ECO423: **Valuation with the binomial model**

Jørgen Haug

Department of Finance, NHH

Spring 2023

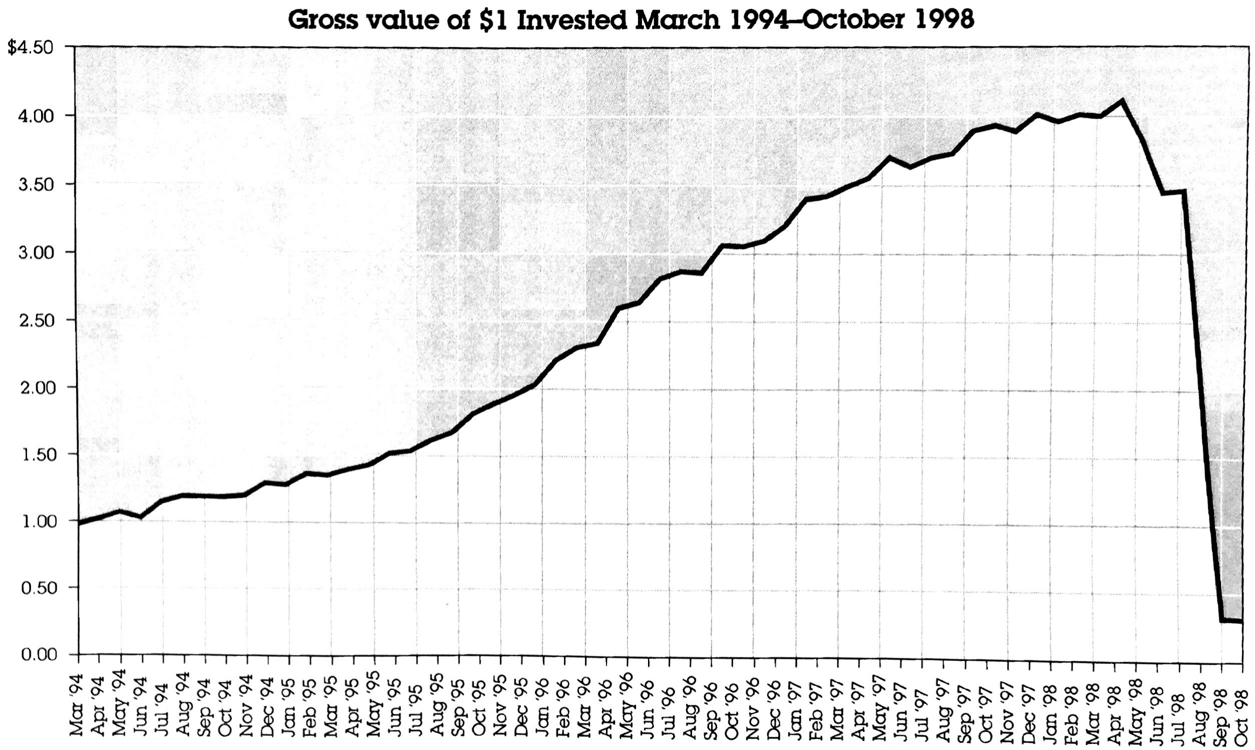
After this lecture you should understand

- the role of *replicating portfolio* and *no arbitrage*
    - that replicating portfolio must be updated over time
- that derivatives pricing simplifies classical present value techniques by
    1. (objectively) adjusting for systematic risk in the probabilities—*risk-adjusted probabilities*
    2. discounting at the riskless rate

- Aim of course
    1. Develop *high precision models* to determine present values/prices of derivatives
    2. Understand and master the wide applications for the toolbox, and the value of precision
    3. Understand the *limitations* of the toolbox
- High precision particularly valuable in hedging/risk management and trading/speculation
- Understanding the economics of derivatives valuable also in low-precision applications like capital budgeting/real options analysis or mutual fund management

- Models we cover in the course staple of investment banks, hedge funds, . . .
    - . . . less prevalent in capital budgeting
- Long-Term Capital Management (LTCM) great example of the value of the models we cover in this course
- LTCM also great illustration of dangers/limitations of the models
- LTCM a "hedge fund" $\sim$ an investment "club"
    - up to 99 "members" can invest a minimum of \$1 million each
    - 11 highly competent partners plus about 30 traders
    - two of the partners Nobel laureates (Merton, Scholes)
- Main strategy: spread trades
    - Identify "very similar" assets
    - Use models to determine "fair price difference" $\sim$ "fair spread"
    - Identify *deviations* from "fair spread"
    - Buy "cheap" leg, sell "expensive" leg
    - Hold until spread converges back to "normal" levels
    - Use models to monitor/manage risk exposure

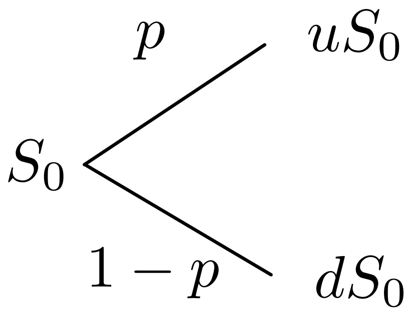
- LTCM started trading in March 1994



- Invested capital quadrupled in four years, with *very low levels of volatility* ($< 15\%$)
- . . . closed down in October 1998, with $> \$1.5$ billion in 2018-assets and \$1.5 trillion in notional value of derivatives! And *lots* of debt!

- Keep in mind that a model may be valuable *because* it leaves out many/most aspects of reality
    - by its very nature a simplification of reality
    - whether the “left out”-part can be ignored depends on the application
    - it is valuable if it helps us solve a problem we want to solve
    - trying to capture all of reality prevents us from gaining valuable insights
    - …but in the end we are confronted with reality
- To gain economic *intuition* for option pricing, it is sufficient to work with a very simplistic model of reality
- Let’s use binomial model to recall the role of
    - (Dynamic) Replicating portfolio
    - No arbitrage
    - Physical versus risk-adjusted probabilities
- Course develops and applies these economic arguments in a *high-precision* environment

- Consider a stock



where uS_0 and dS_0 are realizations of S_1

- Consider next a derivative cash flow

$$X_1 = g(S_1) = \begin{cases} g(uS_0) & = H \\ g(dS_0) & = L \end{cases}$$

- *Q:* What is the 'fair price' of the derivative? I.e., what is $PV_0 = PV(X_1)$?

- Consider first the classical DCF method
- Assume X_1 has required gross rate of return μ^X

$$PV_0 = \frac{1}{\mu^X}[pH + (1-p)L] \equiv \frac{1}{\mu^X}\mathrm{E}^{p}\{X_1\}$$

(Q: Why does this expression for 'fair price' make sense—where does it come from?)

- μ^X determined by some capital asset pricing *model*
- *Difficulty 1:* μ^X hard to determine, often based on subjective considerations
- *Difficulty 2:* μ^X generally time and state dependent in multi-period settings

(which one-period model cannot shed light on!)

- *Idea:*
  - Find portfolio of existing assets that perfectly replicates X_1
  - No arbitrage ($\sim$ Law of One Price) $\implies$
    $PV_0 =$ current price of *replicating portfolio*
- *Q:* Is it possible to find such a portfolio?
  - If markets are *complete*: Yes, always
  - Markets are complete if there are as many independent/unique assets as there are states—as we will see shortly
    - independent $\sim$ cannot be replicated as a portfolio of other assets
- We have two states
- Assume in addition to S, we can invest in a riskless savings account, with gross return
$$R = \mathrm{e}^r \iff r = \ln(R)$$
- Consider portfolio of Δ stocks and θ invested in savings account

- Assume perfect capital markets (a model assumption)
- Want to ensure

$$\Delta S_1 + \theta R = X_1$$

regardless of which state occurs, i.e.,

$$\Delta u S_0 + \theta R = H$$
$$\Delta d S_0 + \theta R = L$$

$\longrightarrow$ two equations (one for each state) in two unknowns (one for each security)

- If $u \neq d$ then the two equations are independent (and we say the assets are independent), and a unique solution exists!

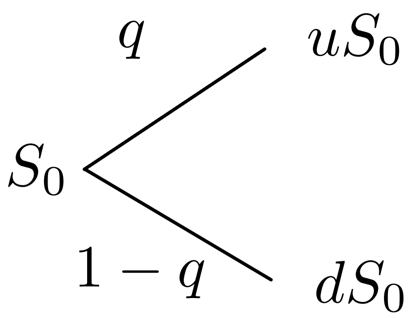
- Assuming $u > R > d$ (consistent with equilibrium demand for the assets), we get
  - From the first equation
$$\theta = \frac{H - \Delta u S_0}{R}$$
  - From the second equation
$$\Delta d S_0 + H - \Delta u S_0 = L$$
    or
$$\Delta = \frac{H - L}{S_0(u - d)}$$
  - Substitute back into expression for θ
$$\theta = \frac{1}{R}\left(H - \frac{uH - uL}{u - d}\right) = \frac{uL - dH}{R(u - d)}$$
  - The time 0 *price of the replicating portfolio* is thus
$$\Delta S_0 + \theta = \frac{1}{R}\left(\frac{R - d}{u - d}H + \frac{u - R}{u - d}L\right) \overset{[\text{LOP}]}{=} PV_0 \qquad (1)$$

- Observe now that

$$q \equiv \frac{R-d}{u-d} > 0, \qquad 1-q = \frac{u-R}{u-d} > 0$$

and $q + (1-q) = 1$

- q and $1-q$ have the properties of probabilities

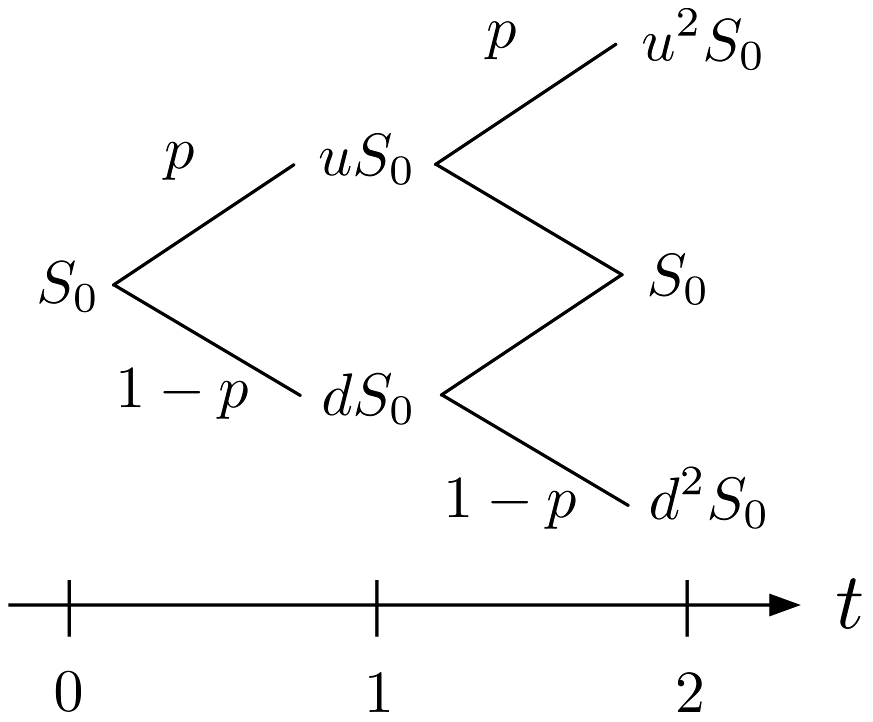


- Can now write

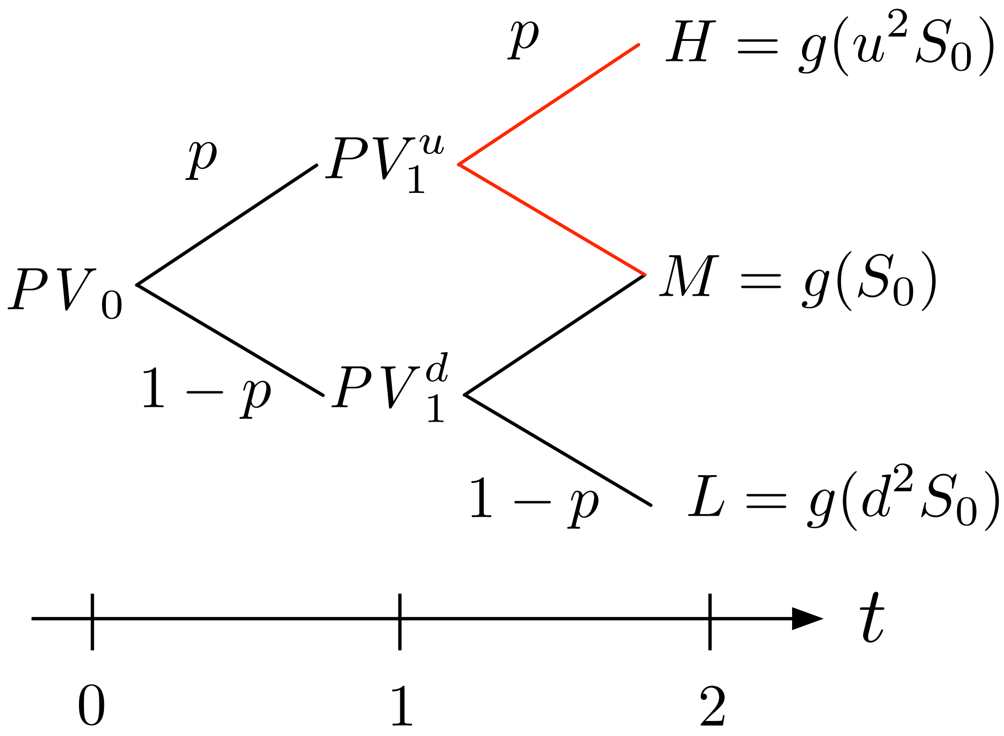
$$PV_0 = \frac{1}{R}[qH + (1-q)L] \equiv \frac{1}{R}\mathrm{E}^{q}\{X_1\} \tag{2}$$

- I.e., using *fictitious* (risk-adjusted) probabilities q we can compute PV_0 by discounting back at R—the riskless rate!
- R objectively observable
- Can use riskless rate for all dates and states (but generally not true that $R_t = R_s$ when $t \neq s$)
- *Important insight:* Option pricing theory is just another *technique* for finding *present values* of cash flows!

- One-period two-state model
  - Great for developing economic intuition
  - Poor for estimating/evaluating real-world present values/prices
- Can increase precision in binomial model by sub-dividing time to expiration into more sub-periods
- *Q:* Anything new with multiple periods?
- Simplest multi-period model of risk with $u = 1/d$



- Derivative now has
  - three date 2 cash flows H, M, and L
  - two possible date 1 present values PV_1^u, PV_1^d
  - one possible date 0 present value PV_0 as before



- Find PV_0 by repeating one-period analysis, by *backwards induction*
  - u, d, and R identical throughout *this* tree $\longrightarrow$ can use same q everywhere
  - PV_1^u date 1 present value of H and M in state $S_1^u = uS_0$:

$$PV_1^u = \frac{1}{R} \mathrm{E}^q \left\{ X_2 | S_1 = S_1^u \right\} = \frac{1}{R} \left[Hq + M(1-q) \right]$$

  - PV_1^d date 1 present value of M and L in state $S_1^d = dS_0$:

$$PV_1^d = \frac{1}{R} \mathrm{E}^q \left\{ X_2 | S_1 = S_1^d \right\} = \frac{1}{R} \left[Mq + L(1-q) \right]$$

  - PV_0 date 0 present value of PV_1^u and PV_1^d

$$PV_0 = \frac{1}{R} \mathrm{E}^q \left\{ PV_1 \right\} = \frac{1}{R} \left[PV_1^u q + PV_1^d (1-q) \right]$$

- Important economic insight from looking at the replicating portfolio $\longrightarrow$ Voluntary Exercise 1
  - To compute derivative's β, use that (realized) gross rate of return on replicating portfolio is given by

$$R_1^X = \frac{PV_1}{PV_0} = \frac{\theta_0 R + \Delta_0 S_1}{\theta_0 + \Delta_0 S_0} = \frac{\theta_0}{\theta_0 + \Delta_0 S_0} R + \frac{\Delta_0 S_0}{\theta_0 + \Delta_0 S_0} R_1^S$$

  - Standard definition of $\beta^S = \frac{\mathrm{Cov}\left(R^S, R^M\right)}{\mathrm{Var}(R^M)}$
  - Similar for returns from date 1 to date 2, R_2^X

- Will in general revise portfolio at each date, and in each state
  - Three possible portfolios in above tree

$$
\begin{aligned}
t = 0,\ S_0 = S_0 : \quad & \theta_0 = \frac{uPV_1^d - dPV_1^u}{R(u-d)}, \quad && \Delta_0 = \frac{PV_1^u - PV_1^d}{S_0(u-d)} \\
t = 1,\ S_1 = S_1^u : \quad & \theta_1^u = \frac{uM - dH}{R(u-d)}, \quad && \Delta_1^u = \frac{H - M}{S_1^u(u-d)} \\
t = 1,\ S_1 = S_1^d : \quad & \theta_1^d = \frac{uL - dM}{R(u-d)}, \quad && \Delta_1^d = \frac{M - L}{S_1^d(u-d)}
\end{aligned}
$$

  - Portfolio typically revised between date 0 and date 1
  - Portfolio typically different in states S_1^u and S_1^d

- *Aside:* Market still complete
  - Each 'local' one-period tree has two branches
  - There are two independent assets throughout the tree
  - Generally, complete if there are as many independent assets as the maximal number of local branches
  - Bottom line: Complete if each 'local' problem of determining the replicating portfolio has a *unique solution*
  - Above case: Two linear equations in two unknowns throughout tree
  - We will work mostly with complete markets in this course
  - Exceptions: Real options with non-traded risk factors; commodities with stochastic convenience yields; stochastic volatility

Exercise

Assume $S_0 = 9$, $u = 1.25$, $d = 1/u$, $R = 1.05$, $p = 0.55$. Consider a put option on S with strike price $K = 10$.

a) Compute Δ, θ, q for the put option.

b) Verify that payoff of portfolio (Δ, θ) coinsides with that of the put.

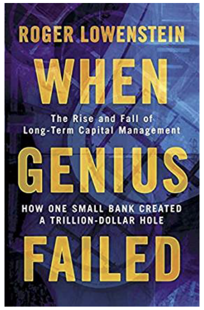
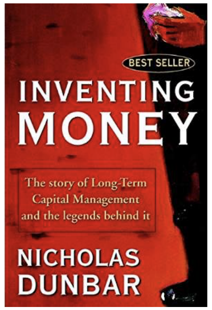
c) Verify that date 0 price of replicating portfolio, LHS of (1), equals 1.185.

d) Compute present value of put as q-expected value discounted at riskless rate R, as in (2).

e) Verify that the q-expected return on S and the put both equal R.

f) Find the value of a call with strike K, as q-expected payoff discounted at R. Important lesson:

- Can compute the price of *any* derivative of S, using the same q.
- Using replicating portfolio requires recomputing (Δ, θ).

- 
ROGER LOWENSTEIN
WHEN
The Rise and Fall of
Long-Term Capital Management
GENIUS
HOW ONE SMALL BANK CREATED
A TRILLION-DOLLAR HOLE
FAILED
- 
BEST SELLER
INVENTING
MONEY
The story of Long-Term
Capital Management
and the legends behind it
NICHOLAS
DUNBAR

ECO423: **Modeling prices as diffusions**

Jørgen Haug

Department of Finance, NHH

Spring 2023

After this section you should

- remember basic properties of and be able to work with
  - the standard Wiener process (W)
  - the generalized Wiener process (X)
  - the geometric Wiener process (S)
- know and understand the role of the equivalence $dW = \sqrt{dt}z$
- know the definition of a diffusion, and be able to judge if a stochastic process is a special case of a diffusion or not
- be able to give economic interpretations of the drift and dispersion terms
  - systematic vs. total risk in the geometric Wiener process

Derivatives

Definition

A derivative asset is a cash flow $g(t, S_t)$ that depends at most on time and the value S_t of one or more underlying assets.

- I'll use 'asset' and 'security' interchangeably throughout the course. The latter is a special case of the former, and is a tradable entitlement to a part of an asset.
- We assume rational behavior, perfect markets throughout course.

Example

A European call option with time to expiration T and strike price K offers a payoff $S_t - K$ at date $t = T$ and zero otherwise:

$$g(T, S_T) = g(S_T) = \max(S_T - K, 0).$$

Example

An American put option with time to expiration T and strike price K offers a one-time payoff $K - S_t$ at date $\tau \leq T$ chosen by the owner, and zero otherwise:

$$g(\tau, S_\tau) = g(S_\tau) = \max(K - S_\tau, 0).$$

- Problem: What is the present value P_0 of $P_T = g(T, S_T)$?
- Ambition: Find high-precision estimate of P_0 for "any" kind of derivative $g(\cdot)$!
- Strategy: Create high-precision model of distribution of S_t
  - Observation: Any valuation problem requires good estimates of cash flow distributions as we must always estimate expected cash flows $\mathrm{E}\{P_t\}$
  - Applies also when we use risk-adjusted discount rates (DCF),

$$P_0 = \frac{\mathrm{E}\{P_t\}}{\mu}$$

Motivation

- We need a model for price *increments* ΔS_t
- Let's start with deterministic growth at *rate* μ per year

$$S_{t+\Delta t} = S_t + \mu S_t \Delta t$$

or, a drift term

$$\Delta S_t = \mu S_t \Delta t$$

- Let's add a random shock to allow for uncertainty

$$\Delta S_t = \mu S_t \Delta t + \epsilon_{t+\Delta t}$$

and let's assume uncertainty is proportional to price level, with the dispersion term

$$\epsilon_{t+\Delta t} = \sigma S_t u_{t+\Delta t}$$

Motivation

- Let's assume
  - $u_{t+\Delta t}$ has a normal distribution—"easy" to work with
  - uncertainty increases with length of time period Δt
- To keep it simple let
$$u_{t+\Delta t} \sim \mathrm{N}(0, \Delta t)$$
- No loss of freedom in having mean zero and variance Δt
  - growth term captures expected changes
  - dispersion term scales the uncertainty
- For convenience and convention, change notation
$$u_{t+\Delta t} = W_{t+\Delta t} - W_t = \Delta W_t$$
- W will serve as basic building block for uncertainty

Motivation

Exercise:

- Enumerate the dates $0, \Delta t, 2\Delta t, \ldots$ simply as $0, 1, 2, \ldots$
- Consider the price changes $\Delta S_t = S_{t+1} - S_t$
- Show that changes imply levels, for instance

$$S_3 = S_0 + \Delta S_0 + \Delta S_1 + \Delta S_2 \tag{1}$$

(1) corresponds to setting $N = 3$ in the more general result that the date T price equals the initial price S_0 plus price changes

$$S_T = S_0 + \sum_{n=0}^{N-1} \Delta S_{n\Delta t}$$

where $\Delta t = T/N$.

The Wiener Process (a.k.a. Brownian Motion)

Definition

A standard Wiener process W has the key property

- $W_s - W_t \sim \mathrm{N}(0, s-t)$, *for* $s \geq t$

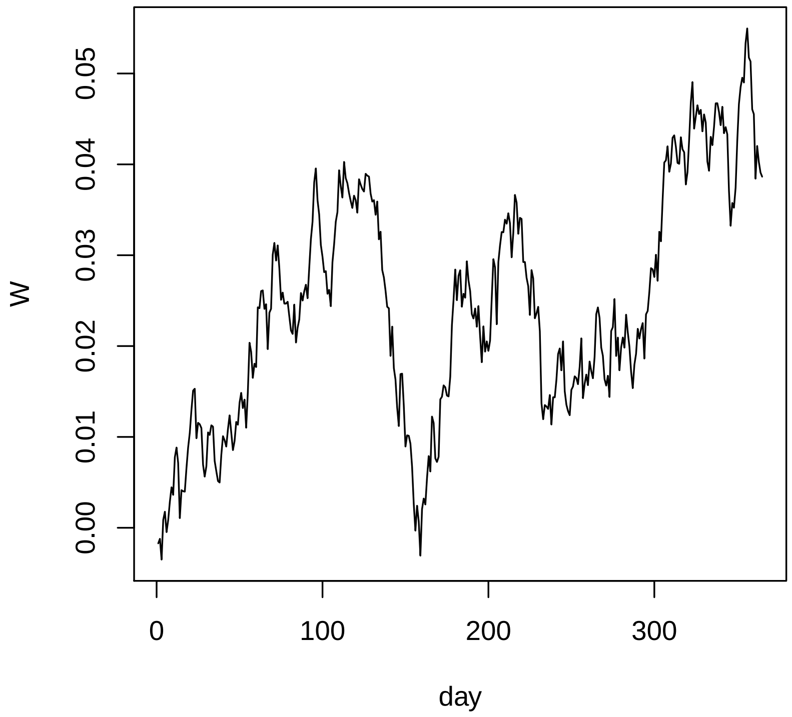
and in addition

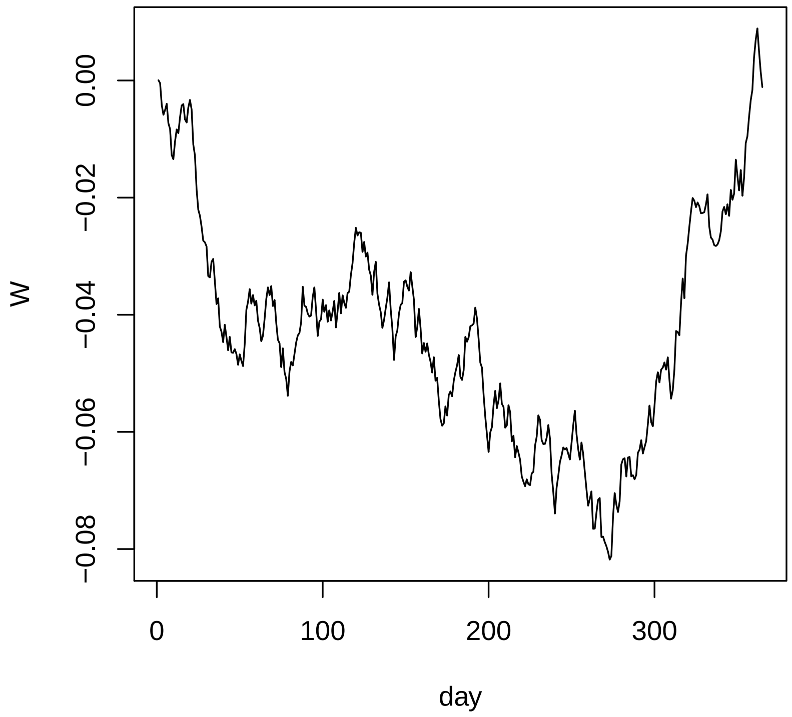
- $W_0 = 0$
- *Increments* $W_s - W_t$ *for non-overlapping intervals* (t, s) *are independent*
- *The sample paths of* W *are continuous*

- We'll use W_t and $W(t)$ interchangeably, and similarly for $S_t = S(t)$ etc (fewer parentheses to keep track of with S_t)
- *Q:* Is W a good model for a stock or oil price?

An important equivalence: equal in distribution

- Computer software do not have a Wiener process function/command (but for instance R has a package you can install)
- Easy to implement a Wiener process by observing
  - $W_s - W_t \sim \mathrm{N}(0, s-t)$ is *equal in distribution* to $\sqrt{s-t}z$ where $z \sim \mathrm{N}(0,1)$, or
    $$W_t \overset{d}{=} \sqrt{t}z \tag{2}$$
    Exercise: Check that (2) is true!
  - All decent software has an implementation of the standard normal distribution function $\mathrm{N}(0,1)$
- We'll use (2) in both numerical and analytical work





R-code

```
set.seed(1) # The right hand figure uses set.seed(10)
dW = rnorm(365, 0, sqrt(1/365))
W = cumsum(dW)
plot(W, type="l", xlab="day")
```

Comments on R and numerical implementation in general

- Always read your software's documentation for how the random number generator is implemented
  - Here **rnorm**(N, m, s); m is the mean and s is the *standard deviation*
  - In other software the last argument may be *variance* rather than standard deviation
  - Using the wrong quantity is hard to spot from plotting W

Stochastic processes in continous time

- A *stochastic process* is simply a list of random variables that are organized according to their date—one random variable per date
  - For instance $\{W_0, W_{\Delta t}, W_{2\Delta t}, \ldots, W_{N\Delta t}\}$, which we graphed two possible realizations of above
- We want to
  - have high precision
  - allow for cashflows at any time
- Achieve both of these objectives by allowing $\Delta t > 0$ to become arbitrarily small
- We introduce the *differential* dt for the "smallest possible" Δt (called "infinitesimal") and similarly

$$dW_t = W_{t+dt} - W_t$$

Q: What is the distribution of dW_t?

The Generalized Wiener Process

Let's use W to build a more general stochastic process, to better capture price properties

- Let
$$\Delta X_t = X_{t+\Delta t} - X_t = \mu_X \Delta t + \sigma_X \Delta W_t \tag{3}$$
starting at $X_0 = x$ (called the *inital condition*).
- *Q:* What is the distribution of X_T?
- *Q:* Is it a good model for stock or oil prices?
- We will work with (3), but allow for two modifications
    - Allow μ_X and σ_X to vary with time: $\mu_X(t)$ and $\sigma_X(t)$
    - Consider the model in continuous time $\Delta t \to dt$, and $dX_t = X_{t+dt} - X_t$
- Allowing parameters to vary deterministically with time is necessary to allow for e.g. seasonalities

The Generalized Wiener Process

Definition

We call X a generalized (a.k.a. an arithmetic) Wiener process if

$$dX_t = \mu_X(t)\,dt + \sigma_X(t)dW_t, \qquad X_0 = x, \tag{X}$$

where μ_X and σ_X are functions of at most time, t.

- Exercise: Assume that $\mu_X(t) = \mu_X$ and $\sigma_X(t) = \sigma_X$ in (X), i.e., are constant. What is the distribution of X_T?
- An equivalent way to write (X), which you may see in journal articles, is

$$X_t = X_0 + \int_0^t \mu_X(s)\,ds + \int_0^t \sigma_X(s)\,dW_s \tag{4}$$

- We'll stick to the *differential form* in (X) rather than the *integral form* in (4).

Definition

- Integral form not so "mystical" when considered as a discrete sum, as in determining the distribution of X_T in (3)

$$X_t = X_0 + \int_0^t \mu_X(s)\,ds + \int_0^t \sigma_X(s)\,dW_s = X_0 + \int_0^t dX_s \approx X_0 + \sum_n \Delta X_{s_n}$$

  where $s_n = n\Delta t$.

- Can generalize (X) by allowing μ_X and σ_X to be stochastic
  - Allows for "arbitrary" distribution of X_t
  - Mathematics kept reasonably simple by restricting how free we are to define μ_X and σ_X

Definition

The stochastic process Y is a diffusion (a.k.a. stochastic differential equation (SDE), or Itô process) if

$$dY_t = \mu_Y(t, Y_t)\,dt + \sigma_Y(t, Y_t)\,dW_t, \qquad Y_0 = y, \tag{Y}$$

where μ_Y and σ_Y are functions of at most t and Y_t.

Properties

- Observe that (X) is a special case of (Y)
- (Y) is the most general framework we consider in this course
- (Y) has the Markov property: Future changes in Y, "dY_t", depend only on time and the current value of Y
  - typically know its contemporary and historical values—the realized *path*—Y_u, $u \leq t$, and its distribution/law of motion (Y)
  - but *historical values* Y_u, $u < t$, are *decision irrelevant*

  $\sim$ history independence
- Cannot solve (Y) using "standard" methods, because W is *nowhere differentiable*, i.e.,

  $$\frac{dW(t)}{dt}$$

  does not exist!
- *Q:* Is non-differentiability of W good or bad economic property?

Our stock price model

- Will use special case of (Y) as our main example—the most common (stock) price model
- Let S be the price of a stock and assume

$$dS_t = \mu S_t \, dt + \sigma S_t \, dW_t, \qquad S_0 = s > 0, \tag{S}$$

  where μ and $\sigma > 0$ are constants.
- *Q:* (S) a special case of (Y) because. . . ?
- Diffusions on the form (S) are called *geometric Wiener processes* (a.k.a. geometric Brownian motion)

Properties

- (S) has the properties
    - $S_0 = s > 0$ causes $S_t > 0$, all t
    - Stock *returns* over non-overlapping time intervals are independent
    - Distribution of future stock *returns* is independent of past returns and S_t
    - It is Markov; all decision relevant info is contained in current stock price (knowing the law of motion (S))
- Project 1 asks you to implement (S) in Excel or R
- For some computations it is useful to be specific about what information we use
- We will use $\mathrm{E}_t\{\cdot\}$ and $\mathrm{Var}_t(\cdot)$ to mean that you compute expectations and variances knowing the date t value of S, and its history (conditional expectation and variance).

Economic interpretation of parameters

- Behavior of (S) determined by μ and σ
- To attain economic interpretation of the parameters, observe first that

$$\frac{dS_t}{S_t} = \mu\, dt + \sigma\, dW_t$$

  is the net *rate of return* of the stock during $[t, t + dt]$
- The annualized *expected rate of return* is

$$\mathrm{E}_t\left\{\frac{dS_t}{S_t}\right\}\frac{1}{dt} \tag{5}$$

- Annualized *total risk* is given by

$$\mathrm{Var}_t\left(\frac{dS_t}{S_t}\right)\frac{1}{dt} \tag{6}$$

Alternative expression for stock price

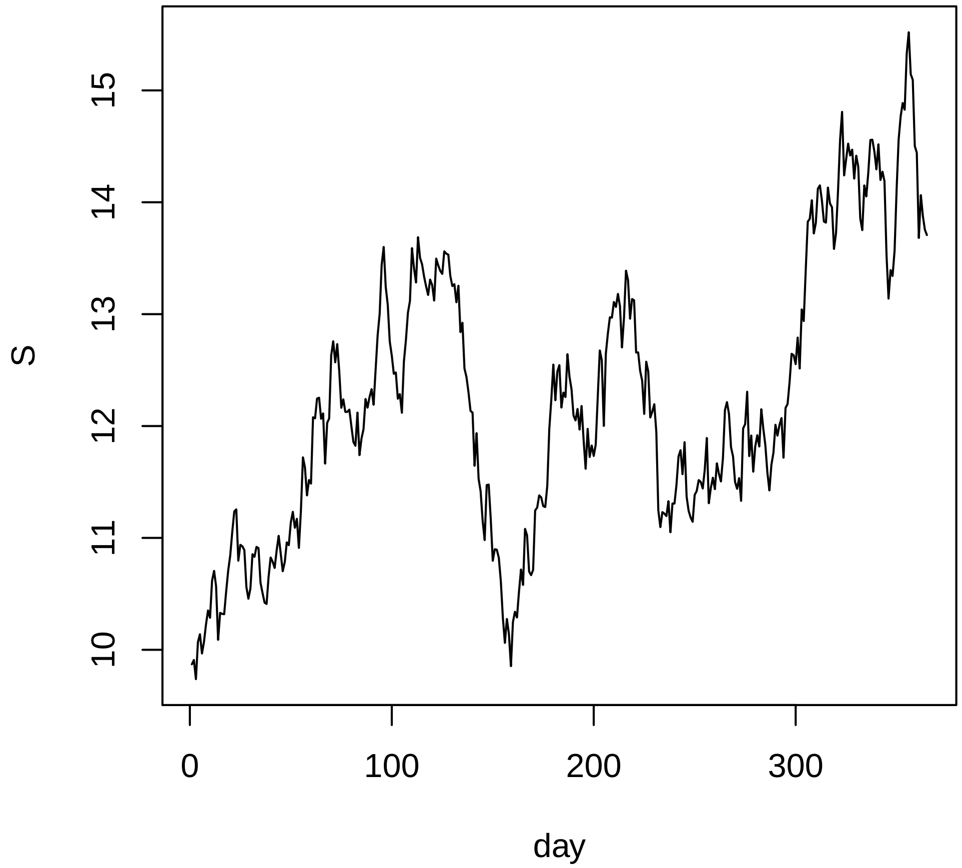
- Exercise: To become more confortable working with diffusions, compute (5) and (6). Reflect on what valuation/price relevant information (S) contains. We will go through these computations later in class.
- We will later show that the solution to (S) is

$$S_t = S_0 \mathrm{e}^{\left(\mu - \sigma^2/2\right)t + \sigma W_t}$$

- Using that $W_t \stackrel{d}{=} \sqrt{t}z$ we can rewrite this as

$$S_t = S_0 \mathrm{e}^{\left(\mu - \sigma^2/2\right)t + \sigma \sqrt{t}z}$$

where $z \sim \mathrm{N}(0, 1)$



R-code

```
dt = 1/365; mu = 0.1; sigma = 0.4; S0 = 10
z = set.seed(1); rnorm(365, 0, 1)
X = exp((mu-sigma^2/2)*dt + sigma*sqrt(dt)*z)
S = S0*cumprod(X)
plot(S, type="l", xlab="day")
```

ECO423: **Gains from trade, Itô's Lemma**

Jørgen Haug

Department of Finance, NHH

Spring 2023

After this section you should

- *understand* economic meaning of and remember definitions of
  - riskless asset
  - gains processes
  - self-financing portfolios
  - arbitrage opportunities
- be able to work with Itô's Lemma: Given access to the formula, carry out the steps
  - identification
  - computation
  - back-substitution
  - simplification
- *remember* the multiplication table

Dividends

- Recall our stock price model

$$dS_t = \mu S_t \, dt + \sigma S_t \, dW_t, \qquad S_0 = s > 0 \tag{S}$$

- Have so far not accounted for possible dividends
- We will assume S is the *ex-dividend* price
- Will allow for a special kind of dividend payout
  - Payout of $\delta_S(t, S_t) \, dt$ during $[t, t + dt]$
  - Will work with the special (annualized) case

$$\delta_S(t, S_t) = \delta S_t \tag{δ}$$

- *Q:* For what types of assets S is (δ) a "good" model for dividend payouts?

Trade

- Recall main insight from binomial model: the role of a replicating portfolio
- Want to use same strategy in continous time: price derivatives via replicating portfolio argument
- Crucial to keep track of
  - total gains from trade: capital gains plus dividend gains
  - portfolios of replicating assets: stock and riskless asset (money market account)
- Consider first a "portfolio" consisting only of the stock
- Let Δ denote number of stocks held
  - generally a stochastic process
  - $\Delta_t = \#$ of stocks held at time t
    (not to be confused with the discrete time increment Δt)

Gains from trade

- Investment in stock yields capital and dividend gains during $[t, t + dt]$
  - Capital gain dS_t
  - Dividend gain $dD_t = \delta S_t \, dt$

Definition

The gains process is given by

$$dG_t = dS_t + dD_t \tag{G}$$

- Let's introduce riskless asset (money market account)

$$A_t = \mathrm{e}^{rt},$$
$$\text{iff}$$
$$dA_t = rA_t \, dt, \qquad A_0 = 1 \tag{A}$$

Portfolios

- Two equivalent interpretations of riskless asset
  1. An asset with constant price of 1, and dividend stream rA_t
  2. An asset with price A_t, and no dividend payouts
- Gains process for riskless asset coincides with A, and we refrain from using a special symbol for it—for now
  - Will later let

$$dG^{(1)} = dS + dD$$

$$dG^{(2)} = dA$$

  - In interpretation (1) $dS^{(2)} \equiv 0$ and $dD^{(2)} = dA \implies dG^{(2)} = dA$
  - In interpretation (2) $dS^{(2)} = dA$ and $dD^{(2)} \equiv 0 \implies dG^{(2)} = dA$
- Let's expand portfolio by investing θ in the riskless asset
- Get the trading strategy (Δ, θ) with profit during $[t, t + dt]$

$$\text{profit} = \underbrace{\Delta_t}_{(\#\text{ stocks})} \underbrace{dG_t}_{(\text{stock gains})} + \underbrace{\theta_t}_{(\#\text{ riskless units})} \underbrace{dA_t}_{(\text{riskless gains})}$$

Self-financing portfolios

Definition

We say a trading strategy (Δ, θ) is self financing if

$$d(\Delta_t S_t + \theta_t A_t) = \Delta_t \, dG_t + \theta_t \, dA_t$$

- The self-financing restriction is an accounting identity, keeping track of sources and destination of funds
- Self-financing property requires that all gains/losses are reinvested in the portfolio
  - No outside source of funds flowing into portfolio
  - No funds "leaking out" of portfolio

Non-self-financing portfolios

Examples

A portfolio that finances consumption rate c is *not* self-financing:

$$d(\Delta_t S_t + \theta_t A_t) = \Delta_t \, dG_t + \theta_t \, dA_t - c_t \, dt,$$

and neither is one that receives influx of oil profits at rate π

$$d(\Delta_t S_t + \theta_t A_t) = \Delta_t \, dG_t + \theta_t \, dA_t + \pi_t \, dt,$$

The role of self-financing portfolios

- *Q: Why are we interested in self-financing portfolios?*
  Hint: The portfolio in the binomial model is also self-financing
- Usefulness of self-financing portfolios rely on assumption of *no arbitrage opportunities*
- An arbitrage opportunity is a trading strategy that
  - Costs zero to set up (after normalization)
  - Has a non-zero payoff in the future with strictly positive probability
- Much stricter requirement than commonly applied "statistical arbitrage" (cf. e.g. Wikepedia)
  - E.g. hedge funds mostly rely on statistical arbitrage strategies, even when they claim they do arbitrage strategies

Arbitrage opportunities

- Principle of arbitrage is simple
  - Buy cheap
  - Sell expensive

Example

Assume riskless rate of zero. Consider a firm whose assets are a plot of land, worth 100, financed by debt and equity with a market values of 50 and 40.
Q: What should you do?

Example

Consider a stock with market price 100. Three month European call and put options with strike 110 sell at 10 and 25 respectively. The riskless rate is zero.
Q: What should you do?

Itô's Lemma

- To carry out replicating portfolio argument, we first need to know what to replicate: how the derivative's value changes across time and states
- Know that present value of European derivative is given by discounted value of $\mathrm{E}\left\{g(T, S_T)\right\}$
- Can show that
$$\mathrm{E}\left\{g(T, S_T)\right\} = f(0, S_0),$$
or more generally,
$$\mathrm{E}_t\left\{g(T, S_T)\right\} = f(t, S_t)$$
- We'll additionally assume *for now* that f is "smooth": $f \in \mathcal{C}^{1,2}$, i.e.,
    - Once differentiable in its first argument: 1
    - Twice differentiable in its second argument: 2
    - Both derivatives are continuous: $\mathcal{C}$

Itô's Lemma: background

- Define the multiplication table

	dt	dW_t
dt	0	0
dW_t	0	dt

- Let

$$
\begin{aligned}
f_1 &= \left.\frac{\partial f(t,y)}{\partial t}\right|_{y=Y_t} \\
f_2 &= \left.\frac{\partial f(t,y)}{\partial y}\right|_{y=Y_t} \\
f_{22} &= \left.\frac{\partial^2 f(t,y)}{\partial y^2}\right|_{y=Y_t}
\end{aligned}
$$

Itô's Lemma

Itô's Lemma gives us an SDE that f = price of derivative must satisfy:

Result

Let Y be the solution to (Y). If $Z_t = f(t, Y_t)$ and f is "smooth" then Z is the solution to

$$dZ_t = f_1\,dt + f_2\,dY_t + \frac{1}{2}f_{22}(dY_t)^2, \qquad Z_0 = f(0, y)$$

- To get intuition for Itô's Lemma, recall the *total* derivative

$$df(x, y) = f_1(x, y)\,dx + f_2(x, y)\,dy$$

- Itô's lemma yields the same expression with $x = t$ and $y = Z_t$, except for the additional term $\frac{1}{2}f_{22}(dY_t)^2$
- The term $\frac{1}{2}f_{22}(dY_t)^2$ is a "correction" for the "wild" behavior of the Wiener process (being nowhere differentiable)

Itô's Lemma

- Could in principle solve the SDE for the unknown price f, but difficult to solve SDEs
- Will use different strategy: replicating portfolio argument
- Let's start by getting familiar with Itô's lemma:

Example

Let $R = ln(S)$ where S satisfies our stock price model (S).
Q1: Characterize dR
Q2: Use the answer to *Q1* to show that the solution to (S) is

$$S_t = S_0 \mathrm{e}^{(\mu - \sigma^2/2)t + \sigma W_t}$$

ECO423: **The Black-Scholes Analysis and the Fundamental PDE**

Jørgen Haug

Department of Finance, NHH

Spring 2023

After this section you should

- Know how to
  - set up a self-financing replicating portfolio
  - derive a PDE for the price of a derivative; understand in what way it is "fundamental"
  - state the Feynman-Kac solution to the PDE (given access to the formula; no need to memorize it)
  - compute the Black-Scholes-Merton formula
    - indicator functions; definition, properties
    - "devide and conquer" approach
    - the standard normal distribution and its basic properties
    - basic integration techniques (will be less important if you master Monte Carlo simulation, which we cover later)
- Have economic intuition for
  - the role of "self-financing" and "no arbitrage" in pricing via a replicating portfolio
  - what the constituent parts of the Feynman-Kac solution represents

Overall aim and solution strategy

- *Aim:* Use Itô's Lemma to derive a partial differential equation (PDE) that the derivative price C_t must satisify,
  - given price models (S) and (A),
  - without introducing arbitrage opportunities
- *Idea:*
  - Use Itô's Lemma to characterize price changes of the derivative we want to price $\sim$ its diffusion $dC_t = \ldots$
  - Set up a self-financing portfolio

$$d(\Delta_t S_t + \theta_t A_t) = \Delta_t \; dG_t + \theta_t \, dA_t$$

  - Choose Δ and θ such that the self-financing portfolio becomes a replicating portfolio

$$d(\Delta_t S_t + \theta_t A_t) = dC_t \qquad \text{for } \textit{all } t \in [0, T]$$

  - No arbitrage implies $C_t = \Delta_t S_t + \theta_t A_t$ (at any date $t \in [0, T]$)

- Consider a European call option with payoff

$$g(T, S_T) = \max(S_T - K, 0)$$

where S is the diffusion:

$$dS_t = \mu S_t \, dt + \sigma S_t \, dW_t, \qquad S_0 > 0 \tag{S}$$

- Assume for now that we can express its date t present value as $C_t = c(t, S_t)$
- By Itô's Lemma we know that

$$dC_t = \left[c_1 + c_2 \mu S_t + \frac{1}{2} c_{22} (\sigma S_t)^2 \right] dt + c_2 \sigma S_t \, dW_t \tag{1}$$

- Recall our remaining assumptions about price dynamics:

$$dD_t = \delta S_t \, dt \tag{D}$$
$$dA_t = rA_t \, dt \tag{A}$$

- Changes in value of self-financing portfolio thus given by

$$d(\Delta_t S_t + \theta_t A_t) = [\Delta_t(\mu + \delta)S_t + \theta_t rA_t] \, dt + \Delta_t \sigma S_t \, dW_t \tag{2}$$

- Let's first match price changes due to changes in W—the dispersion terms $\longrightarrow$ one equation in one unknown!

$$c_2 \sigma S_t = \Delta_t \sigma S_t$$

iff

$$\Delta_t = c_2 = c_2(t, S_t)$$

The Fundamental PDE (FPDE)

- Let's next match price changes due to changes in t—the drift terms
- Given solution for Δ this is also one equation in one unknown!

$$c_1 + c_2 \mu S_t + \frac{1}{2} c_{22} (\sigma S_t)^2 = \Delta_t (\mu + \delta) S_t + \theta_t r A_t$$

iff

$$\theta_t = \frac{1}{r A_t} \left[c_1 - c_2 \delta S_t + \frac{1}{2} c_{22} \sigma^2 S_t^2 \right]$$

- No arbitrage implies that present value $c = C_t$ must satisfy restriction

$$c = \Delta S + \theta A = c_2 S + \frac{1}{rA} \left[c_1 - c_2 \delta S + \frac{1}{2} c_{22} \sigma^2 S^2 \right] A$$

iff

$$rc = c_1 + c_2 (r - \delta) S + \frac{1}{2} c_{22} \sigma^2 S^2 \tag{FPDE}$$

- The present value of the European call is the solution to (FPDE) subject to the boundary condition

$$c(T, S) = \max(S - K, 0)$$

- Notice that we didn't use any property of the call in deriving (FPDE)!
  - we used only (S), (D), and (A),
  - and the assumption $C_t = c(t, S_t) \in \mathcal{C}^{1,2}$, which we can later verify
- *Any* derivative whose payoff depends on at most t and S_t must therefore be priced according to (FPDE)—hence the *Fundamental* PDE
- Prices of derivatives other than the European call determined by using the appropriate boundary condition
  - For instance European put: $c(T, S) = \max(K - S, 0)$

Important insight (worth thinking about)

- The parameter μ appears explicitly neither in (FPDE) nor in the boundary condition!
  $\longrightarrow$ the solution to (FPDE) is independent of μ
  $\longrightarrow$ derivative's price can be expressed/computed without knowledge of a required rate of return!
    - There's an important insight from me being *very* careful in how I state this insight!

The Feynman-Kac solution

- Consider the more general PDE

$$c_1(t,x) + c_2(t,x)a(t,x) + \frac{1}{2}c_{22}(t,x)\sigma^2(t,x) = c(t,x)b(t,x) \quad (3)$$

with boundary condition

$$c(T,x) = g(x) \quad (4)$$

- For $t \in [0,T]$ the Feynman-Kac solution to (3) and (4) is

$$c(t,x) = \mathrm{E}_t\left\{ \mathrm{e}^{-\int_t^T b(u,Z_u)\,du} g(Z_T) \right\},$$

where Z satisfies the SDE

$$dZ_u = a(u,Z_u)\,du + \sigma(u,Z_u)\,dW_u, \qquad Z_t = x.$$

Applying Feynman-Kac

- To apply the Feynman-Kac solution to (FPDE) and its boundary condition we first carry out an identification step
  - $t \sim t$, $x \sim S$
  - $a(t, S) = (r - \delta)S$
  - $b(t, S) = r$
  - $\sigma(t, S) = \sigma S$
  - $g(S) = \max(S - K, 0)$
- Next, compute the constituent parts of the formula
  - Observe first that

$$\int_t^T b(u, Z_u)\, du = \int_t^T r\, du = r \int_t^T 1\, du = ru|_t^T = rT - rt$$

  and thus

$$\mathrm{e}^{-\int_t^T b(u, Z_u)\, du} = \mathrm{e}^{-r(T-t)}$$

- Next, adapt the SDE for Z to our problem

$$\begin{aligned} dZ_u &= a(u, Z_u)\, du + \sigma(u, Z_u)\, dW_u \\ &= (r - \delta) Z_u\, du + \sigma Z_u\, dW_u, \qquad Z_t = S_t \end{aligned}$$

- *Q1:* Why can we conclude that

$$Z_T = Z_t \mathrm{e}^{\left(r - \delta - \frac{\sigma^2}{2}\right)(T-t) + \sigma(W_T - W_t)}$$

  and thus that Z has a known log-normal distribution?
- *Q2:* What economic interpretation may we attach to Z?

- Substituting back into the Feynman-Kac solution yields

$$C_t = \mathsf{E}_t \left\{ \mathrm{e}^{-r(T-t)} \max(Z_T - K, 0) \right\}$$

or

$$C_0 = \mathrm{e}^{-rT} \mathsf{E} \left\{ \max(Z_T - K, 0) \right\} \tag{5}$$

where

$$Z_T = S_0 \mathrm{e}^{\left(r - \delta - \frac{\sigma^2}{2}\right) T + \sigma W_T},$$

- *Q3:* What's the economic interpretation of e^{-rT}?
- *Q4:* What's the economic interpretation of (5)?

- We can use one of two strategies to translate (5) into a *number* (which you'll ultimately need in practice)
  - use a numerical approximation method
  - analytically compute the expectation
- In either case it is useful to utilize that $W_T \stackrel{d}{=} \sqrt{T}z$, $z \sim \mathrm{N}(0,1)$
- We'll first implement the first strategy via a simple implementation of Monte Carlo Simulation
- We'll then consider the second strategy, which leads to the Black-Scholes-Merton formula

$$c(0,S) = S\mathrm{e}^{-\delta T}\mathrm{N}(d_1) - K\mathrm{e}^{-rT}\mathrm{N}(d_2)$$

$$d_1 = \frac{\ln(S/K) + (r - \delta + \sigma^2/2)T}{\sigma\sqrt{T}}$$

$$d_2 = \frac{\ln(S/K) + (r - \delta - \sigma^2/2)T}{\sigma\sqrt{T}}$$

- Given (S), (D), and (A), we *know the solution* Z_T in Feynman-Kac solution for European call
  - We'll later develop a numerical strategy for when we do not know the solution to $dZ = \dots$
- Using $W_T \stackrel{d}{=} \sqrt{T}z$, $z \sim \mathrm{N}(0, 1)$ we have

$$Z_T = S_0 \mathrm{e}^{\left(r-\delta-\frac{1}{2}\sigma^2\right)T+\sigma\sqrt{T}z}$$

- Let's generate N independent observations of z; $z^1, \dots, z^N$—called *random variates* (a.k.a. "scenarios")
- Let's next use the random variates to generate N observations of the call payoff

$$Z^n = S_0 \mathrm{e}^{\left(r-\delta-\frac{1}{2}\sigma^2\right)T+\sigma\sqrt{T}z^n}$$
$$g^n = g(T, Z^n) = \max(Z^n - K, 0)$$

- Because the random variates are *iid* the Law of Large Numbers implies

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} g^n = \mathrm{E}\left\{g(T, Z_T)\right\}$$

- Or, in practical implementations with fixed N

$$C_0 = \mathrm{e}^{-rT} \mathrm{E}\left\{g(T, Z_T)\right\} \approx \mathrm{e}^{-rT} \frac{1}{N} \sum_{n=1}^{N} g^n$$

  for N sufficiently large
  $\longrightarrow$ The call price can be computed as the average payoff, using Z as a fictitious stock price, discounted back at the riskless rate

- *Q:* What is a "sufficiently large N"?

Which in R looks like

```r
# Initialize the parameter values
S0 = 12; r = 0.02; v = 0.3; K = 13; T = 0.65;

# Number of random variates to use
N = 1000

# Generate N random variates
set.seed(1)
z = rnorm(N)

# Generate N observations of Z-process and call payoff
ZT = S0 * exp((r-v^2/2)*T + v*sqrt(T)*z)
hist(ZT) # Visual check that the simulation is implemented correctly
gPV = exp(-r*T)*pmax(ZT-K,0)

# Present value of call
c0 = mean(gPV)
```

- *Trick 1:* Let's simplify notation

$$Z_T = S\mathrm{e}^{a+bz}$$

- *Trick 2:* Let's devide and conquer
  - Introduce the indicator function

$$1_A = \begin{cases} 1 & \text{if } A \text{ is true} \\ 0 & \textit{otherwise} \end{cases}$$

  - Can then write

$$\max(Z - K, 0) = Z1_{\{Z \geq K\}} - K1_{\{Z \geq K\}}$$

  - and thus (with $S_0 = S$ for simplicity)

$$\begin{aligned} C_0 &= \mathrm{e}^{-rT}\mathsf{E}\left\{Z_T 1_{\{Z_T \geq K\}}\right\} - \mathrm{e}^{-rT}K\mathsf{E}\left\{1_{\{Z_T \geq K\}}\right\} \\ &= \mathrm{e}^{-rT}(C_0^1 - KC_0^2) \end{aligned}$$

- Compute C_0^2 by observing that $\mathrm{E}\left\{1_A\right\} = \Pr(A)$, because

$$\begin{aligned}\mathrm{E}\left\{1_A\right\} &= \mathrm{E}\left\{1_A | A\right\} \Pr(A) + \mathrm{E}\left\{1_A | A^{\complement}\right\} \Pr(A^{\complement}) \\ &= 1\Pr(A) + 0\Pr(A^{\complement}) = \Pr(A)\end{aligned}$$

- Compute C_0^1 by evaluating the integral

$$\int_{-\infty}^{\infty} \mathrm{e}^{a+bz} 1_{\{S\mathrm{e}^{a+bz} \geq K\}} \mathrm{n}(z)\, dz$$

where

$$n(z) = \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-z^2/2}$$

- We'll carry out both computations on the White Board

ECO423: **Monte Carlo Simulation**

Jørgen Haug

Department of Finance, NHH

Spring 2023

After this section you should

- Know how to
  - construct an Euler scheme for a diffusion
  - compute an Euler scheme (by hand and on a computer)
  - carry out a Monte Carlo simulation of the Feynman-Kac solution, with an Euler scheme for the diffusion
  - construct antithetic variates
- Have (economic) intuition for
  - possible issues with discretization schemes; may change the model
  - the principle behind generating random variates from $U(0,1)$
  - what antithetic variates are
  - why antithetic variates increases quality/efficiency of Monte Carlo simulations

Recap: The Feynman-Kac solution

- Recall the Feynman-Kac solution

$$c(t,x) = \mathrm{E}_t \left\{ \mathrm{e}^{-\int_t^T b(u,Z_u)\,du} g(Z_T) \right\}, \tag{1}$$

where Z satisfies the SDE

$$dZ_u = a(u,Z_u)\,du + \sigma(u,Z_u)\,dW_u, \qquad Z_t = x. \tag{Z}$$

- We know how to approximate (1) via Monte Carlo simulation when we *know the solution* to (Z)
- Let's consider how to adapt Monte Carlo simulation when we *know only the SDE* (Z), and not its solution

Recap: Basic idea of Monte Carlo simulation

- We want to utilize the Law of Large Numbers (LLN) as before
  - Generate N *iid* observations of possible payoffs

$$g^n = g(T, Z_T^n)$$

  - Approximate the present value expression by

$$C_0 = \mathrm{e}^{-rT}\mathsf{E}\left\{g(T, Z_T)\right\} \approx \mathrm{e}^{-rT}\frac{1}{N}\sum_{n=1}^{N} g^n \tag{2}$$

  - LLN assures us the the right hand side of (2) converges to the left hand side as N increases
  - N "large" enough when the appropriate standard error is economically small (which depends on the problem at hand)
- Remains to determine Z_T^n!

Discretized SDEs

- *Idea:* Return to the discrete-time motivation for SDEs, and replace dt by Δt
- Observe first that we can think of the *generic diffusion* (Z) as

$$dZ_t = Z_{t+dt} - Z_t = a(t, Z_t)\, dt + \sigma(t, Z_t)(W_{t+dt} - W_t)$$

  and hence

$$\Delta Z_t = Z_{t+\Delta t} - Z_t = a(t, Z_t)\Delta t + \sigma(t, Z_t)(W_{t+\Delta t} - W_t) \tag{3}$$

- Let's simplify notation by enumerating the dates

$$0 = 0\Delta t, \Delta t, 2\Delta t, \ldots, M\Delta t = T$$

  by

$$m = 0, 1, 2, \ldots, M$$

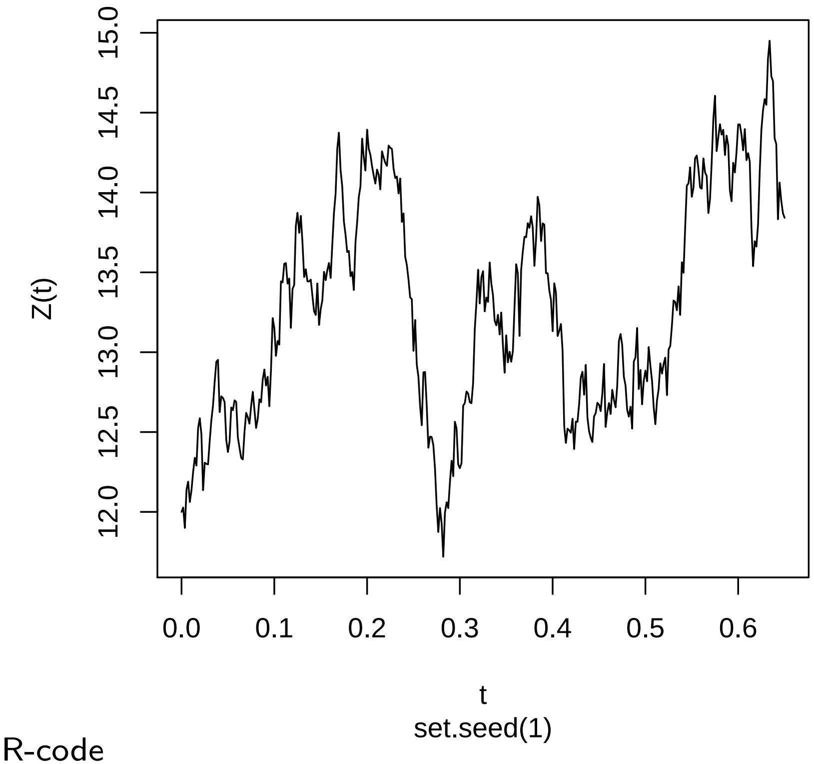
The Euler scheme

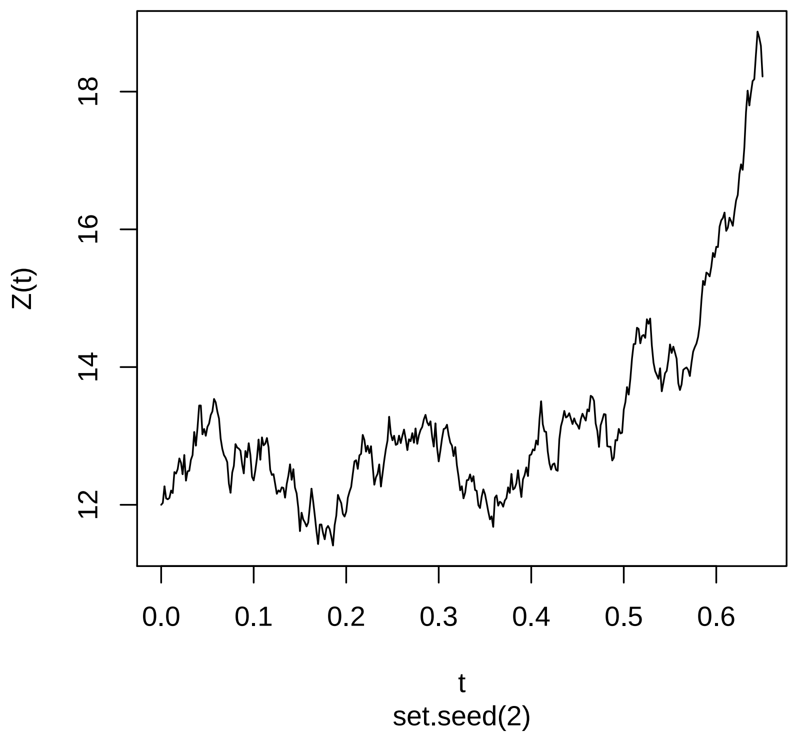
- We can solve (3) for $Z_{t+\Delta t}$, use our enumerated dates, and $\Delta W_t \stackrel{d}{=} \sqrt{\Delta t} z$, $z \sim \mathrm{N}(0,1)$ to arrive at

$$Z_{m+1} = Z_m + a(m, Z_m)\Delta t + \sigma(m, Z_m)\sqrt{\Delta t} z_{m+1} \tag{ES}$$

  where the $z_m \sim \mathrm{N}(0,1)$ are *independent*, $m = 1, \ldots, M$
- For given sequence $\{z_1, \ldots, z_M\}$ we can find Z_T by forward induction
  - We know the initial condition $Z_0 = S_0$
  - Use Z_0 to find $Z_1 = Z_0 + a(0, Z_0)\Delta t + \sigma(0, Z_0)\sqrt{\Delta t} z_1$
  - Use Z_1 to find $Z_2 = Z_1 + a(1, Z_1)\Delta t + \sigma(1, Z_1)\sqrt{\Delta t} z_2$
  - $\ldots$ Use Z_{M-1} to find $Z_T = Z_{M\Delta t} = Z_M = Z_{M-1} + \ldots$
- We call (ES) the (first order) Euler scheme for (Z)

Example of Euler scheme: Geometric Wiener process





R-code

```
S0 = 12; r = 0.02; v = 0.3; T = 0.65;
M = 365; dt = T/(M-1); # 'M' no longer = no. days
set.seed(1); u = rnorm(M+1)
Z = S0
for(m in 1:(M-1))
  Z = c(Z, Z[m] + r*Z[m]*dt + v*Z[m]*sqrt(dt)*u[m+1]);

t = seq(from = 0, to = T, by = dt)
plot(t, Z, type="l", xlab="t", ylab="Z(t)", sub="set.seed(1)")
```

Use of (ES) in Monte Carlo simulation

- Monte Carlo simulation with the Euler scheme:
  1. Generate N sequences $z_1^n, \ldots, z_M^n$, each random variate z_m^n independent $\mathrm{N}(0, 1)$
  2. For each sequence, use (ES) to compute $Z_M^n = Z_T^n$
  3. Compute $g^n = g(T, Z_T^n)$, and proceed with (2) as before
- Monte Carlo simulation
  + Easy to understand and implement
  + Very versatile; can help solve most derivatives problems
  − Computationally costly
    Remember: no need to simulate path when solution to SDE is known—e.g. geometric Wiener process
  − Greater flexibility comes with higher level of complexity—as we'll see when we apply it to American derivatives
    (this drawback applies to all numerical methods)

Example: MC with Euler scheme in R

R-code

```
S0 = 12; r = 0.02; v = 0.3; K = 13; T = 0.65;
N = 1000; M = 100; dt = T/(M-1);

set.seed(1)
g = c() # ensure 'g' is an empty vector
for(n in 1:N) {
   Z = S0;
   z = rnorm(M-1)
   for(m in 1:(M-1))
      Z = c(Z, Z[m] + r*Z[m]*dt + v*Z[m]*sqrt(dt)*z[m]);
      # index of 'noise term' z[m] is m rather than m+1 for
      # coding convenience
   g = c(g, exp(-r*T)*max(K-Z[M], 0));
}

pSE = sd(g)/sqrt(N)
p0 = mean(g); p0
p0 + qnorm(c(0.025, 0.975))*pSE
```

Output

```
> p0 = mean(g); p0
 [1] 1.689907
> p0 + qnorm(c(0.025, 0.975))*pSE
 [1] 1.579786 1.800028
```

Example: More efficient vectorized code

R-code

```
S0 = 12; r = 0.02; v = 0.3; K = 13; T = 0.65;
N = 1000; M = 100; dt = T/(M-1);

set.seed(1)
Z = matrix(c(rep(S0,N)), ncol=1)
z = matrix(rnorm((M-1)*N), nrow=N, byrow=TRUE)
for(m in 1:(M-1))
  Z = cbind(Z, Z[,m] + r*Z[,m]*dt + v*Z[,m]*sqrt(dt)*z[,m])
g = exp(-r*T)*pmax(K-Z[,M], 0)

pSE = sd(g)/sqrt(N)
p0 = mean(g)
p0 = mean(g); p0
p0 + qnorm(c(0.025, 0.975))*pSE
```

Output

```
> p0 = mean(g); p0
[1] 1.689907
> p0 + qnorm(c(0.025, 0.975))*pSE
[1] 1.579786 1.800028
```

Did we change the model?

- Euler scheme an intuitive and simple special case of discretization schemes
  - There are higher order Euler schemes
  - The Milstein scheme is "better" for mean reverting processes
  - ... and more. Glasserman (2004) is a comprehensive reference.
- *General* problem with discretization schemes is that the discretization may change the model
- We have discussed economic rationale for liking the geometric Wiener process
  - Limited liability; $\Pr(S_t \geq 0) = 1$ whenever $S_0 > 0$
- The mean-reverting Ornstein-Uhlenbeck process
  - Allows stochastic interest/growth rates with "limited" (i.e., stationary) variation

Examples of pitfalls with the Euler scheme

- No general theory for how discretization affects a model
- Will illustrate through two examples
  - Geometric Wiener process in class
  - Ornstein-Uhlenbeck in Voluntary Exercise
- Consider the Geometric Wiener process

$$S_1 = S_0 + \mu S_0 \Delta t + \sigma S_0 \sqrt{\Delta t} z_1$$

- *Q:* What's the distribution of S_1?
- Let's investigate a simple example in Excel...

S0	10		t	W(t)	S(t)
mu	0.15		1	0.1	11.9
sigma	0.4		2	-2.95	
T	2		3	1.5	
N	2		4	0.73	
dt	1		5	0.45	

Example of "fix"

- Recall that $R_t = \ln(S_t)$ implies that

$$dR_t = \left(\mu - \frac{\sigma^2}{2}\right) dt + \sigma\, dW_t \tag{4}$$

- Corresponding Euler scheme

$$R_{m+1} = R_m + \left(\mu - \frac{\sigma^2}{2}\right) \Delta t + \sigma \sqrt{\Delta t} z_{m+1} \tag{5}$$

- *Q:* Is the distribution of (5) "consistent" with the distribution of (4)?
- Use definition of R to express

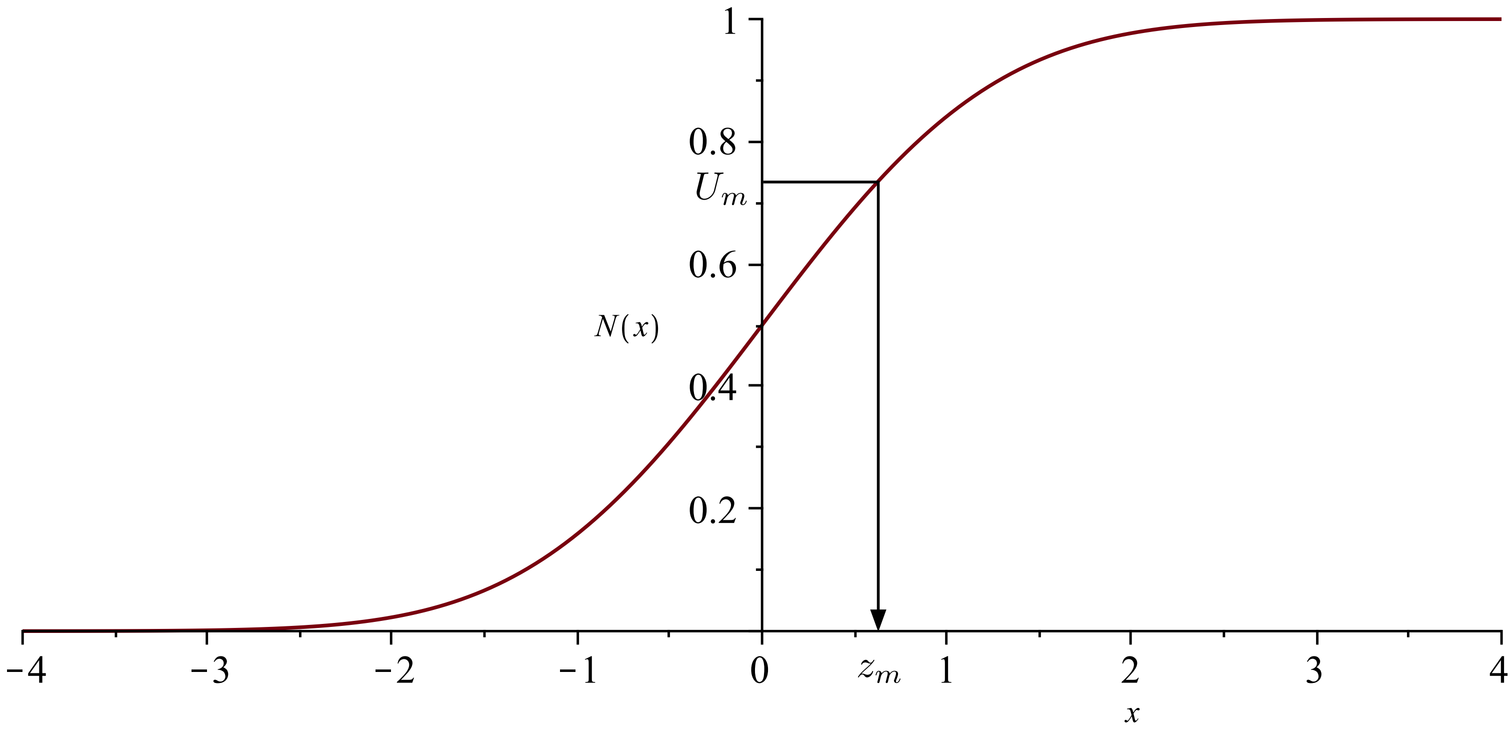
$$S_m = \mathrm{e}^{R_m}$$

- *Q:* Why does this solve our problem?

Objective

- Setting $N = M = 1000$ in R code on slide 7 illustrates weakness of Monte Carlo simulation
- Would be great to reduce number of computations, while retaining accuracy!
  $\longrightarrow$ increase numerical efficiency
- Antithetic variates is a simple approach to increase efficiency
- To understand the technique, let's first consider how we can generate random variates from arbitrary distributions

- To generate a sequence $\{z_1, \ldots, z_M\}$, $z_m \sim F$ we
  - draw $U_m \sim \mathrm{U}(0,1)$ (as in Project 2)
  - *Q:* What's common between U_m and *any* cumulative distribution function (CDF) F?
  - use F^{-1} to transform U_m into z_m



- Any decent software offers selection of inverse CDFs

Examples: Excel and R

- To implement

$$z_m = N^{-1}(U_m)$$

  in Excel

  - Use RAND() to generate $U_m \sim \mathrm{U}(0, 1)$
  - Use NORM.S.INV() to implement $N^{-1}()$
  - Generate z_m by

```
NORM.S.INV(RAND())
```

- To achieve the same in R

```
qnorm(runif(1,0,1))
```

  but of course easier to simply use

```
rnorm(1)
```

Antithetic variates

- Define $\tilde{U}_m = 1 - U_m \sim \mathrm{U}(0,1)$
  - $\tilde{z}_m = F^{-1}(\tilde{U}_m)$ yields sequence of *iid* F-distributed random variates $\tilde{z}_1, \ldots, \tilde{z}_M$, called antithetic variates
  - If density $f(x) = F'(x)$ is symmetric around 0, like e.g. the standard normal $n(x) = \mathrm{N}'(x)$ then

$$\tilde{z}_m = -z_m$$

- Let now

$$\begin{aligned}
g_T^n &= \text{payoff from sequence } z_1^n, \ldots, z_M^n \\
\tilde{g}_T^n &= \text{payoff from sequence } \tilde{z}_1^n, \ldots, \tilde{z}_M^n \\
\bar{g}_T^n &= \frac{1}{2}(g_T^n + \tilde{g}_T^n)
\end{aligned}$$

- By the LLN

$$\frac{1}{N}\sum_{n=1}^{N}\bar{g}_T^n \xrightarrow[N\to\infty]{} \mathrm{E}\left\{g(T,Z_T)\right\}$$

  - *But* z_m and $\tilde{z}_m$ are not independent!
  - *Q:* How come LLN works out?

- Turns out that standard error of

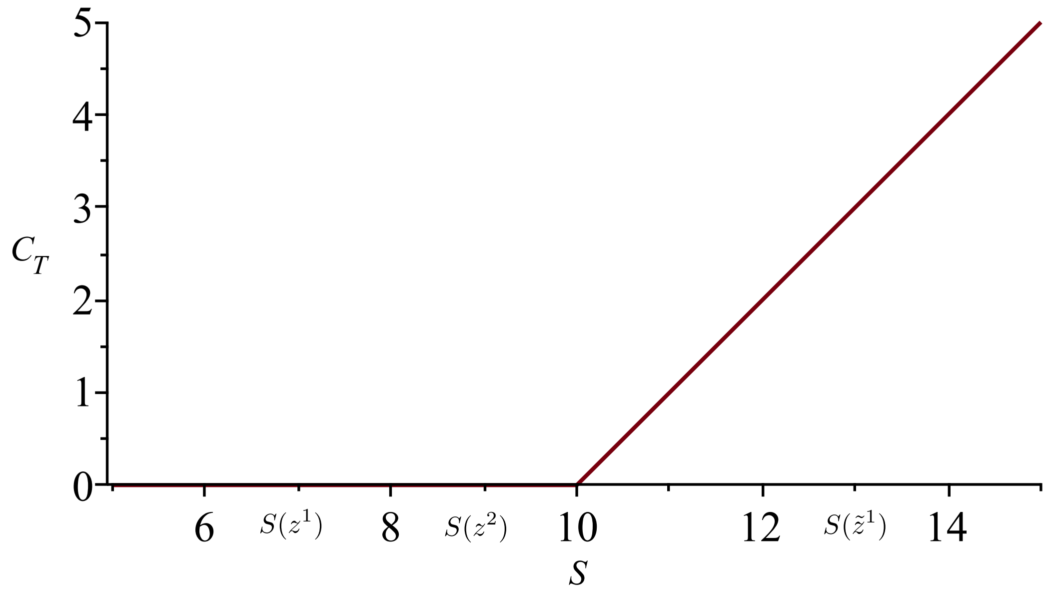
$$\frac{1}{N}\sum_{n=1}^{N}\bar{g}_T^n$$

can be significantly smaller than standard error of

$$\frac{1}{2N}\sum_{n=1}^{2N}g_T^n$$

Intuition for smaller variance

- Consider simple example



- Using z^1 and z^2 we get $c_0 \approx 0$
- Using z^1 and $\tilde{z}^1$ we get $c_0 > 0$
- Antithetic variates ensures more "representative" sampling with same number of total random variates

Intuition for smaller variance

- We can formalize the above intuition
- Consider again two scenarios z^1 and z^2 (with $z^2 = \tilde{z}^1$ for antithetic variates)
- Our estimated (future) value is

$$\hat{c}_0 \mathrm{e}^{rT} = \frac{g^1 + g^2}{2}$$

with variance

$$\mathrm{Var}\left(\hat{c}_0 \mathrm{e}^{rT}\right) = \frac{\mathrm{Var}\left(g^1\right) + \mathrm{Var}\left(g^2\right) + 2\mathrm{Cov}\left(g^1, g^2\right)}{4}$$

  - Standard approach: $\mathrm{Cov}\left(g^1, g^2\right) = 0$
  - Antithetic variates: $\mathrm{Cov}\left(g^1, g^2\right) < 0$

Glasserman, Paul, *Monte Carlo Methods in Financial Engineering*, Springer-Verlag, 2004.

ECO423: **Equivalent Martingale Measures**

Jørgen Haug

Department of Finance, NHH

Spring 2023

After this section you should

- Know
  - the role of the Fundamental Theorem of Asset Pricing
  - how to derive the diffusion for discounted gains processes
  - how to change probabilities from P to Q via the Wiener process, consistent with no arbitrage
  - how to use the martingale property of discounted gains to express the present value of a derivative as a Q-expectation
- Have intuition for
  - the role of "equivalent" and "martingale" in "equivalent martingale measure"
  - how we adjust for systematic risk

Motivation

- We have developed the Feynman-Kac solution to price derivatives
- We have so far worked with physical probabilities, that we'll hereafter carefully denote P
- Will now construct artificial, *risk-adjusted* probabilities, that we'll denote Q (a.k.a. "pseudo" or "risk-neutral" probabilities)
- Why develop another approach?
  - Faster to work with
  - Allows more general results; can price more types of derivatives (no need to assume $c \in \mathcal{C}^{1,2}$)
  - Offers economic insight into the Z "price" diffusion in Feynman-Kac
  - It's part of the "industry standard"

Discounted gains

- Choose a security to discount prices with—the numeraire/deflator
- Deflator must
  - be strictly positive
  - not leave dividends unaccounted for
- $A_t = \mathrm{e}^{rt}$ is the most popular choice
- For instance in the Black-Scholes economy

$$dA_t^* = d\frac{A_t}{A_t} \equiv 0 \tag{A^*}$$

$$dS_t^* = d\frac{S_t}{A_t} \tag{S^*}$$

$$dD_t^* = \frac{\delta S_t}{A_t}\,dt \tag{D^*}$$

$$dG_t^* = dS_t^* + dD_t^* \tag{G^*}$$

- If more securities in the economy, discount in similar fashion

The Equivalent Martingale Measure

- Let's assign one gains process to each asset under consideration, $G^1 = A,\ G^2 = S + D,\ \dots,\ G^N = \cdots$
- *Definition:* We say that Q is an equivalent martingale measure relative to P if (i) $Q(A) = 0$ iff $P(A) = 0$, and (ii) if

$$G_t^{n*} = \mathrm{E}_t^Q \left\{ G_s^{n*} \right\} \text{ for all } s \geq t \tag{EMM}$$

  for *all* assets in the economy, $n = 1, \dots, N$
  - (i) is definition of Q and P being equivalent
  - (ii) is the definition of a martingale
  - Q and P measures the probability of events A
- In the Black-Scholes economy we typically have three assets, the third being the (call) option $G^3 = C$
- As before, we typically simplify notation by reserving the gains process notation for the underlying asset, $G = S + D$

Intuition for EMM concepts

- Equivalence ensures that we do not introduce new arbitrage opportunities when changing from P to Q
  - *Intuition:* One-period binomial model
  - $p = 0.5$, $H = 220$, $L = 0$, $\mu = 1.1$
  - $q = 0$, $R = 1.05$
  - Q: What's the present value/ "quoted price" under p and q?
- The martingale property is simply a consistency condition
  - *Intuition:* "The most important expression in finance"

$$S_t = \frac{\mathrm{E}_t^P \{S_T\}}{\mu^{T-t}}$$

  - $\longrightarrow$ discounted prices should always be martingales; using *consistent combinations of probabilities and discount rates*

The Fundamental Theorem of Asset Pricing

- Remains to determine (a) when an EMM Q exists, (b) how it is useful, and (c) how we can determine Q
- (a) is answered by *the* most important result in asset pricing
- *The Fundamental Theorem of Asset Pricing:* There are no arbitrage opportunities iff there exists an equivalent martingale measure Q
  - (If moreover markets are "complete", then Q is unique)
  - (We always treat Q as unique in this course)
- To answer (b) we'll consider two examples that illustrate how Q can be used to compute present values!

Example: European call

- Assume we know an EMM Q
- Consider $C_T = g(S_T) = \max(S_T - K, 0)$
- Recall that there are no arbitrage opportunities iff all discounted gains are Q-martingales
- The gains process for the call is

$$C_T^* = \frac{g(S_T)}{A_T}$$

- From (EMM) we must have that

$$C_t^* = \mathrm{E}_t^Q \left\{ C_T^* \right\},$$

Example: European call

- which can be expressed as

$$C_t = \mathrm{E}_t^Q \left\{ \frac{A_t}{A_T} C_T \right\}$$

- Using that $A_t = \mathrm{e}^{rt}$ we can rewrite as

$$C_t = \mathrm{e}^{-r(T-t)} \mathrm{E}_t^Q \left\{ \max(S_T - K, 0) \right\} \tag{1}$$

- Since we know Q we can compute (1) $\longrightarrow$ Job (almost) done!

Example: Riskless zero coupon bond

- Payoff

$$g(t, S) = g(t) = \begin{cases} 1, & t = T \\ 0, & \text{otherwise} \end{cases}$$

- Know date T value

$$P(T, T) = 1$$

and want to find date t value $P(t, T)$

- No arbitrage implies (EMM) must hold

$$P^*(t, T) = \mathrm{E}_t^Q \{P^*(T, T)\}$$

iff

$$P(t, T) = \mathrm{e}^{-r(T-t)}$$

Girsanov's Theorem

- Changing from P to Q necessarily changes distribution of W $\longrightarrow$ no longer a standard Wiener process under Q!
- *Girsanov's Theorem:* For (almost) any λ_t

$$d\tilde{W}_t = dW_t + \lambda_t \, dt$$

  is a standard Wiener process relative to *some* appropriately chosen probability measure Q
- How shall we choose Q, and thus answer (c) on page 7?
  - To ensure all discounted gains are Q-martingales!
- *Useful Result:* The diffusion (Y), with W a standard Wiener process relative to some P, is a P-martingale iff $\mu_Y(t, Y_t) \equiv 0$
- The probability measure P in Useful Result is generic; applies also to the Q we're searching for

Idea to determine Q

1. Derive dynamics $dG^* = dS^* + dD^*$ using Itô's Lemma
2. Switch from dW to $d\tilde{W}$
3. Choose λ_t to ensure the drift term $\equiv 0$

Example: Black-Scholes economy

- Assume prices and dividends are given by (A), (S), and (D)
- $S_t^* = \frac{S_t}{A_t} = \mathrm{e}^{-rt} S_t \implies dS_t^* = \cdots$
- $dD_t^* = \cdots$

$$\ldots \implies dG_t^* = (\mu + \delta - r - \sigma\lambda_t) S^* \, dt + \sigma S_t^* \, d\tilde{W}_t \tag{2}$$

- G^* in (2) is a diffusion under Q because of the presence of $\tilde{W}_t$
- To prevent arbitrage G^* must be a Q-martingale, iff

$$(\mu + \delta - r - \sigma\lambda_t) S^* = 0$$

iff

$$\lambda_t = \lambda = \frac{\mu + \delta - r}{\sigma}$$

which nails down $\tilde{W}$ and thus (implicitly) Q!

- Can now use $\tilde{W}$ to risk-adjust S

$$\begin{aligned} dS_t &= \mu S_t\,dt + \sigma S_t\,dW_t \\ &= \mu S_t\,dt + \sigma S_t\,(d\tilde{W}_t - \lambda\,dt) \\ &= (\mu - \sigma\lambda) S_t\,dt + \sigma S_t\,d\tilde{W}_t \end{aligned}$$

- i.e., S under Q is given by

$$dS_t = (r - \delta) S_t\,dt + \sigma S_t\,d\tilde{W}_t \tag{3}$$

- *Notice:* We flag that a diffusion is stated under P by using the standard Wiener process relative to P—W in our case, and under Q by using the standard Wiener process relative to Q—$\tilde{W}$ in our case
- We do not work directly with P and Q, but rather indirectly through our choice of Wiener process

Example: European call option

- Recall that we can write

$$C_0 = \mathrm{e}^{-rT} \mathrm{E}_0^Q \left\{ \max(S_T - K, 0) \right\}$$

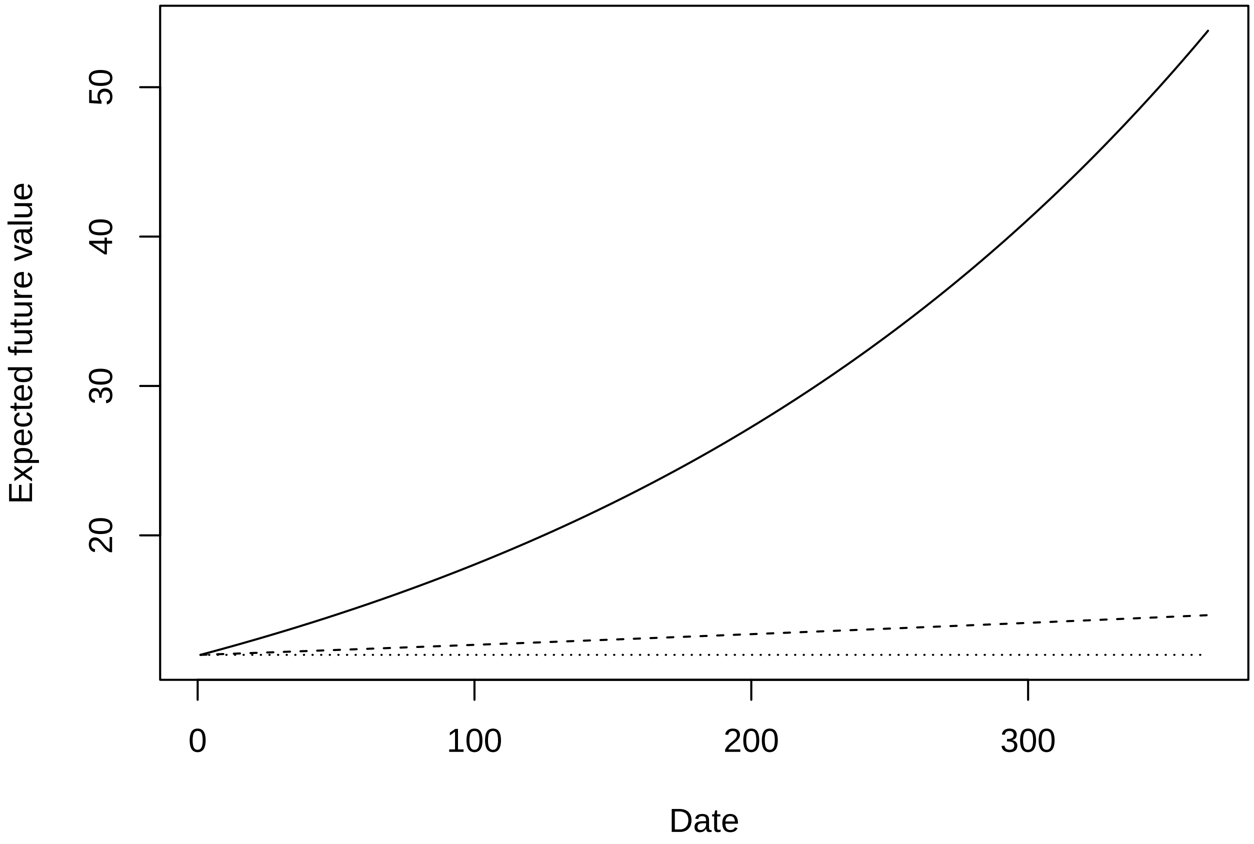
Btw: $\mathrm{E}_0 \{\cdot\}$ is the same as $\mathrm{E} \{\cdot\}$ for our purposes

- The problem is thus to compute the above expectation subject to (3)
- We recognize (3) as a geometric Wiener process, and thus

$$S_T = S_0 \mathrm{e}^{\left(r-\delta-\frac{\sigma^2}{2}\right)T+\sigma \tilde{W}_T}$$

- Problem is identical to the Feynman-Kac problem we solved earlier!
  - Don't get "misled" by the Q-expectation; it simply states that $\tilde{W}$ is a standard Wiener process
  - Can use $\tilde{W}_T \stackrel{d}{=} \sqrt{T} z$, $z \sim \mathrm{N}(0, 1)$, and proceed as before

Intuition for why G^* becomes a Q-martingale:



R-code

```
S0 = 12; mu = 0.15; v = 0.3; r = 0.02; T = 10.0; M = 365; dt = T/M

EPG = function(t) return(S0*exp(mu*t))
EQG = function(t) return(S0*exp(r*t))
EQGstar = function(t) return(S0*exp(r*t)/exp(r*t))

dates = seq(from = 0, to = T, by = dt)
plot(EPG(dates), type = "l", xlab="Date", ylab="Expected future value")
lines(EQG(dates), lty=2); lines(EQGstar(dates), lty=3)
```

Let's solve a problem

- Exam 2013, Problem 1 a) and c)
  - In c) consider a constant payoff M rather than the payoff S_T, when $S_T^2 \geq k$
- Try to solve it yourself before class
- *Hint:* c) can be solved without any complex/fancy calculations
- a) Contains an important economic insight...
- Feel free to try to solve b) and the original version of c) too (although I will not go through it in class)
  - b) is an exercise in Itô's lemma that illustrates how you can work with S rather than A as numeraire
  - c) must then be solved by treating W^S in b) as a standard Wiener process

- There are no arbitrage opportunities iff all discounted gains processes are Q-martingales
- A diffusion is a martingale iff its drift term $\equiv 0$
- Girsanov: $d\tilde{W}_t = dW_t + \lambda_t\, dt$ a standard Wiener process for some Q
- Valuation strategy:
  1. Derive diffusions for discounted gains using Itô's lemma
  2. Change from W_t to $\tilde{W}_t$ in diffusion for discounted gains
  3. Select λ_t such that drift $\equiv 0$
  4. Derive diffusion for underlying asset using $\tilde{W}_t$, and treat $\tilde{W}_t$ as a standard Wiener process
  5. Express present value of derivative using that its discounted gains process is also a Q-martingale
  6. Evaluate Q-expectation: analytically or numerically (e.g. Monte Carlo)

ECO423: **The Greeks**

Jørgen Haug

Department of Finance, NHH

Spring 2023

After this section you should

- Know
  - the generic definition of a Greek as a partial derivative
  - the particular definitions of Delta, Gamma, Vega, Theta, Rho (no need to memorize formulas)
  - how to compute Greeks numerically and analytically
  - how to make a portfolio delta and gamma neutral
  - how to compute the beta of a derivative
- Have intuition for
  - the economic interpretation of a Greek
  - the mathematical interpretation as a local measure
  - the possible pitfalls of using Greeks
  - why derivatives (options etc) are typically more "risky" than their underlying assets

Dynamic hedging

Definition

Hedging refers to investing in the underlying asset, or assets that correlate with the underlying, to manage the risk of future cash flows.

- Let n_t^j be number of contract j, $j = 1, \ldots, J$ in portfolio at date t
  - Date T value = net portfolio exposure

$$\Pi_T = \sum_{j=1}^{J} n_T^j v^j(T, S_T)$$

  - Date t value $\Pi_t = \sum_{j=1}^{J} n_t^j v^j(t, S_t)$
- Perfect markets: hedging j-by-j equivalent to hedging net portfolio exposure
- Transactions costs: j-by-j more expensive to hedge than net portfolio exposure

Greeks: Definition

- Portfolio's sensitivity to its 'input parameters' turns out to be important
- The Greeks are partial derivatives of Π wrt. p—i.e. *ceteris paribus*

$$\frac{\partial \Pi_t}{\partial p} = \frac{\partial}{\partial p} \sum_{j=1}^{J} n_t^j v^j(t, S, \ldots) = \sum_{j=1}^{j} \frac{\partial}{\partial p} \left[n_t^j v^j(t, S, \ldots) \right]$$

$\longrightarrow$ the 'Greek' of a portfolio is the portfolio of Greeks; and if n_t^j independent of p—typically assumed in computing Greeks,

$$\frac{\partial \Pi_t}{\partial p} = \sum_{j=1}^{j} n_t^j \frac{\partial}{\partial p} v^j(t, S, \ldots)$$

- $v^j(t, S, \ldots)$ time dependent makes the Greeks time dependent

Greeks: Examples

- Delta

$$\Delta_t = \frac{\partial \Pi_t}{\partial S}$$

- Gamma

$$\Gamma_t = \frac{\partial \Delta_t}{\partial S} = \frac{\partial^2 \Pi_t}{\partial S^2}$$

- Vega

$$\nu_t = \frac{\partial \Pi_t}{\partial \sigma}$$

  - *Q:* Consistent with Black-Scholes model (S)?

Greeks: Examples cont'd

- Theta (with one expiration date, time to maturity $T - t$)

$$\Theta_t = \frac{\partial \Pi_t}{\partial t} = -\frac{\partial \Pi_t}{\partial T}$$

- Rho

$$\rho_t = \frac{\partial \Pi_t}{\partial r}$$

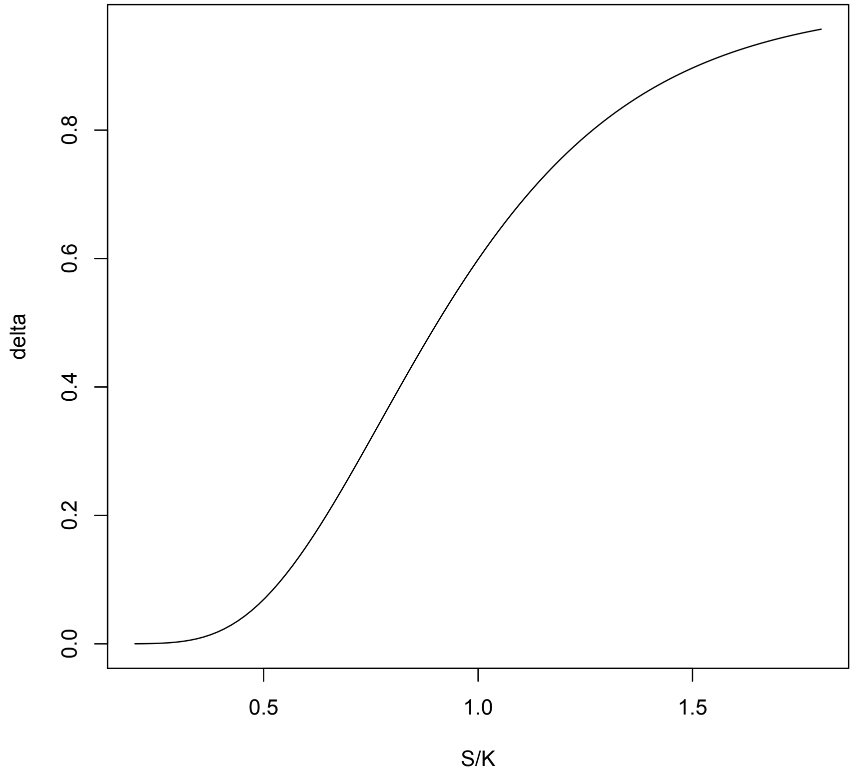
- ...and many, many more $\longrightarrow$ Wikipedia!
- Crucial insight to use Greeks appropriately:
  - Greeks are partial derivatives
  - Partial derivatives are *ceteris paribus*
  - The real economy doesn't work according to *ceteris paribus*!

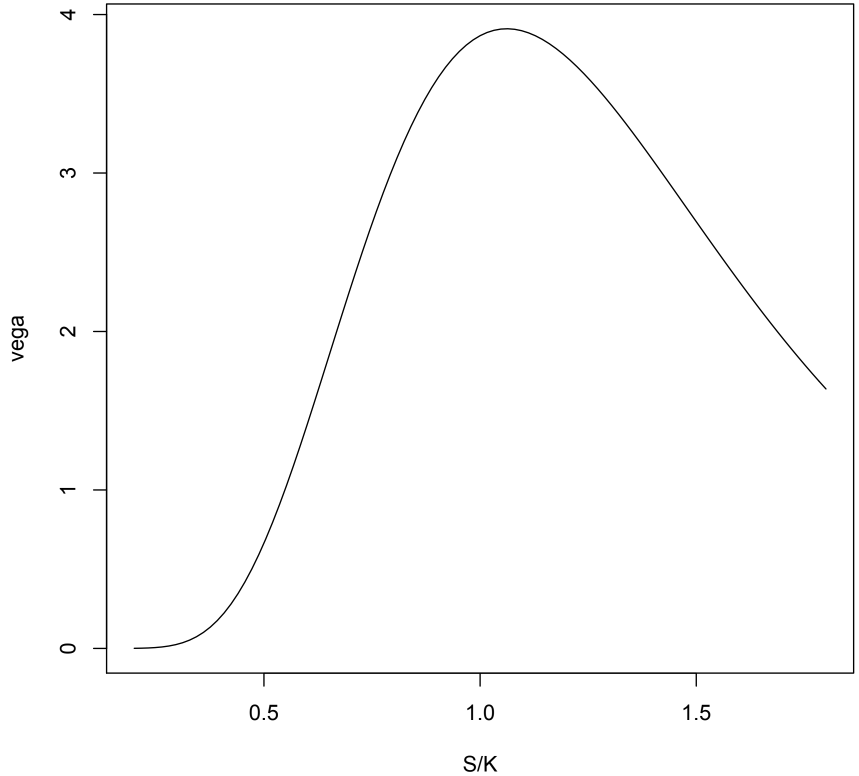
Analytic Greeks

- Standard rules of partial differentiation for contracts with known formulae
- For BSM European call, with $\mathrm{n}(d_1) = \mathrm{N}'(d_1)$, $\tau = T - t$

$$\begin{aligned}
\Delta_t &= \mathrm{e}^{-\delta\tau}\mathrm{N}(d_1)|_{\delta=0} = \mathrm{N}(d_1) \in (0,1) \\
\Gamma_t &= \frac{\mathrm{e}^{-\delta\tau}\mathrm{n}(d_1)}{S\sigma\sqrt{\tau}} > 0 \\
\nu_t &= S\mathrm{e}^{-\delta\tau}\mathrm{n}(d_1)\sqrt{\tau} > 0 \\
\Theta_t &= \delta S\mathrm{e}^{-\delta\tau}\mathrm{N}(d_1) - S\mathrm{e}^{-\delta\tau}\mathrm{n}(d_1)\frac{\sigma}{2\sqrt{\tau}} - rK\mathrm{e}^{-r\tau}\mathrm{N}(d_2)|_{\delta=0} \\
&= -S\mathrm{n}(d_1)\frac{\sigma}{2\sqrt{\tau}} - rK\mathrm{e}^{-r\tau}\mathrm{N}(d_2) < 0
\end{aligned}$$

Delta and Vega for European call






```
delta = function(x,r,sigma,T)
{
  d = (log(x)+(r+sigma^2/2)*T)/sigma
      /sqrt(T)
  return(pnorm(d))
}
curve(delta(x,.02,.4,1.0), 0.2, 1.8,
     xlab="S/K", ylab="delta")
```

```
vega = function(x,K,r,sigma,T)
{
  d = (log(x)+(r+sigma^2/2)*T)/sigma
      /sqrt(T)
  return(x*K*dnorm(d)*sqrt(T))
}
curve(vega(x,10,.02,.4,1.0), 0.2,
     1.8, xlab="S/K", ylab="vega")
```

Analytic Greeks cont'd

- Consider a call option on a non-dividend paying stock
- *Q:* Does $\Theta < 0$ imply that you should always accept an extension of the expiration date?
- *Q:* Does $\nu > 0$ imply that an increase in volatility always increases the value of the contract?
- Hull (2022): "…the BSM differential equation […] does not involve any variables that are affected by the risk preferences of investors."
  - Correct interpretation for pricing ("static" exercise)
  - Misleading/dangerous interpretation for understanding how prices respond to changes in risk ("dynamic" exercise)
- S is a "sufficient statistic" for pricing (but not for risk management)

Numerical Greeks

- May be hard or impossible to compute Greeks analytically, even when valuation formula is available
- General numerical approach
  - Recall definition of derivative
  $$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$$
  - Suggests trying
  $$f'(x) \approx \frac{f(x+h) - f(x)}{h}$$
  for given, *economically* small h
- For partial derivatives
$$\frac{\partial f(x,y)}{\partial x} \approx \frac{f(x+h,y) - f(x,y)}{h}$$

Example I

- How good are numerical approximations?
- Consider European call in Black-Scholes economy (you should do put)
- Analytic delta

```
S = 100; K = 90; r = 0.05; v = 0.4; T = 0.5;
d = (log(S/K)+(r+v^2/2)*T)/v/sqrt(T)
N1 = pnorm(d); N1
```

yields 0.7265179

- Numerical approximation

```
N2 = pnorm(d-sigma*sqrt(T))
c = S*N1-K*exp(-r*T)*N2

h = 1
dh = (log((S+h)/K)+(r+v^2/2)*T)/v/sqrt(T)
N1h = pnorm(dh)
N2h = pnorm(dh-sigma*sqrt(T))
ch = (S+h)*N1h-K*exp(-r*T)*N2h

(ch-c)/h
```

yields 0.732339

Example I cont'd

- *Q:* how much better with $h = 0.01$?
- *Q:* Do numerical approximation perform better or worse for at-the-money and in-the-money calls?
- *Q:* Modify the preceding code to compute analytic and numeric delta for a put

Example II

- Financial firm has long and short positions with clients
  - Long: firm has bought right to exercise from client
  - Short: firm has sold right to exercise to client
- Assume options portfolio has net exposure similar to short position in Voluntary Exercise 3
  - $-\max(S_{t_i} - k_i, 0)$
  - $r = 0.05,\ S_0 = 100,\ \mu = 0.2,\ \sigma = 0.4$
  - $t_1 = 0.5,\ t_2 = 1.0,\ k_1 = 85,\ k_2 = 95$

  $\longrightarrow$ Net portfolio $\sim$ short two European calls
- Theoretical value of portfolio is about 42
  - Based on MC with 10 periods (M) and 10 000 paths (N)
- Assume firm's net income from setting up portfolio is 50—theoretical profit of 8
- Firm would like to secure the profit of 8—get rid of future cash flow obligations

Numerical Greeks cont'd

- To compute Delta in our example
  - With $N = 10000$ and $h = 1$ ($h/S = 0.01$—'small')

$$\Delta = \frac{43.46 - 41.98}{1.00} = 1.49$$

  - Should report confidence interval of Δ
- Make sure you use same sample paths for $f(x+h)$ and $f(x)$!
  - Computing $f(x+h)$ and $f(x)$ with independent observations may easily make $f(x+h) < f(x)$ when economics dictate opposite inequality!
  - Must then use very large K and not too small h

Delta hedging

- Can use Delta-info to manage/reduce risk
- Consider expanding portfolio with n stocks
- Change in value of expanded portfolio

$$\Delta^e = \frac{\partial}{\partial S}(-\Pi + nS) = -\Delta + n$$

- Position insensitive to changes in S if $\Delta^e = 0$, or

$$n = \Delta$$

$\longrightarrow$ position becomes delta neutral

Delta hedging

- Consider our Example, with $\Delta = 1.5$
- Consider a price change of h
- Portfolio increase in value $\approx \Delta h = 1.5h =$ loss
- Stock position increase in value $= \Delta[(S + h) - S] = \Delta h = 1.5h =$ gain
- Loss $=$ gain $\longrightarrow$ no exposure to price changes—delta neutral position

Delta hedging

- Above argument *exact* if $h = 1$—since we computed Δ using $h = 1$
- What if $h = 10$?
  - Gain $1.5h = 15.0$ on stock position
  - Lose $57.6 - 42.0 = 15.6$ on portfolio
  - Net loss $15.6 - 15.0 = 0.6 > 0$
- Loss occurs because Δ is a *local* measure of exposure to changes in S $\longrightarrow$ delta neutral position protects against *small* price movements
- Solution: Revise/rebalance hedging position in S, Δ_t, 'often'

- Discrete Δ-hedging incurs costs—even in perfect markets
- Initial cost of setting up hedge, period t accumulated cost, and expiration date final cost is (short a European call option)

$$
\begin{aligned}
B_0 &= \Delta_0 S_0 \\
B_t &= B_{t-\Delta t} R + (\Delta_t - \Delta_{t-\Delta t}) S_t \\
B_T &= B_{T-\Delta t} R + (\Delta_T - \Delta_{T-\Delta t}) S_T - \begin{cases} K, & \text{ITM} \\ 0, & \text{otherwise} \end{cases}
\end{aligned}
$$

- Two sources of costs
  - Buying and selling at unfavorable prices, $(\Delta_t - \Delta_{t-1}) S_t$
    - After a price increase, we buy at a higher than ideal price
    - After a price decrease, we sell at a lower than ideal price
  - Incurred interest on hedge position, RB_t, $R = \mathrm{e}^{r\Delta t}$

- We have that

$$c_0 \begin{cases} = & \lim_{\Delta t \to 0} \mathrm{e}^{-rT} \mathsf{E}\{B_T\} \\ \leq & \mathrm{e}^{-rT} \mathsf{E}\{B_T\} \end{cases}$$

- The inequality is one reason derivatives providers must charge a mark-up on the theoretical value c_0
- To illustrate the above convergence, consider the hedging performance measure

$$\frac{\sigma(\mathrm{e}^{-rT} B_T)}{c_0}$$

Table: Hedging performance for 20-week European call option

Frequency (weeks)	5	4	2	1	0.5	0.25
Performance	0.42	0.37	0.27	0.19	0.14	0.10

Parameter values as in Hull's Table 19.4 (11th edition). Textbook reports slightly poorer performance: it compares date T costs with date 0 theoretical price c_0. R-code "07Greeks.R" available on Canvas.

Gamma

- Γ measures how Δ changes with S
- Must find 'rule' for how to approximate second order derivatives
  - Naive approach

$$
\begin{aligned}
f''(x) &\approx \frac{f'(x+h) - f'(x)}{h} \\
&= \frac{1}{h^2}\left[f(x+2h) - f(x+h) - f(x+h) + f(x)\right] \\
&= \frac{1}{h^2}\left[f(x+2h) + f(x) - 2f(x+h)\right]
\end{aligned}
$$

  - More representative approximation (using info from both sides of point of approximation)

$$
f''(x) \approx \frac{1}{h^2}\left[f(x+h) + f(x-h) - 2f(x)\right]
$$

- Know $f(x+h)$ and $f(x)$ if we have already computed $f'(x)$ (e.g. Δ)

Delta-Gamma-hedging

- For Example-portfolio

$$\Gamma = \frac{1}{1^2}(43.46 + 40.51 - 2 \cdot 41.98) \approx 0.019$$

- Can use Γ to reduce problem of Δ being local measure

  Making delta and gamma neutral position akin to second order Taylor approximation; delta neutral akin to first order Taylor approximation

- Consider Γ of expanded position

$$\frac{\partial^2}{\partial S^2}\left(-\Pi + nS\right) = -\Gamma$$

  $\longrightarrow$ cannot solve for n to make position gamma neutral, because Gamma of stock is 0

- Solution: introduce position in contract with *nonlinear* payoff in S (ensures second order derivative $\neq 0$)

Delta-Gamma hedging

- Assume now there is a standard European put available
  - $k = 120, \;\; T = 1.0$
  - $d_1 = 0.8829$, $\mathrm{n}(d_1) = \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{d_1^2}{2}} = 0.2702$
  - $\Gamma^p = \frac{\mathrm{n}(d_1)}{s\sigma\sqrt{T}} \approx 0.007$
- Investing in m puts, new position can be made delta-gamma neutral

$$0 = \frac{\partial}{\partial S}(-\Pi + nS + mp) = -\Delta + n + m\Delta^p$$
$$0 = \frac{\partial}{\partial S}(-\Delta + n + m\Delta^p) = -\Gamma + m\Gamma^p$$

- with solution

$$m = \frac{\Gamma}{\Gamma^p} = \frac{0.019}{0.007} \approx 2.7$$
$$n = \Delta - \frac{\Gamma}{\Gamma^p}\Delta^p = 1.5 - 2.7 \cdot 0.2 \approx 1.0$$

Systematic risk

- Have seen that for European call $\Delta = \mathrm{N}(d) \in (0, 1)$
- Implies option price always changes less than underlying
- Q: Consistent with standard claim that options are more risky than stocks?
- A: Yes!
  - Question about distribution of future prices relative to current one
  - Arbitrage restriction $c(t) \leq S(t)$, and typically $c(t)$ much smaller than $S(t)$
  - Implies the *rate of return* from the option may be riskier!

Systematic risk

- The claim of 'more risky than underlying' applies not only to total risk
- Assume for illustration that the CAPM holds for underlying asset
- Can find option's β from replicating portfolio—most transparent in discrete time
    - Value of replicating portfolio 'today' $v_t = \Delta_t S_t + \theta_t A_t$
    - Value of replicating portfolio 'tomorrow' $\Delta_t S_{t+1} + \theta_t A_{t+1}$
    - Gross rate of return on portfolio

$$R^v_{t+1} = \frac{\Delta_t S_{t+1} + \theta_t A_{t+1}}{\Delta_t S_t + \theta_t A_t}$$

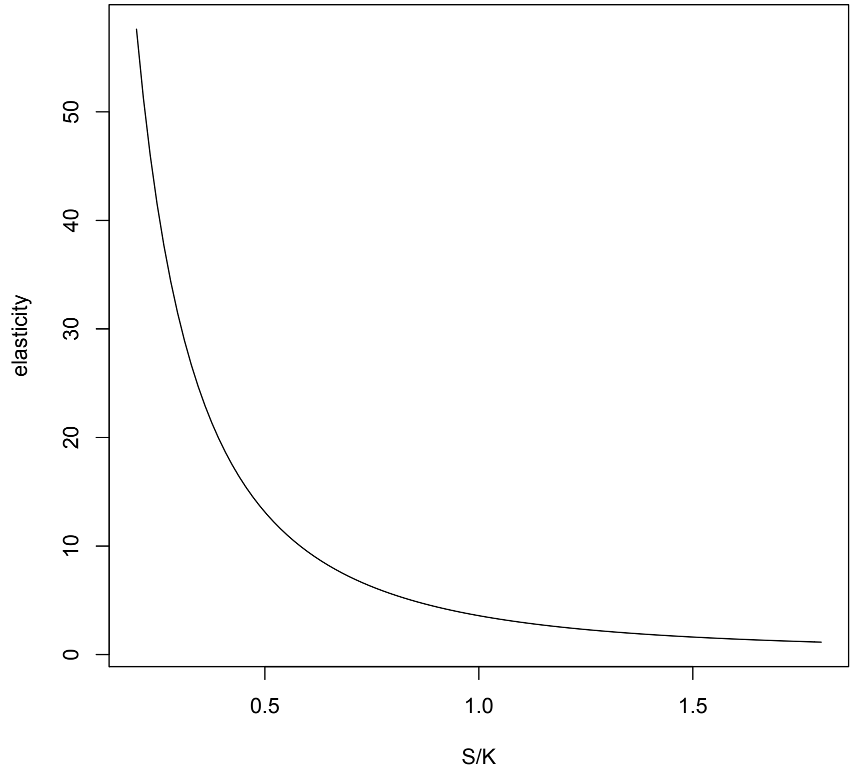
    - Covariance with return on market portfolio

$$\mathrm{Cov}\left(R^v_{t+1}, R^m_{t+1}\right) = \frac{\Delta_t}{v_t}\mathrm{Cov}\left(S_{t+1}, R^m_{t+1}\right) = \frac{\Delta_t S_t}{v_t}\mathrm{Cov}\left(R^S_{t+1}, R^m_{t+1}\right)$$

    - ...which implies that

$$\beta^v_t = \frac{\Delta_t S_t}{v_t}\beta^S = \epsilon_t \beta^S$$

Systematic risk: illustration




```
ela = function(x,r,sigma,T)
{
    d = (log(x)+(r+sigma^2/2)*T)/sigma/
           sqrt(T)
    N1 = pnorm(d)
    N2 = pnorm(d-sigma*sqrt(T))
    c = x*N1-exp(-r*T)*N2

    return(N1/c)
}

curve(ela(x,.02,.4,1.0), 0.2, 1.8, xlab=
      "S/K", ylab="elasticity")
```

Systematic risk

- Several important implications from

$$\beta_t^V = \epsilon_t \beta^S$$

- Option's systematic risk varies
    - with time
    - with the value of the underlying

  even when β^S is constant!
- Very hard to use CAPM to value options, even when valid for the underlying asset
- Very hard to find correct β-estimate for real projects with option-like features
  $\longrightarrow$ Turns out that variations in cost structure is sufficient to make project-betas time and state varying!
  $\longrightarrow$ Motivation for applying option theory to valuation in corporate finance: Real Options

Relevance of option pricing theory for stock markets

- We can express the "payoff" to equity as

$$E = \max(A - K, 0)$$

where E and A are market values of equity and assets, and K is the notional value of debt $\longrightarrow E \sim$ call on A

- We thus have that

$$\beta_t^S = \epsilon_t \beta^A$$

- Stocks's systematic risk varies with
  - time
  - leverage A/K
- Harder than you think(?) to estimate expected equity returns
- Very hard to find correct β-estimate for stocks!
- Option pricing theory useful to understand "stock market anomalies"

Hull, John, *Options, Futures, and Other Derivatives*, 11 ed., Upper Sadle River, New Jersey: Pearson Education, 2022.

ECO423: **Tail Risk**

Jørgen Haug

Department of Finance, NHH

Spring 2023

After this section you should

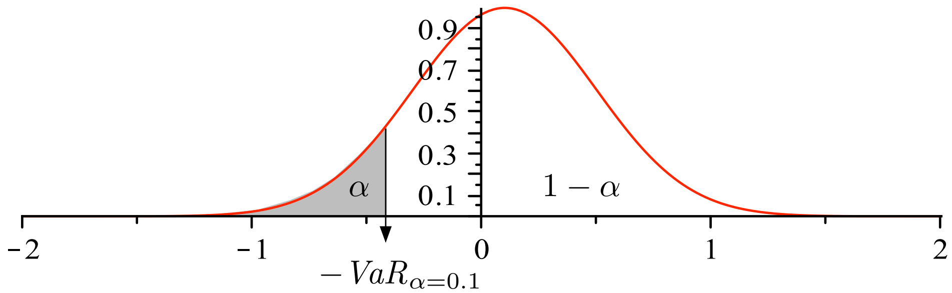
- Know
  - definition of VaR, ES
  - how to illustrate and compute VaR in terms of a portfolio's return/value distribution's pdf, cdf, or empirical/simulated distribution
  - how to compute ES
  - the strengths and weaknesses of VaR, ES
  - the role of P versus Q in VaR and ES analyses
- Have intuition for
  - the interpretation of VaR in capital requirement restrictions for banks
  - why and when the normal distribution is considered an acceptable approximation

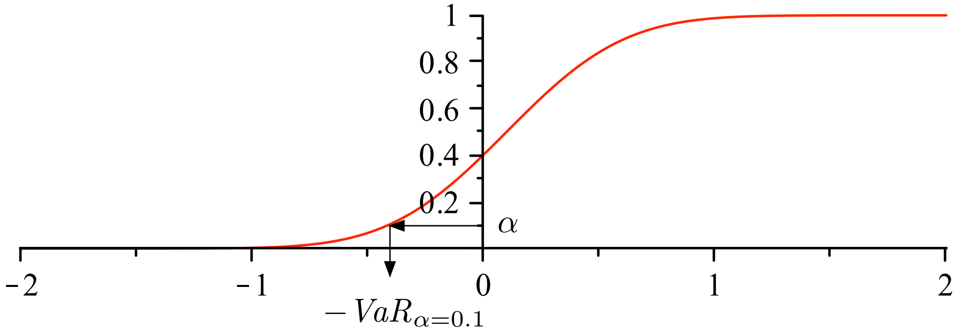
Value at Risk (VaR)

Motivation: Focus on 'bad' aspect of risk—"how bad can it get?"

Definition

Let the stochastic *process* X represent a market value. $-\text{VaR}(T, \alpha; X) = -\text{VaR}_T^\alpha$ is the lower limit of losses (negative outcomes) that occur with probability $1 - \alpha$ within time T.





Value at Risk (VaR)

- Can think of VaR as one measure of risk that aggregates the dimensions captured by the Greeks
  - VaR preferred by "higher level" management
  - Greeks preferred by traders (who typically operate under VaR limits)
- VaR typically computed on a one-day basis
- T-day var typically computed assuming

$$VaR_T^{\alpha} \approx \sqrt{T}\, VaR_1^{\alpha}$$

  (subscript representing days)
- As computers get more powerful, VaR computations get more sophisticated...

Applications of VaR

- Most frequently applied by financial institutions to manage market risk; i.e. to
  - measure downside risk ('passive' risk management)
  - impose risk limits on asset portfolios ('active' risk management)
- Banks apply VaR with $T = 10$ days and $\alpha = 0.01$
  - to report market risk (Basel II; Basel III deals with "bank run" safeguards)
  - to compute capital requirements (European Capital Adequacy Directive—CAD); $k\text{VaR}_T^\alpha$, with $k \geq 3.0$
- Motivation for VaR in minimum capital requirements
  - If the firm has equity of at least VaR_T^α, then the default probability is at most $100\alpha\%$, until time T (given VaR model assumptions)
- Additional reasons for using VaR
  - easy to understand
  - relatively easy to compute

Generic loss distributions

Market risk

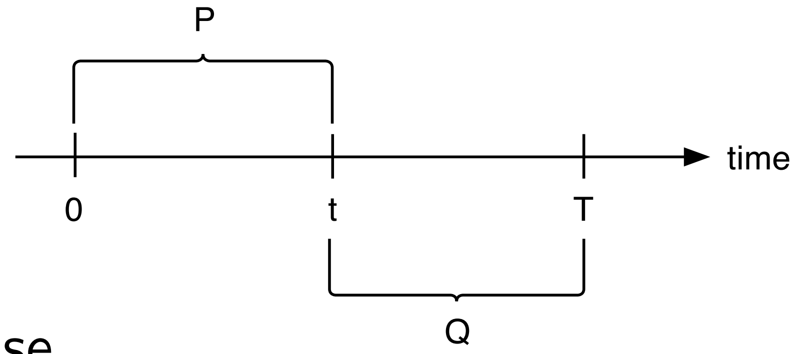
- If loss distribution is not Gaussian, then can always simulate the loss distribution; histogram
- Monte Carlo
    - Have so far computed "scenarios" C_T^n, and focused on

$$C(T, S_T) \approx \frac{1}{N} \sum_{1}^{N} C_T^n$$

    - Instead of taking average, plot histogram for C_T^n $\longrightarrow$ loss distribution for one derivative
    - Similarly for net exposure / portfolio values

Generic loss distributions

- $\text{VaR}_t{}^{\alpha}$ relevant when computed ahead in time, i.e. $t > 0$, assuming 'today' is date 0
- Remember that Q is a *fictitious*, risk-adjusted probability measure
  - Q appropritate for pricing
  - P appropritate for evaluating probability of events



- A bit more precise
  - Simulate price paths until date t, $S^{P,n}$
  - For path n, dates $u > t$
    - Let initial value be $S_t^{P,n}$
    - Simulate future prices $S_u^{Q,k}$
  - Use $S_u^{Q,k}$, $u = t+1, \dots, T$, to compute $C_t^n(S_t^{P,n}) \longrightarrow$ histogram

Problems with VaR

- When used as a descriptive tool ('passive' risk management); Considers only *frequency* of losses to either side of VaR—not the magnitude of losses above VaR (*economic* "tail risk").
- It does not capture the effects of diversification on risks; it's not *sub-additive* (Artzner *et al.*, 1997, 1999)
- When used as a normative tool ('active' risk management); Gives incentives to take on risk with *larger losses* in the tail (Basak and Shapiro, 2001)

Problems with VaR: Example 1

Example (Tail risk)

- Consider portfolio
  - Short $m = 1$ put option with strike $k = 125$ (maximum loss)

$$g_p(T, S) = \max\{k - S, 0\}$$

  - Short $n = 60$ binary options

$$g_b(T, S) = 1_{\{S>k\}}$$

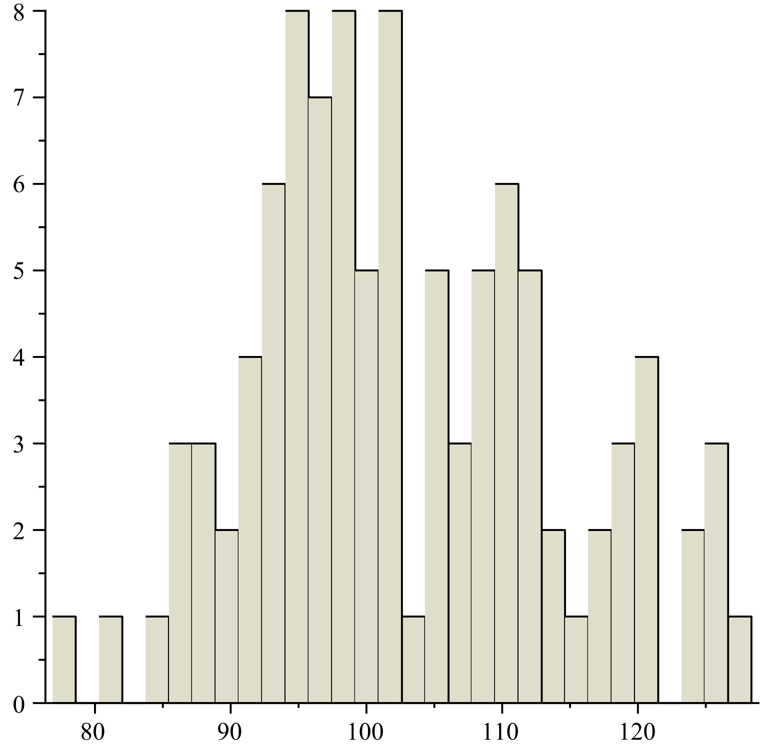
- Want to compute VaR of net portfolio exposure at expiration

$$\mathrm{CF}_T = -\max\{125 - S_T^P, 0\} - 60 \cdot 1_{\{S_T^P > 125\}}$$

- No need to simulate under Q in this case, as no need for present values (purely for illustration)

Problems with VaR: Example 1

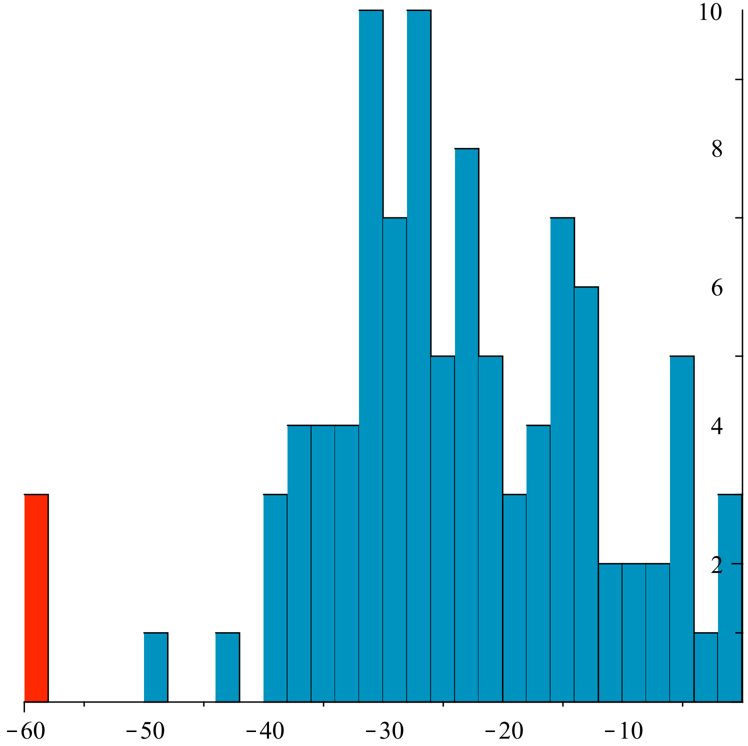
- Assume GBM, with parameter values $S_0 = 100$, $\mu = 0.2$, and $\sigma = 0.4$
- Simulate $K = 100$ paths



- Clearly too few observations to represent possible outcomes (just for illustration)

Problems with VaR: Example 1

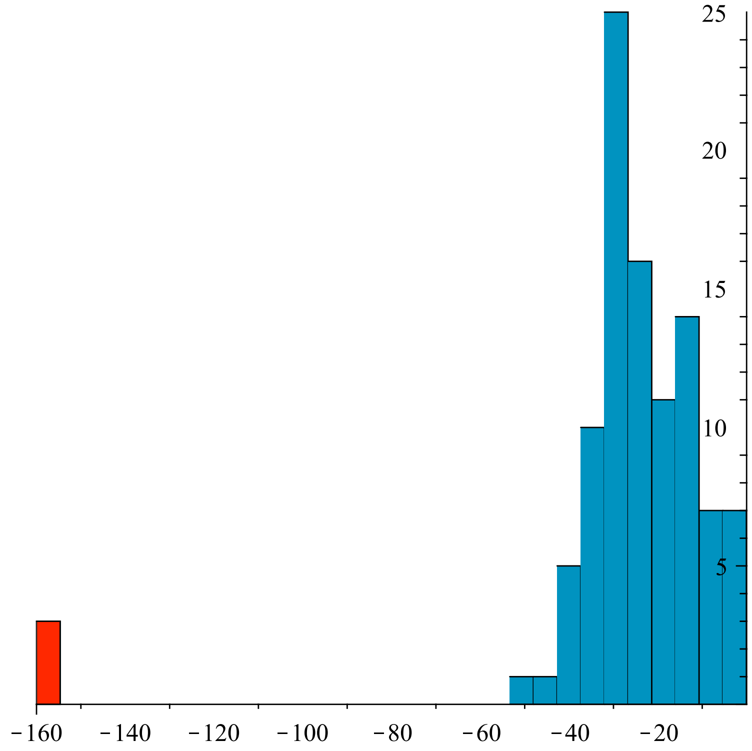
- For each path, compute realized cash flow CF_T



- Visually, $\mathrm{VaR}_T^{0.05} \approx 40$
- Idea: Why not sell 160, rather than 60 binary options?

Problems with VaR: Example 1

- For each path, compute realized cash flow CF_T



- Visually, $\mathrm{VaR}_T^{0.05} \approx 40$, as before!!!
- $\longrightarrow$ Why not sell $160 + 100$ of binary options? ...
- VaR measures only the value at the αth percentile—not the distribution/magnitude of losses on either side of this point

Problems with VaR: Example 2

Example (Credit risk)

Consider 100 different corporate bonds described by

- Maturity of one year
- Coupon of 2%, yield of 2%
- Default probability of 1%
- 0% recovery rate
- Mutually independent defaults

Consider next two alternative portfolios

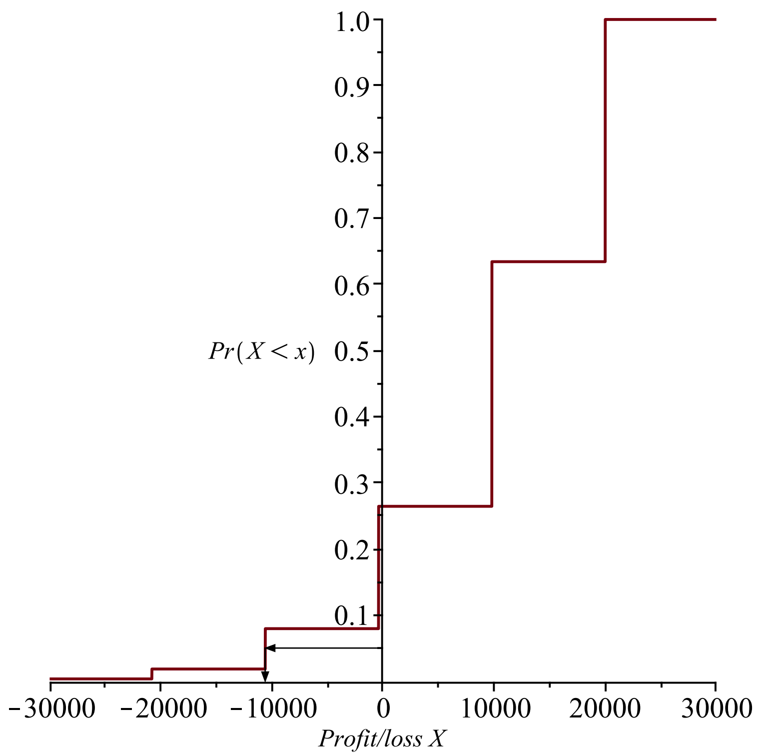
I Invest USD 1 million, with USD 10 000 in each bond

II Invest USD 1 million in one bond

Problems with VaR: Example 2

Example; Portfolio I

- Consider cumulative profit/loss distribution



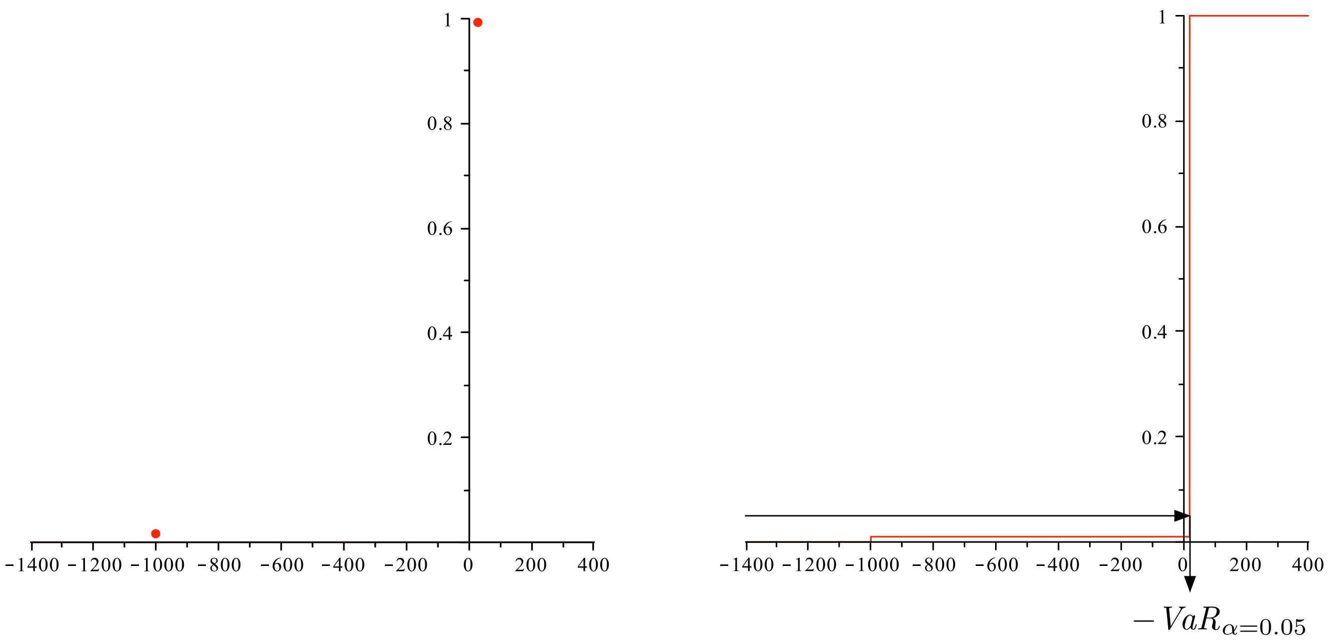
- . . . where we visually observe that $\text{VaR}^{0.05}_{1\text{ year}} \approx 10\ 700$

Problems with VaR: Example 2

Example; Portfolio II

Consider $\text{VaR}_{1\text{ year}}^{0.05}$ of portfolio II by the following considerations

- A loss occurs if *the* one bond defaults
- Probability of a loss, of USD 1 mill., is $0.01 < 0.05 = \alpha$



PDF on the left, CDF on the right.

Problems with VaR: Example 2

Example; Portfolio II

- Must thus have $\text{VaR}^{0.05}_{1\text{ year}} = -20000$

$\longrightarrow$ VaR is smaller in the situation more likely to put the solvency of the firm at risk!

- VaR ignored tail risk
- VaR of portfolio of bonds is *greater* than VaR of single bond

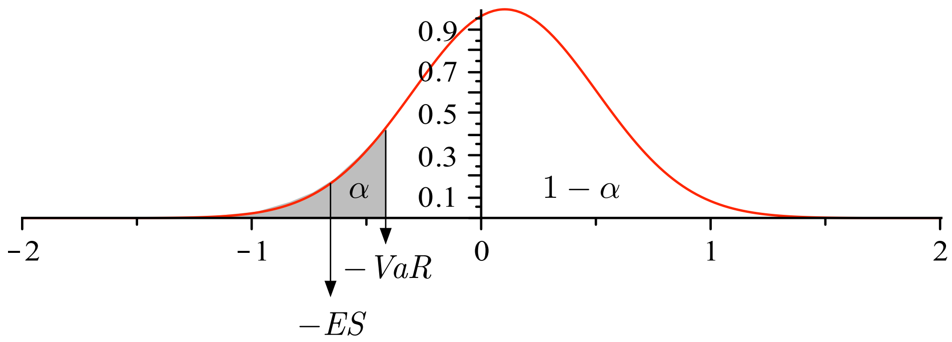
$\longrightarrow$ Fixing these issues is one motivation for considering alternative tail risk measure—like Expected Shortfall

Motivation: "How much can we expect to loose if things get bad?"

Definition (a.k.a. "conditional VaR," "mean excess loss," "beyond VaR," or "tail VaR")

Expected shortfall ES_T^α is the expected loss *conditional* on the loss being larger than VaR_T^α. For a random variable X

$$\mathrm{ES}_T^\alpha = \mathrm{E}\left\{-X \mid -X \geq \mathrm{VaR}_T^\alpha\right\}$$



- Must have $\mathrm{ES}_T^\alpha > \mathrm{VaR}_T^\alpha$, as ES measures losses beyond VaR
- ES thus more conservative when determining capital requirements
- Not possible to relate ES directly to default probability of firm

Expected Shortfall: Example 1

Example (Tail risk with ES, $n = 60$)

- Consider again the portfolio
  - Short 1 put options with strike $k = 125$
  - Short 60 binary options
- Expected Shortfall (here: reading off the histogram)

$$\mathrm{ES}_T^{0.05} \approx \sum_{i:-\mathrm{CF}_T^i > \mathrm{VaR}_T^{0.05}} -\mathrm{CF}_T^i p_{i:\mathrm{VaR}} = \sum_{i:-\mathrm{CF}_T^i > 40} -\mathrm{CF}_T^i p_{i:\mathrm{VaR}}$$

$$\approx 44 \cdot \frac{1}{5} + 50 \cdot \frac{1}{5} + 60 \cdot \frac{3}{5} = 54.8 > 40 = \mathrm{VaR}_T^{0.05}$$

where $p_{i:\mathrm{VaR}}$ is the probability of loss state i, *conditional* on the loss surpassing VaR

Expected Shortfall: Example 1

Example (Tail risk with ES, $n = 160$)

- Consider now the revised portfolio
  - Short 1 put options with strike $k = 125$
  - Short 160 binary options
- Expected Shortfall (here: reading off the histogram)

$$\mathrm{ES}_T^{0.05} \approx \sum_{i:-\mathrm{CF}_T^i > \mathrm{VaR}_T^{0.05}} -\mathrm{CF}_T^i p_{i:\mathrm{VaR}} = \sum_{i:-\mathrm{CF}_T^i > 40} -\mathrm{CF}_T^i p_{i:\mathrm{VaR}}$$

$$\approx 44 \cdot \frac{1}{5} + 50 \cdot \frac{1}{5} + 160 \cdot \frac{3}{5} = 114.8$$

- $\mathrm{VaR}_T^{0.05} = 40$ for both of the portfolios
- ES increased from 54.8 to 114.8 when selling more binary options

Problems with VaR

Tail risk problem does not apply when $X \sim N(\mu_X, \sigma_X)$

- Let q_α be the $100(1-\alpha)$-percentile. Can show (e.g. Hull's textbook; ignoring μ, e.g. for *daily* X returns)

$$\mathrm{VaR}(T, \alpha; X) = q_\alpha \sigma_X,$$

hence, for $\alpha = 0.01$

$$\mathrm{VaR}(T, \alpha; X) \approx 2.33\sigma_X$$

- Can also show that

$$\mathrm{ES}(T, \alpha; X) = \frac{\mathrm{e}^{-\frac{q_\alpha^2}{2}}}{\alpha\sqrt{2\pi}}\sigma_X,$$

hence, for $\alpha = 0.01$

$$\mathrm{ES}(T, \alpha; X) = 2.67\sigma_X$$

- Hence, VaR contains the same information as ES (up to a *known* scalar multiple)

Problems with VaR

Tail risk problem does not apply when $X \sim N(\mu_X, \sigma_X)$

Proof.

We have that

$$
\begin{aligned}
\mathrm{ES}(\alpha) &= \mathrm{E}\left\{-X \middle| -X \geq \mathrm{VaR}(\alpha)\right\} \\
&= \frac{\mathrm{E}\left\{-X 1_{\{X \leq -\mathrm{VaR}(\alpha)\}}\right\}}{\alpha} \\
&= \frac{1}{\alpha \sigma_X \sqrt{2\pi}} \int_{-\infty}^{-\mathrm{VaR}(\alpha)} (-x) e^{-\frac{x^2}{2\sigma_X^2}}\, dx \\
&= \frac{1}{\alpha \sigma_X \sqrt{2\pi}} \left[\sigma_X^2 e^{\frac{-x^2}{2\sigma_X^2}} \right]_{-\infty}^{-\mathrm{VaR}(\alpha)} \\
&= \frac{1}{\alpha\sqrt{2\pi}} e^{\frac{-\mathrm{VaR}^2(\alpha)}{2\sigma_X^2}} \sigma_X
\end{aligned}
$$

where $\mathrm{VaR}(\alpha) = \mathrm{VaR}(T, \alpha; X) = q_\alpha \sigma_X$. □

Problems with VaR

Sub-additivity applies when $X \sim N(\mu_X, \sigma_X)$

- Let X be a *linear combination* of multivariate normal profit-loss distributions
- It still applies that VaR is a scalar multiple of σ_X
- We know from Markowitz (and probability theory) that $\text{Var}(X)$ is sub-additive;

$$\begin{aligned}
\sigma_{X+Y} &= \sqrt{\sigma_X^2 + \sigma_Y^2 + 2\sigma_{XY}} = \sqrt{\sigma_X^2 + \sigma_Y^2 + 2\rho_{XY}\sigma_X\sigma_Y} \\
&\leq \sqrt{\sigma_X^2 + \sigma_Y^2 + 2\sigma_X\sigma_Y} = \sqrt{(\sigma_X + \sigma_Y)^2} = \sigma_X + \sigma_Y
\end{aligned}$$

- Hence, VaR is sub-additive

$\longrightarrow$ Better to compute VaR than ES under normality, as

- they contain same information about tail risk
- they are both sub-additive

Popular version of VaR: The linear model

- Let Y be the price of an 'underlying' asset
- The Euler scheme for $dY_t = Y_t\mu_Y\,dt + Y_t\sigma_Y\,dW_t$ is

$$Y_{t+\Delta t} = Y_t + Y_t\mu_Y\Delta t + Y_t\sigma_Y\sqrt{\Delta t}\epsilon_{t+\Delta t}$$

- For short time periods, e.g. a day, we an approximate the dynamics of Y by

$$Y_{t+\Delta t} - Y_t \approx Y_t\sigma_Y\epsilon_{t+\Delta t}$$

  where σ_Y now is measured from *daily* price changes
- Or

$$R^Y_{t+\Delta t} = \frac{Y_{t+\Delta t} - Y_t}{Y_t} \approx \sigma_Y\epsilon_{t+\Delta t}$$

- 1-day 99% VaR equals $2.33\sigma_Y$ times the dollars invested in this asset

Popular version of VaR: The linear model

- Consider next a contingent claim, $c_t = C(t, Y_t)$
- Can approximate change in value in c_t over a small time interval (a day)

$$
\begin{aligned}
c_{t+\Delta t} &= C(t,y) + C_2(t,y)(Y_{t+\Delta t} - y) + \frac{1}{2} C_{22}(t,y)(Y_{t+\Delta t} - y)^2 + \dots \\
&= C(t,y) + \Delta_t y R^Y_{t+\Delta t} + \frac{1}{2}\Gamma_t \left(y R^Y_{t+\Delta t} \right)^2 + \dots
\end{aligned}
$$

  where we have ignored change in value due to time $C_1(t, Y_t)$, and $y = Y_t$
- If we ignore the Γ-term, then model is clearly still linear in a standard normal term—linear VaR model applies
  $\longrightarrow$ All price and pricing models based on SDEs are 'locally Gaussian' (by Itô's Lemma)

Bootstrapped VaR

- We may be worried that models of underlying assets do not capture extreme market situations
- Can capture realistic scenarios and their frequency by sampling from historical market prices—bootstrapping
- Can capture extreme scenarios
  - use sufficiently long time series
  - stress test with shorter time series that contains relevant events/crises
- Important consideration
  - Cross sectional correlation structure important for multi-asset portfolios
  - Intertemporal correlation structure important for "longer term" VaR
    - Captures effect of cumulative "bad news"
    - Not important for one-day VaR—use recent empirical distribution instead (example in textbook; H21.2)
- Note: Historical prices are observations under P

Example

- Empirical VaR illustrates a situation where software like Excel struggles, while software like R excels
- Consider a naked call on Microsoft stock, beginning of 2021
- Download time series of Microsoft 2000–2020 (from yahoo! finance);

  `PriceData = BatchGetSymbols(tickers, ...)`
  - *Q:* Why would you often prefer a shorter time series?
  - *Q:* What's good about using long time series?
- Extract returns; `R = PriceData$df.tickers$ret.adjusted.prices[-1]`
- Generate bootstrap sample of returns; `Rsample = sample(R, N, replace = TRUE)`
- Generate scenarios for daily call price changes; $c_0 - c_1$
- Conpute 1-day VaR as 1st percentile; `quantile(c0 - c1, probs = 0.01)`
- Scale to desired time horizon

Example; complete R code

```
BSM = function(S,K,r,sigma,T) {
   d = (log(S/K)+(r+sigma^2/2)*T)/sqrt(T)/sigma
   return(S*pnorm(d) - K*exp(-r*T)*pnorm(d-sqrt(T)*sigma))
}

# Returns data to be used in bootstrap sampling
require('BatchGetSymbols')
first.date = "2000-01-01"
last.date = "2020-12-31" # Set to Sys.Date() if you want to use latest observations
freq.data = 'daily'; tic = c('MSFT')
PriceData = BatchGetSymbols(tickers = tic, first.date = first.date,
   last.date = last.date, thresh.bad.data = 0.10, type.return = "arit",
   freq.data = freq.data, cache.folder = file.path(tempdir(), 'BGS_Cache') )
PriceData = PriceData$df.tickers
R = PriceData$ret.adjusted.prices[-1] # Exclude first observation, which equals NA

# Parameters; naked position in one European call
S0 = PriceData[length(PriceData[,1]),]$price.close; v = sd(R[(length(R)-251):length(R)])*
     sqrt(252)
K = 220; r = 0.02; T = 0.5
# Set current price to (latest) market price; "cheating" here, using BSM price
c0 = BSM(S0,K,r,v,T)

# Generate bootstrap sample
N = 10000; set.seed(1); Rsample = sample(R, N, replace = TRUE)

# Scenarios for call price tomorrow and VaR in 10 days
cScenarios = BSM(S0*(1+Rsample), K, r, v, T-1/252)
VaR = quantile(c0 - cScenarios, probs = 0.01)
VaR10 = VaR*sqrt(10); VaR10
```

ECO423: **American Derivatives:** Least Squares Monte Carlo

Jørgen Haug

Department of Finance, NHH

Spring 2023

After this section you should

- Know
  - the definition of an American-style derivative
  - the role of intrinsic and continuation value in determining present value
  - the role of least squares regression in LSMC
  - the logic of the backwards induction process to determine present value
- Have intuition for
  - why we use only ITM paths in the least squares regressions
  - why we do not use the estimated continuation value directly in computing present value

Pricing American derivatives

- Consider American-style derivative
  - expires at time T
  - cash flow $g(t, S_t)$ at time t if exercised, zero otherwise
- Define
  - C_t present/market value time t
  - c_t present value time t if *not* exercised
- Market value given by present value of cash flows, given optimal exercise of contract

$$C_t = \max\{g(t, S_t), c_t\},$$

where

$$c_t = \mathrm{E}_t^Q\left\{\mathrm{e}^{-r} C_{t+1}\right\} = \mathrm{PV}(\textit{future}\ \text{cash flows} | S_t = x)$$

Pricing American derivatives

- Start at time $t = T$ when we *know*

$$C_T = g(T, S_T)$$

- Do backwards induction, $t = T - 1, T - 2, \ldots, 0$
- At each time $t < T$ must compare
  - exercise value $g(t, S_t)$
  - continuation value $c_t = c_t(x) = \mathrm{PV}(\text{future cash flows} | S_t = x)$
- **Problem:** How do we compute

$$\mathrm{PV}(\text{future cash flows} | S_t = x)$$

which depends on future optimal exercise?

Pricing American derivatives

- **Idea:** Approximate conditional expectation by regression equation $a_0 + a_1 x + a_2 x^2 + \ldots + a_p x^p$
  - Assume we have K observations of $S_t^k = X^k$ and $\text{PV}^k(\text{future cash flows}) = Y^k$
  - The conditional expectation $c_t = \text{PV}(\text{future cash flows}|S_t = x)$ is *the mean value of PV$(\cdot)$ given a particular observation x*
  - A regression

$$Y^k = a_0 + a_1 X^k + \epsilon^k$$

evaluated at a particular x,

$$a_0 + a_1 x,$$

is *the average value of Y given $X = x$*

▸ Intuition

- Consider $g(t, S_t) = \max\{0, 1.10 - S_t\}$, $r = 0.06$ per period
- Assume we simulate $K = 8$ price paths, with $M = 3$ time periods

Stock price paths

Path	$t=0$	$t=1$	$t=2$	$t=3$
1	1.00	1.09	1.08	1.34
2	1.00	1.16	1.26	1.54
3	1.00	1.22	1.07	1.03
4	1.00	0.93	0.97	0.92
5	1.00	1.11	1.56	1.52
6	1.00	0.76	0.77	0.90
7	1.00	0.92	0.84	1.01
8	1.00	0.88	1.22	1.34

- Green price indicates option is in-the-money (except date 0)

$t = 3$

- Cash flows from optimal exercise

Cash flow matrix at time 3, C_3

Path	$t=1$	$t=2$	$t=3$
1	—	—	$0.00 = \max\{0, 1.10 - 1.34\}$
2	—	—	0.00
3	—	—	0.07
4	—	—	0.18
5	—	—	0.00
6	—	—	0.20
7	—	—	0.09
8	—	—	0.00

$t = 2$

- Need to consider optimal exercise only for *in-the-money* paths
- Start by comparing immediate exercise with continuing
- Find PV of continuing by regressing $X^k = S_2^k$, for in-the-money-paths, against $Y^k = \mathrm{e}^{-0.06} C_3^k$
- Data for continuation value

Time 2 regression

Path	Y	X
1	$0.00\mathrm{e}^{-0.06}$	1.08
2	—	—
3	$0.07\mathrm{e}^{-0.06}$	1.07
4	$0.18\mathrm{e}^{-0.06}$	0.97
5	—	—
6	$0.20\mathrm{e}^{-0.06}$	0.77
7	$0.09\mathrm{e}^{-0.06}$	0.84
8	—	—

- Regression $Y^k = a_0 + a_1 X^k + a_2 \left(X^k\right)^2 + \epsilon^k$ yields

$$\hat{c}_2(X) = -1.070 + 2.983X - 1.813X^2$$

$t = 2$ cont'd

- Comparing immediate exercise with value of continuing

Optimal exercise

Path	Exercise	Continue, $\hat{c}_2(X)$
1	$0.02 = \max\{0, 1.10 - 1.08\}$	$0.0369 = -1.070 + 2.983 \cdot 1.08 - 1.813 \cdot 1.08^2$
2	—	—
3	0.03	0.0461
4	0.13	0.1176
5	—	—
6	0.33	0.1520
7	0.26	0.1565
8	—	—

$t = 2$ cont'd

- Updated optimal cash flows, given no exercise before time 2

Cash flow matrix at time 2

Path	$t = 1$	$t = 2$	$t = 3$
1	—	0.00	0.00
2	—	0.00	0.00
3	—	0.00	0.07
4	—	0.13	0.00
5	—	0.00	0.00
6	—	0.33	0.00
7	—	0.26	0.00
8	—	0.00	0.00

- Cash flows in red are replaced by time-2 cash flows

$t = 1$

- Repeat the steps at time $t = 2$ for $t = 1$
- Find PV of continuing by regression, $X^k = S_1^k$ against $Y^k = \mathrm{PV}_1(\max\{0, 1.10 - S_\tau^k\})$, S_τ^k the stock price at optimal time of exercise along path k; $\tau \geq 2$ when $t = 1$
- Data for continuation value

Time 1 regression

Path	Y	X
1	$0.00\mathrm{e}^{-0.06(\tau-t)} = 0.00\mathrm{e}^{-0.12}$	1.09
2	—	—
3	—	—
4	$0.13\mathrm{e}^{-0.06}$	0.93
5	—	—
6	$0.33\mathrm{e}^{-0.06}$	0.76
7	$0.26\mathrm{e}^{-0.06}$	0.92
8	$0.00\mathrm{e}^{-0.12}$	0.88

- Regression of X on Y yields

$$\hat{c}_1(X) = 2.038 - 3.335X + 1.356X^2$$

$t = 1$ cont'd

- Comparing immediate exercise with value of continuing

Optimal exercise

Path	Exercise	Continue, $\hat{c}_1(X)$
1	$0.01 = \max\{0, 1.10 - 1.09\}$	$0.0139 = 2.038 - 3.335 \cdot 1.09 + 1.356 \cdot 1.09^2$
2	—	—
3	—	—
4	0.17	0.1092
5	—	—
6	0.34	0.2866
7	0.18	0.1175
8	0.22	0.1533

$t = 1$ cont'd, $t = 0$

- Updated optimal cash flows, given no exercise before time 1

Cash flow matrix at time 1

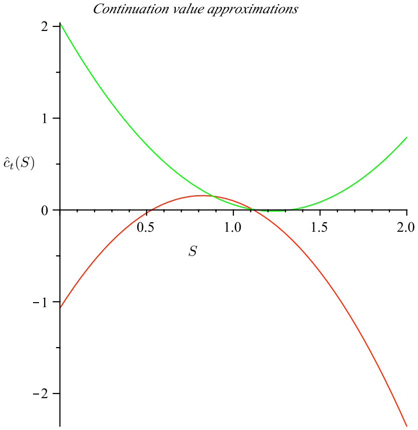
Path	$t=1$	$t=2$	$t=3$
1	0.00	0.00	0.00
2	0.00	0.00	0.00
3	0.00	0.00	0.07
4	0.17	0.00	0.00
5	0.00	0.00	0.00
6	0.34	0.00	0.00
7	0.18	0.00	0.00
8	0.22	0.00	0.00

- Cash flows in red are replaced by time 1 cash flows
- Table contains info needed to determine present value

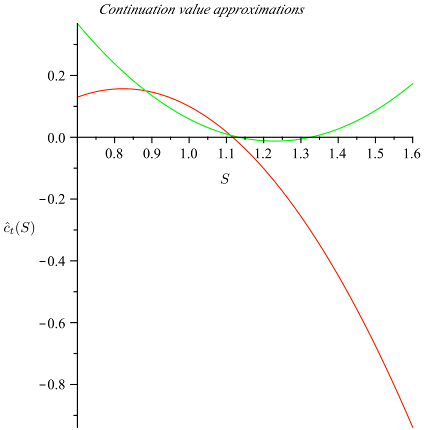
$$C_0 \approx \frac{\mathrm{e}^{-0.06}(0.17 + 0.34 + 0.18 + 0.22) + \mathrm{e}^{-0.06\cdot 3}0.07}{8} = 0.1144$$

- Behavior of approximate continuation value functions

$$\hat{c}_1(X) = 2.038 - 3.335X + 1.356X^2$$
$$\hat{c}_2(X) = -1.070 + 2.983X - 1.813X^2$$



- Important to evaluate approximation on the range of the input data!



- Mostly downward sloping on intervals $[0.76, 1.09]$, $[0.77, 1.08]$

- Owners of American derivatives need to know optimal exercise policy
- At each time t with above approach we ran regression

$$Y_t^k = a_{0,t} + a_{1,t} X_t^k + \ldots + a_{p,t} \left(X_t^k \right)^p + \epsilon_t^k$$

- Approximate continuation value given by function

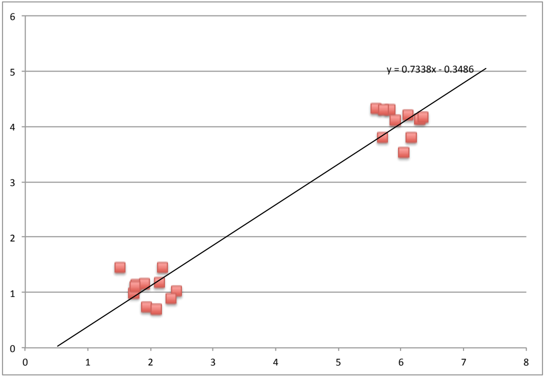
$$\hat{c}_t(x) = a_{0,t} + a_{1,t} x + \ldots + a_{p,t} x^p$$

- Exercise region given by solution x to the equation

$$g(t,x) = \hat{c}_t(x)$$

i.e., all points x such that $g(t,x) > \hat{c}_t(x)$

- Consider regression equation $Y = -0.35 + 0.73X$



- $\mathrm{E}\{Y\} \sim$ average outcome
  - Average outcome $\bar{Y} = 2.5 = -0.35 + 0.73\bar{X}$ $(\bar{X} = 3.88)$
- $\mathrm{E}\{Y|X = x\} \sim$ regression equation evaluated at $X = x$
  - Conditional average for $x = 2$ is $1.1 = -0.35 + 0.73x$

Back

- Purpose of computing conditional expectation: determine *if optimal to exercise*
- Exercise decision not relevant for OTM paths
- Adding OTM paths to regression
  - Similar to adding non-explanatory observations to regression $\sim$ noise; $X + \epsilon$
    $\longrightarrow$ increases standard errors
  - OTM paths also represent increase in domain for approximating function
    $\longrightarrow$ requires more complex approximating function to maintain quality

$\longrightarrow$ Nothing to gain, much to loose

- Why do we use actual payouts in valuation, and not the estimated continuation values?
    - $C_t^k = \max\{g(t, S_t^k), c_t^k\}$, c_t^k *true* continuation value
    - Approximate c_t^k by $\hat{c}_t(S_t^k)$
    - If $\hat{c}_t(S_t^k) > g(t, S_t^k)$ then *do not* update payoff matrix
      $\longrightarrow$ why not instead set $C_t^k \approx \hat{c}_t(S_t^k)$?
- Because $\hat{c}_t(S_t^k) = c_t^k + \epsilon_t^k$, ϵ_t^k mean zero, independent of c_t^k:

$$C_t^k \approx \max\{g(t, S_t^k), \hat{c}_t(S_t^k)\} = \max\{g(t, S_t^k), c_t^k + \epsilon_t^k\}$$

    - tends to pick $\hat{c}_t(S_t^k) = c_t^k + \epsilon_t^k$ when ϵ_t^k large
    - tends to pick $g(t, S_t^k)$ when ϵ_t^k small

$\longrightarrow$ C_t^k upward biased if using $\hat{c}_t(S_t^k)$ as continuation value in final PV computation

ECO423: **Real Options**

Jørgen Haug

Department of Finance, NHH

Spring 2023

After this section you should

- Know
  - what a real option is—definition in this course
  - why constant project beta is unrealistic (cf. lecture on Greeks)
  - how to adjust for systematic risk when risk factors are not traded or commodities
  - the role of convenience yields in valuation
  - how to work with (risk-adjust and simulate) multi-factor models
  - how to use futures/forwards to simplify valuation of projects whose cash flows are linear in commodities prices
- Have intuition for
  - the economic causes for convenience yields

Definition

A *derivative* is a cash flow that is determined only by time and the value of a stochastic process Y, $CF_t = g(t, Y_t)$.

Definition

A *real option* is a derivative whose cash flows stem from an investment project.

- Really nothing new in definition of a real option
- No need to introduce 'flexibility' or 'option like features' to consider a cash flow as a real option
- As is often the case, "The Devil's in the details..."
- The nature of Y determines how convenient option pricing methodology is for estimating present value

Example

An oil installation will generate profits $\pi_t = KS_t - F = I_t - F$

- K maximal output per period
- S_t spot price of oil
- F fixed costs per period

- Why not use the DCF method?
- Assume the CAPM is true, and S has constant beta, $\beta_S \implies I$ has constant beta, $\beta_I = \beta_S$
- Present values are additive; with one period

$$V = PV(\pi) = V^I - V^F$$

- Gross rate of return given by

$$R = \frac{\pi}{V} = \frac{V^I}{V}\frac{I}{V^I} - \frac{V^F}{V}\frac{F}{V^F} = \frac{V^I}{V}R^I - \frac{V^F}{V}R^F$$

- Follows that

$$\beta_R = \frac{KS_0}{V}\beta_S$$

- Factor of proportionality is *state and time dependent*

$$\epsilon = \frac{KS}{V} = \frac{KS}{V^I - V^F}$$

  - (Only) S and V^I varies with level of $S \longrightarrow$ state-dependent
  - V^I and V^F varies with time (less cash flows left to discount, in multi-period setting) $\longrightarrow$ time-dependent

- $\longrightarrow$ Cost structure induces leverage-like effects in β, as we observed in lecture on the Greeks

- Recall that option theory infers correct risk adjustment from underlying 'security'; e.g. $\lambda = \frac{\mu - r}{\sigma}$ in BSM economy
- Crucial that value of underlying 'security' fully reflects investors' attitudes to risk
  - that expected growth rate equals required rate of return
- Relevant to distinguish between three cases
  1. Vague information about the risk factor(s) Y
  2. Good information about Y
     1. not a traded asset or commodity
     2. asset or commodity traded in a well-functioning market

Example

H&M, IKEA, ... (broad collections of consumer products)

- Most important source of uncertainty is Q^i—quantity sold of merchandise i
- Q^i determined by fashion, economic outlook, ...
- Particularly hard to model fashion
- Majority of projects, measured in number of projects, fall into this category!

- Difficult to extract risk adjustment information, λ, directly from risk factors $\longrightarrow$ cannot rely only on option theory
- Can use DCF method to value projects without flexibility—typically assuming
  - no variation in discount rate over time
  - no variation in discount rate over states
  - a capital asset pricing model, e.g. CAPM
- Use result from previous step in building *present value tree*—akin to *stock price tree* for stock option
- Above step allows extraction of λ—or u, d in discrete time, binomial model; $q = \frac{R^f - d}{u - d}$
- Use extracted q to value flexibility in project
  $\longrightarrow$ End up using discrete-time methods
- Case that's typically treated in textbooks in corporate finance
- Copeland, Weston, and Shastri (2005) has a nice Chapter 9 on this methodology (won't test you on this on the exam)

- Traditional project analysis has resulted in the data

R^f	1.05		
R^I	1.11		

Year	0	1	2
$\mathrm{E}\{I_t\}$	0	50.0	50.0
F_t	0	45.0	45.0
$\mathrm{E}\{CF_t\}$	0	5.0	5.0

- The DCF method yields

Year	0	1	2
V_t^I	85.82	95.11	50.00
V_t^F	83.67	87.86	45.00
V_t	2.14	7.26	5.00

- Now want to consider technology that allows temporary shut down of activity—making costs 'semi-fixed'

- Optimal to shut down activity when temporary unprofitable
  $\longrightarrow$ Necessary to model variation in profitability
- Must model variation in present values (a *priced* variable) to be able to infer its market price of risk—and hence appropriate risk adjustment
  - Project value $\sim$ price of underlying asset
  - Cash flows $\sim$ dividend payouts from underlying
- Not realistic with CF independent of present value
  $\longrightarrow$ Need to model relationship
- One simple model is $I_t = a_t V_t^I$ with

$$a_t = \frac{\mathrm{E}\{I_t\}}{V_t^I}$$

Year	0	1	2
a_t	$\frac{0.00}{85.82} = 0.00$	$\frac{50.00}{95.11} = 0.53$	$\frac{50.00}{50.00} = 1.00$

- Assuming a reasonable(?) value for $u = 1.20$ ($d = 1/u$) and using data on a_t yields tree for V^I in the absence of flexibility

Year	0	1	2
			$58.61 = 102.98(1 - a_1)u$
		$102.98 = 85.82(1 - a_0)u$	
	85.82		$40.70 = 102.98(1 - a_1)d$
		71.51	
			28.27

- Net cash flows from project $CF_t = V_t^I a_t - F_t$

Year	0	1	2
			13.61
		$9.14 = 102.98 \cdot 0.53 - 45.00$	
	0.00		-4.30
		$-7.41 = 71.51 \cdot 0.53 - 45.00$	
			-16.73

- Cash flows with possibility of temporary shut down

Year	0	1	2
			13.61
		9.14	
	0.00		0.00
		0.00	
			0.00

- Above data implies $q = \frac{R^f - d}{u - d} = 0.59$
- Present value with flexibility

$$V_0^{\text{flex}} = 4.31 = \frac{9.14q}{R^f} + \frac{13.61q^2}{(R^f)^2}$$

- Value of possibility to temporarily shut down

$$V_0^{\text{flex}} - V_0 = 4.31 - 2.14 = 2.17$$

Steps in analysis

1. Value project without flexibility with standard DCF method
2. Build model of variations in present value
   $\longrightarrow$ risk adjusted probabilities
3. Build model of relationship between cash flows and present value
4. Price cash flows with flexibility, using risk adjusted probabilities

- Schwartz and Moon (2000, 2001) use diffusion model to estimate market value of Amazon.com
- Revenues

$$
\begin{aligned}
dR_t &= \mu_t R_t \, dt + \sigma_t R_t \, dW_t^1 \\
d\mu_t &= \kappa(\bar{\mu} - \mu_t) \, dt + \eta_t \, d(\rho W_t^1 + \sqrt{1-\rho^2} W_t^2) \\
d\sigma_t &= \kappa_1(\bar{\sigma} - \sigma_t) \, dt \\
d\eta_t &= -\kappa_2 \eta_t \, dt
\end{aligned}
$$

where W^1 and W^2 are independent (Authors assume equivalently $d\mu_t = \kappa(\bar{\mu} - \mu_t) \, dt + \eta_t \, dW_t^2$ and $dW_t^1 dW_t^2 = \rho \, dt$)

- Net after-tax profits

$$
\begin{aligned}
Y_t &= (R_t - C_t)(1 - \tau_c) \\
C_t &= (\alpha + \beta) R_t + F
\end{aligned}
$$

where αR_t is cost of goods sold, and $F + \beta R_t$ is “other expenses”

- Value of firm

$$V_0 = \mathrm{E}_0^Q \left\{ \int_0^T \mathrm{e}^{-rt} Y_t \, dt \right\}$$

Different expression in paper, either due to different assumptions about the role of Y in the firm, or due to a typo.

- Remains to determine dynamics of R and μ under Q
- But R and μ are not traded assets! How do we determine risk adjustments?
$\longrightarrow$ Will deal with this new, important issue next!
- Important part of paper is on how to estimate parameters of model from accounting reports etc.

Example

Temperature, goods/services that are not competitively traded, ...

- Cannot extract risk adjustment information, λ, directly from risk factors $\longrightarrow$ cannot rely only on option theory
- Assume we know $CF_t = g(t, Y_t)$
- Assume we have information about time series properties of Y_t
- Can then infer λ from a capital asset pricing model, e.g. the CAPM

- Assume $dY_t = a_Y Y_t\,dt + \sigma_Y Y_t\,dW_t$, a_Y the expected *growth rate* (as opposed to 'required rate of return')
- Collect data on $R^Y_{t+1} = \frac{Y_{t+1}}{Y_t}$ and R^m_{t+1}
- Estimate correlation coefficient $\rho_{Y,m} = \rho(R^Y, R^m)$
- CAPM restriction on *expected rate of return* μ_Y

$$\mu_Y - r = \beta_{Y,m}(\mu_m - r) = \frac{\rho_{Y,m}\sigma_Y\sigma_m}{\sigma_m^2}(\mu_m - r) \tag{1}$$

- It follows from (1) that

$$\lambda = \frac{\mu_Y - r}{\sigma_Y} = \frac{\rho_{Y,m}}{\sigma_m}(\mu_m - r)$$

- With an estimate of λ (a *number*)
  - Adjust the risk factor for systematic risk (now with more precise notation; the Q-superscript)

$$
\begin{aligned}
dY_t^Q &= a_Y Y_t^Q \, dt + \sigma_Y Y_t^Q (dW_t^Q - \lambda \, dt) \\
&= (a_Y - \lambda \sigma_Y) Y_t^Q \, dt + \sigma_Y Y_t^Q dW_t^Q
\end{aligned}
$$

  - Use e.g. Euler + Monte Carlo to estimate PV

$$
PV(CF_t) \approx \mathrm{e}^{-rt} \frac{1}{K} \sum_{k=1}^{K} g(t, Y_t^{Q,k})
$$

- In principle a simple method to use; the same calculations we've always done + correlation estimation
- In reality unclear how good our estimate of λ is—must assume a particular asset pricing model!

- Assume as before that $CF_t = g(t, Y_t)$
- Assume moreover Y_t is a competitive market price
  $\longrightarrow$ Y contains sufficient information about priced/systematic risk to derive risk adjusted probabilities

Example

Statoil, $Y = (\text{oil}, \text{gas})$; Norske Skog, $Y =$ newspaper; Nippon Steel, $Y =$ steel; Norsk Hydro, $Y =$ aluminum; Illy, $Y =$ coffee beans

- Turns out it is very useful if there exist well functioning spot and futures markets—at least the latter

- Can treat project valuation as a standard option pricing problem
- Risk adjust Y so that $Y^* = \frac{Y}{A}$ is a Q-martingale
- Use e.g. MC to estimate project value as

$$PV(CF_t) \approx \mathrm{e}^{-rt} \frac{1}{K} \sum_{k=1}^{K} g(t, Y_t^{Q,k})$$

- (Often) No need for a capital asset pricing model

- Digression on futures pricing
  - Futures contract: a right to buy cash flow Y_T at date T, at a price $F_{t,T}$ determined at date $t < T$
  - By convention $F_{t,T}$ is set to ensure the contract has zero present value,

$$\frac{0}{A_t} = \mathrm{E}_t^Q \left\{ \frac{Y_T - F_{t,T}}{A_T} \right\}, \text{ iff } F_{t,T} = \frac{1}{P_{t,T}} \mathrm{E}_t^Q \left\{ \mathrm{e}^{-\int_t^T r_u \, du} Y_T \right\},$$

    allowing for stochastic riskless rate
  - If r and Y are independent

$$\mathrm{E}_t^Q \left\{ \mathrm{e}^{-\int_t^T r_u \, du} Y_T \right\} = P(t, T) \mathrm{E}_t^Q \left\{ Y_T \right\},$$

    and

$$F_{t,T} = \mathrm{E}_t^Q \left\{ Y_T \right\}$$

- If Y is a geometric Wiener process with constant growth rate r

$$F_{t,T} = Y_t \mathrm{e}^{r(T-t)} \tag{2}$$

- Empirical data shows that for commodities in particular

$$F_{t,T} \neq Y_t \mathrm{e}^{r(T-t)}$$

- Reason for difference
  1. Cost of carry: it is costly to store the asset $\sim$ negative dividend
  2. Convenience yield: it is beneficial to sit on physical supplies for many activities $\sim$ positive dividend

- Approximate cost of carry and convenience yield as *rates*
- Denote net effect at time t by $\delta_t Y_t$, and call it the convenience yield (cost of carry dominates when negative)
- Need a model for δ_t—e.g. $\delta_t = \delta \in \mathbb{R}$
- Given model for δ_t, can define

$$dD_t = \delta_t Y_t \, dt$$

- Ensure $dG_t^* = dY_t^* + dD_t^* = d\frac{Y_t}{A_t} + \frac{\delta_t Y_t}{A_t}\, dt$ is a Q-martingale
- For deterministic r and δ (2) becomes

$$F_{t,T} = Y_t \mathrm{e}^{(r-\delta)(T-t)}$$

  $\longrightarrow$ can infer δ by comparing spot and futures prices, or F_{t,T_1} and F_{t,T_2} for $T_1 \neq T_2$
- Proceed as before, by simulating risk-adjusted price model with "dividends"

Simple exercise (Homework)

Example of calibration of convenience yield to market prices: Assume constant parameters, $r = 0.005$, and observe the following prices from the CME (Chicago Mercantile Exchange; https://www.cmegroup.com/trading/metals/precious/gold.html)

	Market price	Today	Settlement
Gold spot	1473.2	19-Mar-20	19-Mar-20
Mar	1477.5	19-Mar-20	27-Mar-20
Apr	1481.6	19-Mar-20	28-Apr-20
May	1478.4	19-Mar-20	27-May-20

Use Excel or R (or some other method) to show that $\delta \approx -0.023$.

- Consider our simple example

$$\pi(S_t) = KS_t - F$$

- Value of date T cash flow

$$\mathrm{e}^{-r(T-t)}\mathrm{E}_t^Q\left\{KS_T - F\right\} = \mathrm{e}^{-r(T-t)}\left(K\mathrm{E}_t^Q\left\{S_T\right\} - F\right)$$

- Observe now that $\mathrm{E}_t^Q\left\{S_T\right\} = F_{t,T}$
- By treating the forward/futures price as ‘underlying’
  - Avoids having to deal with convenience yields; already reflected in $F_{t,T} = S_t\mathrm{e}^{(r-\delta)(T-t)}$
  - No need to simulate, as Q-expectation reduced to an observable market value $F_{t,T}$!
- Simplification works when cash flows are *linear* in the risk factors

- Two simple and popular models with constant convenience yield *rate* δ, and either

$$dY_t = \mu Y_t \, dt + \sigma Y_t \, dW_t$$

or

$$Y_t = \mathrm{e}^{X_t}, \qquad dX_t = \kappa[\theta - X_t] \, dt + \sigma \, dW_t$$

- Statistical issue which model is better
- However, more complex models necessary to fully capture important dynamics $\longrightarrow$ multi-factor models

- Gibson and Schwartz (1990) suggest the oil price model

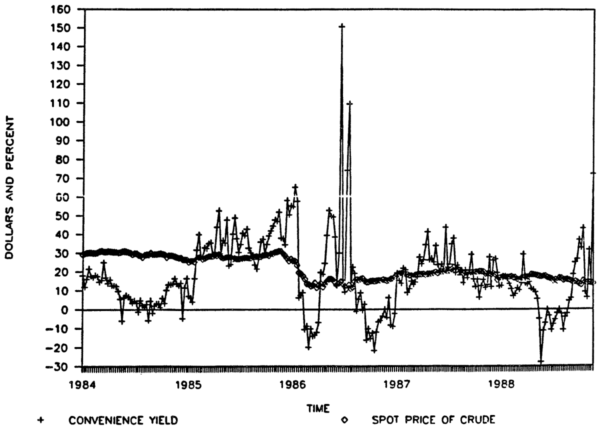
$$dS_t = \mu S_t\,dt + \sigma S_t\,dW_t^S$$
$$d\delta_t = \kappa(\theta - \delta_t)\,dt + \sigma_\delta\,dW_t^\delta$$

  where $\mathrm{corr}(dW_t^S, dW_t^\delta) = \rho$, or $dW_t^S dW_t^\delta = \rho\,dt$.

- Parameter estimates derived from time series regressions

$$\delta_{t+1} - \delta_t = \kappa\theta - \kappa\delta_t + \epsilon_{t+1}^\delta$$
$$\ln\left(\frac{S_{t+1}}{S_t}\right) = a + b\ln\left(\frac{S_t}{S_{t-1}}\right) + \epsilon_{t+1}^S$$

- Time series on δ derived from futures prices as described above



- Unclear from figure whether $\ln(S)$ is normal
- Mean reverting model for δ seems appropriate

- The lognormal model good if b insignificantly different from zero: $b = 0.002$, $t(b) = 0.03$
- $a = \mu - \frac{1}{2}\sigma^2$ not interesting (why?), thus not reported
    - Focus of paper on pricing: μ not important
    - If focus is risk management: μ essential!
- Parameters estimated to be

$$\sigma = 0.35, \quad \kappa = 16.08, \quad \theta = 0.19, \quad \sigma_\delta = 1.12$$

  which confirms time series for δ
    - is very volatile (high σ_δ)
    - quickly reverts to its long term mean (high κ)
- Correlation $\text{corr}(dW_t^S, dW_t^\delta) = \rho$ estimate as $\text{corr}(\epsilon^S, \epsilon^\delta)$

$$\rho = 0.32$$

- One more thing before we can do pricing...

- Observed earlier that non-traded risk factors must be adjusted for systematic risk according to a capital asset pricing model
- Gibson and Schwartz estimate the risk adjustment for δ to be $\lambda^{\delta} = -1.796$
- Risk adjusted model is thus

$$dS_t = (r - \delta_t)\, S_t\, dt + \sigma S_t\, dW_t^{S,Q}$$
$$d\delta_t = \left[\kappa(\theta - \delta_t) - \sigma_\delta \lambda^{\delta}\right] dt + \sigma_\delta\, dW_t^{\delta,Q}$$

where $dW_t^{S,Q} dW_t^{\delta,Q} = \rho\, dt$ (adjusting for risk does not affect covariance structure of Brownian motions)

- The most "sophisticated" extension of Gibson and Schwartz is the three factor model of Casassus and Collin-Dufresne (2005)
- In some markets/at some horizons important to model seasonal variations in prices
  - Typically modelled using trigonometric functions
  - Must ensure they don't get 'lost in translation' from P to Q
    - By for instance incorporating it in the convenience yield
    - For instance as in Cortazar and Naranjo (2006)
  - Ignoring seasonalities when they are important typically show up in 'crazy' parameter estimates, for instance for the CY which often plays the role of a 'free parameter'

Casassus, Jaime and Pierre Collin-Dufresne, "Stochastic Convenience Yield Implied From Commodity Futures and Interest Rates," *Journal of Finance*, October 2005, *60* (5), 2283–2331.

Cortazar, Gonzalo and Lorenzo Naranjo, "An N-factor Gaussian Model of Oil Futures Prices," *Journal of Futures Markets*, March 2006, *26* (3), 243–268.

Gibson, Rajna and Eduardo S. Schwartz, "Stochastic Convenience Yield and the Pricing of Oil Contingent Claims," *Journal of Finance*, July 1990, *45* (3), 959–976.

Schwartz, Eduardo S. and Mark Moon, "Rational Pricing of Internet Companies," *Financial Analysts Journal*, May–June 2000, *56* (3), 62–75.

_ **and** _ , "Rational Pricing of Internet Companies Revisited," *The Financial Review*, November 2001, *36* (4), 7–26.

ECO423 Advanced Derivatives

Stochastic Volatility

Jørgen Haug, NHH

Spring 2023

Learning outcomes

After this module you should

- know
  - the meaning of stochastic volatility and conditional heteroscedasticity
  - basic intuition for the effects of non-constant volatility
- be able to
  - change measure from P to Q for a system of price and volatility, for a given price of volatility risk λ^v
  - assess basic properties of/interpret the NGARCH process, for various values of key parameters
  - simulate Heston and Nandi's model, for risk-management and pricing

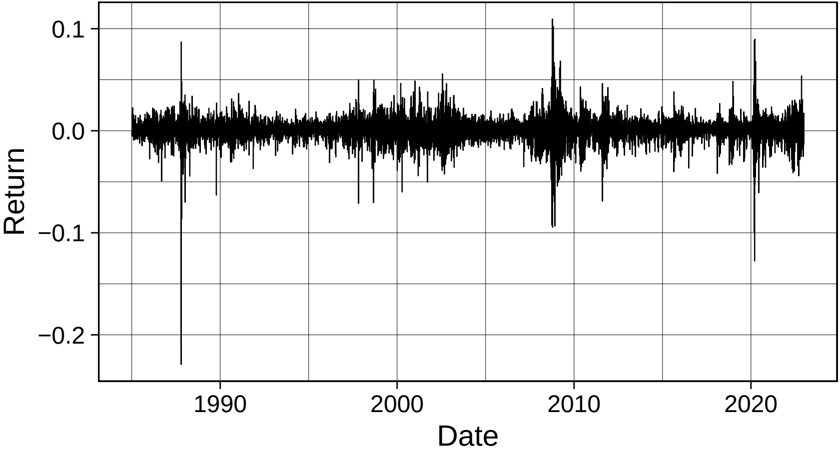
Relevant textbook chapters: H23, 27.2

Motivation

- Volatility forecasts are key inputs in option pricing and risk management
- Precision of volatility estimates increases with the frequency of observations
  - common to use daily observations
- Let's consider the S&P 500 index as an example

```
require(quantmod)
getSymbols("^GSPC", src = "yahoo", from = "1985-01-01",
   to = "2022-12-31")
SP500 = dailyReturn(GSPC$GSPC.Adjusted, type="log")
```

Daily S&P 500 returns



Stochastic volatility

The figure suggests S&P 500 volatility

- varies over time
- volatility clusters in a *random* fashion (no trend or seasonality)
- this kind of time-varying volatility is called *Conditional Heteroscedaticity* (CH)

Moreover

- volatility clusters seems to exhibit growth and decline phases,
- which suggests autocorrelated daily volatilities R_t^2, called Auto Regressive (AR)

Putting it together suggests an ARCH process

Basic estimation

- Common to use (daily) simple or log returns

$$R_{t+1} = \frac{S_{t+1} - S_t}{S_t} \qquad \text{or} \qquad r_{t+1} = \ln\left(\frac{S_{t+1}}{S_t}\right)$$

- Expected return estimate

$$\bar{R} = \frac{1}{T}\sum_{t=1}^{T} R_t$$

- Variance estimate; *unbiased*

$$\hat{v} = \frac{1}{T-1}\sum_{t=1}^{T}(R_t - \bar{R})^2,$$

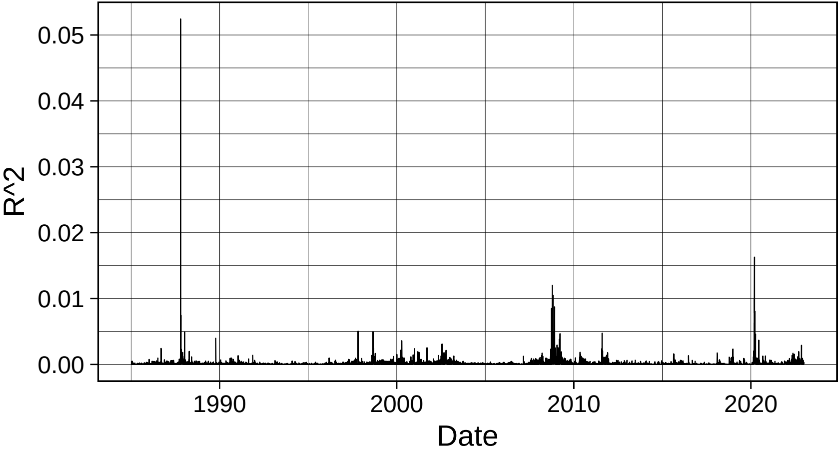
or *maximum likelihood*

$$\hat{v} = \frac{1}{T}\sum_{t=1}^{T}(R_t - \bar{R})^2$$

Estimation of daily volatility

- Daily expected returns are typically negligible
  - for instance S&P 500, `mean(SP500)`: 3.2834747×10^{-4}
- Common to simplify computations with $\bar{R} = 0$ in the maximum likelihood estimator
  - traditional approach, `sd(SP500)`: 0.0116529
  - simplified approach, `sqrt(mean(SP500^2))`: 0.0116569
  - or in annualized terms: 0.1849836 versus 0.1850474
- Thus, with daily $R_t^2 \sim$ realized daily variance

Stochastic volatility



Heston (1993)

The ARCH process allows volatility to be

- stochastic/random
- mean reverting

The following Ornstein-Uhlenbeck process captures these properties (Stein and Stein, 1991; Heston, 1993)

$$d\sigma_t = -\beta\sigma_t \, dt + \frac{\xi}{2} \, dW_t^v \tag{1}$$

This model implies

- long-run average volatility of zero
- strictly positive probability of negative volatility

Heston's stochastic volatility model

Heston suggests a simple fix to the weaknesses of the OU-model for volatility

- define the variance rate as $v_t = \sigma_t^2$, where σ_t is given by (1)
- use Itô's lemma to show that v_t follows the *square-root process*

$$dv_t = \gamma(\bar{v} - v_t)\, dt + \xi \sqrt{v_t}\, dW_t^v \tag{2}$$

  which is non-negative for $v_0 > 0$ because the dispersion term is proportional to $\sqrt{v_t}$
- redefine volatility as $\sigma_t = \sqrt{v_t} > 0$

Heston's stochastic volatility model

- Cannot rule out that volatility risk contains systematic risk
- Must in general allow for risk-adjustment when changing from P to Q
- In principle straight forward, using Girsanov's theorem
  $d\tilde{W}_t^v = dW_t^v + \lambda^v \, dt$

$$dv_t = \gamma(\bar{v} - v_t - \lambda^v \xi \sqrt{v_t})\, dt + \xi \sqrt{v_t}\, d\tilde{W}_t^v$$

- Want to avoid more or less arbitrary choice of λ^v
- Heston and Nandi (2000) offer a straight forward approach

Heston and Nandi (2000)

- Turns out that it is challenging to estimate Heston's model
- Heston and Nandi suggest NGARCH (Nonlinear Generalized ARCH) P-dynamnics, that converge to Heston's model

$$r_{t+1} = r + \lambda v_{t+1} + \sqrt{v_{t+1}} z_{t+1} \tag{3}$$
$$v_{t+1} = \omega + \alpha(z_t - \gamma\sqrt{v_t})^2 + \beta v_t \tag{4}$$

- r_{t+1} one-day *cum dividend* log return, r one-day riskless rate
- $z_{t+1} \sim \mathrm{N}(0, 1)$ observable at date $t + 1$
- v_{t+1} is known at date t
- λ allows $\mu_t = \mathsf{E}_t\{r_{t+1}\} = r + \lambda v_{t+1}$ to vary with volatility level $\sqrt{v_{t+1}}$

Heston and Nandi: parameters

- λ a risk premium parameter
- α determines kurtosis of returns
- γ determines nonlinear/asymmetric response of volatility to shocks z: Negative shocks have larger influence than positive shocks
  - can show that $\mathrm{Cov}_{t-1}(v_{t+1}, \ln(S_t)) = -2\alpha\gamma v_t$
  - γ thus controls the skewness of returns (for $\alpha > 0$)
- Setting $\gamma = 0$ yields the traditional (symmetric) GARCH(1,1) model

$$v_{t+1} = \omega + \alpha z_t^2 + \beta v_t$$

  - common to replace z_t with R_t when assuming $\bar{R} = 0$, as in textbook can view (z_t as a normalization of R_t)

Heston and Nandi: parameters

- Degree of mean reversion

$$\kappa = \beta + \alpha\gamma^2$$

  - (4) stationary if $\beta + \alpha\gamma^2 < 1$

- Annualized steady state/long-run volatility

$$\theta = \sqrt{252(\omega + \alpha)/(1 - \beta - \alpha\gamma^2)}$$

Heston and Nandi: Q dynamics

- $S_{t+1}|S_t$ is lognormal
- *Assume* the value of a European call with one period to expiration equals the BSM value
- The underlying asset must then have a Q-expected return equal to r
  - achieved by introducing $\tilde{z} = z + \lambda + \frac{1}{2}$
  - the appearance of $\frac{1}{2}$ is due to the formula for the expectation of a lognormal
- Get Q-dynamics

$$r_{t+1} = r - \frac{1}{2} v_{t+1} + \sqrt{v_{t+1}} \tilde{z}_{t+1} \tag{5}$$

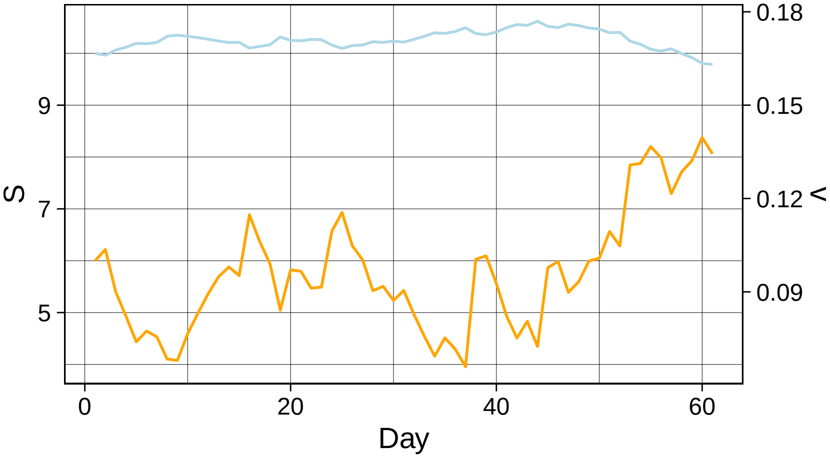
$$v_{t+1} = \omega + \alpha(\tilde{z}_t - \tilde{\gamma}\sqrt{v_t})^2 + \beta v_t \tag{6}$$

where $\tilde{\gamma} = \gamma + \lambda + \frac{1}{2}$

Heston and Nandi: Simulation under P

- Daily observations for the S&P 500 during 1992:01–1994:12 imply

```
omega = 5.02e-6; alpha = 1.32e-6; beta = 0.589
gamma = 421.39; lambda = 0.205
```



Heston and Nandi: Simulation under P

- Figure graphs `S[1,]` and `v[1,]` generated from R code similar to that for earlier Euler schemes:

```
set.seed(1)
N = 10; M = 60 # paths and days
v0 = (0.10)^2/252 # daily initial variance
r = 0.05/252 # riskless rate per day

S = matrix(rep(10, N*(M+1)), nrow = N); K = 12
v = matrix(rep(v0, N*(M+1)), nrow = N)
z = matrix(rnorm(N*M), nrow = N)
rS = matrix(rep(-1, N*(M+1)), nrow = N)

for(m in 1:M) {
   rS[,m] = r + lambda*v[,m] + sqrt(v[,m])*z[,m]
   v[,m+1] = omega + alpha*(z[,m] - gamma*sqrt(v[,m]))^2 + beta*v[,m]
   S[,m+1] = S[,m]*exp(rS[,m])
}
```

Heston and Nandi: Valuation under Q

- ATM put option valuation with `N = 1000`

```
gammaQ = gamma + lambda + 0.5

SQ = matrix(rep(10, N*(M+1)), nrow = N); K = 12
vQ = matrix(rep(v0, N*(M+1)), nrow = N)
rSQ = matrix(rep(-1, N*(M+1)), nrow = N)

for(m in 1:M) {
rSQ[,m] = r - 0.5*v[,m] + sqrt(v[,m])*z[,m]
vQ[,m+1] = omega + alpha*(z[,m] - gammaQ*sqrt(v[,m]))^2 + beta*v[,m]
SQ[,m+1] = SQ[,m]*exp(rSQ[,m])
}

K = 10 # strike price
gT = exp(-r*M/252)*pmax(K-SQ[,M+1], 0)
put.pv = mean(gT)
put.pv + c(qnorm(0.025), 0, qnorm(1-0.025))*sd(gT)/sqrt(N)
```

```
## [1] 0.1186203 0.1325792 0.1465382
```

Heston and Nandi: Valuation under Q

- Heston and Nandi derive formula for European call options
  - put options priced via put-call parity
- Formula given in terms of integral over a relatively complex function
  - can be implemented in Excel or R
- Can simulate (5) and (6) to price other payoff functions, and combine with LSMC to handle American derivatives

Estimation: Maximum Likelihood (MLE)

- Estimation is done under P
- By assumption $r_t \sim \mathrm{N}(r + \lambda v_t, v_t)$
- Likelihood of observing a time series of log returns thus

$$L = \prod_{t=1}^{T} \frac{1}{\sqrt{2\pi v_t}} e^{-\frac{1}{2}\frac{(r_t-\mu_t)^2}{v_t}} = \prod_{t=1}^{T} \frac{1}{\sqrt{2\pi v_t}} e^{-\frac{1}{2}\frac{(r_t-r-\lambda v_t)^2}{v_t}}$$

- Find the parameters that maximizes the likelihood of observed log returns
    - maximizing L equivalent to maximizing the log of L

Estimation: Log-likelihood (MLE)

- Log-likelihood function is

$$l = \ln(L) = \sum_{t=1}^{T} - \ln(\sqrt{2\pi}) - \ln(\sqrt{v_t}) - \frac{1}{2}\frac{(r_t - r - \lambda v_t)^2}{v_t}$$

$$\sim - \sum_{t=1}^{T} \ln(v_t) + \frac{(r_t - r - \lambda v_t)^2}{v_t} = - \sum_{t=1}^{T} \ln(v_t) + z_t^2$$

where the last equality follows from $r_t = r + \lambda v_t + \sqrt{v_t} z_t$ iff

$$z_t = (r_t - r - \lambda v_t)/\sqrt{v_t} \tag{7}$$

- Estimate model by

$$argmax_{\omega,\alpha,\beta,\gamma,\lambda} - \sum_{t=1}^{T} \ln(v_t) + z_t^2 = argmin_{\omega,\alpha,\beta,\gamma,\lambda} \sum_{t=1}^{T} \ln(v_t) + z_t^2$$

Estimation: Fitted volatility (MLE)

- Can use (3) and (7) to express v_t in terms of observable quantities only

$$\begin{aligned} v_{t+1} &= \omega + \alpha(z_t - \gamma\sqrt{v_t})^2 + \beta v_t \\ &= \omega + \alpha\left[(r_t - r - \lambda v_t)/\sqrt{v_t} - \gamma\sqrt{v_t}\right]^2 + \beta v_t \\ &= \omega + \alpha\frac{[r_t - r - (\lambda+\gamma)v_t]^2}{v_t} + \beta v_t \end{aligned}$$

- Given parameter estimates ω, α, β, γ, λ, and initial value $v_1 = \mathrm{Var}\,(r_1, ..., r_T)$ we can determine $v_2, v_3, .., v_T$ by *forward induction*

$$\begin{aligned} v_2 &= \omega + \alpha\frac{[r_1 - r - (\lambda+\gamma)v_1]^2}{v_1} + \beta v_1 \\ v_3 &= \omega + \alpha\frac{[r_2 - r - (\lambda+\gamma)v_2]^2}{v_2} + \beta v_2 \\ &\vdots \end{aligned}$$

Estimation: Implementation

- *Excel:* Must implement preceding estimation "by hand"
  - textbook explains how to estimate GARCH(1,1) model
  - preceding MLE minor variation of GARCH(1,1) approach
- *R:* Can use `fGarch` or `rugarch` packages to avoid programming MLE
  - the `tseries` package is more limited, but does handle for instance GARCH(p,q)

Relevance of non-constant volatility

- With deterministic volatility, $dS_t = rS_t\,dt + \sigma_t S_t\,d\tilde{W}_t$
  - $\ln(S_t)$ has variance $\int_0^T \sigma_t^2\,dt$
  - can rewrite variance as $\bar{\sigma}^2 T = \frac{1}{T}\int_0^T \left(\sigma_t^2\,dt\right) T$
  - European option price equal to BSM price with constant volatility, σ, replaced by average volatility, $\bar{\sigma}$
- S&P 500 volatility is stochastic and negatively correlated with underlying
  - volatility has negative beta $\longrightarrow$ higher drift under Q than under P
  - option-implied volatility higher than P-expected volatility
  - BSM-based analysis may indicate arbitrage opportunity, when in reality there is none