Financial Mathematics

Stock models: The binomial model

A. N. Yannacopoulos

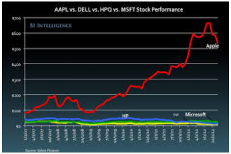
AUEB

Academic year 2022-2023

Stocks

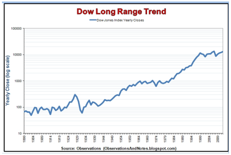
- Stocks are financial assets, traded in organized markets, offering firms capital for development in return to access to future earnings of the firm (e.g. dividends) to stock holders.
- These future earnings are uncertain, and the fact that stock are bought and sold on an everyday basis hence providing an ever changing stochastic demand and supply, leads to stochastic fluctuations for the stock prices.
- Various stochastic models have been proposed to model and predict the stock prices
  - Binomial model
  - Time series models (e.g. Garch)
  - The Black-Scholes model
  - ....

(a) APPL



(b) APPL vs DELL vs HPQ vs MSFT

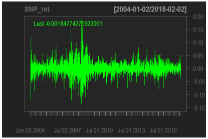
(c) Nokia



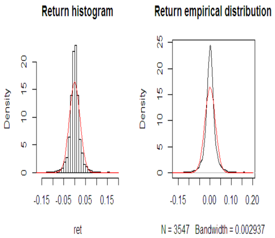
(d) Dow Jones

Figure: Examples of stock prices

(a) SNP prices



(b) Log returns
$R_t := \ln \frac{P_{t+1}}{P_t}$



(c) R_t statistics

Figure: SNP data and statistics

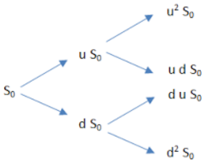
A first model: The binomial model

Keeping it simple (to start with!)

Assumptions

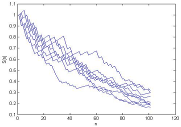
- There is a riskless asset in the market with return r.
- Returns of the risky asset (stock) are independent: What happens between t and $t+1$ stays there!
- Two possible returns each period:

$$R_t = \frac{S_{t+1}}{S_t} = \begin{cases} u & \text{with probability} \quad p \\ d & \text{with probability} \quad 1-p \end{cases}$$

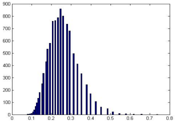


$$S_{n+1} = H_{n+1} S_n, \qquad (1)$$

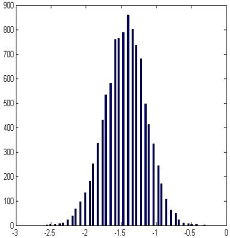
$$\{H_n\} \ i.i.d. \ P(H_n = u) = p, \quad P(H_n = d) = 1-p, \quad n = 1, 2, \cdots$$



(a) Binomial model paths



(b) $R_t := \frac{P_{t+1}}{P_t}$ histogram



(c) $\ln R_t$ histogram

Figure: Binomial model paths and statistics

Information structures for the binomial model

- At any time N the random variable S_N may take the following values

$$\{S_0 u^N, S_0 u^{N-1} d, \cdots, S_0 u d^{N-1}, S_0 d^N\}$$

- $S_N = H_N H_{N-1} \cdots H_1 S_0$ is measurable with respect to the σ-algebra

$$\mathcal{F}_N = \sigma(H_N, H_{N-1}, \cdots, H_1) = \sigma(S_N, S_{N-1}, \cdots, S_1)$$

- $\mathcal{F}_N$ contains the information of the market history up to time N
- Clearly, $\mathcal{F}_1 \subset \mathcal{F}_2 \subset \cdots \subset \mathcal{F}_N \subset \cdots$
- Accumulating increasing information as times goes by: A filtration

Example

Assume that you are interested in the binomial model up to time 2, i.e. you want to describe $S_2 = H_2 H_1 S_0$.

$$\Omega = \{uu, ud, du, dd\}$$

Then,

- $\mathcal{F}_0$ is the information about the market at $t = 0$ (i.e. having observed S_0)

$$\mathcal{F}_0 = \{\emptyset, \Omega\}$$

- $\mathcal{F}_1$ is the information about the market at $t = 1$ (i.e. having observed S_0, S_1)

$$\mathcal{F}_1 = \{\emptyset, \Omega \{uu, ud\}, \{du, dd\}\}$$

- $\mathcal{F}_2$ is the information about the market at $t = 2$ (i.e. having observed S_0, S_1, S_2)

$$\mathcal{F}_2 = \{\emptyset, \Omega \{uu\}, \{ud\}, \{du\}, \{dd\}, \{uu, ud\}, \{du, dd\}, \{uu, ud, du\}, \cdots\}$$

Reminder: σ-algebras and measurability

Definition (σ-algebra)

Let Ω be a set. A σ-algebra $\mathcal{F}$ on Ω is a collection of subsets of Ω such that

1. $\emptyset, \Omega \in \mathcal{F}$.
2. If $A \in \mathcal{F}$ then $A^c \in \mathcal{F}$
3. If $A_1, A_2, \dots \in \mathcal{F}$, then $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$.

Definition (Measurability of a random variable with respect to a σ-algebra)

Let $X : \Omega \to \mathbb{R}$ be a random variable and $\mathcal{F}$ be a σ-algebra on Ω.
X is measurable with respect to $\mathcal{F}$, denoted by $X \in m - \mathcal{F}$ if

$$\forall\, A \subset \mathbb{R} \text{ it holds that } X^{-1}(A) \in \mathcal{F}$$

Measurability of X with respect to $\mathcal{F}$ means that the necessary information to completely describe the random variable X can be found in the σ-algebra (information structure) $\mathcal{F}$.

Conditional expectation

Definition

Let X be a random variable measurable with respect to a σ-algebra $\mathcal{G}$, and $\mathcal{F}$ be a σ-algebra $\mathcal{F} \subset \mathcal{G}$ (i.e. carrying less information)
The conditional expectation of X with respect $\mathcal{F}$, is the best predictor Z for X, that can be determined only using information available in $\mathcal{F}$, i.e. the solution to the problem

$$\mathbb{E}[(X-Z)^2] = \min_{Y \in m-\mathcal{F}} E[(X-Y)^2]$$

We use the notation $Z = \mathbb{E}[X \mid \mathcal{F}]$

It can be shown that Z is the (unique) random variable satifying:

- Z is measurable with respect to $\mathcal{F}$,
- $\int_A Z(\omega)dP(\omega) = \int_A X(\omega)dP(\omega)$ for every $A \in \mathcal{F}$.

Basic properties of conditional expectation

- If $\mathcal{F}_M \subset \mathcal{F}_N$ then $\mathbb{E}[\mathbb{E}[X \mid \mathcal{F}_M] \mid \mathcal{F}_N] = \mathbb{E}[X \mid \mathcal{F}_M]$.
- If $\mathcal{F}_M \subset \mathcal{F}_N$ then $\mathbb{E}[\mathbb{E}[X \mid \mathcal{F}_N] \mid \mathcal{F}_M] = \mathbb{E}[X \mid \mathcal{F}_M]$.
- If $X \in m_{\mathcal{F}}$, then $\mathbb{E}[X \mid \mathcal{F}] = X$
- If X is independent of $\mathcal{F}$ then, $\mathbb{E}[X \mid \mathcal{F}] = \mathbb{E}[X]$

The Markov property for the binomial model

$$\mathbb{E}[S_{n+1} \mid \mathcal{F}_n] = \mathbb{E}[S_n H_{n+1} \mid \mathcal{F}_n] = S_n \mathbb{E}[H_{n+1} \mid \mathcal{F}_n] = S_n \mathbb{E}[H_{n+1}] = S_n (p u + (1-p) d)$$

The best prediction for S_{n+1} given the full history $S_1, \cdots, S_n$, only depends on S_n!

$$\mathbb{E}[S_{n+1} \mid \mathcal{F}_n] = \mathbb{E}[S_{n+1} \mid \sigma(S_1, \cdots, S_n)] = \mathbb{E}[S_{n+1} \mid \sigma(S_n)]$$

The process $\{S_n\}$ "forgets its memory" and is like starting anew on the previous time instance as far as predictions are concerned: Markov property

Long run predictions

Using the law of total probability

$$\mathbb{E}[S_{n+1}] = \mathbb{E}[\mathbb{E}[S_{n+1} \mid \mathcal{F}_n]] = (p\,u + (1-p)\,d)\,\mathbb{E}[S_n]$$

hence by induction

$$\mathbb{E}[S_n] = (p\,u + (1-p)\,d)^n\,S_0$$

Define the discounted price process $S_n^* = (1+r)^{-n}\,S_n$, and work as above to show

$$\mathbb{E}[S_n^*] = \left(\frac{p\,u + (1-p)\,d}{1+r}\right)^n S_0$$

- If $\frac{p\,u+(1-p)\,d}{1+r} > 1$ then the expected value of the asset increases
- If $\frac{p\,u+(1-p)\,d}{1+r} < 1$ then the expected value of the asset decreases

Martingales and asset prices

We observe that

- if $\frac{p\,u+(1-p)\,d}{1+r} > 1$ then

$$\mathbb{E}[S^*_{n+1} \mid \mathcal{F}_n] = \frac{p\,u+(1-p)\,d}{1+r}\,S^*_n \geq S^*_n$$

- if $\frac{p\,u+(1-p)\,d}{1+r} < 1$ then

$$\mathbb{E}[S^*_{n+1} \mid \mathcal{F}_n] = \frac{p\,u+(1-p)\,d}{1+r}\,S^*_n \leq S^*_n$$

- if $\frac{p\,u+(1-p)\,d}{1+r} = 1$ then

$$\mathbb{E}[S^*_{n+1} \mid \mathcal{F}_n] = \frac{p\,u+(1-p)\,d}{1+r}\,S^*_n = S^*_n$$

Martingales

Let $\{X_n\}$ be a stochastic process and $\{\mathcal{F}_n\}$ a filtration ($\mathcal{F}_n \subset \mathcal{F}_{n+1}$)

Definition

$\{X_n\}$ is called adapted to $\{\mathcal{F}_n\}$ if

$$X_n \in m - \mathcal{F}_n, \ \ \forall \ n \in \mathbb{N}.$$

Definition

An adapted process $\{X_n\}$, such that $\mathbb{E}[|X_n|] < \infty$ is called a martingale if

$$\mathbb{E}[X_{n+1} \mid \mathcal{F}_n] = X_n$$

Definition

An adapted process $\{X_n\}$, such that $\mathbb{E}[|X_n|] < \infty$ is called a submartingale if

$$\mathbb{E}[X_{n+1} \mid \mathcal{F}_n] \geq X_n$$

Definition

An adapted process $\{X_n\}$, such that $\mathbb{E}[|X_n|] < \infty$ is called a supermartingale if

$$\mathbb{E}[X_{n+1} \mid \mathcal{F}_n] \leq X_n$$

A martingale is a mathematical model for the earnings of a fair game

Theorem

The following hold for the binomial model:

1. *If* $\frac{p\,u+(1-p)\,d}{1+r} = 1$ *then the discounted price process* $\{S_n^*\}$ *is a martingale*
2. *If* $\frac{p\,u+(1-p)\,d}{1+r} \geq 1$ *then the discounted price process* $\{S_n^*\}$ *is a submartingale*
3. *If* $\frac{p\,u+(1-p)\,d}{1+r} \leq 1$ *then the discounted price process* $\{S_n^*\}$ *is a supermartingale*

The equivalent martingale measure

One may observe the time series for the asset prices (equiv. return) i.e. the paths of the stochastic processes but it is difficult to place probabilities on the possible values.

Hence, one may consider that u and d are known, but the probability p with which these are distributed.

Keeping the paths of the returns but allowing freedom on the probability of returns, choose $p = q$ such that

$$\frac{q\,u + (1-q)\,d}{1+r} = 1,$$

and let Q be the corresponding probability measure of the paths.

Then, by the above arguments

$$\mathbb{E}_Q[S^*_{n+1} \mid \mathcal{F}_n] = S^*_n \iff \mathbb{E}_Q[\frac{S_{n+1}}{S_n} \mid \mathcal{F}_n] = 1 + r,$$

i.e. under the probability measure Q, the discounted price process $\{S^*_n\}$ is a martingale.

If $0 < q < 1$ then under the new measure, we will still have two possible states for the return of the asset: The model P and the model Q have the same null sets i.e. P and Q are equivalent measures.

Q is called an equivalent martingale measure (or risk neutral measure).

$$0 < q < 1 \iff d < 1 + r < u$$

The condition $d < 1 + r < u$ is equivalent to a no arbitrage condition: If it is not satisfied then an investor may have arbitrage opportunities by placing a portfolio in the riskless and the risky asset.

Theorem (Fundamental theorem of asset pricing)

There are no arbitrage opportunities in the binomial model if and only if there exists an equivalent martingale measure

In fact this theorem is true for any model for a financial market (but the proof is more complicated and requires delicate tools from convex duality).

The binomial model in the limit of $N \to \infty$ and the lognormal distribution

Take logs in the binomial model $S_{n+1} = H_{n+1} S_n$:

$$\ln S_{n+1} - \ln S_n = \ln H_{n+1} \iff \ln S_N - \ln S_0 = \sum_{i=1}^{N} \ln H_i$$

Hence,

$$\ln \frac{S_N}{S_0} = \sum_{i=1}^{N} \ln H_i,$$

which is the sum of N independent random variables.

The central limit theorem can be used to approximate the distribution of this sum

The central limit theorem

Theorem

Let $\{X_n\}$ be a sequence of i.i.d. random variables, with $\mathbb{E}[X_n] = \mu$ and $Var(X_n) = \sigma^2$.
Then,

$$Z_n = \frac{\sum_{i=1}^n X_i - \mu n}{\sigma\sqrt{n}} \overset{n\to\infty}{\to} Z \sim N(0,1),$$

with the convergence being understood as convergence in distribution.

Applying the CLT to the binomial model we conclude that

$$\ln \frac{S_N}{S_0} \sim N(\mu N, \sigma^2 N), \quad \text{for } N \text{ large},$$

where

$$\begin{aligned}
\mu &= \mathbb{E}[\ln H_i] = p \ln u + (1-p) \ln d, \\
\sigma^2 &= Var(\ln H_i) = p(\ln u - \mu)^2 + (1-p)(\ln d - \mu)^2
\end{aligned}$$

Hence, in the long horizon limit, S_N follows the lognormal distribution – compare with the empirical data!

This is interesting information, but for inference processes useless since with this parameterization the variance is huge for N large.

A scaling limit and a calibration procedure

Consider data for a time horizon T (e.g. 1 yr) and try to fit a binomial model of N iterations in that.
From data we may calculate

$$\begin{aligned} \mu &:= \frac{1}{T}\,\mathbb{E}_P\left[\ln \frac{S(T)}{S(0)}\right], \\ \sigma^2 &:= \frac{1}{T}\,Var_P\left[\ln \frac{S(T)}{S(0)}\right] \quad \text{volatility} \end{aligned}$$

If the binomial model is a suitable representation for the stock prices then

$$\begin{aligned} \mu = \mathbb{E}[\ln H_i] = p \ln u + (1-p) \ln d, \\ \sigma^2 = Var(\ln H_i) = p(\ln u - \mu)^2 + (1-p)(\ln d - \mu)^2 \end{aligned}$$

Two equations for 3 unknowns (p, u, d).

Fix two of them by setting $ud = 1$ (recombining model) which yields

$$\begin{aligned} \ln u = -\ln d = \sqrt{\frac{T}{N}\,\sigma^2 + \frac{T^2}{N^2}\,\mu^2} \sim \sigma\sqrt{\frac{T}{N}} \\ p = \frac{1}{2} + \frac{1}{2}\frac{\mu}{\sigma}\left(\frac{N}{T} + \left(\frac{\mu}{\sigma}\right)^2\right)^{-1/2} \end{aligned}$$

The above proposes the following calibration procedure:

1. Collect data for a time horizon T and in terms of the log returns calculate the volatility

$$\sigma^2 := \frac{1}{T} Var_P \left[\ln \frac{S(T)}{S(0)} \right] \quad \text{volatility}$$

   A simple estimator for the volatility can be the sample variance.
2. Break $[0, T]$ into N intervals of length $\Delta t = \frac{T}{N}$ each.
3. Then approximate each $S(t_n)$, for $t_n = n\Delta t$ by the binomial model $S_{n+1} = H_{n+1} S_n$ with parameters

$$u = d^{-1} = \exp(\sigma\sqrt{\Delta T}).$$

Remark

We can calculate p as well e.g. through the approximation

$$p = \frac{1}{2} + \frac{1}{2}\frac{\mu}{\sigma}\sqrt{\frac{T}{N}}$$

but this will not be needed if we want to use this binomial model approximation for the data in order to price derivative products (coming up!)

A scaling limit: Take 2

Take $[0, T]$ and break into N intervals of length $\delta t = \frac{T}{N}$, and set $t_n = n\delta t$. Set the parameters (as above)

$$u = d^{-1} = exp(\sigma \sqrt{\delta t}),$$
$$p = \frac{1}{2}\left(1 + \frac{\mu}{\sigma}\sqrt{\delta t}\right)$$

Define the random variables that indicate occurence of high returns

$$X_i = \mathbf{1}_{\{H_i = u\}} = \begin{cases} 1 & H_i = u, \\ 0 & H_i = d \end{cases}$$

The random variables

$$U_n := \sum_{i}^{n} X_i,$$
$$D_n := n - U_n$$

count the number of up and down market moves (respectively) up to time $t := t_n = n\delta t$.

By the binomial model

$$S_n = S(t) = S(t_n) = S(n\delta t) = S(0)\, u^{U_n}\, d^{D_n} = S(0)\, d^n \left(\frac{u}{d}\right)^{U_n}$$

and taking logs (keeping in mind that $n = \frac{t}{\delta t}$) we have

$$\begin{aligned}
\ln\left(\frac{S(t)}{S(0)}\right) &= n \ln(d) + \ln\left(\frac{u}{d}\right) \sum_{i=1}^{n} X_i \\
&= \frac{t}{\delta t} \ln(d) + \ln\left(\frac{u}{d}\right) \sum_{i=1}^{t/\delta t} X_i \\
&= -\frac{\sigma\, t}{\sqrt{\delta t}} + 2\sigma\sqrt{\delta t} \sum_{i=1}^{t/\delta t} X_i
\end{aligned}$$

where we also used the definition of the parameters u, d.

Note that to be completely rigorous we should use $[t/\delta t]$ but we will not complicate the notation.

We will pass to the continuous limit by taking $\delta t \to 0$: By $n = \frac{t}{\delta t}$ this corresponds to $n \to \infty$!

Consider the random sum $Y(t) := -\frac{\sigma t}{\sqrt{\delta t}} + 2\sigma\sqrt{\delta t}\sum_{i=1}^{t/\delta t} X_i$ at this limit.

Important: Because of the chosen scaling even though the above sum has infinite terms it has finite mean and variance!

Indeed

$$\mathbb{E}[Y(t)] = \frac{-\sigma t}{\sqrt{\delta t}} + 2\sigma\sqrt{\delta t}\sum_{i=1}^{t/\delta t}\mathbb{E}[X_i] = \frac{-\sigma t}{\sqrt{\delta t}} + 2\sigma\sqrt{\delta t}\,\frac{t}{\delta t}\,p = \mu\, t$$

(using the definition of p) and

$$Var[Y(t)] = 4\sigma^2\,\delta t\sum_{i=1}^{t/\delta t} Var(X_i) == 4\,\sigma^2\delta t\,\frac{t}{\delta t}\,p(1-p) = 4\sigma^2 p(1-p)t$$
$$\simeq \sigma^2 t \;\; (\text{for } \delta \to 0)$$

Both are finite (since t) is finite and proportional to t!

From the binomial to the Black-Scholes model

In the continuous limit $\delta t \to 0$ ($\delta t = \frac{t}{n}$, t finite) we show that the binomial model yields

$$\mathbb{E}\left[\ln\left(\frac{S(t)}{S(0)}\right)\right] = \mu\, t$$
$$Var\left(\ln\left(\frac{S(t)}{S(0)}\right)\right) = \sigma^2 t$$

Using the CLT with this scaling we can prove that in the limit as $\delta t \to 0$ the stochastic process generated by the binomial model converges (in distribution) to the stochastic process $\{S(t)\ :\ t \in [0, T]\}$ with

$$\ln \frac{S(t)}{S(0)} \sim N(\mu t, \sigma^2 t)$$

which follows the lognormal distribution.

Hence in the limit as $\delta t \to 0$,

$$S(t) = S(0) \exp(\mu t + \sigma\sqrt{t} Z), \quad Z \sim N(0, 1),$$

which is equivalent to the celebrated Black-Scholes model for stock returns!

Recap

The simplest model that can reproduce the basic features of the empirical data for stock prices or indices is the binomial model

$$S(n+1) = H(n+1)S(n), \;\; H(n), \text{ i.i.d. } P(H(n)=u) = p = 1 - P(H(n)=d).$$

This model has the Markov property and there exists an equivalent martingale measure Q that turns the discounted prices $S^*(n) = (1+r)^{-n}S(n)$ into martingales, i.e.

$$\mathbb{E}_Q[S^*(n+1) \mid \mathcal{F}_n] = S^*(n)$$

This model can be calibrated to real data, in terms of the volatility σ, and in the right scaling the binomial model converges to the lognormal distribution (Black-Scholes model)

$$\ln \frac{S(t)}{S(0)} \sim N(\mu t, \sigma^2 t)$$