MGMT MFE 406 – Derivative Markets (4 units)

Part 2: Stochastic Calculus Review & Black-Scholes-Merton Framework

Prof. Eric S. Reiner

Winter 2025

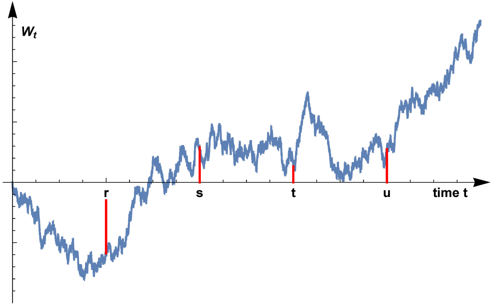
Outline

1. Review of Brownian Motion & Itô Processes
   - Brownian Motion: Construction & Properties
   - Scaling/Order Notation and Variation Properties
   - Itô Processes & Formula
2. Fundamental Pricing Framework & Feynman-Kac
   - Deriving the Fundamental Pricing PDE
   - Interpretation II (Feynman-Kac Theorem)
   - Digression on Delta Functions
3. Brownian Motion Representation & Black-Scholes-Merton Formula
4. Inputs & Sensitivities
   - Inputs
   - Sensitivities

1. Review of Brownian Motion & Itô Processes

Brownian Motion Definitions

- A stochastic process $W_t, t \geq 0$ is called a ***standard Brownian motion (SBM)*** or ***Wiener process*** if the following four conditions are satisfied:



1. it starts at the origin: $W_0 = 0$;
2. it has independent increments: for all $r < s \leq t < u$, $W_u - W_t$ is independent of $W_s - W_r$;
3. it has normally distributed (Gaussian) increments: for all $s < t$, $W_t - W_s \sim N(0, t-s)$;
4. it has continuous sample paths as a function of t: W_t is a continuous function (a.s.)

- A Wiener process is a ***martingale***: for all $s < t$, $\mathbb{E}(W_t | W_s) = W_s$

1. Review of Brownian Motion & Itô Processes (2)

Further Definitions and Notations

- Alternative notations: $W(t)$ (argument instead of subscript), w_t, B_t, z_t, ...
- A ***Brownian motion (BM)*** is a stochastic process, $X_t = X_0 + \mu t + \sigma W_t$, where μ and $\sigma > 0$ are constants, and W_t is a standard Brownian motion.
- An ***exponential*** or ***geometric Brownian motion (GBM)*** is a stochastic process, $Y_t = e^{X_t}$, where X_t is a Brownian motion.

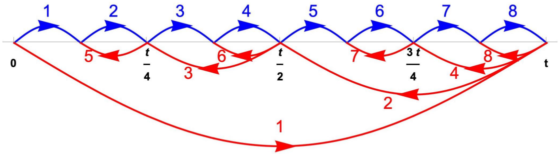
1.1. Brownian Motion: Construction & Properties

Brownian Motion Construction

- We can use a broad variety of underlying distributions (together with the central limit theorem) to construct a[n] SBM W_t as the limit of a discrete process, e.g.:
- Bernoulli increments
  1. Divide t into n increments of length $\Delta t = t/n$;
  2. $\forall i = 1, \ldots, n$: take $\Delta W_i := \pm\sqrt{\Delta t}$, each with probability $1/2$;
  3. Let $W_{t,n} = \sum_{i=1}^{n} \Delta W_i$;
  4. Take the limit $n \nearrow \infty$.
  - This is the strategy behind the binomial model.
- Successive Gaussian increments (the simplest way)
  - Follow steps 1 to 4 as for Bernoulli increments, except in step 2 in which we now take $\Delta W_i \sim N(0, \Delta t) \sim \sqrt{\Delta t}\, N(0,1) = \sqrt{\Delta t}\, z_i \; \forall i = 1, \ldots, n$.
  - This is the (usual) strategy behind Monte Carlo simulation of a Brownian process.
- What are some other possibilities?

1.1. Brownian Motion: Construction & Properties (2)

Brownian Motion Construction, continued



- Gaussian increments with recursive dyadic sub-division
  1. Begin by taking $W_t \sim N(0,t) \sim \sqrt{t}\,N(0,1) = \sqrt{t}\,z_1$ and set $n' = \log_2 n = 0$;
  2. We have $n = 2^{n'}$ increments of length $\Delta t = t/n$;
  3. We know $\{W_{(i-1)\Delta t}, W_{i\Delta t}\}$ at the beginning and end of each such increment $i : 1, \ldots, n$;
  4. $\forall i = 1, \ldots, n$: take an independent $N(0,1)$ sample z_{n+i} and let:
  $$W_{(i-1/2)\Delta t} = \frac{W_{(i-1)\Delta t} + W_{i\Delta t}}{2} + \frac{\sqrt{\Delta t}}{2} z_{n+i} = W_{(i-1)\Delta t} + \frac{\left(W_{i\Delta t} - W_{(i-1)\Delta t}\right) + \sqrt{\Delta t}\, z_{n+i}}{2};$$
  5. Set $n' = n' + 1$ and return to step 2;
  6. Take the limit $n' \nearrow \infty$.

1.1. Brownian Motion: Construction & Properties (3)

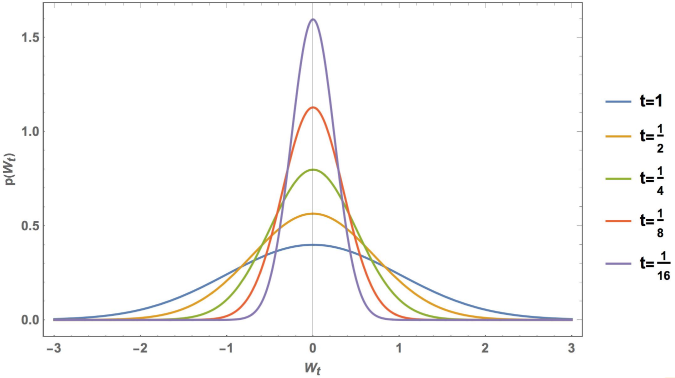
Brownian Motion Construction, continued

- Gaussian increments with recursive dyadic sub-division, continued
  - Even though it's a more complicated approach, it's helpful for investigating the fine structure of a BM since the outer "skeleton" remains static as we recurse deeper-and-deeper.
  - Simplest example of a **Brownian Bridge**: for $\{W_t, W_T\} : t \leq T$, $W_t | W_T \sim N\left(\frac{t}{T} W_T, t\left(1 - \frac{t}{T}\right) = \frac{(T-t)t}{T}\right)$
  - Useful in conjunction with high-dimensional quasi- Monte Carlo methods, where lower dimensions have more robust sampling properties.
- For a given time t commensurate (aligned) with the *partition* $\{i \cdot \Delta t\} : i = 0, ..., n$, convergence to W_t is *weaker* for a discrete (e.g., binomial) vs. a continuous approximation.
- How do we satisfy continuity in time for an approximation with $\Delta t > 0$?
  Linearly interpolate W_t between $(i-1) \cdot \Delta t$ and $i \cdot \Delta t$.
  - But note that this can cause issues if we need to rely on the properties of W_t at intermediate times: Conditional variance is too small (zero) and conditional drift perfectly forecasts values over the interval Δt.

1.1. Brownian Motion: Construction & Properties (4)

Brownian Motion Properties

- Probability density of an SBM W_t: $W_t \sim N(0,t) \Rightarrow p(W_t) = \frac{1}{\sqrt{2\pi t}} \exp(-W_t^2/2t)$



- What is the limiting behavior of $p(W_t)$ as $t \searrow 0$?
  - Strictly positive "function" of infinitely narrow width and infinite height (but finite area $= 1$).
  - Dirac's "delta function"

1.1. Brownian Motion: Construction & Properties (5)

Brownian Motion Properties, continued

- (Auto-) covariance and correlation:
  - $\mathrm{cov}(W_s, W_t) = \min(s,t)$:
    - Assume $s \leq t$; then $\mathrm{cov}(W_s, W_t) = \mathrm{cov}(W_s, W_s) + \mathrm{cov}(W_s, W_t - W_s) = \mathrm{var}(W_s) + 0 = s$
  - $\mathrm{corr}(W_s, W_t) = \rho_{W_s, W_t} = \dfrac{\min(s,t)}{\sqrt{s}\sqrt{t}} = \sqrt{\dfrac{\min(s,t)}{\max(s,t)}} \left(= \sqrt{\dfrac{s}{t}}\ \forall s \leq t \right)$
- Continuity and differentiability; scaling of increments and variations: since $dW \sim N(0, dt)$:
  - $dW \sim dt^{1/2}$
    - A BM is continuous, but not very smooth (locally Hölder continuous only for exponents $\alpha < 1/2$)
    - Almost surely, the sample path W_t of a Brownian motion is nowhere differentiable.
    - W_t is of unbounded variation;
  - $(dW)^2 \sim dt$

    $\Rightarrow$ W_t is of bounded quadratic variation.
    - In fact, $dW^2 = dt$ (a.s.).

1.2. Scaling/Order Notation and Variation Properties

Scaling/Order Notation

- Consider behavior of a function $f(x)$ near $x=x_0$. If there is a Taylor-Laurent series for $f(x)-f(x_0)$ s.t.:

$$f(x)-f(x_0) = a_0(x-x_0)^{\alpha} + a_1(x-x_0)^{\alpha+1} + \ldots$$

then we say that $f(x)-f(x_0) \sim (x-x_0)^{\alpha}$, $f(x)-f(x_0) \sim \mathcal{O}(x-x_0)^{\alpha}$, or f ***scales*** as $(x-x_0)^{\alpha}$ or f ***"big O"*** $(x-x_0)^{\alpha}$.

  - More formally: $\exists c: c \geq 0$ s.t. $\lim_{x\to x_0} \left| \frac{f(x)-f(x_0)}{(x-x_0)^{\alpha}} \right| = c$
  - Intuition: as $x \to x_0$, limiting behavior of deviations from $f(x_0)$ is $a_0(x-x_0)^{\alpha}$
  - Most common usage when $x_0 = 0$ and $f(x_0) = 0$
  - Residual or next/higher order terms: $f(x)-f(x_0) = a_0(x-x_0)^{\alpha} + \mathcal{O}(x-x_0)^{\alpha+1}$
  - Generalization (non-integer powers): $f(x)-f(x_0) = a_0(x-x_0)^{\alpha} + a_1(x-x_0)^{\alpha+\beta} + a_2(x-x_0)^{\alpha+2\beta} + \ldots$
  - More generally, "big O" implies that the first exponent with a non-zero coefficient is ***no smaller than*** α
  - We use a similar notion/notation for the behavior of a function as $x \to \pm\infty$, i.e.: "$f(x)$ approaches zero (for $\alpha < 0$) – or diverges (for $\alpha > 0$) – as (proportional to) x^{α} as $x \nearrow +\infty$"
    - In this case, higher order terms (h.o.t.) will have exponents *less than* α
    - Another generalization is *functional* scaling, e.g., $n(z) \sim \mathcal{O}\left(e^{-z^2/2}\right)$ as $z \to \pm\infty$

1.2. Scaling/Order Notation and Variation Properties (2)

Scaling/Order Notation, continued

- There are many variants of and refinements to this notation.
  - Probably the only one potentially relevant to us is o ("little o") notation, which represents scaling ***of higher order than*** (or, at ∞ divergence slower than/convergence to zero faster than) $(x-x_0)^\alpha$
  - More formally: $\lim_{x\to x_0}\left|\frac{f(x)-f(x_0)}{(x-x_0)^\alpha}\right| = 0$
  - So, if "big O" is analogous to $\leq$, then "little o" is analogous to $<$.

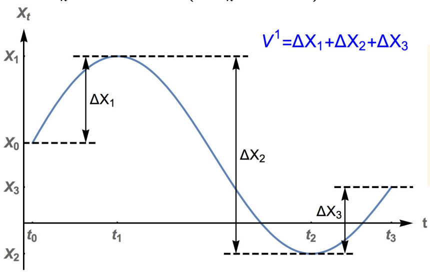
1.2. Scaling/Order Notation and Variation Properties (3)

Variation of a Brownian Motion

- a BM's variation properties are very different to those of standard functions
  - Define the m^{th} ***variation*** between a and b of a function (or process) X, $V^m(a,b) = \mathcal{V}^m[X]$, as:

$$V^m(a,b) = \lim_{\Delta t_k \searrow 0} \sum_{a<t_k<b} |X_{t_{k+1}} - X_{t_k}|^m$$

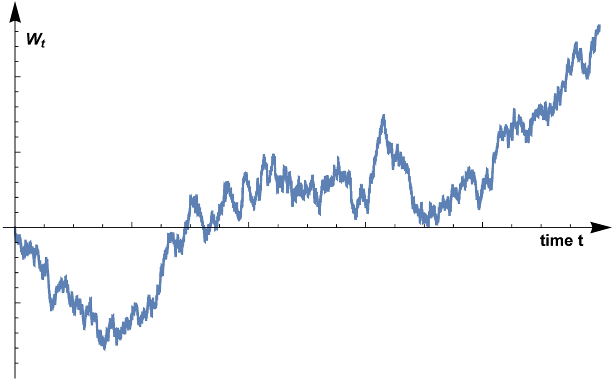
  - If $a=0$, we write $V^m(b)$. V^2 is called the ***quadratic variation***.
  - If $V^1(a,b) < \infty$ for a function X, X is said to be of ***bounded variation***.
  - Smooth functions are of bounded variation and have zero higher variation.
    - $V^1(a,b)$ is the total (absolute) distance traveled vertically when going from $X(a)$ to $X(b)$.
    - For $m \geq 1$, $V^m(t) = \lim V^m_n(t) \approx \sum_k^n |X'_{t_k} \Delta t|^m \sim \left(\sum_k^n |X'_{t_k}|^m \Delta t\right)(\Delta t)^{m-1} \sim \mathcal{O}(\Delta t)^{m-1}$



1.2. Scaling/Order Notation and Variation Properties (4)

Variation of a BM, continued

- Almost surely, the sample path of a Wiener process:
  1. is of infinite variation, $V^1(t) = \mathcal{V}^1[W_t]$ satisfies $V^1(t) \nearrow \infty$ for all $t > 0$,
  2. has quadratic variation function $V^2(t) = t$,
  3. has zero higher order variation(s).



1.2. Scaling/Order Notation and Variation Properties (5)

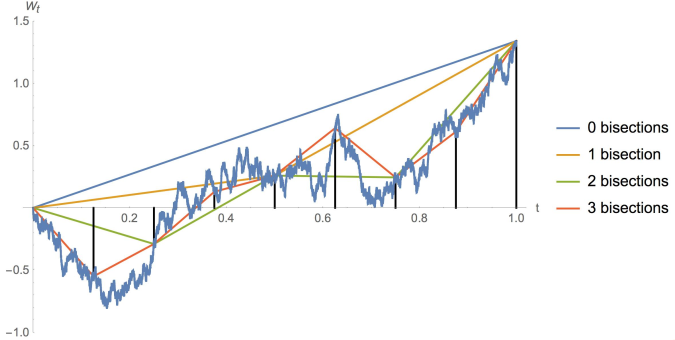
Mathematica code approximating the construction of a[n] SBM

```
T = 1.0; (* Total time interval of 1 *)
Log2n = 20; (* Bisect interval 20 times *)
n = 2 ^ Log2n; (* Yielding a total of n = 2^20 ~ 1 million time steps *)
dT = T / n; (* dT = 1/2^20 *)
TTable = dT * Table[i, {i, 0, n}]; (* Table of times for each time step *)
SeedRandom[0]; (* Initialize random number generator *)
dWt = Sqrt[dT] RandomVariate[NormalDistribution[], n]; (* Table of n Brownian increments
                                                          dWt, each of variance dT *)
Wt = Table[0, {i, 0, n}];
For[i = 1, i <= n, i ++, Wt[[i + 1]] = Wt[[ i ]] + dWt[[ i]] ]; (* Build up Wt from dWt *)
```

- Which construction approach is this taking?
- It is of some practical importance (i.e., for dynamic hedging) to understand the “fine structure” of an asset price process (most simply, a[n] SBM) as it is divided into smaller-and-smaller partitions.
- In particular, we would like to explore the (path-wise and statistical) properties of the variation $\mathcal{V}^m[W_t]$ for various m $(= 1, 2, 3, \ldots)$ as $\Delta t = t/n \searrow 0$.

1.2. Scaling/Order Notation and Variation Properties (6)

- Explore limiting behavior of $V^m(t = T = 1) = \lim_{n \nearrow \infty} \sum_{i=1}^{n} |X_{i/n} - X_{(i-1)/n}|^m$
- Take n as powers of 2 (successive "dyadic" bisections), ranging from 2^0 to 2^{20}:



- Mathematica code for this (given n):

```
Variation[m_, Vec_, Len_, n_] := Sum[Abs[Vec[[1 + i (Len - 1)/n]] - Vec[[1 + (i - 1) (Len - 1)/n]]]^m, {i, 1, n}];
```

1.2. Scaling/Order Notation and Variation Properties (7)

- Since we can write $V_n^m(t)$ along a particular path of W_t as $\sum_{k=1}^n (|z_k|\sqrt{\Delta t})^m$, with z_k being samples from the standard normal distribution $N(0,1)$ and $\Delta t \equiv t/n$, it is fairly clear that, to leading order: $V_n^m(t) \sim \mathcal{O}(n \cdot (t/n)^{m/2}) = \mathcal{O}(t^{m/2} n^{1-m/2})$.
- What is a little less obvious (although the CLT and continuity properties can give us some hints) is whether as $n \nearrow \infty$, that scaling is:
  - ***regular***, i.e. $V_n^m(t) \to c \cdot t^{m/2} n^{1-m/2}$ for some constant $c(W_t)$ along a particular path or not;
  - If so, whether c actually depends on W_t or not.
- Beyond leading order, scaling arguments suggest (the CLT also helps here given that the z_k are presumably independent samples from the standard normal distribution) that the next contribution to $V_n^m(t)$ in powers of Δt or n is (at most) $1/\sqrt{n}$ as large:
  - I.e., (dropping explicit t dependence) $\mathcal{O}(n^{(1-m)/2})$
  - Again, the same questions hold...
- So let's run some experiments!

1.2. Scaling/Order Notation and Variation Properties (8)

Experimental Strategy

- Start with the SBM path we've generated (we'll take $t = 1$ for simplicity, so $dt = 1/n$) and consider $m = 1, 2, 3$ in turn.
- Successively bisect the path and examine V_n^m as a function of n (or its $\log_2$) as n becomes large.
- Pick a comparison function $c_m\, n^{1-m/2}$, with c_m an $\mathcal{O}(1)$ constant, and compare it to V_n^m on our path.
  - Thus, we can determine whether the $n \nearrow \infty$ behavior of V_n^m is regular and, if so, select $c_m(W_t)$ to match the observed asymptotic scaling.
- If we find that the leading-order scaling is regular, elucidate the next order behavior by examining $n^{(m-1)/2}\left(V_n^m - c_m\, n^{1-m/2}\right)$ or $\sqrt{n}\left(\frac{V_n^m}{c_m\, n^{1-m/2}} - 1\right)$.
- We can then repeat the experiment for other sample paths of W_t and observe whether the results are consistent with those for our first path.

1.2. Scaling/Order Notation and Variation Properties (9)

1^{st} Variation

- Table of V^1 vs. number of recursive bisections $\log_2(n)$:

```
Var1Tab = Table[{j, Variation[1, Wt, n + 1, 2^j]}, {j, 0, Log2n}]
```

```
{{0, 1.34034}, {1, 1.34034}, {2, 1.94937}, {3, 3.23239}, {4, 3.97917}, {5, 4.61517}, {6, 6.64896},
 {7, 9.07754}, {8, 13.2017}, {9, 19.0728}, {10, 26.4363}, {11, 36.4464}, {12, 51.5016}, {13, 71.8388},
 {14, 100.996}, {15, 143.477}, {16, 203.219}, {17, 286.754}, {18, 406.942}, {19, 577.706}, {20, 816.57}}
```

- For comparison, table of $\sqrt{\frac{2}{\pi\, dt}}$ vs. number of recursive bisections $\log_2(n)$:

```
Sqrt1Tab = Table[{j, Sqrt[2/π 2^j] // N}, {j, 0, Log2n}]
```

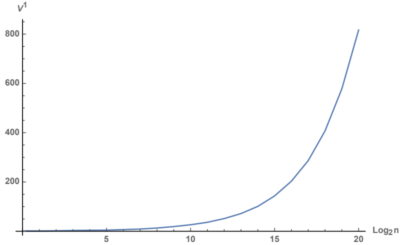
```
{{0, 0.797885}, {1, 1.12838}, {2, 1.59577}, {3, 2.25676}, {4, 3.19154}, {5, 4.51352}, {6, 6.38308},
 {7, 9.02703}, {8, 12.7662}, {9, 18.0541}, {10, 25.5323}, {11, 36.1081}, {12, 51.0646}, {13, 72.2163},
 {14, 102.129}, {15, 144.433}, {16, 204.258}, {17, 288.865}, {18, 408.517}, {19, 577.73}, {20, 817.034}}
```

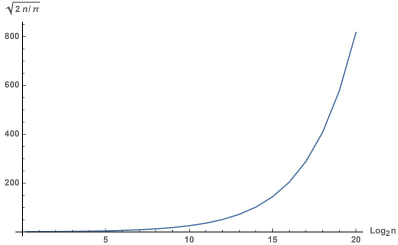
(remember that because $t = T = 1$, $dt = n^{-1} = 2^{-j}$)

1.2. Scaling/Order Notation and Variation Properties (10)

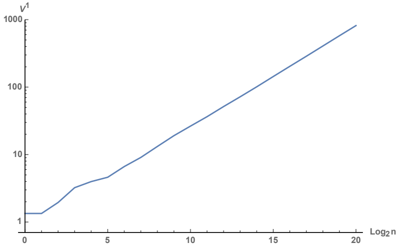
1^{st} Variation (continued)

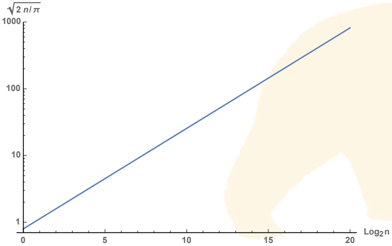
- Plot V^1 and $\sqrt{2/(\pi dt)}$ vs. $\log_2(n)$:





- Log scale is more revealing:

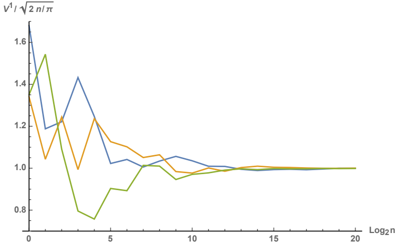


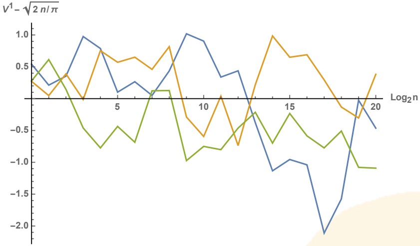


1.2. Scaling/Order Notation and Variation Properties (11)

1^{st} Variation (continued)

- Let's try to validate the leading order behavior of V^1 vs. n and examine the residual $V^1 - \sqrt{2/(\pi dt)}$ (asymptotically as $n \nearrow \infty$, the next term in n) as well:
- Plot ratio of V^1 to $\sqrt{2/(\pi dt)}$; also plot difference between V^1 and $\sqrt{2/(\pi dt)}$:





- (1^{st}) variation diverges as $\sqrt{\frac{2}{\pi dt}} \cdot t = \sqrt{\frac{2\,n\,t}{\pi}}$
- Next order term is stochastic, not regular in dt, and is $\mathcal{O}(t^{1/2} dt^0) = \mathcal{O}(t^{1/2} n^0)$
- (Additional paths shown in yellow and green)

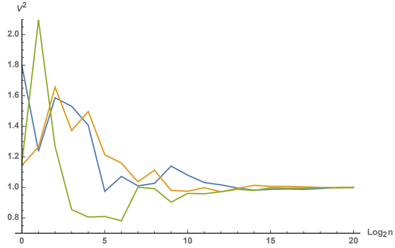
1.2. Scaling/Order Notation and Variation Properties (12)

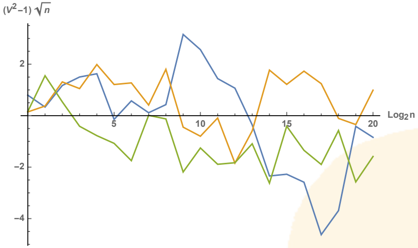
2^{nd} Variation

- Perform the same analysis (leading order behavior + residual) for V^2 vs. n:

```
Var2Tab = Table[{j, Variation[2, Wt, n + 1, 2^j]}, {j, 0, Log2n}]

{{0, 1.7965}, {1, 1.23655}, {2, 1.58659}, {3, 1.53094}, {4, 1.40709}, {5, 0.974038}, {6, 1.07184}, {7, 1.0101},
 {8, 1.02679}, {9, 1.13947}, {10, 1.08014}, {11, 1.03182}, {12, 1.01669}, {13, 0.996045}, {14, 0.981692},
 {15, 0.987428}, {16, 0.989899}, {17, 0.98723}, {18, 0.992793}, {19, 0.999427}, {20, 0.999178}}
```

- Plot V^2 and $(V^2 - 1)/\sqrt{dt}$ (why this choice of scaling?) vs. $\log_2(n)$:





- Quadratic (2^{nd}) variation converges to $t = T$; next order term is $\mathcal{O}(t^{1/2} dt^{1/2} = t\, n^{-1/2})$

- (Additional paths shown in yellow and green)

1.2. Scaling/Order Notation and Variation Properties (13)

3^{rd} Variation

- Table of V^3 vs. number of recursive bisections $\log_2(n)$:

```
Var3Tab = Table[{j, Variation[3, Wt, n + 1, 2^j]}, {j, 0, Log2n}]

{{0, 2.40792}, {1, 1.28213}, {2, 1.50555}, {3, 0.814006}, {4, 0.568063},
 {5, 0.249959}, {6, 0.214042}, {7, 0.144946}, {8, 0.101891}, {9, 0.0866411}, {10, 0.0559643},
 {11, 0.0373588}, {12, 0.0257453}, {13, 0.0177118}, {14, 0.0122179}, {15, 0.00864776},
 {16, 0.00613067}, {17, 0.00434001}, {18, 0.00308552}, {19, 0.0022007}, {20, 0.00155652}}
```

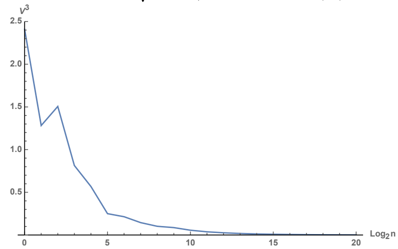
- For comparison, table of $\sqrt{\frac{8\,dt}{\pi}}$ vs. number of recursive bisections $\log_2(n)$:

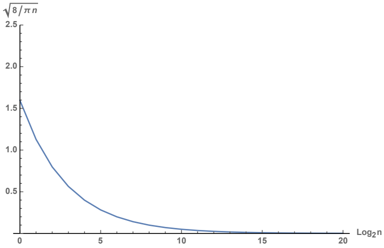
```
Sqrt3Tab = Table[{j, Sqrt[8/π 2^-j] // N}, {j, 0, Log2n}]

{{0, 1.59577}, {1, 1.12838}, {2, 0.797885}, {3, 0.56419}, {4, 0.398942},
 {5, 0.282095}, {6, 0.199471}, {7, 0.141047}, {8, 0.0997356}, {9, 0.0705237}, {10, 0.0498678},
 {11, 0.0352618}, {12, 0.0249339}, {13, 0.0176309}, {14, 0.0124669}, {15, 0.00881546},
 {16, 0.00623347}, {17, 0.00440773}, {18, 0.00311674}, {19, 0.00220387}, {20, 0.00155837}}
```

1.2. Scaling/Order Notation and Variation Properties (14)

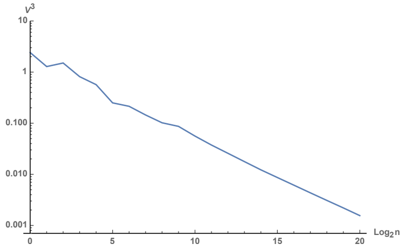
3^{rd} Variation (continued)

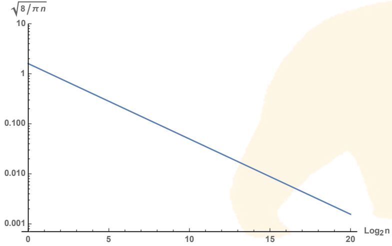
- Plot V^3 and $\sqrt{8\,dt/\pi}$ vs. $\log_2(n)$:





- Again, log scale is more revealing:

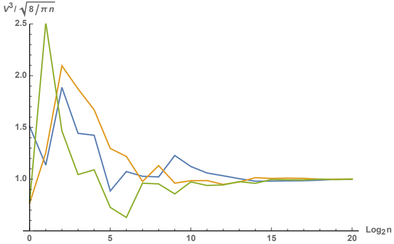


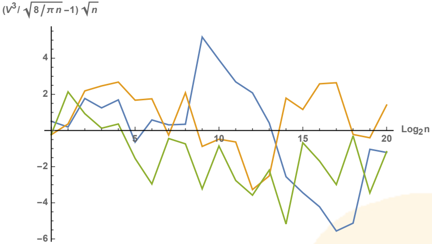


1.2. Scaling/Order Notation and Variation Properties (15)

3^{rd} Variation (continued)

- Again, verify leading-order behavior and examine residual.
- Plot $V^3/\sqrt{8\,dt/\pi}$ and $(V^3/\sqrt{8\,dt/\pi} - 1)/\sqrt{dt}$ (why this scaling?) vs. $\log_2(n)$:





- 3^{rd} variation converges to zero as $\sqrt{\frac{8\,dt}{\pi}} \cdot t = \sqrt{\frac{8\,t^3}{\pi\,n}}$
- (Additional paths shown in yellow and green)

1.2. Scaling/Order Notation and Variation Properties (16)

Why do we care about this stuff?

- (1^{st}) variation measures path length, or sum of (absolute) (log) price changes.
  - Relevant in modeling of transaction costs for rebalancing at fixed Δt.
  - Even for more sophisticated hedging strategies, this implies that there is a mean-variance trade-off between expected transaction costs and hedging variance as Δt (or its equivalent) is varied. See:
    Toft, K.B., "On the Mean-Variance Tradeoff in Option Replication with Transactions Costs," *Journal of Financial and Quantitative Analysis* **31**(2), (Jun, 1996), pp. 233-263. https://doi.org/10.2307/2331181
- Finiteness of quadratic (2^{nd}) variation drives several things:
  - Finite variance of Brownian motion;
  - Quadratic term in Itô's formula;
  - Role of volatility in the Black-Scholes PDE and formula: read Seidenverg, E., "A Case of Confused Identity," *Financial Analysts Journal* **44**(4), (Jul-Aug, 1988), pp. 63-67. www.jstor.org/stable/4479129
  - The residual tells us how we should expect hedging variance to scale with Δt, even in the absence of transaction costs.
- Convergence to zero of higher-order variations is also important:
  - Operation of central limit theorem:
    higher moments of the distribution of increments become irrelevant in the limit $dt \searrow 0$;
  - Itô's formula terminates at second order.

1.3. Itô Processes & Formula

- BM definition: $X_t = X_0 + \mu t + \sigma W_t$, with μ and $\sigma > 0$ constants, and W_t an SBM.
- Integral representation:

$$X_t = X_0 + \mu \int_0^t ds + \sigma \int_0^t dW_s = X_0 + \mu \int_0^t ds + \sigma \fint_0^t dW_s$$

  - $\fint$ or $\fint$ $\Rightarrow$ path(-dependent) integral
- Generalization: μ and $\sigma > 0$ functions of time:

$$X_t = X_0 + \int_0^t \mu_s \, ds + \int_0^t \sigma_s \, dW_s = X_0 + \int_0^t \mu_s \, ds + \fint_0^t \sigma_s \, dW_s$$

  - Definition of an Itô process X_t
  - Technically, μ and σ need not be explicit functions of time but can be any ***predictable/adapted*** functions (processes) of t and X, subject to some boundedness/regularity conditions
  - Differential form:

$$dX_t = \mu(t, X) \, dt + \sigma(t, X) \, dW_t$$

  - Named after Kiyosi Itô (Itō Kiyoshi: 伊藤 清 (いとう きよし))

1.3. Itô Processes & Formula (2)

Formula for changing variables from one Itô process to another

- Ordinary chain rule for derivatives:
  - Given sufficiently well-behaved functions $f(t)$ and $g(t, f(t))$:

$$dg = \frac{\partial g}{\partial t} dt + \frac{\partial g}{\partial f} df \;\Rightarrow\; \frac{dg}{dt} = \frac{\partial g}{\partial t} + \frac{\partial g}{\partial f} \frac{\partial f}{\partial t}$$

  - The problem is that (a function of) a Brownian motion, i.e. a function of unbounded variation, isn't sufficiently well behaved (df is of order $dt^{1/2}$, not dt)
- Taylor expand and collect terms in f & t, making sure we account for all terms of order dt & $dt^{1/2}$:

$$dg = \frac{\partial g}{\partial t} dt + \frac{\partial g}{\partial f} df + \frac{1}{2} \frac{\partial^2 g}{\partial f^2} (df)^2 + \mathcal{O}(dt^{3/2})$$

- In particular, if $df = \mu(t, f) dt + \sigma(t, f) dW_t$ (an Itô process):

$$\begin{aligned} dg &= \frac{\partial g}{\partial t} dt + \frac{\partial g}{\partial f} (\mu dt + \sigma dW_t) + \frac{1}{2} \frac{\partial^2 g}{\partial f^2} (\sigma^2 dt) + \mathcal{O}(dt^{3/2}) \\ &= \left(\frac{\partial g}{\partial t} + \mu \frac{\partial g}{\partial f} + \frac{\sigma^2}{2} \frac{\partial^2 g}{\partial f^2} \right) dt + \sigma \frac{\partial g}{\partial f} dW_t \end{aligned}$$

2. Fundamental Pricing Framework & Feynman-Kac

Standard Model for Asset Prices

- Consider a $1+1$ element (vector) stochastic process:

$$\begin{aligned} \frac{dB}{B} &= \frac{dS_0}{S_0} = r\,dt && \text{(riskless asset, money–market account)} \\ \frac{dS}{S} &= \frac{dS_1}{S_1} = (\mu_1 - y_1)\,dt + \sigma_1\,dW_1 && \\ &= (\mu - y)\,dt + \sigma\,dW && \text{(risky asset)} \end{aligned}$$

  - r, μ, y, σ in principle are functions of S, t (more generally, are ***adapted*** processes).
  - For concreteness, we are considering W in (a)(the) **P** measure: $W^{\mathbf{P}}$

- Alternative process for numéraire (0^{th}) asset:

$$\frac{d1}{1} = (r - r)\,dt + 0\,dW_0$$

  - Cash is an asset that pays a yield equal to the money-market rate.

- There is nothing special or privileged about our choice of money-market account, ZCB, or even \$ cash as numéraire!

2.1. Deriving the Fundamental Pricing PDE

- Seek to value a derivative asset $C(S,t)$
- From Itô's formula: $dC = \left[\frac{\partial C}{\partial t} + (\mu - y)S\frac{\partial C}{\partial S} + \frac{\sigma^2}{2}S^2\frac{\partial^2 C}{\partial S^2}\right]dt + \sigma S\frac{\partial C}{\partial S}dW$
- Assuming that we can calculate Δ_S, form a portfolio Π consisting of: $+1$ unit of C and $-(\Delta_S \doteq \partial C/\partial S)$ units of S.
- The portfolio's return (change in value) $d\Pi$ over dt is:

$$
\begin{aligned}
d\Pi &= \underbrace{\left[\frac{\partial C}{\partial t}+(\mu-y)S\frac{\partial C}{\partial S}+\frac{\sigma^2}{2}S^2\frac{\partial^2 C}{\partial S^2}\right]dt+\sigma S\frac{\partial C}{\partial S}dW}_{\substack{\text{change in}\\ \text{option value}\\ dC}} \overbrace{\underbrace{-\frac{\partial C}{\partial S}\left[(\mu-y)Sdt+\sigma SdW\right]}_{\substack{-\ \text{capital returns}\\ \text{on hedge}\\ \Delta_S dS}}\underbrace{-yS\frac{\partial C}{\partial S}dt}_{\substack{-\ \text{dividends}\\ \text{on hedge}\\ yS\Delta_S dt}}}^{-\ \text{total return on hedge}} \\
&= \left[\frac{\partial C}{\partial t}+(\mu-y)S\frac{\partial C}{\partial S}+\frac{\sigma^2}{2}S^2\frac{\partial^2 C}{\partial S^2}-\mu S\frac{\partial C}{\partial S}\right]dt+\left[\sigma S\frac{\partial C}{\partial S}dW-\sigma S\frac{\partial C}{\partial S}dW\right] \\
&= \left[\frac{\partial C}{\partial t}-yS\frac{\partial C}{\partial S}+\frac{\sigma^2}{2}S^2\frac{\partial^2 C}{\partial S^2}\right]dt
\end{aligned}
$$

2.1. Deriving the Fundamental Pricing PDE (2)

- Π is locally riskless (and independent of drift μ) over dt, so it must earn the riskless rate:

$$d\Pi = r\,\Pi\,dt = r\left(C - S\frac{\partial C}{\partial S}\right)dt$$

- Combining and rearranging:

$$\left(\frac{\partial C}{\partial t} - yS\frac{\partial C}{\partial S} + \frac{\sigma^2}{2}S^2\frac{\partial^2 C}{\partial S^2}\right)dt = r\left(C - S\frac{\partial C}{\partial S}\right)dt$$

$$\Rightarrow \quad \frac{\partial C}{\partial t} + (r-y)S\frac{\partial C}{\partial S} + \frac{\sigma^2}{2}S^2\frac{\partial^2 C}{\partial S^2} = rC$$

  - BACKWARDS (in time) diffusion (or conduction) / convection / reaction equation
  - What's the diffusion coefficient?

- Interpretation I (Economic):

$\dfrac{\partial C}{\partial t}$	$=$	$r\left(C - S\dfrac{\partial C}{\partial S}\right)$	$+$	$yS\dfrac{\partial C}{\partial S}$	$-$	$\dfrac{\sigma^2}{2}S^2\dfrac{\partial^2 C}{\partial S^2}$
change in option value	$=$	interest earned on cash position	$+$	dividends earned on stock position	$-$	hedge slippage (convexity value of information revealed)

2.2. Interpretation II (Feynman-Kac Theorem)

Feynman-Kac: the basic version

- Consider the partial differential equation:

$$\frac{\partial C}{\partial t} + \mu(x,t)\frac{\partial C}{\partial x} + \frac{\sigma^2(x,t)}{2}\frac{\partial^2 C}{\partial x^2} = r(t)C$$

  in an open $\{\mathbb{R}^1 \times \mathbb{R}^-\} = \{x,t\}$ region $\boldsymbol{\Omega} = \{-\infty < x < +\infty, t < T\}$, subject to the Dirichlet condition:

$$C_{\boldsymbol{\partial\Omega}} = C(t{=}T) = C_T(x_T) \text{ on the boundary } \boldsymbol{\partial\Omega} = \{-\infty < x_T < +\infty, t = T\}$$

- The solution to this equation $C(x,t)$ is equal to the expectation:

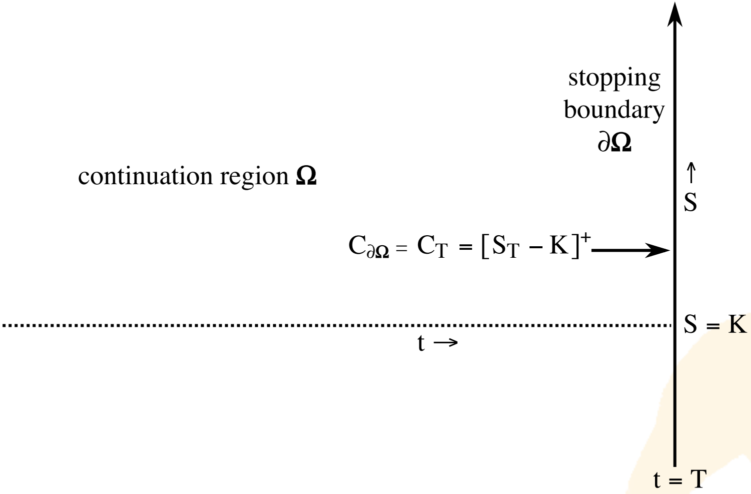
$$\mathbb{E}_t\left[e^{-\int_t^T r(t')dt'} C_T(x_T)\right] = e^{-\int_t^T r(t')dt'}\mathbb{E}_t[C_T(x_T)]$$

  for the Itô process: $dx = \mu(x,t)\,dt + \sigma(x,t)\,dW$ starting at $x(t) = x$

- Terminology:
  - $\boldsymbol{\Omega}$ is called the ***continuation region***.
  - $\boldsymbol{\partial\Omega}$ is called the ***stopping boundary***.
  - T is called (the)(a) ***stopping*** (or ***hitting*** or ***exit***) ***time***.
- Named after Richard Feynman and Mark Kac

2.2. Interpretation II (Feynman-Kac Theorem) (2)

Illustration of Stopping Boundary (1)



2.2. Interpretation II (Feynman-Kac Theorem) (3)

Feynman-Kac Examples

1. Assume parameters are constant. Let $\partial\Omega = \{S_T\colon 0 < S_T < \infty, t = T\}$, $C_{\partial\Omega} = \max[S_T - K, 0]$ (European call option). Then:

$$C(S,t) = e^{-r(T-t)}\,\mathbb{E}_t^{\mathbf{Q}}[\max(S_T - K, 0)]$$

given:

$$dS = (r-y)S\,dt + \sigma S\,dW^{\mathbf{Q}} \quad \text{and} \quad S(t) = S$$

2. Same as 1, but:

$$C_{\partial\Omega} = \mathbf{1}_{S_T-K} = \mathbf{H}_{S_T-K} = \mathbf{I}_{S_T-K} = \boldsymbol{\Theta}_{S_T-K} = \begin{cases} 0, & S_T < K; \\ 1, & S_T \geq K. \end{cases} \text{ (Heaviside step function)}$$

This is a binary (cash-or-nothing) call option struck at K.

$$C(S,t) = e^{-r(T-t)}\,\mathbb{E}_t^{\mathbf{Q}}[\mathbf{1}_{S_T-K}]$$

Interpretation in terms of (discounted) (complementary) cdf of terminal asset price S_T

2.2. Interpretation II (Feynman-Kac Theorem) (4)

3. Same as 1, but:

$$C_{\partial\Omega} = \delta(S_T, S_T^*) = \delta(S_T - S_T^*) \text{ (Dirac delta "function"/distribution)}$$

Then:

$$C(S,t) = e^{-r(T-t)}\, \mathbb{E}_t^{\mathbf{Q}}[\delta(S_T, S_T^*)]$$

- Arrow-Debreu state price density / Green's function / fundamental solution to the B-S PDE
- Formal solution, but what smoothness is required for the Black-Scholes assumptions to apply?

4. Same as 3, but:

$$C_{\partial\Omega} = \max[S_T - K, 0] = \int_K^\infty dS_T^*\; \delta(S_T, S_T^*)\,(S_T^* - K)$$

Then:

$$\begin{aligned} C(S,t) &= \int_K^\infty dS_T^* \left\{ e^{-r(T-t)}\, \mathbb{E}_t^{\mathbf{Q}}[\delta(S_T, S_T^*)] \right\} (S_T^* - K) \\ &= e^{-r(T-t)} \int_K^\infty dS_T^*\, (S_T^* - K)\, \mathbb{E}_t^{\mathbf{Q}}[\delta(S_T, S_T^*)] \end{aligned}$$

- Convolution representation of payoff at time T
- Present value of any "smooth" payoff can be represented using the Green's function or fundamental solution
- In principle, $C(S_T)$ needs to be a continuous function

2.2. Interpretation II (Feynman-Kac Theorem) (5)

- The solution to the Black-Scholes PDE can therefore be interpreted as a discounted expectation with respect to the "risk-neutral" process:
$$\frac{dS}{S} = (r - y)\,dt + \sigma\,dW^{\mathbf{Q}}$$
in which all discounting is done at the risk-free rate r.
- Hence the notion of risk-neutral valuation: the ability to (continuously) hedge away all (instantaneous) risk allows us to transform from:
  - pricing in the original probability space / measure generated by W (or $W^{\mathbf{P}}$ for definitiveness) with drift μ and preference-dependent discounting ("stochastic discount factor")
    to
  - pricing in the probability space / measure generated by $W^{\mathbf{Q}}$ with drift r and risk-free discounting.
- $W^{\mathbf{P}}$ and $W^{\mathbf{Q}}$ are ***equivalent*** ($W^{\mathbf{P}} \sim W^{\mathbf{Q}}$) in the sense that they share the same set of paths (structure of the filtration) and distribution, but they are not necessarily ***identical*** path-by-path for any particular realization of the price process.
- We call $\mathbf{Q}$ an (equivalent) ***martingale*** measure because the re-invested ("total return") risky asset price process is a martingale when normalized by the value of the money-market account (numéraire asset):
$$\mathbb{E}_t^{\mathbf{Q}}\left[\frac{e^{y(T-t)}S(T)}{B(T)}\right] = \frac{S(t)}{B(t)}$$

2.3. Digression on Delta Functions

- In section 1.2, we considered the limiting behavior of the density $p(W_t)$ of an SBM as $t \searrow 0$:
  - $W_t \sim N(0,t) \;\Rightarrow\; p(W_t) = \dfrac{1}{\sqrt{2\pi t}} \exp(-W_t^2/2t)$
  - In the limit, strictly positive "function" of infinitely narrow width (around zero) and infinite height, but finite area $= 1$:
$$\int_{-\infty}^{\infty} dW_t\, p(W_t) \equiv 1$$
- We can extend this to encompass a non-zero starting point:
  - $W_t \sim N(W_0,t) \;\Rightarrow\; p(W_t|W_0) = \dfrac{1}{\sqrt{2\pi t}} \exp\Big(-(W_t - W_0)^2/2t\Big)$
  - Now, $p(W_t|W_0)$ is localized around W_0, but otherwise retains the same properties.
- We define Dirac's "delta function" $\delta(x,x_0) = \delta(x-x_0)$ as this limit:
$$\delta(x,x_0) = \delta(x-x_0) \doteq \lim_{t\searrow 0} \frac{1}{\sqrt{2\pi t}} \exp\Big(-(x-x_0)^2/2t\Big)$$

2.3. Digression on Delta Functions (2)

- An alternative (perhaps ***slightly*** more rigorous) route is to define the delta function in terms of its integration properties with respect to a suitably well-behaved set of "test" functions $f(x)$:

$$\int_a^b dx\,\delta(x-x_0)\,f(x) \doteq f(x_0)\ \forall x_0 : a < x_0 < b$$

- This definition allows us to derive some useful properties:
  - Scaling and symmetry. $\forall\alpha \neq 0$, letting $u \doteq \alpha(x-x_0)$:

$$\int_{-\infty}^{\infty} dx\,\delta(\,\alpha(x-x_0)\,)\,f(x) = \int_{-\infty}^{\infty} \frac{du}{|\alpha|}\,\delta(u)\,f(x_0+u/\alpha) = \frac{1}{|\alpha|}f(x_0) \text{ and so } \delta(\alpha x) = \frac{\delta(x)}{|\alpha|}.$$

    - If α is negative, i.e., $\alpha = -|\alpha|$, then the "sense" of the integration is reversed.
    - Corollaries: $\delta(-x) = \delta(x)$; $\delta(\bullet)$ is an even function and is homogeneous of degree -1.
  - Composition. If $g(x)$ is a function with n distinct roots $\{x_i, i = 1, \ldots, n\}$:

$$\int_{-\infty}^{\infty} dx\,\delta(\,g(x)\,)\,f(x) = \sum_{i=1}^{n} \frac{f(x_i)}{|g'(x_i)|}$$

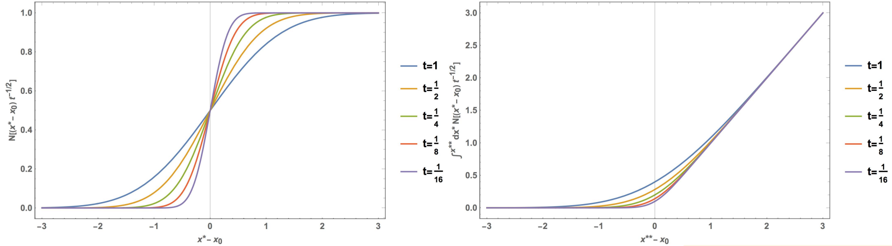
  - ... and even *derivatives* of $\delta(x)$:

$$\int_{-\infty}^{\infty} dx\,\delta'(x-x_0)\,f(x) = \Big[\delta(x-x_0)\,f(x)\Big]_{x=-\infty}^{x=+\infty} - \int_{-\infty}^{\infty} dx\,\delta(x-x_0)\,f'(x) = -f'(x_0)$$

2.3. Digression on Delta Functions (3)

Properties of the delta function, continued

- Integral: $\int_{-\infty}^{x^*} dx\, \delta(x-x_0) = \lim_{t \searrow 0} \int_{-\infty}^{x^*} dx \frac{1}{\sqrt{2\pi t}} \exp\left(-(x-x_0)^2/2t\right) = \lim_{t \searrow 0} \mathcal{N}\left[\frac{x^*-x_0}{\sqrt{t}}\right] = \mathbf{1}_{x^*-x_0}$
  - The integral of a delta function is a step function.
  - Vice-versa: the (first) derivative of a step function is a delta function.



- Iterated integral: $\int_{-\infty}^{x^{**}} dx^* \int_{-\infty}^{x^*} dx\, \delta(x-x_0) = \int_{-\infty}^{x^{**}} dx^*\, \mathbf{1}_{x^*-x_0} = [x^{**}-x_0]^+$
  - The second integral of a delta function is a ramp (ReLU, hockey stick) function.
  - The second derivative of a ramp function is a delta function; its first derivative is a step function.

2.3. Digression on Delta Functions (4)

Properties of the delta function, continued

- Infinite integral times a function:

$$\int_{-\infty}^{\infty} dx\, \delta(x-x_0)\, f(x) = \lim_{t \searrow 0} \int_{-\infty}^{\infty} dx \frac{1}{\sqrt{2\pi t}} \exp\left(-(x-x_0)^2/2t\right) f(x) = f(x_0)$$

  - This just restates the defining integration property of $\delta(x-x_0)$, also called the ***sifting*** property.
- Sum over a (weighted set) of delta functions:

$$f(x) \sum_i \delta(x-x_i) = \lim_{t \searrow 0} \sum_i f(x) \frac{1}{\sqrt{2\pi t}} \exp\left(-(x-x_i)^2/2t\right) = \sum_i f(x_i)\delta(x-x_i)$$

  - By itself, this is "small beer" – we just use the sifting property to build a "comb" consisting of the values of $f(x)$ at the chosen points $\{x_i\}$. This becomes useful when we ***integrate*** over all (continuous) x_i:
- Infinite integral over a (weighted set) of delta functions:

$$\int_{-\infty}^{\infty} dx_i\, f(x_i)\, \delta(x-x_i) = \int_{-\infty}^{\infty} dx_i\, f(x_i)\, \delta(x_i-x) = \lim_{t \searrow 0} \int_{-\infty}^{\infty} dx_i \frac{1}{\sqrt{2\pi t}} \exp\left(-(x-x_i)^2/2t\right) f(x_i) = f(x)$$

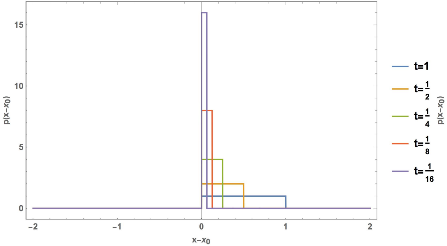
  - This is the ***convolution*** or composition/superposition property of the delta function.
  - It allows us to recover a function from its convolution with $\delta(x)$ and thereby (in a linear operator setting) construct the solution for a general payoff/boundary condition from the Green's function (fundamental solution).

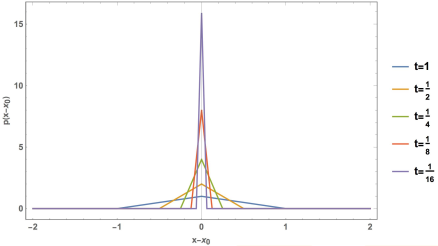
2.3. Digression on Delta Functions (5)

- While our development of the delta function has been informal and intuitive, it ***is*** possible to put it on a fully rigorous footing: Laurent Schwartz's ***Theory of Distributions*** («***Théorie des distributions***»), for which he was awarded the Fields Medal in 1950.

Alternative constructions

- "Box" method: $\delta(x, x_0) = \delta(x - x_0) \doteq \lim_{t \searrow 0} \frac{1}{t}\left(\mathbf{1}_{x-x_0} - \mathbf{1}_{x-(x_0+t)}\right)$ ("right-sided")
  - Also "left-sided" $\left(\lim_{t \searrow 0} \frac{1}{t}\left(\mathbf{1}_{x+t} - \mathbf{1}_{x}\right)\right)$ and centered $\left(\lim_{t \searrow 0} \frac{1}{t}\left(\mathbf{1}_{x+t/2} - \mathbf{1}_{x-t/2}\right)\right)$ boxes (taking $x_0 = 0$ for simplicity)





- "Triangle" method: $\lim_{t \searrow 0} \frac{1}{t}\left((x+t)\mathbf{1}_{x+t} - 2\,x\,\mathbf{1}_{x} + (x-t)\mathbf{1}_{x-t}\right)$

3. Brownian Motion Representation & Black-Scholes-Merton Formula

- Lognormal transformation. If:

$$\frac{dS}{S} = (r - y)\, dt + \sigma\, dW^{\mathbf{Q}}$$

then:

$$d\, \ln(S) = \left(r - y - \frac{\sigma^2}{2}\right) dt + \sigma\, dW^{\mathbf{Q}} = \tilde{\mu}^{\mathbf{Q}} dt + \sigma\, dW^{\mathbf{Q}}$$

with: $\tilde{\mu}^{\mathbf{Q}} \doteq r - y - \sigma^2/2$

- Hence,

$$\ln(S_T) - \ln(S_t) = \ln\left(\frac{S_T}{S_t}\right) \sim N\left[\tilde{\mu}^{\mathbf{Q}}(T-t), \sigma^2(T-t)\right]$$

and the probability density q for $\ln(S_T/S_t)$ is given by:

$$q\left[\ln\left(S_T/S_t\right), T-t\right] = \frac{1}{\sigma\sqrt{2\pi(T-t)}} \exp\left[\frac{-\left(\ln\left(S_T/S_t\right) - \tilde{\mu}^{\mathbf{Q}}(T-t)\right)^2}{2\sigma^2(T-t)}\right]$$

3. Brownian Motion Representation & Black-Scholes-Merton Formula (2)

- Alternatively, we can write:

$$\frac{S_T}{S_t} = \exp\left(\tilde{\mu}^{\mathbf{Q}}(T-t) + \sigma\sqrt{T-t}\,z\right)$$

  with $z \sim N(0,1)$

- Valuation of the payoff $C_T(S_T)$:

$$C(S_t,t) = e^{-r(T-t)} \int_{-\infty}^{\infty} \frac{dz}{\sqrt{2\pi}}\, e^{-z^2/2}\, C_T\left(S_t\, e^{\tilde{\mu}^{\mathbf{Q}}(T-t) + \sigma\sqrt{T-t}\,z}\right)$$

- If $C_T(S_T,T) = \max[S_T - K, 0]$:

$$\begin{aligned} C(S_t,t) &= e^{-r(T-t)} \int_{-\infty}^{\infty} \frac{dz}{\sqrt{2\pi}}\, e^{-z^2/2} \max\left[S_t\, e^{\tilde{\mu}^{\mathbf{Q}}(T-t) + \sigma\sqrt{T-t}\,z} - K, 0\right] \\ &= e^{-r(T-t)} \int_{z^*}^{\infty} \frac{dz}{\sqrt{2\pi}}\, e^{-z^2/2} \left[S_t\, e^{\tilde{\mu}^{\mathbf{Q}}(T-t) + \sigma\sqrt{T-t}\,z} - K\right] \\ \text{with } z^* &= \frac{\ln\left(K/S_t\right) - \tilde{\mu}^{\mathbf{Q}}(T-t)}{\sigma\sqrt{T-t}} \end{aligned}$$

3. Brownian Motion Representation & Black-Scholes-Merton Formula (3)

- Expanding terms in the integral:

$$\begin{aligned}
C(S_t,t) &= \quad e^{-r(T-t)}\int_{z^*}^{\infty}\frac{dz}{\sqrt{2\pi}}\left[S_t\, e^{\tilde{\mu}^{\mathbf{Q}}(T-t)}\, e^{-z^2/2+\sigma\sqrt{T-t}\,z} - K\, e^{-z^2/2}\right] \\
&= S_t\, e^{-y(T-t)}\int_{z^*}^{\infty}\frac{dz}{\sqrt{2\pi}}e^{-z^2/2+\sigma\sqrt{T-t}\,z-\sigma^2(T-t)/2} - K\, e^{-r(T-t)}\int_{z^*}^{\infty}\frac{dz}{\sqrt{2\pi}}\, e^{-z^2/2} \\
&= S_t\, e^{-y(T-t)}\int_{z^*}^{\infty}\frac{dz}{\sqrt{2\pi}}e^{-\left(z-\sigma\sqrt{T-t}\right)^2/2} \qquad\qquad - K\, e^{-r(T-t)}\int_{z^*}^{\infty}\frac{dz}{\sqrt{2\pi}}\, e^{-z^2/2}
\end{aligned}$$

- Using $\int_{z^*}^{\infty}\frac{dz}{\sqrt{2\pi}}\, e^{-z^2/2} = 1-\mathcal{N}[z^*] = \mathcal{N}[-z^*]$ and substituting for z^*:

$$\begin{aligned}
C(S_t,t) &= S_t\, e^{-y(T-t)}\,\mathcal{N}[-z^*+\sigma\sqrt{T-t}] - K\, e^{-r(T-t)}\,\mathcal{N}[-z^*] \\
&= S_t\, e^{-y(T-t)}\,\mathcal{N}\left[\frac{\ln(S_t/K)+\left(r-y-\frac{\sigma^2}{2}\right)(T-t)}{\sigma\sqrt{T-t}}+\sigma\sqrt{T-t}\right] \\
&\quad -K\, e^{-r(T-t)}\,\mathcal{N}\left[\frac{\ln(S_t/K)+\left(r-y-\frac{\sigma^2}{2}\right)(T-t)}{\sigma\sqrt{T-t}}\right]
\end{aligned}$$

3. Brownian Motion Representation & Black-Scholes-Merton Formula (4)

- Simplifying:

$$\begin{aligned} C(S_t,t) &= S_t\,e^{-y(T-t)}\,\mathcal{N}\left[\frac{\ln\left(S_t/K\right)+(r-y)(T-t)}{\sigma\sqrt{T-t}}+\frac{\sigma\sqrt{T-t}}{2}\right] \\ &\qquad - K\,e^{-r(T-t)}\,\mathcal{N}\left[\frac{\ln\left(S_t/K\right)+(r-y)(T-t)}{\sigma\sqrt{T-t}}-\frac{\sigma\sqrt{T-t}}{2}\right] \\ &= S_t\,e^{-y(T-t)}\,\mathcal{N}[z_+] - K\,e^{-r(T-t)}\,\mathcal{N}[z_-] \\ \text{with } z_\pm &= \frac{\ln\left(S_t/K\right)+(r-y)(T-t)}{\sigma\sqrt{T-t}}\pm\frac{\sigma\sqrt{T-t}}{2} \end{aligned}$$

- Put valuation: we must change two things:
  1. Change sign of payoff function:
     $\max[S_T-K,0]=[S_T-K]^+\to\max[K-S_T,0]=[K-S_T]^+$
  2. Change limits on integral: $\int_{z^*}^{\infty}dz\to\int_{-\infty}^{z^*}dz$
  - Result: $P(S_t,t)=K\,e^{-r(T-t)}\,\mathcal{N}[-z_-]-S_t\,e^{-y(T-t)}\,\mathcal{N}[-z_+]$

3. Brownian Motion Representation & Black-Scholes-Merton Formula (5)

- Alternatively, we can use put-call parity:

$$\begin{aligned} P(S_t, t) &= C(S_t, t) \; - S_t\, e^{-y(T-t)} + K\, e^{-r(T-t)} \\ &= S_t\, e^{-y(T-t)}\, \mathcal{N}[z_+] - K\, e^{-r(T-t)}\, \mathcal{N}[z_-] - S_t\, e^{-y(T-t)} + K\, e^{-r(T-t)} \\ &= S_t\, e^{-y(T-t)} (\mathcal{N}[z_+] - 1) \; + K\, e^{-r(T-t)} (1 - \mathcal{N}[z_-]) \\ &= K\, e^{-r(T-t)}\, \mathcal{N}[-z_-] \; - S_t\, e^{-y(T-t)}\, \mathcal{N}[-z_+] \end{aligned}$$

- Symmetric formula (Rubinstein):

$$\begin{Bmatrix} C \\ P \end{Bmatrix} (S_t, t) = \phi \left[S_t e^{-y(T-t)}\, \mathcal{N}(\phi\, z_+) - K e^{-r(T-t)}\, \mathcal{N}(\phi\, z_-) \right]$$

  with $\phi = 1$ for calls and $\phi = -1$ for puts

- Primary references:
  - Fischer Black and Myron Scholes, "The Pricing of Options and Corporate Liabilities," *Journal of Political Economy* **81**(3), (May-Jun, 1973), pp. 637-654. www.jstor.org/stable/1831029
  - Robert C. Merton, "Theory of Rational Option Pricing," *The Bell Journal of Economics and Management Science* **4**(1), (Spring, 1973), pp. 141-183. www.jstor.org/stable/3003143

3. Brownian Motion Representation & Black-Scholes-Merton Formula (6)

Options on Futures/Forwards

- FX trading is based in the forward market, so we re-write the BSM formula in terms of $F_{t,T}$:
  $F_{t,T} = S_t e^{(r-y)(T-t)} \implies S_t e^{-y(T-t)} = F_{t,T} e^{-r(T-t)}$
- Re-write Rubinstein's symmetric formula as:

$$\begin{aligned}\{C,P\}(F_{t,T},t) &= \phi\Big[F_{t,T}e^{-r(T-t)}\mathcal{N}(\phi z_+) - Ke^{-r(T-t)}\mathcal{N}(\phi z_-)\Big]\\ &= \phi\, e^{-r(T-t)}\big[F_{t,T}\,\mathcal{N}(\phi z_+) - K\,\mathcal{N}(\phi z_-)\big]\end{aligned}$$

  with $\{\phi = 1$ for calls, $\phi = -1$ for puts$\}$ and $z_\pm = \dfrac{\ln\big(F_{t,T}/K\big)}{\sigma\sqrt{T-t}} \pm \dfrac{\sigma\sqrt{T-t}}{2}$.
- Since $F_{t,T}$ converges to S_T as $t \nearrow T$, this is identical in content to the BSM formula.
  - This is theoretically a slightly preferable formulation to BSM because all inputs have the same tenor T.
  - Technically, the volatility input σ really should be that of the forward (or futures) price $F_{t,T}$.
- Extension: instead of the forward/futures contract terminating at T, let it mature at $T' > T$.
  - This is how futures options are set up in practice: $T' - T$ is typically a few days to a few months.
  - Black's Model for futures options: use the formula above with $F_{t,T'}$ in place of $F_{t,T}$

4. Inputs & Sensitivities

- Inputs
  1. Risky asset price, S
  2. Strike price, K
  3. Time to maturity, T (or $T-t$)
  4. Riskless interest rate, r
  5. (Dividend or convenience) yield, y
  6. Volatility, σ
- Sensitivities (to all of the above)

4.1. Inputs

- Risky asset (e.g. **S**tock) price: $S = S_t = S(t)$
  - Units: (quantity of numéraire asset / unit of risky asset) e.g. \$/share
  - For currencies, sometimes use X ("e**X**change rate") for clarity
    - Units: (quantity of numéraire or domestic currency / unit of foreign currency) e.g. \$/£, \$/¥, \$/€...
    - When we study currency options, we'll see that quotation conventions appear to be the opposite of this (but are actually completely consistent)
- Stri**K**e price: K
  - Units: same as risky asset
  - Some older texts use X (for e**X**ercise price)
- **T**ime to ma**T**urity: T
  - Units: time (years, months, days...).
    - N.B. conventions associated with volatility and rate inputs (trading days, day counts)
  - We may also use $T - t$ when we want to fix T at a specific point in the future and vary the current time t

4.1. Inputs (2)

- Riskless rate r:
  - Units: time^{-1} (e.g. %/year).
    - N.B. day count and settlement conventions vary between different interest bearing assets, e.g.:
      US Treasury bonds (ACT/ACT), but Treasury bills are ACT/360
      LIBOR and similar indices (usually ACT/360, but also ACT/365 and other variants), etc.

1. Academic literature traditionally references US Treasury (or, for other countries, the corresponding sovereign / central bank) domestic currency yield curve as (most) riskless
   - Closest philosophically to B-S-M assumption of riskless borrowing, but is the US government a.s. default-free? What about other countries?
   - Evidence from rating agencies, credit default swaps (CDS), history: perhaps not.
   - Traditional argument: a government can always print money to inflate its way out of debt in its own currency, so domestically-denominated debt should be riskless while foreign-denominated debt is not
   - Historical studies show that sovereigns default on own-currency debt surprisingly often
   - US Government debt crises (e.g. Oct 2013) have threatened USD treasury defaults (and markets responded demonstrating credibility of threat)
   - What about sovereigns that don't control their own currency (e.g., Euro-zone countries)?
   - For credit-unworthy sovereigns' currencies, could there be another (sovereign, financial, or corporate) counterparty that borrows at a lower (less risky) rate?

4.1. Inputs (3)

- Riskless rate r (continued): **Sovereign CDS Spreads**

Country	Rating S&P	5Y CDS▲	Var 1m	Var 6m	PD (*)	Date
Switzerland	AAA	6.01	+0.33 %	-0.66 %	0.10 %	5 Feb
Australia	AAA	10.05	-0.99 %	-23.63 %	0.17 %	5 Feb
Sweden	AAA	10.22	+4.39 %	-16.30 %	0.17 %	5 Feb
Denmark	AAA	10.36	+18.13 %	+7.14 %	0.17 %	5 Feb
Netherlands	AAA	12.58	+6.70 %	+21.66 %	0.21 %	5 Feb
Germany	AAA	13.52	-2.73 %	+56.30 %	0.23 %	5 Feb
Austria	AA+	15.41	+3.42 %	+16.48 %	0.26 %	5 Feb
Ireland	AA	16.36	+1.05 %	-19.57 %	0.27 %	5 Feb
Japan	A+	19.62	-3.11 %	-19.03 %	0.33 %	5 Feb
Finland	AA+	19.93	+1.94 %	+3.26 %	0.33 %	5 Feb
United Kingdom	AA	23.17	+9.14 %	+1.85 %	0.39 %	5 Feb
Belgium	AA	25.86	-1.45 %	+34.41 %	0.43 %	5 Feb
Portugal	A-	27.09	-3.80 %	-29.03 %	0.45 %	5 Feb
United States	AA+	31.47	+6.03 %	-1.41 %	0.52 %	5 Feb
Spain	A-	33.76	-4.42 %	-9.56 %	0.56 %	5 Feb
South Korea	AA	34.11	-10.52 %	-4.35 %	0.57 %	5 Feb
France	AA-	36.45	-8.69 %	+17.69 %	0.61 %	5 Feb
Canada	AAA	39.60	0.00 %	0.00 %	0.66 %	5 Feb
China	A+	55.74	-12.77 %	-18.70 %	0.93 %	5 Feb
Italy	BBB	57.27	-5.81 %	-16.11 %	0.95 %	5 Feb
Greece	BBB-	57.78	-3.35 %	-14.83 %	0.96 %	5 Feb
Indonesia	BBB	76.67	-2.16 %	-4.22 %	1.28 %	5 Feb
India	BBB-	84.08	0.00 %	0.00 %	1.40 %	5 Feb
Israel	A	87.08	-3.30 %	-30.24 %	1.45 %	5 Feb
Mexico	BBB	130.67	-4.99 %	+15.81 %	2.18 %	5 Feb
Brazil	BB	177.28	-15.16 %	+8.83 %	2.95 %	5 Feb
South Africa	BB-	200.19	+5.27 %	+1.62 %	3.34 %	5 Feb
Turkey	BB-	255.13	-1.02 %	-2.47 %	4.25 %	5 Feb
Egypt	B-	521.34	-7.11 %	-31.11 %	8.69 %	5 Feb
Russia	NR	13775.17	0.00 %	0.00 %	100.00 %	5 Feb

Columns Var 1m through Date: 5 Years Credit Default Swaps

(*) Implied probability of default, calculated on the hypothesis of a 40% recovery rate.

- Source: www.worldgovernmentbonds.com/sovereign-cds/ as of 05 Feb 2024

4.1. Inputs (4)

- Riskless rate r (continued):
  2. Up to the financial crisis, the over-the-counter (OTC) market traditionally used LIBOR or some other, usually ***unsecured***, inter-bank lending rate as the "riskless" interest rate
     - Always questionable, but during the "great moderation" average rating of LIBOR-poll banks was typically a strong AA and TED spreads were in the tens of basis points
     - Some oddities for firms with stronger (or weaker) credit; artifacts in pricing vs. Treasuries
     - Systematic downgrades of banking sector before and during the crisis
     - Explosion of TED spread during the crisis and appearance of significant basis spreads between e.g. different tenors reflecting perceptions of credit (un-)worthiness under stress
     - LIBOR scandals calling into question whether poll actually reflects inter-bank lending rates
  3. Post-crisis: evolution to OIS (Overnight Index Swap) market as benchmark "riskless" rate
     - Based on e.g. (floating) Fed Funds overnight lending rate $\Rightarrow$ like a money-market account
     - Rolling daily settlement and amount at risk equal to (difference between fixed and floating rates) limits credit exposure

4.1. Inputs (5)

- Riskless rate r (continued):
  4. Current state of play:
     - Central bank-calculated reference rates based on either:
       - (secured) overnight (treasury) repo market (e.g.: SOFR for USD, SARON for CHF)
       - (unsecured) overnight interbank market (e.g.: SONIA for GBP, EONIA/€STR for EUR, TONR for JPY, ...)
     - Reference swap rates published by exchanges (e.g., ICE/NYSE: www.theice.com/iba/ice-swap-rate)
     - As of end-2021, new swap issuance based on LIBOR was "strongly discouraged", e.g., US:
       www.federalreserve.gov/supervisionreg/srletters/SR2117a1.pdf
       and even outright banned in some jurisdictions, e.g., UK:
       www.fca.org.uk/publication/libor-notices/article-21a-benchmarks-regulation-prohibition-notice.pdf
     - Publication of LIBOR has been phased out:
       - The final (USD) LIBOR panel ceased to exist as of end-June 2023.
       - A few "synthetic" GBP & USD fall-back rates (SOFR or SONIA + a spread) were updated until Mar/Sep 2024.
       www.fca.org.uk/news/news-stories/fca-announces-decision-synthetic-us-dollar-libor
       www.theice.com/iba/libor
     - Eurodollar futures and options were transitioned/converted to SOFR in April 2023 and delisted in June 2023.
     - Lots of issues (e.g., backward-looking rates), mostly beyond our scope! See:
       BIS paper: *Beyond LIBOR: a primer on the new reference rates*
       Gregory S. Faranello, CFA commentary: *The Path to SOFR and Alternatives*
       NY Fed Alternative Reference Rates Committee: *Transition from LIBOR*
       J.P. Morgan Solutions: *Leaving LIBOR: A Landmark Transition*

4.1. Inputs (6)

- Riskless rate r (continued):
  - Although it wreaks havoc on the theory, a "multi-curve" framework embedding:
    1. a (riskless) money-market-like rate to generate forward prices and
    2. one or more risky rates, perhaps with a benchmark analogous to LIBOR, for discounting

    is becoming standard.
    - The most sophisticated broker-dealers were already moving in this direction before the crisis
    - Accounting standards: Counterparty credit (CVA) and own-credit (DVA) adjustments
    - Active area of research, but results are very messy because discounting rates depend on:

      - collateralization mechanics,
      - correlation between payoff and credit spread, and even
      - entire structures / risk profiles of your and counterparty's books
        (and full sets of counterparty relationships)
  - And then there is the fact that interest rates are actually stochastic...

4.1. Inputs (7)

- yield y (q in some notations):
  - Units: time^{-1} (e.g. %/year).
  - Various interpretations:
    - Dividend yield (also denoted d or δ) for stocks & indices, but also stock borrow/lending rate;
    - Foreign interest rate r_f for currencies;
    - Convenience yield and/or storage rate for commodities and real assets.
  - Stock/index dividend yields:
    - Strong evidence (at least for index aggregates) that dividend ***amounts*** D and asset prices S are co-integrated (M. Lettau and S. Ludvigson, "Expected Returns and Expected Dividend Growth," *JFE* **76**(3), 2005, pp. 583-626) $\Rightarrow$ assumption of fixed dividend yields appears a reasonable place to start, at least for long-maturity options;
    - For single stocks, can get to this assumption by assuming that (in the long run) companies seek to pay out a fixed fraction of their earnings;
    - Most financial firms maintain a dividend database and many analysts attempt to forecast dividends for at least the next year or two.
    - Problem 1: Dividends are paid ***discretely*** (are lumpy) and aren't distributed evenly throughout the year $\Rightarrow$ sometimes need to model $d \sim \sum_i d_i \delta(t - t_i)$
    - Problem 2: In the short term, dividend ***amounts*** D are sticky (signaling, declaration, accruals, etc).
    - Traders often conflate these two issues, but they are both of real importance for both option valuation ***and*** hedging (e.g. UK's 1997 repeal of dividend withholding exemption for broker-dealers)
    - We briefly glanced at some of these issues in Part 1 of the course notes and the first problem set.

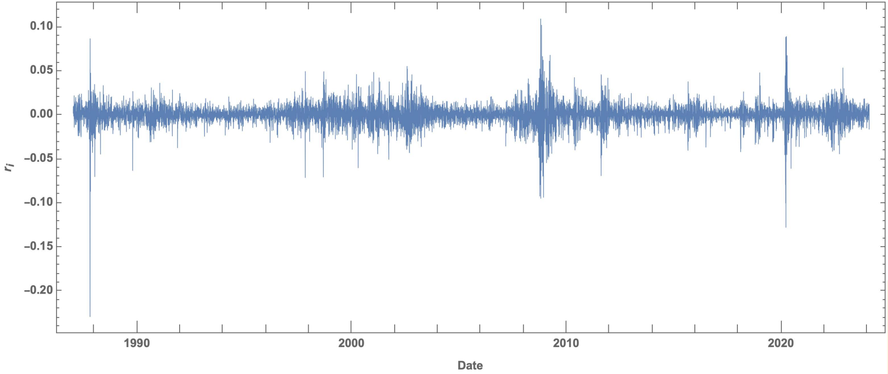
4.1. Inputs (8)

- Volatility σ:
  - Units: time$^{-1/2}$ (e.g. %/year$^{1/2}$). Why?
    Because ***variance*** σ^2 is a ***rate*** and therefore scales as time^{-1}
- σ represents the characteristic size or scale of asset price movements over a period of time:
  - $\ln(S_{t+\Delta t}/S_t) \sim \sigma\sqrt{\Delta t}$
  - Traders' rule: daily vol $\approx$ (annual vol)/16,
    e.g., annual volatility of 16% corresponds approximately to a daily move of $\pm 1\%$
- What does the word "volatility" mean in the context of option pricing and risk management?
  Many (inter-related) answers:
  1. Historical / Realized vol
  2. Forecast vol
  3. Implied vol

4.1. Inputs (9)

Volatility Properties

- SPX daily (log) returns,1987-present:



- Asset price process volatility (in the **P**-world) isn't directly observable
  $\Longrightarrow$ volatility must be *estimated* or *forecast*
- Short-term persistence: volatility *"clusters"* then displays mean reversion-like behavior
- Is volatility a stationary process? If so, over what time horizon?

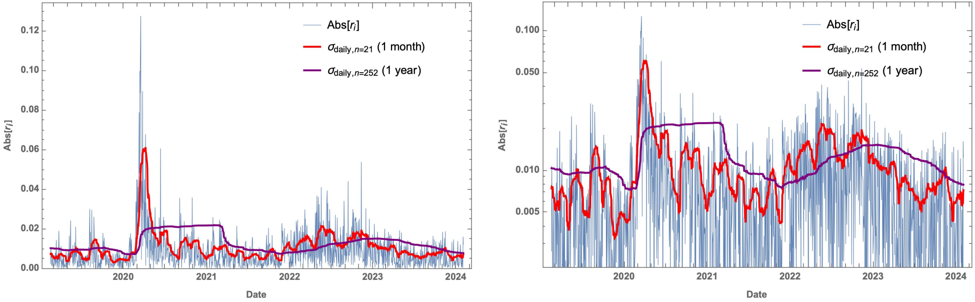
4.1. Inputs (10)

1 Historical / Realized Volatility

- Looking backward, calculate estimator of historical vol (over some interval, with some frequency) from e.g. the root-mean-square formula:

$$\sigma^2_{daily} \sim \sigma^2_{daily;n} \doteq \frac{1}{n-1}\sum_{i=0}^{n-1}(r_{-i} - \langle r \rangle_{-n})^2 \text{ with } r_{-i} = \ln[S_{-i}/S_{-(i+1)}] \text{ and } \langle r \rangle_{-n} = \frac{1}{n}\sum_{i=0}^{n-1} r_{-i}$$

- Depending on perspective, estimator is either unconditional or conditional only on the sampled time range



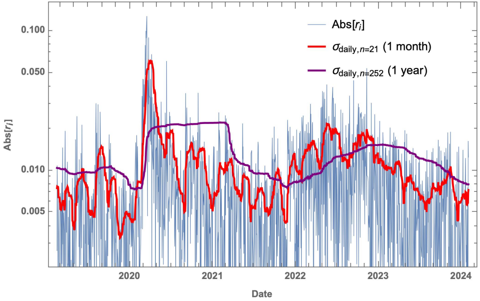
- Trade-off: shorter vs. longer time series of returns
  - Longer time series: better confidence in estimator but may be less responsive to changes, less relevant to future
  - Shorter time series: responsive and more current, but poorer confidence and more volatile
- Possible solutions/extensions: use more than just close-close returns (OHLC, hi-frequency data)

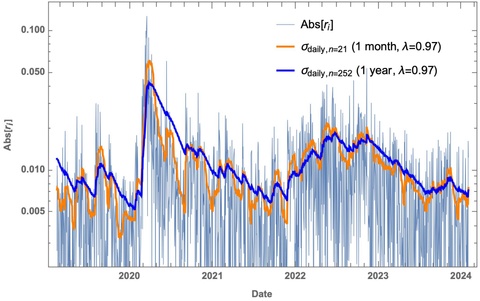
4.1. Inputs (11)

1 Historical / Realized Volatility (continued)

- What we really want is an estimator conditional on a particular (future) time or range of times.
  - Looking forward, seek to predict the future realized volatility (driving the cost of dynamically hedging an option)
  - Backward-looking estimator isn't necessarily a good predictor of future volatility, particularly around times of stress...
- Perhaps weight returns (e.g., exponential decay) so that more recent price moves (with presumably greater value in estimating future volatility) are weighted more heavily:

$$\sigma^2_{daily} \sim \sigma^2_{daily;n,\lambda} \doteq \frac{1-\lambda^2}{2(\lambda-\lambda^n)} \sum_{i=0}^{n-1} \lambda^i (r_{-i} - \langle r \rangle_{-n,\lambda})^2 \text{ with } \langle r \rangle_{-n,\lambda} = \frac{1-\lambda}{1-\lambda^n} \sum_{i=0}^{n-1} \lambda^i r_i;\ n_{eff} = \frac{(1+\lambda)(1-\lambda^n)}{(1-\lambda)(1+\lambda^n)}$$



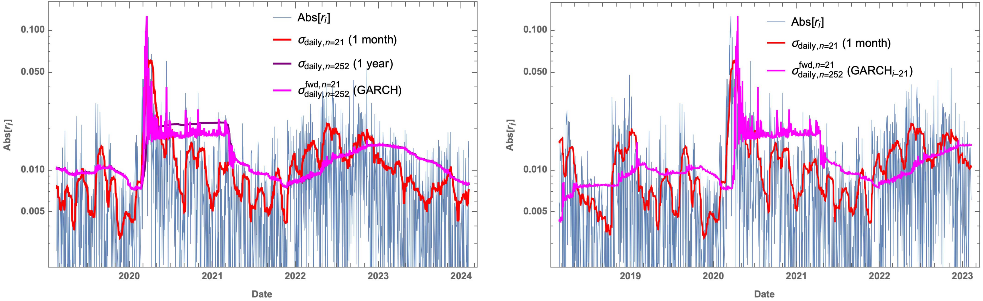


EWMA vol estimator or "poor man's GARCH" (above: $\lambda = 0.97$, $n_{eff} = \{20.3, 65.6\}$)

4.1. Inputs (12)

2 Forecast Volatility

- Try to model stylized properties of volatility (persistence/mean reversion) as conditional heteroskedasticity: E.g., ARCH(1): $r_i = \sigma_i z_i$; $\sigma_i^2 = k + \alpha_1 r_{i-1}^2$; GARCH(1,1): $r_i = \sigma_i z_i$; $\sigma_i^2 = k + \alpha_1 r_{i-1}^2 + \beta_1 \sigma_{i-1}^2$



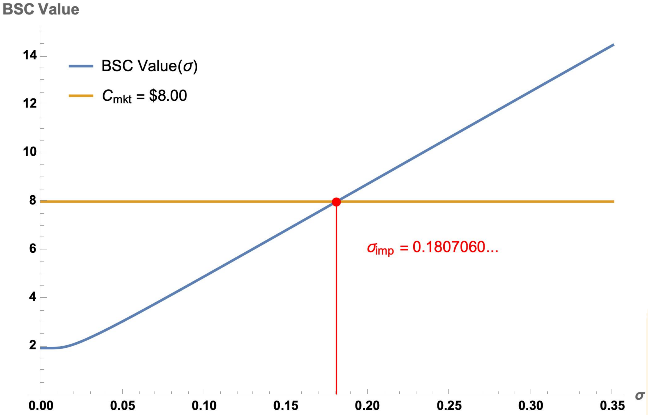
- Empirically has *some* predictive value, particularly for shorter time horizons, particularly using high-frequency data as input (but recognizing that some issues arise at very high frequencies)
- Popular in academic circles, but never really developed much industry traction, at least on the sell side.
- Should we use $\mathbb{E}[\sigma_T]$, $\mathbb{E}[\sigma_T^2]$, or, more generally, some form of term structure involving $\sigma_t : 0 \le t \le T$, e.g., $\mathbb{E}[\int_0^T dt\, \sigma_t^2]$ as input to our pricing model?
  - Problems of convexity of option prices vs. σ and correlation of σ with S
  - It would be better to consider this in a stochastic volatility pricing framework

4.1. Inputs (13)

3 Implied Volatility

Definition of implied vol: $\sigma_{imp} \doteq \boldsymbol{\sigma} : C(S, K, T; r, y, \boldsymbol{\sigma}) = C_{mkt}$

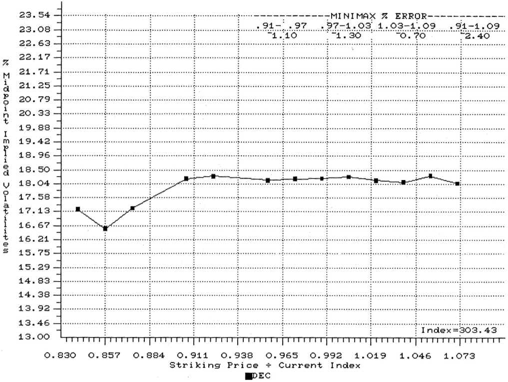
- Example: "Base Case" parameters: $S = K = 100$, $T - t = 1$ yr, $r = 4\%$ p.a., $y = 2\%$ p.a., with $C_{mkt} = 8$:

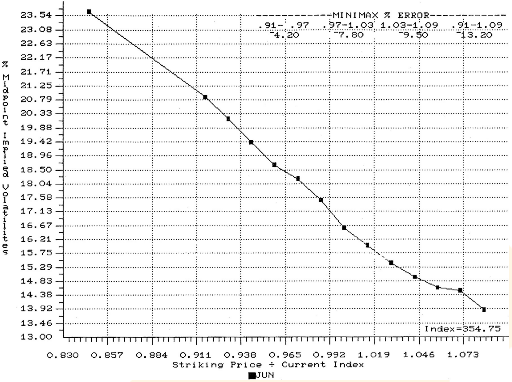


- In practice, this is a non-linear (numerical) root-finding problem, requiring a good initial guess for σ_{imp}.

4.1. Inputs (14)

3. **Implied Volatility** (continued):
   - For index options, significant volatility "skew" emerged rapidly following the 1987 market crash.
     SPX σ_{imp} vs. K/S: Dec '87 expiry as of 1-Jul-87 vs. Jun '90 expiry as of 2-Jan-90 (both $T \approx 5$ months)



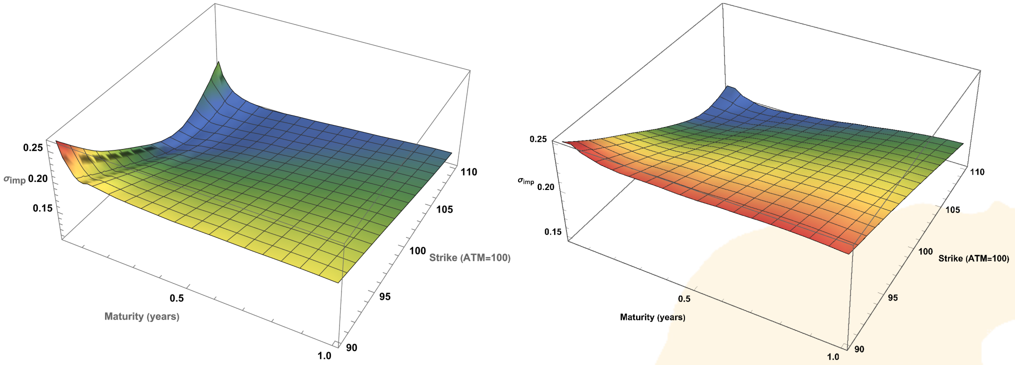


Source: Mark Rubinstein, *Journal of Finance* **49**(3), (Jul 1994), pp. 771-818.

   - In general, $\sigma_{imp} = \sigma_{imp}(T, K)$.
     In addition, σ_{imp} changes (with unpredictable components) as $\{t, S_t\}$ evolve

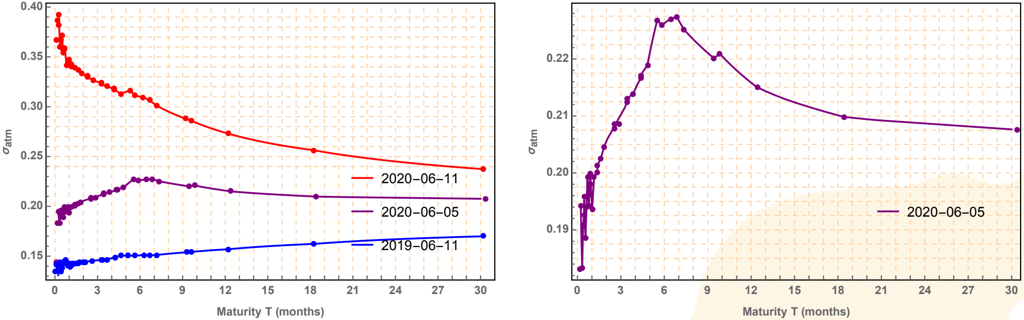
4.1. Inputs (15)

3. **Implied Volatility** (continued):
   - SPX Implied Volatility Surfaces as of 2025-01-17 and 2023-01-20



4.1. Inputs (16)

3. **Implied Volatility** (continued):
   - SPX ATM Implied Volatility Curves as of 2019-06-11, 2020-06-05, and 2020-06-11:

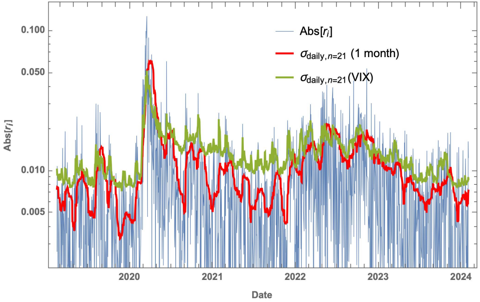


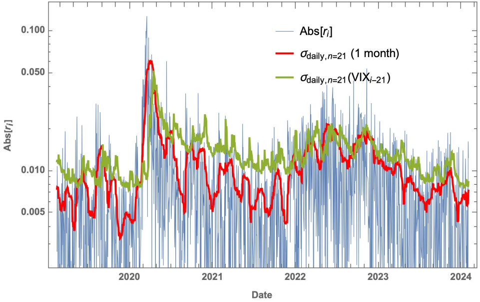
4.1. Inputs (17)

3. **Implied Volatility** (continued):
   - Standard aphorism: σ_{imp} is the market's forecast of future asset price (realized) volatility
   - How should we trade if we want to take a position that the "market's forecast" is too high or too low?
   - If so, *which* implied volatility/instruments?
     - $\sigma_{imp,ATM}$ via, e.g., ATM straddles (where liquidity is usually concentrated)?
     - We've already raised some of the issues with this approach when we discussed straddles in Part 1.
   - Is there, in some sense, an *unconditional* market volatility forecast that doesn't depend on assumptions like remaining at/near the money?
     - Remarkably, the answer is YES! Start with our log-process: $d\ln(S_t) = (r-y-\sigma^2/2)dt + \sigma dW_t^{\mathbf{Q}}$
     - Drift is inconvenient, but is neutralized if we consider the process for $F_{t,T} = e^{(r-y)(T-t)}S_t$:
     $$d\ln(F_{t,T}) = -\sigma_t^2/2\,dt + \sigma_t dW_t^{\mathbf{Q}}$$
     - Integrating and taking expectations:
     $$\ln\left(\frac{F_{T,T}}{F_{0,T}}\right) = -\frac{1}{2}\int_0^T dt\,\sigma_t^2 + \int_0^T \sigma_t\,dW_t^{\mathbf{Q}} \implies \mathbb{E}^{\mathbf{Q}}\left[\ln\left(\frac{S_T}{F_{0,T}}\right)\right] = -\frac{1}{2}\mathbb{E}^{\mathbf{Q}}\left[\int_0^T dt\,\sigma_t^2\right] = -\frac{1}{2}\int_0^T dt\,\mathbb{E}^{\mathbf{Q}}\left[\sigma_t^2\right]$$
   - Hence: $\mathbb{E}^{\mathbf{Q}}\left[\int_0^T dt\,\sigma_t^2\right] = -2\,\mathbb{E}^{\mathbf{Q}}\left[\ln\left(\frac{S_T}{F_{0,T}}\right)\right]$ ... a remarkable result!
     - The LHS is the market's expectation of the total variance to be experienced between now and T. It forms the basis of *realized variance* contracts and *variance swaps*.
     - The RHS contains the $\mathbf{Q}$-expectation of the payoff of a *log contract*. We will see how to replicate (and value) such payoffs in the next set of notes.
     - Together, these form the basis for calculating the VIX index, corresponding to the square-root of the LHS expectation over the next calendar month, times an annualization factor.

4.1. Inputs (18)

3. **Implied Volatility** (continued):
    - VIX calculated & published as of 1990, originally based on ATM implied vols for the S&P 100 (OEX)
        - Updated to the variance contract methodology applied to the SPX in 2003
        - Subsequently, a plethora of listed VIX futures and options have appeared





- Is σ_{imp} actually predictive of future realized vol?
- Does it contain information beyond that contained in measures of historical vol?
- Evidence is equivocal.
    - Early work (Canina & Figlewski, 1993), using ATM vols, says not really
    - Later efforts (e.g., Jiang & Tian, 2005), using measures akin to the VIX, suggest yes, albeit with positive bias.

4.1. Inputs (19)

3. **Implied Volatility** (continued):
   - OTC markets operate almost entirely on the basis of σ_{imp}.
   - Although reference is made to historical and forecast volatilities, and traders (in a proprietary book) may lean one way or another if implied vols are out of line with historical vols, option books are marked (to market) at implied vols.
   - Especially for short maturities, split in roles between:
     - providers of liquidity in volatility/vanilla options (e.g. "locals") and
     - consumers of volatility (vanilla options) to produce more complex products (e.g. broker-dealers)
   - Notion of volatility as building block in manufacture of structured, OTC products
   - (Sign of) spread between realized and implied vols typically reflects who's providing the liquidity in the underlying volatility (vanilla options) market
   - Examples of reversal:
     - Japanese insurance firms selling (naked) Nikkei puts to investment banks in the late 1980s;
     - retail single-stock structured bond market in Europe incorporating embedded (short) puts (GOAL = Geld oder Aktien Lieferung), 2000s - present.

4.1. Inputs (20)

Measurable vs. forecast vs. implied inputs

- The use of implied volatility (and our inability to perfectly forecast future realized volatility) introduces some nuance into our notions of real-world ($\mathbf{P}$) and risk-neutral ($\mathbf{Q}$) measures.
  - Absolute vs. relative pricing: fair (forecast) value of dynamically replicating an option vs. pricing one option relative to another;
  - Arguably, reflects market incompleteness or inability to perfectly hedge $\Rightarrow$ introduction of vanilla options (helps) complete the market
  - Pricing in implied vol space usually involves some form of assumption analogous to having identified a unique EMM (equivalent martingale measure)
  - More subtly, the use of implied vols to parameterize the valuation of other (more complex) options also implies the use of the options underlying those parameters to hedge the manufactured book against changes in implied vols.
  - Does implied vol surface imply completeness of the market? Are there any unhedged risks?
- To some extent, the same notions hold true for interest rate and dividend inputs:
  - "Stripping" a series of swap rates and/or bond values to create a "riskless" yield curve of (actually stochastic) rates
  - Inferring dividends/yields from futures prices
  - To "lock in" (or hedge) rates and yields, what do we need to do?
  - Are there any residual risks?

4.1. Inputs (21)

Classification of Parameter Dependence

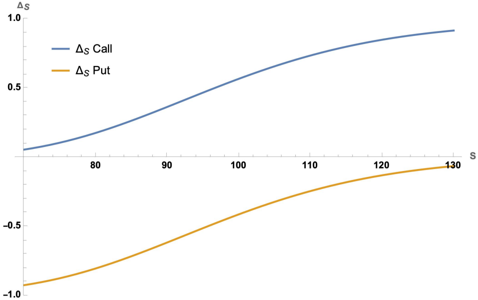
- Black-Scholes- (Merton-) world BS(M):
  - constant rates r, yields y, and volatilities σ
  - Technically, this is a bit of a misnomer since Merton '73 already provided for σ and r as functions of time (and a whole lot more).
    - Merton, R. C., "A Rational Theory of Option Pricing," *Bell Journal of Economics and Management Science* **4**(1) (Spring, 1973), 141-183.
- BS(M)+:
  - $r(t)$, $y(t)$, and $\sigma(t)$ are known, explicit functions of time $t \leq T$ ($\mathbf{P}$ measure)
  - Equivalently, ***implied*** values $r(t,T)$, $y(t,T)$, $\sigma(t,T)$ are functions of (maturity) time T observed at time t ($\mathbf{Q}$ measure)
  - We will spend much of the rest of the course in this world.
- BS(M)++:
  - $r(t,\ldots)$, $y(t,\ldots)$, and $\sigma(t,\ldots)$ are potentially stochastic functions of time t and may depend on S_t, its path, or more generally on the realization of the filtration ($\mathbf{P}$ measure)
  - Alternatively, ***implied*** values of r, y, and in particular $\sigma(t,T;S_t,K)$ are functions of (maturity) time T and strike K observed at time t and spot price S_t ($\mathbf{Q}$ measure)

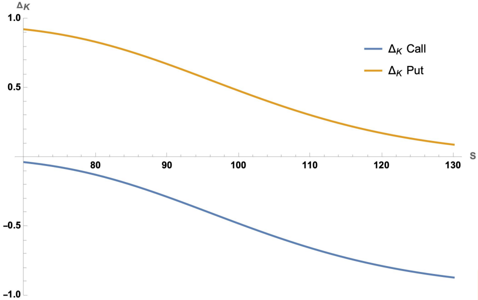
4.2. Sensitivities

1. First-order in risky asset price, S : Δ_S, or Strike price, K : Δ_K
2. First-order in riskless interest rate, r : ρ or ρ_r (or ρ_B)
3. First-order in (dividend or convenience) yield, y : ρ_y or ρ_S
4. Second-order in risky asset price, S : Γ_{SS}, or Strike price, K : Γ_{KK}
5. First-order in calendar time t : Θ, or time to maturity, T
6. First-order in volatility, σ : $\mathcal{V}$
7. Other second (and higher-order) sensitivities...
8. Base-case scenario: $S = K = 100$, $T - t = 1$ yr, $r = 4\%$ p.a., $y = 2\%$ p.a., $\sigma = 20\%$ p.a.
9. (Usually) vary S for calls and puts, holding other scenario parameters constant, to examine moneyness effects.

4.2. Sensitivities (2)

- **Delta**: $\Delta = \Delta_S \doteq \partial\{C,P\}/\partial S$ and "**Dollar Delta**" $= S\Delta_S$ (dropping t for brevity):
  $\Delta_S = \partial_S\{\phi\left[Se^{-yT}\mathcal{N}(\phi z_+) - Ke^{-rT}\mathcal{N}(\phi z_-)\right]\} = \phi\, e^{-yT}\mathcal{N}(\phi z_+)$



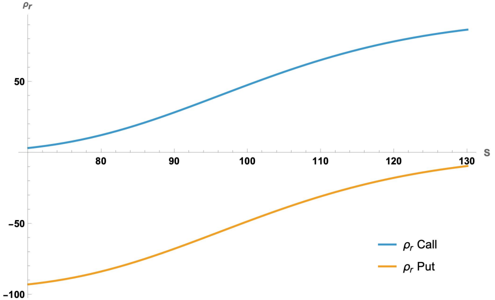


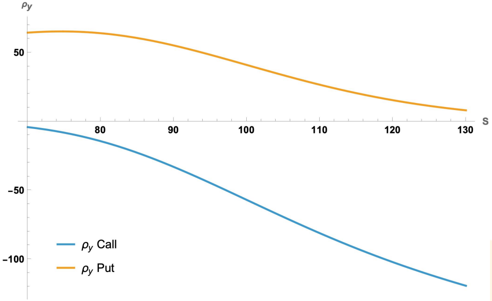
- Also, "**Strike Delta**" $= \Delta_K \doteq \partial\{C,P\}/\partial K$ and "**Cash Delta**" $\Delta_\$ = K\Delta_K$:
  $\Delta_K = \partial_K\{\phi\left[Se^{-yT}\mathcal{N}(\phi z_+) - Ke^{-rT}\mathcal{N}(\phi z_-)\right]\} = -\phi\, e^{-rT}\mathcal{N}(\phi z_-)$
- Equity spot convention ($\Delta = \Delta_S \doteq \partial_S$) vs. FX forward convention: $\Delta = \Delta_F \doteq \partial_{[e^{-r(T-t)}F_{t,T}]}$
- Note that $S\Delta_S + K\Delta_K = S\Delta_S + \$\Delta_\$ = \{C,P\}$
  - Suggests vector notation $\underline{\mathbf{S}}$ for assets, gradient notation $\nabla_{\underline{\mathbf{S}}}$ for Deltas, dot product for value: $C = \underline{\mathbf{S}} \cdot \nabla_{\underline{\mathbf{S}}} C$

4.2. Sensitivities (3)

- Ordinarily, one proceeds from Δ_S to Γ_{SS}, but there is a close relationship between deltas and rhos.
- **Interest rate rho**:

$$\rho = \rho_r \text{ (or } \rho_B) \doteq \frac{\partial\{C,P\}}{\partial r} = \partial_r\left\{\phi\left[Se^{-yT}\mathcal{N}(\phi z_+) - Ke^{-rT}\mathcal{N}(\phi z_-)\right]\right\} = \phi\, K\, T\, e^{-rT}\mathcal{N}(\phi z_-)$$





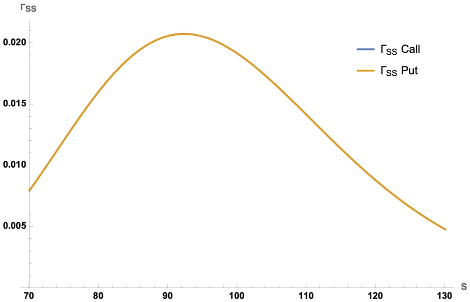
- **Yield rho** (formerly sometimes nu = ν):

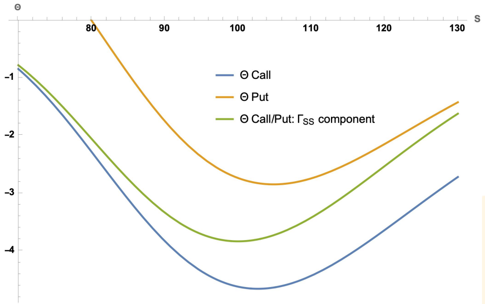
$$\rho_y \text{ (or } \rho_S) \doteq \frac{\partial\{C,P\}}{\partial y} = \partial_y\left\{\phi\left[Se^{-yT}\mathcal{N}(\phi z_+) - Ke^{-rT}\mathcal{N}(\phi z_-)\right]\right\} = -\phi\, S\, T\, e^{-yT}\mathcal{N}(\phi z_+)$$

- Alternatively, if we define rate/yield vector $\underline{\mathbf{r}}$ with, e.g., $r_0 = r$ and $r_1 = y$ (corresponding to assets in $\underline{\mathbf{S}}$), then rho vector can be defined as: $\underline{\rho} \doteq \nabla_{\underline{\mathbf{r}}}\{C,P\}$ with $\rho_i \doteq \partial\{C,P\}/\partial r_i$
- Some firms use scaled rho, e.g. $0.01\frac{\partial\{C,P\}}{\partial r}$ or $0.0001\frac{\partial\{C,P\}}{\partial r}$. Conventions vary...

4.2. Sensitivities (4)

- **Gamma**: $\Gamma = \Gamma_{SS} \doteq \frac{\partial^2\{C,P\}}{\partial S^2} = \partial_S\left\{\phi\, e^{-yT}\mathcal{N}(\phi\, z_+)\right\} = \frac{e^{-yT}}{\sigma\sqrt{T}\,S}\,n(z_+)$
  - Sometimes there are "Dollar Gammas" $= S\,\Gamma_{SS}$ or $0.01S\,\Gamma_{SS}$ (why?). Conventions vary...
  - Similarly, "**Cross Gammas**" $\Gamma_{SK} \doteq \frac{\partial^2\{C,P\}}{\partial S \partial K} = \Gamma_{KS} \doteq \frac{\partial^2\{C,P\}}{\partial K \partial S}$ and $\Gamma_{KK} \doteq \frac{\partial^2\{C,P\}}{\partial K^2}$
    - This framework is most useful in a multi-asset setting. Nevertheless, we will identify an interpretation for Γ_{KK}!

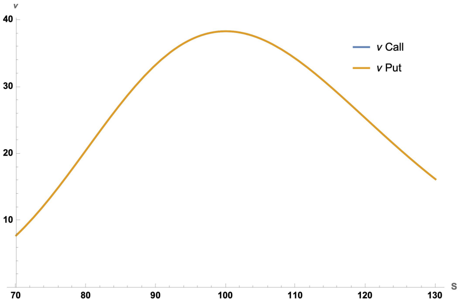


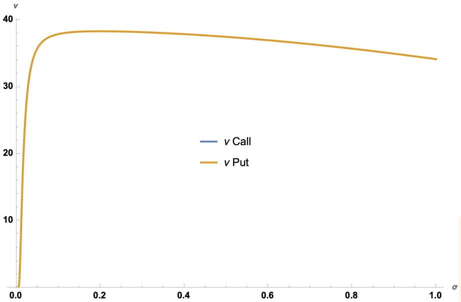


- **Theta**: θ (or Θ) $\doteq \frac{\partial\{C,P\}}{\partial t} = rK\Delta_K + yS\Delta_S - \frac{\sigma^2}{2}S^2\Gamma_{SS}$
  - Time decay: the trader's obsession
    - Similar to (but not necessarily the same as) (-) maturity sensitivity $-\frac{\partial\{C,P\}}{\partial T}$ (why?)
    - Sometimes normalized to passage of one (trading) day.

4.2. Sensitivities (5)

- **Vega**:
$$\mathcal{v} = \nu \doteq \frac{\partial\{C,P\}}{\partial\sigma} = \partial_\sigma\left\{\phi\left[Se^{-yT}\mathcal{N}(\phi z_+) - Ke^{-rT}\mathcal{N}(\phi z_-)\right]\right\} = S\sqrt{T}e^{-yT}n(z_+) = K\sqrt{T}e^{-rT}n(z_-)$$





  - Other terminology: tau (τ) or kappa (κ)
  - Some firms use scaled vega, e.g. $0.01\frac{\partial\{C,P\}}{\partial\sigma}$ (why?). Conventions vary...

4.2. Sensitivities (6)

From: MARK RUBINSTEIN
To: Eric Reiner
Sent: Saturday, June 17, 2017 2:42 PM
Subject: Vega

Nothing that creative here. Delta, Gamma, Beta were taken. Needed to use a Greek letter, and Vega was free. I don't think Vega is actually Greek but is sounds as if it should be. And being the name of a bright star, it had that going for it.

Fischer Black, best to my knowledge, gets full credit for delta and gamma. Perhaps I get credit for vega. But I wouldn't assume that was true.

Mark Rubinstein

- Other second (& higher) order sensitivities: define as needed, e.g.:
  - vanna $\doteq \dfrac{\partial \Delta}{\partial \sigma} = \dfrac{\partial^2 \{C,P\}}{\partial S \partial \sigma} = \dfrac{\partial^2 \{C,P\}}{\partial \sigma \partial S} = \dfrac{\partial v}{\partial S}$
  - volga $=$ vomma $\doteq \dfrac{\partial v}{\partial \sigma} = \dfrac{\partial^2 \{C,P\}}{\partial \sigma^2}$