








# Modelling Financial Markets

## Seminar 11: Continuous-Time Models — Hands-on




Daniel Traian PELE
Drd. Antoaneta Amza, antoaneta.amza@csie.ase.ro

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

## Today's Questions and Seminar Structure

- Questions of the seminar
  - **what can daily data tell us about a continuous-time model?**
  - **what can they not tell us?**
- Deliverables
  - the answers to the A exercises
  - one table and one figure per B exercise
  - the corrected C1 code
  - one interpretation sentence per exercise
- Models and data
  - models: Geometric Brownian Motion (GBM), Ornstein–Uhlenbeck (OU) and the autoregressive model of order one, AR(1)
  - markets: the S&P 500 (the US equity index of Standard & Poor's), BET (Bucharest Exchange Trading, the Bucharest Stock Exchange index) and Bitcoin
  - in C1: an answer generated by Artificial Intelligence (AI)
- Notebook for today: `chapter11_seminar_notebook.ipynb` (Colab)
  - each task names its notebook section

# Exercise Map

| Exercise | Question | Type | Model |
|---|---|---|---|
| A1 | how precise is $\kappa$ from an AR(1)? | Solved | – |
| A2 | what does the extra term in $\int W \, \mathrm{d}W$ measure? | Solved | – |
| A3 | how non-Normal do five jumps a year make daily returns? | Solved | – |
| B1 | is the S&P 500 a GBM?<br>how well do we know $\sigma$ and $m$? | Solved | – |
| B2 | do interest rates really mean-revert? | Solved | – |
| C1 | what is wrong in an AI answer about GBM? | Proposed | A2, B1 |
| B3 | BET and Bitcoin: same tests as B1 (at home) | Proposed | B1 |
| After the lecture | B4: how fast do Euler and Milstein converge?<br>B5: do the data reject GBM in favour of jumps? | Solved | – |
| Extensions | A4–A9; B2 and B5 Extended, B6, B7, B8, B8 Extended, B9; C2, C3 (optional, after the seminar) | both | on each task |

- **[Solved]**: full solution in the slides and the notebook, a model to copy; **[Proposed]**: you apply or extend the method; the solution is discussed in class
- Part A: derivations on paper; Part B: estimation and inference on data, each ending with an interpretation; Part C: open questions and AI critique

## What Does Each Data Series Measure?

- FRED (Federal Reserve Economic Data) is the economic database of the Federal Reserve Bank of St. Louis. DTB3 (3-Month Treasury Bill: secondary market rate) is its code for the three-month United States (US) Treasury bill rate on the secondary market, quoted at an annual rate on a discount basis. It is an interest rate, not a price return.
- VIX (Cboe Volatility Index) measures 30-day implied volatility for the S&P 500, quoted as an annualised percentage. In a volatility model it is a proxy, not the unobserved instantaneous variance itself.
- SPY (SPDR S&P 500 ETF Trust) is the ticker of a fund tracking the S&P 500. Its 5-minute bars serve the intraday-volatility extension; they are different observations from the daily index prices used in the core exercises.
- A trading-day calendar and a calendar-day calendar produce different time steps. Keep the calendar and the units attached to every estimate.

# Data Used

| Series | Symbol, field | Calendar, cleaning | Sample | Returns |
|---|---|---|---|---|
| S&P 500 | `GSPC.INDX`, close | weekdays; stale quote = unchanged close and high = low | 1990 – 2026 | $n = 9\,245$ |
| BET | `BET`, close | weekdays; the BET series has no daily high/low, so every unchanged close counts as stale (none occurs) | 2000 – 2026 | $n = 6\,670$ |
| Bitcoin | `BTC-USD.CC`, close | 7 days a week, $\Delta t = 1/365$ | 2014 – 2026 | $n = 4\,384$ |
| VIX | `VIX.INDX`, close | weekdays | 1990 – 2026 | $n = 9\,283$ (9 284 levels) |
| DTB3 | FRED, percent /100 | no filter | 1954 – 2026 | $n = 18\,170$ levels |

- Daily prices of the course data, one series per asset (field used: close); end date 18 September 2026
- Log returns $r_t = \ln P_t - \ln P_{t-\mathbf{1}}$ on each series' own calendar; genuine zero returns (unchanged close on a day with trading) are kept; extensions also use SPY 5-minute bars and the Chapter 9 realised variance

# Setup: Run and Verify

```
# notebook section "Data" defines
# log_returns(key), load_vix(), read_fred(id)
rets = {k: log_returns(k)
        for k in ['sp500', 'bet', 'btc']}
for k, r in rets.items():
    print(k, len(r), r.index[0].date(),
          r.index[-1].date())
tb = read_fred('DTB3') / 100
print('DTB3', len(tb), tb.index[0].date())
print('VIX', len(load_vix()))
```

- Expected output

| | *n* | first date |
|---|---|---|
| sp500 | 9 245 | 1990-01-03 |
| bet | 6 670 | 2000-01-06 |
| btc | 4 384 | 2014-09-18 |
| DTB3 (levels) | 18 170 | 1954-01-04 |
| VIX (levels) | 9 284 | 1990-01-02 |

- All series end on 2026-09-18; returns start one day after the first price
- If a count differs, stop and check the symbol, the field and the start date

## Which Return and Which Time Step?

- The symbols $P_t$ and $S_t$ both denote an asset price. The price symbol does not change the definition of its return.
- A price rising from 100 to 110 has a simple return of 10 percent. Its log return is about 0.0953, or 9.53 percent when multiplied by 100. The two are close only for small changes.
- Here daily log returns are decimals. Volatility and drift are annualised; multiplying a displayed estimate by 100 expresses it in percent without changing the underlying calculation.
- $\Delta$ and $\Delta t$ both denote the sampling step in years. Daily observations use the calendar of that market; finer sampling is different from observing more calendar years.

## Notation Used in Every Exercise

- $W_t$: standard Brownian motion (Brownian Motion); $P$: real-world measure, $Q$: risk-neutral measure (Extension Packs: Measure Change and Heston); $\Phi$: standard Normal distribution function
- Units: time in years; drifts and volatilities per year; standard errors (SE) of percentages in percentage points; every estimate with its SE or confidence interval (CI)

## Brownian Motion

- Details: Chapter 11
  - proofs and depth in the lecture
- **Random walk**: one step of size $\pm\sqrt{\Delta}$ every $\Delta$ years

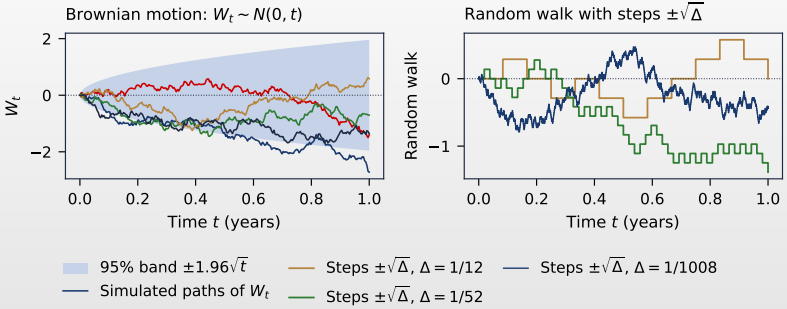
$$X_{t+\Delta} = X_t + \sqrt{\Delta}\,\epsilon_{t+\Delta}, \qquad \epsilon = \pm 1 \text{ with probability } 1/2$$

  - after $t/\Delta$ steps: $\mathrm{Var}(X_t) = (t/\Delta)\cdot\Delta = t$, for any step length $\Delta$
- **Standard Brownian motion** $W_t$, $t \geq 0$: the limit of the random walk as $\Delta \to 0$

$$W_0 = 0, \qquad W_{t+\Delta} - W_t \sim N(0,\ \Delta) \text{ independent of the past}, \qquad W_t \sim N(0,\ t)$$

  - continuous paths; the s.d. of $W_t$ grows like $\sqrt{t}$
  - the s.d. of an increment is $\sqrt{\Delta}$, not $\Delta$: one trading day, $\Delta = 1/252$, has s.d. $\sqrt{1/252} = 0.063$
  - a simulated step of $\sigma W$ is $\sigma\sqrt{\Delta}\,Z$, with $Z \sim N(0,1)$

# Simulated Brownian Paths




- Left: five simulated paths of $W_t$ over one year (daily steps) and the 95% band $\pm 1.96\sqrt{t}$
  - the band widens like $\sqrt{t}$: uncertainty grows with the horizon, but more slowly than the horizon
- Right: random walks with steps $\pm\sqrt{\Delta}$ for $\Delta = 1/12$, $1/52$ and $1/1008$
  - as $\Delta$ shrinks, the staircase becomes a Brownian path: continuous, but nowhere smooth

## Quadratic Variation

- **Quadratic variation**: the limit of the sum of squared increments on a grid $0 = t_0 < t_1 < \cdots < t_n = T$ with step $\Delta = T/n$

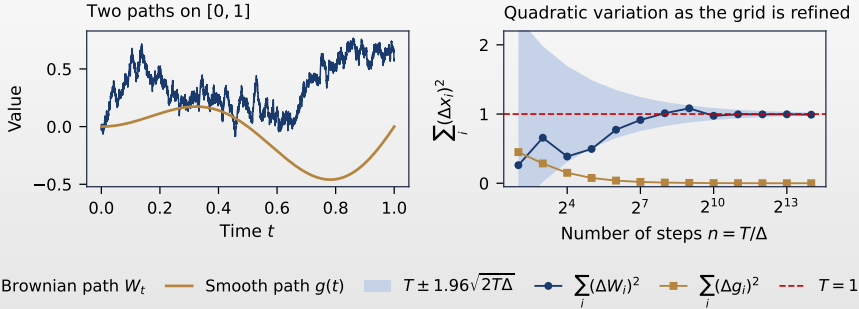
$$[W]_T = \lim_{\Delta \to 0} \sum_{i=0}^{n-1} \left(W_{t_{i+1}} - W_{t_i}\right)^2 = T$$

  - each square has mean $\Delta$ and variance $2\Delta^2$; the $n$ squares are independent
  - so the sum has mean $n\Delta = T$ and variance $n \cdot 2\Delta^2 = 2T\Delta \to 0$: it converges to the constant $T$
- A smooth path $g$ (with a derivative $g'$) has quadratic variation 0

$$\sum_i (\Delta g_i)^2 \approx \sum_i g'(t_i)^2 \, \Delta^2 \le \max_t g'(t)^2 \cdot T\Delta \to 0$$

- Rules for computing with $\mathrm{d}W$: $(\mathrm{d}W)^2 = \mathrm{d}t$, $\mathrm{d}t\,\mathrm{d}W = 0$, $(\mathrm{d}t)^2 = 0$
  - for a price $\mathrm{d}\ln S = m\,\mathrm{d}t + \sigma\,\mathrm{d}W$: $[\ln S]_T = \sigma^2 T$; realised variance (Chapter 9) estimates it

# Quadratic Variation on a Simulated Path




- Left: one Brownian path on $[0, 1]$ and a smooth path $g(t) = 0.6\, t \sin(2\pi t)$
- Right: the sums of squared increments on grids of $n = 4, 8, \ldots, 2^{\mathbf{14}}$ steps
  - Brownian path: the sum settles at $T = 1$ (0.99 with $2^{\mathbf{14}}$ steps), inside the band $T \pm 1.96\sqrt{2T\Delta}$
  - smooth path: the sum falls to 0, like $\Delta$

## The Itô Integral

- **Itô integral** Itô (1944): the limit of sums with the integrand taken at the **left** point of each step

$$\int_0^T f_t \, \mathrm{d}W_t = \lim_{\Delta \to 0} \sum_{i=0}^{n-1} f_{t_i} \left( W_{t_{i+1}} - W_{t_i} \right)$$

  - $f_t$: the integrand, known at time $t$ (e.g. a position chosen before the price moves); $t_i$: the grid points
  - reading: the gain of a strategy holding $f_{t_i}$ units over $[t_i, t_{i+1}]$

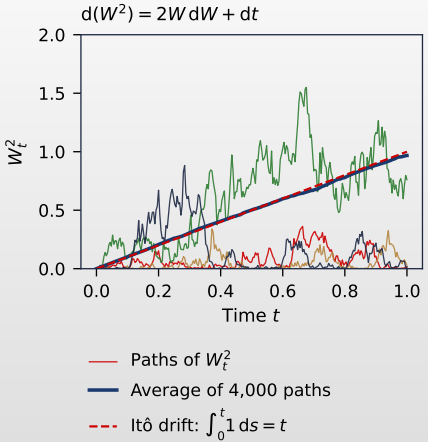
- Two properties: mean zero and the Itô isometry

$$\mathbb{E}\left[ \int_0^T f_t \, \mathrm{d}W_t \right] = 0, \qquad \mathrm{Var}\left( \int_0^T f_t \, \mathrm{d}W_t \right) = \int_0^T \mathbb{E}\left[ f_t^2 \right] \mathrm{d}t$$

  - mean zero: $f_{t_i}$ is fixed before the increment, which has mean 0 and is independent of it

- **Stratonovich integral**: the integrand averaged over the **two endpoints**, $\frac{1}{2}(f_{t_i} + f_{t_{i+1}})$
  - it uses the next move, so in general its mean is not zero (A2)

# Itô's Lemma




- **Itô's lemma**: the chain rule for a process $\mathrm{d}X = a\,\mathrm{d}t + b\,\mathrm{d}W$

$$\mathrm{d}f(X) = \left(af' + \tfrac{1}{2}b^2 f''\right)\mathrm{d}t + bf'\,\mathrm{d}W$$

  - $a$: drift, $b$: diffusion coefficient of $X$; $f'$, $f''$: first and second derivatives of $f(x)$
  - a second-order Taylor expansion in which $(\mathrm{d}X)^2 = b^2\,\mathrm{d}t$ does not vanish
  - the ordinary chain rule misses the term $\frac{1}{2}b^2 f''\,\mathrm{d}t$
- Example: $f(x) = x^2$, $X = W$ ($a = 0$, $b = 1$)

$$\mathrm{d}(W^2) = 2W\,\mathrm{d}W + \mathrm{d}t$$

  - chart: the average of $W_t^2$ grows like $t$ (0.97 at $t = 1$), although $\mathrm{d}W$ has mean 0
- Moments: $\mathbb{E}[W_t^2] = t$, $\mathbb{E}[W_t^4] = 3t^2$, $\mathbb{E}[e^{aW_t}] = e^{a^2 t/2}$

## Geometric Brownian Motion

- **GBM** Samuelson (1965): the simple return $\mathrm{d}S/S$ over $\mathrm{d}t$ has mean $\mu\,\mathrm{d}t$ and variance $\sigma^2\,\mathrm{d}t$

$$\mathrm{d}S_t = \mu S_t\,\mathrm{d}t + \sigma S_t\,\mathrm{d}W_t$$

  - $\mu$: expected return per year (price drift); $\sigma$: volatility per year
- Itô with $f = \ln S$ ($f' = 1/S$, $f'' = -1/S^2$) gives the log price and its law

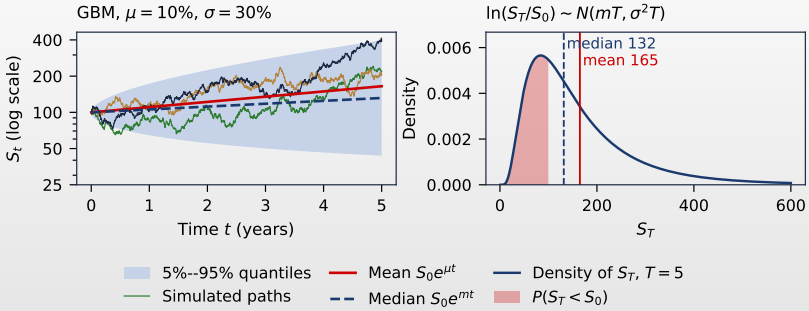
$$\mathrm{d}\ln S_t = m\,\mathrm{d}t + \sigma\,\mathrm{d}W_t, \qquad m = \mu - \tfrac{1}{2}\sigma^2, \qquad \ln\frac{S_T}{S_0} \sim N\left(mT,\ \sigma^2 T\right)$$

  - $m$: the log drift, the growth rate of the median; exact simulation: $\ln S_{t+\Delta} = \ln S_t + m\Delta + \sigma\sqrt{\Delta}\,Z$
- Mean, median and the probability of a loss after $T$ years

$$\mathbb{E}[S_T] = S_0 e^{\mu T}, \qquad \mathrm{Med}(S_T) = S_0 e^{mT}, \qquad P(S_T < S_0) = \Phi\left(-\frac{m\sqrt{T}}{\sigma}\right)$$

  - $\Phi$: the standard Normal distribution function; the median is below the mean: a few very large outcomes pull the mean up
  - example: $\mu = 10\%$, $\sigma = 30\%$: $m = 0.10 - 0.045 = 5.5\%$; $P(S_1 < S_0) = \Phi(-0.18) = 0.43$

# Simulated GBM Paths and the Lognormal Law




- Left: GBM with $\mu = 10\%$, $\sigma = 30\%$, $S_0 = 100$, over 5 years, log scale
  - the mean $100e^{\mu t}$ (red) grows faster than the median $100e^{mt}$ (dashed): after 5 years, 165 against 132
- Right: the lognormal density of $S_5$; the shaded area is $P(S_5 < S_0) = \Phi(-m\sqrt{5}/\sigma) = 0.34$

## Estimating GBM from Daily Returns

- Data: $n$ log returns $r_t = \ln(S_t/S_{t-1})$ with step $\Delta$ (in years), over $T = n\Delta$ years

$$\hat{\sigma}^2 = \frac{1}{n\Delta}\sum_{t=1}^{n}(r_t - \bar{r})^2, \qquad \hat{m} = \frac{\bar{r}}{\Delta}, \qquad \hat{\mu} = \hat{m} + \tfrac{1}{2}\hat{\sigma}^2$$

  - $\bar{r}$: the sample mean of the $r_t$; each $r_t \sim N(m\Delta, \sigma^2\Delta)$ under GBM

- Standard errors under GBM, with Normal returns that are independent and identically distributed (i.i.d.)

$$\mathrm{SE}(\hat{\sigma}) \approx \frac{\hat{\sigma}}{\sqrt{2n}}, \qquad \mathrm{SE}(\hat{m}) = \frac{\hat{\sigma}}{\sqrt{T}}, \qquad \mathrm{CI}_{95\%}(m) = \hat{m} \pm 1.96\,\mathrm{SE}(\hat{m})$$

  - $\hat{\sigma}$ improves with the number of observations $n$; $\hat{m}$ only with the calendar span $T$
  - years for $\pm h$ around $m$ at 95%: $T = (1.96\,\hat{\sigma}/h)^2$

- Without Normality (independent returns with excess kurtosis $k$, defined on the next slide)

$$\mathrm{SE}(\hat{\sigma}) \approx \hat{\sigma}\sqrt{\frac{k+2}{4n}}$$

  - $k = 0$ gives back $\hat{\sigma}/\sqrt{2n}$; fat tails ($k > 0$) make $\hat{\sigma}$ less precise

## Testing Normality: Jarque–Bera

- Sample skewness and kurtosis of the standardised returns $z_t = (r_t - \bar{r})/s$, $s$: the sample s.d.

$$\hat{S} = \frac{1}{n}\sum_{t=1}^{n} z_t^3, \qquad \hat{K} = \frac{1}{n}\sum_{t=1}^{n} z_t^4, \qquad k = \hat{K} - 3 \text{ (excess kurtosis)}$$

- **Jarque–Bera** Jarque & Bera (1980): $H_0$: Normal returns ($S = 0$, $K = 3$)

$$JB = \frac{n}{6}\left(\hat{S}^2 + \frac{(\hat{K}-3)^2}{4}\right) \sim \chi^2_2 \text{ under } H_0$$

  - reject at 5% if $JB > 5.99$; large values: asymmetry or fat tails

## Testing Independence: ACF and Ljung–Box

- Sample autocorrelation function (ACF) at lag $k$ of a series $x_t$ (here $x_t = |r_t|$; $k$: the lag, not the excess kurtosis)

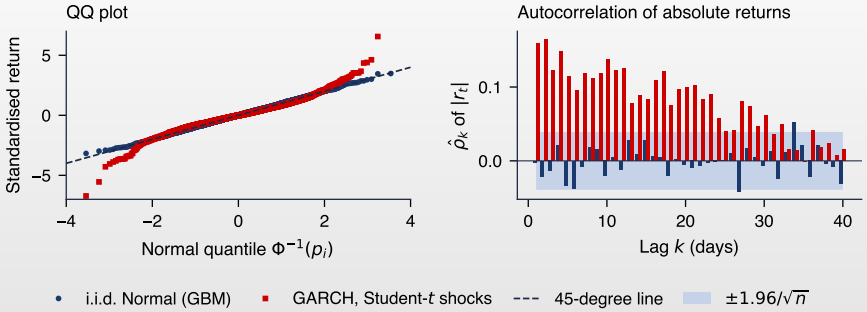
$$\hat{\rho}_k = \frac{\sum_{t=k+1}^{n}(x_t - \bar{x})(x_{t-k} - \bar{x})}{\sum_{t=1}^{n}(x_t - \bar{x})^2}$$

  - for independent data $\hat{\rho}_k \approx N(0, 1/n)$: the band $\pm 1.96/\sqrt{n}$
- **Ljung–Box** Ljung & Box (1978): $H_0$: $\rho_1 = \cdots = \rho_h = 0$

$$Q(h) = n(n+2)\sum_{k=1}^{h}\frac{\hat{\rho}_k^2}{n-k} \;\sim\; \chi^2_h \text{ under } H_0$$

  - $h = 20$: reject at 5% if $Q > \chi^2_{20;\,0.95} = 31.4$; on $|r_t|$ it tests volatility clustering

# QQ Plot and ACF: GBM Against Real-Looking Returns




- **Quantile–quantile (QQ) plot**: the sorted $z_t$ against the Normal quantiles $\Phi^{-1}(p_i)$, $p_i = (i - 0.5)/n$; Normal data lie on the 45-degree line
- Two simulated samples of $n = 2\,500$ days: i.i.d. Normal (GBM) and a Generalized AutoRegressive Conditional Heteroskedasticity (GARCH) model with Student-$t$ shocks
  - GBM: $JB = 0.2$, $Q(20) = 17.5$ on $|r_t|$: neither test rejects
  - GARCH: $JB = 2\,566$ ($\hat{K} = 8.0$), $Q(20) = 688$: both reject; the tails leave the line and $\hat{\rho}_k$ of $|r_t|$ stays positive

## Variance Ratio and the Moving-Block Bootstrap

- **Variance ratio** over $q$ periods: variance of non-overlapping $q$-period sums over $q$ times the one-period variance

$$VR(q) = \frac{\widehat{\mathrm{Var}}\left(r_t + r_{t+1} + \cdots + r_{t+q-1}\right)}{q\,\widehat{\mathrm{Var}}(r_t)}$$

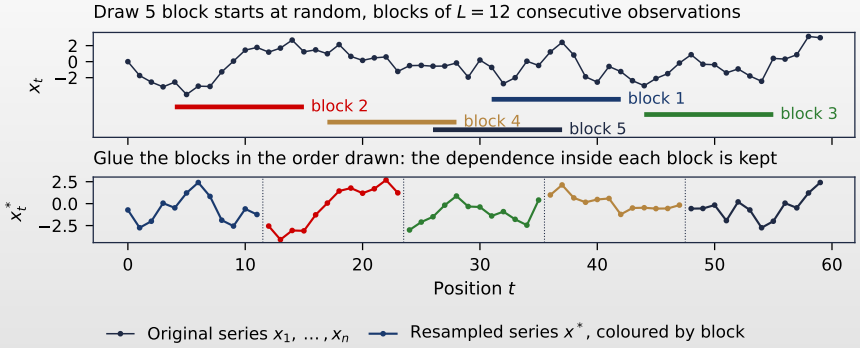
  - uncorrelated returns: $VR = 1$; $VR < 1$: negative autocorrelation (reversals); $VR > 1$: positive autocorrelation

- **Moving-block bootstrap** Künsch (1989): the uncertainty of a statistic $\hat{\theta}$ without a model
  1. draw $\lceil n/L \rceil$ starting points at random; take the blocks $(r_s, \ldots, r_{s+L-1})$ of $L$ consecutive returns
  2. glue them into a series of length $n$ and recompute the statistic: $\hat{\theta}_b^*$
  3. repeat for $b = 1, \ldots, B$

$$\mathrm{SE}^*(\hat{\theta}) = \sqrt{\frac{1}{B-1}\sum_{b=1}^{B}\left(\hat{\theta}_b^* - \bar{\theta}^*\right)^2}, \qquad \mathrm{CI}_{95\%} = \left[q^*_{2.5\%},\ q^*_{97.5\%}\right]$$

  - $\bar{\theta}^*$: the mean of the $\hat{\theta}_b^*$; $q_p^*$: their $p$ quantile; blocks keep the dependence (volatility clustering) shorter than $L$

# The Moving-Block Bootstrap, One Resample




- Top: a persistent series of $n = 60$ observations and five blocks of $L = 12$ consecutive observations, drawn at random (they may overlap)
- Bottom: the resampled series $x^*$; inside each block the original order is kept, the dependence is broken only at the junctions

## Mean Reversion: the OU Process

- **Ornstein–Uhlenbeck (OU) process** Uhlenbeck & Ornstein (1930): pulled towards the level $\theta$ at speed $\kappa > 0$

$$\mathrm{d}x_t = \kappa(\theta - x_t)\,\mathrm{d}t + \sigma\,\mathrm{d}W_t$$

  - above $\theta$ the drift is negative, below it positive; $\sigma$: volatility of the shocks; Vasicek Vasicek (1977): OU for the short interest rate

- Sampled every $\Delta$ years, OU is **exactly** an AR(1)

$$x_{t+\Delta} = a + b\,x_t + \varepsilon_{t+\Delta}, \qquad b = e^{-\kappa\Delta}, \qquad a = \theta(1-b), \qquad \mathrm{Var}(\varepsilon) = s^2 = \frac{\sigma^2(1-b^2)}{2\kappa}$$

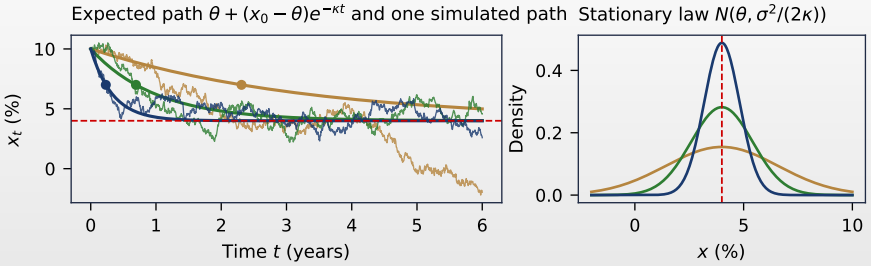
  - $\varepsilon$: i.i.d. Normal errors; the expected path: $\mathbb{E}[x_t \mid x_0] = \theta + (x_0 - \theta)e^{-\kappa t}$

- **Half-life** and the stationary law

$$t_{1/2} = \frac{\ln 2}{\kappa}, \qquad x_\infty \sim N\left(\theta,\ \frac{\sigma^2}{2\kappa}\right)$$

  - $t_{1/2}$: the time for an expected deviation from $\theta$ to halve; a larger $\kappa$: faster reversion, narrower law
  - example: monthly data ($\Delta = 1/12$), $b = 0.95$: $\kappa = -12\ln 0.95 = 0.62$ a year, $t_{1/2} = 1.1$ years

# Simulated OU Paths




- Left: start $x_0 = 10\%$, level $\theta = 4\%$, $\sigma = 0.02$, three speeds $\kappa$: expected paths (thick) and one simulated path each
  - the dots mark the half-lives $\ln 2/\kappa$: half of the distance to $\theta$ is gone
  - OU can become negative (slow path): the Vasicek rate is not bounded below
- Right: the stationary laws $N(\theta, \sigma^2/(2\kappa))$: the faster the reversion, the narrower

## Estimating OU: OLS and the Delta Method

- Ordinary Least Squares (OLS) regression of $x_{t+\Delta}$ on $x_t$ ($n$ transitions), then invert the three relations

$$\hat{b} = \frac{\sum_t (x_t - \bar{x}_0)(x_{t+\Delta} - \bar{x}_1)}{\sum_t (x_t - \bar{x}_0)^2}, \qquad \hat{\kappa} = -\frac{\ln \hat{b}}{\Delta}, \qquad \hat{\theta} = \frac{\hat{a}}{1 - \hat{b}}$$

  - $\bar{x}_0$, $\bar{x}_1$: the means of the regressor and of the response; for a stationary AR(1), $\mathrm{Var}(\hat{b}) \approx (1 - b^2)/n$
- **Delta method**: the SE of a smooth function $g$ of an estimate (first-order Taylor expansion)

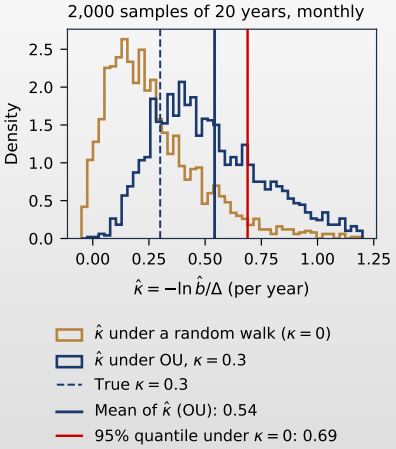
$$g(\hat{\beta}) \approx g(\beta) + g'(\beta)(\hat{\beta} - \beta) \quad \Rightarrow \quad \mathrm{SE}\left(g(\hat{\beta})\right) \approx \left|g'(\hat{\beta})\right| \mathrm{SE}(\hat{\beta})$$

  - Wald interval: $\hat{\kappa} \pm 1.96\,\mathrm{SE}(\hat{\kappa})$; a monotone function maps an interval's ends: $t_{1/2} \in [\ln 2/\kappa_{\max},\ \ln 2/\kappa_{\min}]$ if $\kappa_{\min} > 0$
- In small samples $\hat{b}$ is biased **down**, so $\hat{\kappa}$ is biased **up**

$$\mathrm{bias}(\hat{\kappa}) = \mathbb{E}[\hat{\kappa}] - \kappa \approx \frac{4}{T}$$

  - $T = n\Delta$: the span in years; approximation of Tang & Chen (2009)
  - a random walk ($\kappa = 0$) can look mean-reverting

# The Bias of $\hat{\kappa}$ and a Monte Carlo Test




- 2 000 simulated OU samples of 20 years, monthly, with $\kappa = 0.3$ (blue)
  - mean of $\hat{\kappa}$: 0.54, a bias of 0.24, close to $4/T = 0.20$
- The same under a random walk, $\kappa = 0$ (gold): $\hat{\kappa}$ is positive, mean 0.27
- **Monte Carlo test** of $H_0$: $\kappa = 0$
  - simulate $R$ samples under $H_0$, estimate $\hat{\kappa}^*_1, \ldots, \hat{\kappa}^*_R$

$$p = \frac{1 + \#\{j : \hat{\kappa}^*_j \geq \hat{\kappa}\}}{R + 1}$$

  - $\#$: the number of simulated values at least as large as the observed $\hat{\kappa}$; the $+1$ counts the observed sample itself
  - reject at 5% if $\hat{\kappa}$ exceeds the 95% quantile under $H_0$ (0.69 here), not if it exceeds 0

## Jumps: the Merton Model

- **Poisson process** $N_t$: counts jumps arriving at rate $\lambda$ per year; over $\Delta$ years

$$P(N_\Delta = k) = e^{-\lambda\Delta}\,\frac{(\lambda\Delta)^k}{k!}, \qquad k = 0, 1, 2, \ldots, \qquad \mathbb{E}[N_\Delta] = \mathrm{Var}(N_\Delta) = \lambda\Delta$$

  - $\lambda = 5$: a jump on about $5/252 = 2\%$ of days; no jump in a year with probability $e^{-5}$
- **Merton jump diffusion** Merton (1976): the GBM log return plus the sum of the day's jumps

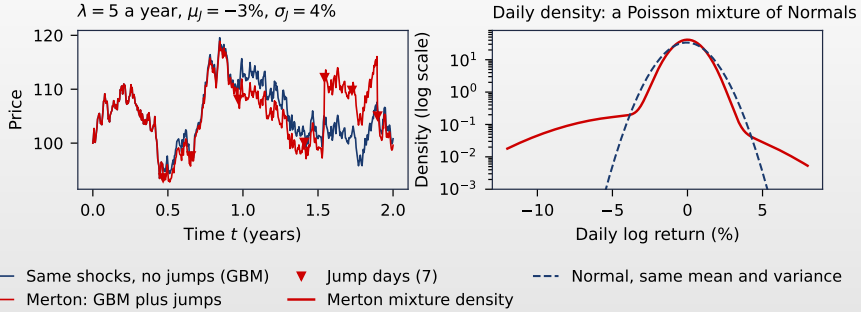
$$r_\Delta = m\Delta + \sigma\sqrt{\Delta}\,Z + \sum_{j=1}^{N_\Delta} Y_j, \qquad Z \sim N(0,1), \qquad Y_j \sim N(\mu_J, \sigma_J^2)$$

  - $\mu_J$: mean jump size (negative: crashes); $\sigma_J$: dispersion of jump sizes; $Z$, $N$, $Y_j$ independent
- Given $k$ jumps the return is Normal, so the density is a Poisson mixture of Normal densities

$$f(r) = \sum_{k=0}^{\infty} P(N_\Delta = k)\;\varphi\left(r;\; m\Delta + k\mu_J,\; \sigma^2\Delta + k\sigma_J^2\right)$$

  - $\varphi(\cdot;\ \text{mean}, \text{variance})$: the Normal density; in practice the sum stops at $k = 12$

# A Simulated Merton Path and the Daily Density




- Left: two years with $\sigma = 15\%$, $\lambda = 5$, $\mu_J = -3\%$, $\sigma_J = 4\%$; the same diffusion shocks without (blue) and with jumps (red)
- Right, log scale: the mixture has a higher peak and far heavier tails than the Normal with the same mean and variance
- **What do you think?** In the left panel, can you spot the 7 jump days among the 504 from the size of the return?
- **Answer**: only some; most jumps look like ordinary large diffusive moves

## Cumulants, Fisher Information, Monte Carlo Error

- **Cumulant generating function** and cumulants $c_k$ of a return $r$

$$K(u) = \ln \mathbb{E}\left[e^{ur}\right], \qquad c_k = K^{(k)}(0), \qquad \text{skewness} = \frac{c_3}{c_2^{3/2}}, \qquad \text{excess kurtosis} = \frac{c_4}{c_2^2}$$

  - $K^{(k)}(0)$: the $k$-th derivative at 0; $c_1$: mean, $c_2$: variance; for independent parts the cumulants add
  - the Normal distribution has $c_k = 0$ for $k \geq 3$: what produces skewness in Merton? (A3)

- **Fisher information** of a parameter $\beta$ in a log-likelihood $\ell(\beta)$, and the SE of the maximum likelihood estimator (MLE) $\hat{\beta}$

$$I(\beta) = -\mathbb{E}\left[\ell''(\beta)\right], \qquad \mathrm{SE}(\hat{\beta}) \approx \frac{1}{\sqrt{I(\hat{\beta})}}$$

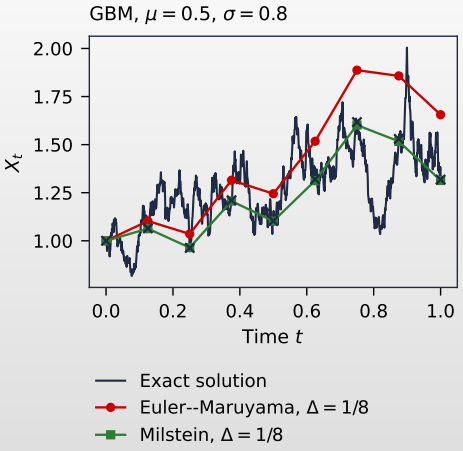
  - a sharply curved log-likelihood: much information, a small SE; with several parameters $I$ is a matrix (the negative expected Hessian)

- **Monte Carlo error** of a simulated probability $\hat{p}$ from $N$ independent paths

$$\mathrm{SE}(\hat{p}) = \sqrt{\frac{\hat{p}(1-\hat{p})}{N}}, \qquad \mathrm{SE}(\bar{X}) = \frac{s_X}{\sqrt{N}}$$

  - $s_X$: the s.d. across paths; the error falls only like $1/\sqrt{N}$: 100 times more paths for one more digit

# Discretisation: Euler–Maruyama and Milstein




- A stochastic differential equation (SDE) $\mathrm{d}X = a(X)\,\mathrm{d}t + b(X)\,\mathrm{d}W$ on a grid with step $\Delta$, $\Delta W \sim N(0, \Delta)$; Euler–Maruyama (EM) and Milstein Higham (2001)

$$\text{EM}: \quad X_{t+\Delta} = X_t + a\Delta + b\,\Delta W$$

$$\text{Milstein}: \quad \text{EM} + \tfrac{1}{2}\,b\,b'\left[(\Delta W)^2 - \Delta\right]$$

  - $a$, $b$, $b'$ evaluated at $X_t$; Milstein adds the Itô term of $(\Delta W)^2 \approx \Delta$

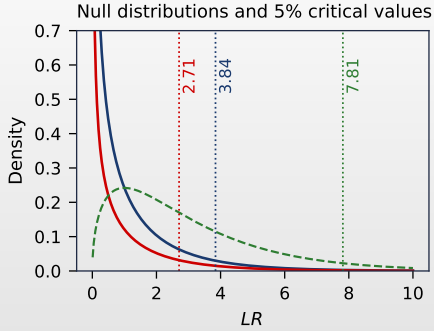
- Errors at $T$ and their orders

$$e_{\mathrm{s}}(\Delta) = \mathbb{E}\left|X_T - X_T^{\Delta}\right| \le C\Delta^{\gamma}$$

$$e_{\mathrm{w}}(\Delta) = \left|\mathbb{E}X_T - \mathbb{E}X_T^{\Delta}\right| \le C\Delta^{\beta}$$

  - $X_T^{\Delta}$: the scheme at $T$; strong: path by path, weak: means; EM: $\gamma = 1/2$, $\beta = 1$; Milstein: $\gamma = 1$
  - the order $\gamma$ is the slope of $\ln e$ on $\ln \Delta$
  - chart: one GBM path, 8 steps: EM ends 0.34 from the exact value, Milstein 0.004

# Likelihood Ratio Tests and a Boundary




- $\chi^2_1$: interior parameter (Wilks)
- $\frac{1}{2}\chi^2_0 + \frac{1}{2}\chi^2_1$ (mass $\frac{1}{2}$ at 0): parameter on the boundary
- $\chi^2_3$

- **Likelihood ratio** (LR): compares the maximised log-likelihoods of a model and of its restricted version

$$LR = 2\left(\ell_{\mathbf{1}} - \ell_{\mathbf{0}}\right) \geq 0$$

  - $\ell_{\mathbf{1}}$: unrestricted (Merton), $\ell_{\mathbf{0}}$: restricted (GBM); large $LR$: the restriction fits much worse
- **Wilks** Wilks (1938): under $H_{\mathbf{0}}$, $LR \sim \chi^{\mathbf{2}}_d$, $d$ = number of restrictions
  - needs the true value inside the parameter space and every parameter identified under $H_{\mathbf{0}}$
  - with $\lambda = 0$ on the boundary the null law changes (chart: red against blue)
- **Parametric bootstrap**: simulate $B$ samples from the fitted $H_{\mathbf{0}}$ model, refit both models, get $LR^*_{\mathbf{1}}, \ldots, LR^*_B$

$$p = \frac{1 + \#\{b : LR^*_b \geq LR\}}{B + 1}$$

## Extension Packs: Measure Change and Heston

- The extension packs (after the lecture) also use three models; the formulas are below
- **Girsanov** Girsanov (1960): GBM under the risk-neutral measure $Q$, with the risk-free rate $r$

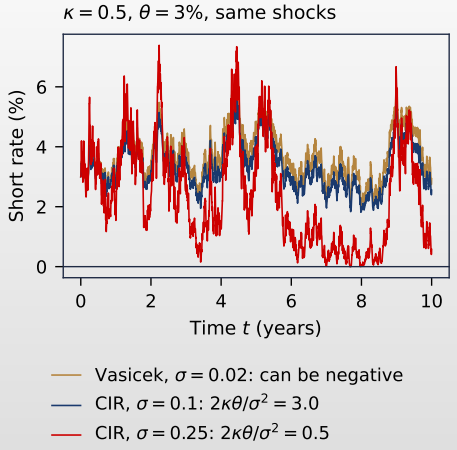
$$\mathrm{d}\tilde{W}_t = \mathrm{d}W_t + \frac{\mu - r}{\sigma}\,\mathrm{d}t, \qquad \mathrm{d}S_t = rS_t\,\mathrm{d}t + \sigma S_t\,\mathrm{d}\tilde{W}_t$$

  - $\tilde{W}$ is a Brownian motion under $Q$; $(\mu - r)/\sigma$: the market price of risk
  - the drift changes from $\mu$ to $r$, the volatility $\sigma$ does not: $P$ forecasts, $Q$ prices
- **Heston** Heston (1993): stochastic volatility, the variance $v_t$ mean-reverts

$$\mathrm{d}S_t = \mu S_t\,\mathrm{d}t + \sqrt{v_t}\,S_t\,\mathrm{d}W_t^{(1)}, \qquad \mathrm{d}v_t = \kappa(\theta - v_t)\,\mathrm{d}t + \xi\sqrt{v_t}\,\mathrm{d}W_t^{(2)}$$

  - $\theta$: long-run variance, $\kappa$: its speed of reversion, $\xi$: volatility of variance
  - $\rho = \mathrm{Corr}(\mathrm{d}W^{(1)}, \mathrm{d}W^{(2)}) < 0$: prices fall when variance rises (the leverage effect)

# Extension Packs: the CIR Process and the Feller Condition




- **Cox–Ingersoll–Ross (CIR)** Cox, Ingersoll & Ross (1985): like OU, but the volatility scales with $\sqrt{r}$

$$\mathrm{d}r_t = \kappa(\theta - r_t)\,\mathrm{d}t + \sigma\sqrt{r_t}\,\mathrm{d}W_t$$

  - near 0 the shocks vanish and the drift $\kappa\theta > 0$ pushes the rate up
  - same conditional mean $\mathbb{E}[r_t \mid r_0]$ as OU (Mean Reversion: the OU Process)

- **Feller condition**: $r_t$ never reaches 0 if

$$2\kappa\theta \geq \sigma^2$$

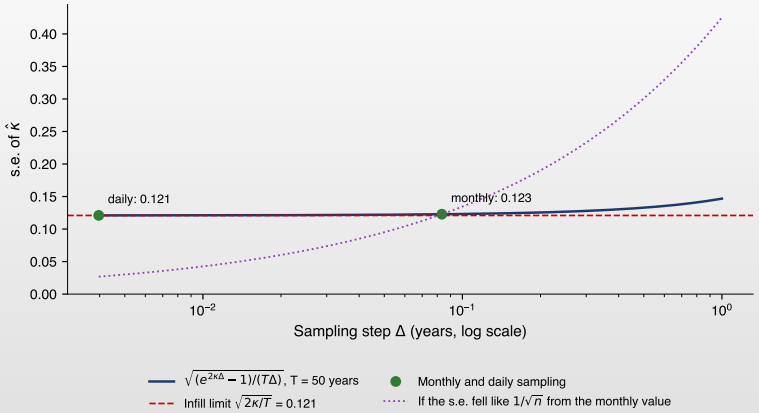
  - chart: same shocks; blue satisfies it ($2\kappa\theta/\sigma^2 = 3$), red does not (0.48) and touches 0; Vasicek (gold) goes negative

- **Chan–Karolyi–Longstaff–Sanders (CKLS)** Chan, Karolyi, Longstaff & Sanders (1992): volatility $\sigma r^\gamma$
  - $\gamma = 0$: Vasicek; $\gamma = 1/2$: CIR; $\gamma$ is estimated from the data

# A1 [Solved]: From an AR(1) to an OU Process — Task

- **Question**: how precisely does a 50-year monthly sample pin down the speed of mean reversion?
- **Inputs**: an AR(1) estimated on $n = 600$ monthly transitions ($T = n\Delta = 50$ years, $\Delta = 1/12$) of a rate in decimals: $x_{t+1} = 0.0012 + 0.97\, x_t + \varepsilon_t$, residual s.d. $\hat{s}_\varepsilon = 0.0025$; the rate is an exactly sampled OU process, with $b$, $a$, $s^2$ as in Mean Reversion: the OU Process
- **Tasks**
  1. Invert the three relations to obtain $\kappa$, $\theta$, $\sigma$ and the half-life $\ln 2/\kappa$.
  2. Use the large-sample $\mathrm{Var}(\hat{b})$ and the delta method of Estimating OU: OLS and the Delta Method to compute the SE of $\hat{\kappa}$.
  3. Build a 95% interval for the half-life by transforming the interval for $\kappa$ (valid only if both ends are positive).
  4. Keep $T = 50$ years but sample daily ($\Delta = 1/252$): compute $b$, the SE of $\hat{\kappa}$ and its limit $\sqrt{2\kappa/T}$.
- **Report**: a table with parameter, formula and value; SE of $\hat{\kappa}$ only; the half-life interval; the daily SE of $\hat{\kappa}$
- **Interpretation**: does daily sampling make $\kappa$ more precise?
- Notebook: section A1

# A1 [Solved]: Mapping, Delta Method and Sampling Frequency (1/2)




MFM_ch11_seminar

# A1 [Solved]: Mapping, Delta Method and Sampling Frequency (2/2)

1. $\kappa = -\ln b/\Delta = -12\ln 0.97 = 0.366$; $\theta = a/(1-b) = 0.0012/0.03 = 4.0\%$; $\sigma = s\sqrt{2\kappa/(1-b^2)} = 0.88\%$; half-life $\ln 2/\kappa = 1.90$ years
2. $\mathrm{SE}(\hat{b}) = \sqrt{(1-0.97^2)/600} = 0.0099$; $\partial\kappa/\partial b = -1/(b\Delta)$, so $\mathrm{SE}(\hat{\kappa}) = 0.0099/(0.97/12) = 0.123$
3. $\kappa \in [0.125, 0.606]$, both ends positive, so the half-life lies in $[\ln 2/0.606, \ln 2/0.125] = [1.14, 5.55]$ years: asymmetric around 1.90
4. Daily: $b = e^{-\kappa/252} = 0.99855$, $n = 12\,600$, SE 0.121; in general $\mathrm{Var}(\hat{\kappa}) \approx (e^{2\kappa\Delta} - 1)/(T\Delta) \to 2\kappa/T$, limit 0.121 (chart)

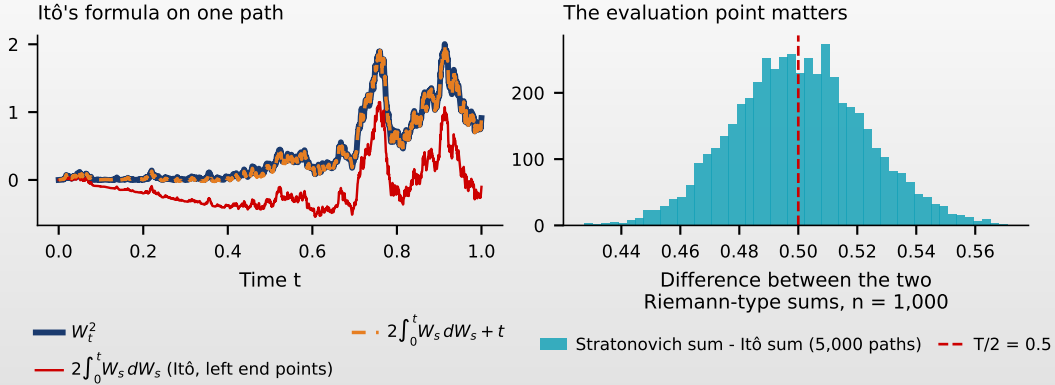
⊡ **Interpretation**: 21 times more observations give the same precision: $\kappa$ is identified by calendar time, not by the number of points; if the interval for $\kappa$ reached zero, the half-life would be unbounded

MFM_ch11_seminar

# A2 [Solved]: Itô Integrals — Task

- **Questions**
  - what does the extra term in $\int_0^T W \,\mathrm{d}W$ measure?
  - why is its mean zero?
- **Inputs**: a standard Brownian motion $W$ on $[0, T]$, $T = 1$ for the numbers; its moments: Itô's Lemma; no data
- **Tasks**
  1. Apply Itô's lemma to $f(x) = x^2$ and show that $\int_0^T W \,\mathrm{d}W = \frac{1}{2}(W_T^2 - T)$.
  2. Compute the mean and the variance of this integral and check the variance with the Itô isometry.
  3. The Stratonovich integral (endpoint averages) equals $\frac{1}{2} W_T^2$: compute its mean.
  4. Apply Itô's lemma to $W_t^3$ and deduce the mean and the variance of $\int_0^1 W^2 \,\mathrm{d}W$.
- **Report**: a table with integral, mean and variance for $\int W \,\mathrm{d}W$, the Stratonovich integral and $\int W^2 \,\mathrm{d}W$
- **Interpretation**: why must an Itô integral have mean zero, while the Stratonovich integral does not?
- Notebook: section A2

# A2 [Solved]: Identity, Moments and a Simulation Check (1/2)




MFM_ch11_ito

# A2 [Solved]: Identity, Moments and a Simulation Check (2/2)

1. $\mathrm{d}(W^2) = 2W\,\mathrm{d}W + \mathrm{d}t$; integrating from 0 to $T$: $W_T^2 = 2\int_0^T W\,\mathrm{d}W + T$, hence $\int_0^T W\,\mathrm{d}W = \frac{1}{2}(W_T^2 - T)$
2. Mean $\frac{1}{2}(\mathbb{E}[W_1^2] - 1) = 0$; variance $\frac{1}{4}\mathrm{Var}(W_1^2) = \frac{1}{4}(3-1) = 0.5$; isometry: $\int_0^1 \mathbb{E}[W_t^2]\,\mathrm{d}t = \int_0^1 t\,\mathrm{d}t = 0.5$
3. Stratonovich: mean $\frac{1}{2}\mathbb{E}[W_1^2] = 0.5 \neq 0$
4. $\mathrm{d}(W^3) = 3W^2\,\mathrm{d}W + 3W\,\mathrm{d}t$; $\mathbb{E}\int_0^1 W^2\,\mathrm{d}W = 0$ and $\mathrm{Var} = \int_0^1 \mathbb{E}[W_t^4]\,\mathrm{d}t = \int_0^1 3t^2\,\mathrm{d}t = 1$
5. Chart (notebook simulation): left, $2\int_0^t W\,\mathrm{d}W + t$ lies approximately on $W_t^2$ (2,000 steps; the exact grid identity has $\sum_i(\Delta W_i)^2$ in place of $t$); right, Stratonovich minus Itô over 5,000 paths has mean 0.500; Itô sums have variance 0.520

- **Interpretation**: the Itô integrand is fixed before the increment, so a strategy cannot anticipate the move; the endpoint average uses the next move and earns $T/2$ on average: half the quadratic variation $[W]_T = T$

MFM_ch11_ito

# A3 [Solved]: Merton: Cumulants and the Information on $\lambda$ — Task

- **Questions**
  - how non-Normal do five modest jumps a year make daily returns?
  - how well can $\lambda$ be learnt?
- **Inputs**: the Merton daily return $r_\Delta$ of Jumps: the Merton Model; $\sigma = 15\%$, $\lambda = 5$ per year, $\mu_J = -3\%$, $\sigma_J = 4\%$, $\Delta = 1/252$; for step 4, $T = 36.7$ years
- **Tasks**
  1. Write the cumulant generating function $K(u) = \ln \mathbb{E}[e^{ur\Delta}]$ and show that the cumulants (Cumulants, Fisher Information, Monte Carlo Error) are $c_k = \lambda\Delta\,\mathbb{E}[Y^k]$ for $k \geq 3$.
  2. Compute $c_2$, split into its diffusion and jump parts, then $c_3$ and $c_4$.
  3. Compute the skewness $c_3/c_2^{3/2}$ and the excess kurtosis $c_4/c_2^2$.
  4. Assume every jump is observed: derive the Fisher information of the Poisson count $N_T$ and the SE of $\hat{\lambda} = N_T/T$ at $\lambda = 5$; repeat at the B5 MLE $\hat{\lambda} = 101.9$ and compare with its MLE SE 10.5.
- **Report**: a table with the diffusion and jump parts of $c_2$, $c_3$, $c_4$, skewness, excess kurtosis; the two standard errors
- **Interpretation**: why is the MLE SE much larger than the SE with observed jumps?
- Notebook: section A3

Back Exercise list

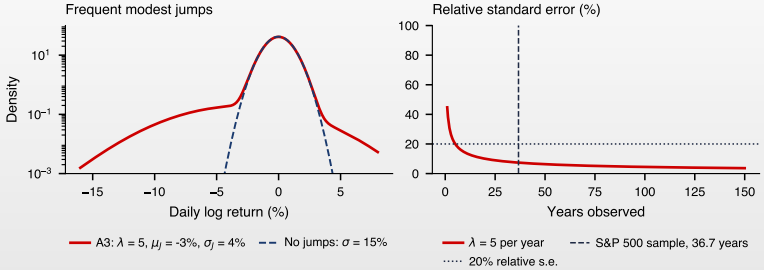
# A3 [Solved]: Cumulants

1. $K(u) = u\,m\Delta + \frac{1}{2}u^2\sigma^2\Delta + \lambda\Delta\left(\mathbb{E}[e^{uY}] - 1\right)$; the $k$-th derivative at 0 gives $c_k = \lambda\Delta\,\mathbb{E}[Y^k]$ for $k \geq 3$
2. $\mathbb{E}[Y^2] = \mu_J^2 + \sigma_J^2$, $\mathbb{E}[Y^3] = \mu_J^3 + 3\mu_J\sigma_J^2$, $\mathbb{E}[Y^4] = \mu_J^4 + 6\mu_J^2\sigma_J^2 + 3\sigma_J^4$ (moments of a Normal variable)
3. Substitute into the table below; skewness $c_3/c_2^{3/2} = -2.07$, excess kurtosis $c_4/c_2^2 = 17.6$

| | Diffusion | Jumps | Total |
|---|---|---|---|
| $c_2$ $(\times 10^{-4})$ | $\sigma^2\Delta = 0.893$ | $\lambda\Delta\mathbb{E}[Y^2] = 0.496$ | 1.389 (35.7% jumps) |
| $c_3$ $(\times 10^{-6})$ | 0 | $-3.393$ | $-3.393$ |
| $c_4$ $(\times 10^{-7})$ | 0 | 3.399 | 3.399 |

- **Interpretation**: five modest jumps a year, one day in 50, make daily returns strongly non-Normal; the diffusion adds variance but no higher cumulants

# A3 [Solved]: Information on $\lambda$




4. Log-likelihood of the Poisson count $\ell(\lambda) = N_T \ln \lambda - \lambda T + \text{const}$, $\ell'' = -N_T/\lambda^2$; Fisher information $I(\lambda) = -\mathbb{E}[\ell'']$, so $I(\lambda) = \mathbb{E}[N_T]/\lambda^2 = T/\lambda$ and $\text{SE}(\hat{\lambda}) = \sqrt{\lambda/T}$
   - Scenario 1, $\lambda = 5$: SE $\sqrt{5/36.7} = 0.37$, a relative SE of 7.4% (red curve)
   - Scenario 2, the B5 fit $\hat{\lambda} = 101.9$: observed jumps would give SE 1.67, but the MLE SE is 10.5, 6.3 times larger

⊡ **Interpretation**: the jumps are latent: the Poisson mixture must separate them from diffusive moves, which destroys most of the information about $\lambda$; the left panel shows how little of the density the jumps occupy

MFM_ch11_seminar

# B1 [Solved]: GBM on the S&P 500 — Task

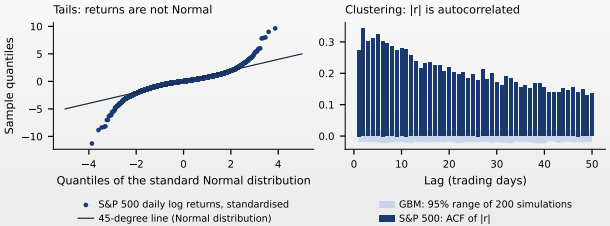
- **Questions**
  - which GBM parameter do 36 years of daily data pin down?
  - which GBM assumption fails?
- **Data**: S&P 500 closes, `GSPC.INDX` (field `close`), 2 January 1990 – 18 September 2026, weekdays, stale quotes removed; $n = 9\,245$ daily log returns, $T = 36.7$ years, $\Delta t = 1/252$
- **Tasks**
  1. Compute $\hat{\sigma}$, $\hat{m}$, $\hat{\mu}$, their GBM SE and a 95% CI for $m$ with the formulas of Estimating GBM from Daily Returns.
  2. Test Normality with Jarque–Bera and independence with Ljung–Box on $|r_t|$ (20 lags), both at 5%, and plot the QQ plot and the ACF of $|r_t|$.
  3. Compute the variance ratio of non-overlapping 20-day sums (blocks from the first return; the incomplete last block dropped) to 20 times the daily variance.
  4. Recompute the SE without GBM: the excess-kurtosis formula of Estimating GBM from Daily Returns and a moving-block bootstrap (Variance Ratio and the Moving-Block Bootstrap; blocks of 63 days, 999 samples).
- **Report**: estimates with GBM and bootstrap SE (percentage points), test statistics with critical values and decisions, two charts
- **Interpretation**
  - which parameter is known well?
  - which one is known badly?
  - can $\hat{\sigma}$ still be used for option pricing?
- Notebook: section B1

Back Exercise list

## B1 [Solved]: Estimates Step by Step

1. $n = 9\,245$, $T = 36.7$: $\hat{\sigma} = 18.0\%$, $\mathrm{SE} = \hat{\sigma}/\sqrt{2n} = 0.13$ percentage points (p.p.); $\hat{m} = 8.3\%$, $\mathrm{SE} = \hat{\sigma}/\sqrt{T} = 3.0$ p.p.
   - $\hat{\mu} = \hat{m} + \hat{\sigma}^2/2 = 10.0\%$; GBM 95% CI for $m$: $[2.5\%, 14.2\%]$; $\pm 1$ point would need $(1.96\,\hat{\sigma}/0.01)^2 = 1\,243$ years
4. Without GBM, SE of $\hat{\sigma}$: 0.34 p.p. with kurtosis $k = 10.9$, 1.15 p.p. by block bootstrap, 8.7 times the GBM value; $m$: block SE 2.5 p.p., CI $[3.5\%, 13.3\%]$ (chart: GBM versus block-bootstrap uncertainty)

# B1 [Solved]: Diagnostics

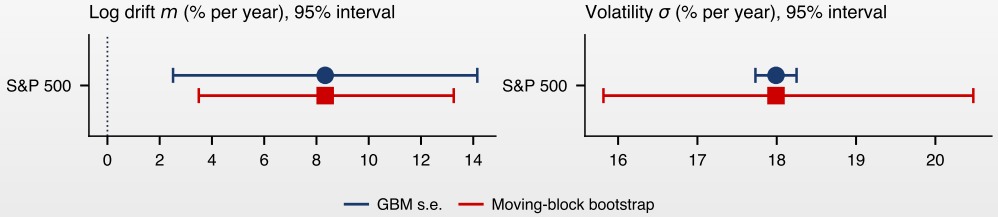



2. QQ plot: tails far beyond the 45-degree line; Jarque–Bera 45 805 against the 5% critical value 5.99: Normality rejected
   - ACF of $|r_t|$: 0.27 at lag 1, far outside the GBM range; Ljung–Box 13 381 against $\chi^2_{20,0.95} = 31.4$: independence rejected
3. Variance ratio 0.75, below 1: a descriptive sign of mild mean reversion in 20-day sums; no uncertainty attached, so no test

⊡ **Interpretation**: both stylised facts of Chapter 1 reject GBM; $\hat{\sigma}$ is an average of a time-varying volatility, a benchmark for pricing, not a calibration

MFM_ch11_gbm_vs_data

# B1 [Solved]: GBM versus Block-Bootstrap Uncertainty




- Blue: GBM intervals; red: moving-block bootstrap; S&P 500 (B1), on the sample used above
- Volatility: the bootstrap interval is several times wider than the GBM one, yet still narrow relative to $\hat{\sigma}$; drift: both intervals are wide
- **Interpretation**: $\sigma$ is known well even without GBM, the drift badly under any method: a long-span problem

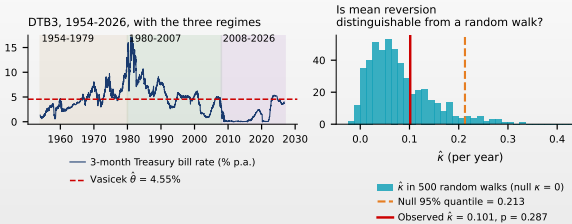
MFM_ch11_seminar

# B2 [Solved]: Vasicek and the Bias of $\hat{\kappa}$ — Task

- **Question**: does the 3-month Treasury bill rate really revert to a mean, once the bias of $\hat{\kappa}$ is accounted for?
- **Data**: FRED series DTB3 (3-month US Treasury bill rate), daily, 4 January 1954 – 18 September 2026, divided by 100 (decimals); $n = 18\,170$ observations, $T = 72.1$ years, $\Delta t = 1/252$
- **Tasks**
  1. Regress $x_{t+1}$ on $x_t$ by OLS (the exact AR(1) of A1), recover $\kappa$, $\theta$, $\sigma$ with delta-method SE, a 95% CI for $\kappa$ and the half-life Vasicek (1977).
  2. Bias experiment: simulate 500 exact OU paths of the same length, started from the stationary law, with the estimates as the truth, and re-estimate $\kappa$ on each.
  3. Compute the bias (mean estimate minus truth), compare it with $4/T$ Tang & Chen (2009), and give the bias-corrected $\kappa$ with an approximate interval.
  4. Null experiment: simulate 500 random walks ($\kappa = 0$, no drift) of the same length from $x_0$, with the estimated residual s.d.; $p$: the Monte Carlo $p$-value of The Bias of $\hat{\kappa}$ and a Monte Carlo Test with $R = 500$.
- **Report**: estimates with SE, the two simulated distributions and the corrected values
- **Interpretation**
  - is mean reversion significant before and after the correction?
  - what does the bias do to long-maturity rate risk?
- Notebook: section B2

Back Exercise list

# B2 [Solved]: Estimates and the Random-Walk Null




1. $\hat{b} = 0.99960$, so $\hat{\kappa} = -\ln \hat{b}/\Delta t = 0.101$ (SE 0.051), Wald 95% CI $[0.001, 0.202]$; $\hat{\theta} = 4.55\%$ (SE 1.55), $\hat{\sigma} = 1.33\%$ (SE 0.007), half-life 6.8 years
4. Null (right panel): under a random walk $\hat{\kappa}$ averages 0.082, not 0, and its 95% quantile is 0.213; the observed 0.101 has $p = 0.287$: a random walk is not rejected

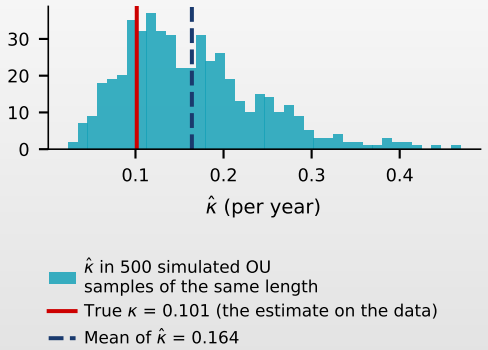
⊡ **Interpretation**: the Wald CI barely excludes zero because it ignores the skewed small-sample law of $\hat{\kappa}$; the three regimes (left) also give very different local estimates

MFM_ch11_seminar

# B2 [Solved]: Bias of $\hat{\kappa}$ and Correction

2. 500 exact OU samples of 72.1 years with $\kappa = 0.101$: mean $\hat{\kappa} = 0.164$ (s.d. 0.078), bias 0.063, close to $4/T = 0.055$; 78% of the estimates exceed the truth
3. Corrected $\kappa = 0.101 - 0.063 = 0.039$, approximate interval $0.039 \pm 1.96 \cdot 0.078 = [-0.114, 0.192]$: it contains zero, so no finite half-life can be reported with confidence

⊡ **Interpretation**: the data are compatible with slow reversion but cannot exclude a random walk; with $\kappa^Q = \kappa$, the upward bias makes Vasicek understate long-maturity rate risk



# C1 [Proposed]: Critique an AI Answer — Task

- **Question**: a student asked an assistant: "S&P 500 as a GBM with $\mu = 8\%$, $\sigma = 20\%$: what is $P(S_1 < S_0)$? Check by simulation."
- **The answer (numbered lines)**

T1. $\mathrm{d}S = \mu S\,\mathrm{d}t + \sigma S\,\mathrm{d}W$
T2. Chain rule: $\mathrm{d}\ln S = \mathrm{d}S/S = \mu\,\mathrm{d}t + \sigma\,\mathrm{d}W$, so $\ln(S_T/S_0) \sim N(\mu T, \sigma^2 T)$
T3. $P(S_T < S_0) = \Phi(-\mu\sqrt{T}/\sigma) = \Phi(-0.40) = 0.345$
T4. $\mathbb{E}[S_T] = S_0 e^{\mu T}$: on average 8.3% higher
T5. "The simulation gives 0.345 and confirms the formula"

```
C1  import numpy as np
C2  mu, sigma, T, n, N = 0.08, 0.20, 1.0, 252, 20_000
C3  dt = T / n
C4  rng = np.random.default_rng(1)
C5  Z = rng.standard_normal((N, n))
C6  lnS = np.cumsum((mu - 0.5*sigma**2)*dt
C7                  + sigma*dt*Z, axis=1)
C8  print((lnS[:, -1] < 0).mean())
```

- **Tasks**
  1. Find three independent errors. Count an inherited error once. Use `T1`-`T5` for text and `C1`-`C8` for code lines (not exercise C1); identify the correct lines too.
  2. For each, give a check that detects it: a formula, a test or a run of the code.
  3. Run the original and the corrected code with seed 1 and compute the Monte Carlo SE
- **Report**: a table: error, line, check, correction; log it in `AI_ERRORS.md`. Model: A2, B1

Back Exercise list

## Calendar Time and Thin Trading

- BET observations follow trading days; Bitcoin observations follow calendar days. The same number of returns can therefore span different lengths of time.
- Thin trading means infrequent transactions: some index constituents retain stale prices while others update. This can induce serial correlation in index returns, but serial correlation alone does not identify the cause.
- In B3 apply the B1 method to BET and Bitcoin and compare blocks of the same calendar length.

# B3 [Proposed]: GBM on BET and Bitcoin — Task

- **Questions**
  - which market is further from GBM?
  - is Bitcoin's drift significantly positive?
- **Data**: BET (`BET`, close, 2000–2026, weekdays, $\Delta t = 1/252$, $n = 6\,670$); Bitcoin (`BTC-USD.CC`, close, 18 September 2014 – 18 September 2026, 7 days a week, $\Delta t = 1/365$, $n = 4\,384$); different sample periods
- **Tasks**
  1. For each market, compute $\hat{\sigma}$ and $\hat{m}$ with GBM and block-bootstrap 95% CIs (blocks of 63 observations) and the years needed for $\pm 1$ point on $m$.
  2. Compute Jarque–Bera, Ljung–Box on $|r_t|$ (20 lags, 5%) and the 20-observation variance ratio, and draw the QQ and ACF charts.
  3. Convert 20 and 63 observations into calendar time for each market. Should block lengths be matched in observations or in calendar time?
- **Report**: one table (rows: $\hat{\sigma}$, $\hat{m}$ with both CIs, years, kurtosis, Jarque–Bera statistic, Ljung–Box, lag-1 ACF, variance ratio; columns: BET, Bitcoin) and the charts
- **Interpretation**
  - what could a variance ratio different from 1 suggest for BET?
  - what else would you need to call it thin trading?
- Model: B1; Notebook: section B3

Back Exercise list

## Checks and Project Menu

- Before you draw a conclusion, check
  - the $-\sigma^2/2$ term (price drift versus log drift)
  - units: annual parameters with the right $\Delta t$ (Bitcoin: $1/365$)
  - no $\chi^2$ critical values for a parameter on the boundary
  - $\hat{\kappa}$ from a short span is biased upwards
- Project menu: C2 (which model fits BET and Bitcoin) or C3 (rough volatility of SPY), in the extension packs

# B4 [Solved]: Orders of Convergence — Task

- **Questions**
  - how fast do the Euler–Maruyama and Milstein schemes converge?
  - what limits a Monte Carlo check?
- **Data**: $\mathrm{d}X = 2X\,\mathrm{d}t + X\,\mathrm{d}W$, $X_0 = 1$, $T = 1$, exact solution $X_T = \exp(1.5 + W_T)$ Higham (2001); 20,000 paths on a fine grid of $2^{11}$ steps; a coarse increment is the sum of $2^{11}\Delta t$ fine increments, so every scheme and the exact value use the same path
- **Tasks**
  1. For $\Delta t = 2^{-11}, \ldots, 2^{-5}$, store the path-by-step array of absolute errors $|X_T - X_T^{\Delta t}|$ for both schemes and average over paths.
  2. Regress ln(error) on $\ln \Delta t$ (the global slope); get its SE and 95% CI by resampling whole paths, all step sizes jointly (1,000 samples).
  3. Compute the weak error of Euler–Maruyama exactly, from $\mathbb{E}[X_T^{\Delta t}] = (1 + 2\Delta t)^{1/\Delta t}$, and by the Monte Carlo mean; estimate both slopes.
- **Report**: a table of the four slopes with bootstrap SE and CI, and the log-log chart
- **Interpretation**
  - do the slopes agree with the orders 1/2 and 1?
  - why does the Monte Carlo weak error give a wrong slope?
- Notebook: section B4

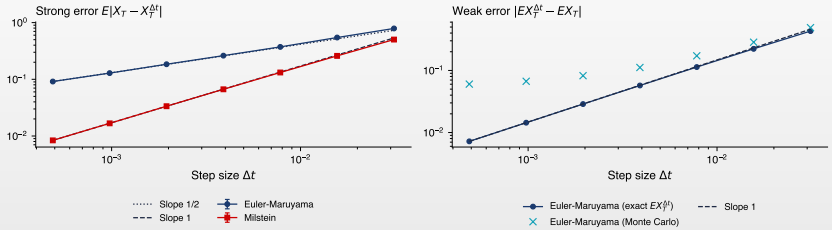
Back Exercise list

# B4 [Solved]: Slopes Step by Step

| | Global slope | SE | 95% CI |
|---|---|---|---|
| Strong, Euler–Maruyama | 0.518 | 0.003 | [0.512, 0.525] |
| Strong, Milstein | 0.986 | 0.001 | [0.984, 0.987] |
| Weak, exact mean | 0.985 | – | deterministic |
| Weak, Monte Carlo | 0.511 | 0.269 | [0.247, 1.289] |

1. Global strong slopes are close to 1/2 and 1; their CIs measure Monte Carlo noise only and can exclude 1/2 and 1, because coarse steps are pre-asymptotic
2. The local slopes between the two finest steps, 0.492 (Euler) and 0.998 (Milstein), are the ones that approach the limiting orders
3. Weak, exact mean: a deterministic slope close to 1 (local 0.998), with no sampling interval

# B4 [Solved]: The Chart and the Monte Carlo Floor




1. Left: strong errors with reference slopes 1/2 and 1; halving $\Delta t$ halves the Milstein error but cuts Euler only by $1 - 2^{-\mathbf{1/2}} = 29\%$
2. Right: the Monte Carlo SE of the mean, $e^{\mathbf{2}}\sqrt{e-1}/\sqrt{20\,000} = 0.068$, exceeds the exact weak error 0.0072 at $2^{-\mathbf{11}}$, so the crosses flatten
3. For this direct-mean estimator, resolving the error at 95% needs $(1.96 \cdot 9.69/0.0072)^{\mathbf{2}} \approx 6\,937\,381$ paths ($9.69 = e^{\mathbf{2}}\sqrt{e-1}$, the s.d. of $X_T$)

⊡ **Interpretation**: for expected values the Monte Carlo sample size, not the time step, is the binding constraint

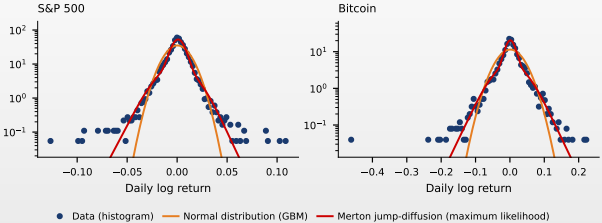
MFM_ch11_convergence

# B5 [Solved]: Merton and a Bootstrap Likelihood-Ratio Test — Task

- **Question**: do daily S&P 500 returns reject GBM in favour of the Merton jump-diffusion?
- **Data**: the S&P 500 sample of B1 ($n = 9\,245$ daily log returns, $\Delta t = 1/252$)
- **Tasks**
  1. Fit $(m, \sigma, \lambda, \mu_J, \sigma_J)$ by maximising $\sum_t \ln f(r_t)$, with the Poisson-mixture density $f$ of Jumps: the Merton Model (log-sum-exp), on $\sigma \geq 5\%$, from six starting points, with Hessian SE Merton (1976)
  2. Compute the likelihood-ratio statistic $LR = 2(\ell_{\text{Merton}} - \ell_{\text{GBM}})$ on this restricted space ($\ell$: maximised log-likelihood; large $LR$ favours Merton).
  3. Simulate 200 GBM samples of the same length from the fitted GBM, refit both models with the same restriction and starts, and compute the 95% critical value and the bootstrap $p$-value of Likelihood Ratio Tests and a Boundary ($B = 200$).
- **Report**: all five parameters with SE, the LR statistic, the bootstrap and $\chi^2_3$ critical values, $p$
- **Interpretation**
  - why is $\chi^2_3$ not valid here?
  - does rejecting GBM prove that prices jump?
- Notebook: section B5

Back Exercise list

# B5 [Solved]: Fitted Parameters and Density




1. $\hat{m} = 24.8\%$ (2.6), $\hat{\sigma} = 9.44\%$ (0.32), $\hat{\lambda} = 101.9$ (10.5) per year, $\hat{\mu}_J = -0.162\%$ (0.034), $\hat{\sigma}_J = 1.44\%$ (0.06); SE in brackets
   - The total log drift is $m + \lambda\mu_J = 24.8\% - 16.5\% = 8.3\%$ per year, consistent with the GBM estimate. Here $m$ describes the diffusion part alone.
   - Implied skewness $-0.31$ (data $-0.36$), excess kurtosis 3.71 (data 10.9): the fitted density (chart, left) follows the centre, but two thirds of the kurtosis is missed

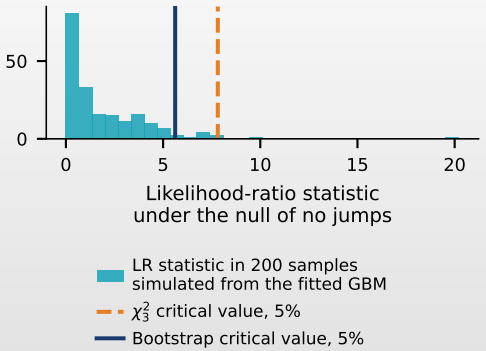
⊡ **Interpretation**: about 100 small "jumps" a year stand in for changing volatility, not for rare crashes

MFM_ch11_merton_density

# B5 [Solved]: Bootstrap Null Distribution and Decision

2. LR $= 2\,401$ ($\sigma \geq 5\%$ not binding); bootstrap 95% critical value 5.62 against $\chi^2_{3,0.95} = 7.81$
3. 0 of the 200 bootstrap values exceed the LR, so $\hat{p} = (1+0)/201 = 0.0050$, the smallest value this design can give (95% Monte Carlo upper bound 0.015): GBM is rejected at 5%

- Under $H_0$, $\lambda = 0$ is on the boundary and $\mu_J$, $\sigma_J$ are unidentified: Wilks' theorem fails, hence the bootstrap
- **Interpretation**: the test rejects i.i.d. Normal returns; it does not identify discontinuous paths, since stochastic volatility would also reject GBM (Chapter 11, mini-case on jumps)




MFM_ch11_seminar

# Extension Packs: What Each One Needs

| Exercise | Question | Type | Model |
|---|---|---|---|
| A4 | GBM under $P$ and $Q$: which probability forecasts? which one prices? | Solved | – |
| A5 | Siegel's paradox: can both sides expect to gain? | Proposed | A4 |
| A6 | the Vasicek bond and yield curve by Feynman–Kac | Solved | – |
| A7 | CIR: moments, transition law, Feller condition | Solved | – |
| A8 | what does the VIX identify in Heston? | Proposed | A7 |
| A9 | rare, large jumps: can crash risk be learnt? | Proposed | A3 |
| B2 Extended | exact CIR, Euler and CKLS on DTB3 | Solved | – |
| B5 Extended | is the Merton likelihood bounded? | Solved | – |
| B6, B7 | OU for the VIX; jumps in Bitcoin | Proposed | B2; B5 |
| B8, B8 Extended | Heston from the VIX; the affine map and realised variance | Proposed | B2, B5; A8 |
| B9 | nonparametric diffusion of the T-bill rate | Proposed | B2 Extended |
| C2, C3 | project seeds: models for BET and Bitcoin; rough and persistent SPY volatility | Proposed | B1, B5, B8; B2, B4 |

# A4 [Solved]: GBM under $P$ and $Q$ (Girsanov) — Task

- **Questions**
  - which probability of a ten-year loss is a forecast?
  - which one is a pricing weight?
- **Inputs**: under the real-world measure $P$, $\mathrm{d}S = \mu S\,\mathrm{d}t + \sigma S\,\mathrm{d}W$ with $\mu = 8\%$, $\sigma = 20\%$; risk-free rate $r = 3\%$; $T = 10$ years; $Q$ is the measure under which $e^{-rt}S_t$ is a martingale; $\Phi$: standard Normal distribution function; no data
- **Tasks**
  1. Find the market price of risk $\theta$ that removes the drift of $e^{-rt}S_t$ and write the density $Z_T = \mathrm{d}Q/\mathrm{d}P$ Girsanov (1960).
  2. Using the moment $\mathbb{E}[e^{aW_T}]$ of Itô's Lemma, show that $\mathbb{E}^P[Z_T S_T] = S_0 e^{rT}$; state whether $\sigma$ changes under $Q$.
  3. Compute $P(S_{10} < S_0)$ and $Q(S_{10} < S_0)$ separately, from the lognormal law under each measure.
- **Report**: $\theta$, $Z_T$, $\mathbb{E}^Q[S_{10}]/S_0$ and the two probabilities
- Notebook: section A4

Back Exercise list

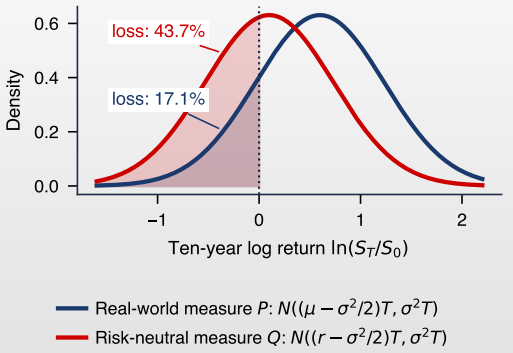
## A4 [Solved]: Girsanov Step by Step

1. Product rule: $\mathrm{d}(e^{-rt}S_t) = e^{-rt}S_t\,[(\mu - r)\,\mathrm{d}t + \sigma\,\mathrm{d}W_t]$; with $\mathrm{d}W = \mathrm{d}\tilde{W} - \theta\,\mathrm{d}t$ the drift is $\mu - r - \sigma\theta = 0$ for $\theta = (0.08 - 0.03)/0.20 = 0.25$
   - Girsanov with constant $\theta$ (Novikov holds): $Z_T = \exp(-\theta W_T - \frac{1}{2}\theta^2 T)$
2. $\mathbb{E}^P[Z_T S_T] = S_0 e^{(\mu - \sigma^2/2 - \theta^2/2)T}\,\mathbb{E}[e^{(\sigma-\theta)W_T}] = S_0 e^{(\mu-\sigma^2/2-\theta^2/2+(\sigma-\theta)^2/2)T} = S_0 e^{(\mu-\sigma\theta)T} = S_0 e^{rT}$, so $\mathbb{E}^Q[S_{10}]/S_0 = e^{0.3} = 1.350$
   - Under $Q$, $\mathrm{d}S = rS\,\mathrm{d}t + \sigma S\,\mathrm{d}\tilde{W}$: $\sigma$ is unchanged, since $[\ln S]_T = \sigma^2 T$ is a property of the path

# A4 [Solved]: Two Loss Probabilities

3. Log drifts $m_P = \mu - \sigma^2/2 = 6\%$ and $m_Q = r - \sigma^2/2 = 1\%$; $\ln(S_{10}/S_0) \sim N(10m, 0.4)$ under each measure
   - $P(S_{10} < S_0) = \Phi(-0.06\sqrt{10}/0.20) = \Phi(-0.949) = 17.1\%$; $Q(S_{10} < S_0) = \Phi(-0.01\sqrt{10}/0.20) = \Phi(-0.158) = 43.7\%$ (shaded areas)

- **Interpretation**: the $P$-probability forecasts a loss; the $Q$-probability is a pricing weight that embeds the risk premium and so overweights bad outcomes; the two densities differ only by their drift




MFM_ch11_seminar

## Numéraire and Foreign Assets

- A numéraire $N_t > 0$ is a traded asset used as the unit of account. Under its pricing measure $Q^N$, traded prices divided by $N_t$ are martingales (under the usual integrability assumptions).
- At a constant rate $r$, a money-market account is $B_t = e^{rt}$. If $X_t$ is lei per euro, a euro deposit is worth $X_t B_t^{EUR}$ lei: it is a traded asset expressed in lei.
- Siegel's paradox: averaging an exchange rate and inverting it do not commute. If $X_0 = 1$ and $X = 0.5$ or $2$ with equal probability, $\mathbb{E}[X] = \mathbb{E}[1/X] = 1.25$, while $1/\mathbb{E}[X] = 0.8$. The positive means of both percentage changes do not create arbitrage.

# A5 [Proposed]: Siegel's Paradox and the Numéraire — Task

- **Question**: if the EUR/RON rate (lei per euro; EUR: euro, RON: Romanian leu) follows GBM, is the drift of the inverse rate RON/EUR (euros per leu) just the opposite?
- **Inputs**: $X_t$ = lei per euro, GBM with $\mu = 2\%$, $\sigma = 4.4\%$ under $P$; $Y_t = 1/X_t$; for steps 3–4: constant rates $r_{RON}$, $r_{EUR}$ (symbolic), money-market accounts $B_t^{RON} = e^{r_{RON}\,t}$ in lei and $B_t^{EUR} = e^{r_{EUR}\,t}$ in euro, worth $X_t B_t^{EUR}$ in lei; given numbers, no data
- **Tasks**
  1. Apply Itô's lemma to $f(x) = 1/x$ and write $\mathrm{d}Y_t$.
  2. Compute the percentage drift of $Y$ to four decimals and compare it with $-\mu$; then compare the log drifts of $X$ and $Y$.
  3. Show that $F = X_0 e^{(r_{RON} - r_{EUR})T} = \mathbb{E}^{Q^{RON}}[X_T]$, using that $X_t B_t^{EUR}/B_t^{RON}$ is a $Q^{RON}$-martingale.
  4. Optional: with the numéraire density $\mathrm{d}Q^{EUR}/\mathrm{d}Q^{RON} = X_T B_T^{EUR}/(X_0 B_T^{RON})$, show that $\mathbb{E}^{Q^{EUR}}[1/X_T] = 1/F$.
- **Report**: the drifts and the two expectations
- **Interpretation**: why is there no arbitrage when both a RON and a euro investor seem to expect a gain?
- Model: A4; Notebook: section A5

Back Exercise list

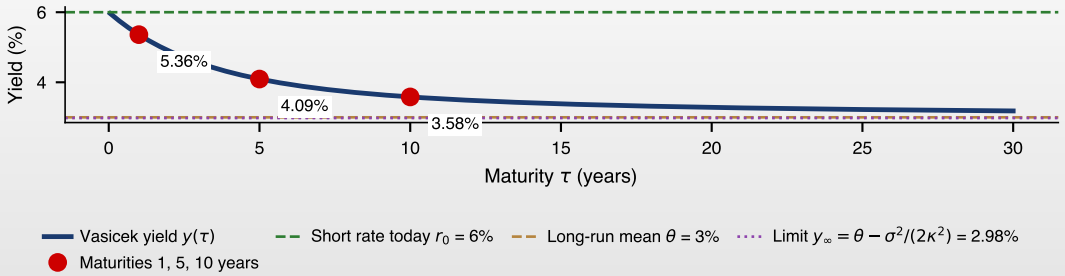
# A6 [Solved]: The Vasicek Bond by Feynman–Kac — Task

- **Question**: what yield curve does a Vasicek short rate imply today?
- **Inputs**: under $Q$, $\mathrm{d}r = \kappa(\theta - r)\,\mathrm{d}t + \sigma\,\mathrm{d}\tilde{W}$, $\kappa = 0.5$, $\theta = 3\%$, $\sigma = 1\%$ (risk premium already in $\theta$), $r_0 = 6\%$ Vasicek (1977); a zero-coupon bond pays 1 at $T$; $p(t, r; T)$ is its price, $\tau = T - t$, and a prime denotes $\mathrm{d}/\mathrm{d}\tau$
- **Tasks**
  1. Write the Feynman–Kac partial differential equation (PDE) and terminal condition for $p(t, r; T) = \mathbb{E}^Q[e^{-\int_t^T r_s \mathrm{d}s} \mid r_t = r]$ Kac (1949).
  2. Insert $p = \exp(A(\tau) - B(\tau)\, r)$, collect terms in $r$ and solve for $B$ and $A = \int_0^\tau [-\kappa\theta B(s) + \sigma^2 B(s)^2/2]\,\mathrm{d}s$.
  3. Compute $y(\tau) = (B r_0 - A)/\tau$ for $\tau = 1, 5, 10$, the limit $y_\infty$, and plot the curve.
- **Report**: $B(5)$, $A(5)$, three yields, $y_\infty$ and the chart
- **Interpretation**
  - why does the curve slope down?
  - why is $y_\infty$ below $\theta$?
- Notebook: section A6

## A6 [Solved]: From the PDE to Two Ordinary Differential Equations

1. Feynman–Kac: $p_t + \kappa(\theta - r)p_r + \frac{1}{2}\sigma^2 p_{rr} - rp = 0$, $p(T, r; T) = 1$
2. With $p = e^{A - Br}$ and $\partial_t = -\mathrm{d}/\mathrm{d}\tau$: $p_t = (-A' + B'r)p$, $p_r = -Bp$, $p_{rr} = B^2 p$; dividing by $p$: $-A' + B'r - \kappa(\theta - r)B + \frac{1}{2}\sigma^2 B^2 - r = 0$ for every $r$
   - Terms in $r$: $B' = 1 - \kappa B$, $B(0) = 0$, so $B(\tau) = (1 - e^{-\kappa\tau})/\kappa$
   - Constants: $A' = -\kappa\theta B + \frac{1}{2}\sigma^2 B^2$; integrating with $\int_0^\tau B = (\tau - B)/\kappa$ and $\int_0^\tau B^2 = (\tau - B)/\kappa^2 - B^2/(2\kappa)$: $A(\tau) = \left(\theta - \frac{\sigma^2}{2\kappa^2}\right)(B - \tau) - \frac{\sigma^2 B^2}{4\kappa}$

# A6 [Solved]: Numbers and the Yield Curve




3. $B(5) = (1 - e^{-\mathbf{2.5}})/0.5 = 1.836$; $A(5) = (0.03 - 0.0002)(1.836 - 5) - 0.0001 \cdot 1.836^{\mathbf{2}}/2 = -0.0945$
   - $y(1) = 5.36\%$, $y(5) = 4.09\%$, $y(10) = 3.58\%$; $y_\infty = \theta - \sigma^{\mathbf{2}}/(2\kappa^{\mathbf{2}}) = 2.98\%$

⊡ **Interpretation**: $r_0 = 6\%$ is above $\theta$, so expected rates fall and the curve slopes down; $y_\infty < \theta$ by convexity (Jensen): rate uncertainty raises bond prices and lowers long yields

MFM_ch11_seminar

# A7 [Solved]: CIR: Moments, Transition Law and the Feller Condition — Task

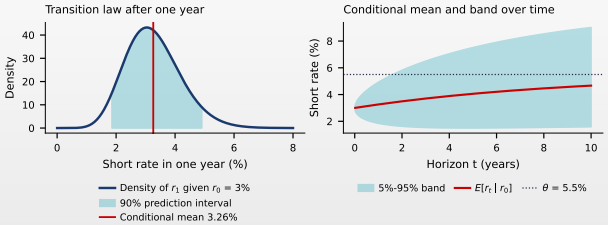
- **Questions**
  - where can a CIR short rate be in one year?
  - can it reach zero?
- **Inputs**: the CIR process of Extension Packs: the CIR Process and the Feller Condition with $\kappa = 0.11$, $\theta = 5.5\%$, $\sigma = 0.056$ (close to the exact CIR fit of B2 Extended) Cox, Ingersoll & Ross (1985); $r_0 = 3\%$, $t = 1$; transition $2c\, r_t \mid r_0 \sim \chi^2_\nu(\delta)$ with $c = 2\kappa/(\sigma^2(1 - e^{-\kappa t}))$, $\nu = 4\kappa\theta/\sigma^2$, noncentrality $\delta = 2c\, r_0 e^{-\kappa t}$
- **Tasks**
  1. Apply Itô's lemma to $e^{\kappa t} r_t$ and derive $\mathbb{E}[r_t \mid r_0]$.
  2. Apply Itô's lemma to $r_t^2$, solve the ordinary differential equation for $m_2(t) = \mathbb{E}[r_t^2]$ with an integrating factor and derive $\mathrm{Var}(r_t \mid r_0)$.
  3. Compute $c$, $\nu$, $\delta$ and the 5% and 95% quantiles of $r_1$ (`scipy.stats.ncx2.ppf`, then divide by $2c$).
  4. Show that $2\kappa\theta \geq \sigma^2 \Leftrightarrow \nu \geq 2$ and evaluate the Feller ratio.
- **Report**: mean, s.d., $c$, $\nu$, $\delta$, the 90% conditional prediction interval (not a parameter CI), the Feller ratio
- **Interpretation**: why is the interval asymmetric around the mean?
- Notebook: section A7

Back Exercise list

## A7 [Solved]: Conditional Moments

1. $\mathrm{d}(e^{\kappa t} r) = \kappa\theta e^{\kappa t}\,\mathrm{d}t + \sigma e^{\kappa t}\sqrt{r}\,\mathrm{d}W$; the stochastic integral has mean zero ($\mathbb{E}[r_s] < \infty$), so $\mathbb{E}[r_t] = \theta + (r_0 - \theta)e^{-\kappa t}$
2. $\mathrm{d}(r^2) = 2r\,\mathrm{d}r + \sigma^2 r\,\mathrm{d}t$, so $m_2' = (2\kappa\theta + \sigma^2)\mathbb{E}[r_t] - 2\kappa m_2$; multiply by $e^{2\kappa t}$ and integrate: $m_2(t) = r_0^2 e^{-2\kappa t} + (2\kappa\theta + \sigma^2)\int_0^t e^{-2\kappa(t-s)}\mathbb{E}[r_s]\,\mathrm{d}s$
   - Subtracting $\mathbb{E}[r_t]^2$: $\mathrm{Var} = r_0\frac{\sigma^2}{\kappa}(e^{-\kappa t} - e^{-2\kappa t}) + \frac{\theta\sigma^2}{2\kappa}(1 - e^{-\kappa t})^2$
   - $\mathbb{E}[r_1] = 0.055 + (0.03 - 0.055)e^{-0.11} = 3.26\%$; s.d. 0.94%

# A7 [Solved]: Transition Law and Feller




3. $c = 673$, $\nu = 7.72$, $\delta = 36.2$ (not the jump intensity); 90% prediction interval $[1.85\%, 4.92\%]$
4. $\nu = 2 \cdot (2\kappa\theta/\sigma^2)$, so $\nu \geq 2 \Leftrightarrow 2\kappa\theta \geq \sigma^2$; Feller ratio 3.86: zero is never reached

⊡ **Interpretation**: volatility $\sigma\sqrt{r}$ grows with the level, so the law is skewed to the right and the band widens upwards (right panel); negative rates are impossible, unlike Vasicek

MFM_ch11_seminar

# A8 [Proposed]: The VIX under Heston — Task

- **Questions**
  - which Heston parameters does a regression on the squared VIX recover?
  - under which measure?
- **Inputs**: instantaneous variance $v_t$ with $\mathrm{d}v = \kappa^Q(\theta^Q - v)\,\mathrm{d}t + \xi_v\sqrt{v}\,\mathrm{d}\bar{W}^{(2)}$ under $Q$ and $\kappa^P$, $\theta^P$ under $P$ (same $\xi_v$); $y_t = (\mathrm{VIX}_t/100)^2 = \frac{1}{\tau}\mathbb{E}^Q[\int_t^{t+\tau} v_s\,\mathrm{d}s \mid \mathcal{F}_t]$, $\tau = 30/365$; $\kappa^Q = 5$, $\theta^Q = 0.044$; a CIR fit to $y$ gives $\hat{\xi}_y = 0.56$
- **Tasks**
  1. Using A7 under $Q$, integrate $\mathbb{E}^Q[v_s \mid v_t]$ over $[t, t+\tau]$ and show $y_t = a + b\,v_t$; compute $b$ and $a$.
  2. From $\mathrm{d}y = b\,\mathrm{d}v$, write the diffusion of $y$ and the ratio $\xi_y/\xi_v$ as a function of $v$; evaluate the local correction $\hat{\xi}_y/b$ at $v = \theta^Q$.
  3. Derive the $P$-dynamics of $y$ and state which of $\kappa^P$, $\theta^P$, $\rho$ a historical regression on $y$ recovers, and what $\rho$ additionally needs.
  4. Compute $b$ at a 60-day horizon and compare it with its 30-day value.
- **Report**: $a$, $b$, the local correction, $b$ at 60 days, the chart of $\xi_y/\xi_v$
- **Interpretation**
  - is 0.56 too high or too low for $\xi_v$?
  - what is needed to estimate $\kappa^Q$?
- Model: A7; Notebook: section A8

Back Exercise list

# A9 [Proposed]: Rare, Large Jumps — Task

- **Question**: can the intensity of rare crashes be estimated from one index history?
- **Inputs**: the A3 model with $\sigma = 15\%$ but $\lambda = 0.5$ per year, $\mu_J = -10\%$, $\sigma_J = 5\%$ (A3: $\lambda = 5$, $\mu_J = -3\%$, $\sigma_J = 4\%$); $\Delta = 1/252$; $T = 36.7$ years; B5 MLE for the S&P 500: $\hat{\lambda} = 101.9$
- **Tasks**
  1. Fill the A3 table (diffusion and jump parts of $c_2$, $c_3$, $c_4$) for the new parameters and compute skewness and excess kurtosis.
  2. Compute the probability of no jump in 1 and in 10 years.
  3. With every crash observed, compute the SE of $\hat{\lambda}$ over 36.7 years and the years needed for a 20% relative SE
- **Report**: the moment table, the two probabilities, the SE and the years
- **Interpretation**
  - what does the B5 MLE capture instead?
  - is the problem only precision or also model uncertainty?
- Model: A3; Notebook: section A9

Back Exercise list

## Why Can Standard Errors Disagree?

- An estimate and its uncertainty answer different questions. A precise-looking coefficient may result from a standard-error formula whose assumptions the data violate.
- The Hessian uses the curvature of the fitted objective. A sandwich covariance also uses the variability of the scores, the contributions of observations to the derivative of that objective.
- HAC (Heteroskedasticity and Autocorrelation Consistent) standard errors allow both unequal variances and temporal dependence in the scores; the lag choice controls which autocovariances enter.
- In the next exercise, compare the standard errors before interpreting the Wald tests. A change in rejection can reflect the uncertainty estimate, even when the fitted coefficient stays the same.

# B2 Extended [Solved]: Exact CIR, Euler and CKLS — Task

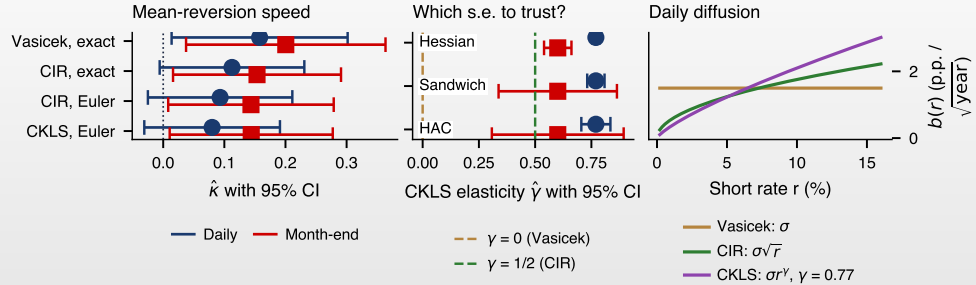
- **Questions**
  - which volatility specification do short-rate data prefer?
  - which test can be trusted?
- **Inputs**: DTB3, 4 January 1954 – 31 December 2007 (after 2008: 13 exact zeros, 8 negative quotes, outside the CIR and CKLS domains), daily ($n = 13\,487$, $\Delta = 1/252$) and month-end ($n = 647$, $\Delta = 1/12$); $\kappa, \theta, \sigma > 0$
- **Tasks**
  1. Maximise the exact CIR log-likelihood $\sum_t [\ln(2c) + \ln f_{\chi^2_{\nu}(\delta_t)}(2c\, r_{t+\Delta})]$ (the $2c$ is the Jacobian, A7) and the Euler Gaussian pseudo-likelihood; compare $\hat{\kappa}$, $\hat{\theta}$ with Hessian SE
  2. Fit CKLS (volatility $\sigma r^{\gamma}$, Extension Packs: the CIR Process and the Feller Condition) by the Euler quasi-likelihood; SE of $\hat{\gamma}$ from $H^{-1}$, the sandwich $H^{-1}\Omega H^{-1}$ with $\Omega$ the score outer product, and HAC $\Omega$ (Bartlett, $\lfloor 4(n/100)^{2/9} \rfloor$ lags) Chan, Karolyi, Longstaff & Sanders (1992); Newey & West (1987).
  3. Test $\gamma = 0$ and $\gamma = 1/2$ by the LR of restricted against unrestricted Euler objectives and by HAC Wald tests, at 5% (3.84), at both frequencies.
- **Report**: a table of $\hat{\kappa}$ (SE), $\hat{\theta}$; $\hat{\gamma}$ with three SE; the test statistics; the chart
- **Interpretation**: why do the LR and the Hessian SE over-reject?
- Notebook: section B2 Extended

# B2 Extended [Solved]: Estimates

| | Daily | | Month-end | |
|---|---|---|---|---|
| Model | $\hat{\kappa}$ (SE) | $\hat{\theta}$ (%) | $\hat{\kappa}$ (SE) | $\hat{\theta}$ (%) |
| Vasicek, exact | 0.158 (0.073) | 5.39 | 0.201 (0.083) | 5.36 |
| CIR, exact | 0.113 (0.060) | 5.48 | 0.153 (0.070) | 5.43 |
| CIR, Euler | 0.093 (0.060) | 5.55 | 0.144 (0.069) | 5.45 |
| CKLS, Euler | 0.080 (0.057) | 5.84 | 0.144 (0.068) | 5.46 |

1. Exact and Euler CIR differ by less than one SE of $\hat{\kappa}$ at both frequencies: discretisation error is small next to long-span uncertainty
2. All $\hat{\kappa}$ lie between 0.08 and 0.20, with SE 0.06–0.08: no likelihood pins down the speed (left panel of the chart, next slide)

# B2 Extended [Solved]: Tests on $\gamma$




1. $\hat{\gamma} = 0.77$ daily (SE: Hessian 0.008, sandwich 0.020, HAC 0.033, 11 lags); 0.60 monthly (0.031, 0.134, 0.149, 6 lags)
2. $\gamma = 0$: LR 7 926.6 and 285.6, HAC Wald 545.1 and 16.1: Vasicek rejected at both frequencies
3. $\gamma = 1/2$: daily LR 1 139.34 and HAC Wald 67.01 reject (22.13 with 55 lags); monthly LR 10.07 rejects, HAC Wald 0.45 does not

⊡ **Interpretation**: volatility rises with the level; omitted clustering makes the scores heteroskedastic and autocorrelated, so trust the HAC Wald; Chan, Karolyi, Longstaff & Sanders (1992) found $\hat{\gamma} \approx 1.5$ on 1964–1989

MFM_ch11_seminar

# B6 [Proposed]: Mean Reversion of the VIX — Task

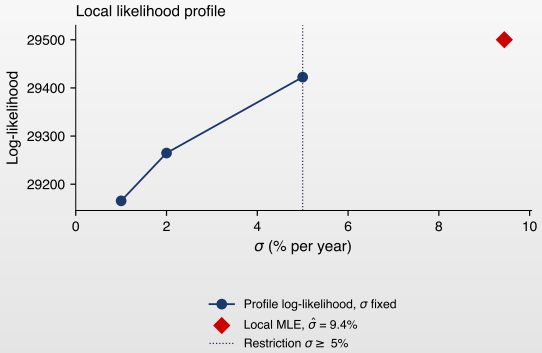
- **Questions**
  - how fast do volatility spikes fade?
  - is one OU factor enough?
- **Data**: `VIX.INDX`, close, 1990–2026, weekdays; $x_t = \ln \mathrm{VIX}_t$, $\Delta t = 1/252$, $n = 9\,283$ transitions
- **Tasks**
  1. Fit the OU process by the exact AR(1) (B2, step 1) and give $\hat{\kappa}$, its SE and the stationary median level $e^{\hat{\theta}}$.
  2. Convert $\hat{\kappa}$ into a half-life in trading days, $252 \ln 2/\hat{\kappa}$, with the interval $[252 \ln 2/\kappa_{\mathrm{upper}}, 252 \ln 2/\kappa_{\mathrm{lower}}]$ (both ends positive).
  3. Estimate the bias with 300 simulated OU paths, and compare the sample ACF with $\hat{b}^k$ up to lag 120 on a chart.
- **Report**: $\hat{\kappa}$, $e^{\hat{\theta}}$, half-life with interval, bias, the ACF chart
- **Interpretation**
  - at which horizons does one OU factor fail?
  - what extension would you try?
- Model: B2; Notebook: section B6

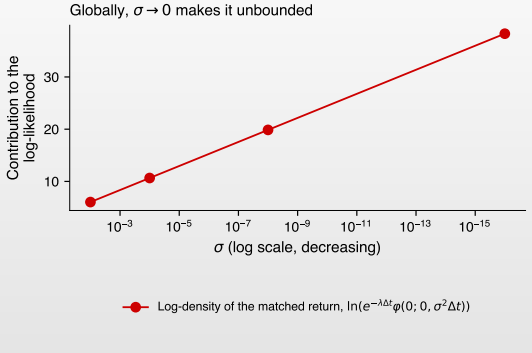
Back Exercise list

# B5 Extended [Solved]: Is the Merton Likelihood Bounded? — Task

- **Question**: what does the reported Merton MLE mean if the likelihood has no maximum?
- **Inputs**: the S&P 500 returns and the B5 fit ($\hat{\sigma} = 9.4\%$); the fitting helper of B5 with $\sigma$ fixed Honoré (1998)
- **Tasks**
  1. Compute the profile log-likelihood with $\sigma$ fixed at 5%, 2% and 1%, refitting the other four parameters.
  2. Set $m\Delta t$ equal to one observed return and let $\sigma \to 0$: compute that return's term $\ln(e^{-\lambda\Delta t}\varphi(0; 0, \sigma^{\mathbf{2}}\Delta t))$ at $\sigma = 10^{-\mathbf{2}}, 10^{-\mathbf{4}}, 10^{-\mathbf{8}}, 10^{-\mathbf{16}}$.
- **Report**: the profile table and the singular sequence, with the chart
- **Interpretation**
  - is the likelihood bounded?
  - why can numerical profiles alone not settle it?
- Notebook: section B5 Extended

# B5 Extended [Solved]: Local Profile and Global Singularity (1/2)




MFM_ch11_seminar

# B5 Extended [Solved]: Local Profile and Global Singularity (2/2)

| $\sigma$ (fixed) | log-likelihood | $\hat{\lambda}$ per year | $\hat{\sigma}_J$ (%) |
|---|---|---|---|
| 9.4% (MLE) | 29 500 | 101.9 | 1.44 |
| 5% | 29 422 | 302 | 0.93 |
| 2% | 29 265 | 555 | 0.70 |
| 1% | 29 165 | 710 | 0.62 |

1. Locally the profile falls as $\sigma$ shrinks while $\hat{\lambda}$ grows (left panel); the MLE satisfies $\sigma \geq 5\%$
2. Globally the matched return's term is 6.0, 10.7, 19.9, 38.3 at $\sigma = 10^{-2}, 10^{-4}, 10^{-8}, 10^{-16}$, while the other returns keep finite densities from the jump components: $\sup = +\infty$ (right panel)

⊡ **Interpretation**: the argument, not the numbers, proves unboundedness; the MLE is an interior local maximum, so the LR test of B5 is defined on $\sigma \geq 5\%$ for data and bootstrap alike

MFM_ch11_seminar

# B7 [Proposed]: Jumps in Bitcoin — Task

- **Question**: how often does Bitcoin jump, by a fitted mixture and by a detection test?
- **Data**: `BTC-USD.CC`, close, 18 September 2014 – 18 September 2026, 7 days a week, $\Delta t = 1/365$; Lee–Mykland: $L_t = r_t/\hat{\sigma}_t$, $\hat{\sigma}_t^2 = \frac{1}{K-2}\sum_{j=t-K+2}^{t-1} |r_j|\,|r_{j-1}|$, $K = 16$; statistic $(|L_t| - C_n)/S_n$ against $-\ln(-\ln(0.99))$ (formulas for $C_n$, $S_n$: Chapter 11, Further Topics) Lee & Mykland (2008)
- **Tasks**
  1. Fit Merton as in B5 and report all five parameters with SE, the implied excess kurtosis and the jump share of variance.
  2. Apply the test: the first $K - 1$ days have no local scale and are not tested, $n$ = number of tested days; count the flagged days per year and their mean absolute size, and plot returns and statistic.
  3. Calibrate the threshold by simulation: the 99% quantile of the sample maximum of the statistic in 2,000 i.i.d. Normal samples of length $n$, same $K$; count the days above it.
- **Report**: the Merton table and the Lee–Mykland counts side by side, with the chart
- **Interpretation**
  - Compare $\hat{\lambda}$ with the detected-jump count. Do they measure the same object?
  - Under a change to $Q$, jump intensity and the jump-size law may also change. Why is neither estimate directly the option-pricing intensity?
- Model: B5; Notebook: section B7

Back Exercise list

# B8 [Proposed]: Heston Parameters from the VIX — Task

- **Questions**
  - what do S&P 500 and VIX data say about $\kappa$, $\theta$, $\xi$ and $\rho$?
  - are these parameters stable?
- **Data**: S&P 500 (as in B1) and `VIX.INDX` joined on common days, 1990–2026, then returns; $y_t = (\mathrm{VIX}_t/100)^2$ in place of $v_t$, $\Delta t = 1/252$ Heston (1993)
- **Tasks**
  1. Regress $(y_{t+\Delta} - y_t)/\sqrt{y_t}$ on $\Delta t/\sqrt{y_t}$ and $-\Delta t\sqrt{y_t}$ without an intercept: $\kappa = \beta_2$, $\theta = \beta_1/\beta_2$, $\xi = s_e/\sqrt{\Delta t}$; give OLS and HAC SE and a 95% CI for $\kappa$.
  2. Estimate $\rho$ as the correlation of $r_{t+1}/\sqrt{y_t}$ with the regression residual over the same interval, with a Fisher-$z$ CI (assumes i.i.d. bivariate Normal shocks) and a moving-block bootstrap CI (63-day blocks of the price–VIX rows, all parameters re-estimated), and the Feller ratio $2\kappa\theta/\xi^2$.
  3. Repeat on 1990–2007 and 2008–2026.
- **Report**: a table with the full sample and the two halves
- **Interpretation**: Compare $\sqrt{\theta}$ from the VIX with realised volatility. What could explain a gap?
- Model: B2, B5; Notebook: section B8

Back Exercise list

# B8 Extended [Proposed]: The Affine Map and Realised Variance — Task

- **Question**: does $\rho$ depend on which variance proxy we use?
- **Inputs**: as in B8, plus the SPY 5-minute realised variance (RV) of Chapter 9 (`ch9_rv_spy.csv`) (total, overnight included, in squared percent): annualise by $\mathrm{RV} \times 10^{-4}/\Delta t$; join S&P 500, VIX and RV prices on common days first (October 2020 – September 2026), then returns
- **Tasks**
  1. On the common days, estimate $\rho$ twice (VIX and RV in place of $v_t$), with the same return timing and Fisher-$z$ 95% CIs; draw both scatter plots.
- **Report**: the two correlations with CIs, the chart
- **Interpretation**
  - Compare $\rho$ from RV and VIX. What could explain the difference?
  - which $\rho$ would you use in a Heston model?
- Model: A8, B8; Notebook: section B8 Extended

## Local Estimates and Bandwidth

- Silverman's bandwidth rule uses $s_r$, the sample standard deviation of the rate levels, in $h = 1.06 s_r n^{-1/5}$. The bandwidth controls how many nearby rate levels contribute appreciably.
- For drift, $y_t = \Delta r_t / \Delta$ and $\hat{m} = \hat{a}(x)$. For variance, $y_t = (\Delta r_t)^2 / \Delta$ and $\hat{m} = \hat{b}^2(x)$. Both are local weighted averages at level $x$.
- In B9 compare the estimates for the requested bandwidths. Distinguish changes in the estimated curve from changes in its uncertainty.

# B9 [Proposed]: Nonparametric Diffusion of the T-Bill Rate — Task

- **Question**: which parametric diffusion, if any, matches the data at every rate level?
- **Data**: DTB3 daily, 1954–2007, $\Delta = 1/252$; the fits of B2 Extended Bandi & Phillips (2003); grid of 60 levels between the 2% and 98% quantiles; Gaussian kernel, $h = 1.06\, s_r\, n^{-1/5}$; weights $w_t(x) = K((r_t - x)/h) / \sum_s K((r_s - x)/h)$
- **Tasks**
  1. Compute $\hat{a}(x) = \sum_t w_t\, \Delta r_t / \Delta$ and $\hat{b^2}(x) = \sum_t w_t\, (\Delta r_t)^2 / \Delta$ with SE $\sqrt{\sum_t w_t^2 (y_t - \hat{m})^2}$; bands for $b$ are the square roots of the $b^2$ bands, truncated at zero.
  2. Overlay the Vasicek, CIR and CKLS diffusions and compute the share of levels inside the pointwise band.
  3. Repeat with $h/2$ and $2h$; report the log-slope of $\hat{b}(r)$ and the share of levels where the drift band contains zero.
- **Report**: the chart and a table by bandwidth
- **Interpretation**
  - can the drift be distinguished from zero?
  - what do the bands ignore?
- Model: B2 Extended; Notebook: section B9

Back Exercise list

# C2 [Proposed]: Which Continuous-Time Model Fits BET and Bitcoin? — Task

- **Question**: which stylised facts does each fitted model reproduce in markets without an implied-volatility index?
- **Data**: BET and Bitcoin daily log returns (Data Used); variance proxy $RV_t = \frac{1}{21\Delta t}\sum_{j=0}^{20} r_{t-j}^2$, an overlapping window: for i.i.d. returns its lag-$k$ autocorrelation is $(21-k)/21$, so persistence is partly built in (control: simulated ACF 0.952 at lag 1)
- **Tasks**
  1. Start with one market: fit GBM (B1), Merton (B5) and Heston by the B8 regression with $RV_t$ in place of $v_t$.
  2. Simulate 100 samples of the data length per model (Heston: Euler with full truncation, daily step, $\theta$ set to the realised variance, $v_0 = \theta$, 500 burn-in steps, fixed seed) and compute excess kurtosis, the Hill statistic $\left[\frac{1}{k}\sum_{i\le k} \ln(L_{(i)}/L_{(k+1)})\right]^{-1}$ of the losses $L = -r$ ($k = 5\%$ of returns), skewness and $\sum_{k=1}^{20}$ ACF of $|r_t|$.
  3. Report $q$ = share of simulated statistics at or below the data value (e.g. $q = 0.97$: data above 97 of 100 samples), a descriptive check since parameters are not refitted.
- **Deliverables**: medians with 90% ranges, the $q$ table, the charts; optional: Bates (Heston plus jumps) as a separate advanced project
- **Assessment criteria**: reproducible fits and simulations (fixed seed); a correct comparison of the data with the simulated distributions; the limits of each conclusion
- **Interpretation**
  - Which features does no simulation reproduce?
  - Is Heston's near-zero skewness a property of the model or of the discretisation?
  - Is Bitcoin a distinct asset class Pele et al. (2023)?
- Model: B1, B5, B8; Notebook: section C2

Back Exercise list

## How Do We Evaluate Rough-Volatility Models?

- Nonlinear Least Squares (NLLS) fits a nonlinear curve by minimising squared residuals. In the variogram exercise it estimates the roughness parameter directly rather than using a log-linear approximation.
- Rough Fractional Stochastic Volatility (RFSV) is a stochastic-volatility model with log-volatility paths that are locally less regular than Brownian motion. It is a forecast competitor, not a conclusion from one estimated slope.
- Mean Squared Error (MSE) averages squared forecast errors. Quasi-likelihood (QLIKE) is a loss for variance forecasts that treats under- and over-prediction differently. Compare models on the same observations and horizons.
- A Model Confidence Set (MCS) retains models whose predictive performance cannot be statistically distinguished under the procedure. It gives a set, not a ranking or proof that every retained model is correct.

# C3 [Proposed]: Is SPY Volatility Rough and Persistent? Stage 1 — Task

- **Question**: is daily SPY log-volatility rough ($\alpha < 0$) and consistent with long memory ($\beta < 1$)?
- **Data**: SPY 5-minute bars, complete 09:30–16:00 New York sessions (78 bars); first return of a session log(close/open) of its first bar, then close to close, no overnight return; the factor $M/(M-1)$ in $\mathrm{BV}_t$ shifts log-volatility by a constant, which cancels in its ACF and variogram; lecture case study Bennedsen, Lunde & Pakkanen (2022)
- **Tasks**
  1. Compute daily bipower variation $\mathrm{BV}_t = \frac{\pi}{2}\frac{M}{M-1}\sum_{j=2}^{M}|r_j||r_{j-1}|$, $M = 78$ Barndorff-Nielsen & Shephard (2004), and the daily log-volatility $x_t = \log\hat{\sigma}_t = \frac{1}{2}\log \mathrm{BV}_t$, with $\hat{\sigma}_t = \sqrt{\mathrm{BV}_t}$ on day $t$; plot the series.
  2. Compute the variogram $\hat{\gamma}_2(k) = \frac{1}{n-k}\sum_t (x_{t+k} - x_t)^2$, which behaves as $c\,k^{2\alpha+1}$ for small $k$; estimate $\alpha$ by OLS of $\log\hat{\gamma}_2(k)$ on $\log k$, $k = 1, \ldots, 6$, $\hat{\alpha} = (\text{slope} - 1)/2$, and by NLLS $\hat{\gamma}_2(k) = b_0 + b_1 k^{2\alpha+1}$ averaged over $m = 10, \ldots, 20$ lags.
  3. Fit the Cauchy ACF $\rho(h) = (1 + h^{2\alpha+1})^{-\beta/(2\alpha+1)}$ and an exponential (OU) ACF to the first $\lceil n^{1/3} \rceil$ lags; plot both against the sample ACF to lag 100.
- **Deliverables**: $\hat{\alpha}_{\mathrm{OLS}}$, $\hat{\alpha}^{*}_{\mathrm{NLLS}}$, $\hat{\beta}$, the OU half-life, three charts
- **Assessment criteria**: a reproducible $\mathrm{BV}_t$ series and variogram; robustness to the lags; an interpretation of $\hat{\alpha}$ and $\hat{\beta}$ that states its limits
- **Interpretation**
  - is "consistent with long memory" the same as "has long memory"?
  - How robust are the estimates to the lags used?
- Model: B2, B4; Notebook: section C3

Back Exercise list

## C3 [Proposed]: Stages 2 and 3 (Advanced, Optional)

- **Stage 2, intraday**: $\Delta = 30$ and 65 minutes ($M = 6$ and 13 returns per bin), $\hat{\sigma}_t^2 = \mathrm{BV}_t/\Delta$ divided by a seasonal factor
  - Seasonal factor: flexible Fourier form for each bin's volatility Andersen & Bollerslev (1997), scaled so that its square averages 1 over the session, estimated from past data only; plot the intraday profile before and after
  - Lags $h$ and $\lambda$ are in steps of $\Delta$: compare with Table 3, Panel A of the paper only in the same unit; bars instead of every trade mean no pre-averaging and more noise
- **Stage 3, forecasting**: at $\Delta = 65$ minutes, forecast the variance summed over $h = 1, 5, 20$ steps with log-HAR (heterogeneous autoregressive model for log volatility) Corsi (2009), RFSV Gatheral, Jaisson & Rosenbaum (2018) and the Cauchy model, rolling window of 200 observations, only data up to each origin
  - Losses MSE and QLIKE, 90% MCS Hansen, Lunde & Nason (2011); plot realised against forecast variance and cumulative QLIKE differences; the MCS does not rank the models it contains
- Assess robustness and reproducibility, not whether the paper's ranges are matched

# Acronyms Used in This Material

- **ACF**: AutoCorrelation Function
- **AI**: Artificial Intelligence
- **AR**: AutoRegressive (model)
- **BET**: Bucharest Exchange Trading (index)
- **CI**: Confidence Interval
- **CIR**: Cox–Ingersoll–Ross (model)
- **CKLS**: Chan–Karolyi–Longstaff–Sanders (model)
- **DTB3**: 3-Month Treasury Bill: secondary market rate (FRED series code)
- **EM**: Euler–Maruyama (discretisation scheme)
- **FRED**: Federal Reserve Economic Data
- **GARCH**: Generalised AutoRegressive Conditional Heteroskedasticity
- **GBM**: Geometric Brownian Motion
- **HAC**: Heteroskedasticity and Autocorrelation Consistent (standard errors)
- **JB**: Jarque–Bera (test)
- **LR**: Likelihood Ratio (test)
- **MCS**: Model Confidence Set
- **MLE**: Maximum Likelihood Estimation
- **MSE**: Mean Squared Error
- **NLLS**: Non-Linear Least Squares
- **OLS**: Ordinary Least Squares
- **OU**: Ornstein–Uhlenbeck (process)
- **PDE**: Partial Differential Equation
- **QLIKE**: Quasi-LIKElihood (loss function)
- **QQ**: Quantile–Quantile (plot)
- **RFSV**: Rough Fractional Stochastic Volatility (model)
- **RV**: Realised Variance
- **SDE**: Stochastic Differential Equation
- **SE**: Standard Error
- **US**: United States
- **VIX**: Cboe Volatility Index

# References (1/4)


- Andersen, T.G., Bollerslev, T. (1997). Intraday periodicity and volatility persistence in financial markets. *Journal of Empirical Finance*, 4(2–3), 115–158.
- Bandi, F.M., Phillips, P.C.B. (2003). Fully nonparametric estimation of scalar diffusion models. *Econometrica*, 71(1), 241–283.
- Barndorff-Nielsen, O.E., Shephard, N. (2004). Power and bipower variation with stochastic volatility and jumps. *Journal of Financial Econometrics*, 2(1), 1–37.
- Bennedsen, M., Lunde, A., Pakkanen, M.S. (2022). Decoupling the short- and long-term behavior of stochastic volatility. *Journal of Financial Econometrics*, 20(5), 961–1006.
- Bollerslev, T., Zhou, H. (2002). Estimating stochastic volatility diffusion using conditional moments of integrated volatility. *Journal of Econometrics*, 109(1), 33–65.
- Borak, S., Härdle, W.K., López-Cabrera, B. (2013). Statistics of Financial Markets: Exercises and Solutions, 2nd ed. Springer.
- Chan, K.C., Karolyi, G.A., Longstaff, F.A., Sanders, A.B. (1992). An empirical comparison of alternative models of the short-term interest rate. *Journal of Finance*, 47(3), 1209–1227.
- Cont, R. (2001). Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance*, 1(2), 223–236.
- Corsi, F. (2009). A simple approximate long-memory model of realized volatility. *Journal of Financial Econometrics*, 7(2), 174–196.

# References (2/4)


- Cox, J.C., Ingersoll, J.E., Ross, S.A. (1985). A theory of the term structure of interest rates. *Econometrica*, 53(2), 385–407.
- Franke, J., Härdle, W.K., Hafner, C.M. (2019). Statistics of Financial Markets: An Introduction, 5th ed. Springer.
- Gatheral, J., Jaisson, T., Rosenbaum, M. (2018). Volatility is rough. *Quantitative Finance*, 18(6), 933–949.
- Girsanov, I.V. (1960). On transforming a certain class of stochastic processes by absolutely continuous substitution of measures. *Theory of Probability and Its Applications*, 5(3), 285–301.
- Hansen, P.R., Lunde, A., Nason, J.M. (2011). The model confidence set. *Econometrica*, 79(2), 453–497.
- Heston, S.L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies*, 6(2), 327–343.
- Higham, D.J. (2001). An algorithmic introduction to numerical simulation of stochastic differential equations. *SIAM Review*, 43(3), 525–546.
- Honoré, P. (1998). Pitfalls in estimating jump-diffusion models. Working paper, SSRN.
- Itô, K. (1944). Stochastic integral. *Proceedings of the Imperial Academy*, 20(8), 519–524.

# References (3/4)


- Jarque, C.M., Bera, A.K. (1980). Efficient tests for normality, homoscedasticity and serial independence of regression residuals. *Economics Letters*, 6(3), 255–259.
- Kac, M. (1949). On distributions of certain Wiener functionals. *Transactions of the American Mathematical Society*, 65(1), 1–13.
- Künsch, H.R. (1989). The jackknife and the bootstrap for general stationary observations. *Annals of Statistics*, 17(3), 1217–1241.
- Lee, S.S., Mykland, P.A. (2008). Jumps in financial markets: a new nonparametric test and jump dynamics. *Review of Financial Studies*, 21(6), 2535–2563.
- Ljung, G.M., Box, G.E.P. (1978). On a measure of lack of fit in time series models. *Biometrika*, 65(2), 297–303.
- Merton, R.C. (1976). Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics*, 3(1–2), 125–144.
- Newey, W.K., West, K.D. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 55(3), 703–708.
- 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.
- Samuelson, P.A. (1965). Rational theory of warrant pricing. *Industrial Management Review*, 6(2), 13–39; reprinted in *Henry P. McKean Jr. Selecta* (2015), Birkhäuser.
- Tang, C.Y., Chen, S.X. (2009). Parameter estimation and bias correction for diffusion processes. *Journal of Econometrics*, 149(1), 65–81.

# References (4/4)


- Uhlenbeck, G.E., Ornstein, L.S. (1930). On the theory of the Brownian motion. *Physical Review*, 36(5), 823–841.
- Vasicek, O. (1977). An equilibrium characterization of the term structure. *Journal of Financial Economics*, 5(2), 177–188.
- Wilks, S.S. (1938). The large-sample distribution of the likelihood ratio for testing composite hypotheses. *Annals of Mathematical Statistics*, 9(1), 60–62.