_j)D_j|\mathcal{F}(s)) = \Delta(t_j)D_j$. For the second summand,

$$\mathbb{E}(\Delta(t_l)D_l|\mathcal{F}(s)) = \Delta(t_l)(\mathbb{E}(W(t_{l+1})|\mathcal{F}(s)) - \mathbb{E}(W(t_l)|\mathcal{F}(s))) = \Delta(t_l)(W(s) - W(t_l)).$$

The third and the fourth summands are similar. For $l+1 \le j \le k$, by the "towering property,"

$$\begin{aligned}&\mathbb{E}(\Delta(t_j)D_j|\mathcal{F}(s)) = \mathbb{E}(\mathbb{E}(\Delta(t_j)D_j|\mathcal{F}(t_j))|\mathcal{F}(s))\\ &\quad = \mathbb{E}(\Delta(t_j)(\mathbb{E}(W(t_{j+1})|\mathcal{F}(t_j)) - \mathbb{E}(W(t_j)|\mathcal{F}(t_j)))|\mathcal{F}(s)) = \mathbb{E}(\Delta(t_j)(W(t_j) - W(t_j))|\mathcal{F}(s)) = 0.\end{aligned}$$

Hence, combining the arguments, we can conclude that $I(t)$ is a martingale because

$$\mathbb{E}(I(t)|\mathcal{F}(s)) = \sum_{j=0}^{l-1} \Delta(t_j)D_j + \Delta(t_l)(W(s) - W(t_l)) = I(s).$$

(Itô isometry) By multinomial expansion, we obtain

$$I^2(t) = \sum_{j=0}^{k} \Delta^2(t_j)D_j^2 + \sum_{0 \le i < j \le k} \Delta(t_i)\Delta(t_j)D_iD_j.$$

Since D_j^2 is independent of $\mathcal{F}(t_j)$ while $\Delta^2(t_j)$ is $\mathcal{F}(t_j)$-measurable,

$$\mathbb{E}\left(\sum_{j=0}^{k} \Delta^2(t_j)D_j^2\right) = \sum_{j=0}^{k-1} \mathbb{E}(\Delta^2(t_j))(t_{j+1} - t_j) + \mathbb{E}(\Delta^2(t_k))(t - t_k) = \int_0^t \mathbb{E}(\Delta^2(u))du.$$

The last equation holds because $\Delta^2(t_j)$ is constant on each subinterval, so the Riemann sum is exactly the corresponding Riemann integral. Also, since D_j is independent of $\mathcal{F}(t_j)$ while $\Delta(t_i)\Delta(t_j)D_i$ is $\mathcal{F}(t_j)$-measurable, $\mathbb{E}(\Delta(t_i)\Delta(t_j)D_iD_j) = \mathbb{E}(\Delta(t_i)\Delta(t_j)D_i)\mathbb{E}(\Delta(t_j)) = 0$. Hence, we can conclude that

$$\mathbb{E}(I^2(t)) = \int_0^t \mathbb{E}(\Delta^2(u))du + 0 = \mathbb{E}\left(\int_0^t \Delta^2(u)du\right).$$

(Quadratic variation) Choose $t_j = s_0 < s_1 < \cdots < s_m = t_{j+1}$ as partition points and consider

$$\sum_{i=0}^{m-1}(I(s_{i+1}) - I(s_i))^2 = \Delta^2(t_j)\sum_{i=0}^{m-1}(W(s_{i+1}) - W(s_i))^2.$$

Since Brownian motion accumulates quadratic variation at rate one per unit time, then on each subinterval $[t_j, t_{j+1}]$, the Itô integral accumulates quadratic variation at rate $\Delta^2(t_j)$ per unit time. Bringing $m \to \infty$, we obtain

$$\lim_{m\to\infty}\sum_{i=0}^{m-1}(I(s_{i+1}) - I(s_i))^2 = \int_{t_j}^{t_{j+1}} \Delta^2(u)du.$$

Adding up all the pieces, we obtain the quadratic variation of an Itô's integral

$$[I, I](t) = \int_0^t \Delta^2(u)du. \qquad \square$$

Now, we want to further extend the Itô's integral with respect to more general stochastic processes than Brownian motions. The stochastics processes we are interested in is called the Itô processes.

Definition 2.3 Let $W(t)$, $t \geq 0$, be a Brownian motion, and let $\mathcal{F}(t)$, $t \geq 0$, be an associated filtration. An *Itô process* is a stochastic process of the form

$$X(t) = X(0) + \int_0^t \Delta(u)dW(u) + \int_0^t \Theta(u)du, \tag{2.2}$$

where $X(0)$ is nonrandom and $\Delta(u)$ and $\Theta(u)$ are adapted stochastics processes. We assume finiteness conditions for both integrals.

Definition 2.4 Let $X(t)$, $t \geq 0$, be an Itô process as described in Definition (2.2) and let $\Gamma(t)$, $t \geq 0$, be an adapted process. We define the *Itô's integral with respect to an Itô process* as

$$\int_0^t \Gamma(u)dX(u) = \int_0^t \Gamma(u)\Delta(u)dW(u) + \int_0^t \Gamma(u)\Theta(u)du.$$

2.2 Itô-Doeblin's Formula

Theorem 2.5 (Itô-Doeblin's Formula for an Itô process) Let $X(t)$, $t \geq 0$, be an Itô process as defined in (2.2), and let $f(t, x)$ be a function for which the partial derivatives $f_t(t, x)$, $f_x(t, x)$, and $f_{xx}(t, x)$ are defined and continuous. Then, for every $T \geq 0$,

$$f(T, X(T)) = f(0, X(0)) + \int_0^T f_t(t, X(t))dt + \int_0^T f_x(t, X(t))\Delta(t)dW(t) + \frac{1}{2}\int_0^T f_{xx}(t, X(t))\Delta^2(t)dt.$$

Proof Fix $T > 0$, and let $\Pi = \{t_0, t_1, \cdots, t_n\}$ be a partition of $[0, T]$. Let $\|\Pi\|$ denote the length of the longest subinterval. Then, bring $\|\Pi\| \to 0$ and apply Taylor's formula, we obtain

$$\begin{aligned}
&f(T, X(T)) - f(0, X(0)) \\
&= \lim_{\|\Pi\|\to 0}\sum_{j=0}^{n-1} f_t(t_j, X(t_j))(t_{j+1} - t_j) + \lim_{\|\Pi\|\to 0}\sum_{j=0}^{n-1} f_x(t_j, X(t_j))(X(t_{j+1}) - X(t_j)) \\
&\quad + \lim_{\|\Pi\|\to 0}\sum_{j=0}^{n-1}\frac{1}{2}f_{tt}(t_j, X(t_j))(t_{j+1} - t_j)^2 + \lim_{\|\Pi\|\to 0}\sum_{j=0}^{n-1}\frac{1}{2}f_{xx}(t_j, X(t_j))(X(t_{j+1}) - X(t_j))^2 \\
&\quad + \lim_{\|\Pi\|\to 0}\sum_{j=0}^{n-1} f_{tx}(t_j, X(t_j))(t_{j+1} - t_j)(X(t_{j+1}) - X(t_j)) + \text{higher order terms.}
\end{aligned}$$

We informally write in differential form, and denote f for $f(t, X(t))$, f_x for $f_x(t, X(t))$, etc. By (1.4) and (1.6), recall that $dW(t)dW(t) = dt$, $dtdt = 0$, and $dtdW(t) = 0$, and we obtain

$$\begin{aligned} df &= f_t dt + f_x dX(t) + \frac{1}{2} f_{tt} dtdt + \frac{1}{2} f_{xx} dX(t)dX(t) + f_{tx} dtdX(t) + \text{higher order terms} \\ &= f_t dt + f_x(\Delta(t)dW(t) + \Theta(t)dt) + \frac{1}{2} f_{xx}(\Delta(t)dW(t) + \Theta(t)dt)^2 + f_{tx} dt(\Delta(t)dW(t) + \Theta(t)dt) \\ &= f_t dt + f_x \Delta(t)dW(t) + f_x \Theta(t)dt + \frac{1}{2} f_{xx} \Delta^2(t)dt. \end{aligned} \tag{2.3}$$

which is exactly our desired result in differential form. □

3 Black-Scholes-Merton Equation

The key idea behind the Black-Scholes-Merton model is to hedge the option by buying and selling the underlying asset in the right way to eliminate risk. We are going to derive the Black-Scholes-Merton equation for the price of an option on an asset modeled as a geometric Brownian motion.

Definition 3.1 Let $W(t)$, $t \geq 0$ be a Brownian motion. Let α and $\sigma > 0$ be constants. We define the *geometric Brownian motion* $S(t)$, $t \geq 0$, as

$$S(t) = S(0) \exp\left(\sigma W(t) + \left(\alpha - \frac{1}{2}\sigma^2\right) t\right), \tag{3.1}$$

where $S(0)$ is nonrandom and positive.

3.1 Computations on the Portfolio and the Option Values

Consider an agent who at each time t has a portfolio valued at $X(t)$, which invests in:

- a money account paying a constant rate of interest r; and
- a stock modeled by the geometric Brownian motion $S(t)$ as defined in (3.1). Rewrite as

$$S(t) = S(0) \exp\left(\int_0^t \sigma dW(t) + \int_0^t \left(\alpha - \frac{1}{2}\sigma^2\right) dt\right).$$

Let $S(t) = f(X(t))$ with $f(x) = S(0)e^x$. Apply Itô-Doeblin formula in differential form (2.3), then

$$\begin{aligned} dS(t) &= S(0)e^{X(t)} \sigma dW(t) + S(0)e^{X(t)} \left(\alpha - \frac{1}{2}\sigma^2\right) dt + \frac{1}{2} S(0)e^{X(t)} \sigma^2 dt \\ &= \sigma S(t)dW(t) + \alpha S(t)dt. \end{aligned} \tag{3.2}$$

Now suppose at each time t the investor holds $\Delta(t)$ shares of stock, which is random but adapted to the filtration associated with the Brownian motion $W(t)$, $t \geq 0$. We can compute the following:

- the capital gain on the stock position: $\Delta(t)dS(t)$;
- the interest earnings on the cash position: $r(X(t) - \Delta(t)dS(t))dt$; and thus
- the portfolio value:

$$dX(t) = \Delta(t)dS(t) + r(X(t) - \Delta(t)dS(t))dt = rX(t)dt + \Delta(t)(\alpha - r)S(t)dt + \Delta(t)\sigma S(t)dW(t).$$

To determine the present value of a stream of payments that is to be received in the future, we shall often consider the discounted stock price $e^{-rt}S(t)$ and the discounted portfolio value $e^{-rt}X(t)$. Apply the Itô-Doeblin formula with $f(t, x) = e^{-rt}x$, we obtain

$$d(e^{-rt}S(t)) = -re^{-rt}S(t)dt + e^{-rt}dS(t) = (\alpha - r)e^{-rt}S(t)dt + \sigma e^{-rt}S(t)dW(t), \tag{3.3}$$

$$d(e^{-rt}X(t)) = -re^{-rt}X(t)dt + e^{-rt}dX(t) = \Delta(t)(\alpha - r)e^{-rt}S(t)dt + \Delta(t)\sigma e^{-rt}S(t)dW(t)$$

$$= \Delta(t) d(e^{-rt} S(t)). \tag{3.4}$$

We also consider a European call option that pays $(S(T) - K)^+$ at time T. The strike price $K \geq 0$. Let $c(t, x)$ denote the value of the call at time t if the stock price at that time is $S(t) = x$. Our goal is to find a profitable $c(t, x)$. By the Itô-Doeblin formula, we obtain

$$\begin{aligned} dc(t, S(t)) &= c_t(t, S(t))dt + c_x(t, S(t))dS(t) + \frac{1}{2} c_{xx}(t, S(t))dS(t)dS(t) \\ &= c_t(t, S(t))dt + \alpha S(t) c_x(t, S(t))dt + \sigma S(t) c_x(t, S(t))dW(t) + \frac{1}{2}\sigma^2 S^2(t) c_{xx}(t, S(t))dt \\ &= \Big(c_t(t, S(t)) + \alpha S(t) c_x(t, S(t)) + \frac{1}{2}\sigma^2 S^2(t) c_{xx}(t, S(t)) \Big) dt + \sigma S(t) c_x(t, S(t)) dW(t). \end{aligned}$$

Then, by the Itô-Doeblin formula with $f(t, x) = e^{-rt} x$, we obtain the discounted option price

$$\begin{aligned} d(e^{-rt} c(t, S(t))) &= -re^{-rt} c(t, S(t))dt + e^{-rt} dc(t, S(t)) \\ &= -re^{-rt} c(t, S(t))dt + e^{-rt} c_t(t, S(t))dt + \alpha e^{-rt} S(t) c_x(t, S(t))dt \\ &\quad + \sigma e^{-rt} S(t) c_x(t, S(t))dW(t) + \frac{1}{2}\sigma^2 e^{-rt} S^2(t) c_{xx}(t, S(t))dt \\ &= e^{-rt}(-rc(t, S(t)) + c_t(t, S(t)) + \alpha S(t) c_x(t, S(t)) + \frac{1}{2}\sigma^2 S^2(t) c_{xx}(t, S(t)))dt \\ &\quad + e^{-rt} \sigma S(t) c_x(t, S(t)) dW(t). \end{aligned} \tag{3.5}$$

3.2 Black-Scholes-Merton Equation

A (short option) hedging portfolio starts with some initial capital $X(0)$ and invests in the stock and money market account so that the portfolio value $X(t)$ at each time $t \in [0, T]$ agrees with the value of call $c(t, S(t))$. Hence, $d(e^{-rt} X(t)) = d(e^{-rt} c(t, S(t)))$ suffices since the initial values $X(0) = c(0, S(0))$. Thus by (3.4) and (3.5), we require

$$\begin{aligned} &\Delta(t)(\alpha - r)S(t)dt + \Delta(t)\sigma S(t)dW(t) \\ &= \Big(-rc(t, S(t)) + c_t(t, S(t)) + \alpha S(t) c_x(t, S(t)) + \frac{1}{2}\sigma^2 S^2(t) c_{xx}(t, S(t)) \Big) dt + \sigma S(t) c_x(t, S(t)) dW(t) \\ &\Longrightarrow \begin{cases} \Delta(t)\sigma S(t) = \sigma S(t) c_x(t, S(t)), & \text{(A)} \\ \Delta(t)(\alpha - r)S(t) = -rc(t, S(t)) + c_t(t, S(t)) + \alpha S(t) c_x(t, S(t)) + \frac{1}{2}\sigma^2 S^2(t) c_{xx}(t, S(t)). & \text{(B)} \end{cases} \end{aligned}$$

From (A) we have $\Delta(t) = c_x(t, S(t))$, and by substituting into (B), we obtain the *Black-Scholes-Merton partial differential equation*:

$$c_t(t, x) + rx c_x(t, x) + \frac{1}{2}\sigma^2 x^2 c_{xx}(t, x) = rc(t, x), \quad t \in [0, T),\ x \in [0, \infty), \tag{3.6}$$

with terminal condition and boundary conditions

$$\text{TC:} \quad c(T, x) = (x - K)^+,\ x \in [0, \infty), \tag{3.7}$$

$$\text{BC:} \quad c(t, 0) = 0,\ \lim_{x \to \infty} \Big(c(t, x) - \Big(x - e^{-r(T-t)} K \Big) \Big) = 0,\ t \in [0, T). \tag{3.8}$$

The corresponding solution [2] to Black-Scholes-Merton equation (3.6) with terminal condition (3.7) and boundary conditions (3.8) is

$$c(t, x) = x\mathcal{N}(d_+(T - t, x)) - Ke^{-r(T-t)}\mathcal{N}(d_-(T - t, x)), \quad t \in [0, T),\ x \in [0, \infty), \tag{3.9}$$

where $\mathcal{N}$ is the cumulative standard normal distribution, and

$$d_{\pm}(\tau, x) = \frac{1}{\sigma\sqrt{\tau}}\left(\log\frac{x}{K} + \left(r \pm \frac{\sigma^2}{2}\right)r\right).$$

4 Risk-Neutral Pricing

In this section, we will use the risk-neutral probability measure to obtain the solution (3.9) to Black-Scholes-Merton Equation. The idea is to find a new probability measure $\tilde{\mathbb{P}}$ such that $V(T)$ is a martingale under $\tilde{\mathbb{P}}$, and $V(t) = \tilde{\mathbb{E}}(V(T)|\mathcal{F}(s))$ for all $0 \leq t \leq T$.

4.1 Change of Probability Measure

Definition 4.1 Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let $\tilde{\mathbb{P}}$ be another probability measure on $(\Omega, \mathcal{F})$ that is equivalent to $\mathbb{P}$, and let Z be an almost surely positive random variable with $\mathbb{E}(Z) = 1$. Suppose

$$\tilde{\mathbb{P}}(A) = \int_A Z(\omega)d\mathbb{P}(\omega) \quad \text{for all } A \in \mathcal{F}, \quad \tilde{\mathbb{E}}(X) = \mathbb{E}(XZ). \tag{4.1}$$

Then Z is called the *Radon-Nikodým derivative* of $\tilde{\mathbb{P}}$ with respect to $\mathbb{P}$, and we write $Z = d\tilde{\mathbb{P}}/d\mathbb{P}$.

Given the filtration $\mathcal{F}(t)$, $0 \leq t \leq T$, we can further define the *Radon-Nikodým derivative process* $Z(t) = \mathbb{E}(Z|\mathcal{F}(t))$, $0 \leq t \leq T$. Notice that the Radon-Nikodým derivative process is a martingale because by the "towering property," $\mathbb{E}(Z(t)|\mathcal{F}(s)) = \mathbb{E}(\mathbb{E}(Z|\mathcal{F}(t))|\mathcal{F}(s)) = \mathbb{E}(Z|\mathcal{F}(s)) = Z(s)$.

4.2 Girsanov's Theorem

In order to recognize a Brownian motion, we state the following theorem without an explicit proof.

Theorem 4.2 (Lévy, one dimension) Let $M(t)$, $t \geq 0$, be a martingale relative to a filtration $\mathcal{F}(t)$, $t \geq 0$. Assume that $M(0) = 0$, $M(t)$ has continuous paths, and $[M, M](t) = t$ for all $t \geq 0$. Then $M(t)$ is a Brownian motion.

Now we investigate the change of a Brownian motion under the change of probability measure.

Theorem 4.3 (Girsanov, one dimension) Let $W(t)$, $0 \leq t \leq T$, be a Brownian motion on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and let $\mathcal{F}(t)$, $0 \leq t \leq T$, be a filtration for this Brownian motion. Let $\Theta(t)$, $0 \leq t \leq T$, be an adapted process. Define

$$Z(t) = \exp\left(-\int_0^t \Theta(u)dW(u) - \frac{1}{2}\int_0^t \Theta^2(u)du\right), \quad \tilde{W}(t) = W(t) + \int_0^t \Theta(u)du,$$

and assume that $\mathbb{E}\left(\int_0^T \Theta^2(u)Z^2(u)du\right) < \infty$. Set $Z = Z(T)$. Then $\mathbb{E}(Z) = 1$ and under the probability measure $\tilde{\mathbb{P}}$ given by (4.1), the process $\tilde{W}(t)$, $0 \leq t \leq T$, is a Brownian motion.

Proof We use Lévy's Theorem (Theorem 4.2). We have

$$\begin{aligned}
&\tilde{W}(0) = W(0) + \int_0^0 \Theta(u)du = 0, \\
&d\tilde{W}(t)d\tilde{W}(t) = (dW(t) + \Theta(t)dt)^2 = dW(t)dW(t) + 2\Theta(t)dW(t)dt + \Theta^2(t)dtdt = dt.
\end{aligned}$$

Thus it remains to prove that $\tilde{W}(t)$ is a martingale under $\tilde{\mathbb{P}}$ to apply Lévy's Theorem. We first observe that $Z(t)$ is a martingale under $\mathbb{P}$. Apply the Itô-Doeblin formula with $f(x) = e^x$ and with

$$X(t) = -\int_0^t \Theta(u)dW(u) - \frac{1}{2}\int_0^t \Theta^2(u)du,$$

we obtain

$$dZ(t) = df(X(t)) = e^{X(t)}(-\Theta(t)dW(t) - \frac{1}{2}\Theta^2(t)dt) + \frac{1}{2}e^{X(t)}\Theta^2(t)dt = -\Theta(t)Z(t)dW(t).$$

Hence, $Z(t)$ is an Itô integral by definition, and thus a martingale. In particular, $\mathbb{E}(Z) = \mathbb{E}(Z(T)) = Z(0) = 1$. Since $Z = Z(T)$, we obtain $Z(t) = \mathbb{E}(Z(T)|\mathcal{F}(t)) = \mathbb{E}(Z|\mathcal{F}(t))$, $0 \leq t \leq T$. This shows that Z(t), $0 \leq t \leq T$ is a Radon-Nikodym derivative process. Next we prove that $\tilde{W}(t)Z(t)$ is a martingale under $\mathbb{P}$. To see this, we first state the Itô's product rule.

Corollary 4.4 (Itô's product rule) Let $X(t)$ and $Y(t)$ be Itô processes. Then

$$d(X(t)Y(t)) = X(t)dY(t) + Y(t)dX(t) + dX(t)dY(t).$$

This corollary can be verified by applying the Itô-Doeblin formula with $f(x, y) = xy$. Applying this Itô's product rule, we obtain

$$\begin{aligned} d(\tilde{W}(t)Z(t)) &= \tilde{W}(t)dZ(t) + Z(t)d\tilde{W}(t) + d\tilde{W}(t)dZ(t) \\ &= -\tilde{W}(t)\Theta(t)Z(t)dW(t) + Z(t)dW(t) + Z(t)\Theta(t)dt + (dW(t) + \Theta(t)dt)(-\Theta(t)Z(t)dW(t)) \\ &= (-\tilde{W}(t)\Theta(t) + 1)Z(t)dW(t). \end{aligned}$$

Since the previous expression has no dt term, the process $\tilde{W}(t)Z(t)$ is a martingale under $\mathbb{P}$. Now let $0 \leq s \leq t \leq T$ be given. By the partial-averaging property of conditional expectations, we obtain

$$\tilde{\mathbb{E}}(\tilde{W}(t)|\mathcal{F}(s)) = \frac{1}{Z(s)}\mathbb{E}(\tilde{W}(t)Z(t)|\mathcal{F}(s)) = \frac{1}{Z(s)}Z(s)\mathbb{E}(\tilde{W}(t)|\mathcal{F}(s)) = \tilde{W}(s).$$

This shows that $\tilde{W}(t)$ is a martingale under $\tilde{\mathbb{P}}$. The proof is complete. □

4.3 Stock and Pricing Under the Risk-Neutral Measure

Let $W(t)$, $0 \leq t \leq T$, be a Brownian motion on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and let $\mathcal{F}(t)$, $0 \leq t \leq T$, be a filtration for this Brownian motion. Here T is a fixed final time, $S(t)$ is the stock price, $D(t)$ is the discounted process, $\alpha(t)$ is the mean rate of return, $\sigma(t)$ is the volatility, and $R(t)$ is the interest rate. Then we obtain the discounted stock process

$$d(D(t)S(t)) = (\alpha(t) - R(t))D(t)S(t)dt + \sigma(t)D(t)S(t)dW(t) = \sigma(t)D(t)S(t)(\Theta(t)dt + dW(t)). \quad (4.2)$$

Define $d\tilde{W}(t) = \dfrac{\alpha(t) - R(t)}{\sigma(t)}dt + dW(t)$, so (4.2) reduces to $d(D(t)S(t)) = \sigma(t)D(t)S(t)d\tilde{W}(t)$. By Girsanov's Theorem (Theorem 4.3), $\tilde{W}(t)$ is a Brownian motion under measure $\tilde{\mathbb{P}}$, and thus a martingale under measure $\tilde{\mathbb{P}}$. Thus, in a generalized form of Definition 3.1, the stock modeled under the risk-neutral measure is the generalized geometric Brownian motion

$$S(t) = S(0)\exp\left(\int_0^t \sigma(s)d\tilde{W}(s) + \int_0^t \left(R(s) - \frac{1}{2}\sigma^2(s)\right)ds\right). \quad (4.3)$$

Now let $X(t)$ and $V(t)$ represent portfolio value and the payoff respectively at time t. Suppose the investor hold $\Delta(t)$ shares of stock at time t, then similar to (3.4), we shall obtain

$$d(D(t)X(t)) = \Delta(t)d(D(t)S(t)) = \Delta(t)\sigma(t)D(t)S(t)d\tilde{W}(t).$$

We shall require $X(t) = V(t)$ almost surely for all $0 \leq t \leq T$ in order to hedge a short position. Also use the fact that $D(t)X(t)$ is a martingale under $\tilde{\mathbb{P}}$, we obtain the general form of the *risk-neutral pricing formula for the continuous-time model* as

$$D(t)V(t) = D(t)X(t) = \tilde{\mathbb{E}}(D(T)X(T)|\mathcal{F}(t)) = \tilde{\mathbb{E}}(D(T)V(T)|\mathcal{F}(t)), \quad 0 \leq t \leq T. \quad (4.4)$$

4.4 Deriving the Black-Scholes-Merton formula

To obtain the Black-Scholes-Merton price of a European call, we assume constant volatility σ and constant interest rate r. Set the payoff to be $V(T) = (S(T) - K)^+$ and $D(t) = e^{-rt}$. Then from (4.3), we shall obtain

$$S(T) = S(t)\exp\left(\sigma(\tilde{W}(T) - \tilde{W}(t)) + \left(r - \frac{1}{2}\sigma^2\right)\tau\right) = S(t)\exp\left(-\sigma\sqrt{\tau}Y + \left(r - \frac{1}{2}\sigma^2\right)\tau\right),$$

where Y is the standard normal random variable

$$Y = -\frac{\tilde{W}(T) - \tilde{W}(t)}{\sqrt{T-t}}, \quad \tau \text{ is the "time to expiration" } T - t.$$

Let $c(t, x)$ denote the value of the call at time t if the stock price at that time is $S(t) = x$. Therefore, (4.4) holds with

$$\begin{aligned} c(t,x) &= \tilde{\mathbb{E}}\left(e^{-r\tau}\left(x\exp\left(-\sigma\sqrt{\tau}Y + \left(r - \frac{1}{2}\sigma^2\right)\tau\right) - K\right)^+\right) \\ &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-r\tau}\left(x\exp\left(-\sigma\sqrt{\tau}Y + \left(r - \frac{1}{2}\sigma^2\right)\tau\right) - K\right)^+ e^{-\frac{1}{2}y^2}dy, \end{aligned} \tag{4.5}$$

where the integrand is positive if and only if

$$y < d_-(\tau, x) = \frac{1}{\sigma\sqrt{\tau}}\left(\log\frac{x}{K} + \left(r - \frac{1}{2}\sigma^2\right)\tau\right)$$

Hence, (4.5) becomes

$$\begin{aligned} c(t,x) &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{d_-(\tau,x)} x\exp\left(-\frac{y^2}{2} - \sigma\sqrt{\tau}Y - \frac{\sigma^2\tau}{2}\right)dy - \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{d_-(\tau,x)} K\exp\left(-\frac{y^2}{2} - r\tau\right)dy \\ &= \frac{x}{\sqrt{2\pi}}\int_{-\infty}^{d_-(\tau,x)} \exp\left(-\frac{1}{2}(y + \sigma\sqrt{\tau})^2\right)dy - e^{-r\tau}K\mathcal{N}(d_-(\tau,x)) \\ &= \frac{x}{\sqrt{2\pi}}\int_{-\infty}^{d_-(\tau,x)+\sigma\sqrt{\tau}} \exp\left(-\frac{z^2}{2}\right)dz - e^{-r\tau}K\mathcal{N}(d_-(\tau,x)) \\ &= x\mathcal{N}(d_+(\tau,x)) - e^{-r\tau}K\mathcal{N}(d_-(\tau,x)), \end{aligned}$$

where $\mathcal{N}$ is the cumulative standard normal distribution, and

$$d_+(\tau,x) = d_-(\tau,x) + \sigma\sqrt{\tau} = \frac{1}{\sigma\sqrt{\tau}}\left(\log\frac{x}{K} + \left(r + \frac{1}{2}\sigma^2\right)\tau\right).$$

This is exactly the same result as we obtained in (3.9).

Acknowledgements We are deeply indebted to Professor Wei Wu for his contributions during the semester. Gratitudes are also extended to Professor Siran Li, Chenlin Gu, and so on, from whom we received precious remarks on our final presentation.

References

[1] Shreve S. E. (2004). Stochastic Calculus for Finance II: Continuous-Time Models. New York: Springer, **11**: 83–234.
[2] Law E. (2019). Finding the Solution to the Black-Scholes-Equation. 3–9.