Mathematical Finance with Heavy-Tailed Distributions

M. Vidyasagar

Cecil & Ida Green Professor
The University of Texas at Dallas
M.Vidyasagar@utdallas.edu

MTNS 2010, Hungary

Outline

Outline

Some Terminology

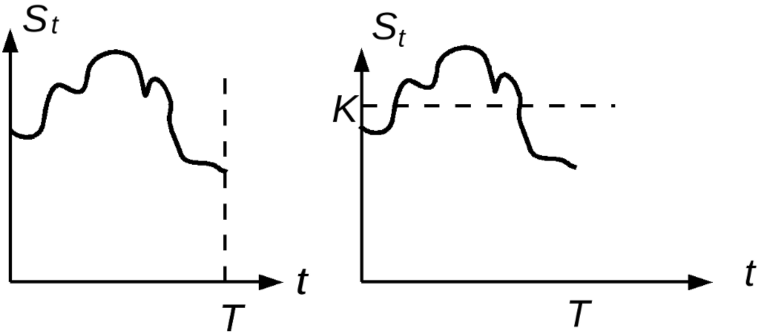
A 'market consists of a 'safe' asset called the 'bond', plus n 'uncertain' assets called 'stocks'.

A 'portfolio' is a set of holdings of the bond and the stocks. It is a vector in $\mathbb{R}^{n+1}$. Negative 'holdings' correspond to borrowing money or 'shorting' stocks.

A 'call option' is an instrument that gives the buyer the right, but not the obligation, to buy a stock a prespecified price called the 'strike price' K. (A 'put' option gives the right to *sell* at a strike price.)

A 'European' option can be exercised only *at* a specified time T. An 'American' option can be exercised *at any time prior to* a specified time T.

European vs. American Options



The value of the European option is $\{S_T - K\}_+$. In this case it is worthless even though $S_t > K$ for some intermediate times. The American option has positive value at intermediate times but is worthless at time $t = T$.

The Questions Studied Here

- What is the minimum price that the seller of an option should be willing to accept?
- What is the maximum price that the buyer of an option should be willing to pay?
- How can the seller (or buyer) of an option 'hedge' (minimimze or even eliminate) his risk after having sold (or bought) the option?

Outline

One-Period, One-Stock Model

Many key ideas can be illustrated via 'one-period, one-stock' model.

We have a choice of investing in a 'safe' bond or an 'uncertain' stock.

$B(0) =$ Price of the bond at time $T = 0$. It increases to $B(1) = (1 + r)B(0)$ at time $T = 1$.

$S(0) =$ Price of the stock at time $T = 0$.

$$S(1) = \begin{cases} S(0)u & \text{with probability } p, \\ S(0)d & \text{with probability } 1 - p. \end{cases}$$

Assumption: $d < 1 + r < u$; otherwise problem is meaningless!

Rewrite as $d' < 1 < u'$, where $d' = d/(1 + r)$, $u' = u/(1 + r)$.

Options and Contingent Claims

An 'option' gives the buyer the right, but not the obligation, to buy the stock at time $T = 1$ at a predetermined strike price K. Again, assume $S(0)d < K < S(0)u$.

More generally, a 'contingent claim' is a random variable X such that

$$X = \begin{cases} X_u & \text{if } S(1) = S(0)u, \\ X_d & \text{if } S(1) = S(0)d. \end{cases}$$

To get an option, set $X = \{S(1) - K\}_+$. Such instruments are called 'derivatives' because their value is 'derived' from that of an 'underlying' asset (in this case a stock).

Question: How much should the seller of such a claim charge for the claim at time $T = 0$?

An Incorrect Intuition

View the value of the claim as a random variable.

$$X = \begin{cases} X_u & \text{with probability } p_u = p, \\ X_d & \text{with probability } p_d = 1 - p. \end{cases}$$

So

$$(1+r)^{-1} E[X, \mathbf{p}] = (1+r)^{-1} [p X_u + (1-p) X_d].$$

Is this the 'right' price for the contingent claim?

NO! The seller of the claim can 'hedge' against future fluctuations of stock price by using a part of the proceeds to buy the stock himself.

The Replicating Portfolio

Build a portfolio at time $T = 0$ such that its value *exactly matches* that of the claim at time $T = 1$ *irrespective* of stock price movement.

Choose real numbers a and b (investment in stocks and bonds respectively) such that

$$aS(0)u + bB(0)(1 + r) = X_u,$$
$$aS(0)d + bB(0)(1 + r) = X_d,$$

or in vector-matrix notation

$$[a \;\; b](1 + r) \begin{bmatrix} S(0)u' & S(0)d' \\ B(0) & B(0) \end{bmatrix} = [X_u \;\; X_d].$$

This is called a 'replicating portfolio', and there is a unique solution for a, b if $u' \neq d'$.

Cost of the Replicating Portfolio

The unique solution for a, b is

$$[a \;\; b] = (1+r)^{-1}[X_u \;\; X_d] \begin{bmatrix} u' & d' \\ 1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1/S(0) & 0 \\ 0 & 1/B(0) \end{bmatrix}.$$

Amount of money needed at time $T = 0$ to implement the replicating strategy is

$$c = [a \;\; b] \begin{bmatrix} S(0) \\ B(0) \end{bmatrix} = (1+r)^{-1}[X_u \;\; X_d] \begin{bmatrix} q_u \\ q_d \end{bmatrix},$$

where with $u' = u/(1+r), d' = d/(1+r)$, we have

$$\mathbf{q} := \begin{bmatrix} q_u \\ q_d \end{bmatrix} = \begin{bmatrix} u' & d' \\ 1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1-d'}{u'-d'} \\ \frac{u'-1}{u'-d'} \end{bmatrix}$$

Martingale Measure: First Glimpse

Note that $\mathbf{q} := (q_u, q_d)$ is a probability distribution on $S(1)$. Moreover it is the *unique distribution* such that

$$E[(1+r)^{-1}S(1), \mathbf{q}] = S(0)u' \frac{1-d'}{u'-d'} + S(0)d' \frac{u'-1}{u'-d'} = S(0),$$

i.e. such that $\{S(0), (1+r)^{-1}S(1)\}$ is a 'martingale' under $\mathbf{q}$.

Important point: $\mathbf{q}$ depends *only* on the returns u, d, and not on the associated 'real world' probabilities $p, 1-p$.

Thus the initial cost of the replicating portfolio

$$c = (1+r)^{-1}[X_u \;\; X_d] \begin{bmatrix} q_u \\ q_d \end{bmatrix}$$

is the *discounted expected value* of the contingent claim X under the (unique) martingale measure $\mathbf{q}$.

Arbitrage-Free Price of a Claim

Theorem: The quantity

$$c = (1+r)^{-1}[X_u \;\; X_d] \begin{bmatrix} q_u \\ q_d \end{bmatrix}$$

is the *unique arbitrage-free price* for the contingent claim.

Suppose someone is ready to pay $c' > c$ for the claim. Then the seller collects c', invests $c' - c$ in a risk-free bond, uses c to implement replicating strategy and settle claim at time $T = 1$, and pockets a risk-free profit of $(1+r)(c' - c)$. This is called an 'arbitrage opportunity'.

Suppose someone is ready to sell the claim for $c' < c$. Then the *buyer* can make a risk-free profit.

Summary

Fact 1. There is a unique 'synthetic' distribution $\mathbf{q}$ on $S(1)$ that makes the process $\{S(0), (1+r)^{-1}S(1)\}$ into a martingale. This distribution $\mathbf{q}$ depends *only* on the two possible outcomes, but *not* on the associated 'real world' probabilities.

Fact 2. The *unique arbitrage-free price* of a contingent claim (X_u, X_d) is the discounted expected value of the claim under $\mathbf{q}$.

Some Interesting Observations

Replicating portfolio is given by

$$[a \;\; b] = (1+r)^{-1}[X_u \;\; X_d] \begin{bmatrix} u' & d' \\ 1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1/S(0) & 0 \\ 0 & 1/B(0) \end{bmatrix}$$

$$= \frac{1}{u-d}[X_u - X_d \quad -d'X_u + u'X_d] \begin{bmatrix} 1/S(0) & 0 \\ 0 & 1/B(0) \end{bmatrix}.$$

Consider a call option with strike price

$$K = S(0)w = S(0)(1+r)w',$$

where $d' < w' < u'$. Then

$$X_u = (1+r)(u'-w')S(0), X_d = 0.$$

Moreover, b *is negative!* What does this mean?

Overhedging

In the case of a call option with strike price $K = S(0)w = S(0)(1+r)w'$, $X_u = (1+r)(u'-w')S(0), X_d = 0$, the initial share holding in the replicating portfolio is

$$aS(0) = \frac{u'-w'}{u'-d'}, c = (1-d')\frac{u'-w'}{u'-d'} < aS(0)!$$

So the seller of the option *needs to borrow money to hedge, because the price of the option is lower than the money needed to procure the shares!*

A Numerical Example

Take $r = 0, S(0) = B(0) = 1$ – everything is normalized and in discounted currency.

Say $u = 1.8, d = 0.8, w = 1.5$ Then

$$[q_u \;\; q_d] = [0.2 \;\; 0.8], c = 0.06, [a \;\; b] = [0.30 \;\; -0.24].$$

The seller of a call option against 100 shares gets \$ 6, but needs to hedge by herself procuring *30 shares!*. So she needs to borrow \$ 24 to implement this replicating strategy. This is called 'overhedging' and can be shown to hold under very general conditions.

Multiple Periods: Binomial Model

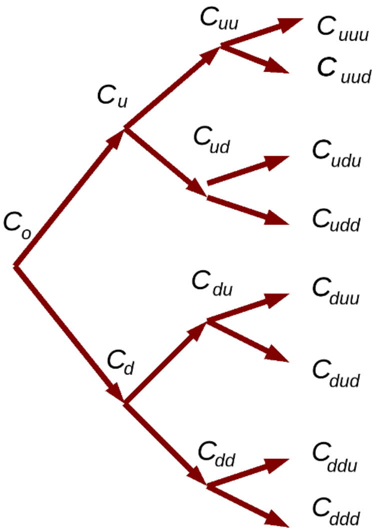
Bond price is deterministic:

$$B_{n+1} = (1 + r_n)B_n, n = 0, \ldots, N - 1.$$

Stock price can go up or down: $S_{n+1} = S_n u_n$ or $S_n d_n$.

There are 2^N possible sample paths for the stock, corresponding to each $\mathbf{h} \in \{u, d\}^N$. For each sample path $\mathbf{h}$, at time N there is a payout $X_{\mathbf{h}}$ due at the end (European claim).

We already know to replicate over one period. Extend argument to N periods. This is called the **binomial model**.


C_o
C_u
C_d
C_{uu}
C_{ud}
C_{du}
C_{dd}
C_{uuu}
C_{uud}
C_{udu}
C_{udd}
C_{duu}
C_{dud}
C_{ddu}
C_{ddd}

Features of Pricing & Hedging in Binomial Model

There is a unique 'synthetic' probability distribution $\mathbf{q}$ under which the discounted stock process

$$\left\{ \prod_{i=0}^{n-1} (1 + r_i)^{-1} \mathrm{S}_n \right\}$$

where the empty product is taken as one, is a martingale. This synthetic probability distribution $\mathbf{q}$ depends *only* on the up and down movements at each time instant, and not on the associated real-world probabilities.

The theory is readily extendable to more than one stock.

Some More Features

Important note: This strategy is **self-financing:**

$$a_0 S_1 + b_0 B_1 = a_1 S_1 + b_1 B_1$$

whether $S_1 = S_0 u_0$ or $S_1 = S_0 d_0$ (i.e. whether the stock goes up or down at time $T = 1$). This property has no analog in the one-period case.

It is also *replicating* from that time onwards.

Observe: Implementation of replicating strategy requires reallocation of resources N times, once at each time instant.

Outline

Continuous-Time Processes: Black-Scholes Formula

Take 'limit' at time interval goes to zero and $N \to \infty$; binomial asset price movement becomes geometric Brownian motion:

$$S_t = S_0 \exp \left[\left(\mu - \frac{1}{2} \sigma^2 \right) t + \sigma W_t \right], t \in [0, T],$$

where W_t is a standard Brownian motion process. μ is the 'drift' of the Brownian motion and σ is the volatility.

Bond price is deterministic: $B_t = B_0 e^{rt}$, where r is the risk-free interest rate.

At time T there is a payout $H(S_T)$ depending on the final asset price (European contingent claim). For an option with strike price K, take $H(x) = \{x - K\}_+$.

Black-Scholes Formula

Define

$$\Phi(x) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp(-y^2/2) dy,$$

the distribution function of a gaussian.

Theorem (Black-Scholes 1973): For an option with strike price K and time to mature T, the unique arbitrage-free option price at time t as a function of the current stock price x is given by

$$u(T-t,x) = x\Phi(g(t,x)) - Ke^{-rt}\Phi(h(t,x)),$$

where

$$g(t,x) = \frac{\ln(x/K) + (r + 0.5\sigma^2)t}{\sigma t^{1/2}},$$

$$h(t,x) = g(t,x) - \sigma t^{1/2}.$$

Black-Scholes PDE

For a general contingent claim of $h(S_T)$ at time T, the unique arbitrage-free price of the option at time t is given by $u(t, S_t)$, where

$$\frac{\partial u}{\partial t} + 0.5\sigma^2 x^2 \frac{\partial^2 u}{\partial x^2} + rx\frac{\partial u}{\partial x} = ru, u(x,T) = h(x).$$

Continuous trading strategy is: Choose $a_\tau = (\partial u/\partial x)(\tau, S_\tau)$ where S_τ is the *actual stock price* at time τ, and invest the rest in bonds.

No closed-form solution in general, but if $h(x) = (x-K)_+$ the solution is as given earlier.

Extensions to Multiple Assets

Binomial model extends readily to multiple assets.

Black-Scholes theory extends to the case of multiple assets of the form

$$S_t^{(i)} = S_0^{(i)} \exp \left[\left(\mu^{(i)} - \frac{1}{2} [\sigma^{(i)}]^2 \right) t + \sigma W_t^{(i)} \right], t \in [0, T],$$

where $W_t^{(i)}, i = 1, \ldots, d$ are (possibly correlated) Brownian motions.

Analog of Black-Scholes PDE: Option price equals $f(0, S_0^{(1)}, \ldots, S_0^{(d)})$ where f satisfies a PDE. But no closed-form solution for f in general.

American Contingent Claims

An 'American' contingent claim can be exercised at any time *up to and including* time T.

So we need a 'super-replicating' strategy: The value of our portfolio must *equal or exceed* the value of the claim at all times.

In the case of American options, $X_t = \{S_t - K\}_+$, then both price and hedging strategy are same as for European claims. In particular, *buyer of American option should wait until expiry.*

Very little known about pricing and exercising general American contingent claims. Theory of 'optimal time to exercise option' is very deep and difficult.

Outline

Modeling Errors in Recent Financial Crisis

In my view, modeling errors *per se* contributed very little to recent financial crisis. Nevertheless, highly desirable to get better models for asset price movements.

Can modeling errors in recent financial crisis be explained by the over-use of Black-Scholes theory, particularly the GBM (geometric Brownian motion) model?

Can observed behavior be better explained by modeling asset returns as heavy-tailed random variables (r.v.s)?

For us: $\mathcal{X}$ is 'heavy-tailed' if $\mathcal{X}$ has finite mean but infinite variance.

Historical Perspective

As far back as 1963, Mandelbrot proposed Pareto distribution (a heavy-tailed distribution) as model for cotton prices.

Later on, Taleb and others claimed 'scale-free' property of asset returns: If $\mathcal{X}$ is the daily return of an asset value (e.g. a stock), then $\Pr\{X \geq c\}/\Pr\{X \geq 1.5c\}$ seems to be pretty constant with respect to c.

Recent results in extremal value theory show that the large deviation behavior of heavy-tailed r.v.s is *qualitatively very different* from those with finite variance.

Some of the theoretical predictions seem to tally with observed phenomena in stock prices.

Observable vs. Unobservable Parts of the Universe

If $\mathcal{X}$ is the daily return on an asset (e.g. a stock), asking 'Does $\mathcal{X}$ have infinite variance?' is silly.

If a stock price shows unusual movements, exchange will halt trading! So all 'real' asset returns are *bounded* r.v.s and have finite moments of all orders.

We are extrapolating from 'observable' universe (stock price movements of a few percent daily) to 'unobservable' universe.

GBM was a useful model because it led to closed-form formulae for option prices. But it is demonstrably 'at variance' with observed data.

Infinite variance will be a useful model *only if* it can explain phenomena in the 'observable' universe!

Some Observed Phenomena

Black-Scholes theory assumes that asset prices follow GBM (Geometric Brownian Motion).

'Real' asset returns don't follow GBM description! Daily returns show fatter tails than Gaussian (kurtosis > 3).

More to the point, extreme events take place far too often!
• For daily returns of the Dow Jones industrial average, $\sigma \approx 0.012$ or 1.2%. So 21% decline in DJA in October 1987 was a 20σ event. The 8% selloff in 1989 was a 7σ event.
• On 24 February 2003, the price of natural gas changed by 42% in one day, a 12σ event.

The Gaussian distribution would tell us that such events should take place at most once within the known age of the universe.

Outline

Laws of Large Numbers

Suppose $\{\mathcal{X}_t\}_{t \geq 1}$ is an i.i.d. sequence of random variables with $E[\mathcal{X}_t] =: \mu < \infty$. Consider the cumulative sums and averages

$$S_l := \sum_{t=1}^{l} \mathcal{X}_t, A_l = \frac{1}{l} S_l = \frac{1}{l} \sum_{t=1}^{l} \mathcal{X}_t.$$

Under mild conditions A_l converges to μ in probability. So if we define

$$\delta(l, \epsilon) := \Pr\{A_l \geq \epsilon\} = \Pr\{S_l \geq l\epsilon\},$$

then $\delta(l, \epsilon) \to 0$ as $l \to \infty$, for each $\epsilon > \mu$. Can we be more precise about the tail behavior?

Large Deviation Behavior of Average vs. Maximum

We have already defined

$$\delta(l, \epsilon) := \Pr\{A_l \geq \epsilon\} = \Pr\{S_l \geq l\epsilon\},$$

Now define

$$\gamma(l, \epsilon) := \Pr\{\max\{\mathcal{X}_1, \ldots, \mathcal{X}_l\} \geq l\epsilon\}.$$

How do $\delta(l, \epsilon)$ and $\gamma(l, \epsilon)$ compare?

If $\mathcal{X}_t$ are *nonnegative*, then

$$\gamma(l, \epsilon) \leq \delta(l, \epsilon), \ \forall l, \epsilon.$$

A huge excursion in one r.v. causes the average to be larger than ϵ, but average can exceed ϵ even through several small excursions in each r.v.

Stable Distributions: Background

(Sloppily speaking), the limit distribution of the cumulative average A_l and cumulate sum S_l *must be stable distributions*, whether or not $\mathcal{X}_t$ has finite variance!

Original work by Paul Lévy, A. N. Kolmogorov etc.

Good (available) reference: *Probability* by Leo Breiman.

Great (but out of print) reference: *Limit Distributions of Sums of Independent Random Variables* B. V. Gnedenko and A. N. Kolmogorov.

Stable Distributions: Definition

A distribution Φ_X of an r.v. X is said to be **stable** if, whenever Y, Z are i.i.d. copies of X, for every pair of real numbers a, b, there exist two other real numbers c, d such that $aY + bZ$ is distributionally equal to $cX + d$.

A distribution Φ_X is **strictly stable** if $d = 0$, and p-**strictly stable** if $c = (a^p + b^p)^{1/p}$.

Every Gaussian distribution is stable, whereas every *zero mean* Gaussian distribution is 2-strictly stable.

Stable Distributions: Characterization

The Gaussian is *the only stable distribution with finite variance!*

Every other stable distribution $F(\cdot)$ is either the Cauchy distribution, or else has a characteristic function $\Phi(\cdot)$ of the form $\Phi(u) = \exp(\psi(u))$ where

$$\psi(u) = \mathbf{i}uc - d|u|^\alpha \left(1 + \mathbf{i}\theta \frac{u}{|u|} \tan \frac{\alpha\pi}{2} \right),$$

where $\alpha \in (0,1) \cup (1,2)$, c is real, d is real and positive, and θ is real with $|\theta| \leq 1$. α is called the 'exponent'.

Unfortunately we cannot invert the Fourier transform to get the distribution function in closed form. But we *can* characterize its asymptotic behavior.

Stable Distributions: Tail Behavior

If F is a non-Gaussian stable distribution function, then there exist constants M_1, M_2, not both zero, and constants α_1, α_2 with $\alpha_i \in (0,2)$ if $M_i \neq 0$, such that

$$F(x) \sim L_1(x) x^{-\alpha_1} \text{ as } x \to -\infty \text{ if } M_1 \neq 0,$$

$$F(x) \sim L_2(x) x^{-\alpha_2} \text{ as } x \to \infty \text{ if } M_2 \neq 0,$$

where L_1, L_2 are 'slowly varying functions', i.e.

$$\lim_{x \to \infty} \frac{L_i(tx)}{L_i(x)} = 1.$$

In short, *all non-Gaussian stable distributions exhibit 'scale-free' tail behavior!*

Outline

Tail Behavior of Cumulative Averages

Recall:

$$\delta(l, \epsilon) := \Pr\{A_l \geq \epsilon\} = \Pr\{S_l \geq l\epsilon\},$$

$$\gamma(l, \epsilon) := \Pr\{\max\{\mathcal{X}_1, \ldots, \mathcal{X}_l\} \geq l\epsilon\}.$$

If X_t has finite variance, then central limit theorem applies and $\gamma(l, \epsilon)/\delta(l, \epsilon) \to 0$ as $l \to \infty$.

Theorem: If $\mathcal{X}_t$ is nonnegative and has scale-free tail behavior, then

$$\gamma(l, \epsilon) \sim \delta(l, \epsilon), \text{ i.e. } \lim_{l \to \infty} \frac{\gamma(l, \epsilon)}{\delta(l, \epsilon)} = 1.$$

In words, if we average heavy-tailed r.v.s, then a tail excursion is *just as likely to occur* through a huge excursion of one variable as through several small excursions of several variables.

Observed Behavior of Asset Prices

• Increase in share price of Akamai from November 2001 until now: 12.10, or a return of 1,110%.

$\Rightarrow$ *Just three days* account for 694% of the return!

• Between 1955 and 2004, S&P average moved up by a factor of 180.

$\Rightarrow$ If we remove ten largest movements (most of which were negative), the increase is 350.

So real asset price movements *do* move in a few large bursts! Is *this* a justification for using heavy-tailed r.v.s in modeling?

If so, how can we fit heavy-tailed models to observed data (some work done; typical values of $\alpha \approx 1.6$).

How can we do option-pricing, hedging, risk assessment, etc.? (completely open field).

Summary

We have examined discrete-time markets, continuous-time markets with the GBM (geometric Brownian motion) model, and heavy-tailed asset distributions.

A great many interesting and open problems remain!

Thank You!