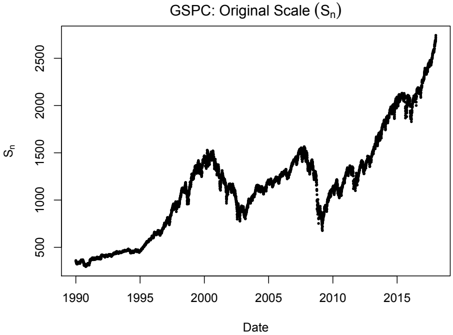
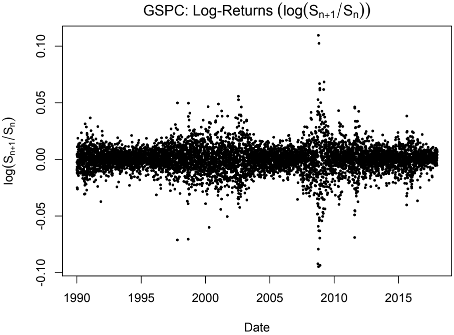
Stochastic Differential Equations

version: 2020-01-09 · 07:55:34

Motivation



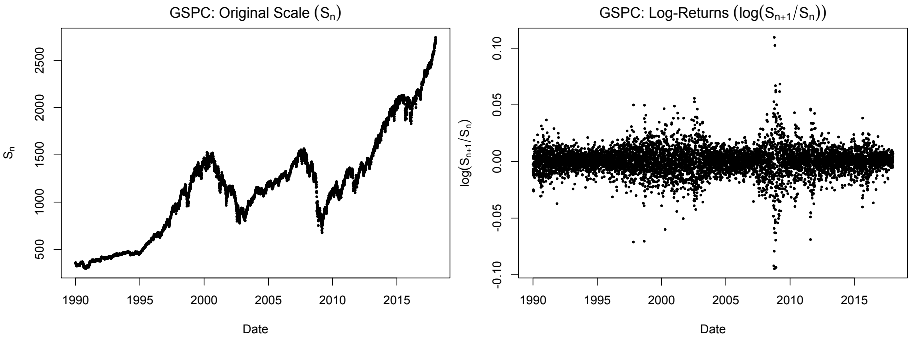


- **Data:** Daily GSPC (S&P500) closing prices between 1990-2018:

$$\boldsymbol{S} = (S_1, \ldots, S_N), \qquad N = 7060.$$

- **Goal:** Fit a model and use it to make forecasts

Motivation

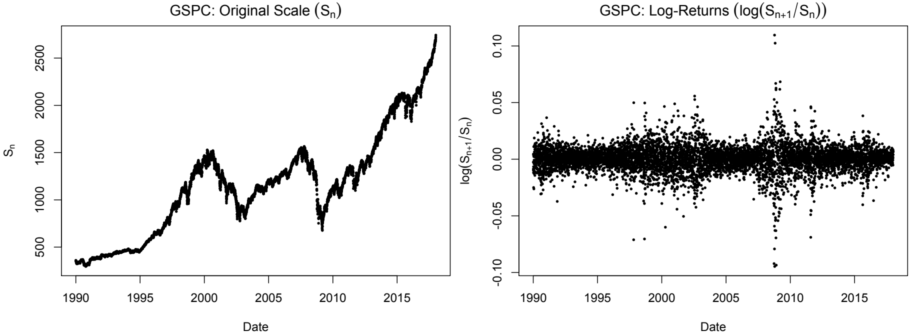


- **Data:** Daily GSPC closing prices $\boldsymbol{S} = (S_1, \ldots, S_N)$
- **Model:** From 1900 till about 1989, this was a stochastic differential equation (SDE) called Geometric Brownian motion:

$$dS_t = \alpha S_t \, dt + \sigma S_t \, dB_t,$$

where S_t is the price at time $t \in \mathbb{R}$ and B_t is Brownian motion.

Motivation



- **Data:** Daily GSPC closing prices $\boldsymbol{S} = (S_1, \dots, S_N)$
- **Model:** From 1900 till about 1989, this was a stochastic differential equation (SDE) called Geometric Brownian motion:

$$\mathrm{d}S_t = \alpha S_t \, \mathrm{d}t + \sigma S_t \, \mathrm{d}B_t, \qquad \text{What do these symbols mean???}$$

where S_t is the price at time $t \in \mathbb{R}$ and B_t is Brownian motion.

Stochastic Differential Equations

- **Definition:** Let X_t be a continuous-time stochastic process for which the following hold:

  1. Almost every sample path is continuous.

  2. X_t has the homogeneous Markov property:

$$
\begin{aligned}
p(X_{t+s} \mid \{X_u = x_u : u \le t\}) &= p(X_{t+s} \mid X_t = x_t) \\
&= p(X_s \mid X_0 = x_t).
\end{aligned}
$$

  Then (up to mathematical singularity) X_t is a **diffusion process** which can be described with a **stochastic differential equation**

$$
\mathrm{d}X_t = \mu(X_t)\,\mathrm{d}t + \sigma(X_t)\,\mathrm{d}B_t,
$$

  where:

  - $\mu(x)$ and $\sigma(x)$ (or $\sigma^2(x)$) are the drift and diffusion functions
  - B_t is **Brownian motion**.

Brownian Motion

- **Diffusion Process:** X_t continuous & Markov
- **SDE:** $\mathrm{d}X_t = \mu(X_t)\,\mathrm{d}t + \sigma(X_t)\,\mathrm{d}B_t$
- **Brownian Motion:** B_t is the *only* process with the following properties:

  1. Almost every sample path is *continuous*.

  2. *Independent increments:* If $t_0 < t_1 < \cdots < t_N$, then

$$B_{t_1} - B_{t_0} \quad \amalg \quad B_{t_2} - B_{t_1} \quad \amalg \quad \cdots \quad \amalg \quad B_{t_N} - B_{t_{N-1}}$$

  3. B_t is a *Gaussian process* with $B_{s+t} - B_s \sim \mathcal{N}(0, t)$. (Typically assumed that $B_0 = 0$)

- **Mean and Variance:** Turns out that properties 1-3 completely determine the distribution of B_t, such that

$$E[B_t] = 0, \qquad \mathrm{cov}(B_t, B_s) = \min(t, s).$$

Brownian Motion

- **Definition:** B_t is a continuous Gaussian process with $E[B_t] = 0$ and $\text{cov}(B_t, B_s) = \min(t, s)$
- **Simulation:** To simulate B_t at times $0 < t_1 < \cdots < t_N$, i.e., to generate $\boldsymbol{B} = (B_{t_1}, \ldots, B_{t_N})$
  - *Method 1:* Let $\boldsymbol{\Sigma} = \text{var}(\boldsymbol{B}) \iff \Sigma_{ij} = \min(t_i, t_j)$, then set $\boldsymbol{B} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma})$.

    However, this requires $\boldsymbol{\Sigma}^{1/2}$, which is $\mathcal{O}(N^3)$.

Brownian Motion

- **Definition:** B_t is a continuous Gaussian process with $E[B_t] = 0$ and $\text{cov}(B_t, B_s) = \min(t, s)$
- **Simulation:** To simulate B_t at times $0 < t_1 < \cdots < t_N$, i.e., to generate $\boldsymbol{B} = (B_{t_1}, \ldots, B_{t_N})$
  - *Method 1:* Let $\boldsymbol{\Sigma} = \text{var}(\boldsymbol{B}) \iff \Sigma_{ij} = \min(t_i, t_j)$, then set $\boldsymbol{B} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma})$.

    However, this requires $\boldsymbol{\Sigma}^{1/2}$, which is $\mathcal{O}(N^3)$.
  - *Method 2:* Use independent increment property

$$B_{t_{i+1}} - B_{t_i} \quad \amalg \quad B_{t_{j+1}} - B_{t_j}, \qquad i \neq j.$$

Therefore, letting $t_0 = 0$, $\Delta B_i = B_{t_{i+1}} - B_{t_i}$, and $\Delta t_i = t_{i+1} - t_i$:

1. Draw $\Delta B_i \overset{\text{ind}}{\sim} \mathcal{N}(0, \Delta t_i)$ for $i = 0, \ldots, N-1$
2. Let $B_{t_n} = \sum_{i=0}^{n-1} \Delta B_i$.

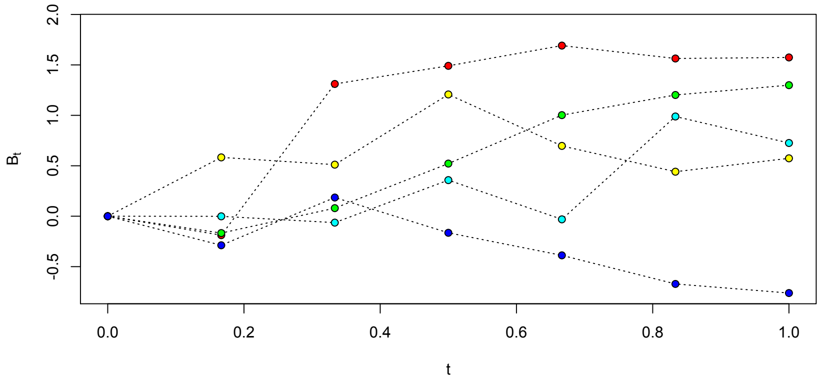
This method of simulation is only $\mathcal{O}(N)$.

Brownian Motion

- **Simulation:** To simulate B_t on $[0, T]$ at intervals of $\Delta t = T/N$:

  1. Draw $\Delta B_i \overset{\text{iid}}{\sim} \mathcal{N}(0, \Delta t)$
  2. Let $B_{n\Delta t} = \sum_{i=0}^{n-1} \Delta B_i$

- $T = 1$, $N = 6$:



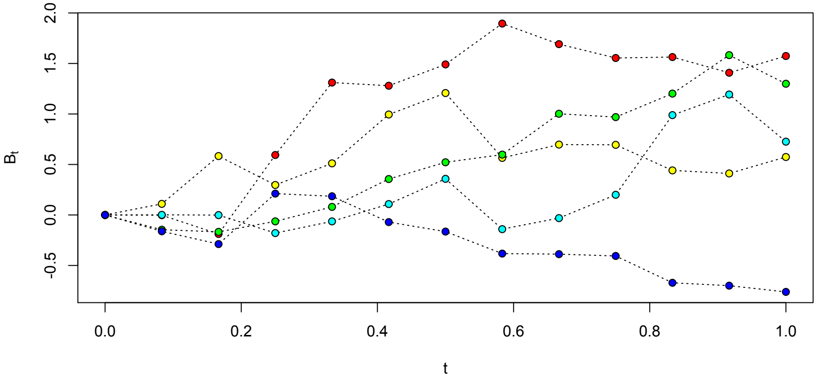
Brownian Motion

- **Simulation:** To simulate B_t on $[0, T]$ at intervals of $\Delta t = T/N$:

  1. Draw $\Delta B_i \overset{\text{iid}}{\sim} \mathcal{N}(0, \Delta t)$ **2.** Let $B_{n\Delta t} = \sum_{i=0}^{n-1} \Delta B_i$

- $T = 1$, $N = 12$: Infill using Markov property with

$$B_t \mid B_{t-s}, B_{t+u} \sim \mathcal{N}(\sigma^2\{B_{t-s}/s + B_{t+u}/u\}, \sigma^2), \qquad \sigma^2 = 1/(1/s + 1/u)$$



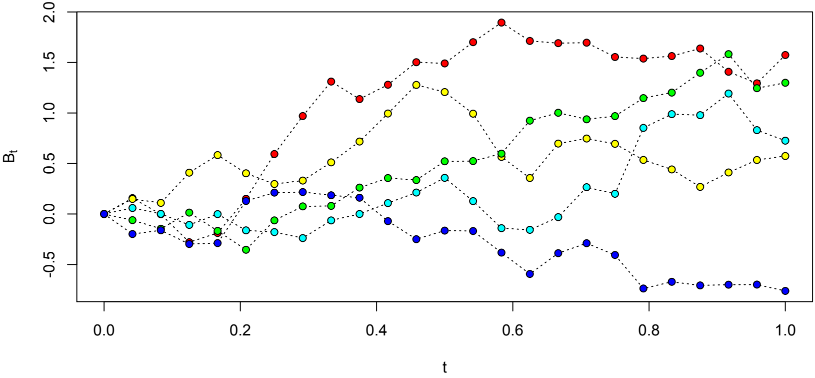
Brownian Motion

- **Simulation:** To simulate B_t on $[0, T]$ at intervals of $\Delta t = T/N$:

1. Draw $\Delta B_i \overset{\text{iid}}{\sim} \mathcal{N}(0, \Delta t)$

2. Let $B_{n\Delta t} = \sum_{i=0}^{n-1} \Delta B_i$

- $T = 1,\ N = 24$:



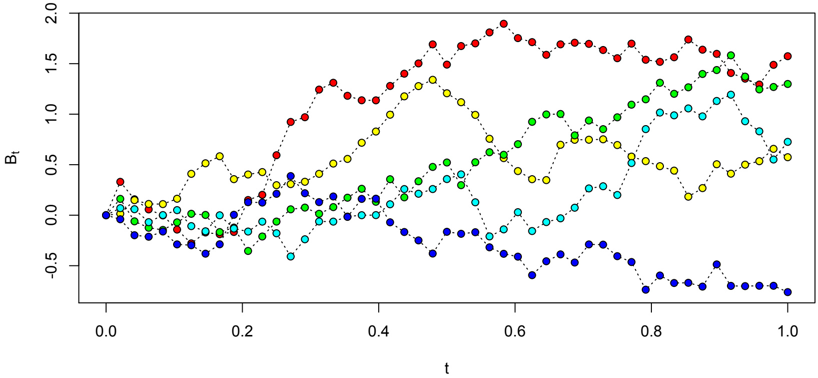
Brownian Motion

- **Simulation:** To simulate B_t on $[0, T]$ at intervals of $\Delta t = T/N$:

1. Draw $\Delta B_i \overset{\text{iid}}{\sim} \mathcal{N}(0, \Delta t)$

2. Let $B_{n\Delta t} = \sum_{i=0}^{n-1} \Delta B_i$

- $T = 1,\ N = 48$: (As $N \to \infty$ get a continuous process but nowhere differentiable)



Stochastic Differential Equations (Continued)

- **Diffusion Process:** X_t continuous & Markov
- **SDE:** $\mathrm{d}X_t = \mu(X_t)\,\mathrm{d}t + \sigma(X_t)\,\mathrm{d}B_t$
- **Example 1:** $\mathrm{d}X_t = \mathrm{d}B_t$. (so $\mu(x) = 0$ and $\sigma(x) = 1$.) Solution is

$$X_t = X_0 + \int_0^t \mathrm{d}X_s = X_0 + \int_0^t \mathrm{d}B_s = X_0 + B_t,$$

  i.e., Brownian motion.

Stochastic Differential Equations (Continued)

- **Diffusion Process:** X_t continuous & Markov
- **SDE:** $\mathrm{d}X_t = \mu(X_t)\,\mathrm{d}t + \sigma(X_t)\,\mathrm{d}B_t$
- **Example 1:** $\mathrm{d}X_t = \mathrm{d}B_t$. (so $\mu(x) = 0$ and $\sigma(x) = 1$.) Solution is

$$X_t = X_0 + \int_0^t \mathrm{d}X_s = X_0 + \int_0^t \mathrm{d}B_s = X_0 + B_t,$$

  i.e., Brownian motion.
- **Example 2:** $\mathrm{d}X_t = \mu\,\mathrm{d}t + \sigma\,\mathrm{d}B_t$. Solution is

$$X_t = X_0 + \int_0^t \mu\,\mathrm{d}s + \int_0^t \sigma\,\mathrm{d}B_s = X_0 + \mu t + \sigma B_t,$$

  i.e., Brownian motion with drift.

Stochastic Differential Equations

- **Diffusion Process:** X_t continuous & Markov
- **SDE:** $\mathrm{d}X_t = \mu(X_t)\,\mathrm{d}t + \sigma(X_t)\,\mathrm{d}B_t$
- **Example 3:** General case.
  - Unfortunately, $p(X_t \mid X_0)$ is typically intractable (solution to a PDE).
  - *Small-Time Approximation:* use the Euler Discretization.

Euler Discretization

- **Ordinary Differential Equation:**

$$\frac{\mathrm{d}X_t}{\mathrm{d}t} = G(X_t), \qquad X_0 = x_0.$$

Typically these are intractable.

- **Euler Discretization:** Replace infinitesimal $\mathrm{d}t$ by time interval Δt:

$$\frac{\mathrm{d}X_t}{\mathrm{d}t} = G(X_t) \quad \Longrightarrow \quad \frac{\Delta X_t}{\Delta t} \approx G(X_t) \quad \Longrightarrow \quad \Delta X_t \approx G(X_t)\Delta t,$$

where $\Delta X_t = X_{t+\Delta t} - X_t$.

- **Approximate Solution:** To approximate ODE on $[0, T]$ at intervals of $\Delta t = T/N$:

  1. Let $\tilde{X}_0 = x_0$.
  2. Let $\tilde{X}_{n+1} = \tilde{X}_n + G(\tilde{X}_n)\Delta t$.

Euler Discretization

- **Ordinary Differential Equation:** $\frac{\mathrm{d}X_t}{\mathrm{d}t} = G(X_t)$.
- **Euler Discretization:** $\frac{\mathrm{d}X_t}{\mathrm{d}t} = G(X_t) \implies \Delta X_t \approx G(X_t)\Delta t$.
- **Approximate Solution:** To approximate ODE on $[0, T]$ at intervals of $\Delta t = T/N$:

  1. Let $\tilde{X}_0 = x_0$.

  2. Let $\tilde{X}_{n+1} = \tilde{X}_n + G(\tilde{X}_n)\Delta t$.

- **Convergence:** If $\tilde{X}_t^{(N)}$ is a linear interpolation of the approximated path skeleton with $\Delta t = T/N$, then

$$\lim_{N\to\infty} \left\{ \sup_{0\le t\le T} (\tilde{X}_t^{(N)} - X_t)^2 \right\} = 0.$$

Euler Discretization

- **Stochastic Differential Equation:**

$$\mathrm{d}X_t = \mu(X_t)\,\mathrm{d}t + \sigma(X_t)\,\mathrm{d}B_t, \qquad X_0 = x_0.$$

  Typically these are intractable.

- **Euler Discretization:** Replace infinitesimal $\mathrm{d}t$ by time interval Δt:

$$\mathrm{d}X_t = \mu(X_t)\,\mathrm{d}t + \sigma(X_t)\,\mathrm{d}B_t \quad \Longrightarrow \quad \Delta X_t = \mu(X_t)\Delta t + \sigma(X_t)\Delta B_t,$$

  where $\Delta X_t = X_{t+\Delta t} - X_t$, and $\Delta B_t = B_{t+\Delta t} - B_t \sim \mathcal{N}(0, \Delta t)$.

- **Approximate Solution:** To approximate SDE on $[0, T]$ at intervals of $\Delta t = T/N$:

  1. Let $\tilde{X}_0 = x_0$.
  2. Let $\tilde{X}_{n+1} = \tilde{X}_n + \mu(\tilde{X}_n)\Delta t + \sigma(\tilde{X}_n)\Delta B_n, \qquad \Delta B_0, \ldots, \Delta B_{N-1} \overset{\text{iid}}{\sim} \mathcal{N}(0, \Delta t)$.

Euler Discretization

- **Stochastic Differential Equation:** $\mathrm{d}X_t = \mu(X_t)\,\mathrm{d}t + \sigma(X_t)\,\mathrm{d}B_t$.
- **Euler Discretization:**
$\mathrm{d}X_t = \mu(X_t)\,\mathrm{d}t + \sigma(X_t)\,\mathrm{d}B_t \implies \Delta X_t = \mu(X_t)\Delta t + \sigma(X_t)\Delta B_t$.
- **Approximate Solution:** To approximate SDE on $[0, T]$ at intervals of $\Delta t = T/N$:

  1. Let $\tilde{X}_0 = x_0$.
  2. Let $\tilde{X}_{n+1} = \tilde{X}_n + \mu(\tilde{X}_n)\Delta t + \sigma(\tilde{X}_n)\Delta B_n, \qquad \Delta B_0, \ldots, \Delta B_{N-1} \overset{\text{iid}}{\sim} \mathcal{N}(0, \Delta t)$.

- **Convergence:** If $\tilde{X}_t^{(N)}$ is a linear interpolation of the approximated path skeleton with $\Delta t = T/N$, then

$$\lim_{N\to\infty} \left\{ \sup_{0\le t\le T} E\left[(\tilde{X}_t^{(N)} - X_t)^2 \right] \right\} = 0.$$

Example: Geometric Brownian Motion

- **Model:** $\mathrm{d}S_t = \alpha S_t \Delta t + \sigma S_t \, \mathrm{d}B_t.$

  ("S" stands for stock: this is the model used for GSPC)

- **Analytic Solution:** $S_t = S_0 \cdot \exp\left\{(\alpha - \tfrac{1}{2}\sigma^2)t + \sigma B_t\right\}.$

  (In fact, $X_t = \log(S_t)$ is Brownian Motion + drift: $\mathrm{d}X_t = (\alpha + \tfrac{1}{2}\sigma^2)\,\mathrm{d}t + \sigma\,\mathrm{d}B_t.$)

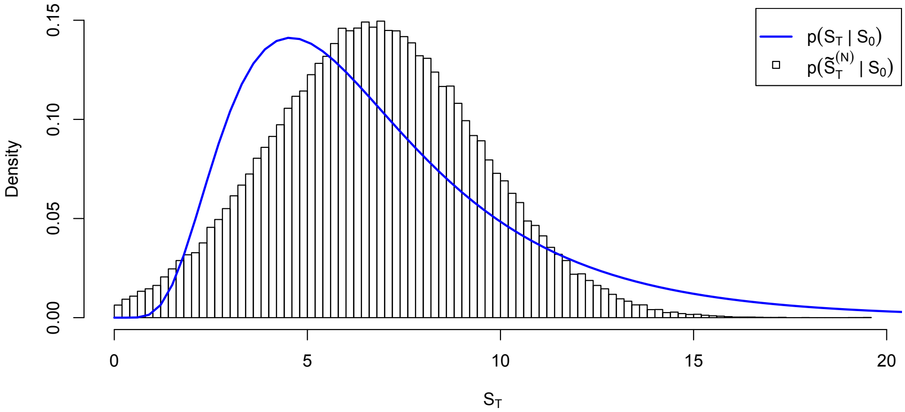
- **Euler Approximation:** $S_t \approx S_0 + \alpha S_0 t + \sigma S_0 B_t.$

- **Simulation Study:** Compare true $p(S_T \mid S_0)$ to approximate $p(\tilde{S}_T^{(N)} \mid S_0)$ for

$$\alpha = .1, \quad \sigma = .3, \quad S_0 = 5.1, \quad T = 3.2.$$

Example: Geometric Brownian Motion

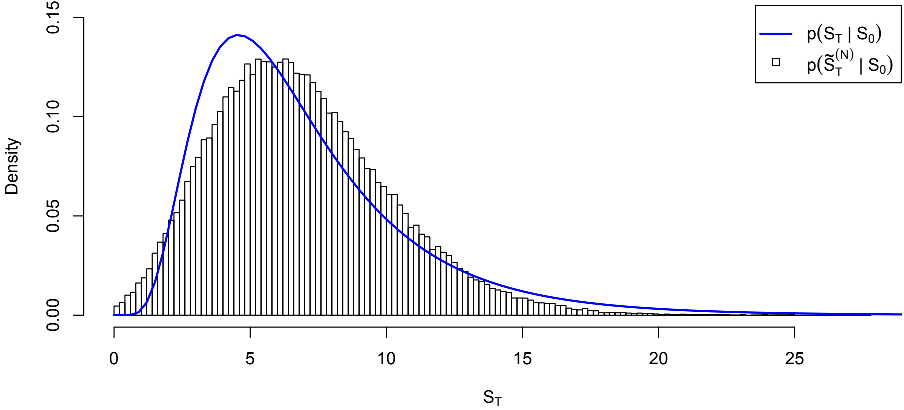
- **Model:** $\mathrm{d}S_t = \alpha S_t \,\mathrm{d}t + \sigma S_t \,\mathrm{d}B_t$
- **Simulation:** $\alpha = .1,\ \sigma = .3,\ S_0 = 5.1,\ T = 3.2$ – $N = 1$ $(\Delta t = T/N)$



(Note that $S_T > 0$ but $\tilde{S}_T^{(N)} \in \mathbb{R}$)

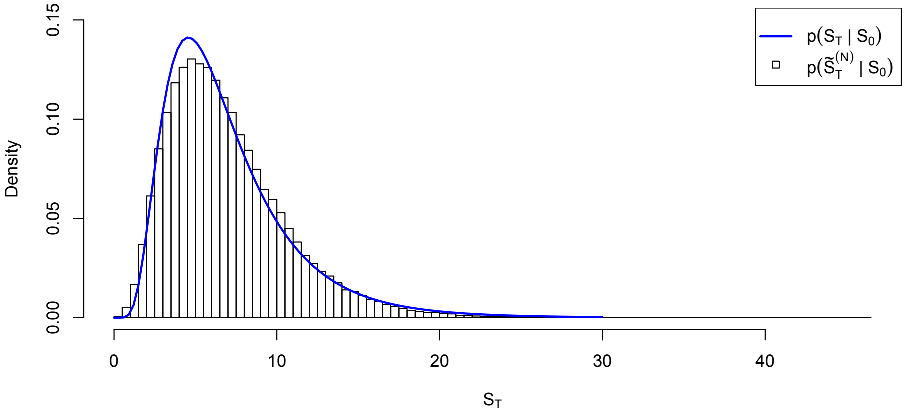
Example: Geometric Brownian Motion

- **Model:** $\mathrm{d}S_t = \alpha S_t \,\mathrm{d}t + \sigma S_t \,\mathrm{d}B_t$
- **Simulation:** $\alpha = .1,\ \sigma = .3,\ S_0 = 5.1,\ T = 3.2$ – $N = 2$ $(\Delta t = T/N)$



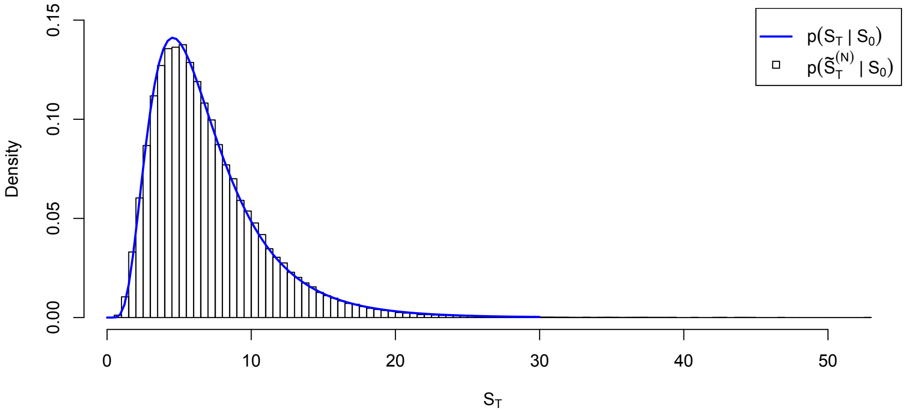
Example: Geometric Brownian Motion

- **Model:** $\mathrm{d}S_t = \alpha S_t \, \mathrm{d}t + \sigma S_t \, \mathrm{d}B_t$
- **Simulation:** $\alpha = .1$, $\sigma = .3$, $S_0 = 5.1$, $T = 3.2$ – $N = 5$ $(\Delta t = T/N)$



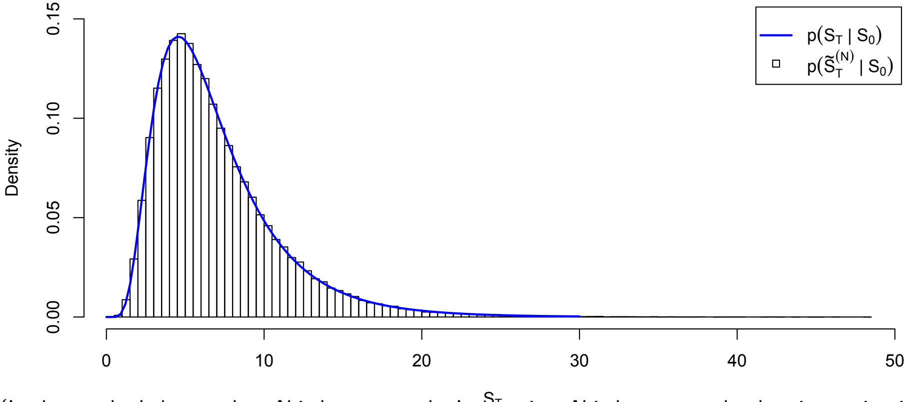
Example: Geometric Brownian Motion

- **Model:** $\mathrm{d}S_t = \alpha S_t \, \mathrm{d}t + \sigma S_t \, \mathrm{d}B_t$
- **Simulation:** $\alpha = .1,\ \sigma = .3,\ S_0 = 5.1,\ T = 3.2$ – $N = 20$ $(\Delta t = T/N)$



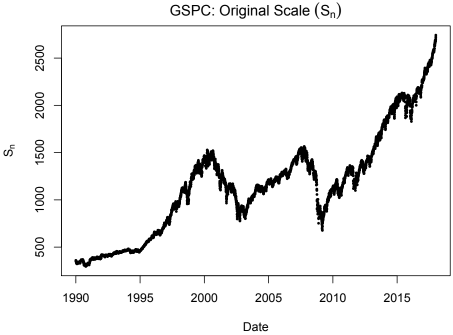
Example: Geometric Brownian Motion

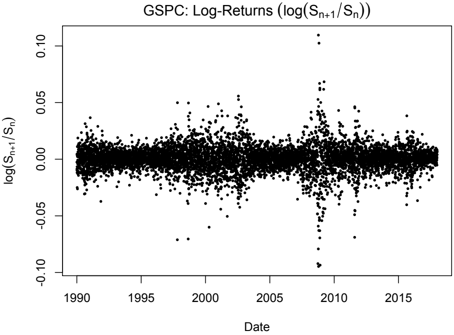
- **Model:** $\mathrm{d}S_t = \alpha S_t \,\mathrm{d}t + \sigma S_t \,\mathrm{d}B_t$
- **Simulation:** $\alpha = .1,\ \sigma = .3,\ S_0 = 5.1,\ T = 3.2$ – $N = 100$ $(\Delta t = T/N)$



(In theory, don't know when N is large enough. In practice, N is large enough when increasing it doesn't change result.)

Motivation (Continued)





- **Data:** Daily GSPC closing prices $\boldsymbol{S} = (S_1, \ldots, S_N)$
- **Model:** From 1900 till about 1989, this was a stochastic differential equation (SDE) called Geometric Brownian motion:

$$\mathrm{d}S_t = \alpha S_t \, \mathrm{d}t + \sigma S_t \, \mathrm{d}B_t,$$

where S_t is the price at time $t \in \mathbb{R}$ and B_t is Brownian motion.

Motivation

- **Data:** Daily GSPC closing prices $\boldsymbol{S} = (S_1, \ldots, S_N)$.

  Time between observations is $\Delta t = 1/252$ years (as there are 252 trading days in one year).

- **Model:** $\mathrm{d}S_t = \alpha S_t \, \mathrm{d}t + \sigma S_t \, \mathrm{d}B_t$.

- **Inference:** Let $X_t = \log(S_t) = (\alpha - \frac{1}{2}\sigma^2)t + \sigma B_t$. Then

$$\Delta X_i \overset{\text{iid}}{\sim} \mathcal{N}\big((\alpha - \tfrac{1}{2}\sigma^2)\Delta t, \sigma^2 \Delta t\big).$$

- **MLE:**
  - Let $\mu = (\alpha - \frac{1}{2}\sigma^2)\Delta t$ and $\tau^2 = \sigma^2 \Delta t$.
  - Thus $\hat{\mu} = \frac{1}{N-1} \sum_{i=1}^{N-1} \Delta X_i$ and $\hat{\tau}^2 = \frac{1}{N-1} \sum_{i=1}^{N-1} (\Delta X_i - \hat{\mu})^2$.

    (There are $M = N - 1$ increments, and MLE of variance divides by sample size M – not $M - 1$, which is unbiased estimate)

Motivation

- **Data:** Daily GSPC closing prices $\boldsymbol{S} = (S_1, \dots, S_N)$, $\Delta t = 1/252$.
- **Model:** $\mathrm{d}S_t = \alpha S_t \,\mathrm{d}t + \sigma S_t \,\mathrm{d}B_t$.
- **Inference:** Let $X_t = \log(S_t)$. Then $\Delta X_i \overset{\text{iid}}{\sim} \mathcal{N}\big((\alpha - \frac{1}{2}\sigma^2)\Delta t, \sigma^2 \Delta t\big)$.
- **MLE:**
  - Let $\mu = (\alpha - \frac{1}{2}\sigma^2)\Delta t$ and $\tau^2 = \sigma^2 \Delta t$.
  - Thus $\hat{\mu} = \frac{1}{N-1}\sum_{i=1}^{N-1} \Delta X_i$ and $\hat{\tau}^2 = \frac{1}{N-1}\sum_{i=1}^{N-1}(\Delta X_i - \hat{\mu})^2$.
  - Invariance Property: If the likelihood function is $\ell(\boldsymbol{\theta} \mid \boldsymbol{X})$ and $\boldsymbol{\eta} = g(\boldsymbol{\theta})$ is an invertible function, then

$$\hat{\boldsymbol{\theta}} = \arg\max_{\boldsymbol{\theta}} \ell(\boldsymbol{\theta} \mid \boldsymbol{X}) \qquad \Longrightarrow \qquad \hat{\boldsymbol{\eta}} = g(\hat{\boldsymbol{\theta}}) = \arg\max_{\boldsymbol{\eta}} \ell(\boldsymbol{\eta} \mid \boldsymbol{X})$$

Proof: Whether you parametrize by $\boldsymbol{\theta}$ or $\boldsymbol{\eta}$, the likelihood function takes on the same set of values, i.e., $\ell(\boldsymbol{\theta} \mid \boldsymbol{X}) = \ell(\boldsymbol{\eta} = g(\boldsymbol{\theta}) \mid \boldsymbol{X})$. Therefore,

$$\ell(\boldsymbol{\eta} \mid \boldsymbol{X}) = \ell(\boldsymbol{\theta} = g^{-1}(\boldsymbol{\eta}) \mid \boldsymbol{X}) \le \ell(\hat{\boldsymbol{\theta}} \mid \boldsymbol{X}) = \ell(\boldsymbol{\eta} = g(\hat{\boldsymbol{\theta}}) \mid \boldsymbol{X}).$$

Motivation

- **Data:** Daily GSPC closing prices $\boldsymbol{S} = (S_1, \ldots, S_N)$, $\Delta t = 1/252$.
- **Model:** $\mathrm{d}S_t = \alpha S_t \, \mathrm{d}t + \sigma S_t \, \mathrm{d}B_t$.
- **Inference:** Let $X_t = \log(S_t)$. Then $\Delta X_i \overset{\text{iid}}{\sim} \mathcal{N}\big((\alpha - \tfrac{1}{2}\sigma^2)\Delta t, \sigma^2 \Delta t\big)$.
- **MLE:**
  - Let $\mu = (\alpha - \frac{1}{2}\sigma^2)\Delta t$ and $\tau^2 = \sigma^2 \Delta t$.
  - Thus $\hat{\mu} = \frac{1}{N-1} \sum_{i=1}^{N-1} \Delta X_i$ and $\hat{\tau}^2 = \frac{1}{N-1} \sum_{i=1}^{N-1} (\Delta X_i - \hat{\mu})^2$.
  - Invariance Property: $\hat{\sigma} = \sqrt{\hat{\tau}^2 / \Delta t}$ and $\hat{\alpha} = \hat{\mu}/\Delta t + \frac{1}{2}\hat{\sigma}^2$.

Motivation

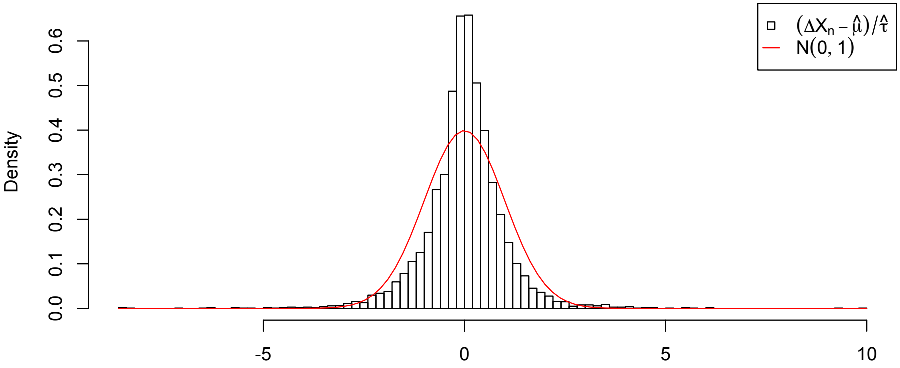
- **Data:** Daily GSPC closing prices $\boldsymbol{S} = (S_1, \ldots, S_N)$, $\Delta t = 1/252$.
- **Model:** $\mathrm{d}S_t = \alpha S_t \,\mathrm{d}t + \sigma S_t \,\mathrm{d}B_t$.
- **Inference:** Let $X_t = \log(S_t)$. Then $\Delta X_i \overset{\text{iid}}{\sim} \mathcal{N}\big((\alpha - \frac{1}{2}\sigma^2)\Delta t, \sigma^2 \Delta t\big)$.
- **MLE:**
  - Let $\mu = (\alpha - \frac{1}{2}\sigma^2)\Delta t$ and $\tau^2 = \sigma^2 \Delta t$.
  - Thus $\hat{\mu} = \frac{1}{N-1}\sum_{i=1}^{N-1} \Delta X_i$ and $\hat{\tau}^2 = \frac{1}{N-1}\sum_{i=1}^{N-1}(\Delta X_i - \hat{\mu})^2$.
  - *Invariance Property:* $\hat{\sigma} = \sqrt{\hat{\tau}^2/\Delta t}$ and $\hat{\alpha} = \hat{\mu}/\Delta t + \frac{1}{2}\hat{\sigma}^2$.
- **Goodness-of-Fit:** If model is correct, then approximately we have

$$Z_i = (\Delta X_i - \hat{\mu})/\hat{\tau} \overset{\text{iid}}{\sim} \mathcal{N}(0, 1).$$

Motivation

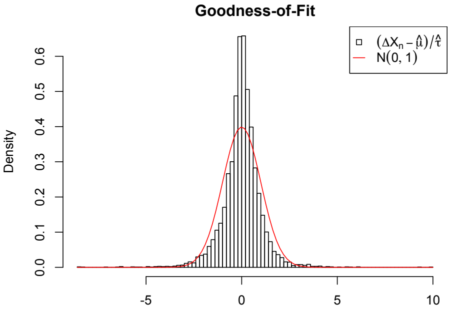
- **Data:** Daily GSPC closing prices $\boldsymbol{S} = (S_1, \ldots, S_N)$, $\Delta t = 1/252$.
- **Model:** $\mathrm{d}S_t = \alpha S_t \,\mathrm{d}t + \sigma S_t \,\mathrm{d}B_t$. Let $X_t = \log(S_t)$. Then $\Delta X_i \overset{\text{iid}}{\sim} \mathcal{N}(\mu, \tau^2)$.
- **Goodness-of-Fit:** If model is correct, then approximately we have

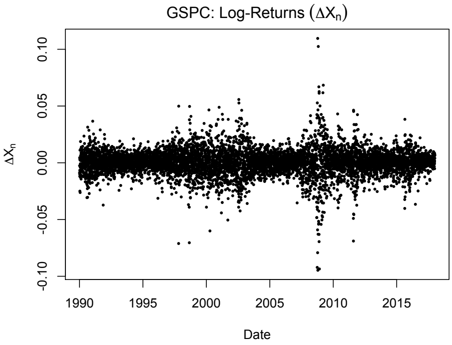
$$Z_i = (\Delta X_i - \hat{\mu})/\hat{\tau} \overset{\text{iid}}{\sim} \mathcal{N}(0, 1).$$



Motivation

- **Data:** Daily GSPC closing prices $\boldsymbol{S} = (S_1, \ldots, S_N)$, $\Delta t = 1/252$.
- **Model:** $\mathrm{d}S_t = \alpha S_t \,\mathrm{d}t + \sigma S_t \,\mathrm{d}B_t$. Let $X_t = \log(S_t)$. Then $\Delta X_i \overset{\text{iid}}{\sim} \mathcal{N}(\mu, \tau^2)$.
- **Goodness-of-Fit:** If model is correct, then $(\Delta X_i - \hat{\mu})/\hat{\tau} \approx \mathcal{N}(0, 1)$.





Assumption of constant volatility (σ) is violated.

Stochastic Volatility Modeling

- S_t: value of GSPC at time t.
- **Constant-Volatility Model:** Let $X_t = \log(S_t)$. Then gBM model on log scale is

$$\mathrm{d}X_t = (\alpha - \tfrac{1}{2}\sigma^2)\,\mathrm{d}t + \sigma\,\mathrm{d}B_t.$$

- **Stochastic Volatility Model:**
  - *Idea:* Replace σ^2 by time-dependent volatility V_t: $\mathrm{d}X_t = (\alpha - \frac{1}{2}V_t)\,\mathrm{d}t + {V_t}^{1/2}\,\mathrm{d}B_t$
  - *Problem:* V_t is an infinite-dimensional parameter, so need to *regularize* to estimate it.
  - *Solution:* Give another SDE to V_t. Thus we get the two-dimensional process

$$\begin{aligned}\mathrm{d}X_t &= (\alpha - \tfrac{1}{2}V_t)\,\mathrm{d}t + V_t^{1/2}\,\mathrm{d}B_{1t},\\ \mathrm{d}V_t &= -\gamma(V_t - \mu)\,\mathrm{d}t + \sigma V_t^{\lambda}\,\mathrm{d}B_{2t}.\end{aligned}$$

Stochastic Volatility Modeling

- S_t: value of GSPC at time t.
- **Stochastic Volatility Model:**
  - *Idea:* Replace σ^2 by time-dependent volatility V_t, and give another SDE to V_t. Thus we get the two-dimensional process

$$
\begin{aligned}
\mathrm{d}X_t &= (\alpha - \tfrac{1}{2} V_t)\,\mathrm{d}t + V_t^{1/2}\,\mathrm{d}B_{1t}, \\
\mathrm{d}V_t &= -\gamma(V_t - \mu)\,\mathrm{d}t + \sigma V_t^{\lambda}\,\mathrm{d}B_{2t}.
\end{aligned}
$$

  - *Parameter Interpretation:*
    - α: "interest" rate: $E[S_t] = S_0\, e^{\alpha t}$
    - $\mu = E[V_t]$: long-run mean
    - γ: mean-reversion: $\mathrm{cor}(V_s, V_{s+t}) = e^{-\gamma t}$
    - σ and λ: scale and shape

Stochastic Volatility Modeling

- S_t: value of GSPC at time t.
- **Stochastic Volatility Model:**
  - *Idea:* Replace σ^2 by time-dependent volatility V_t, and give another SDE to V_t. Thus we get the two-dimensional process

$$\begin{aligned} \mathrm{d}X_t &= (\alpha - \tfrac{1}{2} V_t)\,\mathrm{d}t + V_t^{1/2}\,\mathrm{d}B_{1t}, \\ \mathrm{d}V_t &= -\gamma(V_t - \mu)\,\mathrm{d}t + \sigma V_t^{\lambda}\,\mathrm{d}B_{2t}. \end{aligned}$$

  - Euler Approximation:

$$\begin{bmatrix} \Delta X_t \\ \Delta V_t \end{bmatrix} \sim \mathcal{N}\left\{ \begin{bmatrix} \alpha - \frac{1}{2} V_t \\ -\gamma(V_t - \mu) \end{bmatrix} \Delta t, \begin{bmatrix} V_t & 0 \\ 0 & \sigma^2 V_t^{2\lambda} \end{bmatrix} \Delta t \right\}$$

Parameter Inference

- **Data:** Daily GSPC log-prices $\boldsymbol{X} = (X_1, \dots, X_N)$, $\Delta t = 1/252$.
- **Stochastic Volatility Model:**

$$
\begin{aligned}
\mathrm{d}X_t &= (\alpha - \tfrac{1}{2} V_t)\,\mathrm{d}t + V_t^{1/2}\,\mathrm{d}B_{1t}, \\
\mathrm{d}V_t &= -\gamma(V_t - \mu)\,\mathrm{d}t + \sigma V_t^{\lambda}\,\mathrm{d}B_{2t}.
\end{aligned}
$$

- **Likelihood:** $\ell(\boldsymbol{\theta} \mid \boldsymbol{X}) = ???$ Two complications:

1. $\mathcal{L}(\boldsymbol{\theta} \mid \boldsymbol{X}, \boldsymbol{V}) \propto p(\boldsymbol{X}, \boldsymbol{V} \mid \boldsymbol{\theta}) = \prod_{n=1}^{N-1} p(X_{n+1}, V_{n+1} \mid X_n, V_n, \boldsymbol{\theta})$.

   However, the *transition density* $p(X_{n+1}, V_{n+1} \mid X_n, V_n, \boldsymbol{\theta})$ is not available in closed form (solution to a PDE).

2. $\mathcal{L}(\boldsymbol{\theta} \mid \boldsymbol{X}) \propto p(\boldsymbol{X} \mid \boldsymbol{\theta}) = \int p(\boldsymbol{X}, \boldsymbol{V} \mid \boldsymbol{\theta})\,\mathrm{d}\boldsymbol{V}$.

   That is, the volatility process is *latent* (completely unobserved), and hence must be integrated out.

Parameter Inference

- **Data:** Daily GSPC log-prices $\boldsymbol{X} = (X_1, \dots, X_N)$, $\Delta t = 1/252$.
- **Stochastic Volatility Model:**

$$\begin{aligned} \mathrm{d}X_t &= (\alpha - \tfrac{1}{2} V_t)\,\mathrm{d}t + V_t^{1/2}\,\mathrm{d}B_{1t}, \\ \mathrm{d}V_t &= -\gamma(V_t - \mu)\,\mathrm{d}t + \sigma V_t^{\lambda}\,\mathrm{d}B_{2t}. \end{aligned}$$

- **Likelihood:** $\ell(\boldsymbol{\theta} \mid \boldsymbol{X}) = ???$ Two complications:

1. $\mathcal{L}(\boldsymbol{\theta} \mid \boldsymbol{X}, \boldsymbol{V}) \propto p(\boldsymbol{X}, \boldsymbol{V} \mid \boldsymbol{\theta}) = \prod_{n=1}^{N-1} p(X_{n+1}, V_{n+1} \mid X_n, V_n, \boldsymbol{\theta})$.

   However, $p(X_{n+1}, V_{n+1} \mid X_n, V_n, \boldsymbol{\theta})$ is intractable.

   Solution: Use Euler approximation!

$$\begin{bmatrix} \Delta X_{n+1} \\ \Delta V_{n+1} \end{bmatrix} \sim \mathcal{N}\left\{ \begin{bmatrix} \alpha - \frac{1}{2} V_n \\ -\gamma(V_n - \mu) \end{bmatrix} \Delta t, \begin{bmatrix} V_n & 0 \\ 0 & \sigma^2 V_n^{2\lambda} \end{bmatrix} \Delta t \right\}$$

   (Results are approximate, but reportedly good enough for daily financial data)

Parameter Inference

- **Data:** Daily GSPC log-prices $\boldsymbol{X} = (X_1, \dots, X_N)$, $\Delta t = 1/252$.
- **Stochastic Volatility Model:**

$$\mathrm{d}X_t = (\alpha - \tfrac{1}{2}V_t)\,\mathrm{d}t + V_t^{1/2}\,\mathrm{d}B_{1t},$$
$$\mathrm{d}V_t = -\gamma(V_t - \mu)\,\mathrm{d}t + \sigma V_t^{\lambda}\,\mathrm{d}B_{2t}.$$

- **Likelihood:** $\ell(\boldsymbol{\theta} \mid \boldsymbol{X}) = ???$ Two complications:

1. $\mathcal{L}(\boldsymbol{\theta} \mid \boldsymbol{X}, \boldsymbol{V}) \propto p(\boldsymbol{X}, \boldsymbol{V} \mid \boldsymbol{\theta}) = \prod_{n=1}^{N-1} p(X_{n+1}, V_{n+1} \mid X_n, V_n, \boldsymbol{\theta})$.

   However, $p(X_{n+1}, V_{n+1} \mid X_n, V_n, \boldsymbol{\theta})$ is intractable.

   Solution: Use Euler approximation!

2. $\mathcal{L}(\boldsymbol{\theta} \mid \boldsymbol{X}) \propto p(\boldsymbol{X} \mid \boldsymbol{\theta}) = \int p(\boldsymbol{X}, \boldsymbol{V} \mid \boldsymbol{\theta})\,\mathrm{d}\boldsymbol{V}$.

   That is, the latent volatility process must be integrated out.

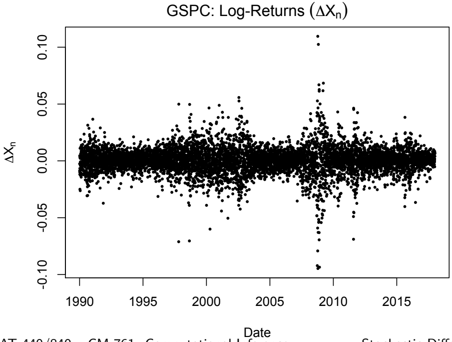
   Solution: Replace V_t by an observable market proxy.

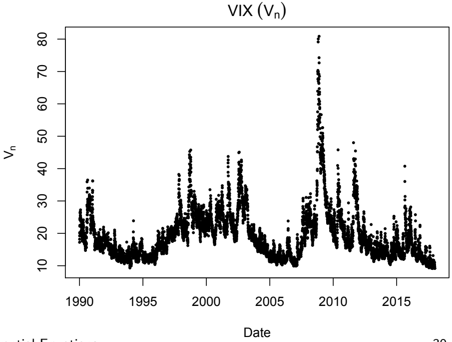
Parameter Inference

- **Data:**
  - X_i: log GSPC value on day i
  - V_i: VIX value on day i (measure of implied volatility determined by CBOE)
- **Stochastic Volatility Model:**

$$\begin{aligned} dX_t &= (\alpha - \tfrac{1}{2}\tau V_t)\,dt + (\tau V_t)^{1/2}\,dB_{1t}, \\ dV_t &= -\gamma(V_t - \mu)\,dt + \sigma V_t^{\lambda}\,dB_{2t}. \end{aligned}$$

(τ is the VIX scale adjustment)





Parameter Inference

- **Data:** $\boldsymbol{Y}_i = (X_i, V_i)$: (log-GSPC, VIX) pair on day i. $\boldsymbol{Y} = (\boldsymbol{Y}_1, \ldots, \boldsymbol{Y}_N)$.
- **Stochastic Volatility Model:**

$$
\begin{aligned}
\mathrm{d}X_t &= (\alpha - \tfrac{1}{2}\tau V_t)\,\mathrm{d}t + (\tau V_t)^{1/2}\,\mathrm{d}B_{1t} \\
\mathrm{d}V_t &= -\gamma(V_t - \mu)\,\mathrm{d}t + \sigma V_t^{\lambda}\,\mathrm{d}B_{2t},
\end{aligned}
$$

- **Loglikelihood:** For $\boldsymbol{\theta} = (\alpha, \gamma, \mu, \sigma, \lambda, \tau)$, Euler approximation gives

$$
\begin{aligned}
\ell(\boldsymbol{\theta} \mid \boldsymbol{Y}) &= \sum_{i=1}^{N-1} \log p(\boldsymbol{Y}_{i+1} \mid \boldsymbol{Y}_i, \boldsymbol{\theta}) \\
&\approx -\frac{1}{2} \sum_{i=1}^{N-1} \left\{ \frac{\left[\Delta X_i - (\alpha - \frac{1}{2}\tau V_i)\Delta t\right]^2}{\tau V_i \Delta t} + \log(\tau V_i \Delta t) \right\} \\
&\quad - \frac{1}{2} \sum_{i=1}^{N-1} \left\{ \frac{[\Delta V_i + \gamma(V_i - \mu)\Delta t]^2}{\sigma^2 V_i^{2\lambda} \Delta t} + \log(\sigma^2 V_i^{2\lambda} \Delta t) \right\}
\end{aligned}
$$

(each term is the log-pdf of a univariate normal)

Parameter Inference

- **Data:** $\boldsymbol{Y}_i = (X_i, V_i)$: (log-GSPC, VIX) pair on day i. $\boldsymbol{Y} = (\boldsymbol{Y}_1, \ldots, \boldsymbol{Y}_N)$.
- **Stochastic Volatility Model:**

$$\begin{aligned} \mathrm{d}X_t &= (\alpha - \tfrac{1}{2}\tau V_t)\,\mathrm{d}t + (\tau V_t)^{1/2}\,\mathrm{d}B_{1t} \\ \mathrm{d}V_t &= -\gamma(V_t - \mu)\,\mathrm{d}t + \sigma V_t^{\lambda}\,\mathrm{d}B_{2t}, \end{aligned}$$

- **Loglikelihood:** Let $\boldsymbol{\theta} = (\alpha, \gamma, \mu, \sigma, \lambda, \tau)$.
  - Euler approximation:

$$\ell(\boldsymbol{\theta} \mid \boldsymbol{Y}) \approx \sum_{i=1}^{N-1} \log \varphi_X(\Delta X_i \mid \boldsymbol{Y}_i, \boldsymbol{\theta}) + \log \varphi_V(\Delta V_i \mid \boldsymbol{Y}_i, \boldsymbol{\theta}),$$

    where $\varphi_X(\cdot \mid \boldsymbol{Y}, \boldsymbol{\theta})$ and $\varphi_V(\cdot \mid \boldsymbol{Y}, \boldsymbol{\theta})$ are univariate normal PDFs specified by SV model.

  - *MLE:* $\hat{\boldsymbol{\theta}} = \arg\max_{\boldsymbol{\theta}} \ell(\boldsymbol{\theta} \mid \boldsymbol{Y})$ does not have an analytic solution, so use R function `optim` for numerical optimization. (See `?optim` for details!)

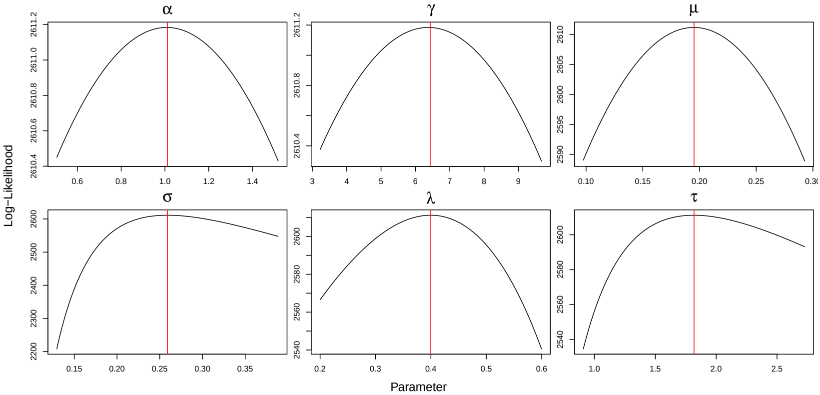
Parameter Inference

- **Stochastic Volatility Model:**

$$\mathrm{d}X_t = (\alpha - \tfrac{1}{2}\tau V_t)\,\mathrm{d}t + (\tau V_t)^{1/2}\,\mathrm{d}B_{1t}$$
$$\mathrm{d}V_t = -\gamma(V_t - \mu)\,\mathrm{d}t + \sigma V_t^{\lambda}\,\mathrm{d}B_{2t}.$$

- **Simulation:**

Parameter	α	γ	μ	σ	λ	τ	X_0	V_0	N	Δt
Value	.1	.5	.2	.3	.5	2	6.5	.2	500	1/252



Good algorithm initialization: $\boldsymbol{\theta}_{\text{init}} = \boldsymbol{\theta}_{\text{true}}$.

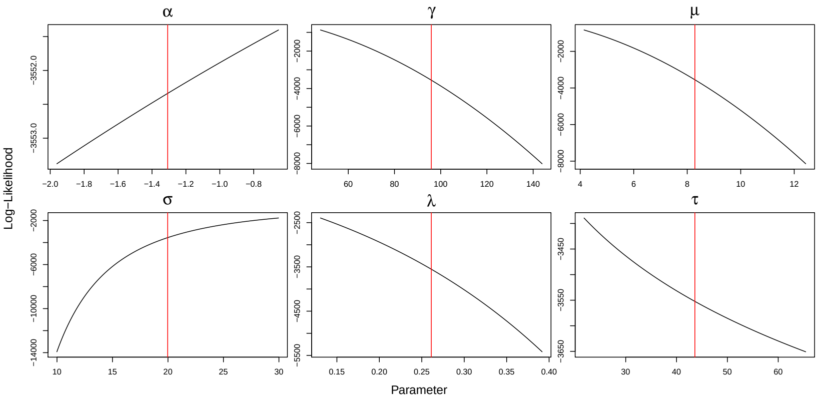
Parameter Inference

- **Stochastic Volatility Model:**

$$\mathrm{d}X_t = (\alpha - \tfrac{1}{2}\tau V_t)\,\mathrm{d}t + (\tau V_t)^{1/2}\,\mathrm{d}B_{1t}$$
$$\mathrm{d}V_t = -\gamma(V_t - \mu)\,\mathrm{d}t + \sigma V_t^{\lambda}\,\mathrm{d}B_{2t}.$$

- **Simulation:**

Parameter	α	γ	μ	σ	λ	τ	X_0	V_0	N	Δt
Value	.1	.5	.2	.3	.5	2	6.5	.2	500	1/252



Bad algorithm initialization: $\boldsymbol{\theta}_{\text{init}} = 10 \times \boldsymbol{\theta}_{\text{true}}$.

Resources

- `msde`: An R/C++ package for simulation/inference with multivariate SDEs having latent variables (e.g., SV model with V_t not proxied by VIX).