

Statistics of Financial Markets

Chapter 4: Probability



Daniel Traian PELE

Bucharest University of Economic Studies
IDA Institute Digital Assets
Blockchain Research Center
AI4EFin Artificial Intelligence for Energy Finance
Romanian Academy, Institute for Economic Forecasting
MSCA Digital Finance

Acronyms used in this material

- **AI**: Artificial Intelligence
- **AR**: AutoRegressive (model)
- **ARCH**: AutoRegressive Conditional Heteroskedasticity
- **BET**: Bucharest Exchange Trading (index)
- **CDF**: Cumulative Distribution Function
- **CLT**: Central Limit Theorem
- **CRR**: Cox–Ross–Rubinstein (binomial model)
- **GARCH**: Generalised AutoRegressive Conditional Heteroskedasticity
- **GBM**: Geometric Brownian Motion
- **IBM**: International Business Machines (company)
- **LCG**: Linear Congruential Generator
- **LLN**: Law of Large Numbers
- **PCG64**: Permuted Congruential Generator, 64-bit (NumPy default)
- **PMF**: Probability Mass Function
- **RANDU**: RANDU (IBM random number generator of the 1960s)
- **SFE**: Statistics of Financial Markets, Exercises (Quantlet collection of the textbook)

Today's Question, Route and Learning Outcomes

- **Question**: which probability tools do we need to describe returns and to simulate prices?
  - every later chapter (tails, volatility, VaR, efficiency) is built on them
- **Route and outcomes**: after this chapter you can
  - compute expectations, variances, covariances and correlations, and the volatility of a portfolio
  - explain why uncorrelated returns can still be dependent, and measure it
  - use conditional expectation and conditional variance, the building blocks of GARCH
  - generate random numbers with the inverse transform and estimate a probability by Monte Carlo, with its standard error
  - define and simulate the binomial model, a random walk, a martingale, white noise, AR(1), the Wiener process and GBM
- Real data, 2000–2026: S&P 500, DAX, BET and Bitcoin (from 2014)

Reading and Tools

- Textbook: Franke, Härdle & Hafner (2019), *Statistics of Financial Markets*, 5th ed., Ch. 3 (Sec. 3.1–3.5) and Ch. 4 (discrete-time stochastic processes)
  - exercises with solutions: Borak, Härdle & López-Cabrera (2013), Ch. 3–4
- Monte Carlo methods in finance: Glasserman (2003)
- Python Quantlets of this chapter: Quantlets/Ch_04
  - ported from the SFE Quantlets of the textbook (SFErandu, SFErangen1, SFErangen2, SFEBinomp, SFEWienerProcess, SFEsimGBM)
  - each chart in these slides links to the Quantlet that draws it
- Lecture notebook: open in Google Colab
- Video course: Statistics of Financial Markets

1654: the Problem of Points

- Two players stop a fair game before the end: how should the stake be split?
- Blaise Pascal and Pierre de Fermat solved it in letters (1654): split by the probability of winning, not by the score
- Example: A needs 1 more win, B needs 2; at most two more rounds
  - B wins only if B wins both rounds: probability $1/4$
  - A should get 75% of the stake: its **expected** gain
- The same idea prices a derivative today: a fair price is an expected payoff



Blaise Pascal (1623–1662)

Portrait: unknown author, after François II Quesnel (c. 1690); public domain; Wikimedia Commons

Kolmogorov's Axioms (1933)

- A **probability space** $(\Omega, \mathcal{F}, P)$ Kolmogoroff (1933)
  - Ω: all possible outcomes (all possible paths of tomorrow's market)
  - $\mathcal{F}$: the events we can assign a probability to (a σ-algebra)
  - P: a function from events to $[0, 1]$
- Axioms, for any event $A \in \mathcal{F}$
  - $P(A) \geq 0$: no probability is negative
  - $P(\Omega) = 1$: some outcome certainly occurs
  - $P(\cup_i A_i) = \sum_i P(A_i)$ for disjoint $A_1, A_2, \ldots$ (no two can occur together): the probability of "one of them" is the sum
- Everything else (expectation, independence, processes) is built on these three rules


ERGEBNISSE DER MATHEMATIK
UND IHRER GRENZGEBIETE
HERAUSGEGEBEN VON DER SCHRIFTLEITUNG
DES
„ZENTRALBLATT FÜR MATHEMATIK"
ZWEITER BAND

GRUNDBEGRIFFE DER
WAHRSCHEINLICHKEITS-
RECHNUNG
VON
A. KOLMOGOROFF

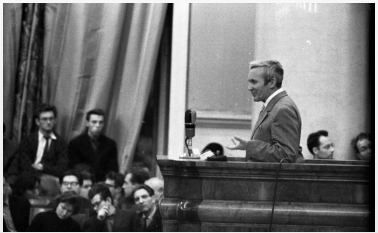
BERLIN
VERLAG VON JULIUS SPRINGER
1933

Title page of the *Grundbegriffe*, Springer 1933

Photo: KurtSchwitters (2012); CC BY-SA 3.0; Wikimedia Commons

Andrey Kolmogorov

- Andrey Nikolaevich Kolmogorov (1903–1987), Moscow State University
- Made probability a branch of measure theory: events are sets, probability is a measure
- Also: conditional expectation in its modern form, Markov processes, turbulence, complexity
- The Kolmogorov–Smirnov test (Chapter 6) carries his name



Kolmogorov lecturing in Moscow, 1963–1964

Photo: Vsevolod Tarasevich (1963–1964); CC BY 4.0; Wikimedia Commons

Events, Conditional Probability, Independence

- **Conditional probability**: $P(A \mid B) = P(A \cap B)/P(B)$, for $P(B) > 0$
  - A, B: two events; $A \cap B$: both occur
  - in words: among the cases in which B occurs, the share in which A also occurs
- **Bayes' rule**: $P(B \mid A) = P(A \mid B)P(B)/P(A)$
  - $P(B)$: the probability of B before we learn A; $P(B \mid A)$: the updated probability once A is observed
  - it reverses the direction of conditioning: from "signal given crisis" to "crisis given signal"
- **Independent** events: $P(A \cap B) = P(A)P(B)$, i.e. $P(A \mid B) = P(A)$
  - learning that B occurred does not change the probability of A

Example: Down Days in a Row

- S&P 500, 2000–2026: $P(\text{down day}) = 46.3\%$
  - after a down day: $P(\text{down} \mid \text{down yesterday}) = 43.4\%$
  - two down days in a row: 20.1% in the data
  - under independence: $P(\text{down})^2 = 21.4\%$, close to the data
- BET: 50.0% after a down day vs 42.6% after an up day
  - the sign of the return persists from one day to the next
  - a consequence of thin trading: prices adjust to news over several days

Recap: Origins of Probability Theory

- Probability space $(\Omega, \mathcal{F}, P)$ and three axioms
- Fair value = expected value: the idea of Pascal and Fermat
- Conditional probability measures how information changes the odds

Random Variables

- A **random variable** $X : \Omega \to \mathbb{R}$ assigns a number to each outcome
  - tomorrow's log return, the number of up days in a week, the loss of a portfolio
- **Discrete**: finitely or countably many values x_i
  - **PMF** (probability mass function): $p(x_i) = P(X = x_i)$, with $\sum_i p(x_i) = 1$
  - $p(x_i)$: the probability of the value x_i; the probabilities of all values add up to 1
- **Continuous**: $P(a < X \leq b) = \int_a^b f(x)\,dx$
  - the probability of an interval $(a, b]$ is the area under f between a and b
  - f is the **PDF** (probability density function): $f \geq 0$ and the total area $\int f = 1$
  - a single point has no area: $P(X = x) = 0$ for every x
- Textbook: Franke, Härdle & Hafner (2019), Sec. 3.1

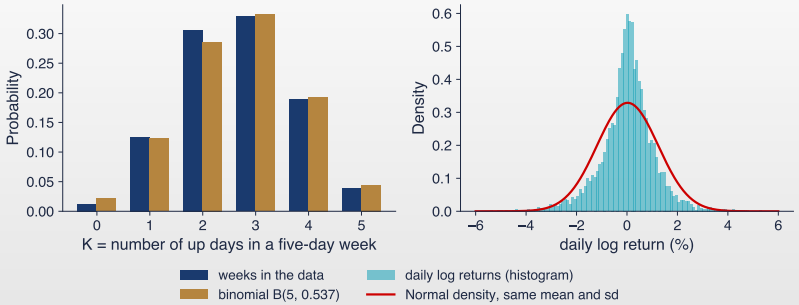
The Distribution Function and the Quantile Function

- **CDF** (cumulative distribution function): $F(x) = P(X \leq x)$
  - the probability that X does not exceed the threshold x
  - non-decreasing, right-continuous, from $F(-\infty) = 0$ to $F(\infty) = 1$
  - for a continuous X the density is its derivative: $f = F'$
- **Quantile** of order α: $q_\alpha = F^{-1}(\alpha) = \inf\{x : F(x) \geq \alpha\}$
  - $\alpha \in (0, 1)$: a probability; q_α is the smallest x with $P(X \leq x) \geq \alpha$
  - F^{-1}: the quantile function, the inverse of the CDF; inf: the smallest value of the set
  - median $= q_{0.5}$; the 1% quantile is the return exceeded on 99% of the days

Value at Risk and the Empirical CDF

- **VaR** (Value at Risk) 1%: $\text{VaR}_{1\%} = -q_{0.01}$
  - $q_{0.01}$: the 1% quantile of the returns, a negative number
  - the minus sign turns it into a loss: the loss exceeded with probability 1%
  - estimation and backtesting in Chapter 10
- The **empirical CDF** of a sample: $\hat{F}(x) = \frac{1}{n}\#\{t : r_t \leq x\}$
  - $r_1, \ldots, r_n$: the observed returns; n: the sample size
  - $\#\{\cdot\}$: the number of days in the set; $\hat{F}(x)$ is the share of days with a return at most x
  - the hat marks an estimate computed from data

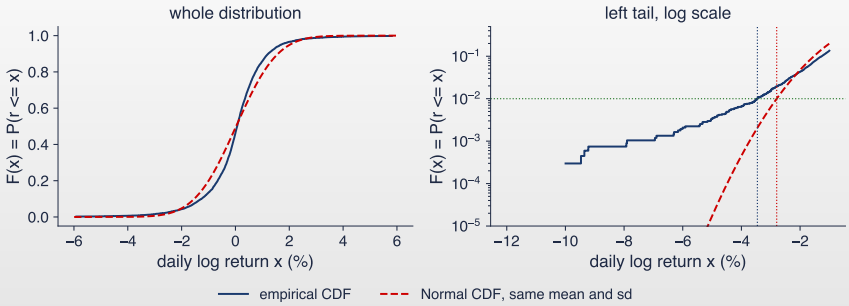
A Discrete and a Continuous Random Variable



- Left: K = number of up days in a full week of the S&P 500 (1 144 weeks)
  - compared with $B(5, p)$, the binomial distribution: 5 days, each up with probability $p = 0.537$
  - mean 2.67 vs 2.69; variance 1.16 vs 1.24
- Right: daily log returns, a continuous random variable; the histogram estimates the PDF
- The binomial fits K well; the Normal density misses the peak and the tails of the returns (Chapter 2)

SFM_ch4_random_variables

Empirical CDF and the 1% Quantile of the S&P 500



- $n = 6\,714$ days; median 0.063%; $\hat{F}(0) = 46.3\%$ of the days are not up days
- VaR 1%: empirical 3.45%, Normal 2.80%; $P(r < -3\%)$: 1.55% vs 0.63%
- $P(r < -5\%)$: 0.34% in the data, 199 times the value under the Normal distribution ($1.7 \times 10^{-3}\%$)

SFM_ch4_random_variables

Recap: Random Variables

- PMF for discrete, PDF for continuous variables; the CDF works for both
- Quantiles invert the CDF; VaR 1% is minus the 1% quantile
- The tails of the empirical CDF are much heavier than the Normal ones

Expectation

- **Expectation** (mean): $E[X] = \sum_i x_i p(x_i)$ or $E[X] = \int x f(x)\, dx$
  - x_i, $p(x_i)$: the values and their probabilities; f: the density
  - the centre of mass of the distribution; it exists only if $E|X| < \infty$ (not for Cauchy, Chapter 3)
- Of a function: $E[g(X)] = \int g(x) f(x)\, dx$, without finding the distribution of $g(X)$
- **Linearity**: $E[aX + bY + c] = aE[X] + bE[Y] + c$, always (no independence needed)
  - a, b, c: constants; X, Y: any two random variables
  - the expected return of a portfolio is the weighted mean of the expected returns
- In general $E[g(X)] \neq g(E[X])$
  - Jensen: for a convex g, $E[g(X)] \geq g(E[X])$; for example $E[e^X] > e^{E[X]}$
  - the source of the volatility drag of Chapter 1

Variance and Moments

- **Variance**: $\mathrm{Var}(X) = E[(X-\mu)^2] = E[X^2] - \mu^2$
  - $\mu = E[X]$; the variance is the mean squared distance from the mean
  - standard deviation $\sigma = \sqrt{\mathrm{Var}(X)}$, in the units of X; for returns it is the volatility
- $\mathrm{Var}(aX+b) = a^2\,\mathrm{Var}(X)$
  - adding a constant b does not change risk; multiplying by a multiplies the variance by a^2
- Moments: $E[X^k]$ for $k = 1, 2, \ldots$
  - skewness $E[(X-\mu)^3]/\sigma^3$: the asymmetry of the distribution (0 for the Normal distribution)
  - kurtosis $E[(X-\mu)^4]/\sigma^4$: equals 3 for the Normal distribution
  - excess kurtosis = kurtosis -3; positive for heavy tails (Chapter 2)

Worked Example: Annual Mean and Volatility

- S&P 500, daily log returns r_t
  - daily mean 0.0247%, daily standard deviation $\sigma = 1.213\%$
  - $A = 252$: the number of trading days per year
- Assume i.i.d. daily returns (independent and identically distributed)
  - the annual log return is the sum of the A daily ones: $\sum_{t=1}^{A} r_t$
- Mean: $E[\sum r_t] = A \times 0.0247 = 6.2\%$ a year
  - linearity of expectation
- Variance: $\mathrm{Var}(\sum r_t) = A\sigma^2$, because the covariances are zero
  - annual volatility $\sigma\sqrt{A} = 1.213\sqrt{252} = 19.3\%$
  - the square-root-of-time rule: volatility grows with $\sqrt{h}$ over h periods

Frequency of 4-Sigma Days: Chebyshev's Bound

- **Chebyshev's inequality**: for any X with finite variance, $P(|X - \mu| \geq k\sigma) \leq 1/k^2$
  - μ, σ: the mean and the standard deviation; $k > 1$: the distance from the mean, in standard deviations
  - it uses only the mean and the variance, so it holds for every distribution
- $k = 4$, S&P 500, 2000–2026
  - Chebyshev bound: at most 6.25% of the days
  - Normal distribution: 0.0063%
  - data: 46 days, 0.69%
- The data lie between the two
  - the Normal distribution is far too optimistic; the worst case of Chebyshev is too pessimistic

Recap: Expectation and Variance

- Expectation is linear; variance scales with the square
- For i.i.d. returns the mean grows with h and the volatility with $\sqrt{h}$
- Chebyshev bounds tail probabilities for any distribution

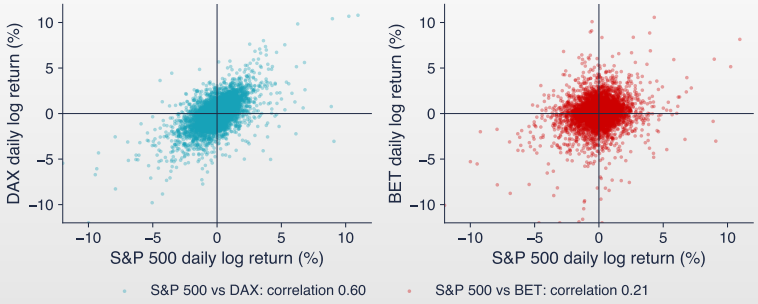
Joint Distribution and Marginals

- **Joint CDF**: $F(x, y) = P(X \leq x, Y \leq y)$
  - the probability that both variables are below their thresholds at the same time
  - for two returns: both markets fall below given levels on the same day
- **Marginal** CDF: $F_X(x) = F(x, \infty)$
  - the distribution of X alone, whatever the value of Y
- **Joint PDF** $f(x, y)$: $P((X, Y) \in B) = \iint_B f(x, y)\, dx\, dy$
  - B: a region of the plane; the probability is the volume under f above B
  - marginal density: $f_X(x) = \int f(x, y)\, dy$, integrating Y out

Covariance, Correlation and the Variance of a Sum

- **Covariance**: $\mathrm{Cov}(X, Y) = E[(X - \mu_X)(Y - \mu_Y)] = E[XY] - \mu_X \mu_Y$
  - μ_X, μ_Y: the means; positive when X and Y tend to be above their means together
- **Correlation**: $\rho = \mathrm{Cov}(X, Y)/(\sigma_X \sigma_Y) \in [-1, 1]$
  - the covariance divided by the two standard deviations: a number without units
  - measures **linear** dependence only; $|\rho| = 1$ iff $Y = a + bX$
- **Variance of a sum**: $\mathrm{Var}(aX + bY) = a^2\sigma_X^2 + b^2\sigma_Y^2 + 2ab\,\mathrm{Cov}(X, Y)$
  - a, b: constants (in a portfolio, the weights); the covariance term can raise or lower the risk
  - many assets: $\sigma_p^2 = w^\top \Sigma w$, with w the vector of weights and Σ the covariance matrix
  - σ_p^2: the variance of the portfolio; $w^\top$: the transposed vector

Joint Distributions of Daily Returns



- Common trading days only: prices are joined first, then differenced ($n = 6\,605$ days for S&P 500 and DAX)
- S&P 500 and DAX: $\rho = 0.60$; S&P 500 and BET: $\rho = 0.21$
- The cloud is elliptic in the centre, but extreme days are joint: crises hit all markets together

SFM_ch4_dependence

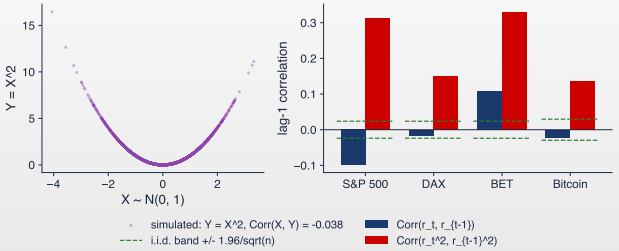
Worked Example: a 50/50 Portfolio of S&P 500 and DAX

- Daily standard deviations: $\sigma_1 = 1.22\%$ (S&P 500), $\sigma_2 = 1.43\%$ (DAX)
  - covariance 1.045; correlation 0.60
  - weights $a = b = 0.5$, so $a^2 = b^2 = 0.25$ and $2ab = 0.5$
- $\sigma_p^2 = 0.25\sigma_1^2 + 0.25\sigma_2^2 + 2 \times 0.25\,\mathrm{Cov} = 0.883 + 0.522 = 1.405$
  - $\sigma_p = 1.19\%$, below the average volatility 1.32%: diversification
- S&P 500 and BET, lower correlation: $\sigma_p = 1.04\%$ vs average 1.33%
  - the lower the correlation, the larger the reduction in risk
- Diversification works only through $\rho < 1$
  - in a crash correlations rise, exactly when diversification is needed

Independence and Uncorrelatedness

- X and Y are **independent** if $F(x, y) = F_X(x)F_Y(y)$ for all x, y
  - the joint CDF is the product of the marginal CDFs
  - then $E[g(X)h(Y)] = E[g(X)]E[h(Y)]$ for **all** functions g, h (for example squares or absolute values)
- Independent $\Rightarrow$ uncorrelated ($g, h =$ identity); the converse is false
- Example: $X \sim N(0, 1)$, $Y = X^2$
  - $\text{Cov}(X, X^2) = E[X^3] = 0$: uncorrelated
  - but Y is a function of X: completely dependent
- Exception: for jointly Normal variables, uncorrelated $\Leftrightarrow$ independent

Autocorrelation of Returns and of Squared Returns



- Left: simulated $Y = X^2$, sample correlation -0.038, yet Y is determined by X
- Right: $\mathrm{Corr}(r_t, r_{t-1})$ is small: S&P 500 -0.099, DAX -0.017, Bitcoin -0.022
  - $\mathrm{Corr}(r_t^2, r_{t-1}^2)$ is far outside the band: 0.31, 0.15, 0.14; BET 0.33
  - band $\pm 1.96/\sqrt{n}$: the values compatible with zero correlation at the 5% level
  - the negative S&P 500 value comes mostly from crisis days
- BET: $\mathrm{Corr}(r_t, r_{t-1}) = 0.109$, an effect of thin trading

SFM_ch4_dependence

Recap: Joint Distributions

- Covariance enters the variance of every portfolio
- Independent implies uncorrelated, not the other way round
- Returns: nearly uncorrelated, but their squares are correlated, so they are not independent

Conditional Expectation

- Conditional density: $f(y \mid x) = f(x, y)/f_X(x)$
  - the distribution of Y once we know that $X = x$: the joint density rescaled by the marginal one
- **Conditional expectation**: $E[Y \mid X = x] = \int y f(y \mid x)\, dy$
  - the mean of Y computed with the conditional density
  - $E[Y \mid X]$ is a random variable, a function of X
- It is the best forecast of Y from X: it minimises $E[(Y - g(X))^2]$ over all functions g
  - $E[(Y - g(X))^2]$: the mean squared error of the forecast $g(X)$
  - for returns: $E[r_t \mid \mathcal{F}_{t-1}]$, the forecast given all information $\mathcal{F}_{t-1}$ up to yesterday
- **Tower property** (law of iterated expectations): $E\big[E[Y \mid X]\big] = E[Y]$
  - averaging the conditional forecasts over all values of X gives back the unconditional mean

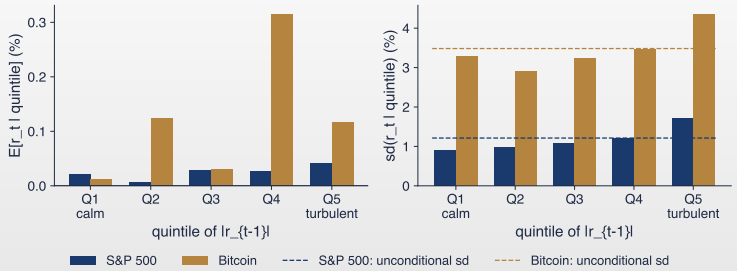
Conditional Variance and the Law of Total Variance

- **Conditional variance**: $\mathrm{Var}(Y \mid X) = E\left[(Y - E[Y \mid X])^2 \mid X\right]$
  - the spread of Y around its conditional mean, once X is known
- **Law of total variance**: $\mathrm{Var}(Y) = E\left[\mathrm{Var}(Y \mid X)\right] + \mathrm{Var}\left(E[Y \mid X]\right)$
  - $E[\mathrm{Var}(Y \mid X)]$: the average variance inside the groups defined by X (within)
  - $\mathrm{Var}(E[Y \mid X])$: how much the group means differ from each other (between)
  - total risk = within-group variance + between-group variance

Worked Example: Two Regimes

- X = the market regime, $Y = r$ = the daily return (in %)
  - calm: probability 0.7, mean 0.05%, standard deviation 0.8%
  - turbulent: probability 0.3, mean -0.10%, standard deviation 2%
- Mean (tower property): $E[r] = 0.7 \times 0.05 + 0.3 \times (-0.10) = 0.005\%$
- The two parts of the variance
  - within: $E[\mathrm{Var}(r \mid X)] = 0.7 \times 0.8^2 + 0.3 \times 2^2 = 1.648$
  - between: $\mathrm{Var}(E[r \mid X]) = 0.7(0.05 - E[r])^2 + 0.3(-0.10 - E[r])^2 = 0.0047$
- $\mathrm{Var}(r) = 1.653$, standard deviation 1.286%
  - almost all risk comes from the different variances of the regimes, not from their different means

Today's Return Given Yesterday's Move



- Quintiles of $|r_{t-1}|$: Q1 = the calmest 20% of yesterdays, Q5 = the most turbulent 20%
- Conditional mean: between 0.006% and 0.042% for the S&P 500, no usable pattern
- Conditional s.d.: S&P 500 from 0.91% (Q1) to 1.71% (Q5), ratio 1.88; Bitcoin ratio 1.33
- Law of total variance, S&P 500: 1.469 = 1.469 + 0.00013; the between part is 0.01%

SFM_ch4_dependence

Conditional Variance and GARCH Models

- Write the return as $r_t = \mu_t + \sigma_t z_t$, with z_t i.i.d., mean 0, variance 1
  - $\mu_t = E[r_t \mid \mathcal{F}_{t-1}]$: the conditional mean, close to constant
  - $\sigma_t^2 = \mathrm{Var}(r_t \mid \mathcal{F}_{t-1})$: the conditional variance, which changes with the news
- ARCH (autoregressive conditional heteroskedasticity) Engle (1982): $\sigma_t^2 = \omega + \alpha r_{t-1}^2$
  - today's variance is a base level plus a share of yesterday's squared return
  - $\omega > 0$: the base level; $\alpha \geq 0$: how strongly a large move yesterday raises today's variance
- GARCH Bollerslev (1986): $\sigma_t^2 = \omega + \alpha r_{t-1}^2 + \beta \sigma_{t-1}^2$ (Chapter 9)
  - $\beta \geq 0$: the persistence; a large β keeps a high variance high for many days
- Exactly the pattern in the chart: a large $|r_{t-1}|$ raises today's conditional s.d.
- The unconditional distribution is a mixture over σ_t
  - it has heavier tails than the Normal distribution, even if z_t is Normal

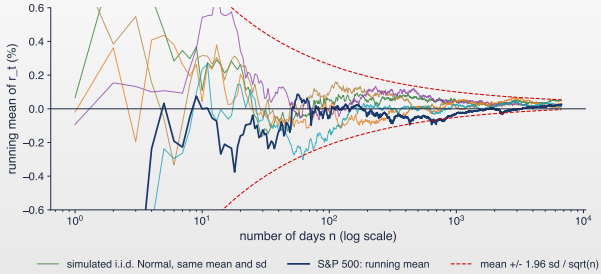
Recap: Conditioning

- $E[Y \mid X]$ is the best forecast; the tower property averages it back to $E[Y]$
- Returns: conditional mean almost flat, conditional variance strongly time-varying
- Modelling σ_t^2 is the job of ARCH and GARCH

Two Limit Theorems

- **LLN** (law of large numbers): for i.i.d. X_i with $E|X| < \infty$, $\bar{X}_n \to \mu$ as $n \to \infty$
  - $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$: the sample mean; $\mu = E[X]$: the true mean
  - the reason why averages, frequencies and Monte Carlo estimates work
- **CLT** (central limit theorem): with finite variance, $\sqrt{n}(\bar{X}_n - \mu)/\sigma \to N(0,1)$ (Chapter 2)
  - the standardised sample mean has, for large n, approximately the standard Normal distribution
  - gives the speed: the error of $\bar{X}_n$ is about $\sigma/\sqrt{n}$
  - infinite variance: sums converge to α-stable laws instead (Chapter 3)
- Both need some form of independence; dependent returns converge more slowly

The LLN for the Mean Return: Slow Convergence



- S&P 500, 26.6 years: mean 0.0247% a day, standard error $\sigma/\sqrt{n} = 0.0148\%$
  - t = mean / standard error $= 1.67$; $|t| > 1.96$ would mean a mean significantly different from 0 at 5%
- Annual mean 6.2%, 95% interval $[-1.1\%, 13.5\%]$
  - even 26 years do not pin down the expected return
  - for $t = 3$ at this mean and volatility we would need about 86 years of data

SFM_ch4_monte_carlo

The Monte Carlo Method

- Estimate $\theta = E[g(X)]$ by $\hat{\theta}_N = \frac{1}{N} \sum_{i=1}^{N} g(X_i)$ with simulated X_i
  - θ: the unknown quantity; g: a function of the simulated variable; N: the number of simulations
  - a probability is an expectation: $P(A) = E[1_A]$, with $1_A = 1$ if A occurs and 0 otherwise
  - the LLN makes $\hat{\theta}_N \to \theta$; the CLT gives the error $\text{sd}(g(X))/\sqrt{N}$
- Named after the Monaco casino by Ulam and von Neumann, Los Alamos, 1940s Metropolis & Ulam (1949)
- In finance Glasserman (2003)
  - prices of path-dependent options, portfolio VaR (Chapter 10), stress scenarios
  - any expectation without a closed form
- Two ingredients: uniform random numbers and a way to turn them into any distribution

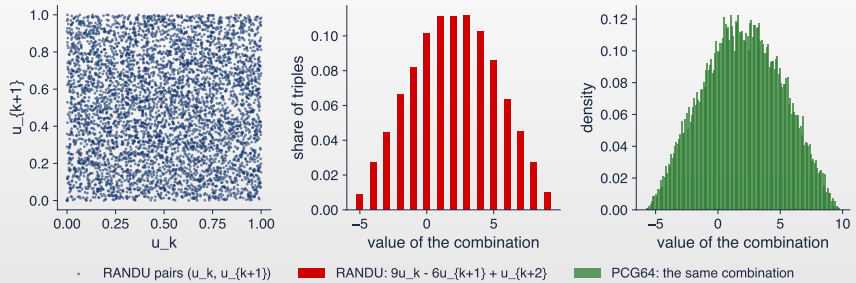
Pseudo-Random Numbers: Linear Congruential Generators

- **LCG** (linear congruential generator): $x_{k+1} = (ax_k + c) \bmod M$, $u_k = x_k/M \in [0, 1)$
  - each integer x_{k+1} is computed from the previous one x_k
  - a: the multiplier; c: the increment; M: the modulus; $\bmod M$: the remainder of the division by M
  - u_k: the uniform number in $[0, 1)$ obtained by dividing by M
- Deterministic: the same seed x_0 gives the same sequence
  - this makes results reproducible
  - the **period**: the number of values before the sequence repeats; at most M
- Real generators use huge M; the period must be far longer than any simulation

Worked Example: an LCG with $M = 16$

- SFErangen1: $a = 5$, $c = 1$, $M = 16$, $x_0 = 7$
  - first step: $x_1 = (5 \times 7 + 1) \bmod 16 = 36 \bmod 16 = 4$
  - the sequence x: 7, 4, 5, 10, 3, 0, 1, 6, 15, 12, 13, 2, 11, 8, 9, 14, 7
- Period 16, the maximum possible, M
  - with $a = 4$ the sequence sticks after two steps (period 1)
- Good uniform marginals are not enough
  - consecutive numbers must also look independent: the next chart

RANDU: Random Numbers Fall Mainly in the Planes (SFErandu)



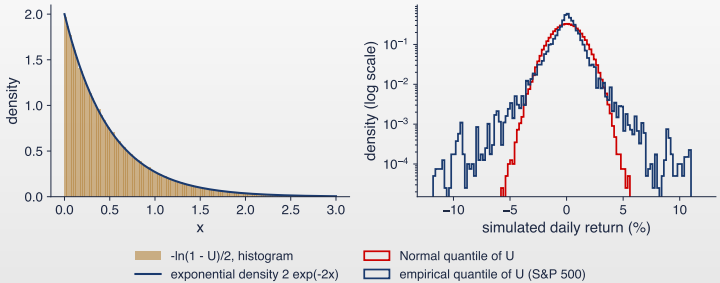
- RANDU, IBM, 1960s: $x_{k+1} = 65539\, x_k \bmod 2^{31}$; pairs look uniform (left), correlation 0.003
- But $9u_k - 6u_{k+1} + u_{k+2}$ takes only 15 integer values, from -5 to 9: all triples lie on 15 planes Marsaglia (1968)
- Modern generators: Mersenne Twister Matsumoto & Nishimura (1998), PCG64 (the NumPy default, right panel)

SFM_ch4_random_numbers

The Inverse Transform Method

- **Theorem**: if $U \sim U(0,1)$ and F is a CDF, then $X = F^{-1}(U)$ has CDF F
  - $U(0,1)$: the uniform distribution on $[0,1]$, so $P(U \leq u) = u$
  - in words: a uniform number, read as a probability, is turned into the quantile of that order
  - proof: $P(X \leq x) = P(F^{-1}(U) \leq x) = P(U \leq F(x)) = F(x)$
- Exponential with rate $\lambda > 0$: $F(x) = 1 - e^{-\lambda x}$, so $X = -\ln(1-U)/\lambda$
  - obtained by solving $u = F(x)$ for x; the mean is $1/\lambda$
  - $\lambda = 2$, $u = 0.9$: $x = -\ln(0.1)/2 = 1.151$
- Normal: $X = \mu + \sigma\,\Phi^{-1}(U)$
  - Φ^{-1}: the quantile function of $N(0,1)$; μ, σ: the desired mean and standard deviation
- Data: use the empirical quantile function $\hat{F}^{-1}$
  - this is resampling past days: historical simulation and the bootstrap

Inverse Transform: Exponential, Normal and Historical



- Left: 200 000 draws $-\ln(1-U)/2$ match the exponential density; sample mean 0.500 (theory $1/\lambda = 0.5$)
- Right: the same uniforms through the Normal and the empirical S&P 500 quantile functions
- Draws below -5%: 0.0035% Normal (7 draws; exact 0.0017%, the gap is simulation noise) vs 0.34% historical (data: 0.34%); worst draw -5.7% vs -12.7%

SFM_ch4_random_numbers

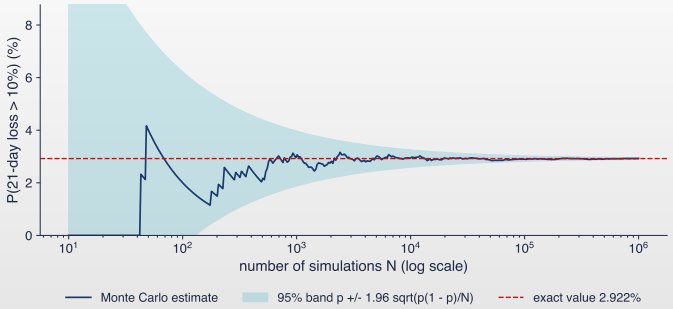
Worked Example: a Monthly Loss Probability by Monte Carlo (1/2)

- Model: S&P 500 daily log returns i.i.d. $N(0.0247, 1.213^2)$, in %
  - $N(m, s^2)$: the Normal distribution with mean m and variance s^2
- Event: the 21-day log return (one month) is below -10%
  - p: its probability, the quantity we want
- Exact: the sum of 21 i.i.d. Normal returns is $N(0.519, 5.56^2)$
  - mean 21×0.0247, standard deviation $1.213\sqrt{21}$
  - standardise: $z = (-10 - 0.519)/5.56 = -1.892$
  - $p = \Phi(z) = 2.922\%$, with Φ the CDF of $N(0, 1)$

Worked Example: a Monthly Loss Probability by Monte Carlo (2/2)

- Monte Carlo: simulate N months; $\hat{p}_N$ = share of simulated months below -10%
  - standard error $\sqrt{p(1-p)/N}$: the typical distance between $\hat{p}_N$ and p
- Results
  - $N = 1\,000$: 3.000% (s.e. 0.533%)
  - $N = 100\,000$: 2.901% (s.e. 0.053%)
  - 10 times more precision needs 100 times more simulations
- For a relative error of 10% at the 95% level: $N \approx 1.96^2(1-p)/(0.1^2 p) = 12\,762$
  - obtained from $1.96 \times \sqrt{p(1-p)/N} = 0.1\,p$: the half-width of the interval is 10% of p
  - a rare event needs many simulations

Convergence Rate of Monte Carlo: $1/\sqrt{N}$



- The estimate wanders for small N and settles inside the 95% band, which shrinks like $1/\sqrt{N}$
- $N = 10^6$: 2.926% vs exact 2.922%, s.e. 0.017%
- Monte Carlo removes the simulation error, not the model error: Seminar 4 (B3) repeats it with resampled days

SFM_ch4_monte_carlo

Recap: Random Numbers and Monte Carlo

- Pseudo-random generators are deterministic; check them in several dimensions
- $F^{-1}(U)$ turns uniforms into any distribution, including the empirical one
- Monte Carlo error $\propto 1/\sqrt{N}$; always report it

The Binomial Process

- Each period the price is multiplied by u (up) with probability p, or by $d < u$ (down) with probability $1 - p$
  - u, d: the up and down factors (e.g. $u = 1.2$: $+20\%$); the periods are independent
- After n steps with K up moves: $S_n = S_0 u^K d^{n-K}$
  - S_0: the initial price; $K \sim B(n, p)$: the number of up moves is binomial
  - $\ln S_n = \ln S_0 + K \ln u + (n - K) \ln d$: a random walk in log prices (Section 9)
- $E[S_n] = S_0 (pu + (1 - p)d)^n$, because the steps are independent
  - $pu + (1 - p)d$: the expected growth factor of one step
  - the basis of the binomial option pricing model Cox, Ross & Rubinstein (1979); textbook: Franke, Härdle & Hafner (2019), Ch. 7

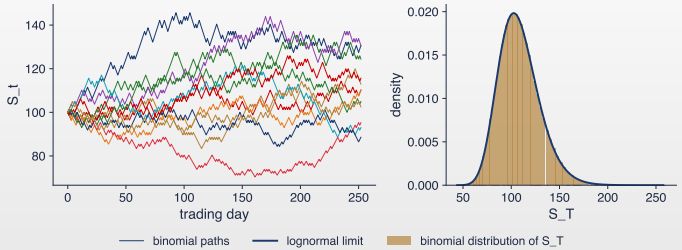
Worked Example: a Two-Step Tree

- $S_0 = 100$, $u = 1.2$, $d = 0.8$, $p = 0.6$
- Terminal values and probabilities
  - uu: 144 with $p^2 = 0.36$; ud or du: 96 with $2p(1-p) = 0.48$; dd: 64 with $(1-p)^2 = 0.16$
- $E[S_2] = 100 \times (0.6 \times 1.2 + 0.4 \times 0.8)^2 = 100 \times 1.04^2 = 108.16$
- $P(S_2 < 100) = 0.48 + 0.16 = 0.64$
  - one up and one down move lose money, because $ud = 0.96 < 1$
- With $p^* = (1-d)/(u-d) = 0.50$ the price is a fair game: $E[S_{t+1} \mid S_t] = S_t$ (Section 9)
  - p^* solves $p^* u + (1-p^*)d = 1$: the expected growth factor of a step equals 1

Calibrating the Tree to a Market (Cox–Ross–Rubinstein)

- **CRR** (Cox–Ross–Rubinstein) with step Δt: $u = e^{\sigma\sqrt{\Delta t}}$, $d = 1/u$, $p = (e^{\mu \Delta t} - d)/(u - d)$
  - Δt: the length of one step, in years; μ: the annual drift; σ: the annual volatility
  - $d = 1/u$: an up move followed by a down move returns to the start
  - chosen so that the mean and the variance of each step match a market with drift μ and volatility σ
- S&P 500, 2000–2026: $\sigma = 19.26\%$, $\mu = 8.08\%$ a year; daily steps, $\Delta t = 1/252$
  - $u = 1.0122$, $d = 0.9879$, $p = 0.5102$
- As $\Delta t \to 0$, $\ln S_T$ becomes Normal (CLT for K)
  - T: the horizon; the tree converges to GBM (Section 10)

Binomial Paths and the Lognormal Limit (SFEBinomp)



- Left: 12 paths over one year of daily steps, $S_0 = 100$
- Right: distribution of S_T after 252 steps ($T = 1$ year); $E[S_T] = 108.42 = 100e^{\mu}$
  - 5% quantile 78.5 vs 77.5 in the lognormal limit
- $P(S_T < 100)$: 34.9% binomial ($P(S_T \leq 100) = 39.7\%$) vs 37.3% lognormal
  - the exact value $S_T = 100$ has probability 4.8% in the tree; the lognormal value lies between the two

SFM_ch4_binomial

Recap: The Binomial Model

- Up or down each period: $S_n = S_0 u^K d^{n-K}$ with binomial K
- CRR matches μ and σ; the limit is lognormal
- A special p^* turns the price into a fair game: the key to option pricing

Stochastic Processes and Information

- A **stochastic process** $\{X_t\}$: a family of random variables indexed by time
  - one realisation = one path; a price series is a single path of the process
- $\mathcal{F}_t$: the information available at time t (past prices, news)
  - it grows with t: nothing is forgotten (a **filtration**)
- **(Weakly) stationary**: $E[X_t]$ and $\mathrm{Cov}(X_t, X_{t+h})$ do not depend on t
  - h: the lag between two dates; the mean and the dependence look the same at every date
  - returns: roughly yes; prices: no
- Textbook: Franke, Härdle & Hafner (2019), Ch. 4 and Sec. 11.3

White Noise and the Random Walk

- **White noise**: $E[\varepsilon_t] = 0$, $\text{Var}(\varepsilon_t) = \sigma^2$, $\text{Cov}(\varepsilon_t, \varepsilon_s) = 0$ for $t \neq s$
  - ε_t: the shock at time t; mean zero, constant variance σ^2, no correlation across dates
  - i.i.d. noise is stronger: the ε_t are independent, not only uncorrelated
- **Random walk**: $S_t = S_{t-1} + \mu + \varepsilon_t = S_0 + \mu t + \sum_{s=1}^{t} \varepsilon_s$
  - each step adds the drift μ and a new shock; S_0: the starting value
  - $E[S_t] = S_0 + \mu t$; $\text{Var}(S_t) = t\sigma^2$ grows with t: not stationary
  - simulated with $\sigma = 1$: variance 10.1, 49.8, 100.9 at $t = 10, 50, 100$
- Log prices as a random walk = returns as white noise
  - the benchmark of Chapter 7 (efficient markets)

Martingales

- $\{X_t\}$ is a **martingale** if $E|X_t| < \infty$ and $E[X_{t+1} \mid \mathcal{F}_t] = X_t$
  - $E|X_t| < \infty$: the mean exists; $\mathcal{F}_t$: the information up to t
  - a fair game: the best forecast of tomorrow is today
- A **martingale difference**: $\varepsilon_t = X_t - X_{t-1}$ with $E[\varepsilon_t \mid \mathcal{F}_{t-1}] = 0$
  - the change cannot be predicted from the past
- Examples
  - a random walk without drift; $S_t^2 - t\sigma^2$ for the same walk; the binomial price under p^*
- Martingale $\neq$ i.i.d. increments: the conditional **variance** may change, as in GARCH
  - the efficient-market idea in its weak form Fama (1970): excess returns (above the risk-free rate) are a martingale difference

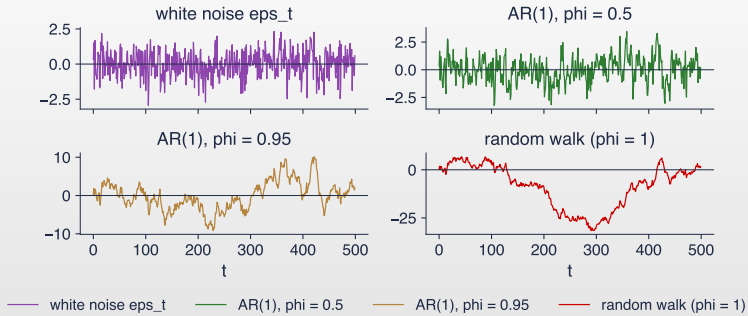
The AR(1) Process

- **AR(1)** (autoregressive of order 1): $X_t = c + \phi X_{t-1} + \varepsilon_t$, ε_t white noise
  - today's value is a constant, plus a share ϕ of yesterday's value, plus a new shock
  - c: the constant; ϕ: the persistence; σ^2: the variance of ε_t
  - $\phi = 0$: white noise; $\phi = 1$, $c = \mu$: random walk
- Stationary iff $|\phi| < 1$; then
  - mean $c/(1-\phi)$, variance $\sigma^2/(1-\phi^2)$
  - autocorrelation at lag h: $\rho(h) = \mathrm{Corr}(X_t, X_{t-h}) = \phi^h$, decaying geometrically
  - half-life of a shock: $\ln 0.5 / \ln \phi$ periods, the time until half of a shock is gone
- Uses
  - interest rates and spreads (mean reversion)
  - the squared returns behave like a persistent AR, which leads to GARCH

The AR(1) Process: Simulation

- Simulated with $\sigma = 1$, $\phi = 0.5$
  - variance 1.34 (theory $1/(1 - 0.5^2) = 1.33$)
  - $\rho(5) = 0.034$ (theory $0.5^5 = 0.031$); half-life 1.0 periods
- Simulated with $\sigma = 1$, $\phi = 0.95$
  - variance 10.20 (theory 10.26)
  - $\rho(5) = 0.772$; half-life 13.5 periods
- The closer ϕ is to 1, the larger the variance and the slower a shock fades

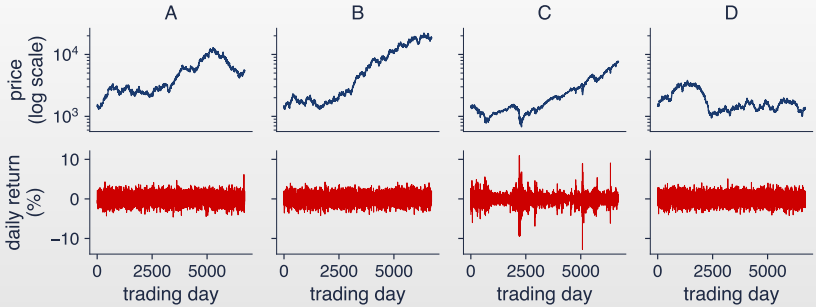
White Noise, AR(1) and Random Walk with the Same Shocks



- The larger ϕ, the longer a shock lives: the AR(1) with $\phi = 0.95$ wanders, but returns to 0
- The random walk never returns on purpose: shocks are permanent and its variance grows with t

SFM_ch4_processes

Question for the Room: Which One Is Real?



- One column is the S&P 500, 2000–2026; three are GBM paths (i.i.d. Normal returns) with the same mean and volatility
- **What do you think?** Which column is real, from the prices alone?
- **What do you think?** Which column is real, from the returns?

SFM_ch4_processes

Answer: Column C

- **Answer**: column C is the S&P 500
- Prices: hard to tell; random walks with drift produce booms and long falls too
  - this is why Bachelier (1900) and Osborne (1959) modelled prices as random walks
- Returns: obvious; the real series has quiet years and violent clusters (2008, 2020)
  - excess kurtosis 10.7 vs at most 0.02 in the simulations; $\mathrm{Corr}(r_t^2, r_{t-1}^2) = 0.31$ vs at most 0.006
  - largest absolute daily move 12.8% vs 6.2%
- Lesson: a model can fit the price path and still be wrong about risk

Recap: Discrete-Time Processes

- White noise: uncorrelated; i.i.d. noise: independent; a martingale difference: unpredictable mean
- Random walk: permanent shocks, variance $t\sigma^2$; AR(1) with $|\phi| < 1$: shocks fade at rate ϕ^h
- Prices look like random walks; returns are not i.i.d.

A History of Brownian Motion

- 1827: the botanist Robert Brown sees pollen particles jiggling in water
- 1900: Louis Bachelier models the Paris bond market with the same motion, five years before physics Bachelier (1900)
- 1905: Albert Einstein explains it by molecular collisions Einstein (1905)
- 1923: Norbert Wiener builds the rigorous mathematical object Wiener (1923)
- 1959: Osborne (1959) applies it to log prices: geometric Brownian motion



Robert Brown (1773–1858)

Photo: Maull & Polyblank (1855); public domain; Wikimedia Commons

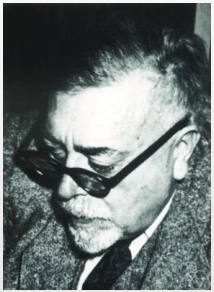


Louis Bachelier (1870–1946)

Photo: unknown author; public domain; Wikimedia Commons

The Wiener Process

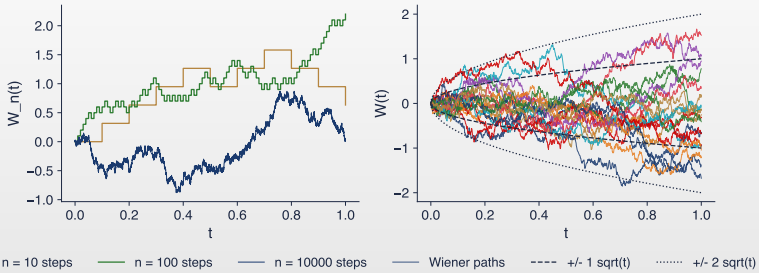
- $\{W(t), t \geq 0\}$ is a **Wiener process** (standard Brownian motion) if
  - $W(0) = 0$
  - increments over disjoint intervals are independent
  - $W(t) - W(s) \sim N(0, t - s)$ for $s < t$: the variance equals the elapsed time
  - the paths are continuous
- Limit of the scaled random walk $W_n(t) = (\varepsilon_1 + \cdots + \varepsilon_{\lfloor nt \rfloor})/\sqrt{n}$ (Donsker)
  - $\varepsilon_i = \pm 1$ with probability $1/2$ each; n steps per unit of time
  - $\lfloor nt \rfloor$: the number of steps up to time t (integer part); dividing by $\sqrt{n}$ keeps the variance equal to t
- Textbook: Franke, Härdle & Hafner (2019), Ch. 5



Norbert Wiener (1894–1964)

Photo: Konrad Jacobs; CC BY-SA 2.0 DE; Wikimedia Commons

The Wiener Process as a Limit of Random Walks (SFEWienerProcess)



- Left: scaled random walks with 10, 100 and 10 000 steps; the steps vanish, the randomness stays
- Right: 20 Wiener paths; $\mathrm{Var}(W(t)) = t$, so they spread like $\pm\sqrt{t}$
  - simulated $\mathrm{Var}(W(1)) = 1.006$; 68.2% of $W(1)$ inside ± 1 (theory 68.3%)

SFM_ch4_wiener_gbm

Properties of Wiener Paths

- Continuous everywhere, differentiable nowhere
  - the slope $(W(t+h) - W(t))/h \sim N(0, 1/h)$ has a variance $1/h$ that grows without bound as $h \to 0$
- **Quadratic variation**: $\sum_k (W(t_{k+1}) - W(t_k))^2 \to t$ on a finer and finer grid
  - $0 = t_0 < t_1 < \cdots < t$: the grid; the sum of squared increments tends to the elapsed time t
  - 1000 steps on $[0, 1]$: mean 0.999, s.d. 0.045 across 10 000 paths
  - the reason why $(dW)^2 = dt$ in Itô calculus, and why realised variance measures volatility (Chapter 8)
- $\mathrm{Cov}(W(s), W(t)) = \min(s, t)$: simulated $\mathrm{Cov}(W(0.5), W(1)) = 0.496$
- W is a martingale; so is $W(t)^2 - t$

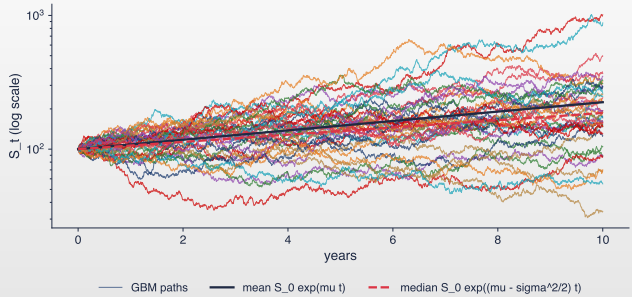
Geometric Brownian Motion

- **GBM** (geometric Brownian motion): $dS_t = \mu S_t\, dt + \sigma S_t\, dW_t$
  - dS_t: the change of the price over a very short interval dt
  - dW_t: the increment of the Wiener process over dt, distributed $N(0, dt)$
  - the relative change dS/S has drift μ and volatility σ, both per year
- Solution: $S_t = S_0 \exp\left((\mu - \sigma^2/2)t + \sigma W_t\right)$
  - $\ln(S_t/S_0) \sim N\left((\mu - \sigma^2/2)t, \sigma^2 t\right)$
  - lognormal prices, i.i.d. Normal log returns
- The model behind Black & Scholes (1973) (textbook: Franke, Härdle & Hafner (2019), Ch. 6)

GBM: Mean, Median and Simulation

- Mean $E[S_t] = S_0 e^{\mu t}$; median $S_0 e^{(\mu - \sigma^2/2)t}$
  - the median is lower: half of the paths end below $S_0 e^{(\mu - \sigma^2/2)t}$
  - the gap, $\sigma^2/2$ a year in log terms, is the volatility drag
- Exact simulation on a grid: $S_{t+\Delta t} = S_t \exp\left((\mu - \sigma^2/2)\Delta t + \sigma\sqrt{\Delta t}\, Z\right)$
  - Δt: the grid step, in years; $Z \sim N(0,1)$, a new independent draw at each step
  - exact because each log increment is Normal with the right mean and variance

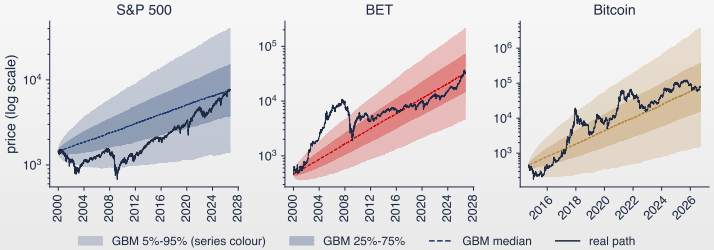
GBM Calibrated to the S&P 500 (SFEsimGBM)



- Calibration, 2000–2026: $\sigma = 19.26\%$, log drift $\mu - \sigma^2/2 = 6.23\%$, so $\mu = 8.08\%$ a year
- After $T = 10$ years: mean 224.4, median 186.4
  - 62.0% of the paths end below the mean; theory: $\Phi(\sigma\sqrt{T}/2) = 62.0\%$, with Φ the $N(0,1)$ CDF
- Probability of a loss after 10 years: 15.3%

SFM_ch4_wiener_gbm

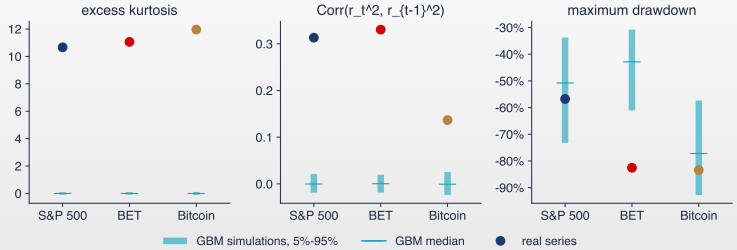
Real Paths in the GBM Fan



- Each fan: 4000 GBM paths with the drift and volatility of the series itself, started on its first day
- S&P 500: the path reached the 0.5% quantile on 23 July 2002, then the median at the end
  - outside the 90% band 7% of the time
- BET: outside the band 24% of the time
  - the 2000–2007 boom and the 2008 crash are not GBM-like
- Bitcoin: $\sigma = 66.6\%$ a year makes the fan enormous
  - annualised with $A = 365$: crypto trades every calendar day ($A = 252$ for the indices)

SFM_ch4_wiener_gbm

Limitations of the GBM Model



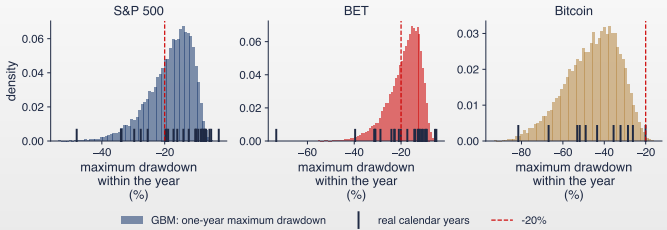
- Excess kurtosis: S&P 500 10.7, BET 11.1, Bitcoin 12.0; 90% of 300 GBM simulations lie within $[-0.10, 0.09]$
- Volatility clustering: $\text{Corr}(r_t^2, r_{t-1}^2) = 0.31, 0.33, 0.14$ vs at most 0.03 under GBM
- Maximum drawdown $\min_t \left(S_t / \max_{s \le t} S_s - 1 \right)$: the largest fall from a previous peak
  - S&P 500 -57% is ordinary for GBM; BET -83% is beyond all but 5% of the paths (5% quantile -61%)
- 4-sigma days: 46 (S&P 500) vs at most 1 in 95% of the GBM paths; crypto has the same problem Pele et al. (2023)

SFM_ch4_wiener_gbm

Recap: Wiener Process and GBM

- Wiener process: independent Normal increments with variance equal to the elapsed time
- GBM: lognormal prices, median below mean, the benchmark of option pricing
- Real returns have heavy tails and volatility clusters that GBM cannot produce

An Open Research Question: Drawdown Risk under GBM



- Drawdown (Limitations of the GBM Model), here within a calendar year
  - beyond 20%: S&P 500 in 6 of 26 years (23%) vs 34% under GBM; BET 46% vs 34%
- Beyond 40%: GBM expects 0.14 such years for the S&P 500 and 0.18 for the BET
  - both had 1 (2008: −48% and −73%)
- Open question
  - does GBM overstate moderate drawdowns and understate extreme ones because calm and crisis years cluster?

SFM_ch4_drawdown_open

Possible Uses of AI Tools

- **Literature**: find studies of drawdown distributions and of their link to volatility clustering, and summarise them
- **Code**: draft simulations of yearly drawdowns under GBM, the i.i.d. bootstrap and a GARCH model
- **Design**: propose a test that compares the observed number of bad years with the model probability
- Example prompt
  - `Write Python code that simulates 10,000 one-year paths of daily S&P 500 returns under GBM and under an i.i.d. bootstrap of 2000-2026 returns, and returns the probability of a drawdown beyond 20% and 40% within the year.`

What to Check

- Definitions: drawdown from the running peak within the year, not from the first day of the year
- Calibration: the same μ and σ, the same number of trading days a year as the data
- Inference: 26 years is a small sample; compute a binomial p-value, not only shares
- Mechanism: the i.i.d. bootstrap keeps heavy tails but removes clustering; the difference isolates the effect of clustering
- References: every cited paper must exist; check the DOI

Project Idea

- **Question**: how wrong is GBM about the probability of a bad year, and why?
  - data: S&P 500, DAX, BET and Bitcoin from the course data
- Steps
  - compute within-year maximum drawdowns for every calendar year
  - simulate the same statistic under GBM, the i.i.d. bootstrap and a GARCH(1,1) model (Chapter 9)
  - compare the shares beyond 20% and 40% with binomial tests
- Deliverable: one table, one chart, and a paragraph on what the data can and cannot show
- Declare any AI use, and list the errors of the AI that you corrected Wang et al. (2023)

Key Formulas

Quantity	Formula
Variance of a sum	$\mathrm{Var}(aX + bY) = a^2\sigma_X^2 + b^2\sigma_Y^2 + 2ab\,\mathrm{Cov}(X, Y)$
Correlation	$\rho = \mathrm{Cov}(X, Y)/(\sigma_X\sigma_Y)$
Tower property	$E[E[Y \mid X]] = E[Y]$
Total variance	$\mathrm{Var}(Y) = E[\mathrm{Var}(Y \mid X)] + \mathrm{Var}(E[Y \mid X])$
Inverse transform	$X = F^{-1}(U),\ U \sim U(0, 1)$
Monte Carlo error	$\sqrt{p(1-p)/N}$
Random walk	$S_t = S_{t-1} + \mu + \varepsilon_t,\ \mathrm{Var}(S_t) = t\sigma^2$
AR(1)	$\mathrm{Var}(X_t) = \sigma^2/(1-\phi^2),\ \rho(h) = \phi^h$
GBM	$S_t = S_0 \exp((\mu - \sigma^2/2)t + \sigma W_t),\ E[S_t] = S_0 e^{\mu t}$

Self-Assessment

- **Question**: $\sigma_1 = 2\%$, $\sigma_2 = 3\%$, $\rho = 0.5$; what is the volatility of the 50/50 portfolio?
  - **Answer**: $\sqrt{0.25 \times 4 + 0.25 \times 9 + 2 \times 0.25 \times 0.5 \times 2 \times 3} = \sqrt{4.75} = 2.18\%$
- **Question**: GBM with $\mu = 8\%$, $\sigma = 20\%$, $S_0 = 100$; what are the mean and the median of S_1?
  - **Answer**: mean $100e^{0.08} = 108.33$, median $100e^{0.06} = 106.18$
- **Question**: AR(1) with $\phi = 0.8$ and $\sigma = 1$; what is its variance?
  - **Answer**: $1/(1 - 0.64) = 2.78$
- Next: Chapter 5, heavy tails and extreme value theory

References (1/2)

- Bachelier, L. (1900). Théorie de la spéculation. *Annales scientifiques de l'École Normale Supérieure*, 17, 21–86.
- Black, F., Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy*, 81(3), 637–654.
- Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3), 307–327.
- Borak, S., Härdle, W.K., López-Cabrera, B. (2013). *Statistics of Financial Markets: Exercises and Solutions*, 2nd ed. Springer.
- Cont, R. (2001). Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance*, 1(2), 223–236.
- Cox, J.C., Ross, S.A., Rubinstein, M. (1979). Option pricing: a simplified approach. *Journal of Financial Economics*, 7(3), 229–263.
- Einstein, A. (1905). Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen. *Annalen der Physik*, 322(8), 549–560.
- Engle, R.F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, 50(4), 987–1007.
- Fama, E.F. (1970). Efficient capital markets: a review of theory and empirical work. *Journal of Finance*, 25(2), 383–417.
- Franke, J., Härdle, W.K., Hafner, C.M. (2019). *Statistics of Financial Markets: An Introduction*, 5th ed. Springer.

References (2/2)

- Glasserman, P. (2003). *Monte Carlo Methods in Financial Engineering*. Springer.
- Kolmogoroff, A. (1933). *Grundbegriffe der Wahrscheinlichkeitsrechnung*. Springer, Berlin.
- Marsaglia, G. (1968). Random numbers fall mainly in the planes. *Proceedings of the National Academy of Sciences*, 61(1), 25–28.
- Matsumoto, M., Nishimura, T. (1998). Mersenne twister: a 623-dimensionally equidistributed uniform pseudo-random number generator. *ACM Transactions on Modeling and Computer Simulation*, 8(1), 3–30.
- Metropolis, N., Ulam, S. (1949). The Monte Carlo method. *Journal of the American Statistical Association*, 44(247), 335–341.
- Osborne, M.F.M. (1959). Brownian motion in the stock market. *Operations Research*, 7(2), 145–173.
- Pele, D.T., Wesselhöfft, N., Härdle, W.K., Kolossiatis, M., Yatracos, Y.G. (2023). Are cryptos becoming alternative assets? *European Journal of Finance*, 29(10), 1064–1105.
- Wang, H., Fu, T., Du, Y., Gao, W., et al. (2023). Scientific discovery in the age of artificial intelligence. *Nature*, 620, 47–60.
- Wiener, N. (1923). Differential-space. *Journal of Mathematics and Physics*, 2(1–4), 131–174.