Probability and Stochastic Calculus

1 Intervals

$$[a, b] = \{a \leq x \leq b$$
$$(a, b) = \{a < x < b\}$$

2 Basic Probability - Bayes

Ω = sample space, set of all total possible outcomes of an experiment or trial
ω = an element of the sample space, an outcome or realisation of the trial
A = an event, for example rolling even number on a die such that $A = \{2, 4, 6$
$A \in \Omega$, the event belongs to or is an element of the sample space
Mutually disjoint = if A happens B cannot happen, no overlap in events

$$P(\cup A) = \sum P(A)$$

$$if A \in \Omega, P(A) > 0$$
$$Conditional = \frac{joint}{marginal}$$
$$Posterior = \frac{likelihood \times prior}{evidence} = P(model|data) = \frac{P(data|model)\, P(model)}{P(data)}$$
$$P(A \cap B) = P(B|A)\, P(A)$$
$$P(A|B) = \frac{P(B|A)\, P(A)}{P(B)}$$
$$P(B) = P(B|A)\, P(A) + P(B|A^c)\, P(A^c)$$

marginal probability of B contains all possible combinations of B

3 Sigma Algebra

σ algebra = subset of events, collecting events with determinable length

$$a\sigma\ F\ on\ \Omega if \Omega \in F\ and \varnothing \in F$$

$$A \in F\ and\ A^c \in F,\ A^c = \Omega$$

4 Measurable

σ is measurable if: $\mu : F \to [0, \infty]$ μ is a function that maps the set $F \to [0, \infty]$ μ is countably additive if $\mu(\varnothing) = 0$ Closed under complements Closed under countable unions Therefore μ is a measure on (Ω, F) to form the measure space (Ω, F, μ) for example, disjoint sets of the sigma algebra F denoted F and G which means sets of events in F that have no element in common are countably additive or $\mu(F \cup G) = \mu(F) + \mu(G)$ and $\mu(F \cap G) = \varnothing$ because if no element in common, probability of F and G is impossible finitely additive means can add probability of pairwise independent events for a finite set countably additive means can do the above for infinite set

5 Probability Space

Therefore a PROBABILITY SPACE is a measure space (Ω, F, P) where $P : F \to [0, 1]$ P, or probability is a function that maps events in the sigma algebra, which belong to the sample space (the entire universe of outcomes) to a value in the set [0,1]. It 'measures' the events of the sigma algebra according to the numerical set between [0,1].

$$P(\Omega) = 1$$

$$A \in F,\ P(A) > 0$$

$$P(\cup A) = \sum P(A)$$

6 Borel Algebra

Includes all subsets of R, real numbers or smallest sigma algebra containing all open intervals in R. Sets in the Borel algebra are borel sets.

7 Random Variable and Generated Algebras

A subset of events $C \in \Omega$ The sigma algebra generated by C is the smallest algebra containing all elements of C, $C \in F$ A random variable is F measurable function on the probability space (Ω, F, P) mapping onto

$$X : \Omega \to R$$

The sigma algebra generated by X is measurable, F measurable.

8 F Measurable

Function on (Ω, F, P) maps $f : \Omega \to R$

$$f^{-1}(B) = \{\omega \in \Omega : f(\omega) \in B\} \in F$$

The inverse of the function, for the set of outcomes which belong to the sample space, and for which the function belongs to the borel set B, belongs to the sigma algebra F Does this hold?

$$f(\omega) = \begin{cases} 1 \text{ if } \{\omega_1, \omega_2, \omega_3 \\ -1 if \{\omega_4, \omega_5, \omega_6 \end{cases}$$

$$f^{-1}(\omega) = \{\omega \in \Omega \ : f(\omega) \in B = \{\omega_1, \omega_2, \omega_3 \ \ or \ \ \{\omega_4, \omega_5, \omega_6$$

$$If \ \ \{\omega_1, \omega_2, \omega_3 \ \ \{\omega_4, \omega_5, \omega_6 \in \text{F, F } measurable$$

$$y \in R \Rightarrow \{\omega \in \Omega : f(\omega) \leq y\} \in F$$

If y is a real number, for an outcome of sample space, if the function of the outcome is less than or equal to y, is belongs to the sigma algebra F. Everything is measurable technically, but needs to be done so in a meaningful way – need to have a sigma algebra that contains and allows us to measure the information we care about / events / outcomes we care about, that you can ‘measure’ the probability of. What is this saying? The function f maps the sample space to the set of real numbers. The inverse of the function, where the function of an outcome of the sample space belongs to the borel set (and therefore to set of real numbers) belongs to the sigma algebra F. In other words, a function is F measurable if it maps to the real numbers (borel set), and the inverse of the function maps it back to the set of events (f sigma algebra). Once we know the random value $X(\omega)$ we know which of F has occurred.

9 Filtration

Information increases and does not decrease with time. Each sigma algebra at time t where s < t contains the sigma algebra s. A collection of sigma algebras $\{F_{t\, t \geq 0}$ is a filtration, for every

$s \leq t, F_s \subseteq F_t$ Random variables are adapted to the filtration if for every t, X_t is measurable with respect to F_t In other words, the function of the random variable maps the sigma algebra at time t to the set of real numbers, and the inverse of the function maps the borel set (real numbers) back to the sigma algebra of events. At every t, we know all information dependent on what has happened up to time t. We do not know the future, but we can measure the random variable at time t depending only on what has happened in the time up to t. At time t, the random variable X at time t can assign values to the sets of outcomes we know are possible at time t. These outcomes are dependent on what came before. It cannot assign values to sets of outcomes in the previous sigma algebra, because they do not map to one another.

10 Martingale and Stochastic Processes / Random Variables

A stochastic process is a 'collection of random variables' indexed by time (either discrete or continuous). What does that actually mean? It describes a process where at every point in time t, we have a realisation of the random variable (which as we have seen before, is a process on the probability space that maps a sigma algebra of events to the set of real numbers). So for example, a stock price following a stochastic process means at every point in time, we have $S_0, S_1, S_2 \ldots$ where the price at every point in time is a random variable. Formally, a collection of random variables: $\{X_t, t \in T$

$$X(t, \omega) : T \times \Omega \to R$$

The random variable is a function of time and element of the sample space (outcome) that maps time and the outcomes of a trial to the set of real numbers. For discrete time $T = [0, +\infty]$ For continuous time $T = [0, T]$ Martingale IF:

$$E(X_{t+1}|F_t) = E[X_t]$$

Adapted to the F_t You could say that the expectation of the future random variable is anchored by the value or the information that we have today

11 Order of Infinitesimals

infinitesimal is a function s.t. $lim : x \to a$ and $f(x) = 0$ infinitesimal functions u and v if $\frac{u}{v} = +ve$ and $0 < |x| < \epsilon$ u and v are of the same order $O(u) = O(v)$, u is $O(v)$ If $\frac{u}{v} = 0$ or $\frac{v}{u} = \infty$ the function v is greater than the function u v is of lower order than u (not as small as?) u is of higher order than v (smaller than?) $O(u) > O(v)$, u is o (v) x^n is O (x) :when are they the same order? if x tends to 0, and n is<0 $x^n \to \infty$

12 Jensen's Inequality

$$E[g(X)] \geq g(E[X])$$

The expectation of a function is greater than the function of the expectation If this is true, it is a **convex function** Convex if midpoint of interval of a function, does not exceed arithmetic mean of interval bounds

$$f\left(aK + (1-a)K'\right) \leq af(K) + (1-a)f\left(K'\right)$$

if a call is convex, by PCP put is also convex

$$C\left(\frac{1}{2}K + d + \frac{1}{2}K - d\right) \leq \frac{1}{2}C(K+d) + \frac{1}{2}C(K-d)$$

$$= 2C(K) \leq C(K+d) + C(K-d)$$

To prove, do the payoff tables for the above:

$$0 \leq C(K+d) + C(K-d) - 2C(K)$$

Positive second derivative also proves convexity Function increasing or decreasing in K, S, whatever parameter also determined through the first derivative (or just by looking at the function)

13 Mean Value Theorem

If you have two points on a graph, the secant (line that joins them) is parallel to a tangent point on the graph if: The graph is continuous between [a,b] (greater than or equal) The graph is differentiable between (a,b) (greater than) random variables with mean 0 and variance AA' = sigma

$$f'(x) = \frac{f(b) - f(a)}{b - a}$$

$$f'(x)(b-a) = f(b) - f(a)$$

14 Total and Quadratic Variation

If we have a function, and we split it into steps, we can find the total variation of the function over an interval. Say we have an interval of b, and we want to split it into n half sub intervals.

$$stepsize = \frac{b}{2^n}$$

$$variationininterval = f\left(x_{\frac{kb}{2^n}}\right) - f\left(x_{\frac{(k-1)b}{2^n}}\right)$$

$$quadratic variation = |f\left(x_{\frac{kb}{2^n}}\right) - f\left(x_{\frac{(k-1)b}{2^n}}\right)|^2$$

$$mean value theorem = f\left(x_{\frac{kb}{2^n}}\right) - f\left(x_{\frac{(k-1)b}{2^n}}\right) = f'(x)\left(x_{\frac{kb}{2^n}} - x_{\frac{(k-1)b}{2^n}}\right)$$

as these are summations, they can be written as Riemann integrals

$$\int f'(x)\,dx$$

we know that dx is the length of the interval, which is a constant

$$(dx)^2 = \left(x_{j+1} - x_j\right)^2 = \left(\frac{b}{2^n}\right)^2 = \left(x_{j+1} - x_j\right)\left(\frac{b}{2^n}\right) = \frac{b}{2^n}dx$$

$$\frac{b}{2^n}\int f'(x)^2\,dx = \int f'(x)^2\,dx^2$$

a more intuitive way of looking at this is assuming we have a function f (x)

want to define the total quadratic variation between $[0, T]$

can split this into intervals, which are $\left(t_{j+1} - t_j\right)$ which do not have to be equally spaced

we can see that as we increase number of intervals, they become infinitely small

$$\max\left(t_{j+1} - t_j\right) \to 0 as n \to \infty,$$

$$\left[f\left(t_{j+1}\right) - f\left(t_j\right)\right]^2 = f'\left(t'_{j+1}\right)^2\left(t_{j+1} - t_j\right)^2$$

$$\left(t_{j+1} - t_j\right)^2 = \left(t_{j+1} - t_j\right)\left(t_{j+1} - t_j\right)$$

$$\left(t_{j+1} - t_j\right)\left(t_{j+1} - t_j\right) \leq \left(t_{j+1} - t_j\right)\ max\left(t_{j+1} - t_j\right)$$

because we are specifying the maximum interval length, $\left(t_{j+1} - t_j\right)$ has to be equivalent

$$\max\left(t_{j+1} - t_j\right)\int f'(x)^2\,dt$$

$$n \to \infty, \max\left(t_{j+1} - t_j\right) = 0 therefore quadratic variation is 0$$

15 Quadratic Variation in Brownian Motion

$$dW_t = W_{j+1} - W_j \sim N(0, dt)$$

$$X = \left(W_{j+1} - W_j\right)$$

$$E[X] = 0$$

$$E\left[\left(W_{j+1} - W_j\right)^2\right] = E\left[X^2\right] = E\left[X^2\right] - E[X]^2 \text{ because E}[X] \text{ is 0}$$

$$= Var(X) = Var\left(W_{j+1} - W_j\right) = dt$$

$$Var\left(\left(W_{j+1}-W_j\right)^2\right) = E\left[X^2\right] - E\left[X\right]^2 =$$

$$E\left[\left(W_{j+1}-W_j\right)^4\right] - E\left[\left(W_{j+1}-W_j\right)^2\right]^2 =$$

$fourthmomentisdefinedas 3 \times Var^2$ for normal r.v. mean 0

$$E\left[\left(W_{j+1}-W_j\right)^2\right] = dt$$

$$3\,(dt)^2 - (dt)^2 = 2\,(dt)^2$$

$$(dt)^2 = (dt)\,(dt) \;\; and \;\; (dt) = \max\left(t_{j+1}-t_j\right)$$

sampled quadratic variation therefore has expectation T and variance 0

therefore expected quadratic variation or d$W_t dW_t = dt$

$as\;\left(t_{j+1}-t_j\right)$ grows smaller, the variance $2\,(dt)^2$ grows infinitesimally small

$therefore\;(dW_t)^2$ becomes increasingly deterministic (variance negligible) approaching mean

law of large numbers dictates that as increments increase $(dW_t)^2$ reverts to mean

$$\left(W_{j+1}-W_j\right)^2 \approx dt$$

$$\frac{\left(W_{j+1}-W_j\right)^2}{dt}$$

$\dfrac{dW_t}{\sqrt{dt}} = \dfrac{x-\mu}{\sigma} = N\,(0,1)$ because E$\,(dW_t)$ is 0, and Var$\,(dW_t)$ is dt

$$t_j = \frac{jT}{n}\;therefore\;t_{j+1}-t_j = \frac{(j+1)\,T}{n} - \frac{jT}{n} = \frac{T}{n}$$

$$\frac{Y^2T}{n} = \left(W_{j+1}-W_j\right)^2$$

LLN dictates this reverts to mean, which is 1 times T (squaring a standard normal)

Brownian motion accumulates quadratic variation at rate one unit per time

$$\left(W_{j+1}-W_j\right)\left(t_{j+1}-t_j\right) = 0$$

$$\left(t_{j+1}-t_j\right)\left(t_{j+1}-t_j\right) = 0$$

tends to 0 as maximum length of subinterval goes to 0

Look at p104 of Shreve Stochastic Calculus II

16 Reflection Principle

Let's assume there is a maximum level for the Brownian motion.

$$M_t = \max\{W_s\ for 0 \leq s \leq t$$

$$P\{M_t \geq a \text{ for a>0}$$

$$P\{W_t \geq a \text{ can be split accordingly}$$

$$P\{W_t \geq a = P\{W_t \geq a, M_t \geq a + P\{W_t \geq a, M_t < a$$

these are mutually exclusive and impossible, if $M_t < a$ W_t cannot be above a

$$P\{W_t \geq a, M_t \geq a = P(AB)$$

$$P(AB) = P(A|B)\,P(B)$$

$$P\{W_t \geq a, M_t \geq a = P\{W_t \geq a|M_t \geq a\ P\{M_t \geq a$$

$$P\{W_t \geq a|M_t \geq a + P\{W_t < a|M_t \geq a = fullsetofpossibleoutcomesgivenM_t \geq a$$

$$thereforethismustsumto 1 - W_t \text{ must be somewhere if } M_t \geq a$$

we can see that the probability of being above a or below a is the same by looking at reflected paths

$$\text{therefore P}\{W_t \geq a|M_t \geq a = \frac{1}{2}P\{M_t \geq a$$

Joint distribution of the stock hitting the maximum (or minimum) (price of an option is its discounted expected payoff, this distribution required to calculate the expected payoff) Use Girsanov twice – changing to risk neutral measure Q, and changing to a measure that eliminates drift A vanilla option can be expressed as equivalent to combination of its down and out and down and it versions.

17 Brownian Motion

Probability distribution over the set of continuous functions $B : R_{\geq 0} \rightarrow R$

$$P(B_0 = 0) = 1, thedistributionstartsat 0, nonnegative \text{ (stock prices lower bounded)}$$

Independent increments, with each increment normally distributed

$$B(t) - B(s) \sim N(0, dt)$$

What does that mean? There EXISTS a probability distribution that DESCRIBES the path of a Brownian Motion. This distribution says that for intervals, the values of the distribution have a mean of 0 and variance of dt. Also, the distributions of the independent increments are independent. Probability distribution is over PATHS. Not over POINTS because it is a continuous processes, not discrete. Comparable to taking the 'limit' of a simple random walk.

If we take as many observations of a discrete time process as we can and it converges to a Brownian motion, then we can model it as such. The distribution crosses the t axis infinitely often. Does not deviate much from $t = y^2$

mean 0, variance t, sd t

Not differentiable. According to quadratic variation, the total variation of a function is the total change in value over the time period. The quadratic is the sum of the variation across all intervals (which can be infinitesimally small and limitless in number) squared. When we take the limit of this, which is the integral of the differential of the function, it will be 0. This is not the case with a Brownian motion, where as t tends to infinity, variation only increases. The result is:

$$(dB)^2 = dt$$

https://www.youtube.com/watch?v=AptlhWEgOto What if we want to differentiate the function:

$$f(B_t)$$

$$\text{If f was } B_t, df = \frac{dB_t}{dt}\text{dt but we know differentiation is impossible}$$

https://www.youtube.com/watch?v=PPl-7_RL0Ko 47:32 for differentiation not possible We discussed in the workshop that differentiation is not possible because path is constantly varying at every infinitesimal interval. No matter where you attempted to take a differential, there would be an error as the actual path will deviate from the tangent you draw, this is the element of randomness. We correct for this with Ito's Lemma which allows us to take the 2nd order term by accounting for quadratic variation (which does not disappear for a Brownian motion) What about:

$$df = f'(B_t)\, dB_t$$

$$\text{We can obtain d}B_t\ (\delta B_t)\ \text{ and f}\,(B_t)\ \text{ is a differentiable function}$$

$$\text{Not possible because }\ (dB_t)^2 = dt$$

$$f(t+x) \approx f(t) + f'(t)(x) + \frac{f''(t)}{2}(x)^2 \ldots$$

$$f(t+x) - f(t) \approx f'(t)(x) + \frac{f''(t)}{2}(x)^2 \ldots \approx x, youuseTaylorexpansiontofinddifference$$

$$f(B_t) - f(B_{t-1}) = f'(B_t)(B_t - B_{t-1}) + f''(B_t)(B_t - B_{t-1})^2$$

$$(dB_t)^2 = dt$$

$$df = f'(B_t)\, dB_t + \frac{f''(B_t)}{2}dt = Ito' \text{s Lemma, because 'x' is d}B_t \text{ in this case}$$

Mnemonics: dt dt = 0 dt dWt = 0 dWt dWt = dt

18 Ito's Lemma

$$\text{Say we have f}\,(t,x)\, where t = t and x = B_t$$

$$f(t,x) + f_t'(t,x)\,dt + f_x'(t,x)\,dx + \frac{1}{2}f_{tt}''(t,x)\,dt^2 + f_{tx}''(t,x)\,dxdt + \frac{1}{2}f_{xx}(t,x)\,dx^2\ldots$$

$$f_t = partial derivative with respect to t\,\text{(differentiate t)}$$

$$f_x = partial derivative with respect to x which is B_t\ \text{in this case}$$

$$f_{xx} = 2nd order partial derivative with respect to x, dx^2 = dt because\,(dB_t)^2 = dt$$

$$f_{tx} = 2nd order partial derivative with respect to t then x$$

Essentially we can drop all second order terms except for dx^2 because they are 0 (quadratic variation)

$$f(t+\Delta, x+\Delta) = f(t,x) + f_t'(t,x)\,dt + f_x'(t,x)\,dx + \frac{1}{2}f_{xx}(t,x)\,dx^2$$

$$df = f(t+\Delta, x+\Delta) - f(t,x) = f_t'(t,x)\,dt + f_x'(t,x)\,dx + \frac{1}{2}f_{xx}(t,x)\,dx^2$$

$$= \left(f_t'(t,x) + \frac{1}{2}f_{xx}'(t,x)\right)dt + f_x'(t,x)\,dx$$

$$\text{because d}x^2 = (dB_t)^2 = dt$$

Now consider a stochastic process with drift and diffusion term:

$$X_t = \mu t + \sigma B_t$$

$$dX_t = \mu dt + \sigma dB_t$$

The variable is dependent on time as well as the Brownian motion

If we replace B_t *with* X_t in Ito$'$s Lemma and using mnemonics above:

$$df = \left(f_t'(t,x) + f_x'(t,x)\,\mu + \frac{1}{2}f_{xx}'(t,x)\,\sigma^2\right)dt + f_x'(t,x)\,\sigma dB_t$$

$$Remember : (dX_t)^2 = (\mu dt)^2 + 2\,(\mu\sigma dt dB_t) + (\sigma dB_t)^2$$

$$= 0 + 2\,(0) + \sigma^2 dt$$

19 Geometric Brownian Motion

$$dX_t = \mu X_t dt + \sigma X_t dB_t$$

$$\text{Appling Ito}'\text{s Lemma}$$

$$(\mu X_t dt + \sigma X_t dB_t)^2 = \sigma^2 X_t^2 dt$$

$$df = \left(f_t'(t,x) + f_x'(t,x)\mu X_t + \frac{1}{2} f_{xx}'(t,x)\sigma^2 X_t^2\right) dt + f_x'(t,x)\sigma X_t dB_t$$

$$\text{Taking f}(X_t)\ \ as\ln(X_t)$$

$$df = \left(0 + \frac{1}{X_t}\mu X_t + \frac{1}{2}\frac{-1}{X_t^2}\sigma^2 X_t^2\right) dt + \frac{1}{X_t}\sigma X_t dB_t$$

$$dln(X_t) = \left(\mu - \frac{1}{2}\sigma^2\right) dt + \sigma dB_t$$

19.1 Stock Prices

We want to model the percentile difference in stock prices, and for them to be normally distributed and following a Brownian motion.

$$\frac{dS_t}{S_t} = dB_t$$

$$\frac{dS_t}{dB_t} = S_t$$

$$\frac{dy}{dx} = e^x = y$$

$$S_t = e^{B_t}$$

$$\text{Implies d}S_t = e^{B_t} dB_t = S_t dB_t$$

$$\frac{dS_t}{S_t} = \sigma dB_t$$

We want to model the return on stocks as a Brownian motion, with some variance – but when we take the differential, we have a drift term: we assume the expected value should be 0

$$S_t = e^{\sigma B_t}$$

$$\text{Using Ito}'\text{s Lemma, f}(t, B_t) = e^{\sigma B_t}$$

$$dS_t = 0 + \sigma e^{\sigma B_t} dB_t + \frac{1}{2}\sigma^2 e^{\sigma B_t} dt$$

$$We have \sigma dB_t + \frac{1}{2}\sigma^2 dt, an additional drift term - we can remove drift with e^{-\frac{1}{2}\sigma^2 t + \sigma B_t}$$

For a geometric Brownian motion, if we instead have $e^{\mu - \frac{1}{2}\sigma^2 t + \sigma B_t}$ this will give us:

$$dS_t = 0 + \sigma e^{\sigma B_t} dB_t + \frac{1}{2}\sigma^2 e^{\sigma B_t} dt - \frac{1}{2}\sigma^2 e^{\sigma B_t} dt + \mu e^{\sigma B_t} dt$$

20 Ito Process

Can be defined as a stochastic process with $X_t = X_0 + \mu t + \sigma B_t$

$$dX_t = \mu dt + \sigma dB_t$$

Has properties of Riemann integrals – for each interval which is a partition of time $[0, T]$, $X_T = X_0 + \sum \mu (t_i - t_{i-1}) + \sum \sigma (t_i - t_{i-1})$ Which is the same as:

$$X_T = X_0 + \int \mu dt + \sigma dB_t =$$

$$X_0 + \int \mu dt + \int \sigma dB_t \; where \; \int \sigma dB_t \text{ is stochastic integral}$$

$$\text{Taking dln}(X_t) = \left(\mu - \frac{1}{2}\sigma^2\right) dt + \sigma dB_t$$

$$\ln X_t - lnX_0 = \int \left(\mu - \frac{1}{2}\sigma^2\right) dt + \int \sigma dB_t$$

$$= lnX_0 + \left(\mu - \frac{1}{2}\sigma^2\right) t + \sigma B_t$$

$$= X_0 \exp\left[\left(\mu - \frac{1}{2}\sigma^2\right) t + \sigma B_t\right]$$

Basic properties are: Linearity and the Martingale Property

$$E[B_t|F_s] = E[(B_t - B_s) + B_s|F_s]$$

but because $B_t - B_s$ is independent of F_s, $\mathrm{E}[B_t - B_s|F_s] = E[B_t - B_s] = 0$

$$E[(B_t - B_s) + B_s|F_s] = E[B_t - B_s|F_s] + E[B_s|F_s] = B_s$$

Replacing $W_t \; with \; W_t - W_s + W_s$ allows us to prove more easily?

$$E[W_s W_t] = min\{s, t\}$$

$$E\left[\int_0^t X_u dB_u | F_s\right] = \int_0^s X_u dB_u$$

Ito isometry The expectation of the Ito integral of the Ito process eg. σdB_t which is an adapted process

$$E\left[\left(\int X_u dB_u\right)^2\right] = E\left[X_u^2\right] du$$

Therefore we can calculate E(X) and E(Xˆ 2) and therefore variance or sigma When is an Ito process a martingale? No tendency to go up or down but remain fixed in expectation. Therefore if the drift term is 0, it is a martingale – because if it had a drift term, it would move away from the value we currently know in the future, the whole point of a martingale and stochastic process is that we do not know the future, hence 'fair game', therefore all we can expect for the future is what we've obtained today. Another way is to think of it as Brownian motion with expected value 0 and variance dt, introducing a drift term changes the expected value. Or, in terms of stochastic process, the drift process is deterministic, therefore if it remained, we could predict the future. By removing it, process depends only on the randomness of the Brownian Say I have a financial derivative that follows stochastic process, if it has a drift term and I can engineer it to remove variance, I can make an arbitrage – fixed return for no variance or risk

Proof of Martingale:

$$E\left[S_T\right] = S_t$$

$$\text{or prove E}\left[dS_t\right] = 0$$

$$E\left[dW_t\right] = 0 because dW_t \sim N(0, dt)\text{, drift must be 0 for E}\left[dS_t\right] = 0 to hold$$

20.1 Integration

Integral of adapted process (process that depends only in information from time 0 to t) If we have Brownian motion it is normally distributed:

$$B_t \sim N(0, t)$$

$$\text{With a constant:}$$

$$X_t = \sigma B_t \sim N\left(0, \sigma^2 t\right)$$

$$\int \sigma dB_t$$

$$At each time when you take the integral, you will have a normal distribution - sum of normal distributions is a normal$$

$$\text{If f is a function dependent only on t } (NOT\ B_t):$$

$$X_t = \int f(t)\, dB_t \text{ has a normal distribution}$$

21 Stochastic Calculus

To find the derivatives, you must always represent the stochastic process as a function, then use Ito's lemma to differentiate the **FUNCTION** not the process.

$$d(X_tY_t) : think of the function f(X_t, Y_t) = X_tY_t, \text{ then differentiate}$$

$$Y_t dX_t + X_t dY_t + 0 + 0 + 1 dX_t dY_t$$

$$d\left(\frac{X_t}{Y_t}\right) = f(X_t, Y_t) = \frac{X_t}{Y_t} = X_t Y_t^{-1}$$

$$Y^{-1} dX_t - XY^{-2} dY_t + \frac{1}{2} 2XY^{-3} dY_t^2 - Y^{-2} dX_t dY_t$$

$$d(X_t)^a = f(t, X_t) = X_t^a$$

Remember to substitute in the dynamics or the dX_t, dY_t terms after

Remember to always define the function you are differentiating first, and the relevant parameters.

22 Differential Equations

When we are saying 'solve' the differential equation or find the solution to it, what are we asking for? If a stock follows a geometric Brownian motion such that:

$$dS_t = \mu S_t dt + \sigma S_t dB_t$$

What is the function of S_t such that when differentiated, gives us the geometric Brownian motion formula? If we can find this function, we have our solution to the differential equation. To verify something is a solution to a stochastic differential, the differential of the function (using Ito's Lemma) should give the original stochastic differential.

23 Ornstein-Uhlembeck Model

$$dX_t = \theta(\mu - X_t) dt + \sigma dW_t$$

$$\mu and \sigma are constants - the unconditional mean and instantaneous volatility$$

$$this models the short term interest rate, and is a mean - reverting process$$

$$\theta is the rate of mean reversion$$

24 Radon-Nikodym Derivative – Change of Measure (RNP)

If I have two Brownian motions, one with drift term and one without, is there a measure such that:

$$P(\omega) \text{ is with and } \widetilde{P}(\omega) \text{ is without}$$

$$Probabilityspace\,(\Omega, P)$$

$$Z(\omega)\,\widetilde{P}(\omega) = P(\omega)$$

$$\text{If P}(\omega)\ \ and\ \widetilde{P}(\omega) \text{ are equivalent, yes}$$

$$P(\omega) > 0 \Leftrightarrow \widetilde{P}(\omega) > 0$$

$$Z(\omega) = \frac{d\widetilde{P}(\omega)}{dP(\omega)} = e^{-\mu\omega(T)-\mu^2\frac{T}{2}}$$

25 Girsanov Theorem

Say we have a random, normally distributed variable:

$$\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

$$\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2} whereN \sim (0, 1)$$

$$p(x) = \int_a^b f(x)\,dx = \frac{1}{\sqrt{2\pi}}\int_a^b e^{-\frac{1}{2}x^2}dx$$

assuming we shift the boundaries but keep the variable constant

$$\frac{1}{\sqrt{2\pi}}\int_{a-\mu}^{b-\mu} e^{-\frac{1}{2}x^2}$$

using change of variable $(substitution)$, so that the boundaries remain constant

$$whenvariablexislowerboundary = a - \mu$$

$$thevariabley = a$$

$$x = y - \mu, y = x + \mu$$

$$whenxisupperboundary = b - \mu$$

$$thevariabley = b$$

$$x = y - \mu, y = x + \mu$$

$$\text{Therefore shift from N}(0, 1)\,toN(\mu, 1)$$

$$equatedx = dy, andsubinyforx$$

$$\frac{1}{\sqrt{2\pi}}\int_a^b e^{-\frac{1}{2}(y-\mu)^2}dy$$

$$\frac{1}{\sqrt{2\pi}}\int_a^b e^{-\frac{1}{2}\left(y^2-y\mu+\mu^2\right)}dy$$

$$\frac{1}{\sqrt{2\pi}}\int_a^b e^{-\frac{1}{2}y^2}e^{y\mu+\frac{1}{2}\mu^2}dy$$

$$\frac{1}{\sqrt{2\pi}}\int_a^b e^{-\frac{1}{2}y^2} = p(y)$$

$$\frac{1}{\sqrt{2\pi}}\int_a^b p(y)\, e^{y\mu+\frac{1}{2}\mu^2}dy$$

$$p(y)\, dy = dP(y)$$

$$z(y) = \frac{dQ(y)}{dP(y)}$$

$$Remembertheintegralistheantiderivative - sowhenyouintegrateafunction$$

the function has itself is a differential

$$\int_a^b f(y) = F(y) - F(a)$$

$$f(y) = \frac{dF(y)}{dy}$$

$\frac{dQ}{dP}$ is therefore a density, taking the integral gives a probability

dQ is over dP, because we are taking the antiderivative (so the function is a derivative under Q)

over dP, because the change is in respect to the original probability distribution $(boundaries)$ of P

we denote $\frac{dQ}{dP} = z = e^{y\mu+\frac{1}{2}\mu^2}$

$$\int_a^b z(x)\, dP(x) = \int_a^b \frac{dQ(x)}{dP(x)}dP(x) = E_P[Z(x)]$$

same as expected value of continuous variable $\int_a^b xf(x)\, dx$

under P because dP $(x) = p(x)\, dx$

$$\int_a^b h(x)\, z(x)\, dP(x) = \int_a^b \frac{dQ(x)}{dP(x)}dP(x)\, h(x) = E_Q[h(x)]$$

We can express the integral

$$\int f(x)$$

$$A$$

$$\int_{-\infty}^{\infty} I_{[a,b]}z(x)\,dP(x) = E_P\left[I_{[a,b]}\frac{dQ(x)}{dP(x)}\right] = E_P[Z(x)]$$

the indicator function means all integrated values outside boundaries a and b set to 0

The most intuitive way of why I think that the Q can be equal to the P is that Q tells us how the variable shifts under the Q measure, then allows us to take the new probability of that, but overall the outcome is the same. All we've done is shift the boundaries of the distribution (but the probabilities will also be shifted to account for that..?)

25.1 For a Brownian Motion

recall that it is the CHANGE in Brownian we are interested in

$$dB_t \sim N(0, dt)$$

shifts in the boundaries will also vary with dt

$$\frac{1}{\sqrt{2\pi}}\int_{(a-\mu)dr}^{(b-\mu)dt} e^{-\frac{1}{2dt}(x_i-x_i-1)^2}$$

$$x = y - \mu dt$$

$$\frac{dQ}{dP} = e^{\mu B_t - \frac{t}{2}\mu^2} = z(t)$$

replace the normal distribution with a standard brownian motion

$$dP(x) = \frac{1}{\sqrt{2\pi}}\int_a^b e^{\frac{1}{2t}B_t^2}$$

25.2 Changing Drift

$$dX_t = \mu dt + \sigma dB_t$$

we want to change this to a process with no drift under measure Q

$$\sigma d\widetilde{B}_t = \mu dt + \sigma dB_t$$

$$\sigma d\widetilde{B}_t = \sigma\left(\frac{\mu}{\sigma}dt + dB_t\right)$$

$$d\widetilde{B}_t = \frac{\mu}{\sigma}dt + dB_t$$

$$d\widetilde{B}_t - \frac{\mu}{\sigma}dt = dB_t$$

$$\widetilde{B}_t - \frac{\mu}{\sigma}t = B_t$$

$$we shift the boundaries by \mu t or Y(t) = \mu t$$

$$\text{in this case, we have shifted by } \Upsilon(t) = -\frac{\mu}{\sigma}t$$

$$\text{sub into} \frac{dQ}{dP} = e^{\mu\left(\widetilde{B}_t - \frac{\mu}{\sigma}t\right) - \frac{1}{2}\left(\frac{\mu}{\sigma}\right)^2 t}$$

$$\frac{r}{\sigma}dt + d\widetilde{B}_t = \frac{\mu}{\sigma} + dB_t$$

26 Deriving Black Scholes with Girsanov

For n(d1) we are changing the numeraire from the MMA to the stock itself. We know what using the MMA as numeraire creates a martingale. The same applies when using the stock as a numeraire. Which gives us the different probability for d1 vs d2.

$$\frac{1}{(2\pi dt)^{\frac{n}{2}}} \int_{(a-\gamma)dr}^{(b-\gamma)dt} e^{-\frac{1}{2dt}(dx)^2} \text{ where x is a standard brownian motion with gaussian dist}$$

$$\text{this is technically a multivariate gaussian, time and the brownian are changing}$$

$$wecanshiftthebrownianmotionby\gamma dt$$

$$\int e^{-\frac{1}{2dt}(dx+\gamma dt)^2}$$

$$+\gamma dtasweshiftRNPupwards$$

$$\int e^{-\frac{1}{2dt}\left(dx^2+2dx\gamma dt+\gamma^2 dt^2\right)}$$

$$\int e^{-\frac{1}{2dt}(dx)^2+\left(-dx\gamma-\frac{1}{2}\gamma^2 dt\right)}$$

$$\int e^{-\frac{1}{2dt}dx^2} = standardbrownian$$

$$-\int dx\gamma - \int \frac{1}{2}\gamma^2 dt = x\gamma - \frac{1}{2}\gamma^2\text{t as integral removes d terms}$$

$$= \frac{dQ}{dP} = changeofmeasureshiftingbrownianmotionbyprocess\gamma$$

$$therefore$$

$$dB_t + \gamma dt = d\widetilde{B}_t$$

$$B_t - B_0 + \int_0^t \gamma ds = \widetilde{B}_t - \widetilde{B}_0$$

$$\text{we also want to prove that the shifted brownian is still a Martingale, prove 0 drift}$$

$$\sigma d\widetilde{B}_t = \mu dt + \sigma dB_t$$
$$\widetilde{B}_t = \frac{\mu}{\sigma} t + Bt$$
$$\gamma = \frac{\mu}{\sigma} = processshiftingbrowniantoamartingaleunderQmeasure$$
$$itismartingaleunderQmeasurebecausetheQmeasureshiftsbrownianby\gamma$$
this shifted brownian has 0 drift, meaning it is a martingale
$$therefore \frac{dQ}{dP} = \exp\left(-\frac{(\mu - r)}{\sigma} B_t - \frac{1}{2}\frac{(\mu - r)^2}{\sigma^2} t\right)$$
$$the - rcomesfromthefactthatwearelookingattheDISCOUNTEDSTOCKPRICE$$

26.1 Novikov Condition

A process that satisfies:
$$E\left[\exp\left(\frac{1}{2}\int_0^T \theta^2 ds\right)\right] < \infty$$
$$M_t^\theta = \exp\left(-\int_0^t \theta_s dX_s - \int_0^t \theta_s^2 ds\right) \text{ is an exponential martingale}$$
$$\text{therefore we see that the discounted stock price } \exp\left(-\frac{(\mu - r)}{\sigma} B_t - \frac{1}{2}\frac{(\mu - r)^2}{\sigma^2} t\right)$$
is an exponential martingale

26.2 Applying to Stock Pricing

we want the discounted stock price process to be a martingale
$$dS_t = \mu S_t dt + \sigma S_t dB_t$$
let Z_t be the discounted stock price process
$$Z_t = e^{-rt} S_t$$
using the product rule
$$dZ_t = e^{-rt} dS_t + de^{-rt} S_t$$
sub in the stock price process
$$dZ_t = e^{-rt}\left[\mu S_t dt + \sigma S_t dB_t\right] - re^{-rt} S_t dt$$
$$= e^{-rt}\left[(\mu - r) S_t dt + \sigma S_t dB_t\right]$$
sub in Z_t $again$

$$dZ_t = (\mu - r) Z_t dt + \sigma Z_t dB_t$$

to make this a martingale it must have 0 drift

$$\sigma Z_t d\widetilde{B}_t = (\mu - r) Z_t dt + \sigma Z_t dB_t$$

$$\sigma Z_t d\widetilde{B}_t = \sigma Z_t \left[\frac{(\mu - r)}{\sigma} dt + dB_t \right]$$

$$d\widetilde{B}_t = \frac{(\mu - r)}{\sigma} dt + dB_t$$

therefore we shift the brownian by $\frac{(\mu - r)}{\sigma}$ to get a martingale

$$\frac{(\mu - r)}{\sigma} = \gamma$$

$$\sigma \widetilde{B}_t - (\mu - r) t = \sigma B_t$$

therefore we can now substitute $B_t \ for \ \widetilde{B}_t$ under any circumstances

sub into undiscounted stock price process

$$dS_t = \mu S_t dt + \sigma S_t dB_t$$

$$dS_t = \mu S_t dt + \sigma S_t \left(d\widetilde{B}_t - \frac{(\mu - r)}{\sigma} dt \right)$$

$$= r S_t dt + \sigma S_t d\widetilde{B}_t$$

sub into our function for S_t obtained via It$o^{'}$s Lemma

$$= S_0 \exp \left[\left(\mu - \frac{1}{2} \sigma^2 \right) t + \sigma \widetilde{B}_t - (\mu - r) t \right]$$

$$S_0 \exp \left[\left(r - \frac{1}{2} \sigma^2 \right) t + \sigma \widetilde{B}_t \right]$$

discounted asset process is a Q martingale

$$S_t = \frac{S_0 \exp \left[\left(r - \frac{1}{2} \sigma^2 \right) t + \sigma \widetilde{B}_t \right]}{\exp (rt)} = S_0 \exp \left[-\frac{1}{2} \sigma^2 t + \sigma \widetilde{B}_t \right]$$

using It$o^{'}$s Lemma

$$S_t = S_0 \sigma \exp \left[-\frac{1}{2} \sigma^2 t + \sigma \widetilde{B}_t \right] \text{ which has no drift}$$

intuitively we can get the result r$S_t dt + \sigma S_t d\widetilde{B}_t$ through martingale property

$$E [Z_t] = E [e^{-rt} S_t] = e^{-rt} E [S_t]$$

$$e^{rt} e^{-rt} E [S_t] = E [S_t]$$

$$e^{rt} E [Z_t] = E [S_t]$$

$$because\ Z_t \text{ is a martingale, } \mathrm{E}\left[Z_t\right] = Z_s$$

$$Z_s = e^{-rs} S_s$$

$$e^{rt} e^{-rs} S_s = E\left[S_t\right] = e^{r(t-s)} S_s$$

this shows the return on the stock is simply r

$$therefore in the undiscounted stock process, we could intuitively replace \mu with r$$

to satisfy no arbitrage, in Q world cannot earn more than r

27 Call Option and Black Scholes Formula

$$\text{Call option is } e^{-r(T-t)} E_Q\left(\left[S_T - K\right]^+ | F_t\right)$$

$$S_T = S_t \exp\left[\left(r - \frac{1}{2}\sigma^2\right)(T-t) + \sigma\left(\widetilde{B}_T - \widetilde{B}_t\right)\right]$$

$$express T - t as \tau$$

$$express - \frac{\widetilde{B}_T - \widetilde{B}_t}{\sqrt{\tau}} \text{ as Y}$$

$$S_T = S_t \exp\left[\left(r - \frac{1}{2}\sigma^2\right)\tau - \sigma\sqrt{\tau}Y\right]$$

Y is a random variable, independent of F_t because brownian increment it is gaussianly distributed

$$S_t\ is\ F_t\ measurable$$

$$\left[S_T - K\right]^+ \text{ expires at the money when } S_T = K$$

$$K = S_t \exp\left[\left(r - \frac{1}{2}\sigma^2\right)\tau - \sigma\sqrt{\tau}Y\right]$$

rewrite in terms of Y

$$Y = \frac{\ln\left(\frac{S_t}{K}\right) + \left(r - \frac{1}{2}\sigma^2\right)\tau}{\sigma\sqrt{\tau}} = d_2$$

$$when Y \le d_2,\ S_T \ge K$$

$$so our call option is only positive when Y \le d_2$$

therefore the integral, which is taking the expectation

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-r\tau}\left(S_t \exp\left[\left(r - \frac{1}{2}\sigma^2\right)\tau - \sigma\sqrt{\tau}y\right] - K\right)^+ e^{-\frac{1}{2}y^2} dy$$

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{d_2} e^{-r\tau}\left(S_t \exp\left[\left(r - \frac{1}{2}\sigma^2\right)\tau - \sigma\sqrt{\tau}y\right] - K\right) e^{-\frac{1}{2}y^2} dy$$

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{d_2} S_t \exp\left[-\frac{1}{2}\sigma^2\tau - \sigma\sqrt{\tau}y - \frac{1}{2}y^2\right] - e^{-r\tau}Ke^{-\frac{1}{2}y^2}dy$$

Riemann integrals are linear

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{d_2} S_t \exp\left[-\frac{1}{2}\sigma^2\tau - \sigma\sqrt{\tau}y - \frac{1}{2}y^2\right]dy - \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{d_2} e^{-r\tau}Ke^{-\frac{1}{2}y^2}dy$$

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{d_2} e^{-r\tau}Ke^{-\frac{1}{2}y^2}dy = e^{-r\tau}KN(d_2)$$

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{d_2} S_t \exp\left[-\frac{1}{2}\sigma^2\tau - \sigma\sqrt{\tau}y - \frac{1}{2}y^2\right]dy$$

$$\frac{S_t}{\sqrt{2\pi}}\int_{-\infty}^{d_2} \exp\left[-\frac{1}{2}\left(\sigma^2\tau + 2\sigma\sqrt{\tau}y + y^2\right)\right] = -\frac{1}{2}\left(y + \sigma\sqrt{\tau}\right)^2$$

$$e^{-\frac{1}{2}\left(y+\sigma\sqrt{\tau}\right)^2} = changeofmeasure, shiftingthedistributionupby\sigma\sqrt{\tau}$$

$$\frac{S_t}{\sqrt{2\pi}}\int_{-\infty}^{d_2+\sigma\sqrt{\tau}} \exp\left[-\frac{1}{2}z^2\right]\text{dz where z is a new gaussian variable with the shifted boundaries}$$

$$= S_tN(d_1) \text{ where we define } d_1 = d_2 + \sigma\sqrt{\tau}$$

this gives us the Black Scholes formula!

$$C_t = S_tN(d_1) - e^{-r\tau}KN(d_2)$$

Look at p221 of Shreve, Stochastic Calculus II

28 Change of Numeraire

If instead we look at it through expectation:

$$C = e^{-rt}E_Q\left[(S_T - K)^+\right] = e^{-rt}E_Q\left[(S_T - K)\,1_{S_T>K}\right]$$

$$= e^{-rt}E_Q\left[S_T 1_{S_T>K}\right] - e^{-rt}E_Q\left[K1_{S_T>K}\right]$$

$$secondterm = Ke^{-rt}E_Q\left[S_T > K\right]$$

$$firstterm = e^{-rt}E_Q\left[S_T 1_{S_T>K}\right]$$

$$S_T = S_0 \exp\left[\left(r - \frac{\sigma^2}{2}\right)t + \sigma dW_t^Q\right]$$

multiplied by e^{-rt} taken inside the expectation

$$S_0 \exp\left[-\frac{\sigma^2}{2}t + \sigma dW_t^Q\right]$$

$\exp\left[-\frac{\sigma^2}{2}t+\sigma dW_t^Q\right]$ is similar to Girsanov $\exp\left[\mu y+\frac{1}{2}\mu^2\right] where \mu is the boundary shift$

in our case we have y as d$W_t^Q and \mu as \sigma because of Q expectation we have p\left(Q\right)dQ$

$\left[-\mu y-\frac{1}{2}\mu^2\right]$ also acceptable

$$recall\ \int_a^b \frac{dQ\left(x\right)}{dP\left(x\right)}dP\left(x\right)h\left(x\right)=E_Q\left[h\left(x\right)\right]$$

define a new Q measure that will rearrange into desired Girsanov change of measure

$$-\frac{\sigma^2}{2}t+\sigma dW_t^Q=\sigma dW_t^{Q^*}+\frac{1}{2}\sigma^2 t$$

$$dW_t^Q-\sigma t=dW_t^{Q^*}$$

$$dW_t^{Q^*}+\sigma t=dW_t^Q$$

$$\sigma dW_t^{Q^*}+\sigma^2 t=\sigma dW_t^Q$$

substitute in we have

$$\exp\left[\sigma dW_t^{Q^*}+\frac{\sigma^2}{2}t\right]=\exp\left[\mu y+\frac{1}{2}\mu^2\right]=\frac{dQ^*}{dQ}$$

$$\frac{dQ^*}{dQ}=\frac{S_t}{S_0 B_t}$$

also get stock dynamics under Q^* $dS_t=\left(r+\sigma^2\right)dt+\sigma dW_t^{Q^*}$

$$\frac{dQ^*}{dQ}dQ\left(x\right)h\left(x\right)=E_{Q^*}\left[h\left(x\right)\right]$$

$$first term=e^{-rt}E_Q\left[S_T 1_{S_T>K}\right]$$

$$=S_0 E_{Q^*}\left[S_T e^{-rt}>Ke^{-rt}\right]$$

we know that $S_T e^{-rt}$ can now be expressed with d$W_t^{Q^*}$ by subbing in for dW_t^Q

$$S_0 E_{Q^*}\left[S_0\exp\left[\sigma dW_t^{Q^*}+\frac{\sigma^2}{2}t\right]>Ke^{-rt}\right]$$

$$\frac{S_0\exp\left[\sigma dW_t^{Q^*}+\frac{\sigma^2}{2}t\right]\exp\left[rt\right]}{K}>1$$

$$=\frac{S_0}{K}\exp\left[\left(r+\frac{\sigma^2}{2}\right)t+\sigma dW_t^{Q^*}\right]>1$$

$$\ln\left(\frac{S_0}{K}\right)+\left(r+\frac{\sigma^2}{2}\right)t>-\sigma dW_t^{Q^*}$$

$$\frac{\ln\left(\frac{S_0}{K}\right)+\left(r+\frac{\sigma^2}{2}\right)t}{\sigma\sqrt{t}}>-\frac{\sigma dW_t^{Q^*}}{\sigma\sqrt{t}}$$

$$= N(d_1)$$

giving us the equation:

$$S_0 N(d_1) - Ke^{-rt} N(d_2)$$

Why is this change of measure a 'change of numeraire'? Note that under the risk neutral measure, the asset (strike price) is discounted by the risk free rate. Under the new measure, we have the stock price multiplied by the N(d1) term, meaning that assets are discounted by the stock price. A numeraire measures assets in terms of other assets. An alternative proof:

$$\text{expectation of exponential } e^{yZ-\frac{1}{2}y^2} 1_{S_T>K} = N(y+a)$$

$$1_{y>a}$$

shift in the mean of standard normal from 0 to y

29 Black Scholes PDE

We want to model the dynamics of a call option. We know that a call option is a function, dependent on the underlying stock and time:

$$C = C(S,t)$$

We can differentiate to find the dynamics using Ito's lemma:

$$dC = \frac{dC}{dt}dt + \frac{dC}{dS}dS + \frac{1}{2}\frac{d^2C}{dS^2}dS^2$$

recall that dt^2 and dS dt terms cancel due to quadratic variation

$$substituteindS = \mu S\,dt + \sigma S\,dW$$

$$\left(\frac{dC}{dt} + \frac{dC}{dS}\mu S + \frac{1}{2}\frac{d^2C}{dS^2}\sigma^2 S^2\right)dt + \frac{dC}{dS}\sigma S\,dW$$

assume that we have a portfolio that combined with the call, hedges the stochastic element

$$\Pi = \Delta S + aB$$

$$\Delta = numberofshares$$

$$a = numberofriskfreebonds$$

$$dB = Brdt$$

$$d\Pi = (\Delta\mu S + aBr)\,dt + \Delta\sigma S\,dW$$

the dynamics include the stochastic element due to holding stock we want the stochastic elements to cancel out

$$\Delta\sigma S\,dW + \frac{dC}{dS}\sigma S\,dW = 0$$

$$\Delta\sigma S + \frac{dC}{dS}\sigma S = 0$$

$$\Delta\sigma S = -\frac{dC}{dS}\sigma S$$

$$\Delta = -\frac{dC}{dS}$$

now combine the call and the hedging portfolio to remove stochastic element

$$dC + d\Pi = \left(\frac{dC}{dt} + \frac{dC}{dS}\mu S + \frac{1}{2}\frac{d^2C}{dS^2}\sigma^2 S^2\right)dt + \frac{dC}{dS}\sigma S\,dW + -\frac{dC}{dS}\mu S\,dt + aBrdt + -\frac{dC}{dS}\sigma S\,dW$$

$$dC + d\Pi = \left(\frac{dC}{dt} + \frac{1}{2}\frac{d^2C}{dS^2}\sigma^2 S^2 + aBr\right)dt$$

without any stochastic element, this is deterministic

by no arbitrage the rate of change must be the riskless rate

$$d\left(C + \Pi\right) = \left(C + \Pi\right)rdt$$

$$\left(C - \frac{dC}{dS}S + aB\right)rdt = \left(\frac{dC}{dt} + \frac{1}{2}\frac{d^2C}{dS^2}\sigma^2 S^2 + aBr\right)dt$$

$$\left(C - \frac{dC}{dS}S + aB\right)r = \left(\frac{dC}{dt} + \frac{1}{2}\frac{d^2C}{dS^2}\sigma^2 S^2 + aBr\right)$$

$$Cr - \frac{dC}{dS}Sr + aBr = \frac{dC}{dt} + \frac{1}{2}\frac{d^2C}{dS^2}\sigma^2 S^2 + aBr$$

$$0 = \frac{dC}{dt} + \frac{1}{2}\frac{d^2C}{dS^2}\sigma^2 S^2 + \frac{dC}{dS}Sr - Cr$$

$$Cr - \frac{dC}{dS}Sr = \frac{dC}{dt} + \frac{1}{2}\frac{d^2C}{dS^2}\sigma^2 S^2$$

riskless portfolio of C and $\frac{dC}{dS}$ number of S

on the RHS we have a theta term (decay of call value with time) and a convexity $(gamma)$ $term$

this is the BS PDE, we solve it by implementing bounds $(payoffs)$ to obtain pricing formula

30 Feynman-Kac Representation

What does Feynman – Kac tell us? That given a parabolic PDE and a stochastic process, it is possible to represent the solution to a stochastic problem through conditional expectation (which is deterministic):

$$u(x,t)$$

$$\psi(x)$$

$$V(x,t)$$

$$f(x,t)$$

$$\frac{du}{dt}(x,t) + \mu(x,t)\frac{du}{dx}(x,t) + \frac{1}{2}\sigma^2(x,t)\frac{d^2u}{dx^2}(x,t) - V(x,t)\,u(x,t) + f(x,t) = 0$$

$$u(x,T) = \psi(x)$$

$$u(x,t) = E^Q\left[e^{-V(X,T)}f(X,T) + e^{-V(X,T)}\psi(X)\,|X = x\right]$$

$$dX = \mu(X,t)\,dt + \sigma(X,t)\,dW^Q$$

we see that this is analogous to the BS PDE under RNP measure Q

$$\max(S_T - K, 0) = \psi(x)$$

$$r = V(x,t)$$

$$0 = f(x,t)$$

$$C = u(x,t)$$

subbing in the BS PDE, we can therefore express it as discounted expected payoff

given the conditional payoff, we can then just solve with basic Girsanov/numeraire/etc

31 Heat Equation

It seems the intuition is that the BS equation has similarities to the heat equation – diffusion of the distribution of prices (heat) over time. It is another way for us to obtain an analytical (PDE) solution to BS. Useful because it allows for pricing of barrier options, but for vanillas it provides no new intuition.

32 Gamma Scalping

What does the PDE prove? That we can create a riskless portfolio by purchasing a call and by shorting delta stocks. The profit/loss on this strategy can be broken down into a theta and a gamma component. We constructed the strategy using: a call, a short stock, borrowing on a MMA. Looking at:

$$\left(C - \frac{dC}{dS}S + aB\right) rdt = \left(\frac{dC}{dt} + \frac{1}{2}\frac{d^2C}{dS^2}\sigma^2 S^2 + aBr\right) dt$$

$$\left(C - \frac{dC}{dS}S + aB\right) \text{rdt is a riskless hedging portfolio, or d}V_t$$

$$\text{it must be equal to 0 to satisfy no arbitrage}$$

$$therefore aBrdt = \left(\frac{dC}{dS}S - C\right) rdt$$

$$Cr - \frac{dC}{dS}Sr = \frac{dC}{dt} + \frac{1}{2}\frac{d^2C}{dS^2}\sigma^2 S^2$$

$$\text{if LHS is our riskless portfolio, it equals 0}$$

$$0 = \frac{dC}{dt} + \frac{1}{2}\frac{d^2C}{dS^2}\sigma^2 S^2$$

$$\text{express theta, or} \frac{dC}{dt} \text{ in terms of gamma}$$

$$-\frac{1}{2}\frac{d^2C}{dS^2}\sigma^2 S^2 = \frac{dC}{dt} = gamma$$

$$\text{also works if we express in terms of the entire portfolio, or V}$$

$$-\frac{1}{2}\frac{d^2V}{dS^2}\sigma^2 S^2 = \frac{dV}{dt}$$

Gains made on the ‘gamma scalping’ (risk has been completely hedged out or delta hedged, so this is constantly rebalancing) are balanced by the theta decay. This aspect, or the profit or loss is known exactly, it is deterministic. We know the volatility (constant) and the difference in squared stock term today by stochastic calculus:

$$dS^2 = S^2\sigma^2 dt$$

We could also assume that, the expected returns from gamma scalping are:

$$E\left[\frac{d\Delta}{dS}dS \times \frac{S}{2} + \frac{S + dS}{2}\right] = E\left[\frac{d\Delta}{dS}\frac{(dS)^2}{2}\right]$$

$$S drops out because when r = 0, E[S] = 0$$

$$\frac{d\Delta}{dS} = gamma$$

$$(dS)^2 = S^2\sigma^2 dt$$
$$E[R] = \Gamma\frac{S^2\sigma^2 dt}{2}$$
$$dollargamma = \frac{1}{2}\Gamma S^2$$
$$= \Gamma_{\$}\sigma^2$$

So the amount we make is the dollar gamma multiplied by squared price change per unit of time:

$$dS^2 = S^2\sigma^2 dt = \frac{\frac{dS^2}{S^2}}{dt} = \sigma^2$$

Or, dollar gamma per multiplied by realised variance. Whilst $dS^2 = S^2\sigma^2 dt = deterministicinthelimit$ (infinites

$$fornon-infinitesimalinterval \text{ (discrete time)}$$

$(\Delta S)^2$ is still random, so gains in discrete time are:

$$\frac{1}{2}\frac{d^2V}{dS^2}\left[(\Delta S)^2 - \sigma^2 S^2\Delta t\right]$$

because realised variance will be random, implied will be known

$$butthetableedwilldependon\Delta t$$

also equivalent to

$$\frac{1}{2}\frac{d^2V}{dS^2}S^2\left[\left(\frac{\Delta S}{S}\right)^2 - \sigma^2\Delta t\right]$$

realised variance take away implied variance

33 Locally Quadratic

Assuming the value of our 'bowl' is:

$$V(S + 2\Delta S)$$

to approximate by It$o^{'}$s Lemma

$$dV = V + \frac{dV}{dS}2\Delta S + \frac{1}{2}\frac{d^2V}{dS^2}(2\Delta S)^2$$
$$V = 0bydefinition, costlessportfolio$$
$$\frac{dV}{dS} = 0bydefinition, becauseitisadeltahedgedportfolio$$

$\frac{1}{2}\frac{d^2V}{dS^2}(2\Delta S)^2$ can be rearranged

$$taking\ S^2 \text{ out of the } (2\Delta S)^2 \text{ term gives}$$

$$\frac{1}{2}\frac{d^2V}{dS^2}S^2\left(2\frac{\Delta S}{S}\right)^2$$

$$\frac{1}{2}\frac{d^2V}{dS^2}S^2 = \Gamma_{\$}$$

$$= \Gamma_{\$}(2\Delta x)^2 where \Delta x = \frac{\Delta S}{S}$$

$$2^2\Gamma_{\$}(\Delta x)^2 = 2^2V(S + \Delta S)$$

$$therefore a move in 2\Delta S gives 4 times the return, not 2 times$$

Variance (square term) rather than standard deviation is important

34 Homogeneity

Option prices are homogeneous functions, to the first degree (power). Therefore:

$$C(S,K) = S - K$$

$$\lambda C(S,K) = C(\lambda S, \lambda K) = \lambda S - \lambda K$$

$$\text{in terms of partial derivatives}$$

$$\lambda\frac{dC}{dt} + \lambda\frac{dC}{dS}$$

35 ODE and Self Financing

$$S_t = stockprice$$

$$y = dividendyield$$

$$\theta = unitsheld$$

$$t = time$$

$$moneyspentonchangeisalwayschangeinunitsatthetime$$

$$\text{makes sense, if I had 100 stocks at 1p each}$$

$$\text{they change to 90 stocks at 1p each}$$

$$\text{amount spent must be } (100 - 90) * 1p$$

$$S_t d\theta_t = \theta_t y S_t dt$$

$$d\theta_t = \frac{\theta y S_t}{S_t} dt$$

change in number of shares held, is the dividend yield from shares held rolled into stocks

$$d\theta = \theta y dt$$

$$\frac{d\theta_t}{dt} = \theta_t y$$

$$we want our variable, \theta, to be separated from y$$

$$\frac{d\theta_t}{\theta_t} = y dt$$

$$\text{integral of both sides, and because} \frac{d\theta}{\theta} = \frac{1}{\theta} d\theta, the integral is the natural log$$

y is a constant, and the integral is the antiderivative

$$\text{result is } \ln\theta_T = yt + C$$

$$e^{\ln \theta_T} = e^{yt+C}$$

$$\theta_T = e^{yt+C}$$

$$\theta_T = e^{yt} e^C$$

$$at t = 0$$

$$\theta_t = e^{y0} e^C$$

$$\theta_t = e^C$$

$$\theta_T = \theta_t e^{yt}$$

36 Forwards

Components of the discount / return are:

$$r = risk free rate$$

$$c = cost of carry\,(storage)$$

$$y = convenience yield$$

$$F_t = S_t e^{(r+c-y)(T-t)}$$

For delivery means you PAY / GIVE whatever needs to be delivered. For example, a forward for delivery of the stock means you lose the stock at time t, and receive the forward price. Vice versa. Why backwardation and contango? For downward slope, when I enter the forward contract I am aware that by deferring payment, I am losing out on a benefit that outweighs the associated costs I have passed onto the counterparty. Therefore price will go down. For upwards slope, when I enter the forward contract I am aware that counterparty is taking on costs on my behalf. They have to store the underlying, and because they aren't receiving my payment until later, they are missing out on investing the money in the risk free rate. Therefore, I pay them a premium and price goes up.

$$F_t = S_t e^{(r-y)(T-t)}$$

$$F_t e^{-r(T-t)} = S_t e^{-y(T-t)}$$

$$V_t = S_t e^{-y(T-t)} - F_t e^{-r(T-t)}$$

Remember to keep a distinction between the VALUE of a forward and the forward PRICE. If I enter a forward at time 0, as the underlying changes the forward accrues value:

$$F_0 = S_0 e^{(r-y)(T-0)}$$

at time t

$$V_t = e^{-r(T-t)} E_Q [S_T - F_0]$$

$$e^{-r(T-t)} E_Q [S_T] = e^{-r(T-t)} S_t^{(r-y)(T-t)} = e^{-y(T-t)} S_t$$

$$V_t = e^{-y(T-t)} S_t - e^{-r(T-t)} F_0$$

This is the same as the mark to market value of the forward.

37 Options

Function of the price at time t is a smooth curve. Function of the price at time T or maturity is kinked. The 'kinked' line indicates the intrinsic value of the option, the difference between stock at maturity and the strike. The smooth curve indicates the time value of the option. This can be decomposed into: Option to defer payment – rather than paying the strike now, I defer it to time T. This increases the discounting of the strike payment, which increases the price. Potentiality or cherry-picking – I have the option to exercise later at maturity. Doing so, with positive volatility on the underlying increases the change for upside, without change in downside. Or, we can view options as forwards with additional value due to downside protection – lower bounded by zero, where forwards can have negative payoffs.

$$V_t = S_t e^{-y(T-t)} - F_t e^{-r(T-t)}$$

simply substitute F_t for K

$$C_t = V_t = S_t e^{-y(T-t)} - K e^{-r(T-t)} + insurance$$

The insurance can be seen as the value of the option to default on the forward, which you cannot usually do. For a put option, simply reverse the signs. No arbitrage bounds:

$$C_t \le S_t$$

$$C_t \ge 0$$

$$P_t \le K$$

$$P_{EU} \le PV(K)$$

$$C_{AM} \ge C_{EU}$$

$$C_{AM}(T) \ge C_{AM}(t) \text{ longer maturity worth more}$$

$$\text{does not hold for European due to dividends}$$

$$C_t \ge \max(0, S_t - PV(K))$$

$$P_t \ge \max(0, PV(K) - S_t)$$

38 Option Pricing – Binomial Tree

We know we can use binomial probability distribution (which converges to a Gaussian) to price an option tree with n periods:

$$\left[\sum_{j=0}^{n} \binom{n}{j} Q^j (1-Q)^{n-j} \left[0, u^j d^{n-j} S_T - K\right]^+\right] (1+r)^{-n}$$

$$\text{there is a minimum value of j such that option always expires ITM}$$

$$j > a,\ S_T - K > 0$$

$$j < a,\ S_T - K < 0$$

$$Q' = Q\frac{u}{r}$$

$$where r = 1 + r$$

$$\text{this transforms}$$

$$Q^j (1-Q)^{n-j} \frac{u^j d^{n-j}}{r^n} = Q'^j \left(1 - Q'\right)^{n-j}$$

$$\text{factor out } Q^j \frac{u^j}{r^j}$$

$$\left[Q\frac{u}{r}\right]^j \left[(1-Q)\frac{d}{r}\right]^{n-j} \left(seethatyouhaver^j\ and\ r^{n-j} \text{ which is still } r^n overall\right)$$

$$\text{rewrite Q and } Q' \text{in terms of} \frac{r-d}{u-d}$$

$$Q'^{j}\left(1-Q'\right)^{n-j}$$

also notice that the coefficient $\frac{u^j d^{n-j}}{r^n}$ only applies to S_T

therefore rewrite binomial formula as:

$$\left[\sum_{j=0}^{n}\binom{n}{j}Q'^{j}\left(1-Q'\right)^{n-j}\right]S_T-\left[\sum_{j=0}^{n}\binom{n}{j}Q^{j}\left(1-Q\right)^{n-j}\right]K\left(1+r\right)^{-n}$$

$$a>\frac{\ln\left(\mathrm{K}/\mathrm{S}d^{n}\right)}{\ln\left(\mathrm{u}/\mathrm{d}\right)}$$

you can get this from $\max\left[0,u^j d^{n-j}S_T-K\right]>0$

take logs and rearrange for j

reverse for puts

How can we infer u, d and p from the mean and variance of the stock?

$$S_T=S_t u^j d^{n-j}$$

$$\frac{S_T}{S_t}=u^j d^{n-j}$$

$$\ln\left(\frac{S_T}{S_t}\right)=jln\left(u\right)+\left(n-j\right)\ln\left(d\right)=jln\left(\frac{u}{d}\right)+nln\left(d\right)$$

$$E\left[\ln\left(\frac{S_T}{S_t}\right)\right]=E\left[j\right]\ln\left(\frac{u}{d}\right)+nln\left(d\right)$$

$$Var\left[\ln\left(\frac{S_T}{S_t}\right)\right]=\left[\ln\left(\frac{u}{d}\right)\right]^2 Var\left[j\right]$$

$$Var\left[j\right]=np\left(1-p\right)$$

$$\frac{E\left[j\right]}{n}=p$$

$$E\left[j\right]=np$$

$$npln\left(\frac{u}{d}\right)+nln\left(d\right)=n\mu$$

$$np\left(1-p\right)\left[\ln\left(\frac{u}{d}\right)\right]^2=n\sigma^2$$

we want the mean and variance of the continuously compounded return on stock to be equal

$$\mu\left(T-t\right)=n\mu$$

$$\sigma^2\left(T-t\right)=n\sigma^2$$

$$u = e^{\sigma\sqrt{\frac{T-t}{n}}} = e^{\sigma\sqrt{h}}$$
$$d = \frac{1}{u} = e^{-\sigma\sqrt{h}}$$
$$p = \frac{1}{2} + \frac{1}{2}\left(\frac{\mu}{\sigma}\right)\sqrt{\frac{T-t}{n}}$$
$$q = \frac{r-d}{u-d}$$
$$e^{rh} = r$$
$$h = \frac{T-t}{n}$$
$$q = \frac{e^{rh}-d}{u-d}$$

This is the **Cox-Ross Rubenstein** calibration which allows the binomial tree to approximate the continuous time BS model where n tends to infinity. The reason why ud = 1 is that so the distribution is centred – if you go up then down, you always end up at the same place regardless of how many steps in you are. To account for the fact that actually the movement will drift upwards or downwards depending on the expected return (loss), the physical probability is adjusted by mu and sigma, and is greater or less than ½. **Jarrow-Rudd** says:

$$p = \frac{1}{2}$$

$$u = e^{\mu h + \sigma\sqrt{\frac{T-t}{n}}}$$
$$d = e^{\mu h - \sigma\sqrt{\frac{T-t}{n}}}$$

Jarrow's approach accounts for drift by incorporating it directly into upward and downward movements, rather than adjusting the probability. **Boyle** says:

$$u = e^{\sigma\sqrt{2h}}$$

$$d = \frac{1}{u}$$

Binomial models have the no-arbitrage bound:

$$u > e^{(r-d)h} > d$$

$$u > e^{rh} > d$$

If the risk free return is higher than u, no need to invest in anything risky, riskless return will always be higher. If risk free return lower than d, always invest in risky stock for higher return. CRR violates this boundary, as $e^{\sigma\sqrt{h}}$ for high values of h or low value of sigma causes problems. rh can become bigger than u. Also, if erh becomes big and d becomes small, there are times where you get probabilities greater than 1 or less than 0, which should not be possible. JR solves this and makes things easier by making probabilities not dependent on sigma and h. In the limit, the binomial tree converges to a normal distribution. This means that continuously compounded returns are normal, which by extension implies that returns follow a lognormal distribution.

$$e^{rh} > e^{\sigma\sqrt{h}} \text{ if h is large, sigma low}$$

$$e^{-\sigma\sqrt{h}} > e^{rh} \text{ if h is large}$$

RNP also goes to negatives

$$\frac{(1+r)-d}{u-d} = e^{rh} - e^{-\sigma\sqrt{h}}$$

$$E[C] = pC_u + (1-p)C_d$$

$$Var[C] = E[C - E[C]]^2$$

$$= p(C_u - E[C])^2 + (1-p)(C_d - E[C])^2$$

$$= p(1-p)(C_u - C_d)^2$$

$$E[S] = puS + (1-p)dS$$

$$Var[S] = E[S - E[S]]^2$$

$$= p(uS - E[S])^2 + (1-p)(dS - E[S])^2$$

$$= p(1-p)(u-d)^2$$

Elasticity is simply defined as the percentage change in a function of something with respect to the percentage change of that something. So in the context of options, the percentage change in the call price (which is a function of the underlying) over the percentage change of the underlying.

same applies to stocks

$$var_s\Omega > var_s = var_c$$

in other words, variance is greater for an option compared to stock because it is levered

$$where\Omega = \frac{S\Delta}{C} = \frac{dC}{\frac{C}{\frac{dS}{S}}} \text{ (work from it backwards)}$$

$$= \frac{S(C_u - C_d)}{CS(u-d)}$$

$$\Delta = \frac{C_u - C_d}{S(u-d)}$$

$$B = \frac{uC_d - dC_u}{u-d}$$

because $uC_d - dC_u \leq 0, C = \Delta S + BwhereB \leq 0, thereforeshorting = leverage$

$$ifC = \Delta S - B$$

$$and leverage = \frac{S\Delta}{C}$$

$$leverage > 1 if S\Delta > C, or S\Delta > \Delta S - B which is absolutely true$$

$$\Omega or elasticity shows change in percentage of call for percentage in stock$$

$$\Omega \geq 1 means leverage, because movement greater than 1 : 1$$

$$delta stock + short bond in hedging portfolio for call = leverage$$

consider a call option as a balance sheet

$$equity = C \text{ (our investment)}$$

$$debt or leverage = -B$$

$$assets = \Delta S$$

$$leverage ratio = \frac{D}{E} = -\frac{B}{C}$$

$$gross leverage = \frac{Assets}{E} = \frac{\Delta S}{C} = \Omega$$

Always remember that equity is the value of our investment. So for call option, call price. For forward, value of the forward (0). Debt is what we must eventually pay back, our liability – strike, forward price. Assets are the underlying – stock. Why is OTM more levered? Because the elasticity (percentage change in call price for underlying price) is much greater when OTM, than when ITM. This makes sense – if you buy cheap, likelihood of ITM lower and initial investment lower, gains you make will be bigger on a percentage basis than if you bought in at more expensive, ITM price. Leverage increases when far OTM and also when time to maturity decreases. In the binomial, when too far out the money you reach a point of no return where no possible number of steps remaining before maturity to take you ITM, so you are worth nothing (and require no leverage). This can't happen in continuous time where there is always a possibility of ending ITM, even if infinitesimal. Shorter maturities have more leverage (OTM or just ATM) because option price will not move much if you are already ITM, but will move for big swings around the money or going from OTM to ITM. **Recall:**

$$\Delta S + \frac{1}{1+r}B = C$$

$$uS\Delta + B = C_u$$

$$dS\Delta + B = C_d$$

$$C_u - C_d = uS\Delta - dS\Delta$$

$$\frac{C_u - C_d}{uS - dS} = \Delta$$

$$\Delta = \frac{C_u - B}{uS}$$

$$\frac{dS\,(C_u - B)}{uS} + B = C_d$$

$$\frac{d\left(C_u - B\right)}{u} + B = C_d$$

$$dC_u - dB + uB = C_d u$$

$$(u - d)\, B = C_d u - C_u d$$

$$B = \frac{uC_d - dC_u}{u - d}$$

Deriving RNP

$$\Delta S + \frac{1}{1+r} B = \frac{S\,(C_u - C_d)}{S\,(u - d)} + \frac{uC_d - dC_u}{(1+r)(u-d)} = \frac{1}{1+r}\frac{Cu\,(1+r-d)}{u-d} + \frac{Cd\,(u-1-r)}{u-d}$$

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

39 Put Call Parity

With continuous dividends, recall that the price of the stock is constantly being eroded by the value of the dividend. You can therefore calculate the present value of the stock with the dividend being removed:

$$C = P + S\ - PV\,(K) - PV\,(D)$$

$$P = C - S\ + PV\,(K) + PV\,(D)$$

$$C\ = P + S_0 e^{-yt} - Ke^{-rt}$$

$$P\ = C - S_0 e^{-yt} + Ke^{-rt}$$

$$S_0 e^{-(r-d)t}$$

$\int_0^t De^{-dt}$ Why? Because the dividend at time 0 is D. However, the stock price is constantly eroded by the dividend yield, and therefore the dividend is also constantly reduced as time goes on. Why is the dividend not discounted at r? In RNP, the expected rate of return has to be r (including dividends). Therefore if the stock price decreases by -d, the dividend payouts must equal d, so that r-d+d = r for the total return.

40 Put Call Symmetry

For a put option and a call option with the same underlying, strike and maturity, when you set the risk free rate to 0 and the strike and stock price to each other, a put and call are exactly the same in price.

$$d_1\ becomes: \frac{\ln\left(\frac{S_t}{K}\right) + \frac{1}{2}\sigma^2 t}{\sigma\sqrt{t}}$$

$$d_2 \; becomes: \; \frac{\ln\left(\frac{S_t}{K}\right) - \frac{1}{2}\sigma^2 t}{\sigma\sqrt{t}}$$

$$d_1 = -d_2$$

$$d_2 = -d_1$$

$$S = K$$

$$r = 0$$

$$Put = Ke^{-0t}(-d_2) - S(-d_1)$$

$$= K(d_1) - S(d_2)$$

$$= S(d_1) - K(d_2) = Call$$

41 Greeks

41.1 Delta

Recall that it is the sensitivity of the option price to the underlying. If you hold the underlying, you need to sell 1/delta calls to hedge. If you hold the option, sell delta stocks to hedge. Why is it that we need to sell 1/delta? If:

$$C = \Delta S_t + aB_t$$

$$dC = \Delta dS_t + adB_t$$

$$-\frac{a}{\Delta}dB_t = dS_t - \frac{1}{\Delta}dC$$

Δ = the delta, or number of shares held in the equivalent hedging portfolio for the call Therefore, by rearranging to make the underlying and call equivalent to the bond or riskless element, we see how we hedge 1 unit of the underlying to remove the stochastics Delta Call = Delta Put + 1 Call delta ranges between 0 and 1. 0 delta, means 0 sensitivity to change in underlying. This makes sense as if a call is far OTM, small changes in the stock won't do much. Whereas when a call is very far ITM, it becomes increasingly stock-like. Put delta ranges between -1 and 0. -1 delta, increase in the stock decreases option price. Makes sense, because you want the price to be lower. Goes to 0 if very ITM, because small changes do nothing. Quick rules of thumb:

$$dN(d_2) = dN\left(d_1 - \sigma\sqrt{t}\right) = dN\left(d_1 - \sigma\sqrt{t}\right)\, dd_1 - d\sigma\sqrt{t}$$

$$= N'(d_1)\; e^{-\frac{\sigma^2}{2} + d_1\sigma\sqrt{t} + \ln\left(\frac{S}{K}\right)}$$

$$\frac{dx}{dx^2} = \frac{1}{\frac{dx^2}{dx}} = \frac{1}{2x}$$

$$N(\infty) = 1$$

$$N(-\infty) = 0$$

$$inotherwords, thec.d.fofinfinityis1, -infinityis0$$

When volatility goes to infinity, the value of the option tends to the stock price (call) or strike (put)
With infinite volatility, probability S > K tends to 0 (because of d2). So the RN probability to expire ITM tends to 0. However, the expected value of the terminal stock price ST (which is also the 1st term in BS formula) must still equal S when discounted due to no arbitrage. So although the majority of the probability density is stuck at 0, there are few but exceedingly high values of ST which average out to the current stock price S. This occurs mechanically as a result of the BS formula, rather than economic intuition.

41.2 Gamma

Change in delta for change in underlying price. Gamma put = Gamma call, always positive. Underlying has no gamma, because think of what the delta for a stock is – either 1 or -1, long or short, there is no intermediate change. Gamma hedging is usually carried out through options trades. Negative gamma occurs when you are writing options, short options. When time to maturity increases, Gamma is flatter. When time to maturity decreases, gamma is highest ATM. Gamma is strictly positive for (long) put and call, because a positive move in the stock means a 'positive' move in the delta. Even for puts, stock going up means delta becomes less negative.

41.3 Dollar Gamma

The volatility trade is derived from the fact that, based on the BS PDE, we can create a riskless portfolio that involves a call option and a short position on the underlying. This also makes sense in the context of delta hedging. A call price function is convex, whereas the static delta is linear. Therefore, profit from a call will always be above that of the underlying, and going long the call and short the underlying will always result in profit (difference of curved function above linear function). Delta will always be neutral with the correct rehdege, thereby removing all directional risk. However, largest gamma (pointiness of P& L bowl) also means largest theta decay by BS PDE.

$$\frac{1}{2}\Gamma S^2$$

Second order derivative hence why there is a half and squared term… ? For convexity? We have established that dollar gamma represents the gains we make from continuously hedging or gamma scalping a portfolio comprised of option and underlying. However, in practice unrealistic to do so – what if instead we only rebalanced periodically at intervals? Where the hold the position for an interval Δt our portfolio returns are:

$$\frac{dV}{dt}\Delta t + \frac{1}{2}\frac{d^2V}{dS^2}(\Delta S)^2$$

$$\text{we know that}\frac{dV}{dt} = -\frac{1}{2}\frac{d^2V}{dS^2}\sigma^2 S^2\text{dt, whereas}\frac{d^2V}{dS^2} = \frac{1}{2}\frac{d^2V}{dS^2}(dS)^2$$

$$soworeplacedtwith\Delta tanddS\,with\Delta S$$

$$\frac{1}{2}\frac{d^2V}{dS^2}(\Delta S)^2 - \frac{1}{2}\frac{d^2V}{dS^2}\Delta t\sigma^2 S^2$$

$$= \frac{1}{2}\frac{d^2V}{dS^2}\left[(\Delta S)^2 - \Delta t\sigma^2 S^2\right]$$

$$= \frac{1}{2}\frac{d^2V}{dS^2}S^2\left[\left(\frac{\Delta S}{S}\right)^2 - \Delta t\sigma^2\right]$$

In other words, gain is dollar gamma multiplied by difference in realised variance $\left(\frac{\Delta S}{S}\right)^2$ and implied variance embedded in options $\Delta t\sigma^2$ This allows us to trade based on our views of future volatility. However, even if realised variance on average is positive and greater than implied, the covariance (multiplication by) dollar gamma can cause negative returns. This is as gamma is not constant – if it is high when realised variance is low, you have huge negative payoff. If it is low when realised variance is high, you have small positive payoff. Gamma can be constant, so the 2nd derivative needs to be $\frac{1}{S^2}$

41.4 Theta

Why is it asymmetrical for OTM and ITM? Because ITM, you have intrinsic value. Even as time goes on and time value fades, still left with intrinsic value. OTM you only have extrinsic (time) value. If it decays as time goes on, your theta goes completely to 0, impossible to end up ITM so does not matter if time goes on. Same applies for both puts and calls. For puts you can have positive theta when extremely ITM because when stock price is close to 0, and TTM decreases, increasingly likely to exercise at close to intrinsic value. Also, because there is less possibility that stock will shoot back up / vary enough to end up OTM again. Discount factor means that as time decreases, K discounted less increasing payoff. As TTM decreases, so does volatility. If put is far in the money, decreasing volatility means less chance for the put to go ITM, and cannot go much lower because bounded by 0. Theta could be positive if option has negative time value. It has to converge to intrinsic value at maturity. So in the case of a call option, all the terms of theta are negative, leading theta to be negative in all circumstances. For a put, if the option is far enough in the money, theta will actually be positive. Be careful that theta is derivative with respect to t, where time to maturity is: $T - t$, so t increasing means we are getting closer to expiry. K is being discounted less!!!! Not discounted more.

The derivative of a cumulative distribution function is:

$$\frac{d}{dt}N(d1) = \frac{dd1}{dt}N^{'}(d1)$$

$$N^{'} = standardnormaldensity = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}d1^{2}\right)$$

$$soyougofromN \sim \left(\mu, \sigma^{2}\right) toN \sim (0,1)\, withx = d1$$

so to take derivatives of the BS pricing formula, you need to use the chain rule

$$\frac{dv}{du}u + \frac{du}{dv}v$$

for theta:

$$Se^{-dt}N(d1) - Ke^{-rt}N(d2) = C$$

$$\frac{dC}{dt} = \frac{d}{dt}\left(Se^{-dt}N(d1)\right) - \frac{d}{dt}\left(Ke^{-rt}N(d2)\right)$$

$$\frac{d}{dt}\left(Se^{-dt}N(d1)\right) \text{ according to chain rule}$$

$$N^{'}(d1)\frac{dd1}{dt}Se^{-dt} = \frac{dv}{du}u$$

$$dSe^{-dt}N(d1) = \frac{du}{dv}v$$

$$becauset = T - t$$

$$\frac{d}{dt}\left(Ke^{-rt}N(d2)\right) \text{ according to chain rule}$$

$$N^{'}(d2)\frac{dd2}{dt}Ke^{-rt} = \frac{dv}{du}u$$

$$= N^{'}\left(d1 - \sigma\sqrt{t}\right)\frac{d\left(d1 - \sigma\sqrt{t}\right)}{dt}Ke^{-rt}$$

$$= N^{'}\left(d1 - \sigma\sqrt{t}\right)\left(\frac{dd1}{dt} + \frac{\sigma}{2\sqrt{t}}\right)Ke^{-rt}$$

$$N^{'}\left(d1 - \sigma\sqrt{t}\right) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(d1 - \sigma\sqrt{t}\right)^{2}\right)$$

$$= \exp\left(-\frac{1}{2}d1^{2}\right)\exp(r-d)\,t\frac{S}{K}$$

multiply by Ke^{-rt}

$$N^{'}(d1)\,Se^{-dt}$$

$$Se^{-dt}N^{'}(d1)\frac{dd1}{dt} + Se^{-dt}N^{'}(d1)\frac{\sigma}{2\sqrt{t}}$$

$$Kre^{-rt}N\left(d1 - \sigma\sqrt{t}\right) = \frac{du}{dv}v$$

putting everything together

$$S e^{-dt} N'(d1) \frac{dd1}{dt} + dS e^{-dt} N(d1) - S e^{-dt} N'(d1) \frac{dd1}{dt} - S e^{-dt} N'(d1) \frac{\sigma}{2\sqrt{t}} - Kre^{-rt} N\left(d1 - \sigma\sqrt{t}\right)$$

$$= dS e^{-dt} N(d1) - S e^{-dt} N'(d1) \frac{\sigma}{2\sqrt{t}} - Kre^{-rt} N\left(d1 - \sigma\sqrt{t}\right) = Theta$$

41.5 Vega

First order sensitivity of option price to volatility. Is strictly positive – greater volatility always leads to greater probability of ending up in the money, and vice versa. When time to maturity is greater, vega is also greater – there is more time for the increased volatility to push the price into the money, and therefore the option has greater value. With low time to maturity, large movements in vol required to end up in the money, therefore sensitivity lowered. Higher volatility also leads to higher chance of ending OTM though? But that gives the 'cherry picking' optionality value more oomph.

41.6 Rho

Sensitivity to interest rates. Simple as positive for call (discounting cash outflow, the strike) and negative for put (discounting cash inflow, the strike).

41.7 Volga

Second order sensitivity of option price to volatility, or first order sensitivity of vega. Volatility of volatility, which becomes relevant once we take into account stochastic volatility.

41.8 Vanna

Sensitivity of vega to the underlying, or sensitivity of the delta to the volatility. Recall that volatility is multiplied by time to maturity in the derivatives, hence why vega and vanna are greater longer away from expiration. If vanna is negative, it means that as volatility increases, delta decreases. As volatility decreases, delta increases. In the case of stock options and feedback loops: if volatility pushes a stock price ATM into the money, where gamma is highest, market makers will have to buy more of the underlying to hedge. As the share price increases, we head further ITM, decreasing gamma and the implied volatility (due to skew). Vanna is 0 ATM? and negative ITM, therefore the decreasing volatility increases the delta and hedging even more. Think about ITM puts and OTM calls. If a call is OTM, increasing volatility increases likelihood of ending ITM. If a put is ITM, increasing vol increases OTM likelihood.

Therefore, delta will be increasing if vol increases, as we move from 0 towards 1 and -1 towards 0, so positive movements only. So Vanna is positive. For OTM puts and ITM calls. As vol increases, ITM and OTM likelihood increases. Put delta will get more negative and call delta will also go down, moving from 0 to -1 and 1 to 0 respectively. So Vanna is negative. Incorporating the skew, where lower strikes have higher IV, and higher strikes have lower IV, going further ITM for a call decreases vol, increasing delta.

41.9 Charm

Measures the change of delta with time. We see that delta becomes more extreme as we approach TTM, with it being flatter when far from expiry. Therefore for options where the spot is right of K, charm is positive. For options where the spot is left of K, charm is negative.

42 Digital Options

A digital option has a binary payoff, which is either:

$$1_{S_T<K}$$

$$1_{S_T>K}$$

The Heaviside step function is equivalent to the first derivative of the option. We can see this as

$$1_{S_T>K} = N(d_2)$$

Or the probability of ending up in the money under risk neutral probability. We also can see that this is clearly the derivative of the call with respect to the strike:

$$C = SN(d_1) - Ke^{-rt}N(d_2)$$

$$\frac{dC}{dK} = e^{-rt}N(d_2)$$

Which happens to be equivalent to the digital price. We can also see that we can replicate the option using an option spread (long and short option at different strikes):

$$\lim_{\epsilon} \to 0 \frac{C(K) - C(K+\epsilon)}{\Delta K} = \frac{dy}{dx} = \frac{dC}{dK}$$

$$\text{number of options in total therefore is} \frac{1}{\epsilon}$$

You cannot actually set the interval in strikes to 0 as you cannot make any money, but ideally you would want to set the interval as small as possible. We must note that the digital is in fact dependent on the volatility skew, and not only on strike however. Looking at Leibniz, and defining the call price as a function of K:

$$C = e^{-rt} E_Q \left[\max \left(S_T - K, 0\right)\right]$$

$$C(K) = e^{-rt} \int_K^\infty (S_T - K)\, q(S)\, dS$$

$$\frac{d}{dx}\int_b^a f(x,y)\,dy = \int_b^a \frac{d}{dx} f(x,y)\,dy + f(x,x)$$

$$\frac{dC}{dK} = \frac{d}{dK}(S_T - K)\, q(S)\, dS\, + f(K,K) = -q(S)\, dS\, + f(K,K) = -\int_K^\infty q(S)\, dS\, - (K-K)$$

$$= -e^{-rt}\int_K^\infty q(S)\, dS\, = -\frac{dC}{dK} = -e^{-rt} N(d_2)$$

If we now definite the call price as a function of strike and volatility, which is NOT DEPENDENT FROM STRIKE as per IV skew (whereas the underlying is) and redo the above:

$$-\frac{dC}{dK} - \frac{dC}{d\sigma}\frac{d\sigma}{dK}$$

comes from chain rule also:

$$z = f(x,y)\, y = g(x)$$

$$\frac{dz}{dx} = \frac{df}{dx}\frac{dx}{dx} + \frac{df}{dy}\frac{dy}{dx}$$

$$C(K,\sigma) \text{ and } \mathrm{K}(\sigma)$$

$$-\frac{dC}{dK}\frac{dK}{dK} + \frac{dC}{d\sigma}\frac{d\sigma}{dK} = -\frac{dC}{dK} - \frac{dC}{d\sigma}\frac{d\sigma}{dK}$$

$$\frac{dC}{d\sigma} \text{ is vega}$$

$\frac{d\sigma}{dK}$ is skew, can quantify by how much vol changes per change in strike, it is negative

Therefore $\frac{dC}{d\sigma}\frac{d\sigma}{dK}$ represents how much we misprice digital options by simply pricing them as $\frac{dC}{dK}$

43 Barrier

First create a sharkfin payoff option. Then create portfolio of options that takes into account the removal of the option payoff upon touching the barrier. Using the put call symmetry, translate these options into actual tradeable securities. The result is the static hedge portfolio. We can see that when r is 0, an option with payoffs:

$$\frac{1}{S} C(S,L) = C\left(\frac{L^2}{S}, L\right)$$

$$P\left(\frac{L^2}{S}, K\right) = \frac{K}{S} C\left(S, \frac{L^2}{K}\right)$$
$$whereris0, andK = S$$

You can prove this by calculating the d1 and d2 for these, with the assumption r is 0 and K = S.

44 Implied Volatilities

We know that in the Black Scholes model, returns are normally distributed:

$$\frac{dS_t}{S_t} = rdt + \sigma dW_t$$

And we know that prices are lognormally distributed:

$$\ln S_t = lnS_0 + exp\left(r - \frac{\sigma^2}{2}\right)t + \sigma dW_t$$

Lognormal distributions have a very specific shape. They have a negative skew (cannot have log of negative values, so everything is skewed towards the right). They also have lower kurtosis (thinner tails, less outliers). Remember: negative skew = big positive values and vice versa High kurtosis = fatter tails In reality, we see that there is an implied volatility surface – when volatility should be flat and constant across all different strike prices for options and stock prices in the model, they are in fact skewed. Generally, the real distribution of volatilities is higher in the left tail and thinner in the right tail – high volatilities, mean higher prices for options. This suggests that options which protect against the downside, ie movement towards the left tail, are more expensive as 'bad' shocks occur more often than you would expect in a lognormal world. This also suggests that the dQ/dP for the left tail is greater, as you value payoffs in the downside more which makes intuitive sense. There is generally a 'variance risk premium' – implied volatility is greater than realised volatility. Why could this be? Because we are willing to pay more 'vol' to protect against potential losses. However, by construction if implied vol is higher for OTM put, it is more expensive. By put call parity, an OTM put is equivalent to a call being ITM. But the higher volatility at that strike means the ITM call is also more expensive – why should that be the case? If we only pay more for downside protection, why should an ITM call be more expensive? The reason is that when the distribution is changed to accommodate fatter tails, we must fit the distribution so that the total probability is still equal to 1. We can always solve for the implied vol of the market price of options. We can do this because vega is strictly positive, that is to say that option prices increase monotonically as a function of volatility. Given a root lower bounded by 0 (as variance cannot be negative), for example you can solve for the implied volatility using numerical methods such as Newton-Raphson method:

$$\text{assume that } \sigma_0 = 0$$

$$\sigma_1 = \sigma_1 - \frac{f(\sigma_0)}{f'(\sigma_0)} \text{ for } \mathrm{f}(\sigma_i) = C_{BS}(\sigma_i) - C^*$$

$$\textit{where } C^* \text{ is market price of call, and } C_{BS} \text{ is BS price}$$

We measure the skew against 'moneyness'. If we only looked at K and how volatilities changed with the strike, we wouldn't get an accurate image because we'd also need to know where the stock price is relative to K. So it's better to assess implied vol against K/S – regardless of how S changes, we will always know what the vol is for a fixed level of moneyness. Another way is also looking at the forward price rather than S – because it 'centres' the surface around an expectation, the forward being an expectation of future moneyness. Note: Implied volatility surface is flatter when there is a long time to maturity. As we approach maturity, skew becomes greater. Reasons for the skew: as stock prices, and equity decreases, the leverage of the firm increases mechanically. This increases the probability of default and further increases probability of further downward movement, or an increase in volatility (as measured by risk). Reason for the 'smile': for currencies, you will want to pay protection against your currency depreciating against another. For your currency to depreciate, another will have to appreciate. Therefore, paying for downside protection of your currency is the same as paying upside protection of the other currency, and vice versa. Hence why there is a smile. For gold, there is a positive skew – safe haven assets appreciate in bad times, hence we pay more for protection against it rising. We also cannot fit the IV surface too loosely – we only observe a limited number of strikes and options, so creating a surface involves filling in gaps and smoothing. If this is done incorrectly, there may be arbitrage.

45 Implied Volatility Dynamics and Regimes

There cannot be only parallel shifts in the implied volatility surface (entire term structure moves up and down by the same amount given a change in underlying price or moneyness). A parallel shift means that when the term structure / surface moves, it moves up and down perfectly without any changes in the skew/overall shape, perfect parallel translation of the curve. This would give rise to arbitrage, through barbell type convexity strategies: purchasing a mixture of long-dated and short-dated bonds always has greater convexity than buying straight debt, meaning you will always make more money than a portfolio of the same duration. So the barbell, you go long the spread of maturities and short the 'bullet' (fixed intermediate maturity) and make arbitrage profits (any changes in interest rates, you profit because the convex barbell outperforms the less convex bullet). How does this work? Longer dated tends to have higher duration and convexity (more sensitivity to changes in interest rates) than short dated and intermediate. But, given that duration is linear, the intermediate dated duration is the same as combination of long dated and short dated. So you can create a portfolio with exact same duration but higher convexity, which represents an arbitrage if parallel shifts occur.

Applying the chain rule or what we learnt with $\frac{dC}{dK}$ simply apply the same but wrt to the underlying: Delta:

$$\frac{dC}{dS} = \frac{dC}{dS}_{BS} + \frac{dC}{d\sigma}_{BS} \times \frac{d\sigma_{IV}}{dS}$$

$$where\ \frac{d\sigma_{IV}}{dS} \text{ is unobservable}$$

The true delta in a world with implied volatility skews, is the Black Scholes delta, with an adjustment due to the fact that volatility shifts. We can observe $\frac{d\sigma_{IV}}{dK}$, or the skew/slope of implied volatilities and use this to deduce $\frac{d\sigma_{IV}}{dS}$ If we look only at the implied volatility of a constant level of moneyness, where K will shift as S does to keep the level of moneyness constant:

$$z = f(x, y)\, y = g(x)$$

$$\frac{dz}{dx} = \frac{df}{dx}\frac{dx}{dx} + \frac{df}{dy}\frac{dy}{dx}$$

$$\sigma_{CM}(m, S) = \sigma_{CM}(K, S) \text{ because moneyness is just} \frac{K}{S}$$

$$\sigma_{IV}\left(\frac{K}{S}S, S\right) = \sigma_{IV}(K, S)$$

$$withaskew, K(\sigma_{IV})$$

$$\frac{d\sigma_{CM}}{dS} = \frac{d}{dS}\sigma_{IV}(K, S) = \frac{d\sigma_{IV}}{dS}\frac{dS}{dS} + \frac{d\sigma_{IV}}{dK}\frac{dK}{dS}$$

$$= \frac{d\sigma_{CM}}{dS} = \frac{d\sigma_{IV}}{dK}m + \frac{d\sigma_{IV}}{dS}$$

$$R = \frac{\frac{d\sigma_{CM}}{dS}}{\frac{d\sigma_{IV}}{dK}m}$$

$$\frac{dC}{d\sigma_{BS}} \times \frac{d\sigma_{IV}}{dS} = \frac{dC}{d\sigma_{BS}} \times \left[\frac{d\sigma_{CM}}{dS} - \frac{d\sigma_{IV}}{dK}m\right]$$

$$= \left[\frac{\frac{d\sigma_{CM}}{dS}\frac{d\sigma_{IV}}{dK}m}{\frac{d\sigma_{IV}}{dK}m} - \frac{\frac{d\sigma_{IV}}{dK}m\frac{d\sigma_{IV}}{dK}m}{\frac{d\sigma_{IV}}{dK}m}\right] \text{ (making the denominator the same)}$$

$$= \frac{d\sigma_{IV}}{dK}m[R - 1] \left(\text{factoring out } \frac{d\sigma_{IV}}{dK}m\right)$$

$$assumingm = 1$$

To clarify: Delta with Vol Skew:

$$\frac{dC}{dS} = \frac{dC}{dS}_{BS} + \frac{dC}{d\sigma_{BS}} \times \frac{d\sigma_{IV}}{dK}m[R - 1]$$

$$\text{by incorporating R we were able to move from } \frac{d\sigma_{IV}}{dS}\ (unobservable)\ to \frac{d\sigma_{IV}}{dK}\ (observable)$$

Change in Constant Moneyness Implied Vol:

$$\frac{d\sigma_{CM}}{dS} = \frac{d\sigma_{IV}}{dK}m + \frac{d\sigma_{IV}}{dS}$$

$$R = \frac{\frac{d\sigma_{CM}}{dS}}{\frac{d\sigma_{IV}}{dK}m}$$

$$\frac{d\sigma_{CM}}{dS} = R\frac{d\sigma_{IV}}{dK}m$$

We then define R as either 0, 1 or 2. If R is 0, we get that we must subtract the adjustment with vega and IV from BS delta. In other words, as S changes the implied vol stays constant. Thus the moneyness vol curve stays constant regardless of changes in the underlying, so long as you are at a given level of moneyness, your implied vol will be the same. So ATM implied vol remains the same always, but the vol surface shifts downwards as the stock price falls (and upwards as it rises?). Delta will be lower? than BS delta, because $\frac{d\sigma_{IV}}{dS}$ is negative. The curve remaining constant wrt moneyness means you are always equally unhappy to lose money. If the implied vol for ATM is 0.2, and implied vol for 0.8 moneyness is 0.3, and these are constant for whatever the stock price is, we are always equally unhappy to lose 20% , regardless of the absolute value of the stock. The skew with respect to K/S is the same. With respect to K, becomes steeper.

$$IV = a - b\frac{K}{S}$$

If R is 1 the implied vol for constant moneyness is the same as the implied vol for a given strike. So the implied vol surface remains the same, and delta is also the same as that in Black Scholes. Or, the skew with respect to K is constant.

$$IV = a - bmS$$

If R > 1 the implied vol for constant moneyness is greater. Sticky local vol means $\frac{d\sigma_{IV}}{dK} = \frac{d\sigma_{IV}}{dS}$ so depending on K+S? The implied vol is the same for a given level of K+S/2. If S tumbles, the implied vol must increase for the same K. IV for K=100, S=100 is the same as K=120, S=80. Because for $IV = a - b\left(\frac{K+S}{2}\right)$ as S decreases IV increases Stickiness:

45.1 Sticky Delta / Moneyness:

For a given level of moneyness or delta, the IV does not change with the underlying. Plotted against moneyness: Same for different S. This means we pay the same downside protection regardless of how the underlying moves, always unhappy to lose the same percentage of money. Plotted against strike: Upward shift for higher S and flatter. Downward shift for lower S and steeper. Delta greater than BS delta. The subtraction of the adjustment term is positive, because the skew is negative (change is absolute?)

$$IV = a - bK\frac{1}{S}$$

See that decreasing S, the coefficient (slope) becomes more negative. Increasing S, less negative. This is the ‘moderate’ market. Violent swings average out over time to a sticky delta regime.

45.2 Sticky Strike:

For a given strike, IV does not change with underlying. As S decreases, the slope flattens, as it increases, it steepens. Plotted against moneyness: Downward shift for higher S, as slope becomes more negative, upward shift for lower S, as slope becomes less negative. Specifically, ATM vol falls for stock increase. Plotted against strike: Same for different S. Delta is BS delta

$$IV = a - b\frac{K}{S}S$$

Derman calls this complacency or greed – in good times as stocks or the index keeps rising, markets still keep volatility at strikes the same. ATM vol therefore decreases – these liquid options become cheaper, and the market acts as if good times will continue. The vol for ATM should increase/stay in line with sticky delta at least, but because everyone is optimistic about index levels continuing to rise, they neglect to account for a future fall. Options are too cheap.

45.3 Sticky Local Vol:

For a given (K+S/2) IV does not change with underlying. Plotted against moneyness: shifts downwards twice as fast as S increases. And vice versa? Plotted against strike: Entire surface shifts upwards for lower S and lower K. Entire surface shifts downwards for higher S and higher K. Delta is lower than BS delta

$$IV = a - \frac{b(K+S)}{2}$$

As S or K increases, slope becomes more negative. As K or S decreases, becomes less negative. Assuming that local vol is constant. As we lose money, we pay more for insurance, we want to avoid further falls even more as opposed to sticky delta where we are equally unhappy to lose at all levels. Market must compensate for ATM vol being left too low, by rapidly increasing vol at twice the rate the index falls.

46 Stochastic Volatility

Corrections needed for the BS model: BS model has excessively thin tails in the left, and excessively fat tails in the right.
BS model has constant volatility.

$$dS_t = rS_t dt + \sigma_t S_t dW_t$$

$$d\sigma_t = a(\sigma_t, t)\, d_t + \gamma(\sigma_t, t)\, dZ_t$$

$$whereaand\gamma aredriftantdiffusiontermswhichdependon\sigma_t \text{ and t}$$

We want to set the Brownian motions to have negative correlation. When there is positive movement in the stock, and price increases, volatility decreases. This means that it is less likely to end up in the right tail when the stock performs well. When there is negative movement in the stock, and price decreases, volatility increases. This means that it is more likely to end up in the left tail when the stock performs badly. If correlation is 0, using law of iterated expectations:

$$E_Q\left[e^{-rt}\left(S_T - K, 0\right)^+\right]$$

$$= E_Q\left[E_Q\left[e^{-rt}\left(S_T - K, 0\right)^+ \mid \sigma_t : t \in [0, T]\right]\right.$$

$$\text{expectation given we know } \sigma_t\text{, which if we know, means we are in BS model}$$

$$= E_Q\left[C_{BS}\right]$$

So taking the average of the BS prices is all we need. This does not work for equities (where we need a strong negative correlation), because taking the condition that we know sigma, means that the stock price process Brownians become 'known' by construction of correlation. Then they will no longer be random, and we cannot apply this stochastic model by taking expectations.

47 Heston Model

The Heston model is a mean-reverting model for stochastic volatility:

$$d\nu_t = -\lambda\left(\nu_t - \overline{\nu}\right)dt + \eta\sqrt{\nu_t}dZ_t$$

$$\overline{\nu} = longrunaverage$$

This is a mean reverting process. When the actual variance exceeds the long run average, the negative constant lambda brings it back down. When the long run average exceeds the actual, the lambda brings it back up. Variance must be positive, which is accounted for through the square root term. If actual variance is small then the square root tends to 0. Remember because you can have negative Brownians, as volatility approaches 0 (and negative numbers), square root approaches 0 and removes the Brownian. If lambda and long run average a larger than eta, this will always remain positive.

48 Local Volatility

Stochastic volatility can be difficult to implement due to multiple Brownian motions. However, only one of these Brownians (the stock) can actually be traded, meaning we cannot exactly replicate stochastic volatility with a money market account and a risky stock. We know that simply allowing volatility to depend on time does not give an implied volatility skew. There can be a term structure, but at each point in time volatility will still be constant across strikes. But we can allow it to depend on time and the stock price, without introducing another Brownian. The diffusion term and the underlying will be directly correlated with one another. We then say that the market believes the volatility at certain points will be at a specified level when 'hovering' around a strike K. This is what makes the volatility 'local', it is the market expectations of volatility 'local' to a particular time point/price. With the Black Scholes, we assumed we had derivatives with respect to St and t, which represent infinitesimal changes. In reality, these are unobservable – but what we do observe is maturities and strikes. At time of calibration, we have knowledge of drift and future probability densities and strikes from the existing range of options. We can use this to obtain the infinitely many diffusion terms. We are essentially fitting volatilities to the current spectrum of option prices. So what local volatility gives a model for us how the volatility will change or look like as the underlying price evolves. Looking at the local vol surface for example, we are looking at what the volatility is expected to be given a particular stock price path such that it ends up at a certain moneyness at a certain time T. This is more dynamic than the static implied volatility surface, which simply gives volatilities based on the stock price today for a set time to maturity. There is no implication of what happens as the stock price moves and hovers around certain strikes at certain times. Just a cross sectional view based on the information we have today. Local vol must be recalibrated at each point in time however, as it only perfectly fits option price one period into the future. We have all the marginal distributions for the local volatility for each strike such that:

$$q\left(T_i|T_0\right)$$

However, we do not have the conditional distributions for subintervals for example:

$$q\left(T_2|T_1\right)$$

You can therefore have all individual marginal distributions but not have the total joint distribution without knowing the process (copula) that binds them all together. Therefore, local vol is not good at pricing derivatives dependent on joint events – it can't accurately reflect the market's view on dependent distributions, conditional distributions. We can interpret the local volatility to be the market expectation of the actual variance at time T local to strike K. We can also assume, in the BS case that implied variance is the average local variance over time:

$$\frac{1}{T}\int_0^T \sigma_L^2\left(t\right)dt = \sigma_{IV}^2$$

However, as in reality local vol depends not only on time but also the underlying this cannot hold. Instead, approximately the following is true:

$$\frac{1}{T}\int_0^T \sigma_L^2\left(\widetilde{S}\left(t\right), t; S_0\right)dt = \sigma_{IV}^2$$

So the implied variance is the average of the local variance around a particular underlying path. In other words, along the path that spends most time around gamma, which produces high value from scalping. **The implied volatility surface flattens over time, due to 2 main factors: firstly, there is a no arbitrage condition. As we look at the volatility of prices which depend on:**

$$e^{-rt}E_Q\left[S_T - K\right] \ or \ E_Q\left[e^{-rt}S_T - e^{-rt}K\right]$$

For large ttm, K tends to 0. For ST>K, or $e^{-rt}S_T > Ke^{-rt}$, $C = E_Q\left[M_T 1_{M_T>0}\right]$ **Therefore for large ttm, the price cannot depend on K, and the volatility cannot depend on K, meaning no skew is present. The next condition is that there is a degree of mean reversion over time. We can see empirically there is a long-run average level of volatility in markets. Hence depending on the speed of reversion and the presence of a long-run mean of volatility, the curve will lose its skew. Local volatility INHERITS the implied volatility skew, it fits according to the observed prices and implied skew we have today. Therefore it must also flatten with time.**

49 Dupire

An equation that gives us the local volatilities, or the implied volatilities the market expects to have conditional on the stock price being around a particular strike for a given future time period. We know that the implied volatilities depend on time and also that the surface varies with strike, so we want to obtain the implied local volatility given we have observable derivatives of the option prices available. We can then price any options, even those that are not quoted. It is like the inverse of the Fokker Planck, because we are using future probability densities to get the diffusion coefficient. So how does Dupire work? Should be similar to Fokker-Planck whereby we take the SDE, apply Ito's Lemma and take expectations to remove the stochastic integral. Expectation by integrals also introduces the probability densities into the equation. We then integrate by parts – though in the case of option prices, the prices ARE the integral, the market has already done the integration with respect to the probability density for us. **The first derivative of an option payoff is the HEAVISIDE (step) FUNCTION. The second derivative of an option payoff, or the derivative of the Heaviside function is the DIRAC DELTA FUNCTION.**

$$C = (S - K)^+$$

$$dC = \frac{dC}{dS}dS + \frac{1}{2}\frac{d^2C}{dS^2}dS^2$$

$$\frac{dC}{dS} = \frac{d}{dS}(S - K)^+ = 1_{S_T>K}\frac{d}{dS}(S - K) = 1_{S_T>K}$$

$$\frac{d^2C}{dS^2} = \frac{d}{dS}1_{S_T>K} = d(S - K) \text{ where d is Dirac delta}$$

$$1_{S_T>K}dS + \frac{1}{2}d(S - K)dS^2$$

$$= 1_{S_T>K}(rS_t dt + \sigma S_t dB_t) + \frac{1}{2}d(S - K)\sigma^2 S^2 dt$$

$$= \left(1_{S_T>K} r S_t + \frac{1}{2} d(S-K)\sigma^2 S^2\right) dt + 1_{S_T>K}\sigma S_t dB_t$$

$$taking expectation as in Fokker - Planck$$

$$E[dC] = E\left[\left(1_{S_T>K} r S_t + \frac{1}{2} d(S-K)\sigma^2 S^2\right)\right] dt$$

$$\frac{d}{dt} E[(S-K)] = E\left[1_{S_T>K} r S_t + \frac{1}{2} d(S-K)\sigma^2 S^2\right]$$

$$= E[1_{S_T>K} r S_t] + E\left[\frac{1}{2} d(S-K)\sigma^2 S^2\right]$$

49.1 First Term

$$rE[1_{S_T>K} S_t] \text{ when expressed in terms of call price}$$

$$1_{S_T>K} S_t - 1_{S_T>K} K = C$$

$$C + 1_{S_T>K} K = 1_{S_T>K} S_t$$

$$E_Q[C + 1_{S_T>K} K] = E_Q[C] + K E_Q[1_{S_T>K}]$$

$$1_{S_T>K} = \frac{d}{dS}(S-K)^+$$

$$-1_{S_T>K} = \frac{d}{dK}(S-K)^+$$

$$Ce^{rt} - Ke^{rt}\frac{dC}{dK} = E_Q[1_{S_T>K} S_t]$$

$$rCe^{rt} - rKe^{rt}\frac{dC}{dK} = rE_Q[1_{S_T>K} S_t]$$

49.2 Second Term

$$rE_Q[1_{S_T>K} S_t] = rCe^{rt} - rKe^{rt}\frac{dC}{dK}$$

$$E\left[\frac{1}{2} d(S-K)\sigma^2 S^2\right] is positive only for S = K$$

$$\frac{1}{2} K^2 E_Q\left[\sigma^2 d(S-K)\right]$$

$$\text{using Baye}'s: P(A|B) = \frac{P(\text{A and B})}{P(B)}$$

$$E_Q\left[\sigma^2|S_T = K\right] = \frac{E_Q\left[\sigma^2 d(S-K)\right]}{E_Q\left[d(S-K)\right]}$$

$$E_Q\left[\sigma^2|S_T = K\right] \text{ is the local vol}$$

$$d(S-K) = \frac{d^2C}{dK^2} = e^{-rt}E_Q\left[\frac{d^2}{dK^2}(S-K)^+\right] = e^{-rt}E_Q\left[d(S-K)\right]$$

$$\frac{1}{2}K^2E_Q\left[\sigma^2|S_T = K\right]e^{rt}\frac{d^2C}{dK^2}$$

49.3 Expectation

$$E_Q\left[(S-K)^+\right] = e^{rt}\text{C or } C^u\text{, undiscounted call}$$

$$\text{rewrite everything in terms of the undiscounted}$$

$$\frac{dC^u}{dt} = C^u - rK\frac{dC^u}{dK} + \frac{1}{2}K^2E_Q\left[\sigma^2|S_T = K\right]\frac{d^2C^u}{dK^2}$$

$$\text{now we simply rearrange for the local vol}$$

$$E_Q\left[\sigma^2|S_T = K\right] = \frac{\left(\frac{dC^u}{dt} - C_u + rK\frac{dC^u}{dK}\right)}{\frac{1}{2}K^2\frac{d^2C^u}{dK^2}} = Dupire$$

50 Fokker-Planck Equation

Purpose of this equation is to go from a drift and diffusion term/process to a probability density. We can compute the future probability distribution of a process? Integration by parts:

$$\int u\frac{dv}{dx}dx = [uv]_a^b - \int v\frac{du}{dx}dx$$

$$where\ [uv]_a^b \text{ is evaluating uv and b and a}$$

Assume we have a standard Brownian:

$$X_t = B_t$$

$$dX_t = dB_t$$

$$\text{using It}o^{'}sLemmaonf(X_t)$$

$$dX_t = f_x dB_t + \frac{1}{2}f_{xx}dB_t^2 = f_x dB_t + \frac{1}{2}f_{xx}dt$$

taking the expectation

$$E\left[f_x dB_t + \frac{1}{2} f_{xx} dt\right] = E\left[\frac{1}{2} f_{xx} dt\right] \text{ because Brownian has expected value 0}$$

expectation is taking the integral with respect to probability distribution of x

$$dE[f] = \frac{1}{2} E[f_{xx}]\, dt$$

$$\frac{d}{dt} E[f] = \frac{1}{2} E[f_{xx}]$$

$$\frac{d}{dt} \int_{-\infty}^{\infty} f(x)\, p(x,t)\, dx = \frac{1}{2} \int_{-\infty}^{\infty} f_{xx} p(x,t)\, dx$$

where $\mathrm{p}(x,t)$ is the probability density of x

now we integrate RHS by parts, choosing u as $\mathrm{p}(x,t)$ and v as f_{xx}

$$[f_x p(x,t)]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} f_x \frac{dp(x,t)}{dx}\, dx$$

the first term $\mathrm{p}(x,t)$ goes to 0

integrate by parts again

$$-\left[f(x) \frac{dp(x,t)}{dx} - f(x) \frac{d^2 p(x,t)}{dx}\right]$$

first term is 0 again

$$\frac{d}{dt} \int_{-\infty}^{\infty} f(x)\, p(x,t)\, dx = \frac{1}{2} \int_{-\infty}^{\infty} f(x) \frac{d^2 p(x,t)}{dx}$$

$$\int_{-\infty}^{\infty} f(x) \frac{dp(x,t)}{dt} dx = \frac{1}{2} \int_{-\infty}^{\infty} f(x) \frac{d^2 p(x,t)}{dx}$$

$$\int_{-\infty}^{\infty} f(x) \left(\frac{dp(x,t)}{dt} - \frac{d^2 p(x,t)}{dx}\right) dx = 0$$

if the integral is 0, density must equal 0

$$\left(\frac{dp(x,t)}{dt} - \frac{d^2 p(x,t)}{dx}\right) = 0 = Fokker - Planck$$

we can do the same with an SDE

$$dX_t = \mu dt + \sigma dB_t$$

$f(X_t)$ is an arbitrary function

$$df(x) = f_x dX_t + \frac{1}{2} f_{xx} dX_t^2$$

$$= f_x (\mu dt + \sigma dB_t) + \frac{1}{2} f_{xx} \sigma^2 dt$$

$$df(x) = \left(f_x\mu + \frac{1}{2}f_{xx}\sigma^2\right)dt + f_x\sigma dB_t$$

$$E[df(x)] = E\left[\left(f_x\mu + \frac{1}{2}f_{xx}\sigma^2\right)dt\right]$$

$$\frac{d}{dt}E[f(x)] = E\left[\left(f_x\mu + \frac{1}{2}f_{xx}\sigma^2\right)\right]$$

$$\int_{-\infty}^{\infty} f_x\mu(x,t)\,p(x,t)\,dx + \frac{1}{2}\int_{-\infty}^{\infty} f_{xx}\sigma^2(x,t)\,p(x,t)\,dx$$

$$u = \sigma^2 p(x,t), dv = f_{xx}$$

$$-\int_{-\infty}^{\infty} f(x)\frac{d\mu(x,t)\,p(x,t)}{dx}dx + \frac{1}{2}\int_{-\infty}^{\infty} f(x)\frac{d^2\sigma^2(x,t)\,p(x,t)}{dx}dx$$

$$= \left(\frac{d}{dt}p(x,t) + \frac{d}{dx}\mu(x,t)\,p(x,t) - \frac{1}{2}\frac{d^2}{dx}\sigma^2(x,t)\,p(x,t)\right) = 0$$

can simply substitute in for the drift and diffusion functions and terms

51 Breeden-Litzenberger

We can derive the implied Q probability density from the function of observed option prices. Using FTAP (present value of discounted cash flows under RNP):

$$C = e^{-rt}E_Q[\max(S_T - K, 0)]$$

we know that the payoff function of option can be expressed in terms of density

$$C = e^{-rt}\int_K^{\infty}(S_T - K)\,q(S)\,dS$$

We want to back out the probability density, q(S). Using **Leibniz**:

$$\frac{d}{dx}\int_b^a f(x,y)\,dy = \int_b^a \frac{d}{dx}f(x,y)\,dy + f(x,x)$$

$$\frac{dC}{dK} = \frac{d}{dK}(S_T - K)\,q(S)\,dS + f(K,K) = -(K-K) - \int_K^{\infty} q(S)\,dS$$

By the **Fundamental Theorem of Calculus:**

$$\frac{d}{dx}\int_b^a f(t)\,dt = f(x)$$

$$\frac{d}{dK} - \int_K^{\infty} q(S)\,dS = -\frac{d}{dK}q(\infty) - q(K) = --q(K)$$

Because the integral is the antiderivative, so the operations essentially cancel? As a result:

$$\frac{d^2C}{dK^2} = e^{-rt}q(S)$$

52 Log Contract

Suppose we wanted to price a derivative with the payoff:

$$g(S_T)$$

$$E_Q[g(S_T)] = \int_0^\infty g(K)\, q(K)\, dK$$

the payoff is expressed in terms of K and the risk neutral density in K

$$wewantprobabilitythatK = S_T \text{ at T}$$

$$fromBreeden - Litzenbergerweknowriskneutraldensityis2ndderivativewrtK$$

can express in terms of both Calls and Puts

use continuum of OTM options which are more liquid and mathematically simpler

$$= \int_0^F g(K)\, \frac{d^2P}{dK^2}\, dK + \int_F^\infty g(K)\, \frac{d^2C}{dK^2} dK$$

if K<F OTM, K>F OTM, strike is too low or too high compared to underlying

$$integratebypartstwice, uv - \frac{du}{dv} v$$

PCP for undiscounted calls and puts is $\text{C}(K) - P(K) = F - K$

$$\left[\frac{dP}{dK} g(K)\right]_0^F - \left[P(K)\, g'(K)\right]_0^F + \left[\frac{dC}{dK} g(K)\right]_F^\infty - \left[C(K)\, g'(K)\right]_F^\infty + \int_0^F g''^{(K)} P(K)\, dK + \int_F^\infty g''^{(K)} C(K)\, dK$$

$$\left[\frac{dP}{dK} g(K)\right]_0^F + \left[\frac{dC}{dK} g(K)\right]_F^\infty = g(F)$$

the derivative of the option payoff (ramp function) is the heaviside function

$$[1_{S_T<K}\, g(K)]_0^F = g(F)\, 1_{S_T<F} - g(0)\, 1_{S_T<0}$$

$$[-1_{S_T>K}\, g(K)]_F^\infty = -g(\infty)\, 1_{S_T>\infty} + 1_{S_T>F}\, g(F)$$

$$= g(F)\, 1_{S_T<F} + 0 - 0 + g(F)\, 1_{S_T>F} = g(F)$$

$$-\left[P(K)\, g'(K)\right]_0^F - \left[C(K)\, g'(K)\right]_F^\infty$$

$$-P(F)\, 1 + P(0)\, 1 - -1C(\infty) - -1C(F) = C(F) - P(F) = F - F = 0$$

$$orequalsF - KunderPCP$$

$$orreplace = PwithC - Vtocancelthingsout?$$

$$E_Q[g(S_T)] = g(F) + \int_0^F g''(K)\, P(K)\, dK + \int_F^\infty g''(K)\, C(K)\, dK$$

See that we can express a product in terms of a portfolio of calls and puts, the number which we need to hold is g”(K). The log contract comes in because we consider a payoff which is:

$$\ln\left(\frac{S_T}{F}\right) = \int_0^F \ln\left(\frac{S_T}{F}\right)\frac{d^2P}{dK^2}\,dK + \int_F^\infty \ln\left(\frac{S_T}{F}\right)\frac{d^2C}{dK^2}dK$$

integration by parts

$$\left[\frac{dP}{dK}\ln\left(\frac{S_T}{F}\right)\right]_0^F - \int_0^F \frac{1}{S_T}\frac{dP}{dK}\,dK + \left[\frac{dC}{dK}\ln\left(\frac{S_T}{F}\right)\right]_0^F - \int_0^F \frac{1}{S_T}\frac{dC}{dK}\,dK$$

$$\ln\left(\frac{S_T}{F}\right) = lnS_T - lnF$$

lnF is a constant, derivative of a constant is 0

$$g''\left(\ln\left(S_T\right)\right) = -\frac{1}{S_T^2}$$

K is set to S?, everything is taken wrt K not S

$$E_Q\left[g\left(\ln\left(\frac{S_T}{F}\right)\right)\right] = \ln\left(\frac{F}{F}\right) - \int_0^F P(K)\,\frac{dK}{K^2} - \int_F^\infty C(K)\,\frac{dK}{K^2}$$

so the number of options to buy is $\frac{1}{K^2}$

recall for constant gamma, the 2nd derivative needed to be $\frac{1}{S_T^2}$ which is possible with this contract

using this we can trade volatility with no direction and no comovement

53 Variance Swaps

Technically a variance swap would be a forward on realised variance, with a forward price known as the variance swap strike that is equal to the expected future realised variance. Since the purpose of the variance swap, as a forward, is the pay the expected no arbitrage amount in return for the speculative realised amount. The strike is therefore set so that the contract is worth nothing and only pays at expiration. Recall that the formula for variance:

$$E\left[X^2\right] - E\left[X\right]^2$$

$$E\left[X\right] = \frac{1}{N}\sum X$$

$$N \times A \times \left\{\frac{1}{I}\sum\left[\ln\left(\frac{S_i}{S_{i-1}}\right)\right]^2 - \left[\frac{1}{I}\ln\left(\frac{S_I}{S_0}\right)\right]^2\right. - N \times K_{var}$$

Because returns are log returns, we have this expression for the variance of returns. However, in reality the $E[X]^2$ or $\left[\frac{1}{I}\ln\left(\frac{S_I}{S_0}\right)\right]^2$ term is removed, which keeps the payoff linear like a forward. Otherwise, it would be convex and difficult to manage. Note that N is vega notional, simply a term to translate variance into a meaningful number.
A annualises, given that I is daily frequency so usually 252. However, although the payoff is linear in terms of VARIANCE it is not linear in terms of VOLATILITY. To get the variance swap strike, we need the Q expectation of future realised DEMEANED variance, because by FTAP, the forward price is the risk neutral expected payoff:

$$\frac{1}{T-t}\int_t^T [dlnS_u]^2 - \left[\frac{1}{T-t}\int_t^T dlnS_u\right]^2$$

$$\text{using It}o^{'}\text{s Lemma on }\ln S_u$$

$$\left[\frac{dS_u}{S_u} - \frac{\sigma^2}{2}dt\right]^2 = \frac{\sigma^2 S_u^2}{S_u^2} - dt^2 = \sigma^2 dt??$$

$$getthisfromdlnSt = \left(r - \frac{\sigma^2}{2}\right)dt + \sigma dW_t$$

$$\frac{1}{T-t}\int_t^T \sigma^2 du - \left[\frac{1}{T-t}\int_t^T dlnS_u\right]^2$$

$$secondtermwhichismeansquaredlog - returnisdifficulttoobtain$$

remove second term due to demeaning and we see that the variance is just average realised variance

$$\text{so variance swap payoff is } \frac{1}{T-t}\int_t^T \sigma^2$$

$$\text{strike or forward price must therefore be } E_Q\left[\frac{1}{T-t}\int_t^T \sigma^2\right]$$

$$\text{we know }\ln\left(\frac{S_T}{S_t}\right) = dlnS_u = \frac{dS_u}{S_u} - \frac{\sigma^2}{2}$$

$$\frac{\sigma^2}{2} = \frac{dS_u}{S_u} - dlnS_u$$

$$\sigma^2 = 2\frac{dS_u}{S_u} - 2dlnS_u$$

$$\sigma^2 = 2\frac{dS_u}{S_u} - 2\ln\left(\frac{S_T}{S_t}\right)$$

$$K_{var} = E_Q\left[\frac{1}{T-t}\int_t^T \sigma^2\right] = \frac{1}{T-t}E_Q\left[\int_t^T \sigma^2\right]$$

$$= \frac{2}{T-t}E_Q\left[\int_t^T \frac{dS_u}{S_u}\right] - \frac{2}{T-t}E_Q\left[\ln\left(\frac{S_T}{S_t}\right)\right]$$

$$\frac{dS_u}{S_u} = (r - d)$$

$$dlnS_u = lnS_T - S_t$$

$$= -\frac{2}{T-t} E_Q [\ln S_T - lnS_t] + \frac{2}{T-t} E_Q \left[\int_t^T (r-d) \right]$$

$$= -\frac{2}{T-t} E_Q \left[\ln \left(\frac{S_T}{S_t \exp\left(\int_t^T (r-d) \right)} \right) \right]$$

$$because lnX + a = lnX + \exp(a)$$

$$F = S e^{\int_t^T (r-d)}$$

subbing in F

$$-\frac{2}{T-t} E_Q \left[\ln \left(\frac{S_T}{F_t} \right) \right] = K_{var}$$

we know the logcontract or geometric returns are replicated by $\frac{1}{K^2}$ calls and puts

we have:

$$-\frac{2}{T-t} E_Q \left[\ln \left(\frac{S_T}{F_t} \right) \right] = -2 \left[-\int_0^F P(K) \frac{dK}{K^2} - \int_F^\infty C(K) \frac{dK}{K^2} \right] = 2 \left[\int_0^F P(K) \frac{dK}{K^2} + \int_F^\infty C(K) \frac{dK}{K^2} \right] = K_{var}$$

$$= \frac{2}{T-t} \int_t^T \frac{dS_u}{S_u} - \frac{2}{T-t} \ln \left(\frac{S_T}{S_t} \right) - 2 \left[\int_0^F P(K) \frac{dK}{K^2} + \int_F^\infty C(K) \frac{dK}{K^2} \right]$$

We now have our variance swap, and we can see that the difference between arithmetic and geometric returns is the realised variance. We can hedge the log contract with a portfolio of calls and puts. We can hedge the arithmetic returns using a self-financing portfolio with MMA and stocks:

we need $\int_t^T \frac{dS_u}{S_u}$

borrow £1 in our MMA

$$b_t = -1$$

invest in a_t *stocks*

$$a_t = \frac{1}{S_t} = \frac{investment}{\text{share price}}$$

$$db_t = -(S_t + dS_t)\, da_t$$

$$theMMAwillchangebytheamountwegain * bytheamountofstockwebuyorsell$$

if we sell, da_t which is change in units held is negative

$$\text{we sell units at the new price which is } S_t + dS_t$$

$$\text{proceeds go into the MMA which is d}b_t$$

$$da_t \text{ under Ito}'\text{s Lemma:}$$

$$-\left(-\frac{1}{S_t^2}dS_t + \frac{1}{S_t^3}dS_t^2\right)(S_t + dS_t) = -\left(\frac{dS_t^2}{S_t^2} - \frac{dS_t^2}{S_t^2} + \frac{dS_t^3}{S_t^3} - \frac{dS_t}{S_t}\right)$$

$$dS_t^3 \text{ will be 0 because all terms will be multiplied by dt}$$

$$db_t = \int_t^T \frac{dS_t}{S_t}$$

$$b_T = b_t + \int_t^T \frac{dS_t}{S_t}$$

$$a_T S_T = 1$$

$$V_T = -1 + \int_t^T \frac{dS_t}{S_t} + 1 = \frac{dS_t}{S_t}$$

We assume we always keep the same amount invested, with all excess put into MMA, and all deficits borrowed. We will always end up with the arithmetic return.

53.1 Relationship between Kvar and IV

First relationship is that the variance swap strike is a Gaussian average of all the implied volatilities on the volatility smile. Weight is highest for ATM IVs, and decays further ITM or OTM. Therefore the Kvar is centred around the ATM IV Second relationship that the Kvar is the initial implied variance of twice the log contract. Assuming the variance swap is equivalent to shorting 2 log contracts. In this case, by assuming that the BS IV is correct at two points in time: at inception and at maturity. We set up a self-financing portfolio, shorting the log contracts, delta hedging based on BS delta of the contracts, and borrowing from an MMA. As we have delta stocks, and we gain based on the change in the log contracts, we are delta neutral:

$$\frac{dV}{dt}dS_t - \frac{dV}{dS}dS_t - \frac{dV}{dt}dt - \frac{1}{2}\frac{d^2V}{dS^2}dS^2$$

$$\text{recall that} \frac{dV}{dt} = -\frac{1}{2}\frac{d^2V}{dS^2}\sigma^2 S^2$$

$$\text{therefore we have } \frac{1}{2}\frac{d^2V}{dS^2}S^2\left[\sigma_{IV}^2 - \sigma^2\right]$$

$$\frac{d^2V}{dS^2} = -\frac{1}{2}S^2 \text{ therefore final payoff is}$$

$$\left[\sigma^2 - \sigma_{IV}^2\right]$$

$$\text{we wanted the second derivative to cancel out} \frac{1}{2} S^2 \text{ to give us a constant payoff}$$

$$\text{no initial cost with payoff, same as a forward or a variance swap}$$

$$K_{var} = \sigma^2_{IV}$$

We can see that the more options we are able to purchase on a continuum of strikes, the closer we are to having a perfect volatility trade that depends: Not on the underlying
Not on direction
Not on co-movement with gamma

53.2 Mark to Market

Recall that the mark to market value of a forward is:

$$V_t = e^{-r(T-t)} E_Q \left[S_T - F_0 \right]$$

Same can be applied to a variance swap:

$$V_t = e^{-r(T-t)} E_Q \left[S^2 (0, T) - K_0 \right]$$

$$S^2 (0, T) = \frac{t}{T-0} S^2 (0, t) + \frac{T-t}{T-0} S^2 (t, T)$$

You can therefore infer that the mark to market value of a variance swap is the gains made on realised variance vs the variance swap strike, plus the expected future variance you would make going forward. We know that the expected future variance, or implied volatility is equal to the variance swap strike decided at the time of writing the contract. Pb set 8, b:

$$K_{var} = E_Q \left[\frac{1}{T-t} \int_t^T \sigma^2 \right]$$

$$K (t, T_1, T_2) = E_Q \left[\frac{1}{T_2 - T_1} \int_{T_1}^{T_2} \sigma^2 \right]$$

$$\text{looking at the integral, } \int_{T_1}^{T_2} \sigma^2 = \int_t^{T_2} \sigma^2 - \int_t^{T_1} \sigma^2$$

$$\frac{1}{T_2 - T_1} E_Q \left[\int_t^{T_2} \sigma^2 \right] - \frac{1}{T_2 - T_1} E_Q \left[\int_t^{T_1} \sigma^2 \right]$$

$$\text{STRIKES for swaps K} (t, t, T_2) \text{ and K} (t, t, T_1) \text{ would be}$$

$$\frac{1}{T_2 - t} and \frac{1}{T_1 - t}$$

$$\text{multiply all by } T_2 - tandT_1 - ttoconvertthestrikestothepureQexpectationterms$$

$$\text{because we need } E_Q \left[\frac{1}{T_2 - T_1} \int_{T_1}^{T_2} \sigma^2 \right] \text{, not K} (t, t, T_2) - K (t, t, T_1)$$

54 OTM Puts and ITM Calls

Comparative statics: If left tail is fatter (lower stock prices are more likely), IV is higher for lower strikes, and OTM puts will be more expensive. But lower strikes for calls mean they are further ITM, and will also have higher IV and be more expensive. Why should we pay more for both? When we redistribute the probability density of the stock return tails, we make the left tails fatter (aka redistribute mass to the left tail). However, doing this would shift the expected return (the centre of the distribution) to the left. But we need to keep the expected return the same as in BSM. In which case, the right tail must also become fatter to maintain the expected value of S the same as before. Normal distribution mass decays exponentially as you approach tails. In reality, people tend to pay more for protection against downsides rather than upsides. Can be viewed in the FX vol smile (my currency doing well means other currency doing bad, which will lead to more protective buying on both ends), gold skew (protection against increase rather than decrease due to safe haven), equities smirk (downside protection).

55 Pinning

Why does it happen? Incentives of traders may differ – attempt to make a final push up or down to ensure options end ITM. If there is a lot of open interest for calls and puts, the trading will keep pushing the price up and down around the ATM point. Or, if FIs are LONG gamma (they have bought options), they must hedge. If stock prices go up, delta increases, they must sell bringing price back down. If stock prices go down, delta decreases, they must sell less (buy more) bringing price back down. When this happens en masse, price can be pinned to ATM point. If FIs are SHORT gamma (they have net sold options), feedback occurs. To hedge, when delta goes up they buy more. When delta goes down, they sell more. This pushes prices in the direction they were already moving. Why might expiration pin risk cause put call parity to not hold? A synthetic position in options may be physically settled rather than cash settled. If you want to take advantage of mispriced options, you cannot be sure if counterparty physically settles options you are short on. You will end up being either naked short or naked long, if you don't also physically settle your positions, leaving you exposed until the next trading day (during which prices may move and exposure you to losses).

56 PPN

How can you give a participation rate over 100% with dividends? We promise them back only gains on the index/stock, or principal – not principal with interest, or gains + dividends. So we 'collect' the dividends ourselves. What that actually means is that the options are cheaper because price is reduced by PV of dividends, so we can buy more, thereby offering higher upside.

57 Structured Product Features

Floor rate – guarantees a minimum level of payoff to be delivered to the client. Participation rate – guarantees a proportion of the upside that will be delivered to client. There is a trade-off between floor rate and participation rate. Both cannot be higher – either you get more protection, or you get more upside. Higher yield in the form of coupons – you may believe / have a certain view on direction, but not the magnitude of direction. Therefore you want compensation for the view but do not believe things will necessarily move in large magnitude. Is there downside (principal) protection built in or not? May be getting some yield but no downside protection. Increased yield may be derived from the issuer not giving dividends, buying barrier or Asian options (lower premium) Barrier options – may only want to hedge at a particular price? Cheaper because offers less than a normal option. Allows for options to reprice (employee compensation). Same as cliquets?