Communication

Information Theory and the Pricing of Contingent Claims: An Alternative Derivation of the Black–Scholes–Merton Formula

Thomas P. Davis (ID)

FactSet Research Systems, Broadgate Quarter, London EC2A 2DQ, UK; todavis@factset.com

Abstract: This paper seeks to determine the best subjective probability to use to carry out expectation values of uncertain future cash flows with the smallest number of assumptions. This results in the unique distribution that guarantees no more information is present other than the stated assumptions. The result is a novel derivation of the well-known Black–Scholes equation without the need to introduce high-level mathematical machinery. This formalism fits nicely into introductory courses of finance, where the value of any financial instrument is given by the present value of uncertain future cash flows.

Keywords: Information Theory; derivative pricing; quantitative finance

1. Introduction

Early on in finance, we learn that the fair value of a security is the present value of expected future cash flows:

$$V(t) = \sum_i Z(t, T_i) \mathbb{E}[c(T_i)], \tag{1}$$

where $Z(t, T)$ is the price of a zero-coupon bond at time t, maturing at time T. Many of the formulae for fixed-income security pricing result from Equation (1) when the cash flows $c(T_i)$ are known with certainty Fabozzi and Mann (2021). However, when the cash flows are uncertain, it is natural to ask what probabilities should be used to determine the average implied by the expectation value in Equation (1).

The overwhelmingly accepted answer is, of course, the famous "risk-neutral probabilities" introduced by Black and Scholes (1973), in which many assumptions were made in order to justify the approach. The assumptions were those of the capital asset pricing model (CAPM) Lintner (1965); Sharpe (1964), which depends upon general economic equilibrium and also upon secondary assumptions such as a constant interest rate (or at least non-stochastic interest rate) and the variance of the stock over the lifetime of the option.

Robert Merton did not agree with many of these assumptions, especially the condition of economic equilibrium, saying of the derivation, "the portfolio weights are chosen to eliminate all 'market risk'. By the assumptions of the CAPM, any portfolio with zero ('beta') market risk must have an expected return equal to the risk-free rate. Hence an equilibrium condition is established between the expected return on the option, the expected return on the stock and the risk-less rate" Merton (1973). Merton then set out to relax some of these assumptions by adding stochastic interest rates and, more importantly, proving that the expectations of the investors play no part in the fair value of an option. This insight came from a precise cancellation of terms when solving the equations for the portfolio to be arbitrage-free, that is, the initial capital investment is all that is required to hedge the portfolio without any inflows or outflows; the gains and losses of the portfolio suffice. Thus, the equilibrium assumptions of the CAPM are not required.

These equations are now known as the "market price of risk" and form the basis of dynamic asset pricing theory, as developed by Harrison and Pliska (1981), who developed

the fundamental theorem of asset pricing by specifying the mathematical and economic requirements for arbitrage freedom in a given financial market.

The mathematical machinery needed to price a simple European option is, therefore, quite complex, requiring the knowledge of stochastic calculus, dynamic asset pricing theory and the solution of partial differential equations—very intimidating for a young student in finance. This paper seeks to determine a financially justified probability distribution without relying on any of this advanced mathematical machinery. The frameworks used are those of Information Theory Shannon (1948) and the principle of maximum entropy developed by Jaynes (1957b, 2002).

2. Information Theory and Maximum Entropy

In 1948, Claude Shannon introduced Information Theory in his landmark paper "A Mathematical Theory of Communication" Shannon (1948). Here, Shannon introduced the concept of informational entropy of a discrete probability distribution:[1]

$$H = -\sum_i p_i \ln p_i,$$ (2)

as a measure of information. It has intuitive properties that the information about an event should satisfy, such as being zero when the outcome is certain—i.e., only one of the probabilities is 1.0—and achieves a maximum when all outcomes are equally likely. From this fundamental quantity, he determined optimal ways to encode communication messages into abstract codes, where the environment can adversely affect the message (i.e., what is transmitted is not what is received).

Later, in 1957, Ed Jaynes noted that the definition of entropy in statistical physics could be thought of as informational entropy Jaynes (1957a) and introduced a way to generate a probability distribution that produces the empirical data without adding any additional assumptions, i.e., a distribution that "...has the important property that no possibility is ignored; it assigns a positive weight to every situation that is not absolutely excluded by the given information".

We can apply this concept in finance, as it is an open question whether asset price fluctuations can have an objective probability distribution. The maximum entropy framework does not assume one, whereas an objective probability distribution is required by the Black–Scholes–Merton framework—all models that attempt to go beyond the Black–Scholes–Merton assumptions (Cox and Ross 1976; Dupire 1994; Heston 1993; Merton 1976). Stocks are not physical particles subject to the immutable laws of physics; they are, rather, a human construct with prices traded in a market driven by human emotion. Thus, we seek to obtain a *subjective* probability distribution that not only relies on all of the information that exists in the market, but is "maximally noncommittal to missing information" Jaynes (1957a), meaning this distribution, once found, does not (cannot) contain any further assumptions than those that are used in its derivation. In this paper, we show that only two assumptions are required to derive this unique distribution:

1. There is a forward contract on the stock, which is fairly priced in the market;
2. The distribution has a variance; otherwise, statistical measures of risk are difficult to quantify.

Using only these assumptions, the subjective probability can be determined.

In the first section of this paper, we use the method of Lagrange multipliers to apply the constraints and derive the subjective probability. In the second section, we use this unique probability distribution to determine the price of a European call option and show that the Black–Scholes equation is obtained.

3. Subjective Probability

We assume that there is a liquidly traded market in forward contracts for this stock. A forward contract is a contract struck at time *t*, where both parties are obliged to trade

the stock at a later time T for an agreed-upon price, regardless of whether or not the buyer could obtain a lower price in the market at that future time. We denote the forward price by $F(t,T)$.

Suppose that the forward price can be obtained by summing over all possible future values that the stock could possibly attain, weighted by a probability

$$F(t,T) = \sum_\alpha p_\alpha S_\alpha \equiv \mathbb{E}[S(T)]. \tag{3}$$

We can think of this probability as containing the totality of information that all market participants collectively possess. It is this probability that we seek to determine.

The second piece of information reflects the fact that the market participants know, almost certainly, that the stock price will not exactly attain the forward price at time T; there is some uncertainty in the future outcome. They must decide on a measure of uncertainty in the forward price $F(t,T)$, and, more importantly, they must decide how this measure of uncertainty is determined and quoted. Once chosen, the participant can calculate the subjective distribution, which represents the least-biased estimate based on these two pieces of information. By following this prescription, we show how to arrive at the Black–Scholes formula.

To motivate the difference between this approach and the traditional approach, we examine what we mean by a subjective probability distribution rather than an objective one. By subjective, we mean "the sense that it describes only a state of knowledge, and not anything that could be measured in a physical experiment" Jaynes (2002), rather than a proscriptive probability distribution. For instance, one could posit a stochastic differential equation for the underlying stock price:

$$dS(t) = \mu(S,t)dt + \sigma(S,t)dW(t), \tag{4}$$

which a priori describes the dynamics of the stock price and, in principle, gives complete information on the statistics of the movements. A posteriori, these statistics can be empirically tested and (4) can be augmented (or rejected) if need be.

On the other hand, subjective probability distributions do not proscribe any dynamics or constrain the system at all; they are an inferential tool used when the underlying dynamics of the system are too complex to determine (or do not even exist at all).

To begin, we specify that volatility is a measure of the distribution of returns on the underlying asset. Mathematically, we write that the variance is an average of the square of the logarithm of the future underlying price relative to the current price:

$$v(T) = \mathbb{E}\left[\left(\ln \frac{S(T)}{S(t)}\right)^2\right]. \tag{5}$$

By using the maximum entropy framework to determine the probabilities, we guarantee that the results do not implicitly contain any further assumptions or biases.

The derivation of the subjective distribution proceeds by maximization of the Shannon entropy Shannon (1948):

$$H = -\sum_\alpha p_\alpha \ln p_\alpha, \tag{6}$$

subject to the constraints

$$g_0 = 1 - \sum_\alpha p_\alpha, \tag{7}$$

$$g_1 = \mu(t,T) - \sum_\alpha \ln \frac{S_\alpha}{S(t)} p_\alpha, \tag{8}$$

$$g_2 = v(t,T) - \sum_\alpha \left(\ln \frac{S_\alpha}{S(t)} \right)^2 p_\alpha. \tag{9}$$

These particular constraints enforce normalization of the probability, i.e., that the forward price will be priced correctly (the functional form of μ will be fixed a posteriori to enforce the average (3)) and that the volatility is given by Equation (5).

Functional maximization subject to constraints is carried out by the method of Lagrange multipliers:

$$\delta H + \lambda \delta g_0 + \gamma_1 \delta g_1 + \gamma_2 \delta g_2 = 0, \tag{10}$$

leading to the probability

$$p_\alpha = \exp\left(-(1+\lambda) - \gamma_1 \ln \frac{S_\alpha}{S(t)} - \gamma_2 \left(\ln \frac{S_\alpha}{S(t)} \right)^2 \right). \tag{11}$$

The Lagrange multipliers are determined by the set of constraint functions $\{g_i = 0\}$, $i = 1, 2, 3$. Instead of fixing λ, we introduce the "partition function" $Z \equiv \exp(1 + \lambda)$, where

$$Z = \sum_\alpha \exp\left(-\gamma_1 \ln \frac{S_\alpha}{S(t)} - \gamma_2 \left(\ln \frac{S_\alpha}{S(t)} \right)^2 \right). \tag{12}$$

With this definition, the probability in (11) becomes

$$p_\alpha = \frac{1}{Z} \sum_\alpha \exp\left(-\gamma_1 \ln \frac{S_\alpha}{S(t)} - \gamma_2 \left(\ln \frac{S_\alpha}{S(t)} \right)^2 \right). \tag{13}$$

In order to make contact with the Black–Scholes derivation, we make the substitution $\sum_\alpha \to \int_0^\infty dx$. The partition function now has a closed-form solution:

$$Z = S(t) \sqrt{\frac{\pi}{\gamma_2}} \exp\left(\frac{1}{4} \frac{(\gamma_1 - 1)^2}{\gamma_2} \right). \tag{14}$$

The further two constraints $\{g_i = 0\}$, $i = 1, 2$ are enforced by choosing the γ_is such that

$$\mu(t,T) = -\frac{\partial}{\partial \gamma_1} \ln Z, \tag{15}$$

$$v(t,T) - \mu(t,T)^2 = -\frac{\partial}{\partial \gamma_2} \ln Z - \left(\frac{\partial}{\partial \gamma_1} \ln Z \right)^2. \tag{16}$$

These equations have solution

$$\gamma_1 = 1 - \frac{\mu(t,T)}{v(t,T)}, \tag{17}$$

$$\gamma_2 = \frac{1}{2v(t,T)}. \tag{18}$$

With the constraints enforced, the probability in (11) becomes

$$p(x,t,T) = \frac{1}{x\sqrt{2\pi v(t,T)}} \exp\left(-\frac{1}{2v(t,T)^2} \left(\ln \frac{x}{S(t)} - \mu(t,T) \right)^2 \right). \tag{19}$$

The next step is to identify the averages $v(t, T)$ and $\mu(t, T)$ with their corresponding financial parameters. The realized volatility is identified as the integral of the volatility term structure that the market participants wish to use to model the underlying

$$v(t, T) = \int_t^T \sigma^2(s)ds; \tag{20}$$

the maximum entropy distribution does not limit this choice. [2]

The average quantity $\mu(t, T)$ is fixed by the forward price (3):

$$F(t, T) = \int_0^\infty xp(x, T)dx, \tag{21}$$

which is an implicit equation fixing $\mu(t, T)$ in terms of the forward price $F(t, T)$ with solution

$$\mu(t, T) = \ln \frac{F(t, T)}{S(t)} - \frac{1}{2}v(t, T). \tag{22}$$

We are now in a position to price any contingent claim armed with the probability distribution

$$p(x, t, T) = \frac{1}{x\sqrt{2\pi v(t, T)}} \exp\left(-\frac{1}{2v(t, T)}\left(\ln \frac{x}{F(t, T)} + \frac{1}{2}v(t, T)\right)^2\right). \tag{23}$$

In particular, we can now price the European call option by explicitly calculating the present value of the uncertain future cash flow using the probability distribution that is guaranteed to not have any more information, or assumptions, than the two that we imposed: it prices forward prices exactly, and the distribution has a variance.

4. Re-Deriving the Black–Scholes Formula

We look at the simplest security with uncertain cash flows, that of a European call option on a stock with a current traded market price $S(t)$. The option gives the buyer the right, but not the obligation, to purchase the stock at the strike price K at some point in the future time $T > t$. That is, we seek to determine the unique value of $C(t, T)$ that is the present value of the uncertain future cash flow:

$$C(t, T) = Z(t, T)\mathbb{E}[\max(S - K, 0)]. \tag{24}$$

Inserting the subjective probability (23) into this payoff formula, i.e.,

$$C(t, T) = Z(t, T)\left(\int_K^\infty xp(x, t, T)dx - K\int_K^\infty p(x, t, T)dx\right) \tag{25}$$

the price can be calculated exactly:

$$C(t, T) = Z(t, T)(F(T)\mathcal{N}(d_+) - K\mathcal{N}(d_-)), \tag{26}$$

where $\mathcal{N}(x)$ is the cumulative normal distribution and

$$d_\pm = \frac{\ln \frac{F(t, T)}{K} \pm \frac{1}{2}v(t, T)}{\sqrt{v(t, T)}}. \tag{27}$$

Equation (26) is the celebrated Black–Scholes formula for the price of a European call option with a volatility term structure.

5. Conclusions

The standard way of teaching the pricing of contingent claims in finance classes relies on the Black–Scholes–Merton framework, stochastic calculus and dynamic asset pricing theory. While mathematically sound, this approach has two major pedagogical drawbacks. First, the student may not have been exposed to such advanced mathematics, and second, more fundamentally, a stock price may not have an objective probability distribution such as, say, a diffusing pollen particle. Stocks are human creations and their trading depends on human emotion. Further, if a model becomes accepted in the market, this can fundamentally alter the way an asset trades.

In this paper, we have derived the same equation without the use of stochastic calculus or dynamic asset pricing theory; therefore, this approach could be better suited to early finance courses. Furthermore, the Black–Scholes equation is derived with only two assumptions: that a forward contract is traded in the market, and that the probability distribution used has some measure of dispersion and, importantly, a guarantee that no other information has been used. A recent paper extends the analysis in this paper by using the maximum entropy formulation and all available option prices to infer higher moments of the distribution Ardakani (2022).

Although we do not believe this supplants the standard treatment of pricing-contingent claims, this method offers a new perspective on the age-old problem, and one that fits easily into an introductory framework of finance without the need to complicate matters by the introduction of advanced concepts such as filtrations, stochastic calculus and numeraires.

Notes

[1]   Shannon used the logarithm of base 2 in the paper, as he was concerned with error-correcting codes of binary digits. Since we are interested in the continuous case, we use the natural logarithm.

[2]   In the standard treatment of Black–Scholes, the volatility of the underlying asset is assumed to be a constant, resulting in the realized variance, as given by the formula volatility, becoming $v(t, T) = \sigma^2 T$.

References

Ardakani, Omid M. 2022. Option pricing with maximum entropy densities: The inclusion of higher-order moments. *Journal of Futures Markets* 42: 1821–36. [CrossRef]

Black, Fischer, and Myron S. Scholes. 1973. The Pricing of Options and Corporate Liabilities. *Journal of Political Economy* 81: 637–54. [CrossRef]

Cox, John C., and Stephen A. Ross. 1976. The valuation of options for alternative stochastic processes. *Journal of Financial Economics* 3: 145–66. [CrossRef]

Dupire, Bruno. 1994. Pricing with a smile. *Risk* 7: 18–20.

Fabozzi, Frank J., and Steven V. Mann. 2021. *The Handbook of Fixed Income Securities*, 9th ed. New York: McGraw-Hill Education.

Harrison, J. Michael, and Stanley R. Pliska. 1981. Martingales and stochastic integrals in the theory of continuous trading. *Stochastic Processes and their Applications* 11: 215–60. [CrossRef]

Heston, Steven L. 1993. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies* 6: 327–43. [CrossRef]

Jaynes, Edwin T. 1957a. Information theory and statistical mechanics. *Physical Review* 106: 620–30. [CrossRef]

Jaynes, Edwin T. 1957b. Information theory and statistical mechanics. ii. *Physical Review* 108: 171–90. [CrossRef]

Jaynes, Edwin T. 2002. *Probablity Theory: The Logic of Science*. Cambridge: Cambridge University Press.

Lintner, John. 1965. The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *Review of Economics and Statistics* 47: 13–37. [CrossRef]

Merton, Robert C. 1973. Theory of Rational Option Pricing. *Bell Journal of Economics* 4: 141–83. [CrossRef]

Merton, Robert C. 1976. Option Pricing When Underlying Stock Returns are Discontinuous. *Journal of Financial Economics* 3: 125–44. [CrossRef]

Shannon, Claude. 1948. A mathematical theory of communication. *The Bell Systems Technical Journal* 27: 379–423. [CrossRef]

Sharpe, William F. 1964. Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance* 19: 425–42.