Financial Derivatives

18.095: IAP Mathematics Lecture Series 2000
Notes for Lecture 5

Martin Z. Bazant

Figure 1: Payoffs $y(x)$ for derivatives on an underlying asset $x$ with strike price $K$. (a) Short and long positions in a forward contract. (b) Long positions in three different kinds of option contracts.

1 Introduction

A derivative security is a financial instrument whose value is “derived” from the value of one or more underlying assets, which could be commodities (e.g. pork bellies), stocks (e.g. shares of AOL), foreign exchange rates (e.g. the $/yen buying rate), interest rates (e.g. the US prime lending rate) or more exotic variables (e.g. the average snowfall in Aspen, CO) [1]. The simplest derivative is a forward contract (or its publicly traded alter-ego, a futures contract) on a stock or commodity, which gives the holder of the long position (the buyer of the contract) the right to purchase the underlying from the holder of the short position (the seller of the contract) for a fixed strike price at a fixed time in the future (the maturity of the contract). Clearly, as shown in Fig. 1(a), the value of the long position in a forward contract $y(x)$ at maturity (ignoring transaction costs, taxes, etc.) is

$$y(x) = x - K \quad \text{(long forward)}$$  (1)
where $x$ is the value of the underlying and $K$ is the strike price\footnote{The value of the short position is $K - x$, right?}.

An **option** is a more sophisticated, harder to price, version of a forward contract, in which the holder may choose whether or not to exercise the contract\footnote{A **European option** can only be exercised at maturity, but an **American option** can be exercised (at the same strike price) at any earlier time as well. In this lecture, we only consider European options, which are much easier to price.}. Of course, the contract will be exercised only if it is profitable to do so. Hence, at maturity the value of a **call option**, which is the right (but not the obligation) to buy the underlying $x$ at the strike price $K$, is

$$y(x) = \max\{x - K, 0\} \quad \text{(long call)}$$

as shown in Fig. 1(b), since the option will not be exercised if $x < K$. A **put option** is the right to sell the underlying $x$ at price $K$ at maturity\footnote{Do you see that a long forward is equivalent to a long call plus a short put? This is an example of **put-call parity**.}. There are many other kinds of options, such as the **digital option** which is a pure bet, i.e. the short pays the long a fixed amount if $x > K$ and vice versa if $x < K$. Many complex derivative contracts can be expressed as linear combinations of put, call, and digital option positions\footnote{See Problem 1.}

The difficulty in pricing derivatives is that they involve risk, which is also their **raison d’être** \cite{2}. Derivative securities exist so that investors can either **hedge**, i.e. reduce their exposure to a certain risk, or **speculate**, i.e. bet on expected outcomes. For example, a shareholder of a stock worth $100 could purchase a put option struck at $50 to insure against losing more than half of her money. Conversely, a speculator who expects the stock to rise could buy a call option struck at $100. In both cases, the crucial question is: What is a fair price for the derivative security?

The answer to this question depends on the statistics of the underlying asset, which are generally not known. Nevertheless, the “rational” approach to derivative pricing, pioneered by Bachelier a century ago \cite{3}, is to assume a certain probability law for the underlying (perhaps reflecting historical data or personal intuition) and determine from it a fair price on which all investors would agree, given the same assumptions. Although the rational-pricing concept is ubiquitous today, the definition of “fair price” has evolved considerably since 1900. In this lecture, we will compare and contrast three major theories of rational option pricing due to Bachelier \cite{3}, Black-Scholes-Merton \cite{4, 5, 6}, and Bouchaud-Sornette \cite{7}, in chronological order.

## 2 Pricing by a Fair-Game Argument

It is common to assume that the underlying asset follows a “Markov process”, which means that the **transition probability** $p(x, t|t_o, t_o)$ that the value is $x$ at time $t$, given that it is $x_o$ at time $t_o < t$, depends only on $(x, t, x_o, t_o)$ and not upon other historical data\footnote{In this lecture, we will only consider a single time interval $(t_o, t)$ and hence adopt the simpler notation $p(x|x_o)$.}. Consider a derivative with payoff $y(x)$. At first it seems very reasonable that the price $w_B(x_o)$ at time $t_o$
be given by the expected payoff at maturity\textsuperscript{6}

\[ w_B(x_o) = \langle y(x) \rangle \]  

(3)

where

\[ \langle y(x) \rangle \equiv \int y(x)p(x|x_o)dx \]  

(4)

for the case of continuous outcomes in the interval \((a, b)\) or

\[ \langle y \rangle \equiv \frac{1}{N} \sum_{i=1}^{N} p_i y_i \]  

(5)

for the multinomial\textsuperscript{7} case of \( N \) discrete outcomes \( x_i \). This “fair-game” approach, which amounts to requiring that there should be no expected profit from a fairly priced options position, was first proposed in 1900 by Bachelier [3]. For example, the Bachelier price of a call option \( c_B(x_o) \) in the continuous case is given by

\[ c_B(x_o) = \int_{K}^{\infty} (x - K)p(x|x_o)dx. \]  

(6)

An important feature of Bachelier’s theory is that it involves residual risk: The option contract can lead to financial gain or loss for either position. A convenient measure of the residual risk \( R_B^* \) is given by the variance of the payoff

\[ R_B^2 = \langle (y - \langle y \rangle)^2 \rangle = \langle y^2 \rangle - \langle y \rangle^2 = \sigma_y^2 \]  

(7)

where \( \sigma_y \) is the volatility (or standard deviation) of the payoff.

Bachelier’s theory also has a variational formulation. Suppose we define residual risk for an arbitrary option price \( w(x_o) \) as the standard deviation \( R_B \) of the payoff

\[ R_B^2 = \langle (y - w)^2 \rangle \]  

(8)

which is a quadratic function of \( w \). The Bachelier price \( w_B \) is obtained by setting \( dR_B^2/dw = 0 \), which yields the minimum residual risk \( R_B^* \).

The geometrical interpretation of (3) is easiest to see in the multinomial case (5). The Bachelier price \( w_B \) is a weighted average of the possible payoffs \( y_i \), with weights given by the transition probabilities \( p_i \). This corresponds to a “best fit” horizontal line going through the \((x_i, y_i)\) point, which is optimal in the “least-squares” sense of minimal variance \( \sigma_y^2 \).

As illustrated in Fig. 2, this means that the Bachelier price is sensitive to properties of the underlying probability distribution, such as its expected return \( \langle x \rangle - x_o \). This makes sense. For example, an investor who expects the underlying value to increase at maturity would pay more for an at-the-money \((x_o = K)\) call option, and take more risk, than an investor who expects the underlying value to decrease. This dependence of the derivative price on the expected return of the underlying is, however, (remarkably) not true when riskless hedging is possible.

\textsuperscript{6}In this lecture, we ignore interest-rate effects, taxes, dividends, transaction costs, etc., which are very important in practice, but not fundamental to the theoretical pricing problem.

\textsuperscript{7}We adopt the shorthand notation \( p_i \equiv p(x_i|x_o) \) and \( y_i = y(x_i) \).
3 Pricing with Riskless Hedging

Although it seems quite reasonable, the Bachelier price is generally not “fair”. To see this, consider the case of a forward or futures contract. The Bachelier price is

$$f_B = \langle x - K \rangle = \langle x \rangle - K$$

(9)

with residual risk

$$R_B^2 = \langle (x - \langle x \rangle)^2 \rangle = \sigma_x^2$$

(10)

where $\sigma_x$ is the volatility of the underlying asset. The problem with Bachelor’s theory is that it neglects the possibility of hedging (i.e. reducing) risk by cleverly trading the underlying asset along with the derivative. For example, an investor who is short a pork-belly future (with any strike price or maturity) can perfectly hedge her risk by immediately buying a pig. In that case, she would pay the initial price $x_o$ for the pig and then would possess a pork belly at maturity to deliver for a price $K$, no matter what happens to the market price $x$ of pork bellies! In other words, by constructing a riskless hedge, the investor will receive a guaranteed payoff of $K - x_o$ at maturity. Therefore, the fair price $f_f(x_o)$ for the long position in the forward contract is uniquely given by

$$f_f = x_o - K.$$  

(11)

In this simple case, the Bachelier price is fair if and only if the expected return is zero, $\langle x \rangle = x_o$.

The preceding analysis suggests that we consider the possibility of hedging a long options position by selling $\phi$ units of the underlying, for a net initial investment

$$u_o = w - \phi x_o.$$  

(12)

---

8The existence of a unique price should make you wonder why futures exchanges exist at all, since the value of a futures position appears to be known with certainty! However, this is not true when interest-rate fluctuations, which affect the value of a future cash flow, are considered [1].
Assuming a **static hedge** (fixed ϕ), the value of the net investment at maturity is

\[ u = y(x) - \phi x. \]  

(13)

The fair-game criterion now requires

\[ u_o = \langle u \rangle = \langle y \rangle - \phi \langle x \rangle \]  

(14)

in which case the price of the derivative is

\[ w(x_o) = \langle y \rangle + \phi (x_o - \langle x \rangle), \]  

(15)

which is a **linear** function of the underlying-asset value \( x_o \). Note that the Bachelier price \( w_B = \langle y \rangle \) is always fair when the expected return is zero, but otherwise, the pricing problem is reduced to finding an appropriate **hedge ratio** \( \phi \). The correction to the Bachelier price is equal to the expected return on the hedging strategy.

There is another way to view (13), which has profound practical consequences: Rearranging the terms, we find that the derivative position is equivalent (albeit, perhaps only in some statistical sense) to a **replicating portfolio** consisting of an amount \( u_o \) in cash and an amount \( \phi \) of the underlying asset

\[ w = u_o + \phi x_o. \]  

(16)

Whenever a riskless hedge is possible, this equivalence is exact, and there is no difference between the derivative position and the replicating portfolio! In other words, in such cases the derivative is not an independent financial instrument, a situation which economists refer to as a **complete market**.

A riskless hedge is always possible for a forward contract because the payoff function is linear\(^9\). In general, however, \( u \) is a random variable dependent on the random variable \( x \) through (13), and the real market is of course **incomplete**. Nevertheless, for certain special probability laws, a riskless hedge is still possible such that \( u \) is non-random, for any choice of \( y(x) \), which implies a unique fair price for the derivative security.

This powerful insight was first made and applied to options pricing by Black and Scholes [4] and Merton [5] in 1973. Their ideas have had a profound impact on financial theory and practice, which recently led to their sharing a Nobel Prize in Economics in 1997. The Black-Scholes-Merton theory involves **dynamic hedging** in continuous time with “diffusive” underlying dynamics (see below). A much simpler, but equivalent version of the same ideas, more in the spirit of this lecture, was later given by Cox, Ross and Ingersoll [6] in 1979, who considered a static hedge with binomial probabilities.

In the **binomial model**, the underlying can assume only two possible values \( x_1 \) and \( x_2 \) at maturity, with probabilities \( p_1 \) and \( p_2 = 1 - p_1 \), respectively. At maturity, the hedged portfolio (13) has value

\[ u = \begin{cases} 
  y_1 - \phi x_1 & \text{with probability } p_1 \\
  y_2 - \phi x_2 & \text{with probability } p_2 
\end{cases} \]  

(17)

Note that the two random outcomes can be equated by choosing \( \phi \) to be the **riskless hedge ratio** given by

\[ \phi = \frac{y_2 - y_1}{x_2 - x_1}. \]  

(18)

---

\(^9\)See problem 2.
Since the final value \( u \) of the portfolio is nonrandom, it must be equal to the initial value \( u_o \) (ignoring interest-rate effects) by the fair-game argument\(^\text{10}\). Therefore, from (18) the unique fair price of the option \( w_f(x_o) \) is

\[
w_f(x_o) = u_o + \phi x_o = u + \phi x_o = y_1 + \phi(x_o - x_1) = \frac{(x_2 - x_o)y_1 + (x_o - x_1)y_2}{x_2 - x_1}.
\]  

(19)

Geometrically, the fair price simply corresponds to a linear interpolation between the two points \( \{(x_i, y_i)\} \), as shown in Fig. 3. This result also corresponds to a risk-neutral valuation\(^\text{11}\): \( w_f \) is equal to the average payoff when \( p_1 \) is adjusted so that \( \langle x \rangle = x_o \). Note that the Bachelier price with the binomial model depends on \( p_1 \), while the fair price (19) does not, except in the risk-neutral case \( \langle x \rangle = x_o \).

![Figure 3: The Bachelier (gray circle, dashed line) and risk-neutral (gray square, solid line) prices in the binomial model with a nonzero expected return \( \langle x \rangle \neq x_o \).](image)

The Black-Scholes-Merton dynamic-hedging model can be obtained by successive applications of (19) on a binary tree of equal spacing \( \delta x \) in underlying price and \( \delta \tau \) in time. If we let \( w_i^{(n)} \) denote the option price at “backward time” \( \tau_n = t - n\delta \tau \) and underlying value \( x_i = i\delta x \), then the risk-neutral pricing formula (19) applied to each “branch” or the tree takes the form

\[
w_i^{(n+1)} = \frac{1}{2}(w_{i+1}^{(n)} + w_{i-1}^{(n)})
\]  

(20)

which is to be solved recursively (backward in time) starting with the payoff function at maturity \( w_i^{(0)} = y_i \). Rewriting (20) as

\[
\frac{w_i^{(n+1)} - w_i^{(n)}}{\delta \tau} = \frac{\delta x^2}{2\delta \tau} \cdot \frac{w_{i+1}^{(n)} - 2w_i^{(n)} + w_{i-1}^{(n)}}{\delta x^2}
\]  

(21)

and taking the continuum limit \( \delta x \to 0, \delta \tau \to 0 \) with \( \delta x^2/\delta \tau = \sigma^2/2 \) held constant, we obtain the celebrated Black-Scholes equation\(^\text{12}\)

\[
\frac{\partial w}{\partial \tau} = \frac{\sigma^2 \partial^2 w}{2 \partial x^2}
\]  

(22)

---

\(^{10}\)In economic terms, this is amounts to requiring that there be no arbitrage opportunities, in which an investor could make a nonzero, riskless return.

\(^{11}\)See Problem 3.

\(^{12}\)This is the case of zero interest rate \( r = 0 \), and “normal” \( (\alpha = 0) \) rather than “lognormal” \( (\alpha = 1) \).
with initial (or rather, final) condition \( w(x, \tau = 0) = y(x) \). You may recognize (22) as the classical **diffusion equation** from physics. The “Black-Scholes miracle” is that if the underlying obeys a classical diffusion process (a “random walk” composed of equal binomial steps), then a riskless hedge can be constructed, as in the discrete binomial case, and hence a **unique** fair price can be calculated by solving (22). Note that as in the binomial case the Black-Scholes price does not depend on the expected rate of return of the underlying asset (only on the volatility rate \( \sigma \)). Taking the continuum limit of (18), we find that at a riskless hedge can be constructed dynamically by selling an amount

\[
\phi = \frac{\partial w}{\partial x}
\]

of the underlying asset for every unit held of the derivative.

Although risk-neutral valuation and the Black-Scholes equation were immensely successful in the 1970s and 1980s, it became clear after the stock market crash of 1987 that theory has serious trouble with **rare events**, which do not fit into the Gaussian diffusion (or binomial tree) framework. In fact, it has been determined that excessive confidence in the Black-Scholes-Merton theory at least enhanced the amplitude of the crash, if not caused it [1, 2]. Specifically, the culprit seems to have been equation (23), which led nervous investors to quickly sell falling stocks in the hope of replicating protective put options via the Black-Scholes-Merton “riskless” hedging strategy. Unfortunately, when the statistical assumptions of the theory began to break down, this strategy of **portfolio insurance** may have had a nonlinear feedback with the market that actually caused stock prices to plummet even faster!

## 4 Pricing and Hedging with Residual Risk

Although it is still a subject of controversy in theoretical circles (but not on Wall Street!), the Black-Scholes-Merton theory is missing the rather important element of pricing, which is present in the original Bachelier theory: namely, **residual risk**. Riskless hedges only exist in certain special cases (which albeit, happen to be remarkably robust\(^{13}\)). A more general pricing theory must somehow deal with residual risk. Unfortunately, in such a theory, the option price is no longer unique (see below).

In 1994, Bouchaud and Sornette [7] (following several others) devised a pricing theory for **risky options** that generalizes the Black-Scholes-Merton theory by finding the **optimal hedge ratio** \( \phi^* \) that minimizes, but not necessarily eliminates, risk. Here we apply the Bouchaud-Sornette ideas to a single static hedge. A convenient definition of risk is the variance of the hedged position (13):

\[
R^2 = \sigma^2_a = \langle (u - \langle u \rangle)^2 \rangle.
\]

\(^{13}\)Thanks in part to the Central Limit Theorem, which explains why simple diffusion processes are so common.
where \( u = y - \phi x \). Since \( R \) is a quadratic function of \( \phi \),

\[
R^2 = \sigma_y^2 - 2\phi(\langle xy \rangle - \langle x \rangle \langle y \rangle) + \phi^2 \sigma_x^2
\]  

(25)

this definition is sometimes called **quadratic risk**. Setting \( dR^2 / d\phi = 0 \) yields the optimal hedge ratio

\[
\phi^* = \frac{\langle xy \rangle - \langle x \rangle \langle y \rangle}{\langle x^2 \rangle - \langle x \rangle^2}
\]  

(26)

in which case the residual quadratic risk \( R^* \) can be written as

\[
R^* = \sigma_y^2 - \phi^* \sigma_x^2 = R^2 - \phi^* \sigma_x^2.
\]  

(27)

The appropriate investment in the hedged position \( u^*_o \) is again determined by the fair-game argument

\[
u^*_o = \langle u \rangle = \langle y \rangle - \phi^* \langle x \rangle = \frac{\langle x^2 \rangle \langle y \rangle - \langle x \rangle \langle xy \rangle}{\langle x^2 \rangle - \langle x \rangle^2}
\]  

(28)

which yields the Bouchaud-Sornette price

\[
w^*(x_o) = u^*_o(x_o) + \phi^*(x_o)x_o.
\]  

(29)

Note that \( u^*_o \) could also be determined by minimizing \( R \) with respect to \( u_o = \langle u \rangle \), i.e. setting \( dR^2 / du_o = 0 \).

Figure 4: The Bachelier price (gray circle, dashed line) and Bouchaud-Sornette price (gray square, solid line) for an at-the-money call option in a multinomial model. The price and residual risk are shown for cases of (a) negative and (b) positive expected returns.

There is an interesting probabilistic interpretation of the Bouchaud-Sornette theory. In terms of the **correlation coefficient** of \( x \) and \( y \)

\[
\rho_{xy} = \frac{\langle (x - \langle x \rangle)(y - \langle y \rangle) \rangle}{\sigma_x \sigma_y}
\]  

(30)

the optimal hedge ratio can be expressed as

\[
\phi^* = \frac{\rho_{xy} \sigma_y}{\sigma_x}
\]  

(31)
and the residual risk as
\[ R^2 = \sigma_y^2 (1 - \rho_{xy})^2. \] (32)

Now it is clear that what is missing in Bachelier’s theory is the effect of correlations between the underlying asset and the derivative price. It is precisely such correlations that make hedging possible. The quality of the hedge, measured by a reduced \( R^2 \), improves as \( \rho_{xy} \) increases from 0 (the Bachelier limit, \( R^* = \sigma_y \)) to 1 (the Black-Scholes-Merton limit, \( R^* = 0 \)). Whenever the payoff function is linear (e.g. for a forward contract), a riskless hedge exists, regardless of the transition probability law, because in that trivial case the derivative and the underlying-asset values are perfectly correlated at maturity. For any nonlinear payoff function, however, there are nontrivial correlations \( 0 < \rho_{xy} < 1 \), which cause the risk to be nonzero and the fair price to depend on the full transition probability law\(^{14}\).

There is also a nice geometrical interpretation of the Bouchaud-Sornette theory. Rewriting (24) as
\[ R^2 = \langle [y - (u_o + \phi x)]^2 \rangle \] (33)
we see that \( u_o^* \) and \( \phi^* \), and hence the fair price \( w^*(x_o) \), are determined by a weighted linear regression of the payoff \( y \) on the underlying \( x \) at maturity. In other words, \( u_o^* \) and \( \phi^* \) are chosen to obtain the best fit of the equation
\[ y \approx u_o^* + \phi^* x \] (34)
in the least-squares sense, with the weight of the point \((x, y)\) given by the transition probability \( p(x|x_o) \) of the underlying asset. For students of probability theory, it is perhaps more transparent to display the optimal coefficients in the multinomial case
\[
\phi^* = \frac{N \sum x_i y_i - (\sum x_i)(\sum y_i)}{N \sum x_i^2 - (\sum x_i)^2} \\
u_o^* = \frac{(\sum x_i^2)(\sum y_i) - (\sum x_i)(\sum x_i y_i)}{N \sum x_i^2 - (\sum x_i)^2} 
\] (35)
(36)

which are the familiar formulae of (weighted) linear regression.

By now it should be clear that changing the definition of risk \( R \) simply amounts to changing the type of linear regression (from least-squares to something else). For example, we could determine the linear coefficients \( u_o^* \) and \( \phi^* \) which minimize the “absolute-value risk”
\[ R' = \langle |u - \langle u \rangle| \rangle = \langle |y - (u_o + \phi x)| \rangle \] (37)
rather than the quadratic risk\(^{15}\). In any case, we conclude that options pricing simply boils down to linear fitting.

Our conclusions about pricing options in the presence of residual risk are clearly valid for the case of a static hedge (one trading interval), but it is natural to ask what happens when dynamic hedging is allowed. If hedging is allowed only at discrete moments in time, then

---

\(^{14}\) The special case of a binomial transition probability (Cox, Ross Rubinstein \cite{6}) does not really violate our general conclusion because a nonlinear payoff function sampled at only two points is effectively linear.

\(^{15}\) In this case, if the transition probability were also uniform over a certain interval, then the fair price would be given by an “equal-area” fit of the payoff function. Do you see why? If so, try problem 4.
repeated applications of the static-hedge formulae backward in time starting with the known payoff function \( y(x) \) at maturity will give the price \( w^*(x, t) \), optimal hedge ratio \( \phi^*(x, t) \) and optimal cash investment \( u^*_0(x, t) \) at any earlier time \( t \). For the \( n \)th backward-time increment \( \delta \tau \), the recursion is,

\[
w^{(n+1)}(x) = w^{(n)}(x)
\]

(38)

analogous to (20), where \( w^{(n)}(x) \) is a linear fit of the local option value at backward time \( \tau_n \) (weighted by the transition probabilities as above)

\[
w^{(n)}(x) = \langle w^{(n)} \rangle + \phi^{(n)}(x - \langle x \rangle).
\]

(39)

This recursion can be solved numerically for any choice of transition probability function(s) and measure(s) of risk. Analytically, the Bouchaud-Sornette formalism can be used to derive perturbative corrections to the Black-Scholes theory which account for residual risk when fluctuations are nearly diffusive\(^{16}\).

Since there is generally nonzero residual risk, the price of an option is not unique, even if everyone agrees upon the transition probability and the appropriate measure of risk. The emerging theory of pricing risky options (or more precisely, real options!) involves many features contrary to the classical Black-Scholes-Merton paradigm [2]:

- A riskless hedge typically does not exist, and neither does a unique “rational” option price. (It is typically not possible to find a unique line passing through more than two data points.)

- Option prices typically depend the expected return and other aspects of the underlying probability law. (The result of linear fitting depends on the weights of the data points.)

- Option prices and appropriate hedging strategies depend on the market’s perception of residual risk. (The result of linear fitting depends on how one measures “goodness of fit”.)

Dealing with these issues is a major ongoing challenge to theoretical modeling [8] which has fueled an exodus of physicists and applied mathematicians to Wall Street in the past decade [9].

### References


\(^{16}\)See the Appendix for details, which are beyond the scope of this introductory lecture. Similar ideas are developed in Refs. [2] and [7].


