The Random Behaviour of Assets

In this lecture. . .

- Different types of financial analysis
- Examining time-series data to model returns
- Are prices random?
- The need for probabilistic models
- The Wiener process, a mathematical model of randomness
- The lognormal random walk—The most important model for equities, currencies, commodities and indices

By the end of this lecture you will be able to

- analyze stock price data statistically
- understand and justify the lognormal random walk for assets
- explain where this simple model goes wrong

Introduction

In this lecture we start with some very simple analysis of equity price data, and then using some common sense we build up a **discrete-time** asset price model.

We often like to work in **continuous time**, so we will see how a continuous-time model can be based on our discrete-time model.

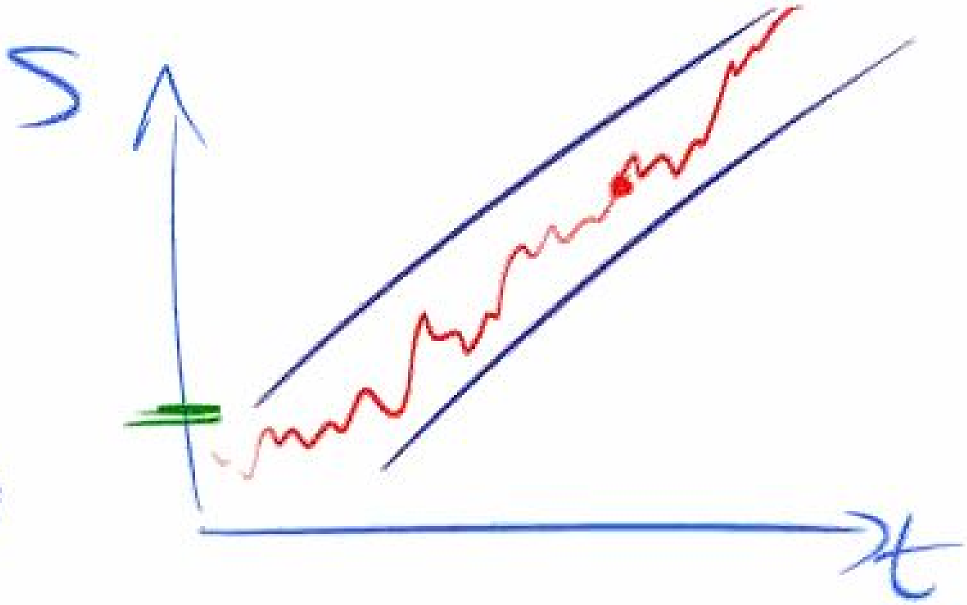
This will be our first look at stochastic calculus and Wiener processes. (This will be very important in most of the CQF!)

The three main types of ‘analysis’ used in finance

1. ~~Fundamental Analysis~~

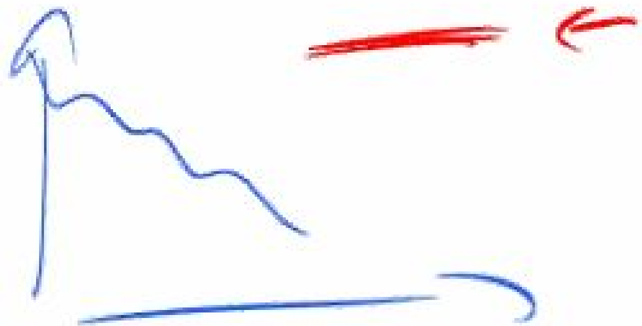
2. ~~Technical Analysis~~

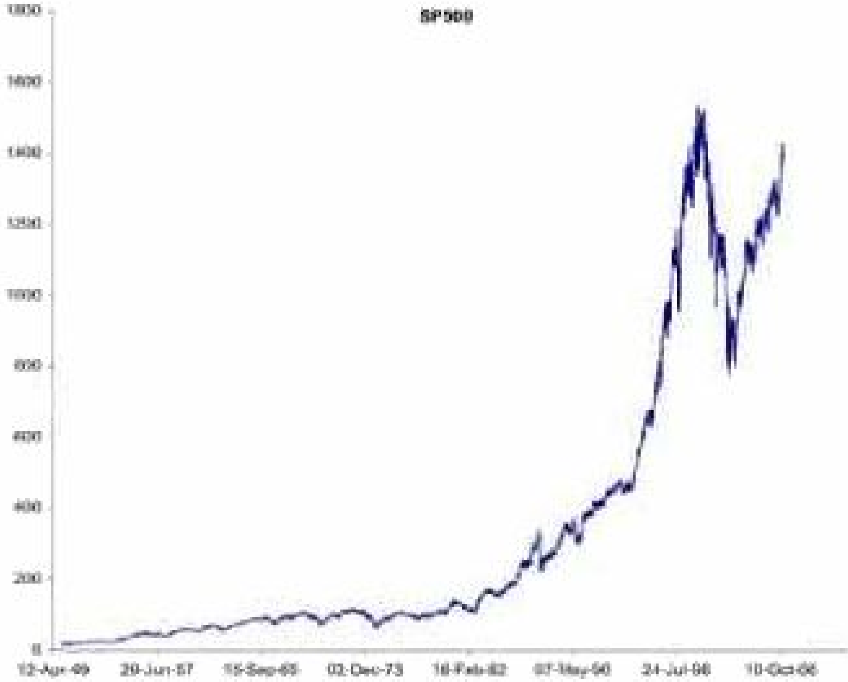
3. **Quantitative Analysis**



μ

σ





The unpredictability that is seen in this figure is the most important feature of financial modelling. Because there is so much randomness, *the most successful mathematical models of financial assets have a probabilistic foundation.*

Why equities, currencies, commodities and indices can be modelled in the same way

Your goal when investing in something is to make a good return. You will be interested in return whether the investment is a stock, commodity, work of art or a case of fine wine.

The *absolute value* of the investment (i.e. in $) is of less interest!

Return means the 'relative' growth in the value of an asset, together with accumulated cashflows (such as dividends), over some period, based on the value that the asset started with:

$$\text{Return} = \frac{\text{Change in value of the asset} + \text{accumulated cashflows}}{\text{Original value of the asset}}.$$

$S_i \quad R_i$

Examining returns

This suggests that instead of examining equity prices directly, we should look at returns over some period.

Often one looks at returns over a period of a day.

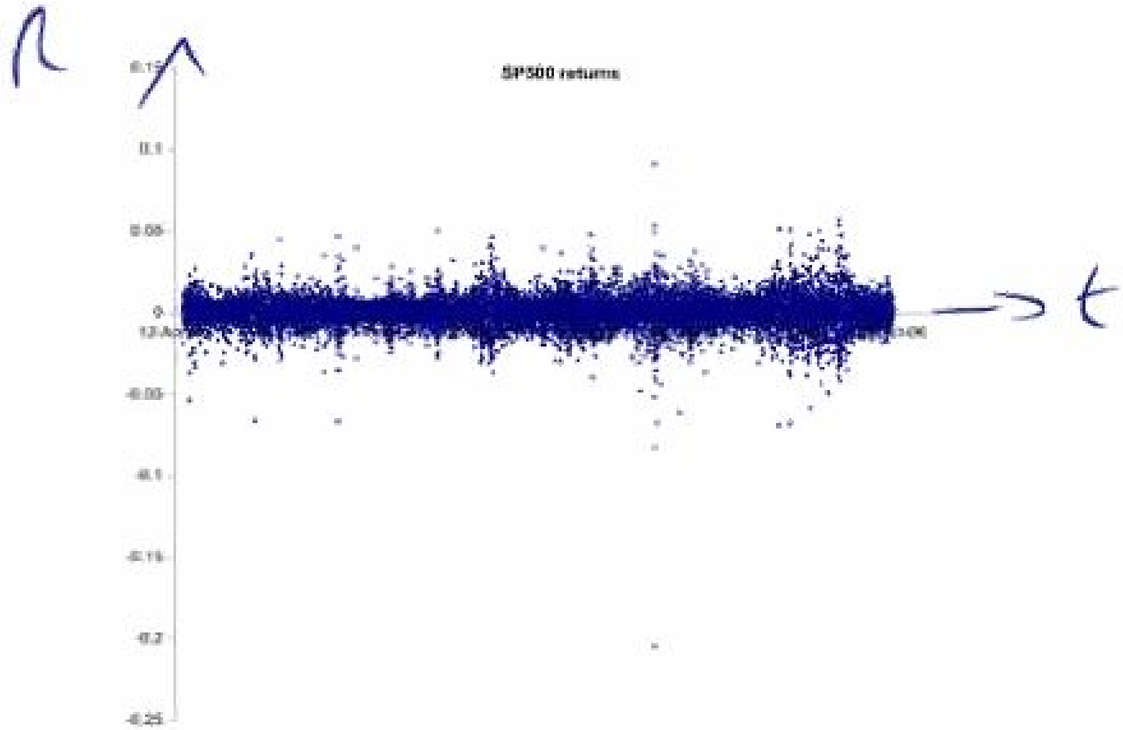
If the asset value on the ith day is denoted by S_i, then the return from day i to day $i+1$ is given by

- $$\frac{S_{i+1} - S_i}{S_i} = R_i.$$

Let's see this on a spreadsheet.

	A	B	C	D	E	F	G	H	I
1	**Date**	**SP500**	**Return**						
2	03-Jan-50	16.66							
3	04-Jan-50	16.85	0.011405		**Average return**	0.00035		=AVERAGE(C:C)	
4	05-Jan-50	16.93	0.004748		**Standard deviation**	0.008909			
5	06-Jan-50	16.98	0.002953						
6	09-Jan-50	17.08	0.005889						
7	10-Jan-50	17.03	-0.002927						
8	11-Jan-50	17.0[illegible]	0.003523				=STDEV(C:C)		
9	12-Jan-50	16.76	-0.01931						
10	13-Jan-50	16.67	-0.00537						
11	16-Jan-50	16.72	0.002999						
12	17-Jan-50	16.86	0.008373		=(B8-B7)/B7				
13	18-Jan-50	16.85	-0.000593						
14	19-Jan-50	16.87	0.001187						
15	20-Jan-50	16.9	0.001778						
16	23-Jan-50	16.92	0.001183						
17	24-Jan-50	16.86	-0.003546						
18	25-Jan-50	16.74	-0.007117						
19	26-Jan-50	16.73	-0.000597						
20	27-Jan-50	16.82	0.00538						
21	30-Jan-50	17.02	0.011891						
22	31-Jan-50	17.05	0.001763						
23	01-Feb-50	17.05	0						
24	02-Feb-50	17.23	0.010557						

This same data was used in the following plot of the daily returns for S&P500 versus time. In the following pages we will model the returns each day as random, and independent from one day to the next.



The mean of the returns is

$$\bar{R} = \frac{1}{M} \sum_{i=1}^{M} R_i = \texttt{AVERAGE(} \cdot \texttt{)}$$

and the sample standard deviation is

$$\sqrt{\frac{1}{M-1} \sum_{i=1}^{M} (R_i - \bar{R})^2} = \texttt{STDEV(} \cdot \texttt{)},$$

where M is the number of returns in the sample. (The expressions on the right are the Excel equivalents.)

From the data in this S&P500 example we find that **the mean is 0.00035** and **the standard deviation is 0.008909**.

Now we know some numbers associated with the (random) return, but what about the shape of the distribution?

We are going to use the data to plot the probability density function for returns. And then we will compare the result with a very famous and important distribution.

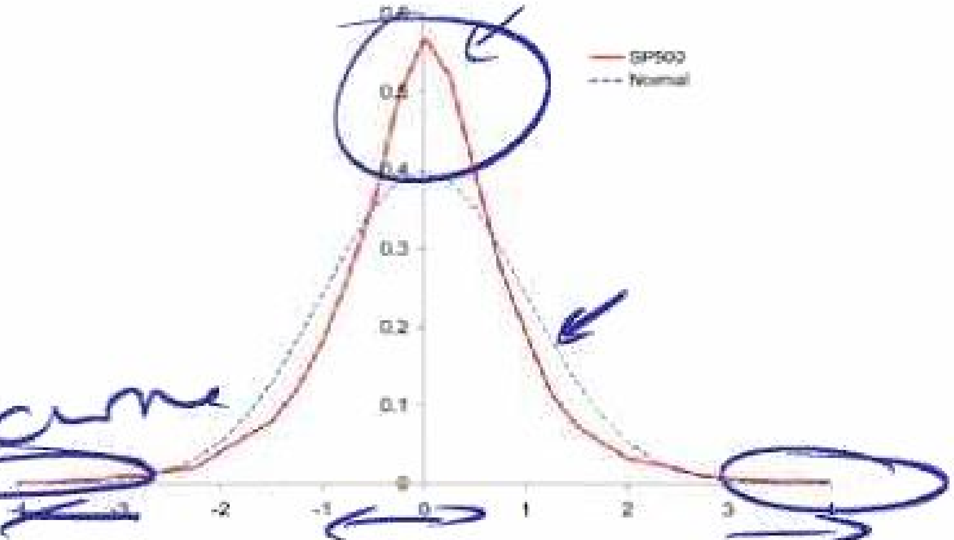
But first we must standardize the distribution to give it a mean of zero and a standard deviation of one.

How to normalize. . .

	A	B	C	D	E	F	G	H
1	**Date**	**SP500**	**Return**	**Scaled rtns**				
2	03-Jan-50	16.66						
3	04-Jan-50	16.85	0.011405	1.2408		**Average return**	0.00035	
4	05-Jan-50	16.93	0.004748	0.493593	**Standard deviation**		0.008909	
5	06-Jan-50	16.98	0.002953	0.292172				
6	09-Jan-50	17.08	0.005889	0.621724				
7	10-Jan-50	17.03	-0.002927	-0.367924				
8	11-Jan-50	17.09	0.003523	0.356137				
9	12-Jan-50	16.76	-0.0193	-2.206775				
10	13-Jan-50	16.67	-0.00537	-0.642092	=(C9-G3)/G4			
11	16-Jan-50	16.72	0.002999	0.297343				
12	17-Jan-50	16.86	0.008373	0.900538				
13	18-Jan-50	16.85	-0.000593	-0.105908				
14	19-Jan-50	16.87	0.001187	0.0939				
15	20-Jan-50	16.9	0.001778	0.160278				
16	23-Jan-50	16.92	0.001183	0.093505				
17	24-Jan-50	16.86	-0.003546	-0.437372				

And here is the distribution!

Gaussian
Normal
Bell-shaped curve



The probability density function for the standardized **Normal distribution** (also having zero mean and standard deviation of one) is also shown:

- $$\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}\phi^2}.$$

Why have I plotted the Normal (or **Gaussian**) distribution? Why do we like it?

Assuming that the empirical returns can be modelled by a Normal distribution then we have our first model!

With

- ϕ **as a random variable drawn from a Gaussian distribution**

our model will represent the returns as a random variable drawn from a **Normal distribution** with a **known, constant, non-zero mean** and a **known, constant, non-zero standard deviation**:

- $$R_i = \frac{S_{i+1} - S_i}{S_i} = 0.00035 + 0.008909 \times \phi_i.$$

$0.00035 + 0.008909 \times \phi$

$N(0,1)$

More generally, i.e. for other indices than S&P500, or for stocks, currencies, commodities, etc.,

- $$R_i = \frac{S_{i+1} - S_i}{S_i} = \text{mean} + \text{standard deviation} \times \phi.$$

This model has two, easily understood parameters.

Aside: Commodities may show seasonal behaviour, so the mean and standard deviation may vary with time.

Goal: How can we get to a **continuous-time** model? At the moment this model is in **discrete time**.

(Maths is easier in continuous time!)

Preliminary question: How do the parameters (the mean return and the standard deviation) vary with the time step we are using (here one day)?

Moving towards continuous time

We need to figure out how the mean and standard deviation of the returns' time series scale with the time step between asset price measurements?

In our example the data is sampled with a time step of one day. We could have used weekly or monthly intervals, or hourly (more data needed, and harder to get). How would this affect the mean and standard deviation?

How does the mean return scale with time?

If the average return in one day is 1%, what is the average return over one week?

The average return scales with the size of the time step.

So the answer is 5% (if there are five business days in a week.)

Obvious!?

Let's do the maths. . .

Call the time step δt. This is going to be a very small number, a tiny fraction of a year.

$$\delta t = \frac{1}{252}$$

I claim we can write

$$\bullet \qquad \text{mean} = \mu \, \delta t,$$

for some μ. (We will assume this to be constant, even if it's not the argument doesn't change much.)

I.e. average return is proportional to the length of the period over which it is measured..

In our S&P500 example we have

$$\text{mean} = 0.00035 = \mu \, \delta t = \mu \times \frac{1}{252},$$

since there are approximately 252 business days in a year.

So $\mu = 0.0882 = 8.82\%$.

Ignoring randomness for the moment while we focus on the mean, our model is simply

mean ~~+ SD~~

$$R_i \quad \frac{S_{i+1} - S_i}{S_i} = \text{mean} = \mu\,\delta t.$$

This can be written as

$$\boxed{S_{i+1} = S_i(1 + \mu\,\delta t).}$$

If the asset begins at S_0 at time $t = 0$ then after one time step $t = \delta t$ and

$$S_1 = S_0(1 + \mu\,\delta t).$$

After two time steps $t = 2\,\delta t$ and

$$S_2 = S_1(1 + \mu\,\delta t) = S_0(1 + \mu\,\delta t)^2.$$

After M time steps $t = M\,\delta t$ and

$$S_M = S_0(1 + \mu\,\delta t)^M.$$

But we can write

δt

$$S_M = S_0 (1 + \mu\, \delta t)^M$$

as

$$S_M = S_0 e^{M \log(1+\mu\, \delta t)}$$

because logarithms and exponentials are the inverses of each other.

And we can approximate the logarithm function...

$$\log(1 + \mu\,\delta t) \approx \mu\,\delta t.$$

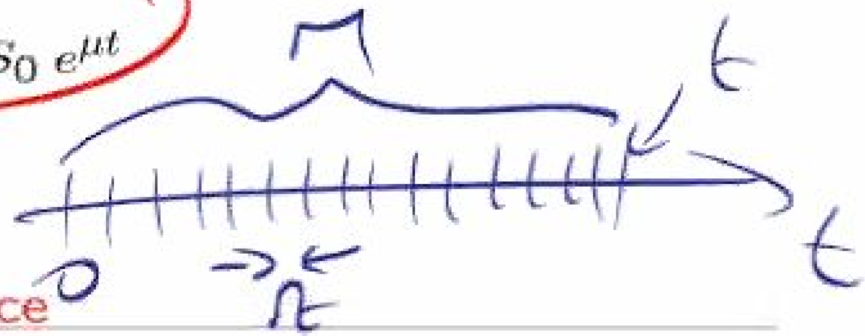
$\ln(1+x) \simeq x$

So

$$S_M \approx S_0 e^{\mu M\,\delta t}.$$

But $M\,\delta t$ is just t, therefore we have

$$S(t) \approx S_0\, e^{\mu t}$$



In the limit as the time step $\delta t \to 0$ S as a function of t becomes

$$S(t) = S_0\, e^{\mu t}.$$

And there aren't any δts in this! Which means that we have something interesting and meaningful when the time step is infinitesimal.

This also shows why the answer to the question about the 1% mean over one day etc. is only approximate. The scaling with time step is actually exponential, it's just that for small enough periods things look linear.

$$(1+0.01)^5 > 1+0.05$$

Aside: What if the mean didn't scale linearly with the time step? What if instead

$$\text{mean} = \mu \, \delta t^{\alpha}$$

with $\alpha \neq 1$? Go through the same analysis to see what happens!

- In the absence of any randomness the asset exhibits exponential growth, just like money in the bank.

- The model is meaningful in the limit as the time step tends to zero.

  To expand on the second point, had we chosen to scale the mean of the returns distribution with any other power of δt it would have resulted in either a boring model ($S(t) = S_0$) or a silly model, with infinite values for the asset.

μ is called the **growth rate** or **drift rate**.

Now let's turn our attention to the standard deviation. . .

How does the standard deviation of returns scale with time?

If the standard deviation of returns is 1% over one day, what is the standard deviation of returns over one week?

This one is not so obvious!

$\sqrt{5}$

Clue: When you have independent random numbers (such as returns from one day to the next) you cannot add standard deviations. Oh, no!

But you can add *variances*. I.e. when X and Y are independent $\mathsf{Var}[X+Y] = \mathsf{Var}[X] + \mathsf{Var}[Y]$.

We will need this in what follows.

Let's suppose that the standard deviation scales with δt^{α}. Therefore the variance scales with $\delta t^{2\alpha}$.

From time zero to time t how many random returns are there? Easy, just $t/\delta t$.

So we have a number $t/\delta t$ of variables each with variance of size $\delta t^{2\alpha}$. Add up this many variances and you'll get a total variance from time zero to time t of size...

$$\frac{t}{\delta t} \times \delta t^{2\alpha}.$$

We want to have a finite, non-zero, variance after a finite period of time *in the limit as* $\delta t \to 0$ therefore we must have

$$\alpha = \tfrac{1}{2}.$$

Therefore the standard deviation of returns scales with the **square root of the time step**.

So the answer to the question is $\sqrt{5}\,\%$ (if there are five business days in a week).

And generally we can write:

- $$\text{standard deviation} = \sigma\, \delta t^{1/2},$$

where σ is a parameter measuring the amount of randomness.

σ is the **volatility**.

It is the **annualized standard deviation of returns**.

In our S&P500 example we have

$$\text{standard deviation} = 0.008909 = \sigma\ \delta t^{\frac{1}{2}} = \sigma \times \frac{1}{\sqrt{252}}.$$

So $\sigma = 0.141 = 14.1\%$.

Volatility

What units do the drift, μ, and the volatility, σ, have?

$1/\text{Time}$

$1/\sqrt{\text{Time}}$

$e^{\mu t}$

$\sigma^2 t$

$e^{\sigma t}$

Back to the full model

Our asset return model, in words, is

ϕ

$$R_i = \frac{S_{i+1} - S_i}{S_i} = \text{mean} + \text{standard deviation}.$$

And using only symbols,

$$R_i = \frac{S_{i+1} - S_i}{S_i} = \mu\,\delta t + \sigma\phi\,\delta t^{1/2}.$$

This can be written as

- $$\boxed{S_{i+1} - S_i = \mu S_i \, \delta t + \sigma S_i \phi \, \delta t^{1/2}.} \qquad (1)$$

This is a model!

(But it is still in discrete time.)

$$S_{i+1} = S_i \left(1 + \mu \, \delta t + \sigma \sqrt{\delta t} \right.$$

You can also write it as

$$S_{i+1} = (1 + \mu\,\delta t)S_i + \sigma S_i \phi\,\delta t^{1/2}.$$

Equations in this form are the basis for **Monte Carlo simulations**.

This is a discrete-time model for a **random walk** of the asset.

We know exactly where the asset price is today but tomorrow's value is unknown.

Because of their different scalings with time, the growth and volatility have different effects on the asset path.

- The growth is not apparent over short timescales. The volatility dominates in the short term.

- Over long timescales, for instance decades, the growth becomes important.

Path of the logarithm of SP500, also showing its expected path (the straight line) and one standard deviation above and below (parabolas).

The Wiener process

We have still not reached our goal of continuous time, we still have a discrete time step.

We will now see a brief introduction to the continuous-time limit. (You'll be seeing the more rigorous side in later CQF lectures!)

We now introduce some more, very standard, notation: $d\cdot$ means 'the change in' some quantity.

So dS is the 'change in the asset price.'

$\delta t \to dt$

But *this change will be in continuous time.*

- In effect, we will go to the limit $\delta t = 0$.

$\frac{dy}{dx}$ $\quad \frac{ds}{dt}$ $\quad \frac{d^2 s}{dt^2} = -g$

The first δt on the right-hand side of

$$S_{i+1} - S_i = \mu S_i \, \delta t + \sigma S_i \phi \, \delta t^{1/2}.$$

becomes dt, but the second term is more complicated.

$$dS = \mu S \, dt + \sigma S \, dX$$

$$\phi \qquad \frac{dt^{1/2}}{dt} = \frac{1}{dt^{1/2}}$$

We cannot straightforwardly write $dt^{1/2}$ instead of $\delta t^{1/2}$.

Why not?

We are going to write the term $\phi\,\delta t^{1/2}$ as

$$dX.$$

You can think of dX as being a random variable, drawn from a Normal distribution with mean zero and variance dt:

$$E[dX] = 0 \quad \text{and} \quad E[dX^2] = dt.$$

This is not exactly what it is, but it is close enough to give the right idea. (More later!)

- This is called a **Wiener process**.

We can build up a continuous-time theory using Wiener processes instead of a discrete-time theory using Normal distributions.

The most important model for equities, currencies, commodities and indices

Using the Wiener process notation, the asset price model can be written as

- $$dS = \mu S\, dt + \sigma S\, dX.$$

This is a **stochastic differential equation** (or **sde**). It is a continuous-time model of an asset price.

It is the most widely accepted model for equities, currencies, commodities and indices, and the foundation of much finance theory.

(And sdes play a BIG role in all of quantitative finance!)

Summary

Please take away the following important ideas

- The return on an investment is the natural quantity to analyze
- In finance theory this return is usually treated as being random
- The random return is often assumed to be Normally distributed. This is not perfect but is a good starting point
- The asset price can then be modeled as a lognormal random walk
- This random walk is the most popular asset price model, and is in the form of a stochastic differential equation