NOTES ON STOCHASTIC CALCULUS AND THE BLACK–SCHOLES–MERTON EQUATION

IORDAN GANEV

Contents

New appendix!

1. Introduction

There are informal notes on stochastic calculus with the aim of understanding the Black–Scholes–Merton differential equation. The main references are [Shr04, Rom12]. We claim no originality. Some basic background in mathematics is assumed, and we start with a brief summary of the necessary concepts from probability theory.

2. Probability basics

2.1. **Probability spaces.** Let Ω be a non-empty set. A *σ-algebra* on Ω is a collection of subsets $\mathcal{F}$ that contains $\emptyset$ and is closed under complements and countable unions. The pair $(\Omega, \mathcal{F})$ is known as a *measurable space*. For example, the Borel σ-algebra $\mathcal{B}$ on $\mathbb{R}$ is defined as the smallest σ-algebra containing the closed intervals $[a, b]$. We can similarly define the Borel σ-algebra $\mathcal{B}([0, 1])$ of the unit interval, or of any interval in $\mathbb{R}$.

A *measure* on $(\Omega, \mathcal{F})$ is a function $\mu : \mathcal{F} \to \mathbb{R}_{\geq 0}$ that is additive on countable collections of disjoint subsets belonging to $\mathcal{F}$. A measure μ is said to be a probability measure if $\mu(\Omega) = 1$. We may write $\mu = \mathbb{P}$ for a probability measure. Let $\mathsf{M}(\Omega, \mathcal{F})$ (resp. $\mathsf{PM}(\Omega, \mathcal{F})$) be the set of all measures (resp. probability measures) on $(\Omega, \mathcal{F})$. A *probability space* is a triple $(\Omega, \mathcal{F}, \mathbb{P})$ where $\mathbb{P}$ is a probability measure on $(\Omega, \mathcal{F})$.

2.2. **Random variables.** A *random variable* on a measurable space $(\Omega, \mathcal{F})$ is a function $X : \Omega \to \mathbb{R}$ such that the inverse image of any Borel subset of $\mathbb{R}$ belongs to $\mathcal{F}$, that is:

$$X^{-1}(B) \in \mathcal{F} \qquad \text{for all } B \in \mathcal{B}.$$

The set of all random variables on $(\Omega, \mathcal{F})$ forms a vector space (see Exercise 6.1), which we denote $\mathsf{RV}(\Omega, \mathcal{F})$ and may abbreviate to RV when $(\Omega, \mathcal{F})$ is clear from context. Any random variable defines a pullback map:

$$\begin{aligned} \mathsf{PM}(\Omega, \mathcal{F}) &\to \mathsf{PM}(\mathbb{R}, \mathcal{B}) \\ \mathbb{P} \mapsto \mu_X^{(\mathbb{P})} &= \left[B \mapsto \mathbb{P}(X^{-1}(B)) \right] \qquad \text{for any } B \in \mathcal{B}. \end{aligned}$$

The probability measure $\mu_X^{(\mathbb{P})}$ on $\mathbb{R}$ is called the *probability distribution* of X with respect to $\mathbb{P}$. Note that different random variables (potentially defined on different probability spaces) may have the same distribution, while the same random variable may have a different distribution under a change of the probability measure.

We say that X has a *density* if there is a non-negative function $f : \mathbb{R} \to \mathbb{R}$ such that[1]

$$\mu_X^{(\mathbb{P})}([a,b]) = \int_a^b f(t)dt \qquad \text{for all } a \leq b.$$

A random variable X on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is said to be a *normal random variable* with mean μ and variance σ^2 if its distribution is the standard normal distribution; in other words, if

$$\mu_X^{(\mathbb{P})}([a,b]) = \frac{1}{\sigma\sqrt{2\pi}} \int_a^b e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt.$$

The inverse images of Borel subsets under a random variable X form a σ-subalgebra of $\mathcal{F}$, denoted $\sigma(X) = \{X^{-1}(B) \mid B \in \mathcal{B}\}$. We say that two random variables X and Y on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ are *independent* if, for any $A \in \sigma(X)$ and $B \in \sigma(Y)$, we have:

$$\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B)$$

2.3. **Lebesgue integral.** Let X be a random variable on $(\Omega, \mathcal{F})$ and let μ be a measure on this space. Assume for the moment that $X(\omega) \geq 0$ for any $\omega \in \Omega$. Let $\Pi = \{y_0, y_1, y_2, \ldots\}$ where $0 = y_0 < y_1 < y_2 < \cdots$. The *lower Lebesgue sum* of X with respect to Π is defined as:

$$\mathrm{LS}_\Pi^-(X) = \sum_{k=0}^{\infty} y_k \mu(X^{-1}[y_k, y_{k+1})).$$

In other words, we take a sum of the y_k, each weighted by the measure of points that X sends to the interval $[y_k, y_{k+1})$[2]. Taking the limit over finer and finer partitions Π, we

[1] This is a slightly simplified version. The function f must be Borel measurable function, i.e., $f^{-1}(B) \in \mathcal{B}$ for every $B \in \mathcal{B}$, and the condition is that $\mu_X^{(\mathbb{P})}(B)$ is the Lebesgue integral $\int_B f(t)dt$ for every $B \in \mathcal{B}$.

[2] Hence, we are partitioning the output space (y-axis) as opposed to the partition of the input space (x-axis) that occurs when defining the Riemann integral.

obtain the *Lebesgue integral* of X, which may be infinity:

$$\text{Lebesgue integral:} \qquad \int_\Omega X d\mu := \lim_{|\Pi|\to 0} \mathrm{LS}^-_\Pi(X)$$

where $|\Pi|$ indicates the maximum distance between successive points in a partition. This procedure can be easily extended in the case where X is not necessarily positive; in this case the Lebesgue integral may not exist. One can also readily make sense of the integral over any subset $A \in \mathcal{F}$ by considering $\mu(A \cap X^{-1}[y_k, y_{k+1}))$.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be probability space. The *expectation* of a random variable X on this space is defined as the Lebesgue integral $\mathbb{E}(X) = \int_\Omega X d\mathbb{P}$. Suppose X has a density f_X, and suppose $g : \mathbb{R} \to \mathbb{R}$ be a Borel-measurable function[3], then the expectation can be computed as an ordinary Riemann integral:

$$\mathbb{E}(g(X)) = \int_{-\infty}^{\infty} g(t) f_X(t) dt$$

provided that the quantity $\int_{-\infty}^{\infty} |g(t)| f_X(t) dt$ is finite.

3. Stochastic calculus

Stochastic calculus is an adaptation of ordinary calculus to the study of stochastic processes that may be nowhere differentiable. A *stochastic process* is, roughly speaking, a continuous path in the space of random variables. More precisely, it is a map $\Delta : \mathbb{R}_{\geq 0} \to \mathsf{RV}(\Omega, \mathcal{F})$ such that $\Delta^{(\omega)} : t \mapsto \Delta(t)(\omega)$ is continuous as a map $\mathbb{R} \to \mathbb{R}$, for every $\omega \in \Omega$. A fundamental stochastic process is Brownian motion, which we now define.

3.1. **Brownian motion.** Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. A stochastic process $W : \mathbb{R}_{\geq 0} \to \mathsf{RV}$ is said to be a *Brownian motion* if $W(0)$ is the constant random variable at zero, and, for any $0 = t_0 < t_1 < t_2 < \cdots < t_n$, the increment random variables $W(t_{i+1}) - W(t_i)$ are independent, and the i-th is normally distributed with mean 0 and variance $t_{i+1} - t_i$. Thus:

$$\mu_{W(t_{i+1}) - W(t_i)} \sim N(0, t_{i+1} - t_i).$$

For the remainder of this section, we fix a Brownian motion[4] W.

Proposition 3.1. *With probability one, we have:*

$$\lim_{|\Pi|\to 0} \sum_{j=0}^{n-1} \left[W(t_{j+1}) - W(t_j)\right]^p = \begin{cases} T & \text{if } p = 2 \\ 0 & \text{if } p > 2 \end{cases} \tag{3.1}$$

[3] That is, $g^{-1}(B) \in \mathcal{B}$ for all $B \in \mathcal{B}$

[4] We also implicitly fix a filtration $\mathcal{F}(t)$, for $t \in \mathbb{R}_{\geq 0}$, which is an increasing sequence of σ-algebras on Ω, all contained in $\mathcal{F}$, also satisfying (i) $[W(t)]^{-1}(B) \in \mathcal{F}(t)$ for any Borel subset B of $\mathbb{R}$, and (ii) $W(t) - W(s)$ is independent of $\mathcal{F}(s)$ for all $t > s$. See [Shr04, Chapters 2 and 3] for more details about filtrations.

We leave the full verification of the above result as an exercise (Exercise 6.2). Briefly, in the case $p = 2$, one can consider the standard normal random variable $Z_{j+1} = \frac{W(t_{j+1})-W(t_j)}{\sqrt{t_{j+1}-t_j}}$, and apply the law of large numbers, noting that $\mathbb{E}(Z_{j+1}^2) = 1$.

3.2. **The Itô integral.** Let Δ be a stochastic process defined on $[0, T]$. The *Itô integral* of Δ is defined as:

$$I_\Delta(T)(\omega) = \lim_{|\Pi|\to 0} \sum_{j=0}^{n-1} \Delta(t_j)(\omega) \cdot \left[W(t_{j+1}) - W(t_j)\right](\omega),$$

where $\Pi = \{t_0 = 1, t_1, \ldots, t_n = T\}$ be a partition of $[0, T]$, with $t_j < t_{j+1}$ and norm $|\Pi| = \max_j\{t_{j+1} - t_j\}$, so the limit is over finer and finer partitions[5]. It is common to suppress ω, and to write the Itô integral as $I(T) = \int_0^T \Delta(u)dW(u)$. A particular evaluation would require a choice of ω.

Theorem 3.2 (Itô–Doeblin formula)**.** *For any twice differentiable function $f : \mathbb{R} \to \mathbb{R}$, we have:*

$$f(W(T)) = f(0) + \int_0^T f'(W(u))dW(u) + \frac{1}{2}\int_0^T f''(W(u))du$$

Sketch of the proof. We fix a partition of $[0, T]$ and consider the Taylor expansion (truncating if necessary):

$$\begin{aligned} f(W(T)) - f(W(0)) &= \sum_{j=0}^{n-1} \left[f(W(t_{j+1})) - f(W(t_j))\right] \\ &= \sum_{j=0}^{n-1} \sum_{p=1}^{\infty} \frac{1}{p!} f^{(p)}(W(t_j)) \left[W(t_{j+1}) - W(t_j)\right]^p \end{aligned}$$

Next, we take the limit as $|\Pi| \to 0$. Proposition 3.1 implies that the Taylor expansion terms with $p > 2$ limit to zero. For $p = 1$, by definition, we obtain the Itô integral $\int_0^T f'(W(t))dW(t)$. For $p = 2$, Proposition (3.1) implies that we obtain the ordinary (Lebesgue) integral $\frac{1}{2}\int_0^T f''(W(t))dt$. The result follows, noting that $W(0) = 0$. ☐

The result $\sum_{j=0}^{n-1} \left[W(t_{j+1}) - W(t_j)\right]^2 \overset{|\Pi|\to 0}{\longrightarrow} T$ implies that Brownian motion accumulates quadratic variation at the rate of one per unit time. An informal way to summarize this result is as $dW(t)dW(t) = dt$, and we may write the Itô formula in differential form:

$$df(W(t)) = f'(W(t))dW(t) + \frac{1}{2}f''(W(t))dt$$

[5] A slightly more rigorous definition involves defining the integral for a simple process first, and then approximating any process with a sequence of simple processes.

3.3. **Itô processes.** An Itô process is one of the form:

$$X(T) = X(0) + \int_0^T \Delta(t) dW(t) + \int_0^T \Theta(t) dt \tag{3.2}$$

where $X(0)$ is non-random and $\Delta(t)$ and $\Theta(t)$ are adapted processes to the fixed filtration for Brownian motion W. (Again, we suppress the ω.) Hence, an Itô process consists of a non-random initial value, an Itô integral, and an ordinary integral. Using similar methods as above, one can verify the Itô–Doeblin formula for an Itô process:

$$f(T, X(T)) - f(0, X(0)) = \int_0^T \left[f_t(t, X(t)) + f_x(t, X(t))\Theta(t) + \frac{1}{2} f_{xx}(t, X(t))\Delta^2(t) \right] dt + \int_0^T f_x(t, X(t))\Delta(t) dW(t)$$

We informally write $dX(t) = \Delta(t)dW(t) + \Theta(t)dt$ and $dX(t)dX(t) = \Delta^2(t)dt$, and so an Itô process $X(t)$ satisfies the equation:

$$df(t, X(t)) = \left[f_t(t, X(t)) + f_x(t, X(t))\Theta(t) + \frac{f_{xx}(t, X(t)}{2}\Delta^2(t) \right] dt + f_x(t, X(t))\Delta(t) dW(t)$$

A *geometric Brownian motion* is a process of the form

$$S(t) = S(0)e^{X(t)}$$

where $X(t)$ is an Itô process as in (3.2). The process $S(t)$ satisfies the equation:

$$dS(t) = \alpha(t)S(t)dt + \sigma(t)S(t)dW(t)$$

where $\alpha(t) = \Theta(t) + \frac{1}{2}\Delta^2(t)$ is the *instantaneous mean rate of return* and $\sigma(t) = \Delta(t)$ is the *volatility*.

4. Black-Scholes-Merton equation

Our standing assumptions are (1) the existence of a money market account paying a constant interest rate of r, (2) there are no arbitrage opportunities, (3) there is sufficient market liquidity and no transaction fees, and (4) the stock is infinitely divisible with continuous time evolution (no jumps). The derivation of the Black–Scholes–Merton equation relies on argument using *hedging*, which is a general term for the practice of constructing a portfolio with less overall random fluctuation than the individual instruments in the portfolio. In this way, hedging provides insurance for investments at the cost of purchasing the off-setting contracts and possible reduced future payoff.

4.1. **Derivative pricing.** A *derivative* of a stock is a contract that derives its value from the performance of the stock. Examples of derivatives include call options, put options, and forward contracts; we discuss these in more detail below. Suppose we are working with a derivative whose payoff at time T depends only on the price $S(T)$ of the stock at time T. For example, a call option with strike K has payoff $\max(0, S(T) - K)$ at expiration time T. In this case, it can be shown that the price of the derivative at any time $t \in [0, T]$ depends only on the time t and the price $S(t)$ of the stock at that time. We write $f(t, x)$

for the price of the derivative at time $t \in [0, T]$ if the stock price is $x = S(t)$. We write f_t, f_x, f_{xx} for the indicated partial derivatives of f.

Theorem 4.1 (Black–Scholes–Merton Equation)**.** *If $S(t)$ is modeled by geometric Brownian motion with constant mean rate of return and constant volatility σ, then function f satisfies:*

$$rf(t,x) = rxf_x(t,x) + f_t(t,x) + \frac{1}{2}\sigma^2 x^2 f_{xx}(t,x).$$

Proof. Consider an agent who would like to hedge the derivative. The agent can do this with a portfolio consisting of some amount of this stock, as well as some amount of cash with interest rate r. If $X(t)$ is the total value of the portfolio, then the agent seeks to have

$$X(t) = f(t, S(t))$$

This equation holds if and only if the differentials of each side are equal. We find expressions for each of these differentials.

We first compute $dX(t)$. Let $\Delta(t)$ be the number of units of stock that the agent owns at time t, so that the agent's cash position at time t is $X(t) - \Delta(t)S(t)$. The stock component evolves as $\Delta(t)dS(t)$. Using the geometric Brownian motion assumption, we have $dS(t) = \alpha S(t)dt + \sigma S(t)dt$, where α is the (constant) mean rate of return. Meanwhile, the cash evolves as $r(X(t) - \Delta(t)S(t))dt$. Hence:

$$\begin{aligned} dX(t) &= \Delta(t)dS(t) + r(X(t) - \Delta(t)S(t))dt \\ &= [rX(t) + \Delta(t)(\alpha - r)S(t)]\, dt + \sigma\Delta(t)S(t)dW(t). \end{aligned}$$

Next, using the Itô–Doeblin formula, the evolution of $f(t,x)$ is given by:

$$\begin{aligned} &df(t, S(t)) = \\ &\left[f_t(t,S(t)) + \alpha S(t) f_x(t,S(t)) + \frac{1}{2}\sigma^2 S^2(t) f_x x(t,S(t)) \right] dt + \sigma S(t) f_x(t,S(t)) dW(t) \end{aligned}$$

Finally, setting the differentials equal to each other, i.e. $dX(t) = df(t,S(t))$, we first equate the dt terms to obtain:

$$\Delta(t) = f_x(t, S(t))$$

This is known as the delta-hedging rule. Equating the $dW(t)$ terms, we obtain:

$$rX(t) + \Delta(t)(\alpha - r)S(t) = f_t(t,S(t)) + \alpha S(t) f_x(t,S(t)) + \frac{1}{2}\sigma^2 S^2(t) f_{xx}(t,S(t))$$

Substituting $\Delta(t) = f_x(t, S(t)$ and $X(t) = f(t, S(t))$, rearranging terms, and writing x instead of $S(t)$, we arrive at the Black–Scholes–Merton differential equation. □

Remark 4.2. The easiest example of a derivative is a single unit of stock, so that $f(t,x) = x$. Alternatively, one can invest K in the money market account, giving $f(t,x) = e^{rt}K$, and this also counts as a derivative even though there is no dependency on $x = S(t)$. Both these functions satisfy the Black–Scholes–Mertion equation with initial conditions $f(0,x) = x$ and $f(0,x) = K$, respectively.

4.2. **Forward contracts.** A *forward contract* is an agreement obligating its holder to buy one share of the stock at the time of expiry T for the delivery price K. Hence, it is an example of a derivative. Let $f(t,x)$ be the value of the forward contract at time $t \in [0,T]$ if the stock price at time t is $x = S(t)$. Hence, at expiration, we have $f(T,x) = S(T) - K$. Furthermore:

Lemma 4.3. *The value of a forward contract with expiration time T and delivery price K is:*

$$f(t,x) = x - e^{-r(T-t)}K \qquad \text{for } t \in [0,T]$$

Proof. An agent can perform a static hedge by selling the forward contract at $S(0) - e^{-rT}K$ and purchasing one unit of the stock. This requires a loan of $e^{-rT}K$ from the money market account. At time t, the value of the agent's portfolio is $S(t) - e^{-r(T-t)}K$; in particular, at expiration the value matches that of the forward contract. Hence it is a replicating portfolio for the forward contract, and the result follows from the no-arbitrage assumption. □

Remark 4.4. Lemma 4.3 does not require that the stock follows geometric Brownian motion. Hence it applies more broadly to situations where there is sufficient liquidity and constant interest rate. At the same time, the function for the forward value is still a solution to the Black–Scholes–Merton equation with terminal condition $f(T,x) = x - K$.

Remark 4.5. The *forward price* is the value of K which makes the forward contract have value zero at time t, i.e., $\mathrm{For}_S(t,T) = e^{r(T-t)}S(t)$.

4.3. **Options.** A *call option* (resp. *put option*) is a contract allowing the holder to purchase (resp. sell) the stock at the time of expiry T for the strike price of K. The holder of the option is not obligated to buy or sell at the time of expiry. We see that options are derivatives whose price at expiration are given by $\max(0, S(T) - K)$ in the case of calls and $\max(0, K - S(T))$ in the case of puts. Let $c(t,x)$ and $p(t,x)$ be the price of a call option (resp. put option) at time t if the stock price is $x = S(t)$.

Proposition 4.6 (Put-Call Parity)**.** *For any $t \in [0,T]$, we have:*

$$c(t,x) - p(t,x) = x - e^{-r(T-t)}K$$

Proof. A portfolio consisting of a purchased call and a sold put is equivalent to a forward contract. In particular, both portfolios have value at expiration equal to $\max(0, S(T) - K) - \max(0, K - S(T)) = S(T) - K$. The no-arbitrage assumption implies that their prices will be the same for any $t \in [0,T]$, and we apply Lemma 4.3. □

Since options are derivatives of the stock, the Black–Scholes–Merton equation applies:

Proposition 4.7. *If $S(t)$ is modeled by geometric Brownian motion with constant mean rate of return and constant volatility, then the call and put option prices satisfy the Black–Scholes–Merton equation with terminal conditions:*

$$c(T,x) = \max(0, x-K) \qquad p(T,x) = \max(0, K-x)$$

5. Black–Scholes–Merton formula

We explain a method of finding a solution to the Black–Scholes–Merton differential equation. A key step in this method involves changing the original 'real-world' probability measure to a 'risk-neutral' probability measure. The terminology is somewhat unfortunate since since both measures apply to the same sample space Ω, which, in financial applications, represents all possible future states of the markets. Both measures agree on what is impossible; specifically, a subset of Ω has probability 1 under one measure if and only if it has probability 1 under the other measure. This is enough to construct hedges, since a hedge that works almost surely under one measure also works almost zsurely under the other measure.

5.1. **Change of measure.** Fix $(\Omega, \mathcal{F})$. One can use random variables to change probability measures via Lebesgue integration. To be precise, consider all pairs $(Z, \mathbb{P})$ consisting of a random variable and a probability measure such that the expectation of Z under $\mathbb{P}$ is equal to 1, that is: $\int_\Omega Z d\mathbb{P} = 1$. Denoting this set of pairs as $(\mathsf{RV} \times \mathsf{PM})_{\text{unit}}$, we have a map:

$$(\mathsf{RV}(\Omega, \mathcal{F}) \times \mathsf{PM}(\Omega, \mathcal{F}))_{\text{unit}} \longrightarrow \mathsf{PM}(\Omega, \mathcal{F}), \qquad (X, \mathbb{P}) \mapsto \mathbb{P}^{(Z)} = [A \mapsto \int_A Z d\mathbb{P}]$$

Next we consider changing the measure to turn a Brownian motion with drift into a Brownian motion with no drift.

Proposition 5.1 (Simple version of Girsanov Theorem)**.** *Let $W(t)$ be Brownian motion on $(\Omega, \mathcal{F}, \mathbb{P})$, defined for $t \in [0, T]$. Fix a constant θ, and consider the random variables:*

$$Z = \exp\left(-\theta W(T) - \frac{1}{2}\theta^2 T\right) \qquad \text{and} \qquad \tilde{W}(t) = \theta t + W(t)$$

Then $\mathbb{E}(Z) = 1$ and $\tilde{W}(t)$ is a Brownian motion on $[0, T]$ under the probability measure $\mathbb{P}^{(Z)}$.

Sketch of proof. To show that $\mathbb{E}(Z) = 1$, let $Z(t) = \exp\left(-\theta W(t) - \frac{1}{2}\theta^2 t\right)$. A bit of Itô calculus shows that $Z(t)$ is a martingale, so that the expectation of the difference $Z(t) - Z(s)$ is zero for all $0 \leq s < t \leq T$. Hence the expectation of $Z = Z(T)$ is equal to the expectation of $Z(0)$, which is 1.

Next, one must show that, for the probability measure $\mathbb{P}^{(Z)}$ and any $0 = t_0 < t_1 < t_2 < \cdots < t_n$, the increment random variables $\tilde{W}(t_{i+1}) - \tilde{W}(t_i)$ are independent, and the i-th is normally distributed with mean 0 and variance $t_{i+1} - t_i$. It is straightforward to verify the assertions of independence, normal distribution, and variance. The claim of mean 0 is equivalent to showing that:

$$\mathbb{E}(\tilde{W}(t) Z) = 0$$

To see this, first write $\tilde{W}(t) Z = \tilde{W}(t) Z(t) \frac{Z}{Z(t)}$ and observe that $\tilde{W}(t) Z(t)$ depends on $\omega \in \Omega$ only through $W(t)$. On the other hand, $\frac{Z}{Z(t)} = \exp\left(-\theta\left(W(T) - W(t)\right) - \frac{1}{2}\theta^2(T - t)\right)$ depends on $\omega \in \Omega$ only through $W(T) - W(t)$. It follows that $\tilde{W}(t) Z(t)$ and $\frac{Z}{Z(t)}$ are independent random variables, so the expectation of their product is the product of their expectations. Finally, an application of Exercise 6.5 implies that $\mathbb{E}(\tilde{W}(t)Z(t)) = 0$. (One can additionally use the martingale property of $Z(t)$ to that that the expectation of the quotient $Z/Z(t)$ is equal to 1.) □

Now consider geometric Brownian motion used to model a stock with instantaneous rate of return α and volatility σ, both of which are constant functions of t:

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

Set $\theta = \frac{\alpha - r}{\sigma}$ and let $\tilde{W}(t) = \theta t + W(t)$, so that

$$S(t) = S(0)\exp\left(\sigma \tilde{W}(t) + \left(r - \frac{1}{2}\sigma^2\right)t\right)$$

Applying Proposition 5.1, we see that there is a random variable Z such that $\tilde{W}$ is a Brownian motion under the probability measure $\mathbb{P}^{(Z)}$, and the random variable $Y(T) = \ln\left(\frac{S(T)}{S(0)}\right)$ is normally distributed with mean $\left(r - \frac{\sigma^2}{2}\right)T$ and variance $\sigma^2 T$. The probability measure $\mathbb{P}^{(Z)}$ is called the risk-neutral measure, since both the stock and the money market account have the same rate of return r.

5.2. **The formula.** Suppose we have a call option with strike price K expiring at time T. Let t be the current time, with $0 \le t \le T$, and let $x = S(t)$ be the current price of the underlying asset. The time until expiration is given by $\tau = T - t$. Set $d_\pm(\tau, x) = \frac{1}{\sigma\sqrt{\tau}}\left(\ln\left(\frac{x}{K}\right) + \left(r\tau \pm \frac{\sigma^2}{2}\right)\right)$.

Theorem 5.2. *With the above notation, and under the standing assumptions, solutions to the Black–Scholes–Merton equation for calls and puts is given by:*

$$c(t,x) = xF_{std}\left(d_+(\tau,x)\right) - Ke^{-r\tau}F_{std}\left(d_-(\tau,x)\right)$$
$$p(t,x) = Ke^{-r\tau}F_{std}\left(-d_-(\tau,x)\right) - xF_{std}\left(-d_+(\tau,x)\right)$$

where F_{std} is the cumulative density function of a standard normal random variable.

Proof. The geometric Brownian motion assumption implies that, under the risk-neutral probability measure, the price $S = S(T)$ of the underlying at expiration is a log-normal random variable with mean $\left(r - \frac{\sigma^2}{2}\right)\tau$ and variance $\sigma^2\tau$, i.e., $Y = \ln\left(\frac{S}{x}\right) \sim \mathcal{N}\left(r\tau - \frac{\sigma^2\tau}{2}, \sigma^2\tau\right)$. The current value of the option is given by:

$$\begin{aligned}\mathbb{E}\left(e^{-r\tau}\max(0, S-K)\right) &= e^{-r\tau}\int_K^\infty (s-K)f_S(s)ds \\ &= e^{-r\tau}\int_K^\infty sf_S(s)ds + Ke^{-r\tau}\int_K^\infty f_S(s)ds\end{aligned}$$

where f_S is the probability density function of S. We first analyze the integral appearing in the second term:

$$\int_K^\infty f_S(s)ds = P(S \geq K) = P\left(Y \geq \ln\left(\frac{K}{x}\right)\right) = P\left(Z \geq \frac{\ln\left(\frac{K}{x}\right) - r\tau + \frac{\sigma^2\tau}{2}}{\sigma\sqrt{\tau}}\right)$$
$$= F_{\text{std}}\left(-\frac{\ln\left(\frac{K}{x}\right) - r\tau + \frac{\sigma^2\tau}{2}}{\sigma\sqrt{\tau}}\right) = F_{\text{std}}\left(\frac{\ln\left(\frac{x}{K}\right) + r\tau - \frac{\sigma^2\tau}{2}}{\sigma\sqrt{\tau}}\right) = F_{\text{std}}\left(d_-(\tau, x)\right)$$

where Z is a standard normal random variable with cumulative density function F_{std}, and we use the fact that $P(Z \geq d) = F_{\text{std}}(-d)$. Next we analyze the integral appearing in the first term:

$$\int_K^\infty s f_S(s)ds = \int_K^\infty \frac{s}{s\sqrt{2\sigma\tau}} \exp\left(-\frac{\left(\ln\left(\frac{s}{x}\right) - \left(r\tau - \frac{\sigma^2\tau}{2}\right)\right)^2}{2\sigma^2\tau}\right) ds$$
$$= \int_{\ln\left(\frac{K}{x}\right)}^\infty \frac{1}{\sqrt{2\sigma\tau}} \exp\left(-\frac{\left(y - \left(r\tau - \frac{\sigma^2\tau}{2}\right)\right)^2}{2\sigma^2\tau}\right) xe^y dy$$
$$= \frac{xe^{r\tau}}{\sqrt{2\sigma\tau}} \int_{\ln\left(\frac{K}{x}\right)}^\infty \exp\left(-\frac{\left(y - \left(r\tau + \frac{\sigma^2\tau}{2}\right)\right)^2}{2\sigma^2\tau}\right) dy$$
$$= xe^{r\tau} P\left(Z \geq \frac{\ln\left(\frac{K}{x}\right) - r\tau - \frac{\sigma^2\tau}{2}}{\sigma\sqrt{\tau}}\right) = xe^{r\tau} F_{\text{std}}\left(-\frac{\ln\left(\frac{K}{x}\right) - r\tau - \frac{\sigma^2\tau}{2}}{\sigma\sqrt{\tau}}\right)$$
$$= xe^{r\tau} F_{\text{std}}\left(\frac{\ln\left(\frac{x}{K}\right) + r\tau + \frac{\sigma^2\tau}{2}}{\sigma\sqrt{\tau}}\right) = xe^{r\tau} F_{\text{std}}\left(d_+(\tau, x)\right)$$

where the first equality uses the probability density function of a log-normal random variable; the second makes the substitution $y = \ln\left(\frac{s}{x}\right)$; the third results from completing the square (see Exercise 6.4); and the last steps are elementary manipulations with the standard normal random variable Z and its cumulative density function. Combining the above calculations of the two integrals, we obtain the desired result for call options. The result for put options follows from the put-call parity formula (Proposition 4.6). ☐

Remark 5.3. Under a change of variables, the Black–Scholes–Merton equation becomes the one-dimensional heat equation. Namely, first let $y = \ln(x)$, which makes the equation have constant coefficients. Then set $\tau = T - t$ (the time remaining until expiration) and $g(\tau, x) = e^{r\tau} f(\tau, x)$, which gets rid of the zero-th order term. Finally, set $z = y + \left(r - \frac{1}{2}\right)\tau$, thus eliminating the first-order term and arriving at:

$$g_\tau = \frac{1}{2}\sigma^2 g_{zz}$$

The solutions presented in Theorem 5.2 match those obtained via standard methods for solving the heat equation.

5.3. **The Greeks.** The "*Greeks*" refers to the derivatives of the call and put option formulas with respect to various variables. For example, we have:

- Delta:
$$c_x(\tau, x) = N(d_+(\tau, x))$$
- Gamma:
$$c_{xx}(\tau, x) = \frac{N'(d_+(\tau, x))}{x\sigma\sqrt{\tau}}$$
- Theta:
$$c_t(\tau, x) = -rKe^{-r\tau}N(d_-(\tau, x)) - \frac{\sigma x N'(d_+(\tau, x))}{2\sqrt{\tau}}$$
- Vega:
$$\frac{\partial c(t, x \mid \sigma)}{\partial \sigma} = SN'(d_+(\tau, x))\sqrt{\tau}$$

One sees that delta and gamma are always positive for a long call. This means that the price of a call increases as the price of the underlying rises, and that the pricing function of the call is convex. Additionally, theta is always negative, that is, the price of an option decreases with time. Equivalently, the price of a call option falls as the time to expiration decreases. Finally, vega is positive, so higher volatility implies higher prices.

6. Exercises

6.1. **Space of random variables.** Show that the space of random variables on $(\Omega, \mathcal{F})$ forms a vector space. Hint: consider the countable union:

$$(X + Y)^{-1}([0, \infty)) = \bigcup_{r \in \mathbb{Q}} \left(X^{-1}([r, \infty]) \cap Y^{-1}([-r, \infty)) \right)$$

6.2. **Partitions and limits.** Let Π be a partition as above. For $p \geq 2$, we have the following convergences as $|\Pi| \to 0$:

- $\sum_{j=0}^{n-1} \left(t_{j+1} - t_j\right)^p \to 0$
- $\mathbb{E}\left[\sum_{j=0}^{n-1} \left(W(t_{j+1}) - W(t_j)\right)^p\right] \to T$ if $p = 2$, otherwise $\to 0$.
- Var $\left[\sum_{j=0}^{n-1} \left(W(t_{j+1}) - W(t_j)\right)^p\right] \to 0$.
- $\sum_{j=0}^{n-1} \left(W(t_{j+1}) - W(t_j)\right)^p \to T$ if $p = 2$, otherwise $\to 0$.

Hint: for the first identity, factor out $|\Pi|^{p-1}$. For the second and third identities, note that $W(t_{j+1}) - W(t_j)$ is normal random variable with mean 0 and variance $t_{j+1} - t_j$, so one can easily compute (or look up) the higher moments. The final identity follows from the previous two.

6.3. **Different types of integrals.** Let g be a continuous function on the interval $[0, T]$. Show that:

$$\lim_{|\Pi|\to 0} \sum_{j=0}^{n-1} g(W(t_j)) \left[W(t_{j+1}) - W(t_j)\right]^p = \begin{cases} \int_0^T g(W(t))dW(t) & \text{if } p = 1 \\ \int_0^T g(W(t))dt & \text{if } p = 2 \\ 0 & \text{if } p \geq 3 \end{cases}$$

Note that the $p = 1$ case is by definition.

6.4. **Completing the square.** Show that $\exp\left(\frac{(x+a)^2}{c}\right) e^{2bx} = \exp\left(\frac{(x+a+bc)^2}{c}\right) e^{-b(2a+bc)}$. Apply this identity with $a = -r\tau + \frac{\sigma^2\tau}{2}$, $b = \frac{1}{2}$, and $c = -2\sigma^2\tau$.

6.5. **Change of measure.** Let X be a standard normal random variable on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Let $\theta \in \mathbb{R}$ be a constant and define $Y = X + \theta$. Under $\mathbb{P}$, the random variable Y normal with mean θ and variance 1. Set

$$Z = \exp\left(-\theta X - \frac{1}{2}\theta^2\right)$$

Show that $\mathbb{E}(Z) = 1$ and that Y is standard normal under $\mathbb{P}^Z$.

References

[Rom12] Steven Roman, *Introduction to the Mathematics of Finance: Arbitrage and Option Pricing*, Undergraduate Texts in Mathematics, Springer, New York, NY, 2012.

[Shr04] Steven Shreve, *Stochastic Calculus for Finance II*, Springer Finance Texbook, Springer, 2004 (en).

Appendix A. From densities to random variables

Above, we described how one arrives at a probability measure on $\mathbb{R}$ from a random variable on some measurable space $(\Omega, \mathcal{F}, \mathbb{P})$. This probability measure often has a density. In this appendix, we describe the opposite procedure:

Task: given a density function $f : \mathbb{R} \to [0,1]$, produce a random variable whose probability distribution admitting f as its density function f.

Step 1. Let $F(x) = \int_{-\infty}^{x} f(t)dt$ be the cumulative density function. Note that F is nondereasing with derivative f.

Step 2. We claim that F is injective on the support of f. Indeed, we argue the contrapositive. Let $a \in \mathbb{R}$ and suppose there exists $b \in \mathbb{R}$ such that $a \neq b$ and $F(a) = F(b)$. We show that $f(a) = 0$. Without loss of generality, assume $a < b$. Then, since F is nondecreasing, we have that $F(a) = F(c)$ for all $c \in [a,b]$. It follows that $f(c) = F'(c) = 0$ for all $c \in [a,b]$. In particular, $f(a) = 0$.

Step 3. Let $G : F(\text{supp}(f)) \to \mathbb{R}$ be the inverse of F on $F(\text{supp}(f))$. So $G(F(x)) = x$ for $x \in \mathbb{R}$ such that $f(x) \neq 0$, and $F(G(z)) = z$ for $z \in F(\text{supp}(f)) \subseteq [0,1]$.

Step 4. Note that $f^{-1}(0)$ is a union of closed intervals and F is constant on each connected component of $f^{-1}(0)$. We see that $F(f^{-1}(0))$ has measure zero under the usual measure on $[0,1]$.

Step 5. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, and let $Y : \Omega \to \mathbb{R}$ be a random variable such that the distribution function of Y is the uniform distribution on $[0,1]$ (see [Shr04, Examples 1.2.4 and 1.2.5]). Observe that $F(f^{-1}(0))$ has measure zero in $[0,1]$ under the uniform measure. It follows that $Y^{-1}(F(f^{-1}(0)))$ has measure zero in Ω. Now let $\Omega^\circ = \Omega \setminus Y^{-1}(F(f^{-1}(0)))$, and let $\mathcal{F}^\circ = \{A \cap \Omega^\circ \in \mathcal{F} \mid A \in \mathcal{F}\}$. Then the restriction of $\mathbb{P}$ to $\mathcal{F}^\circ$ is a probability measure. Define a new random variable $X : \Omega^\circ \to \mathbb{R}$ by $X(\omega) = G(Y(\omega))$. Then one can show that the distribution of X has density function f (see [Shr04, Example 1.2.6]).