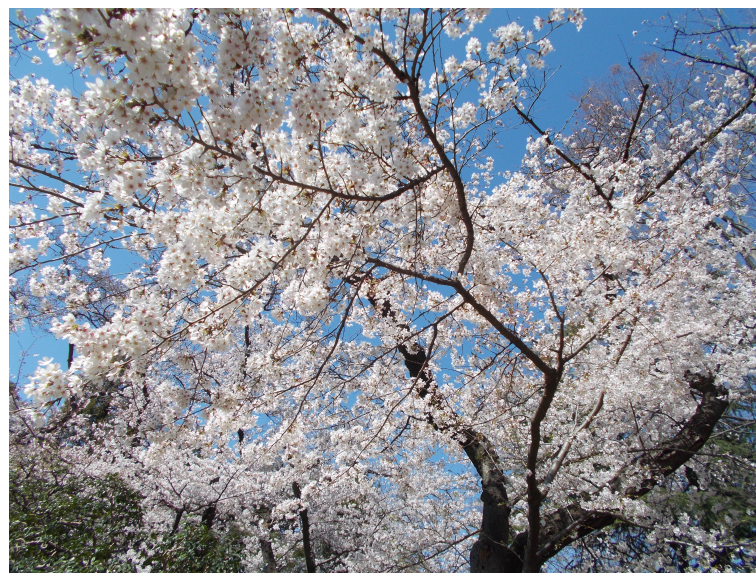


2023 年度公開講座「統計と数学」 令和 5 年 11 月 2 5 日 東京大学大学院数理科学研究科

極限定理

Limit theorems

吉田朋広
東京大学大学院数理科学研究科
JST CREST



極限定理と統計学

- 中心極限定理
- 中心極限定理と確率過程の統計学
- 極限定理研究の発展

中心極限定理

理想的なサイコロを投げる

- サイコロを振る

$$U_1 = i \quad (\text{目の数が } i \text{ であったとき}) \quad (i = 1, 2, \dots, 6)$$

- 理想的なサイコロ

$$P[\text{目の数が } i \text{ である確率}] = P[U_1 = i] = \frac{1}{6} \\ (i = 1, 2, \dots, 6)$$

- U_1 : 確率変数

- U_1 の 確率分布

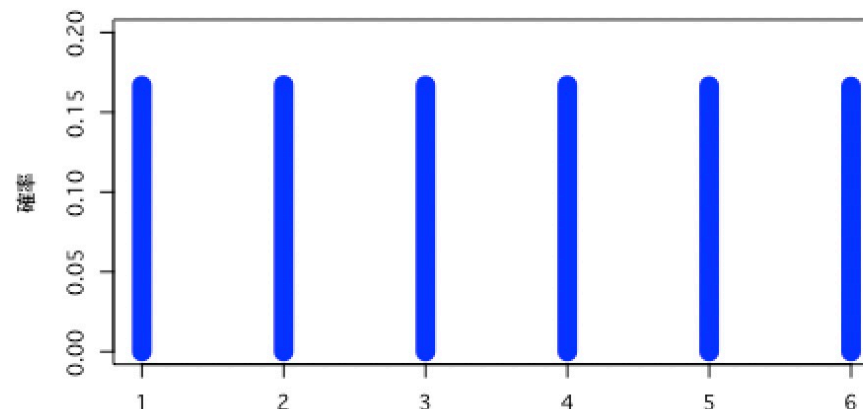


図 1: U_1

理想的なサイコロを 2 回投げる

- U_1 : 1 回目に出た目の数、 U_2 : 2 回目に出た目の数
- 出た目の数の和 $S_2 = U_1 + U_2$ が k になる確率:

$$P[S_2 = k] = P[U_1 + U_2 = k] = \frac{1}{36}(6 - |k - 7|).$$

ただし, $k = 2, \dots, 12$.

- 確率分布

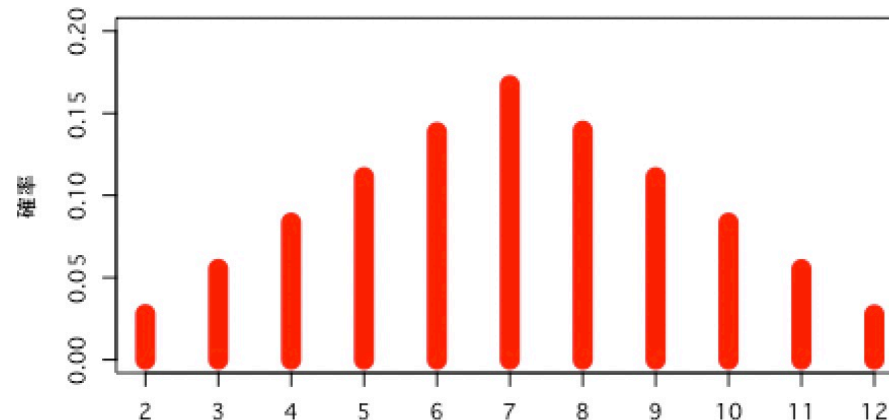
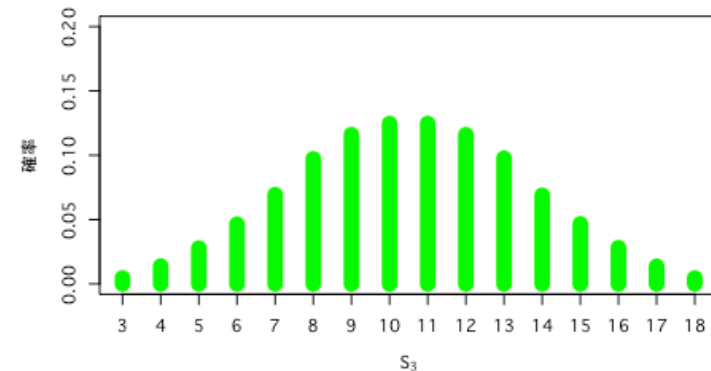
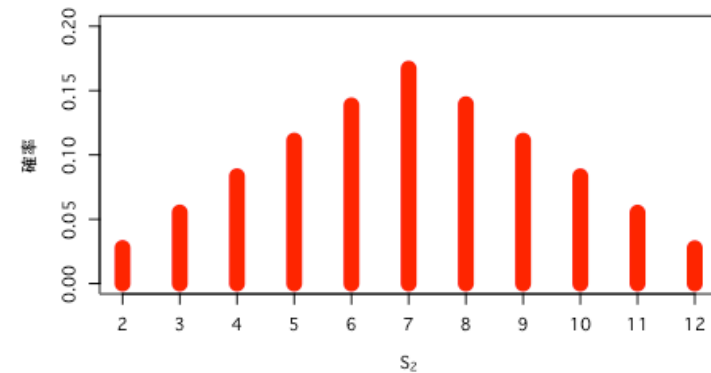
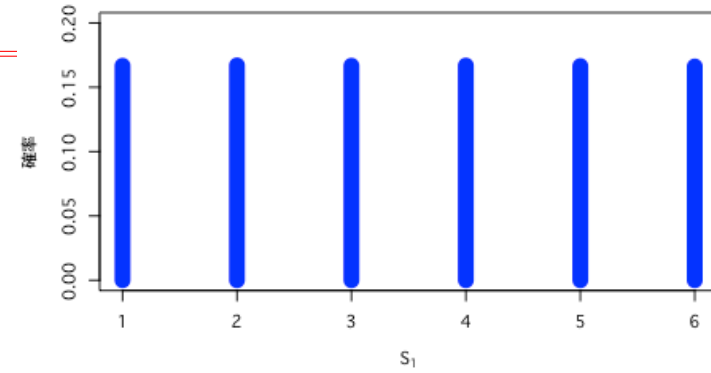


図 2: $U_1 + U_2$

サイコロの目の数の和の分布

- U_1, U_2, U_3 : サイコロの目の数
- $S_1 = U_1$
 $S_2 = U_1 + U_2$
 $S_3 = U_1 + U_2 + U_3$
- 各々 10^6 回 ($\doteq \infty$ 回) 振る
- S_1, S_2, S_3 の確率分布
 (ヒストグラム)
- 分布の形 四角→三角→山型



独立確率変数の和の分布

- つぎの分布を持つ確率変数 U_1 を考える：

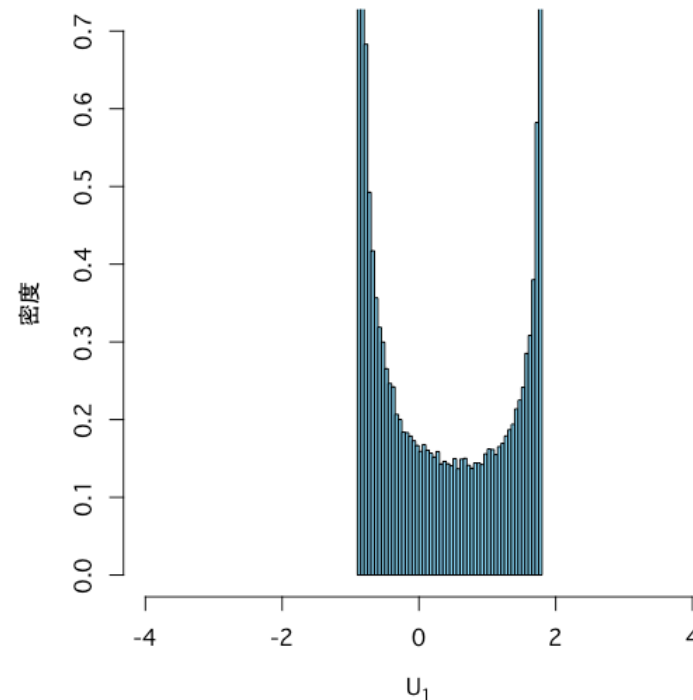


図 3: U_1

- 分布は汚いが，期待値 $E[U_1] = 0$ ，分散 $\text{Var}[U_1] = 1$ となるように調整されている。

独立確率変数の和の分布

- つぎの分布を持つ確率変数 U_1 を考える：

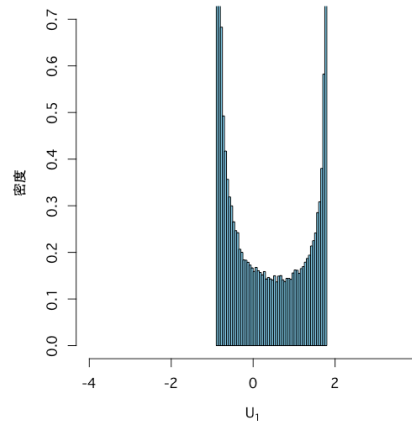


図 3: U_1

- U_1, U_2, \dots を独立な確率変数とし, U_j の分布は U_1 の分布と同じであるとする.
- $S_n = \sum_{j=1}^n U_j = U_1 + U_2 + \dots + U_n$ とする.
- $Z_n := \frac{1}{\sqrt{n}} S_n$ によって S_n を基準化する.

分布の収束

- Z_1, Z_3, Z_5, Z_{30} の確率分布 (ヒストグラム)

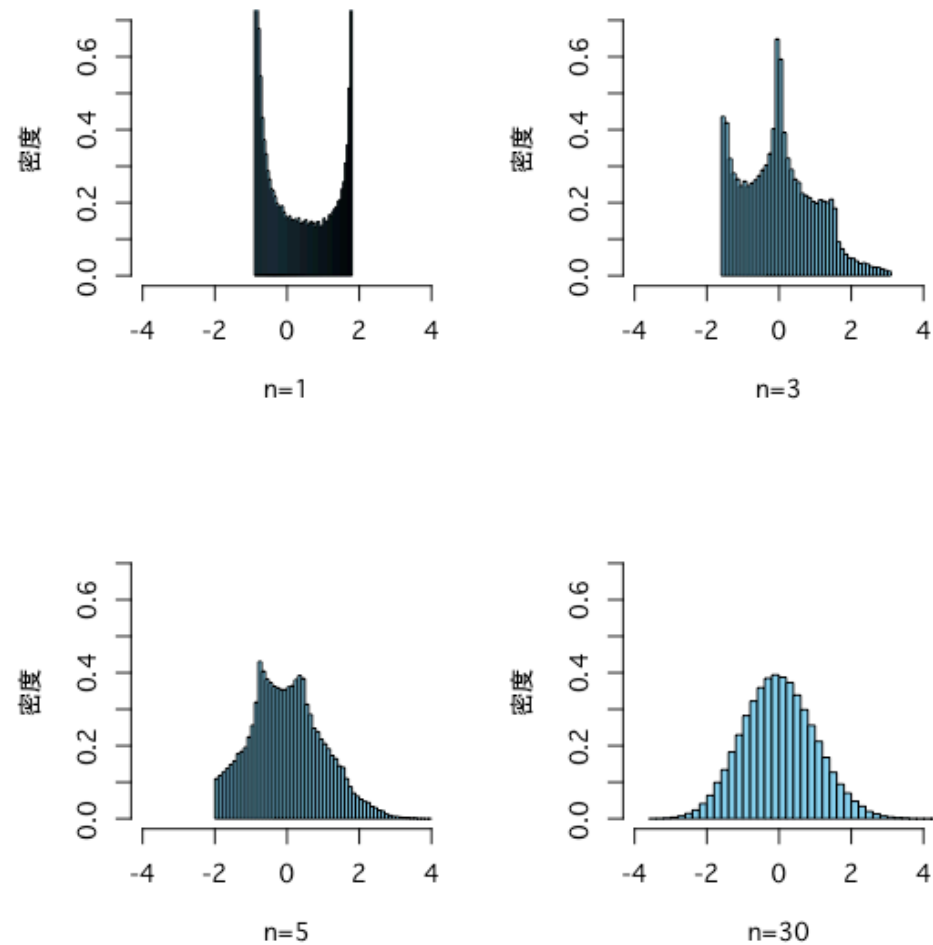


図 4: $Z_n = S_n/\sqrt{n}$

正規分布への収束

- Z_1, Z_3, Z_5, Z_{30} の確率分布 (ヒストグラム)

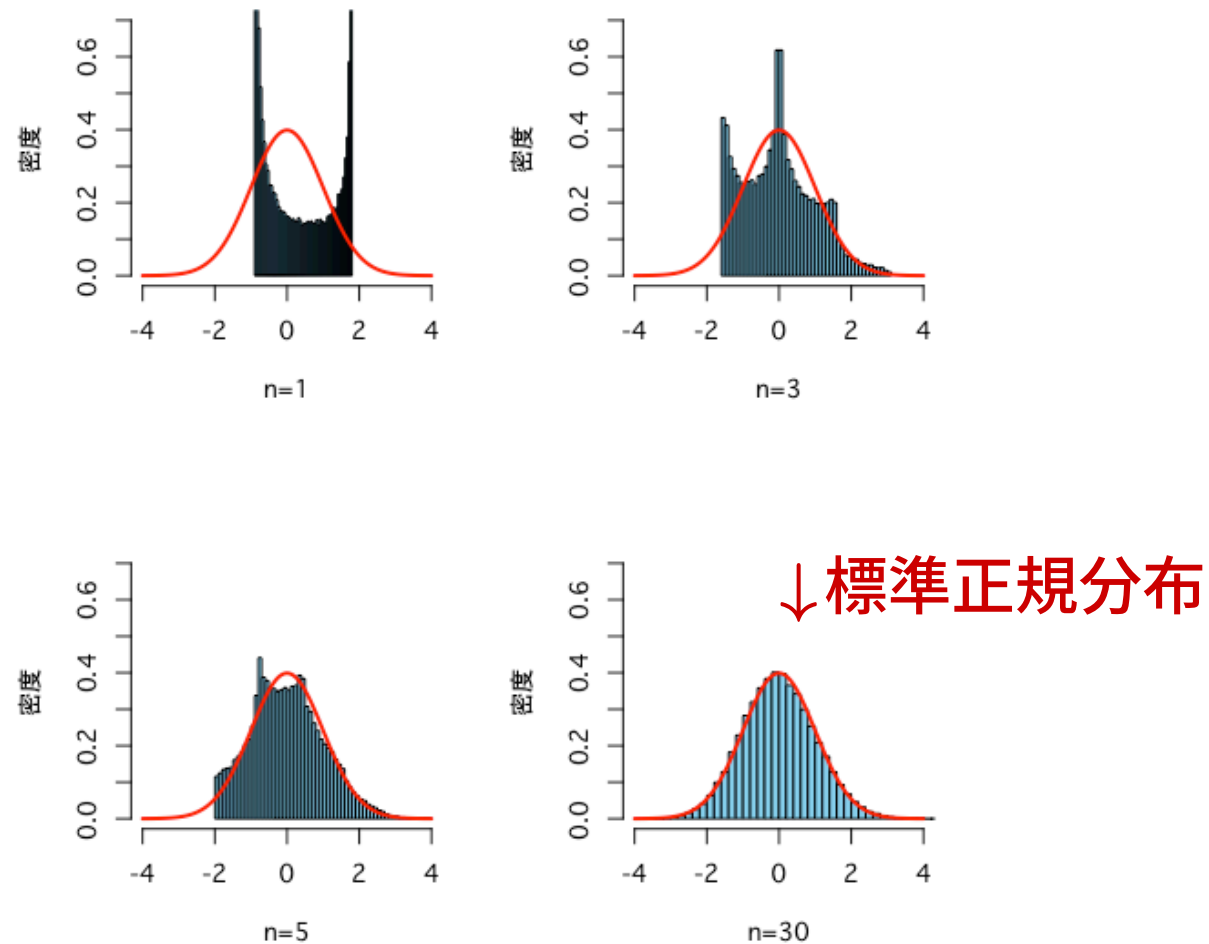


図 5: $Z_n = S_n / \sqrt{n}$

中心極限定理

定理 1. (中心極限定理 central limit theorem)

U_1, U_2, \dots : 独立で同一の分布に従う確率変数列、

$E[U_1] = 0, \text{Var}[U_1] = 1$ とする。

このとき

$$Z_n = \frac{1}{\sqrt{n}} S_n = \frac{1}{\sqrt{n}} (U_1 + U_2 + \dots + U_n)$$

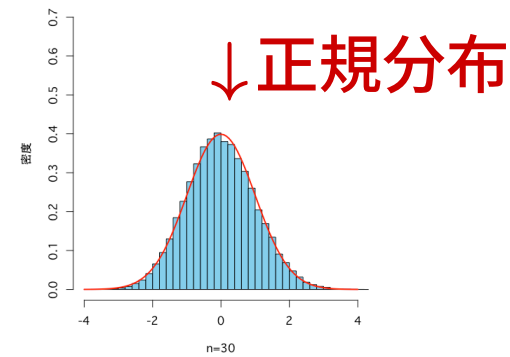
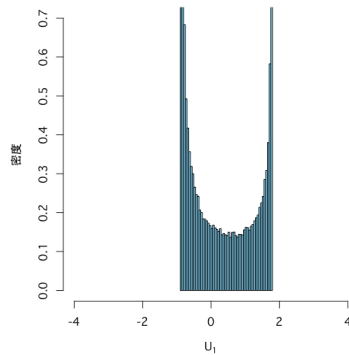
の確率分布は、 $n \rightarrow \infty$ のとき標準正規分布に収束する。

- この収束を $Z_n \xrightarrow{d} N(0, 1)$ ($n \rightarrow \infty$) と表す。
- 標準正規分布の確率密度関数 $\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}$
- ヒストグラムが収束するということ：任意の $a, b \in \mathbb{R}$ ($a < b$) に対して、

$$P[a < Z_n \leq b] \rightarrow \int_a^b \phi(z) dz \quad (n \rightarrow \infty)$$

中心極限定理 $Z_n = \frac{1}{\sqrt{n}}(U_1 + U_2 + \cdots + U_n) \xrightarrow{d} N(0, 1)$

- U_1 の分布はなんでもよい。



- 泥中の蓮 (でいちゅうのはちす, 維摩経)

中心極限定理 $Z_n = \frac{1}{\sqrt{n}}(U_1 + U_2 + \cdots + U_n) \rightarrow^d N(0, 1)$

- 中心極限定理（上記のものとその拡張）へ向けた研究の長い歴史
 - de Moivre (1738): 二項分布. 標準正規分布の密度関数 $\phi(z) = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}z^2}$
 - Gauss (1809) “de Moivre curve”
 - “中心”(central)の命名は Polyá (1920)
 - Laplace (1810): 有界格子点分布. フーリエ変換（特性関数）を使った.
 - Poisson (1824): 反例と一般化
 - Liapounov (1900)
 - Markov (1900): “Markov 連鎖”の場合
 - Bernstein (1922): Markov の結果の拡張
 - Lindeberg (1922): “初等的証明”
 - Lévy (1925): マルチンゲール中心極限定理のはしり
 - Khinchin, Kolmogorov, Feller, ...
-

L. Le Cam: The central limit theorem around 1935. *Statistical Science*, 1, 1, 78-96 (1986)

中心極限定理 $Z_n = \frac{1}{\sqrt{n}}(U_1 + U_2 + \cdots + U_n) \rightarrow^d N(0, 1)$

- 歴史に名を残す大数学者も間違いを犯す。
 - 概念も確立しておらず，数学的な道具もない状況であったことは理解すべき。
- 中心極限定理の証明（特性関数による方法）
 - 標準正規分布の特性関数

$$\varphi_{N(0,1)}(u) = \int_{-\infty}^{\infty} e^{iuz} \phi(z) dz = e^{-\frac{1}{2}u^2} \quad (u \in \mathbb{R})$$
 - Z_n の特性関数 $\varphi_{Z_n}(u) = E[e^{iuZ_n}]$ ($i = \sqrt{-1}$)
 - 各 $u \in \mathbb{R}$ に対して， $n \rightarrow \infty$ のとき， $\varphi_{Z_n}(u) \rightarrow \varphi_{N(0,1)}(u)$.
これは簡単な計算によってわかる。
 - 特性関数が収束すれば，対応する分布が収束することが知られている。
 - したがって， $Z_n \rightarrow^d N(0, 1)$.

三角列に対する中心極限定理

- $Z_n = \frac{1}{\sqrt{n}}(U_1 + U_2 + \cdots + U_n) \rightarrow^d N(0, 1)$

- $\frac{U_1}{\sqrt{n}} + \frac{U_2}{\sqrt{n}} + \cdots + \frac{U_n}{\sqrt{n}} \rightarrow^d N(0, 1)$

- 独立確率変数の三角列: 各 $n \in \mathbb{N}$ に対して,

$$H_1^n, H_2^n, \dots, H_{k_n}^n \text{ は独立}$$

- つまり,

$$H_1^1, H_2^1, \dots, H_{k_1}^1 \text{ は独立,}$$

$$H_1^2, H_2^2, \dots, H_{k_2}^2 \text{ は独立,}$$

...

$$H_1^n, H_2^n, \dots, H_{k_n}^n \text{ は独立,}$$

...

三角列に対する中心極限定理

定理 2. $(H_j^n)_{j=1, \dots, k_n}$ ($n \in \mathbb{N}$) は独立確率変数の三角列とする.

(i) $E[H_j^n] = 0$, $E[(H_j^n)^2] < \infty$, 定数 v が存在して,
 $v_n := \sum_{j=1}^{k_n} \text{Var}[H_j^n] \rightarrow v$ ($n \rightarrow \infty$).

(ii) 任意の $\epsilon > 0$ に対して,

$$(L) \quad \lim_{n \rightarrow \infty} \sum_{j=1}^{k_n} E \left[|H_j^n|^2 \mathbf{1}_{\{|H_j^n| \geq \epsilon\}} \right] = 0.$$

このとき,

$$H_1^n + H_2^n + \cdots + H_{k_n}^n \xrightarrow{d} N(0, v) \quad (n \rightarrow \infty).$$

- 定理の証明は 10 行程度

三角列に対する中心極限定理：リンデベルク条件 (L) の意味

- ひとつずつの H_j^n の全体への影響は小さい.

三角列に対する中心極限定理：例

- L_1^n, \dots, L_n^n ($n \in \mathbb{N}$) は独立確率変数の三角列
- L_j^n の分布が正規分布 $N(0, 1/n)$ (平均0, 分散 n^{-1} の正規分布) に従うとする:

$$P[a < L_j^n \leq b] = \int_a^b \frac{1}{\sqrt{2\pi n^{-1}}} e^{-\frac{z^2}{2n^{-1}}} dz \quad (a < b)$$

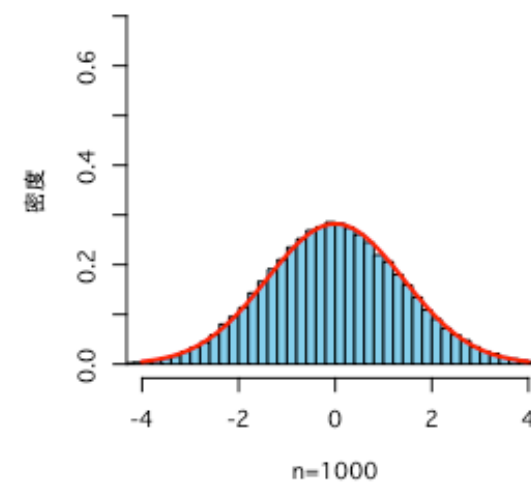
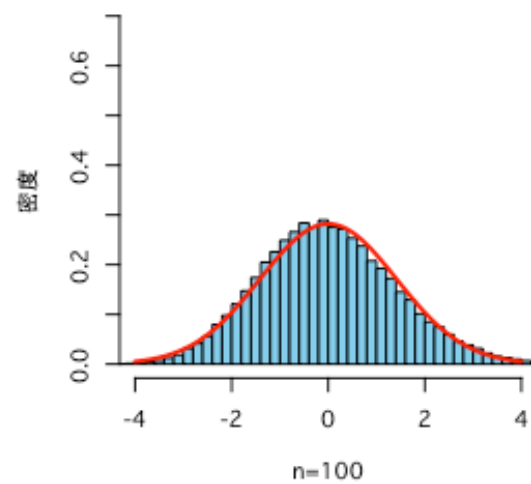
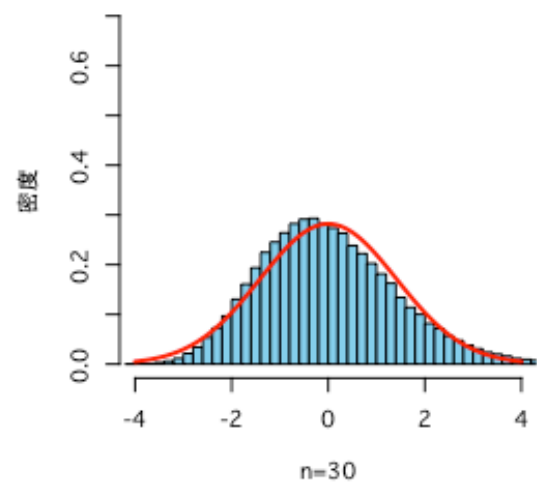
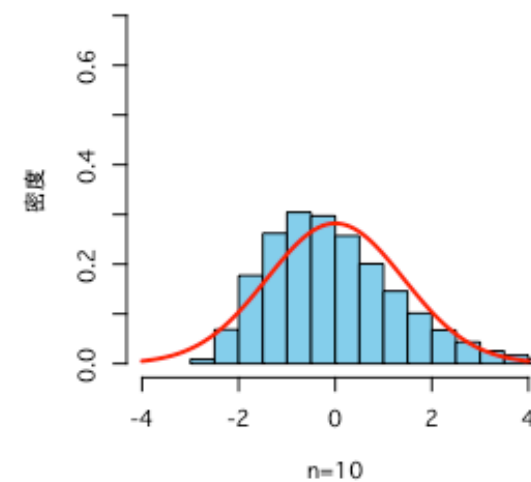
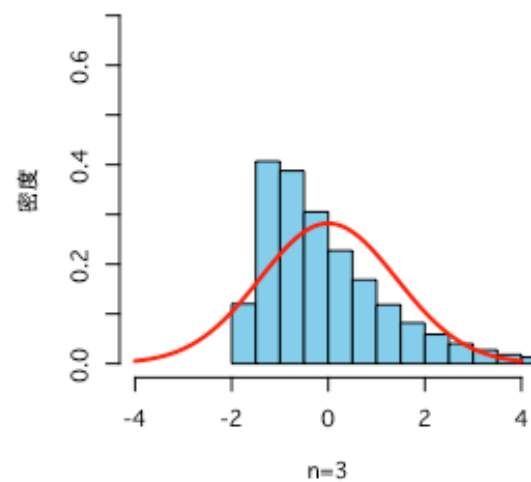
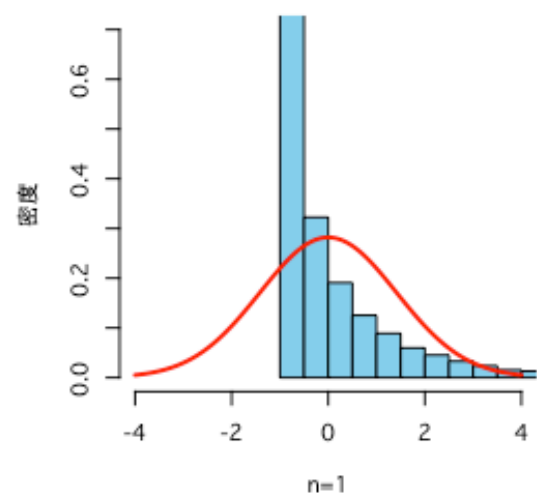
- $H_j^n = \sqrt{n}\{(L_j^n)^2 - n^{-1}\}$ とすると,
 $(H_j^n)_{j=1, \dots, n}$ ($n \in \mathbb{N}$) は独立確率変数の三角列になる.
- $E[H_j^n] = E[\sqrt{n}\{(L_j^n)^2 - n^{-1}\}] = 0$, $\text{Var}[H_j^n] = 2n^{-1}$,
 $\text{Var}[H_1^n] + \dots + \text{Var}[H_n^n] = 2$.
- 個々の H_j^n の全体への影響は小さいので, 中心極限定理によって,

$$H_1^n + H_2^n + \dots + H_n^n \xrightarrow{d} N(0, 2) \quad (n \rightarrow \infty).$$

すなわち,

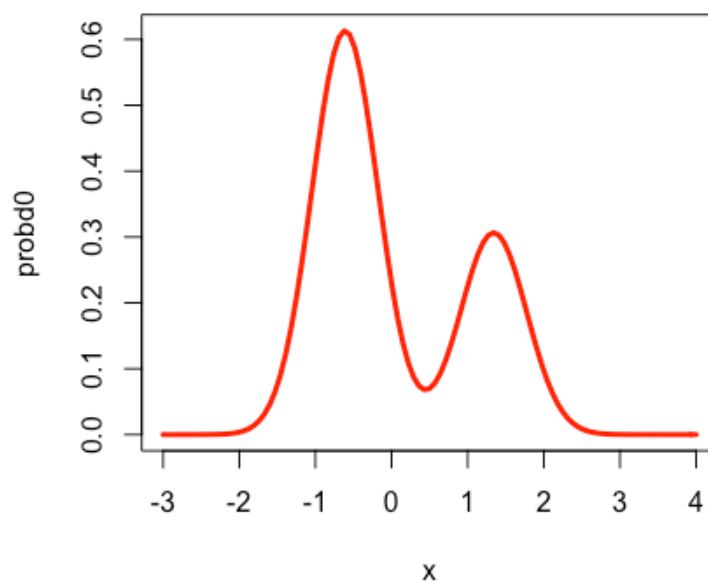
$$\sum_{j=1}^n \sqrt{n}\{(L_j^n)^2 - n^{-1}\} \xrightarrow{d} N(0, 2) \quad (n \rightarrow \infty).$$

三角列に対する中心極限定理：例



中心極限定理からブラウン運動へ

乱数のサンプリング



$\Rightarrow U_1$

$\Rightarrow U_2$

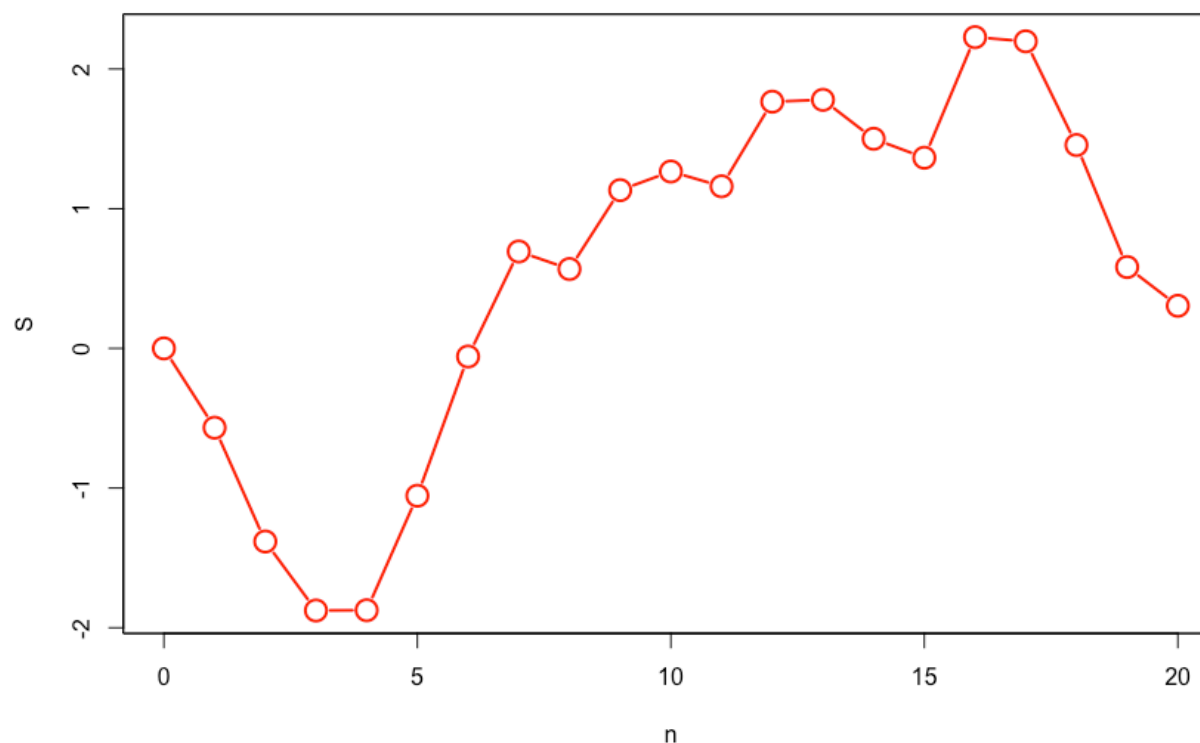
$\Rightarrow U_3$

...

- 任意の、平均0, 分散1の分布からのサンプリング

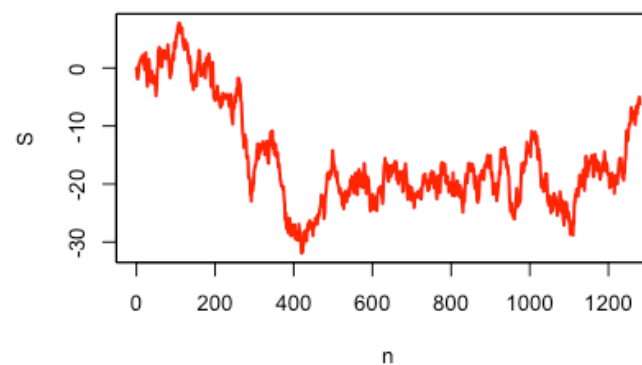
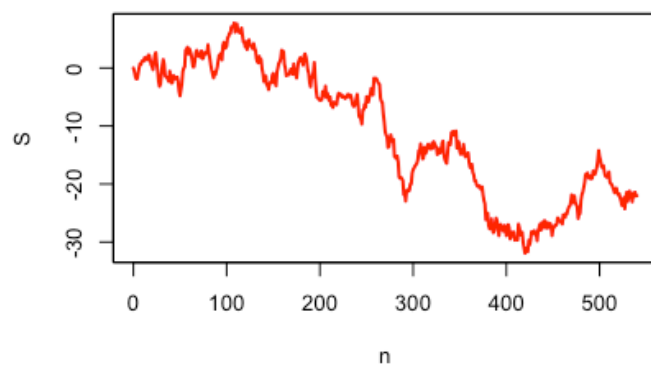
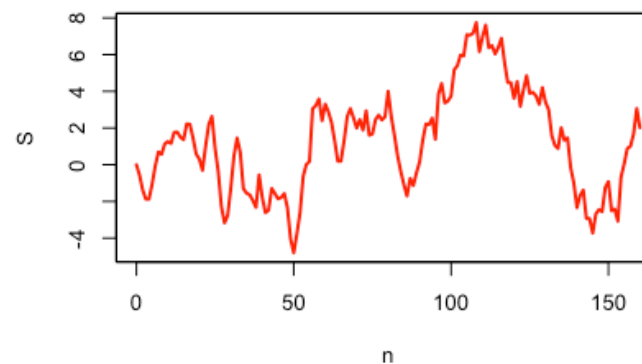
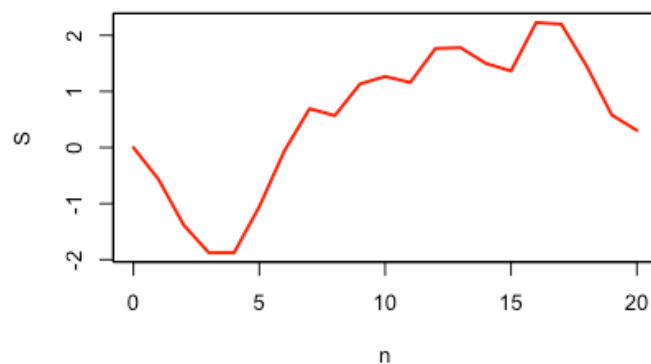
ランダムウォーク

- $(U_j)_{j \in \mathbb{N}}$: 独立で同一の分布に従う確率変数列
- $S_0 = 0, S_n = \sum_{j=1}^n U_j = U_1 + U_2 + \cdots + U_n$ とする



ランダムウォーク

- もっとデータを取り、ズームアウトする



- $n = \infty$ のとき現れるもの

Donskerの定理 (汎関数型中心極限定理)

- $(U_j)_{j \in \mathbb{N}}$: 独立で同一の分布に従う確率変数列。 $E[U_1] = 0$, $E[U_1^2] = 1$ とする。 $S_0 = 0$, $S_n = \sum_{j=1}^n U_j$
- 時空間スケーリングと線形補間

$$X_t^n = \frac{1}{\sqrt{n}} S_{\lfloor nt \rfloor} + \frac{(nt - \lfloor nt \rfloor)}{\sqrt{n}} U_{\lfloor nt \rfloor + 1}$$

$$\left(X_1^n = \frac{1}{\sqrt{n}} S_n \text{ である.} \right)$$

定理 $X^n \rightarrow^d w \quad (n \rightarrow \infty)$

標準 ウィナー過程 (ブラウン運動) w が極限として現れる。

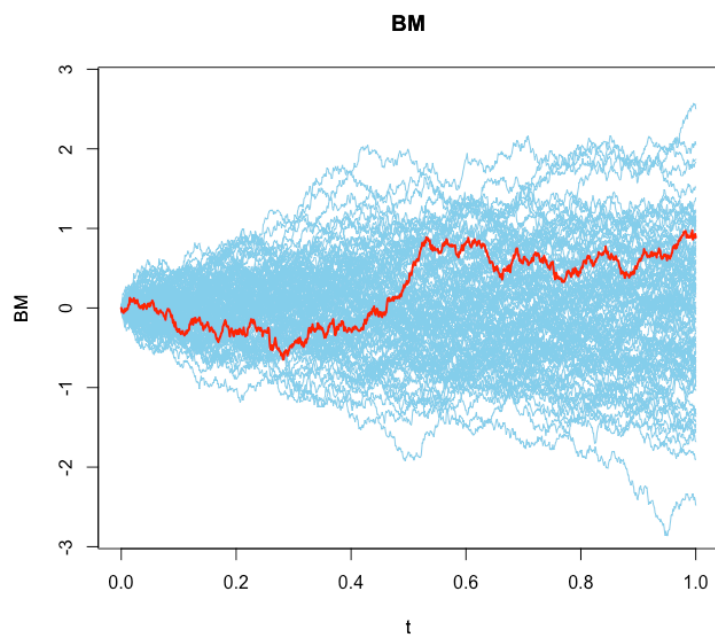
- 収束は空間 C で考える。
 C 連続関数 $f : \mathbb{R}_+ \rightarrow \mathbb{R}$ のある空間, 適当な位相を考える。 $\mathbb{R}_+ = [0, \infty)$.

ウィナー過程のイメージ

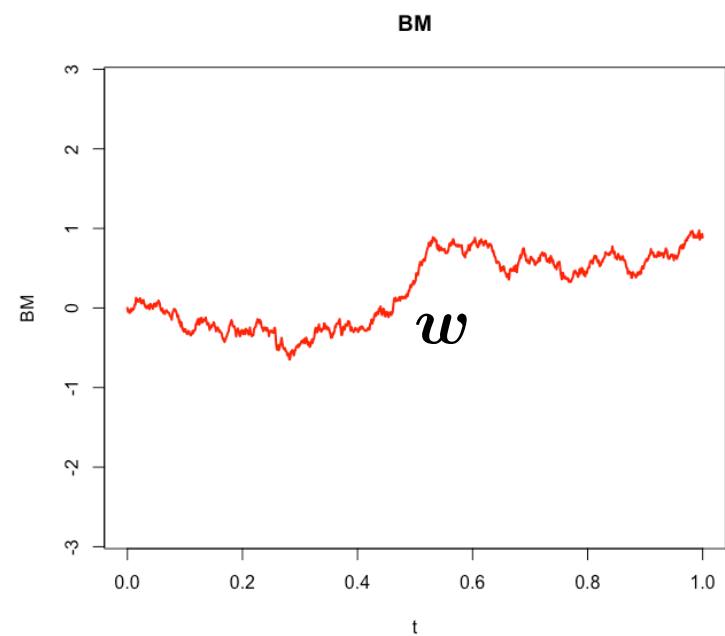
- 一様乱数

区間 $[0, 1]$ $\xrightarrow{\text{抽出}}$ x

- ウィナー過程



$\xrightarrow{\text{抽出}}$



ウィナー過程 (ブラウン運動)

- $w = (w_t)_{t \in \mathbb{R}_+}$: 標準ウィナー過程
- w は次の性質を満たす.

(i) $w_0 = 0$.

(ii) $w_t - w_s \sim N(0, t - s)$ for $s < t$.

(iii) $0 = t_0 < t_1 < \dots < t_k$ ($k \in \mathbb{N}$) に対して, 確率変数

$$w_{t_1}, w_{t_2} - w_{t_1}, \dots, w_{t_k} - w_{t_{k-1}}$$

は独立. (独立増分性)

(iv) 各パス $t \mapsto w_t$ は連続.

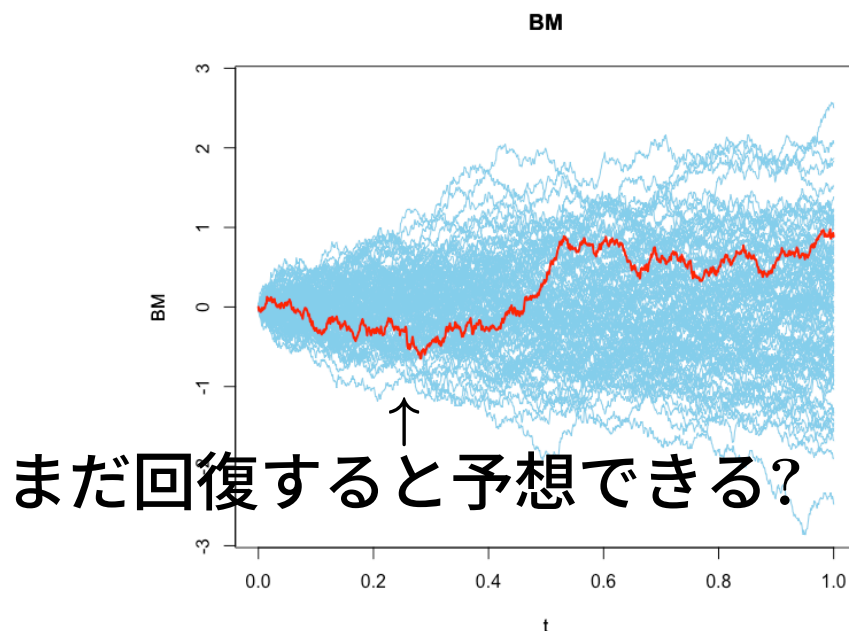
- $X_{t_2}^n - X_{t_1}^n \sim \frac{1}{\sqrt{n}} S_{[nt_2]} - \frac{1}{\sqrt{n}} S_{[nt_1]} = \frac{1}{\sqrt{n}} \sum_{j: [nt_1] < j \leq [nt_2]} U_j$

- w は“ランダムウォーク”. 時刻 t から先の増分はそれ以前の動きとは独立.

ボラティリティの推定（統計の話）

ボラティリティの推定

- $t \in \mathbb{R}_+ = [0, \infty)$ でラベルされた確率変数 X_t を束ねた $X = (X_t)_{t \in \mathbb{R}_+}$ を 確率過程 と呼ぶ。
- 標準 Wiener 過程 $w = (w_t)_{t \in \mathbb{R}_+}$ は確率過程の例である。
- $X_t = \sigma w_t$ であるとする。 σ は正の定数。
- e.g. 株価の収益率のモデル。 σ はその変動の大きさを表す。



ボラティリティの推定

- $w = (w_t)_{t \in \mathbb{R}_+}$: 標準 Wiener 過程.
- $X_t = \sigma w_t$. σ は正の定数.
- 統計的問題
 - データ $(X_{t_j^n})_{j=0, \dots, n}$ ($t_j^n = j/n$) から σ^2 の値を推定する.
 - 2次変動 (リアライズド・ボラティリティ)
 $V_n = \sum_{j=1}^n (X_{t_j^n} - X_{t_{j-1}^n})^2$ で σ^2 を推定する.
- V_n の性質
 - $V_n \rightarrow \sigma^2$ ($n \rightarrow \infty$) ?
 - V_n の推定精度 ?

ボラティリティの推定

- $X_t = \sigma w_t$
 - $t_j^n = \frac{j}{n}$
 - $L_j^n := w_{t_j^n} - w_{t_{j-1}^n} \sim N(0, 1/n)$
 - L_1^n, \dots, L_n^n は独立
-

標準 Wiener 過程 (ブラウン運動) $w = (w_t)_{t \in \mathbb{R}_+}$ の性質

- (i) $w_0 = 0$.
- (ii) $w_t - w_s \sim N(0, t - s)$ for $s < t$.
- (iii) $0 = t_0 < t_1 < \dots < t_k$ ($k \in \mathbb{N}$) に対して、確率変数

$$w_{t_1}, w_{t_2} - w_{t_1}, \dots, w_{t_k} - w_{t_{k-1}}$$

は独立. (独立増分性)

(再掲) 三角列に対する中心極限定理：例

- L_1^n, \dots, L_n^n ($n \in \mathbb{N}$) は独立確率変数の三角列
- L_j^n の分布が正規分布 $N(0, 1/n)$ (平均0, 分散 n^{-1} の正規分布) に従うとする:

$$P[a < L_j^n \leq b] = \int_a^b \frac{1}{\sqrt{2\pi n^{-1}}} e^{-\frac{z^2}{2n^{-1}}} dz \quad (a < b)$$

- $H_j^n = \sqrt{n}\{(L_j^n)^2 - n^{-1}\}$ とすると,
 $(H_j^n)_{j=1, \dots, n}$ ($n \in \mathbb{N}$) は独立確率変数の三角列になる.
- $E[H_j^n] = E[\sqrt{n}\{(L_j^n)^2 - n^{-1}\}] = 0$, $\text{Var}[H_j^n] = 2n^{-1}$,
 $\text{Var}[H_1^n] + \dots + \text{Var}[H_n^n] = 2$.
- 個々の H_j^n の全体への影響は小さいので, 中心極限定理によって,

$$H_1^n + H_2^n + \dots + H_n^n \rightarrow^d N(0, 2) \quad (n \rightarrow \infty).$$

すなわち,

$$\sum_{j=1}^n \sqrt{n}\{(L_j^n)^2 - n^{-1}\} \rightarrow^d N(0, 2) \quad (n \rightarrow \infty).$$

ボラティリティの推定

- $X_t = \sigma w_t,$
- $L_j^n := w_{t_j^n} - w_{t_{j-1}^n} \sim N(0, 1/n), L_1^n, \dots, L_n^n$ は独立
- $X_{t_j^n} - X_{t_{j-1}^n} = \sigma(w_{t_j^n} - w_{t_{j-1}^n}) = \sigma L_j^n$
- だから,

$$\begin{aligned} \sqrt{n}(V_n - \sigma^2) &= \sqrt{n} \left\{ \sum_{j=1}^n (X_{t_j^n} - X_{t_{j-1}^n})^2 - \sigma^2 \right\} \\ &= \sigma^2 \sqrt{n} \sum_{j=1}^n \left\{ (L_j^n)^2 - \frac{1}{n} \right\} \end{aligned}$$

- したがって,

$$\sqrt{n}(V_n - \sigma^2) \rightarrow^d N(0, 2\sigma^4) \quad (n \rightarrow \infty)$$

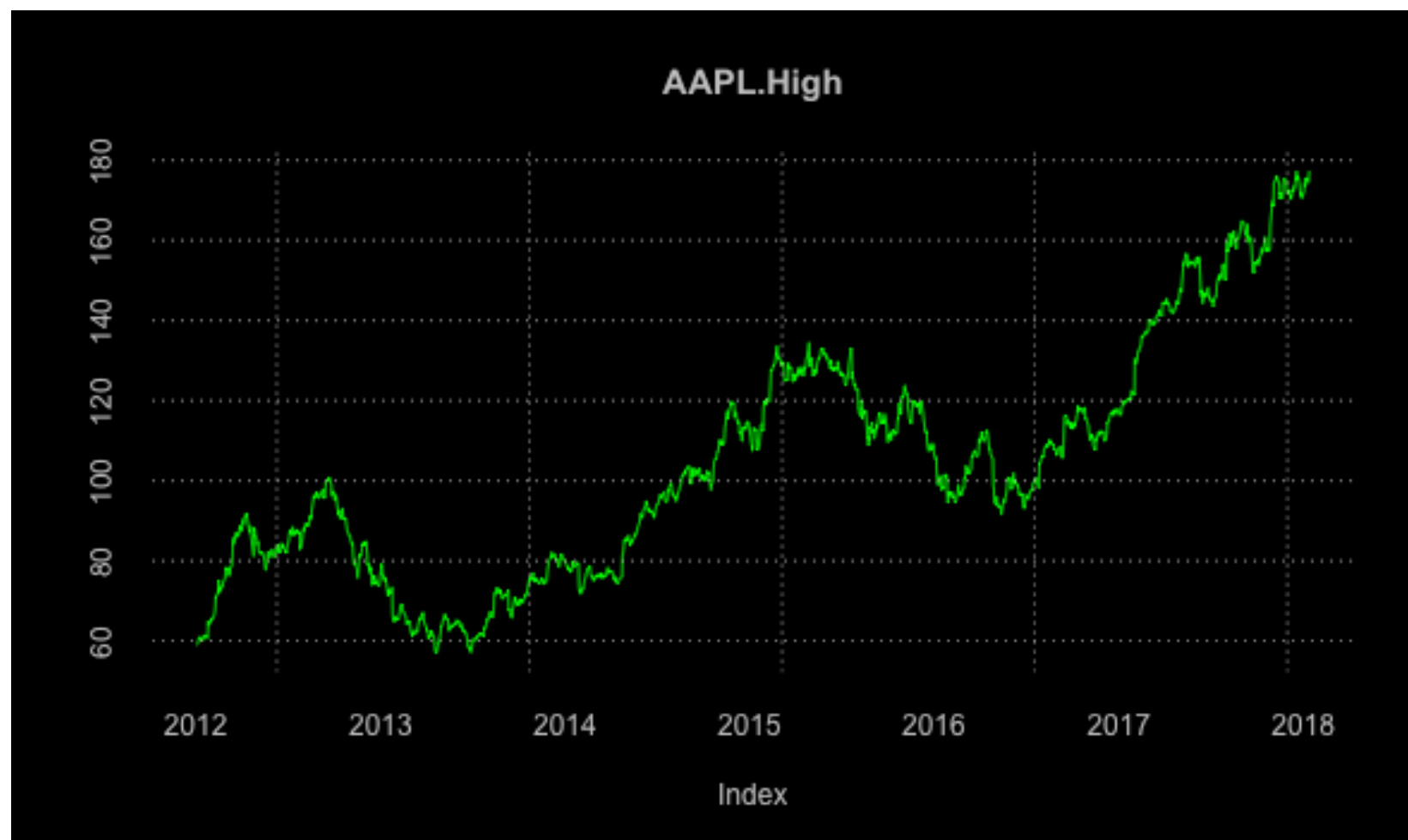
$$\sum_{j=1}^n \sqrt{n} \{ (L_j^n)^2 - n^{-1} \} \rightarrow^d N(0, 2) \quad (n \rightarrow \infty).$$

ボラティリティの推定：まとめ

- $w = (w_t)_{t \in \mathbb{R}_+}$: 標準 Wiener 過程.
- $X_t = \sigma w_t$. σ は正の定数.
- データ $(X_{t_j^n})_{j=0, \dots, n}$ ($t_j^n = j/n$)
- 2次変動 (リアライズド・ボラティリティ)
 $V_n = \sum_{j=1}^n (X_{t_j^n} - X_{t_{j-1}^n})^2$ で σ^2 を推定する.
- V_n の性質
 - $$\sqrt{n}(V_n - \sigma^2) \rightarrow^d N(0, 2\sigma^4) \quad (n \rightarrow \infty)$$
 - データ数 n が大きくなるとき V_n は σ^2 に収束する.
 - 推定量 V_n の誤差は漸近的に正規分布に従う.
- データは区間 $[0, 1]$ 上で採られ, n が大きくなるとき区間 $[0, 1]$ 上で頻繁に観測される. 高頻度観測

確率微分方程式

株価



確率微分方程式によるモデリング

- ブラウン運動 $w = (w_t)$ Louis Bachelier (1900)
- 幾何ブラウン運動 $S = (S_t)$ Black and Scholes (1973),
....

$$S_t = S_0 \exp \left(\left(\mu - \frac{\sigma^2}{2} \right) t + \sigma w_t \right)$$

- “確率微分方程式” で書くと

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

- dS_t は S_t の瞬間的な増分
- dw_t は w_t の瞬間的な増分
- σ : ボラティリティ (volatility)、
 S_t の変動の大きさを表す指標
- 自己回帰モデル

確率微分方程式

- 確率微分方程式 (stochastic differential equation)

$$dX_t = b_t dt + \sigma_t dw_t, \quad X_0 = x_0 \quad (1)$$

積分形 $X_t = x_0 + \int_0^t b_s ds + \int_0^t \sigma_s dw_s$ (伊藤積分)

- 解釈. X_t は漸化式の解の極限と見なせる：

$$\begin{aligned} X_0 &= x_0 \\ X_{\frac{1}{n}} - X_{\frac{0}{n}} &= b_{\frac{0}{n}} \times \frac{1}{n} + \sigma_{\frac{0}{n}} \times (w_{\frac{1}{n}} - w_{\frac{0}{n}}) \\ X_{\frac{2}{n}} - X_{\frac{1}{n}} &= b_{\frac{1}{n}} \times \frac{1}{n} + \sigma_{\frac{1}{n}} \times (w_{\frac{2}{n}} - w_{\frac{1}{n}}) \\ &\dots \\ X_{\frac{j}{n}} - X_{\frac{j-1}{n}} &= b_{\frac{j-1}{n}} \times \frac{1}{n} + \sigma_{\frac{j-1}{n}} \times (w_{\frac{j}{n}} - w_{\frac{j-1}{n}}) \\ &\dots \end{aligned} \quad (2)$$

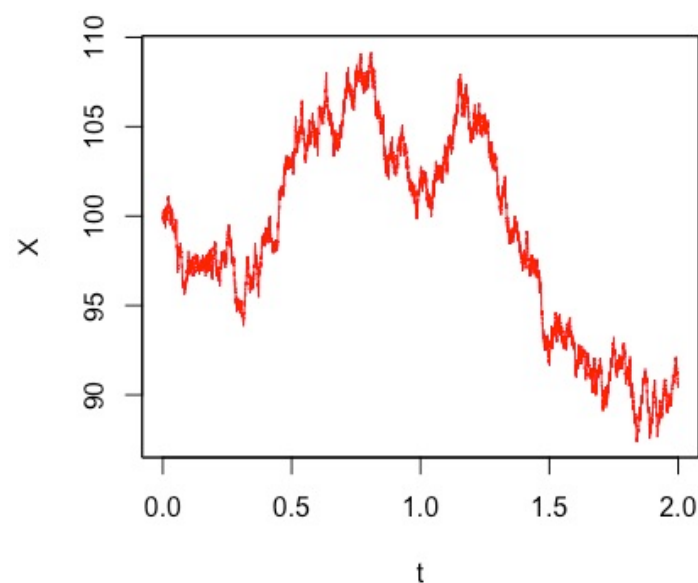
そして $n \rightarrow \infty$ とすると、 X_t が現れる。(Euler-丸山近似)

- 大雑把には、確率微分方程式 (1) は漸化式 (2) で理解してもよい。

Euler-丸山近似で確率微分方程式の解のパスを発生させる

- 確率微分方程式

$$\begin{cases} dX_t = 0.01X_t dt + 0.1X_t dw_t, & t \in [0, 2] \\ X_0 = 100 \end{cases}$$



- ウイナー過程で駆動されている系なので、パス $t \mapsto X_t$ の性質はそれに似ている。
- YUIMA

ボラティリティの推定（再考）

ボラティリティの推定

- 確率微分方程式 (stochastic differential equation)

$$dX_t = b_t dt + \sigma_t dw_t, \quad X_0 = x_0$$

- X_t を 伊藤過程 (Itô process) とよぶ. e.g. $\sigma_t = \sigma(X_t)$
- X_t の変動の大きさをどのように計測するか?
- リアライズド・ボラティリティ (realized volatility)

$$V_n = \sum_{j=1}^n (X_{t_j^n} - X_{t_{j-1}^n})^2, \quad t_j^n = j/n$$

- 事実:

$$V_n \xrightarrow{p} \langle X \rangle_1 = \int_0^1 \sigma_s^2 ds \quad (n \rightarrow \infty)$$

- e.g. $\sigma_t = \sigma$ (定数) ならば, $\int_0^1 \sigma_s^2 ds = \sigma^2$ の推定になる.
- 高頻度データ解析

ボラティリティの推定

- 統計推測で、 V_n の誤差 $V_n - \langle X \rangle_1$ の分布が重要になる。
- $\sigma_t = \sigma$ (定数) の場合には独立な三角列に対する中心極限定理から誤差の極限分布が正規分布であることを知った。

$$dX_t = b_t dt + \sigma_t dw_t, \quad X_0 = x_0,$$
$$V_n = \sum_{j=1}^n (X_{t_j^n} - X_{t_{j-1}^n})^2, \quad t_j^n = j/n$$

ボラティリティの推定

- 統計推測で、 V_n の誤差 $V_n - \langle X \rangle_1$ の分布が重要になる。
- $\sigma_t = \sigma$ (定数) の場合には独立な三角列に対する中心極限定理から誤差の極限分布が正規分布であることを知った。
- σ_t が一般のランダムな場合には新たな極限定理が必要になる。

定理 3. (混合型中心極限定理)

$$Z_n = \sqrt{n}(V_n - \langle X \rangle_1) \xrightarrow{d} G^{1/2}\zeta \quad (n \rightarrow \infty)$$

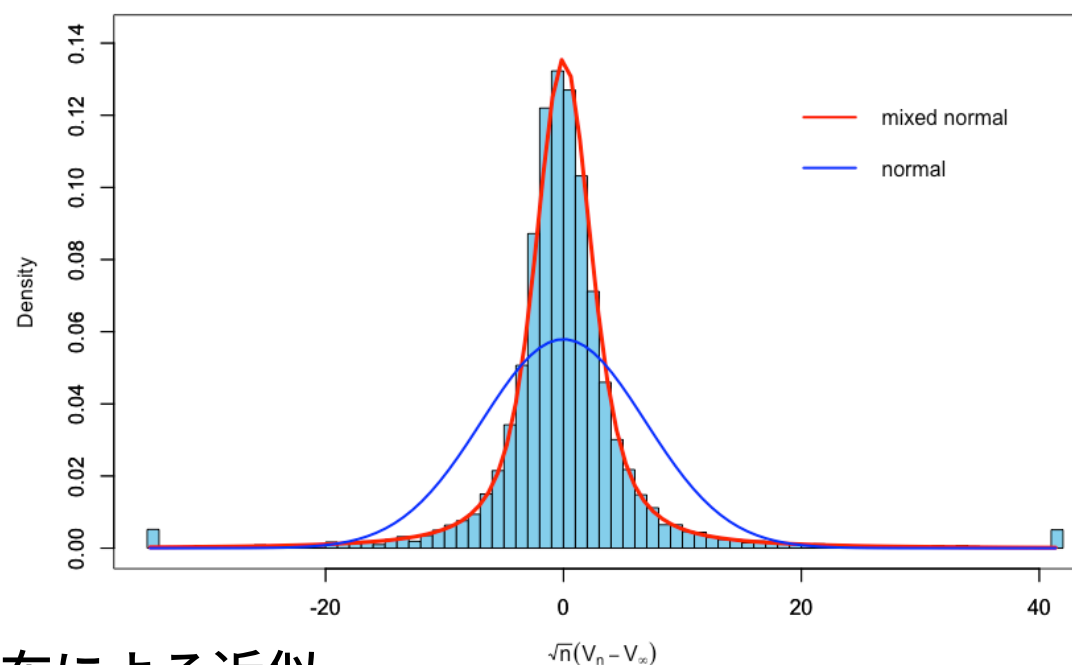
$$G = 2 \int_0^1 \sigma_t^4 ds, \quad \zeta \sim N(0, 1), \quad \zeta \text{ と } G \text{ は独立.}$$

- 極限 $G^{1/2}\zeta$ の分布は正規分布の分散に関する混合
 - 一般に、 σ_t はランダムであるから、 $G^{1/2}\zeta$ の分布は正規分布ではない。
 - 非エルゴード的統計 (non-ergodic statistics)

$$dX_t = b_t dt + \sigma_t dw_t, \quad V_n = \sum_{j=1}^n (X_{t_j^n} - X_{t_{j-1}^n})^2, \quad t_j^n = j/n, \quad \langle X \rangle_1 = \int_0^1 \sigma_s^2 ds$$

ボラティリティの推定： $\sqrt{n}(V_n - \langle X \rangle_1)$ の分布

- モデル $dX_t = -0.1X_t dt + \sqrt{1 + X_t^2} dw_t, \quad X_0 = 1$
- データ $(X_{\frac{j}{500}})_{j=0, \dots, 500}$
- 混合正規分布による近似 vs. 正規近似
MC = 10^4



赤線：混合正規分布による近似

青線：ヒストグラムのデータと同じ分散を持つ正規分布による近似

混合型中心極限定理

- 混合型中心極限定理

$$Z_n \rightarrow^{d_s} G^{1/2} \zeta \quad (n \rightarrow \infty)$$

$$\zeta \sim N(0, 1) \perp\!\!\!\perp \mathcal{F}$$

- \mathcal{F} は情報の集まりを表し, Z_n や G は \mathcal{F} の情報で表現できる.

混合型中心極限定理

- 混合型中心極限定理

$$Z_n \rightarrow^{d_s} G^{1/2}\zeta \quad (n \rightarrow \infty), \quad \zeta \sim N(0, 1) \perp\!\!\!\perp \mathcal{F}$$

- \mathcal{F} は情報の集まりを表し, Z_n や G は \mathcal{F} の情報で表現できる.

- 混合型中心極限定理の証明

- セミマルチンゲール理論 (\supset 伊藤解析) を駆使
- 近年, マリアバン解析 (無限次元確率解析) の方法で直線的に証明できるようになった

漸近展開

漸近展開

- 極限定理 $Z_n \rightarrow^d$ 密度 p_0 の確率分布 \Leftrightarrow

$$E[f(Z_n)] - \int f(z)p_0(z)dz \rightarrow 0 \quad (n \rightarrow \infty)$$

注. $f(z) = 1_{(a,b]}(z)$ のとき,

$$P[a < Z_n \leq b] = E[f(Z_n)] \rightarrow \int_{-\infty}^{\infty} f(z)p_0(z)dz = \int_a^b p_0(z)dz.$$

これはヒストグラムの収束を意味する.

- 漸近展開

$$\begin{aligned} E[f(Z_n)] - \int f(z) \left\{ p_0(z) + n^{-1/2}p_1(z) + \right. \\ \left. \dots + n^{-k/2}p_k(z) \right\} dz \\ = o(n^{-k/2}) \end{aligned}$$

注. 一般に, f は微分可能でなく, 連続でもなくてよい.

オプション価格の計算

- 確率微分方程式

$$\begin{cases} dX_t = (0.1(1 + X_t^2)^{1/2} + 0.05X_t)dt + 0.5(1 + X_t^2)^{1/2}dw_t \\ X_0 = 100 \end{cases}$$

- エイジアン・コールオプションの価格を計算せよ ($r = 0$) :

$$C = E \left[\max \left\{ \int_0^1 X_t dt - K, 0 \right\} \right]$$

- モンテカルロ法 (計算時間、統計的誤差大)
- 漸近展開 (高速、高精度)

$$C \sim c_0 + \epsilon c_1 + \epsilon^2 c_2 + \cdots + \epsilon^k c_k \quad (\epsilon \downarrow 0)$$

$$(0.5 = \epsilon = 1/\sqrt{n})$$

実験 1 オプション価格計算. 漸近展開 vs. モンテカルロ法

K	Order 1	Order 2	Order 3	Order 4	Monte Carlo
80	29.876	29.851	29.873	29.865	29.919 +- 0.030
90	22.496	22.420	22.424	22.419	22.470 +- 0.028
100	16.458	16.376	16.375	16.383	16.413 +- 0.025
110	11.779	11.698	11.698	11.711	11.721 +- 0.022
120	8.300	8.218	8.218	8.220	8.226 +- 0.019
130	5.770	5.707	5.698	5.690	5.701 +- 0.016
140	3.928	3.944	3.910	3.906	3.915 +- 0.014
150	2.583	2.729	2.663	2.672	2.670 +- 0.011
160	1.612	1.890	1.814	1.819	1.814 +- 0.010

Table 5: Evaluation of the expectation in (47) with the dynamics in (51) for several strike prices K . Comparison between the results of the 1st-4th order asymptotic expansion, using a grid of 10000 points to solve the scheme, and Monte Carlo, based on 1 million simulations.

by Emanuele Guidotti

実験2 オプション価格計算. 漸近展開 vs. モンテカルロ法

- 確率微分方程式

$$\begin{cases} dX_t = (0.1(1 + X_t^2)^{1/2} + 0.05X_t)dt + 0.5(1 + X_t^2)^{1/2}dw_t \\ X_0 = 1 \end{cases}$$

- ヨーロピアン・コールオプションの価格 ($r = 0$) :

$$C = E[\max\{X_t - 1.41, 0\}]$$

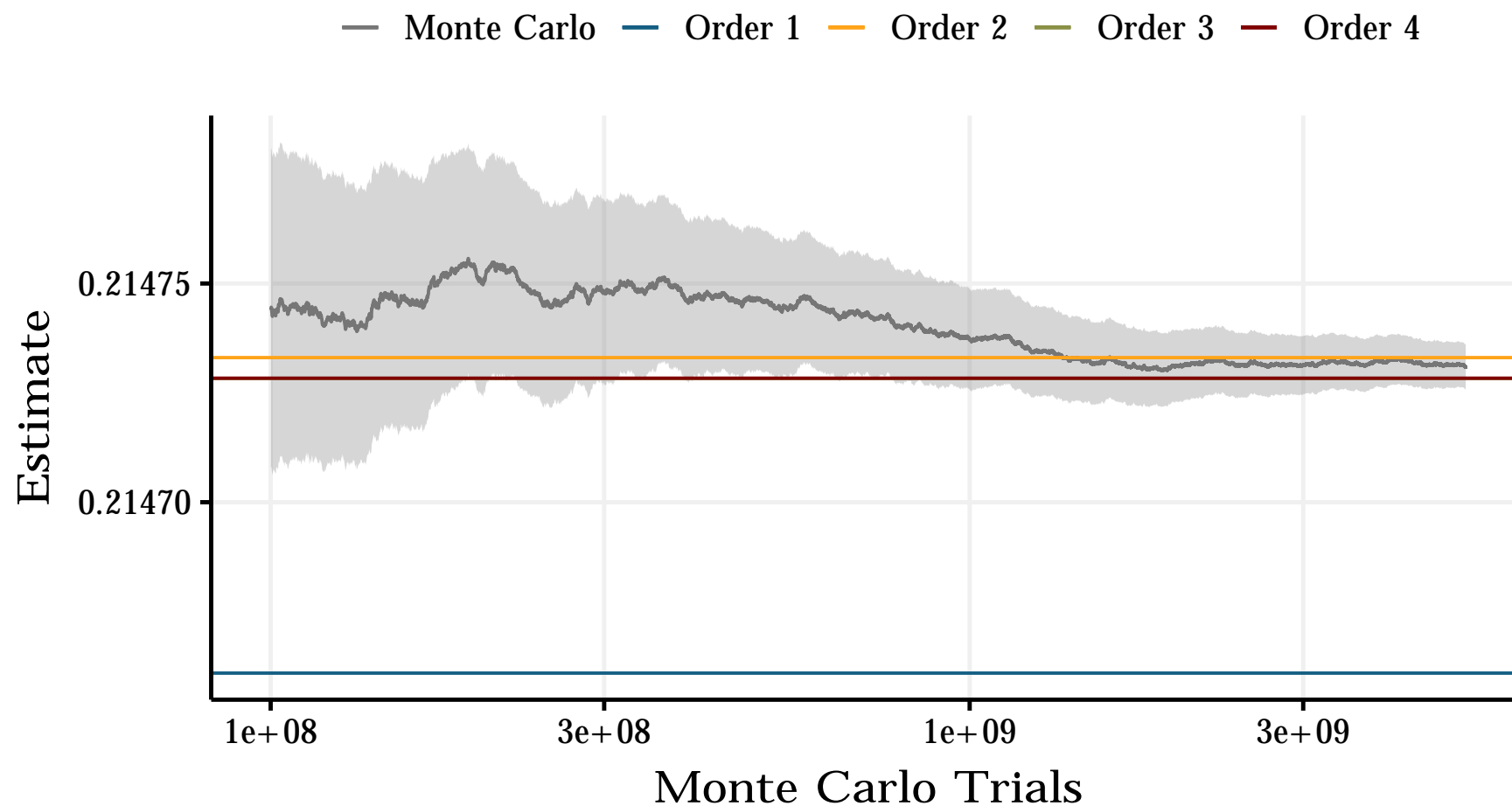
-
- 漸近展開

$$C \sim c_0 + \epsilon c_1 + \epsilon^2 c_2 + \cdots + \epsilon^k c_k \quad (\epsilon \downarrow 0)$$

$$(0.5 = \epsilon = 1/\sqrt{n})$$

実験2 オプション価格計算. 漸近展開 vs. モンテカルロ法

モンテカルロ法の誤差の方が問題！



by Emanuele Guidotti

マリアバン解析と極限定理

- 上の漸近展開は，マリアバン解析 (Malliavin calculus) の渡辺理論によって正当性が保証される．
- マリアバン解析による（分布論的な）極限定理，漸近展開へ発展している．

マリアバン解析，極限定理，統計理論

- マリアバン解析は標本空間 (w が住んでいる空間，無限次元空間) 上の微分積分学．
- マリアバン解析と漸近展開
 - 漸近展開は Z_n の分布のなめらかさに深く関わる．
 - Z_n の分布のなめらかさを定量的に評価できる．
 - 極限定理研究において，漸近展開と結びついた．
(通常分布論的な極限定理においても.)
- マリアバン解析による極限定理研究が盛んになっている．
- マリアバン解析と確率過程の統計学は密接に関係している．
 - 極限定理，漸近展開，確率数値解析
 - 漸近決定理論，非エルゴード的統計，推定量の漸近展開
 - 非整数ブラウン運動

マリアバン解析 (Malliavin calculus)

- $w = (w_t)_{t \in [0, T]}$: 標準 Wiener 過程 (Brown 運動)
- w の関数としての確率変数: $F(w)$
- 例. 確率微分方程式の解 $F(w) = X_t(w)$

$$dX_t = b(X_t)dt + \sigma(X_t)dw_t, \quad X_0 = x_0$$

- マリアバン微分 $DF = \frac{dF}{dw} = \left(\frac{\partial F}{\partial w_t} \right)_{t \in [0, T]}$ (直感的に)
- 双対性 (duality). 確率過程 $u = (u_t)_{t \in [0, T]}$ に対して、

$$E[\langle u, DF \rangle] = E[\delta(u)F] \quad (\forall \text{微分可能な } F)$$
- $\delta(u) = D^*u$ を u の Skorohod 積分 と呼ぶ。
- Skorohod 積分は伊藤積分の一般化になっている。

伊藤解析と極限定理

- ボラティリティ推定の問題

$$dX_t = b_t dt + \sigma_t dw_t, \quad V_n = \sum_{j=1}^n (X_{t_j^n} - X_{t_{j-1}^n})^2, \quad t_j^n = j/n,$$

$$\langle X \rangle_1 = \int_0^1 \sigma_s^2 ds$$

- リアライズド・ボラティリティは伊藤積分を使って表現できる.

$$\begin{aligned} V_n - \langle X \rangle_1 &= \sum_{j=1}^n (X_{t_j^n} - X_{t_{j-1}^n})^2 - \langle X \rangle_1 \\ &= \sum_{j=1}^n \int_{t_{j-1}^n}^{t_j^n} \left(2\sigma_{t_{j-1}^n}^2 \int_{t_{j-1}^n}^s dw_r \right) dw_s + \dots \end{aligned}$$

- 伊藤積分に対する混合型中心極限定理

$$\Rightarrow \sqrt{n}(V_n - \langle X \rangle_1) \rightarrow^{d_s} G^{1/2} \zeta$$

Skorohod 積分に対する混合型中心極限定理

- Skorohod 積分は伊藤積分の一般化になっている：
- $u = (u_t)_{t \in [0, T]}$ において $u_t = u_t((w_r)_{r \in [0, t]})$ であれば、

$$\delta(u) = \int_0^T u_t dw_t \quad (\text{伊藤積分})$$

- $M_n = \delta(u_n)$ ($n \in \mathbb{N}$) Skorohod 積分の列

定理 4. (Nourdin, Nualart and Peccati)

$$E[|\langle DM_n, u_n \rangle - G|] \rightarrow 0 \quad (n \rightarrow \infty)$$

(と付加条件) のもとで、

$$M_n \rightarrow^{d_s} G^{1/2} \zeta \quad (n \rightarrow \infty)$$

ここで、 ζ は G と独立な標準正規確率変数。

- 最近漸近展開へ発展している。(Nualart and Y)

本日のまとめ

- 独立確率変数列の和に対する中心極限定理
- 独立な三角列に対する中心極限定理
- Wiener過程（ブラウン運動）
- 確率微分方程式
- ボラティリティの推定問題
- 混合正規型の極限定理
- 漸近展開
- マリアバン解析と中心極限定理の一般化