

移民項のある分枝過程の統計的逐次解析

Sequential Analysis of
Branching Processes with Immigration

横浜国立大学

国際社会科学研究科

王 歆

目次

はじめに

第一章 予備知識と論文に使用された主要な定理と性質のまとめ

1.1 微分積分学の知識

1.1.1 数列と関数	6
1.1.2 リーマン積分	8

1.2 測度論の基礎

1.2.1 測度空間	8
1.2.2 確率変数について	9
1.2.3 積分の知識	12
1.2.4 平均	14
1.2.5 独立性とマルコフ連鎖	15
1.2.6 大数の法則	16
1.2.7 直積測度とフビニの定理	17

1.3 マルチンゲールの理論

1.3.1 条件付き期待値	17
1.3.2 マルチンゲール	19
1.3.3 一様可積分性	20
1.3.4 停止時	22
1.3.5 離散時間マルチンゲール	23

1.4 確率過程

1.4.1 ブラウン運動	27
1.4.2 確率積分と伊藤の公式	27
1.4.3 Bessel 過程	28

第二章 単一変量分枝過程について

2.1 分枝過程の母関数について	29
2.2 パワーシリーズ分布	31
2.3 移民項のある分枝過程の逐次確率比検定の結論	
2.3.1 モデルと特性値の推定	33
2.3.2 最尤推定のシミュレーションと結果	34
2.3.3 逐次確率比検定	38
2.3.4 逐次確率比検定のシミュレーションと結果分析	40

第三章 臨界値に対する分枝過程の逐次検定

3.1 モデルの基本	43
3.2 Diffusion Approximation	
3.2.1 非逐次検定の近似結果	44
3.2.2 Diffusion Approximation の定理	44
3.3 尤度関数と Fisher 情報量	
3.3.1 尤度関数について	51
3.3.2 Fisher 情報量	53

3.4 停止時刻 (Stopping Time) と逐次検定 (Sequential Test)	
3.4.1 Stopping Time の定義	52
3.4.2 重要な定理	52
3.4.3 定理の証明	53
3.4.4 ベッセル過程 (Bessel Process)	55
3.5 シミュレーションと結果	
3.5.1 シミュレーション理論的な分析	56
3.5.2 シミュレーション理論的な分析結果	59
第四章 SPRT 検定	
4.14.1 Wald の SPRT 検定	
4.1.1 Wald の SPRT 検定の意味	66
4.1.2 閾値の決定	61
4.2 分枝過程の SPRT 検定	68
4.3 分枝過程に対する SPRT 検定のシミュレーション	71
第五章 移民項のある 2 変量分枝過程	
5.1 モデルについての解説	
5.1.1 モデルと直感の意味	72
5.1.2 各記号の意味	73
5.1.3 criticality parameter について	74
5.2 移民項のある 2 変量分枝過程の推定	
5.2.1 CLS について	75
5.2.2 CLS より移民項のある 2 変量分枝過程の推定	75
5.2.3 CLS より移民項のある 2 変量分枝過程の推定 のシミュレーション	77
第六章 参考文献	78

はじめに

本論文のテーマは「移民項のある分枝過程の統計的逐次解析」である。

移民項のある分枝過程モデルに対して、単一変量と二変量の二つケースを察した。統計的な手法で、移民項のある分枝過程モデルに対し逐次解析理論を厳密に証明したうえ、シミュレーションを通じ、主張する観点を直観的検証した。

第1章では、本論文に使われる理論根拠として数学的な定義と定理をまとめておる。微分積分学の理論、測度空間理論、マルチンゲールの理論と確率過程主な基礎となる知識である。

第2章では、単一変量モデルについて直観的な解説するうえ、移民項のある分枝過程の逐次確率比検定を行った。この部分は筆者の修士時代の研究題目で

ある。分枝過程モデル $z_0 = 1, z_n = \sum_{i=1}^{z_{n-1}} \xi_{n,i} + Y_n$ において、 Z_n と Y_n それぞれの期待値

を μ と λ を推定して、シミュレーションよりヒストグラムで表示した。そして帰無仮説 $\mu = 1$ と対立仮説 $\mu > 1$ を設定し、逐次確率比検定をおこなった。統計

計量は関数尤度比 $L = \frac{P_1(z_1, y_1 | z_{n-1})}{P_0(z_1, y_1 | z_{n-1})}$ で決められる。停止時刻

$T = \min \left\{ t \geq 1 : \log \frac{f_1(z_1) f_1(z_2) L}{f_2(z_1) f_2(z_2) L} \notin [a, b] \right\}$ と定義される。この基で一定条件

に満たす時刻 t を取り出し、サンプリングをストップする。このような停止時刻をシミュレーションより観測された。そしてこの検定の二種類誤りをシミュレーションより算出した。理論値と大体一致の結論を取得したうえ、各特性値の特徴を明確にした。

第3章で確率母関数など性質を利用し、移民項のいある分枝過程モデルに対して逐次解析をおこなった。そして臨界値 $m = 1$ において逐次検定シミュレーションのを行った。この部分での理論は I. V. Basawa (1991) Sequential Estimation For Branching Processes With Immigration, と Keiji Nagai (2007) Sequential Estimation Test For The Criticality Of Branching Processes を参考した。Stopping Time を定義するためフィッシャー情報量の意味とパワーシリーズを詳しく紹介した。Keiji Nagai の論文で定義された定理を数学的な厳密に証明をしたうえ、Bessel 過程よりシミュレーション

の考え方を提案した。 $T_c = \inf \left\{ n \geq 0; \sum_{i=1}^n Z_{i-1} / \hat{\sigma}_n^2 \geq c^2 \right\}, c \rightarrow \infty$ で定義されるとき、

この方法ではサンプリングの規模を想定したら Bessel 過程の知識よりサンプリングのとまる条件 c を数量的に考察できる。シミュレーションで観測された StoppingTime が想定したサンプリングの規模大体一致であるので目的を実現した。さらに統計値 $c(\hat{m}_T - 1)$ の分布をヒストグラムで観測された。理論的ではこの統計値が標準ブラウン運動に従うが、シミュレーションで大体正規分布のヒストグラムが出ていたため、モデルの逐次解析理論を検証した。

第4章では第3章の解析理論のうえ SPRT 検定を行った。単一変量移民項のあ

る分枝過程において、Abraham Wald 1947 年が開発した特定の逐次仮説検定の手段で SPRT 検定を行い、シミュレーションで検定力について考察する。二種類誤りの算出方法を明確したうえ、シミュレーション値を計算した。そして、本論文では SPRT 検定の上、Stopping Time の停止条件に可変量の付く場合のも考えて、シミュレーションを行った。簡単に説明すると停止時刻は

$T_{a,b}^c = \inf \{n \geq 1 : \log \Lambda_n^c \notin (a,b)\}$ で定義される。この場合、希望なサンプリング規模を得るため、モデル特性値である m のみの力なら不十分であるとき、さらに v を用いて、精密な停止条件を見つけることができる。物理系の研究などによく使われる分枝過程モデルに対しては、 v の存在が検定の精度を大幅に改善できるのではないかと考えておる。例えば地震の予報や機械のコントロールなど。この部分に関する理論は未成熟であるが n 、今後の研究方向として努力する欲がある。

第 5 章では二変量モデル $\begin{bmatrix} X_{k,1} \\ X_{k,2} \end{bmatrix} = \sum_{j=1}^{X_{k-1,1}} \begin{bmatrix} \xi_{k,j,1,1} \\ \xi_{k,j,1,2} \end{bmatrix} + \sum_{j=1}^{X_{k-1,2}} \begin{bmatrix} \xi_{k,j,2,1} \\ \xi_{k,j,2,2} \end{bmatrix} + \begin{bmatrix} \varepsilon_{k,1} \\ \varepsilon_{k,2} \end{bmatrix}$ に対して、具

体的な例で解説した上逐次検定をした。

二変量モデルの停止時刻問題をシミュレーションより探索してみた。変量モデルの停止時刻の定義、条件設定などのことについて提案した。

本論文でパラメータの推定、停止時刻問題の探索、逐次検定など問題を考察することより、移民項のある単一変量と 2 変量分枝過程に関する特性を明確にした。逐次解析は統計的推測に必要なデータを必要最小限得ることで、早い時点で条件を満たすサンプルを獲得し、早期警報システムを提供できる点になる。逐次解析の方法は一変量モデルから二変量モデルへ拡張して、新たな考えである。

自然界では生物の出生と死亡が自然なものだ。子どもを産みだし続けるので物種が絶滅せず、各世代が一つの集団になる。科学では電子信号の転送するように、一本信号源から出ていく電子が電子板を通る時何本の新しい電子を生み出す。我々は子のような数の増え方に子どもから直観的印象がある。本論文で、数理統計の方法を通じ、この増え方を定量に研究を行う後、分枝過程に対する認識が深くなった。特性値の計算より、分枝過程に従うある集団の規模、この先どうのように変化するのが明らかに知った。さらにシミュレーションが分枝過程を量的に再現し、理論推定の結果を検証することが出来る。コンピュータ技術の先進性を実感できる。

集団規模の縮小して行ってほしいか、拡大して行ってほしいか、各分野の研究に対して希望も異なる。逐次解析を用いて、希望条件を満たすサンプリングを探ることができる。この探査が実現できたら、我々の社会生活では深刻の影響を与える。また本論文の例を言うと、逐次解析より各分野の研究及び社会生活に対する積極的な意義がある。コンピュータは二つの領域をつなぎ合わせ、いろいろな研究成果を共通に利用するようになった。

第一章 予備知識と論文に使用された主要な定理と性質のまとめ

第一章では、本論文に関連である数学、統計学の定義・定理・命題・補題をまとめておる。これらの基礎知識は勉強の重要内容の一つとして、研究の基盤となる。次の本を参考した：Billingsley, P(1999) [参考文献 15], David Williams(1991) [参考文献 16], 笠原皓司 (1974) [参考文献 17]。

1.1 微分積分学の知識

微分積分学は 17 世紀頃から次第にその力強い姿を現し始め、ニュートン・ライプニッツによる体系化を経て以後数世紀にわたり自然科学の最も基本的な道具として用いられ、同時にまたそれ自身も大きく発展してきた人類共通の財産である。微分積分学の導関数関係、漸近展開、数列の収束、リーマン積分などの問題は研究によく用いられたので、整理しておく。

1.1.1 数列と関数

定義 1.1 数列 $\{a_n\} (n=1,2,\dots)$ が $n \rightarrow \infty$ の時 $a \in R$ に収束するとは、 a のどんな ε 近傍 $U_\varepsilon(a)$ をとっても、有限個の番号 n を除いてすべての $a_n \in U_\varepsilon(a)$ となることである。この時、 $\lim_{n \rightarrow \infty} a_n = a$ また $a_n \rightarrow a (n \rightarrow \infty)$ と書く。

定義 1.2 数列 $\{a_n\} (n=1,2,\dots)$ が $n \rightarrow \infty$ がコーシー列であるとは、任意 ε に対して有限個の番号 n を除いてすべての a_n の相互距離が ε 以下になることである。すなわち任意

$\varepsilon > 0$ に対し、ある番号 n_0 から先のすべての p, q に対し、 $|a_p - a_q| < \varepsilon$ となるように n_0 が決められることである。

定理 1.1 収束列がコーシー列である。また、逆にコーシー列は収束列である。

定理 1.2 $\lim_{x \rightarrow +\infty} f(x)$ が存在するための必要十分条件は、任意の $\varepsilon > 0$ に対し $G > 0$ を適当にえらんで、どんな $x_1 > G, x_2 > G$ についても $|f(x_1) - f(x_2)| < \varepsilon$ となるようにできることである（この条件をコーシーの条件と呼ぶ）。

定理 1.3 (中間値定理) $f(x)$ を有界閉空間 $I = [a, b]$ において連続な関数とすると、 $f(a)$ と $f(b)$ の中間の任意の値 l に対し $l = f(\xi)$ となる値 $\xi \in I$ が存在する。

定義 1.3 集合 A 上で定義された連続関数が A で一様連続とは、任意の $\varepsilon > 0$ に対し、 $\delta > 0$ を (ε に応じて) 適当に選ぶと、相互距離が δ 以下の任意の 2 点 $x \in A, y \in A, |x - y| < \delta$ についてつねに $|f(x) - f(y)| < \varepsilon$ が成立することをいう。敷

で書くと、 $\lim_{|x-y| \rightarrow 0} |f(x) - f(y)| = 0$ である。普通の連続性が x_0 を固定するごとに、 ε に対応する $\delta > 0$ をきめて、 $f(U_\delta(x_0) \cap U_\varepsilon(y_0))$ となるのに対し、一様連続性はこの δ に x_0 に無関係にA全体を通じて同じ値に取れることを示している。

定理 1.4 有界集合A上で定義された一様連続関数はAまで連続関数としては拡張できる。そしてその拡張は一意的である。逆に有界閉集合K上で連続な関数はKで一様連続である。

定理 1.5 (ライプニッツ)

$$f(x), g(x) \text{が } x_0 \text{で } n \text{回微分可能のとき, } (fg)^{(n)}(x_0) = \sum_{k=0}^n \binom{n}{k} f^{(k)}(x_0) g^{(n-k)}(x_0).$$

定義 1.4 テイラーの公式

$f(x)$ が x_0 で連続なら $f(x) - f(x_0)$ は $x \rightarrow x_0$ の時無限小である。 $f(x) - f(x_0)$ を典型的無限小の1次結合で表すことができれば、 $f(x)$ が x_0 に近づく時の様子がよくわかることになる。典型的な無限小としてはいろいろな関数が考えられるが最も基本的なものとして、 $(x - x_0)^\alpha (\alpha > 0)$ があげられるであろう。そこで

$$f(x) = f(x_0) + a_1(x - x_0)^{a_1} + \dots + a_n(x - x_0)^{a_n} + o((x - x_0)^{a_n})$$

$(0 < a_1 < a_2 < \dots < a_n)$ とできるか、という問題を考えよう。もし、 $f(x)$ が x_0 で微分可能なら、 $f(x) = f(x_0) + f'(x_0)(x - x_0) + o(x - x_0) \quad (x \rightarrow x_0)$ だから、そのような典型的な無限小の一つ $(x - x_0)$ が確かに取れる。もっと微分可能性の回数が増やせば、このような項は増えるだろう。すなわち次の定理が出る：

(テイラー公式) $[x_0, x]$ で $f^{(n-1)}$ が連続であって、 (x_0, x) で $f^{(n)}$ が存在するならば、

$$f(x) = f(x_0) + \frac{f'(x_0)}{1!}(x - x_0) + \frac{f''(x_0)}{2!}(x - x_0)^2 + \dots + \frac{f^{(n-1)}(x_0)}{(n-1)!}(x - x_0)^{(n-1)} + R_n(x),$$

ここでは $R_n(x) = \frac{f^{(n)}(\xi)}{n!}(x - x_0)^{(n)}$ となるような $\xi \in (x_0, x)$ が存在する。

定理 1.6 ディニの定理 有界閉集合K上で $\{f_n(x)\}_{n=1,2,\dots}$ がnと共に単調に増加(または減少)して $f(x)$ に各点収束し、 $\{f_n(x)\}_{n=1,2,\dots}$ も $f(x)$ も連続関数ならば、この収束はK上で一様収束でなければならない。

定理 1.7 有界な空間 $I = [a, b]$ 上で $\{f_n(x)\}$ が連続関数列であって、 $f(x)$ に一様収束すれば、 $\lim_{n \rightarrow \infty} \int_a^b f_n(t) dt = \int_a^b f(t) dt$.

1.1.2 Riemann 積分

定理 1.7 微分積分学の基本定理

区間上の連続関数 $f(t)$ に対し、定積分が原始関数の一つを与える：

$\frac{d}{dt} \int_a^t f(s) ds = f(t)$ 。言葉を換えていると、 $F(t) = \int_a^t f(s) ds$ とおくと、 $F(t)$ は微分可能であり、 $F'(t) = f(t)$ となる。

定理 1.8 $f(x)$ 、 $g(x)$ がある空間で微分可能で、 $f'(x)$ 、 $g'(x)$ が連続なら、その区間内の任意の a, b について

$$\int_a^b f'(t)g(t)dt = [f(t)g(t)]_a^b - \int_a^b f(t)g'(t)dt.$$

定理 1.9 $f(x)$ は空間 I 上で連続とし、 $\varphi(s)$ は空間 J 上で微分可能で、 $\varphi'(s)$ は連続とする。また、 $\varphi(s) \in I$ 、 $s \in J$ とする。このとき、 $x = \varphi(s)$ とすると、任意の

$$\alpha, \beta \in J \text{ に対し、 } \int_{\varphi(\alpha)}^{\varphi(\beta)} f(x) dx = \int_{\alpha}^{\beta} f(\varphi(s))\varphi'(s) ds.$$

1.2 測度空間理論の基礎

1.2.1 測度空間

定義 2.1 (S, Σ) を可測空間、すなわち、 Σ は S 上の σ -加法族であるとする。写像 $\mu : \Sigma \rightarrow [0, \infty]$ は、 μ が可算加法的であるとき (S, Σ) 上の測度と呼ばれる。三つ組 (S, Σ, μ) は測度空間と呼ぶ。

$\mu(s) < \infty$ であるとき、測度 (μ あるいは測度空間 (S, Σ, μ)) は有限であるという。

Σ の要素の列 $(S_n : n \in \mathbb{N})$ で $\mu(S_n) < \infty (\forall n \in \mathbb{N})$, かつ $\bigcup_n S_n = S$ となるものが存

在するとき、測度 μ を σ -有限とよぶ。

測度 μ は $\mu(S) = 1$ であるとき確率測度と呼ばれ、 (S, Σ, μ) は確率測度と呼ぶ。

定理 2.1 カラテオドリの拡張定理

S を集合、 Σ_0 を S 上の一つの加法族とし、 $\Sigma = \sigma(\Sigma_0)$ としよう。もし μ_0 が可算加法的な写像 $\mu_0 : \Sigma_0 \rightarrow [0, \infty]$ であるとする、 (S, Σ) 上の測度 μ であって、 Σ_0 の上で $\mu = \mu_0$ となるものが存在する。

補題 2.1 基本的な不等式

(S, Σ, μ) を測度空間としよう。このとき、

$$(1) \mu(A \cup B) \leq \mu(A) + \mu(B) \quad (A, B \in \Sigma).$$

$$(2) \mu\left(\bigcup_{i \leq n} F_i\right) \leq \sum_{i \leq n} \mu(F_i) \quad (F_1, F_2, \dots, F_n \in \Sigma).$$

(3)さらに $\mu(s) < \infty$ を仮定すると

$$\mu(A \cup B) = \mu(A) + \mu(B) - \mu(A \cap B) \quad (A, B \in \Sigma).$$

$$(4) F_1, F_2, \dots, F_n \in \Sigma \text{ に対して } \mu\left(\bigcup_{i \leq n} F_i\right) = \sum_{i \leq n} \mu(F_i) - \sum_{i < j \leq n} \mu(F_i \cap F_j) \\ + \sum_{i < j < k \leq n} \mu(F_i \cap F_j \cap F_k) - \dots + (-1)^{n-1} \mu(F_1 \cap F_2 \cap \dots \cap F_n)$$

補題 2.2 測度の単調収束性

(1) $F_n \in \Sigma (n \in \mathbb{N})$ であり、かつ $F_n \uparrow F$ とすると、 $\mu(F_n) \uparrow \mu(F)$ が成立。

(3)可算個の μ -零集合の和はまた μ -零集合である。

1.2.2 確率変数

定義 2.2 (S, Σ) を可測空間とする。 Σ -可測関数、 $h: S \rightarrow \mathbb{R}$ としよう。 $A \subseteq \mathbb{R}$ に対して $h^{-1}(A) = \{s \in S : h(s) \in A\}$ と定義する。 h が Σ -可測であるとは：

$h^{-1}: \mathbb{B} \rightarrow \Sigma$. すなわち $\forall A \in \mathbb{B}$ に対し $h^{-1}(A) \in \Sigma$ であること。ゆえに、 Σ -可測関数 h について次の式が成立：

$$S \xrightarrow{h} \mathbb{R}, \\ \Sigma \xleftarrow{h^{-1}} \mathbb{B}$$

補題 2.3 分布関数の性質

ある確率変数 X の分布関数 $F = F_x$ としよう。この時

(1) $F: \mathbb{R} \rightarrow [0, 1], F \uparrow$ (すなわち $x \leq y \Rightarrow F(x) \leq F(y)$).

(2) $\lim_{x \rightarrow \infty} F(x) = 1, \lim_{x \rightarrow -\infty} F(x) = 0$

(3) F は右連続である。

定理 2.2 単調族定理

H を S から \mathbb{R} への有界関数のつくる一つの類(クラス)で、いかの性質を満たすものとする。

(1) H は \mathbb{R} 上のベクトル空間である。

(2) 定数関数 1 は H の要素である。

(3) (f_n) が H の非負関数列で、 $f_n \uparrow f$ 。ここに f は S 上の有界関数とすると、

Hである。

定義 2.3～定義 2.7 は確率変数の収束についてである。

数学の確率論の分野において、確率変数の収束 (convergence of random variables) に関しては、いくつかの異なる概念がある。確率変数列のある極限への収束は、確率論や、その応用としての統計学や確率過程の研究における重要な概念の一つである。より一般的な数学において同様の概念は確率収束 (stochastic convergence) として知られ、その概念は、本質的にランダムあるいは予測不可能な事象の列は、その列から十分離れているアイテムを研究する場合において、しばしば、本質的に不変な挙動へと落ち着くことが予想されることがある、という考えを定式化するものである。異なる収束の概念とは、そのような挙動の特徴づけに関連するものである：すぐに分かる二つの挙動とは、その列が最終的に定数となるか、あるいはその列に含まれる値は変動を続けるがある不変な確率分布によってその変動が表現される、というようなものである。

定義 2.3 弱収束

F_n と F が分布関数であり、 F が連続関数である。任意 x にたいしては

$\lim_n F_n(x) = F(x)$ であるならば、 F_n が F に弱収束すると言ひ、 $F_n \Rightarrow F$ と書く。

μ_n と μ は確率測度であり、それぞれ分布関数 F_n と F を持つとき、 $F_n \Rightarrow F$ になる条件は $\lim_n \mu_n(A) = \mu(A)$ が成立すること。ここでは $A = (-\infty, x]$, $\mu(x) = 0$.

$F_n(u_n + v_n) \Rightarrow F(x)$ と $F_n(a_n + b_n) \Rightarrow G(x)$ と仮定する、 $u_n > 0, a_n > 0$. そうすると次の式に満たす a, b が存在する。

$$\frac{a_n}{u_n} \rightarrow a, \frac{(b_n - v_n)}{u_n} \rightarrow b, F(ax + b) = G(x)$$

この式に従って次の五つ結論がある：

(i) $F_n \Rightarrow F, a_n \rightarrow a, b_n \rightarrow b$ ならば $F_n(a_n x_n + b_n) \Rightarrow F(ax + b)$.

(ii) $F_n \Rightarrow F, a_n \rightarrow \infty$ ならば $\lim_n F_n(a_n x) = \begin{cases} 1 & x > 0 \\ 0 & x < 0 \end{cases}$

(iii) $F_n \Rightarrow F, b_n$ 無収束ならば $F_n(x + b_n)$ が弱収束にならない.

(iv) $F_n \Rightarrow F, F_n(a_n x_n + b_n) \Rightarrow G(x)$ ならば $0 < \inf_n a_n \leq \sup_n a_n < \infty, \sup_n |b_n| < \infty$.

(v) すべての x においては $F(x) = F(ax + b)$ が成立ならば $a = 1, b = 0$.

定義 2.4 分布収束

このタイプの収束により、ある与えられた確率分布によってより良くモデル化されるようなランダム実験の列における結果を期待することが出来る。分布収束は、この記事内で述べられる全ての他のタイプの収束も意味するという点において、最も弱い収束である。しかしながら、実際の現場において、分布収束

は非常によく利用される；最もよく現れるのは、中心極限定理の応用においてである。 $F_n \Rightarrow F$ ならば X_n は分布収束また法則収束と言ひ、 $X_n \Rightarrow X$ と書く。 X_n と X は確率変数で、それぞれ分布関数 F_n と F をもつ。

定義 2.5 確率収束

確率空間 (Ω, \mathbf{F}, P) での確率変数 X, X_1, X_2, \dots に対して、 X_n は確率 1 で X に収束すると

$$P[|X_n - X| \geq \varepsilon] = 0 \quad (\varepsilon > 0).$$

つまり

$$\lim_n P[|X_n - X| \geq \varepsilon] = 0.$$

この時 X_n が X 確率収束と呼ぶ。 $X_n \rightarrow_p X$ と書く。確率収束が次の性質を持つ

- (1) 確率収束は、分布収束を意味する。
- (2) $X_n \Rightarrow X$ そして $X_n - Y_n = 0$ ならば $Y_n \Rightarrow X$
- (3) 任意 u ($u \rightarrow \infty$) に対し、 $X_n^{(u)} = X^{(u)}$ ($n \rightarrow \infty$), $X_n^{(u)} \Rightarrow X$ そして

$$\lim_u \limsup_n P[|X_n^{(u)} - Y_n| \geq \varepsilon] = 0 \quad (\varepsilon > 0) \text{ ならば } Y_n \Rightarrow X.$$

定義 2.6 概収束

確率変数列 X_n が X へと概収束あるいはほとんど確実に収束、ほとんど至る所で収束、確率 1 で収束あるいは強収束するとは次の式を成立すること。

$$P\left(\lim_{n \rightarrow \infty} X_n = X\right) = 1$$

が上式は、 X_n が X へと収束しない事象が起きる確率が 0 であるという意味で、 X_n の値が X の値へと近付くことを意味する。確率空間 (Ω, \mathbf{F}, P) を定め、 Ω から \mathbb{R} への関数としての確率変数の概念を利用することで、上式は次の式と同値になる：

$$P\left(\omega \in \Omega : \lim_{n \rightarrow \infty} X_n(\omega) = X(\omega)\right) = 1.$$

また概収束の同値な定義には、以下もある：

$$P\left(\liminf \{\omega \in \Omega : |X_n(\omega) - X(\omega)| < \varepsilon\}\right) = 1 \quad (\varepsilon > 0).$$

概収束は、しばしば、収束を表す矢印の上に記号 *a. s.* (almost surely の略) を付け加えることによって表現される：

$$X_n \xrightarrow{a.s.} X.$$

距離空間 (S, d) 上の一般的な確率要素 $\{X_n\}$ に対しても、同様に概収束が定義される：

$$P(\omega \in \Omega : d(X_n(\omega), X(\omega)) \rightarrow 0) = 1.$$

概収束は確率収束を意味し、したがって分布収束を意味する。大数の強法則で用いられる概念は、概収束である。

定義 2.7 確実収束

ある確率空間上定義される列あるいは確率変数 (X_n) (すなわち、確率過程) が X

へ確実収束 (sure convergence) あるいは各点収束するとは、

$$\lim_{n \rightarrow \infty} X_n(\omega) = X(\omega), \forall \omega \in \Omega。$$

が成立することを言う。ここで Ω は、確率変数が定義される確率空間に含まれる標本空間である。これは、関数列の各点収束の概念を確率変数の列へと拡張したものである (確率変数はそれ自身が関数であることに注意されたい)。

$$\left\{ \omega \in \Omega \mid \lim_{n \rightarrow \infty} X_n(\omega) = X(\omega) \right\} = \Omega$$

確率変数の確実収束は、上述の他の全ての収束を意味する。しかし、概収束の代わりに確実収束を用いることのメリットは確率論においてはあまり無い。それら二つの収束の違いは、確率ゼロの集合に関する点のみに存在する。このことが、確実収束の概念が滅多に用いられることの無い理由である。

1.2.3 積分の知識

定義 2.8 一般の可測関数 $f: \Omega \rightarrow [-\infty, +\infty]$ については、その正部分 f^+ と負部分 f^- をそれぞれ $f^+ = f \vee 0, f^- = f \vee 0$ と定義する。これらが非負値可測関数とことはよい。そして、 $\int f^+ d\mu < \infty$ または $\int f^- d\mu < \infty$ であるとき f の積分を

$$\int f d\mu = \int f^+ d\mu - \int f^- d\mu < \infty \text{ と定義する。}$$

非負可測関数対しのこの積分の定義は任意の単関数 ϕ についてもこの積分の定義は $I(\phi)$ と一致することは容易に確認できる。更に $\int f^+ d\mu < \infty$ かつ

$\int f^- d\mu < \infty$ であるとき f は積分可能 (可積, ルベーク積分, ルベーク可積) であるという。また、任意の可測集合 A について f の A 上の積分を $\int_A f d\mu = \int 1_A f d\mu$ と

定義する。可測集合 A に対し、 $1_A f$ が積分可能であるとき f は A 上で積分可能であるという。

$f: \Omega \rightarrow [-\infty, +\infty]$ は積分可能ならば $\int_{\sum_{n=1}^{\infty} A_n} f d\mu = \sum_{n=1}^{\infty} \int_{A_n} f d\mu$ が成立する。

補題 2.4

f, g を任意の可測関数, c を任意の実数とする。

(1) $f \geq 0$ ならば、 $\int f d\mu \geq 0$ である。

(2) $f \geq 0$ かつ $\int f d\mu = 0$ ならば、 $\mu(\{\omega \in \Omega: f(\omega) > 0\}) = 0$ である。

(3) f が積分可能ならば、 cf も積分可能であり、 $\int cf d\mu = c \int f d\mu$ 。

(4) f, g が積分可能ならば、 $f + g$ も積分可能であり、 $\int (f + g) d\mu = \int f d\mu + \int g d\mu$

が成立する。

定理 2.3 ルベーグの収束定理

関数列 f_n が f に概収束し、積分可能関数 g が存在し、ほとんど至るところで $|f_n(\omega)| \leq g(\omega)$ が成立するならば、 f は積分可能で $\lim_{n \rightarrow \infty} \int f_n d\mu = \int f d\mu$ が成立する。

定理 2.4 単調収束定理 (Monotone-Convergence Theorem, MON)

(1) (f_n) を $(m\Sigma)^+$ の要素の列で、 $f_n \uparrow f$ を満たすとしよう。このとき、

$$\mu(f_n) \uparrow \mu(f) \leq \infty,$$

または他の記号では $\int_s f_n(s) \mu(s) \uparrow \int_s f(s) \mu(ds)$ が成立する。

この定理は積分論のまったくすべてがここあるといえる定理である。

(2) 与えられた $f \in (m\Sigma)^+$ に対して、単純関数列 $f^{(r)}$ で $f^{(r)} \uparrow f$ となうものを番目の階段関数 $\alpha^{(r)}: [0, \infty] \rightarrow [0, \infty]$ を次のように定義する。

$$\alpha^{(r)}(x) = \begin{cases} 0 & x = 0 \\ (i-1)2^{-r} & ((i-1)2^{-r} < x \leq i2^{-r} \leq r (i \in N)) \\ r & x > r \end{cases}$$

このように定めると、 $f^{(r)} = \alpha^{(r)} g f$ は $f^{(r)} \in SF^+$ を満たし、 $f^{(r)} \uparrow f$ で、(MON) によって

$$\mu(f) = \uparrow \lim \mu(f^{(r)}) = \uparrow \lim \mu_0(f^{(r)})$$

ここで、 $\alpha^{(r)}$ は左連続であるので、 $f_n \uparrow f$ ならば $\alpha^{(r)}(f_n) \uparrow \alpha^{(r)}(f)$ である。

(3) $f, g \in (m\Sigma)^+$ で $f = g$ (a.e.) ならば $\mu(f) = \mu(g)$ である。

(4) $f \in (m\Sigma)^+$ で、 (f_n) は $(m\Sigma)^+$ の列で μ -零集合 N 上を除いて $f_n \uparrow f$ としよう。この時 $\mu(f_n) \uparrow \mu(f)$ が成立する。

定理 2.5 優収束定理 (Dominated-Convergence Theorem, DOM)

$f_n, f \in m\Sigma$ は S の各点 s で $f_n(s) \rightarrow f(s)$ とし、関数列 (f_n) は $L^1(S, \Sigma, \mu)^+$ の一つ要素 g によって押さえられている。すなわち

$$|f_n(s)| \leq g(s) \quad (\forall s \in S, \forall n \in N)$$

とする。ただしここで $\mu(g) < \infty$ である。このとき、 $f_n \rightarrow f$ が $L^1(S, \Sigma, \mu)$ で成立する。すなわち $\mu(|f_n - f|) \rightarrow 0$ である。なので $\mu(f_n) \rightarrow \mu(f)$ となる。

補題 2.4 FATOU の補題

h を可積分な関数、 $\{f_n\}$ を可測関数列とする。

$$f_n \geq h \quad (\mu - a.e.) \text{ のとき } \int_X \liminf_{n \rightarrow \infty} f_n d\mu \leq \liminf_{n \rightarrow \infty} \int_X f_n d\mu$$

$$f_n \leq h \quad (\mu - a.e.) \text{ と } \int_X \liminf_{n \rightarrow \infty} f_n d\mu \geq \liminf_{n \rightarrow \infty} \int_X f_n d\mu$$

1.2.4 平均

定義 2.9 マルコフの不等式

マルコフの不等式は、測度論的には、 (X, Σ, μ) を測度空間とし、 f を拡張実数値（無限大もとりうる）可測関数とし、 $t > 0$ とすれば、

$$\mu(\{x \in X \mid |f(x)| \geq t\}) \leq \frac{1}{t} \int_X |f| d\mu \text{ であることを述べる。}$$

空間の測度が 1 である特別な場合（つまり確率空間である）には、次のように言い換えられる： X を任意の確率変数とし、 $a > 0$ とすると、

$$P(|X| \geq a) \leq \frac{E(|X|)}{a}$$

定義 2.10 L^p ($1 \leq p < \infty$) 空間の定義

(1) $p \geq 1$ を固定する。 X 上 $a.e.$ で定義された可測関数 $f: X \rightarrow \mathbb{C}$ で、 $\int |f|^p d\mu < \infty$ を満たすものの集合を L^p ($L^p(X)$) と書く。

(2) $f \in L^p$ に対し、 $\|f\|_p = \left\{ \int |f|^p d\mu \right\}^{1/p}$ を f の L^p -ノルムという。

(3) $f \in L^p$, $n \in \mathbb{N}$, が $\lim_{n \rightarrow \infty} \|f_n - f\|_p = 0$ を満たすとき、 f_n は f に p 次平均収束するという。また、 $\lim_{n \rightarrow \infty} f_n(x) = f(x)$, x - $a.e.$, を満たすとき、 f_n は f に概収束するという。

定義 2.11 Schwarz の不等式

$$f, g \in L^2 \text{ ならば } \int |fg| d\mu \leq \sqrt{\int |f|^2 d\mu} \sqrt{\int |g|^2 d\mu}$$

定義 2.12 Hölder の不等式

$$p > 1, \frac{1}{p} + \frac{1}{q} = 1 \text{ とする。 } f \in L^p, g \in L^q \text{ ならば}$$

$$\int_X |fg| \mu(dx) \leq \left(\int |f|^p d\mu \right)^{1/p} \left(\int |g|^q d\mu \right)^{1/q}$$

(1) スカラー倍と和について閉じている空間（集合）を線型空間（ベクトル空間）という。

(2) 線型空間 X に実数値関数 $\|\cdot\| : X \rightarrow \mathbb{R}$ が定義されていて、
非負： $\|f\| \geq 0$,
一意： $\|f\| = 0$ ならば $f = 0$,

一次： $\|af\| = \|a\| \|f\|$,
 三角不等式： $\|f+g\| \leq \|f\| + \|g\|$,
 が成り立つとき、この空間 $(X, \|\cdot\|)$ を線型ノルム空間、 $\|\cdot\|$ をノルムという。

- (3) 非負： $\rho(f, g) \geq 0$,
 一意： $\rho(f, g) = 0$ ならば $f = g$,
 対称： $\rho(f, g) = \rho(g, f)$,
 三角不等式： $\rho(f, h) \leq \rho(f, g) + \rho(g, h)$,

が成り立つ2変数関数を距離という。 $\|\cdot\|$ がノルムのとき、 $\rho(f, g) = \|f-g\|$ は距離である。これをノルムが定義する距離という。

(4) 線型ノルム空間 $(X, \|\cdot\|)$ が、ノルムが定義する距離に関して完備なとき、 $X = (X, \|\cdot\|)$ を *Banach* 空間という。

定理 $(L^p, \|\cdot\|_p)$ は *Banach* 空間である。

1.2.5 独立性とマルコフ連鎖

定義 2.12 独立性に関する諸定義

(1) 独立な σ -加法族

F の部分 σ -加法族 G_1, G_2, \dots が独立であるとは、 $G_i \in \mathcal{G}_j (i \in \mathbb{N})$ で i_1, \dots, i_n が異なるとき、つねに

$$P(G_{i_1} \cap \dots \cap G_{i_n}) = \prod_{k=1}^n P(G_{i_k})$$

が成り立つことである。

(2) 独立な確率変数

確率変数 X_1, X_2, \dots が独立であるとは、 σ -加法族 $\sigma(X_1), \sigma(X_2), \dots$ が独立であることである。確率論の重要な諸問題の多くは、確率変数の列で独立同分布 (independent and identically distributed, IID) であるものに関わっている。

(3) 独立な事象

事象 E_1, E_2, \dots が独立であるとは、 σ -加法族 $\varepsilon_1, \varepsilon_2, \dots$ が独立であることである。ただしここで、 ε_n は σ -加法族 $\{\emptyset, E_n, \Omega \setminus E_n, \Omega\}$ である。

定義 2.13 マルコフ連鎖

マルコフ連鎖とは、確率過程の一種であるマルコフ過程のうち、とりうる状態が散的 (有限または可算) なもの (離散状態マルコフ過程) をいう。また特に、時間が離散的なもの (時刻は添え字で表される) を指すことが多い (他に連続時間マルコフ過程というものもあり、これは時刻が連続である)。マルコフ連鎖は、未来の挙動が現在の値だけで決定され、過去の挙動と無関係である (マルコフ性)。各時刻において起こる状態変化 (遷移または推移) に関して、マルコフ連鎖は遷移確率が過去の状態によらず、現在の状態のみによる系列である。特に重要な確率過程として、様々な分野に応用される。

マルコフという名前はロシアの数学者アンドレイ・マルコフによる (綴りは

Markov)。物理学や統計学のモデルに使われる他、情報学においては外すことのできない重要な考え方である。

マルコフ連鎖は、一連の確率変数 X_1, X_2, X_3, \dots で、現在の状態が決まっていれば、過去および未来の状態は独立であるものである。形式的には、

$$P(X_{n+1} = x | X_n = x_n, \dots, X_1 = x_1, X_0 = x_0) = P(X_{n+1} = x | X_n = x_n)$$

X_i のとりうる値は、連鎖の状態空間と呼ばれ、可算集合 S をなす。マルコフ連鎖は有向グラフで表現され、エッジにはある状態から他の状態へ遷移する確率を表示する。マルコフ連鎖の一例に有限状態機械がある。これは、時刻 n において状態 y にあるとすると、それが時刻 $n+1$ において状態 x に動く確率は、現在の状態にだけ依存し、時刻 n には依存しない。

時間的に均一な（斉時的）マルコフ連鎖とは、すべての n に対し

$$P(X_{n+1} = x | X_n = y) = P(X_n = x | X_{n-1} = y)$$

であるような過程をいう。一般の、時間的に均一でないマルコフ連鎖は、この等式を満たさない。初期状態 i から時刻 n で状態 j に移る確率は、

$$P_{ij}^{(n)} = P(X_n = j | X_0 = i)$$

で定義され、単一段階の遷移は

$$P_{ij} = P(X_1 = j | X_0 = i)$$

で定義される。 n -段階遷移は、任意の $0 < k < n$ に対して次のチャップマン・コルモゴロフの等式を満たす：

$$P_{ij}^{(n)} = \sum_{r \in S} P_{ir}^{(k)} P_{rj}^{(n-k)}$$

時刻 n での状態に関する確率（周辺確率）は次のように書ける：

$$P(X_n = j) = \sum_{r \in S} P_{rj} P(X_{n-1} = r) = \sum_{r \in S} P_{rj}^{(n)} P(X_0 = r)$$

ここで右上付き添え字 (n) は整数値である。もしマルコフ連鎖が時間に対して定常的ならば、この添え字は“ n 乗”という意味にとってもよい。

1.2.6 大数の法則

大数の法則 (law of large numbers) は、確率論・統計学における極限定理のひとつで、「経験的確率と理論的確率が一致する」という、素朴な意味での確率を意味付け、定義付ける法則である。

厳密には、ヤコブ・ベルヌーイによる大数の弱法則 (WLLN: Weak Law of Large Numbers) と、エミール・ボレルやアンドレイ・コルモゴロフによる大数の強法則 (SLLN: Strong Law of Large Numbers) とがある。単に「大数の法則」と言った場合、どちらを指しているのかは文脈により判断する必要がある

定理 2.6

期待値 μ であるような可積分独立同時分布確率変数列 X_1, X_2, \dots の算術平均

$$[X_n] = \frac{X_1 + X_2 + \dots + X_n}{n}$$

のとり値は、十分大きな n まで考えれば、ほとんどの n でおおよそ μ である ($[X_n]$ が μ から大きく外れるような n の現れる確率は n を無限に大きくすると 0 に近づく) : $\lim_{n \rightarrow \infty} [X_n] = \mu$ in probability.

これを大数の弱法則という。また同じ条件下で、 $n \rightarrow \infty$ とするとき、 $[X_n]$ は μ にほとんど確実に (almost surely, 確率 1 で) 収束する : $\lim_{n \rightarrow \infty} [X_n] = \mu$ a.s.

これを大数の強法則という

定理 2.7 チェビシエフの不等式

チェビシエフの不等式は、不等式で表される、確率論の基本的な定理である。パフヌティ・チェビシエフにより初めて証明された。

標本あるいは確率分布は、平均のまわりに、ある標準偏差をもって分布する。この分布と標準偏差の間に、どのような標本・確率分布でも成り立つ関係を示したのが、チェビシエフの不等式である。例えば、平均から 2 標準偏差以上離れた値は全体の 1/4 を超えることはなく、一般に n 標準偏差以上離れた値は全体の $1/n^2$ を超えることはない X を、期待値が μ 、有限の分散が σ^2 である確率変数とすると、任意の実数 $k > 0$ に対して $P(|X - \mu| \geq k\sigma) \leq \frac{1}{k^2}$, ただし $k > 1$ の場合にだけ意味がある。

1.2.7 直積測度とフビニの定理

(X_1, F_1, μ_1) および (X_2, F_2, μ_2) は各各 σ -有限と仮定する。

(i) $f: X_1 \times X_2 \rightarrow [0, \infty]$ は $F_1 \otimes F_2 / B(\bar{R})$ で可測とする。このとき

$$\left\{ \begin{array}{l} x \rightarrow \int_{x_2} f(x, y) \mu_2(dy) \text{ は } F_1 / B(\bar{R}) \text{ 可測} \\ y \rightarrow \int_{x_1} f(x, y) \mu_1(dx) \text{ は } F_2 / B(\bar{R}) \text{ 可測} \end{array} \right. *$$

かつ

$$\int_{X_1 \times X_2} f d(\mu_1 \otimes \mu_2) = \int_{x_1} \left(\int_{x_2} f(x, y) \mu_2(dy) \right) \mu_1(dx) = \int_{x_2} \left(\int_{x_1} f(x, y) \mu_1(dx) \right) \mu_2(dy) **$$

(ii) $f: X_1 \times X_2 \rightarrow \bar{R}$ は $\mu_1 \otimes \mu_2$ 可積分とする。この時*と**が成立。

1.3 マルチンゲールの理論

1.3.1 条件付き期待値

定義 3.1 確率空間 (Ω, \mathcal{F}, P) で X は確率変数で $E(|X|) < \infty$ とし、 \mathcal{G} は \mathcal{F} の部分 σ -加法族とする。このとき一つの確率変数 Y で次の諸性質を満たすものが存在する。

すなわち、

(1) \mathcal{G} 可測。

(2) $E(|Y|) < \infty$

(3) \mathcal{G} のどの集合 A に対しても $\int_A Y dP = \int_A X dP$

が成り立つ。

さらにもし Y が他の確率変数でこれらの諸性質を満たすならば、 $Y = X$ (a.s.)。すなわち、 $P[Y = X] = 1$ である。(1) ~ (3) の諸性質を満たす一つの確率変数 Y を、 \mathcal{G} を与えたときの X の条件付き平均 $E(X|\mathcal{G})$ の一つの変形とよび、 $Y = E(X|\mathcal{G})$ (a.s.) と表す。

条件付き期待値の性質：

(X, Y, X_n は可積分である。 a, b は実数である。)

(1) $E(aX + bY|\mathcal{G}) = aE(X|\mathcal{G}) + bE(Y|\mathcal{G})$ a.s.

(2) $X \geq 0$ であれば $E(X|\mathcal{G}) \geq 0$ a.s.

(3) $E(1|\mathcal{G}) = 1$ a.s.

(4) X は \mathcal{G} で可測、 X, Y 可積ならば、 $E(X|\mathcal{G}) = X$ a.s.

(5) $H \subset \mathcal{G}$ のとき、 $E(E(X|\mathcal{G})|H) = E(X|H)$ a.s.

(6) L_1 で $X_n \rightarrow X$ ならば、 $E(X_n|\mathcal{G}) \rightarrow E(X|\mathcal{G})$ 。

(7) (Jensen 不等式) $\psi: \mathbb{R}^1 \rightarrow \mathbb{R}^1$ は凸であり、 $\psi(X)$ 可積分ならば、

$$\psi(E(X|\mathcal{G})) \leq E(\psi(X)|\mathcal{G}) \quad \text{a.s.}$$

一般的では、 $|E(X|\mathcal{G})| \leq E(|X||\mathcal{G})$

そして $|E(X|\mathcal{G})|^2 \leq E(|X|^2|\mathcal{G})$ 。

(8) X が \mathcal{G} に依存しない。ボレル可測関数 $f(x)$ が可積分の時のみ、

$$E(f(x)|\mathcal{G}) = E(f(x))。$$

(9) Z は \mathcal{G} 可測ならば、 $E(zX|\mathcal{G}) = zE(X|\mathcal{G})$ 。

(10) 独立性 $X_n \in L$, \mathcal{G}, H が σ -alg, で $\sigma(X, \mathcal{G})$ 独立である。

そのとき $E(X) = E[E(X|\mathcal{G})] = E[X|\mathcal{G}]$ 。

定義 3.2 正則条件付確率と確率密度関数 (pdf)

$\{p(\omega, A)\}_{\omega \in \Omega, A \in \mathcal{F}}$ が次の三つ条件を満たすならば、正則条件付確率と呼ぶ。

(1) ω を固定して、 $A \rightarrow p(\omega, A)$ は (Ω, \mathcal{F}) ;

- (2) A を固定して、 $A \in \mathcal{F}$, $\omega \rightarrow p(\omega, A)$ が可測；
 (3) 任意 $A \in \mathcal{F}$, $B \in \mathcal{G}$, $P(A|B) = \int_B p(\omega, A) P(d\omega)$ 。

つまり、任意非負 X と $B \in \mathcal{G}$, $(X : B) = \int_{\Omega} \left\{ I_B(\omega) \int_{\omega} X(\omega') P(\omega, d\omega') \right\} P(d\omega)$
 $\int_{\omega} X(\omega') P(\omega, d\omega')$ が $E(X|G)_{(\omega)}$ と一致する a. s.。

1.3.2 マルチンゲール

定義 3.3 離散時間のマルチンゲール

時刻の集合は $T = \{1, 2, 3, \dots\}$ とし、情報増大系 $\{F_n\}_{n \in T}$ が与えられたとき、実数値離散時間確率過程 X_n , $n \in T$ がマルチンゲールであるとは

- (1) X は適合,
 (2) 任意の時刻 n について $E(|X_n|) < \infty$,
 (3) 任意の時刻 n について $E[X_n | F_{n-1}] = X_{n-1}$ a. s. ($n > 1$)

が成立することである。

マルチンゲール理論が大変に重要であるのは、マルチンゲールが非常に多くの場面に姿を現すということに根拠がある。簡単な例を見る。

(1) 独立で平均 0 の確率変数の和がマルチンゲールである。 X_1, X_2, \dots を独立な確率変数の列で、任意 n に対して $E(|X_k|) < \infty$ であり、かつ $E(X_k) = 0$ とする。 $S_n = X_1 + X_2 + \dots + X_n$ ($S_0 = 0$), $F_n = \sigma(X_1 + X_2 + \dots + X_n)$ とする。このとき $n \geq 1$ に対し a. s.。

(2) 平均 1 の非負独立確率変数の積がマルチンゲールである。 X_1, X_2, \dots を独立な非負確率変数とする。 $M_0 = 1, F_0 = \{0, \Omega\}$ として

$M_n = X_1 X_2 \dots X_n, F_n = \sigma(X_1, X_2, \dots, X_n)$ と定義する。この時、 $n \geq 1$ に対し、

$$E(M_n | F_{n-1}) = E(M_{n-1} X_n | F_{n-1}) = M_{n-1} E(X_n | F_{n-1}) = M_{n-1} E(X_n) = M_{n-1}$$

を (a. s.) の意味で得るから M はマルチンゲールである。

(3) 一つの確率変数に関するデータの累積がマルチンゲールである F_n をフィルトレーションとし、 $\xi \in L^1(\Omega, \mathcal{F}, P)$ とする。適当な変形をとって、

$M_n = E(\xi | F_n)$ と定義しよう。塔の性質によって

$$E(M_n | F_{n-1}) = E(E(\xi | F_n) | F_{n-1}) = E(\xi | F_{n-1}) = M_{n-1} \quad \text{a.s.}$$

を得る。したがって、 M はマルチンゲールである。

定義 3.4 優マルチンゲールと劣マルチンゲール

$(\{F_n\}, P)$ について優マルチンゲールは、定義 3.2 の (1) (2) が成り立つ上、(3) を $E[X_n | F_{n-1}] \leq X_{n-1}$ a. s. ($n > 1$) に置き換えて定義される。そして劣マルチンゲ

ールは(3)を $E[X_n|F_{n-1}] \geq X_{n-1}$ a. s. ($n > 1$)に置き換えて定義される。

定義 3.5 マルチンゲール変換

$f_n, n \in N$ を非負可予測過程とする。劣 (優) マルチンゲールに対し、

$$Y_0 = 0, Y_n = \sum_{k=1}^n f_k (X_k - X_{k-1}), n \geq 1$$

とおく。右辺の各項が可積分ならば Y_n は劣 (優) マルチンゲールである。

定理 3.1 $X_n, n \in N$ を (劣) マルチンゲールとする。 $F(x): R \rightarrow R$ を (増大する) 凸関数とする。 $Y_n = F(X_n)$ がすべての n で可積分ならば、 $Y_n, n \in N$ は劣マルチンゲールである。

この定理は Jensen の不等式より言える。

$$(f(X_n))_{(s)} = f(E[X_{n+1}|F_n]) \leq E[f(X_{n+1})|F_n].$$

1.3.3 一様可積分性

本節では、マルチンゲールの諸性質について考察する。以下、通常条件を満たす情報増大系付き確率空間 $(\Omega, F, G, P), G = \{F_t; t \geq 0\}$ と (Ω, F, G, P) 上で定義された確率過程 $X = \{X(t); t \geq 0\}$ を所与とする。

確率過程の一様可積分性を定義する前に、その理解を容易にするために確率変数の可積分性の必要十分条件について示す。

補題 3.1 確率変数 X が可積分、すなわち、 $E[|X|] < \infty$ であるならば、

$$\forall \epsilon > 0, \exists \delta > 0; E[|X|1_G] < \epsilon; \forall G \in F; P(G) < \delta$$

この補題を証明するには、背理法を用い、Borel-Cantelli の補題と Fatou の補題 (補題 2.4) を利用する。

定理 3.1 確率変数 X が可積分、すなわち、 $E[|X|] < \infty$ であるための必要十分条件は、

$$\lim_{n \rightarrow \infty} E[|X|1_{\{|X|>n\}}] = 0$$

となることである。

証明： X が可積分ならば、 $|X|1_{\{|X|>n\}} \leq |X| < \infty$ a. s.、したがって $\lim_{n \rightarrow \infty} |X|1_{\{|X|>n\}} = 0$ a. s. であるから、Lebesgue の有界収束定理より、定理 3.1 が成立する。

定義 3.5 確率過程 $\{X(t); t \in R^+\}$ が、 $\sup_{t \in R^+} E[X(t)^p] < \infty, p \geq 1$ を満たす時、 p 乗可積分 (square integrable) であるという。特に $p = 1$ のときは、可積分であるという。

定義 3.6 確率過程 $\{X(t); t \in R^+\}$ が、

$$\lim_{n \rightarrow \infty} \sup_{t \in R^+} E[|X(t)|1_{\{|X(t)| > n\}}] = 0$$

となるとき、 X は、一様可積分 (uniformly integrable) であるという。

定理 3.2 $X = \{X(t); t \in R^+\}$ が、一様可積分であれば、 $\sup_{t \in R^+} E[|X(t)|] < \infty$ 、すなわち、 X は可積分である。

定理 3.3 確率過程 $X = \{X(t); t \in R^+\}$ に対して、

$$|X(t)| < Y, t \in R^+, E[|Y|] < \infty$$

となる確率変数 Y が存在ならば、 X は一様可積分である。

証明： $|X(t)| < Y, t \in R^+$ より

$$\sup_{t \geq 0} E[|X(t)|1_{\{|X(t)| > n\}}] \leq E[|Y|1_{\{|Y| > n\}}]$$

かつ、 Y に関して定理 3.1 を用いれば、

$$\lim_{n \rightarrow \infty} \sup_{t \geq 0} E[|X(t)|1_{\{|X(t)| > n\}}] \leq \lim_{n \rightarrow \infty} E[|Y|1_{\{|Y| > n\}}] = 0$$

定理 3.3 より、直ちに次の系を得る：

$E[\sup_{t \in R^+} |X(t)|] < \infty$ であるならば、 $X = \{X(t); t \in R^+\}$ は一様可積分となる。

定理 3.4 $p > 1$ として $X = \{X(t); t \in R^+\}$ が p 乗可積分ならば、一様可積分となる。

定理 3.4 より、 $X = \{X(t); t \in R^+\}$ が 2 乗可積分ならば、 X は一様可積分となる。

定理 3.5 (Doob のマルチンゲール) Y を可積分な確率変数として、

$$X(t) = E[Y|F_t], t \in R^+$$

とおくと、 $X = \{X(t); t \in R^+\}$ は一様可積分なマルチンゲールとなる。

証明： X がマルチンゲールとなることは明らか。したがって、一様可積分性について証明する。

$Y \geq 0$ の場合について証明する。 $X(t_n) \uparrow \infty (t_n \uparrow \infty)$ となる単調増加列 $\{t_n; n \in N\}$ が存在したとすると、単調収束定理より、 $E[X(t_n)] \uparrow \infty$ となるが、これは、 $E[X(t_n)] = E[Y] < \infty$ に矛盾する。したがって、 $X^* = \sup_{t \in R^+} X(t) < \infty$ 。

$\{X(t) > n\} \subset \{X^* > n\}$ と条件付き期待値の定義より、

$$E[|X(t)|1_{\{|X(t)| > n\}}] = E[Y1_{\{|X(t)| > n\}}] \leq E[Y1_{\{|X^*| > n\}}]$$

ここで、 X^* が有限であることと、 Y が可積分であることに注意すると、Lebesgue の有界収束定理より、

$$\lim_{n \rightarrow \infty} E[Y1_{\{|X^*| > n\}}] = 0$$

したがって、 X は一様可積分となる。

定理 3.6 (有界収束定理) $X, X_n, n \in N$, を確率変数とする、このとき、

$$\exists K \in R; |X_n| \leq K, \forall n \in N$$

かつ $\lim_{n \rightarrow \infty} X_n = X$ ならば、

$$\lim_{n \rightarrow \infty} E[|X_n - X|] = 0$$

証明: $\text{plim}_{n \rightarrow \infty} X_n = X$ から、任意の $\epsilon > 0$ に対して、

$$\exists n_0 \in N; P(|X_n - X| > \frac{\epsilon}{3}) < \frac{\epsilon}{3K} \quad \forall n \geq n_0$$

また、 $K \geq |X_n|$ より、

$$\{|X| - K > \frac{1}{k}\} \subset \{|X| - |X_n| > \frac{1}{k}\} \subset \{|X - X_n| > \frac{1}{k}\} \quad \forall k, n \in N$$

なので $P\{|X| - K > \frac{1}{k}\} \leq P\{|X - X_n| > \frac{1}{k}\}$

よって $\text{plim}_{n \rightarrow \infty} X_n = X \Rightarrow P\{|X| > K + \frac{1}{k}\} = 0$

よって $P(\bigcup_{k=1}^{\infty} \{|X| > K + \frac{1}{k}\}) = P(|X| > K) = 0$

すなわち、 $K \geq |X|$ a. s. 以上により

$$\begin{aligned} E[|X_n - X|] &= E[|X_n - X| 1_{|X_n - X| > \frac{\epsilon}{3}}] + E[|X_n - X| 1_{|X_n - X| \leq \frac{\epsilon}{3}}] \leq 2KP(|X_n - X| \\ &> \frac{\epsilon}{3}) + \frac{\epsilon}{3} \leq \epsilon \end{aligned}$$

定理 3.7 (平均収束の必要十分条件) $X, X_n \in L^1(\Omega, F, P), n \in N$, とする。このとき次が成立する。

$$\begin{aligned} &\lim_{n \rightarrow \infty} E[|X_n - X|] = 0 \\ \Leftrightarrow &\begin{cases} \text{plim}_{n \rightarrow \infty} X_n = X \\ \{X_n; n \in N\} \text{は、一様可積分} \end{cases} \end{aligned}$$

1.3.4 停止時

定義 3.6 $\tau: \Omega \rightarrow [0, \infty]$ が

$$\{\tau \leq t\}: \{\omega \in \Omega; \tau(\omega) \leq t\} \in F_t, t \geq 0$$

を満たす時、 τ を $\{F_t\}$ -停止時 (stopping time) といい、情報増大系 $\{F_t\}$ が自明であるときには、単に停止時という。また

$$F_\tau := \{A \subset \Omega; A \cap \{\tau \leq t\} \in F_t, t \geq 0\}$$

で定義される F_τ を τ -停止時前可算加法族 (pre- τ algebra) という。

定理 3.8

- (1) $\tau_1 \leq \tau_2 \Rightarrow F_{\tau_1} \subset F_{\tau_2}$
- (2) $F_{\tau_1 \wedge \tau_2} = F_{\tau_1} \cap F_{\tau_2}$
- (3) $A \in F_{\tau_1 \wedge \tau_2} \Rightarrow A \cap \{\tau_1 \leq \tau_2\} \in F_{\tau_2}$
- (4) $F_{\tau_1 \vee \tau_2} = \sigma(F_{\tau_1}, F_{\tau_2})$
- (5) $F_{t+} := \bigcap_{u>t} F_u$ とすると次が成立する

$$\tau \text{ が } F_{t+} \text{-停止時} \Leftrightarrow \{\tau < t\} \in F_t, t \geq 0$$

$$\begin{aligned} F_{t+} &:= \{A \subset \Omega; A \cap \{\tau \leq t\} \in F_{t+}, t \geq 0\} \\ &= \{A \subset \Omega; A \cap \{\tau < t\} \in F_t, t \geq 0\} \end{aligned}$$

定理 3.9 $\{\tau_n; n \in N\}$ を $\{F_t\}$ -停止時の列とすると、次が成立する

- (1) $\tau_n \uparrow \tau \Rightarrow \tau$ は $\{F_t\}$ -停止時
- (2) $\tau_n \downarrow \tau \Rightarrow \tau$ は $\{F_{t+}\}$ -停止時で $F_{t+} = \bigcap_{n=1}^{\infty} F_{\tau_n+}$

証明：

- (1) $\tau_n \uparrow \tau \Rightarrow \{\tau \leq t\} = \bigcap_{n=1}^{\infty} \{\tau_n \leq t\}$ より明らか
(2) $\tau_n \downarrow \tau \Rightarrow \{\tau < t\} = \bigcap_{n=1}^{\infty} \{\tau_n < t\} \in F_t \quad \therefore \tau$ は $\{F_{t+}\}$ - 停止時
 $E \in \bigcap_{n=1}^{\infty} F_{\tau_n+}$
 $\Leftrightarrow E \cap \{\tau_n < t\} \in F_t, n \in N$ かつ $\tau_n \downarrow \tau \Rightarrow \{\tau < t\} = \bigcup_{n=1}^{\infty} \{\tau_n < t\}$
 $\Leftrightarrow E \cap \{\tau < t\} = \bigcup_{n=1}^{\infty} (E \cap \{\tau_n < t\}) \in F_t$

定理 3.10 X が発展的可測、 τ が停止時ならば、 $\omega \in \Omega \mapsto X(\tau) := X(\tau(\omega), (\omega)) \in R$ は F_{τ} - 可測となる。

1.3.5 離散時間マルチンゲール

本節では、時刻の集合を Z^+ とする。即ち時刻の集合が離散であるとして、離散時間のマルチンゲール $\{X(t); t \in Z^+\}$ について定理をまとめる。

定理 3.11 (Doob の横断数定理 (Doob's upcrossing theorem))

$X = \{X(t); t \in Z^+\}$ を優マルチンゲールとする。時刻 N までにサンプルパスが $[a, b]$ を下から上へと横断する回数を $U_N(X; [a, b])$ とする。すなわち、 $U_N(X; [a, b])$ は

$$X(s_i) < a, X(t_i) > b, i = 1, \dots, k; 0 \leq s_1 < t_1 < \dots < s_k < t_k \leq N$$

となる最大の k である。このとき、次が成立する。

$$E[U_N(X; [a, b])] \leq \frac{E[(X(N) - a)^-]}{b - a}$$

系 3.2 $X = \{X(t); t \in Z^+\}$ を一様可積分なマルチンゲールとし、

$$U_{\infty}(X; [a, b]) := \lim_{N \rightarrow \infty} U_N(X; [a, b])$$

とする。このとき、

$$(b - a)E[U_{\infty}(X; [a, b])] \leq |a| + \sup_{n \in Z^+} E[|X(n)|] < \infty$$

とする。

証明：定理 3.11 と一様可積分性より

$$(b - a)E[U_N(X; [a, b])] \leq |a| + E[|X(N)|] \leq |a| + \sup_{n \in Z^+} E[|X(n)|] < \infty$$

よって、単調収束定理により得る。

定理 3.12 (マルチンゲール収束定理) 優マルチンゲール $X = \{X(t); t \in Z^+\}$ が可積分、すなわち、 $\sup_{t \in Z^+} E[|X(t)|] < \infty$ であれば、確率 1 で有限な極限 $X(\infty) := \lim_{t \rightarrow \infty} X(t)$ が存在する。

証明：

$$\Lambda := \{\omega \in \Omega; \{X(t, \omega); t \in Z^+\} \text{ が } [-\infty, +\infty] \text{ で極限を持たない}\}$$

とおくと、

$$\Lambda = \{\omega \in \Omega; \liminf_{t \rightarrow \infty} X(t, \omega) < \limsup_{t \rightarrow \infty} X(t, \omega)\}$$

$$= \bigcup_{a, b \in Q; a < b} \{\omega \in \Omega; \liminf_{t \rightarrow \infty} X(t, \omega) < a < b < \limsup_{t \rightarrow \infty} X(t, \omega)\}$$

ここで、

$$\Lambda_{a, b} = \{\omega \in \Omega; \liminf_{t \rightarrow \infty} X(t, \omega) < a < b < \limsup_{t \rightarrow \infty} X(t, \omega)\}$$

とし、 $U_\infty(X; [a, b])$ を系 3.2 で定義したものとすると、

$$\Lambda_{a,b} \subset \{\omega \in \Omega; U_\infty(X; [a, b]) = \infty\}$$

となる。一方、系 3.2 より、 $P(U_\infty(X; [a, b]) = \infty) = 0$ であるから $P(\Lambda_{a,b}) = 0$ となり、 $P(\Lambda) = 0$ を得る。さらに、Fatou の補題(補題 2.4)より、

$$E[\lim_{t \rightarrow \infty} |X(t)|] = E[\liminf_{t \rightarrow \infty} |X(t)|] \leq \liminf_{t \rightarrow \infty} E[|X(t)|] \leq \sup_{t \in Z^+} E[|X(t)|] < \infty$$

となるから、 $[X(\infty)] < \infty$ a. s. となる。

系 3.3

- (1) 一様可積分マルチンゲールは、確率 1 で収束する。
- (2) 2 乗可積分マルチンゲールは、確率 1 で収束する。
- (3) 正值マルチンゲールは、確率 1 で収束する。

定理 3.13 $X = \{X(t); t \in Z^+\}$ が、一様可積分優マルチンゲールならば、 $X(t) \xrightarrow{L^1} X(\infty)$ ($t \rightarrow \infty$)となり、

$$X(t) \geq E[X_\infty | F_t] \text{ a. s.}, \quad t \in Z^+$$

特に X がマルチンゲールならば、等号が成立する。

定理 3.14 (Levy-Doob の下方定理) 所与の確率空間 (Ω, F, P) に対して、 $\{G_n; n \in -N\}$ を F の部分可算加法族の族で、

$$G_{-\infty} := \bigcap_{k=1}^{\infty} G_{-k} \subset \cdots \subset G_{-2} \subset G_{-1}$$

となるものとする。 $X = \{X(n); n \in -N\}$ を $\{G_n; n \in -N\}$ -優マルチンゲールとする。すなわち、

$$E[X(n) | G_m] \leq X(m) \text{ a. s.}, \quad m \leq n \leq -1, \quad E[|X(n)|] < \infty \quad \forall n \in -N$$

とする。このとき、 X は、一様可積分であり、極限 $X(-\infty) := \lim_{n \rightarrow -\infty} X(n)$ が概収束かつ平均収束の意味で存在し、

$$E[X(n) | G_{-\infty}] \leq X(-\infty) \text{ a. s.}, \quad n \in -N$$

が成立する。また、 X がマルチンゲールならば、最後の不等式は等号が成立する。

定理 3.15 (Doob の劣マルチンゲール不等式) $X = \{X(t); t \in Z^+\}$ を非負値劣マルチンゲールとする。このとき、任意の $c > 0$ と $n \in Z^+$ に対して次に不等式が成立する。

$$cP(\sup_{k \leq n} X(k) \geq c) \leq E[X(n) 1_{\{\sup_{k \leq n} X(k) \geq c\}}] \leq E[X(n)]$$

定理 3.16 (Doob 分解)

- (1) $X = \{X(n); n \in Z^+\}$ が $X(n) \in L^1(\Omega, F, P), n \in Z^+$ 、となる適合過程とする。

このとき、 $M = \{M(n); n \in Z^+, M(0) = 0\}$ をマルチンゲール、 $A = \{A(n); n \in Z^+, A(0) = 0\}$ を可予測過程として、

$$X(n) = X(0) + M(n) + A(n), \quad n \in Z^+$$

というように分解できる。また、この分解は、マルチンゲール \tilde{M} と可予測過程 \tilde{A} を

$$P(M(n) = \tilde{M}(n), A(n) = \tilde{A}(n), n \in Z^+) = 1$$

として、

$$X(n) = X(0) + \tilde{M}(n) + \tilde{A}(n), \quad n \in Z^+$$

となるという意味で一意である。

(2) X が劣マルチンゲールとなるとは、

$$P(A(n) \leq A(n+1), n \in Z^+) = 1$$

となるときであり、また、その時に限られる。

証明：

$$A(n) := \sum_{k=1}^n E[X(k) - X(k-1) | F_{k-1}] \quad a.s., \quad n \in N, A(0) := 0$$

とすると A は可予測、かつ、

$$E[(X(n) - A(n)) - (X(n-1) - A(n-1)) | F_{k-1}] = 0$$

となる。ここで、 $M(n) := X(n) - X(0) - A(n)$ とすれば、 M はマルチンゲールとなり、Doob の分解を得る。また、(2) の劣マルチンゲールについては、その定義より自明である。

補題 3.2 $X = \{X(n); n \in Z^+\}$ を優マルチンゲールもしくは劣マルチンゲールとすると、任意の $c > 0$ に対して、次が成立する。

$$cP(\sup_{k \leq N} |X(k)| \geq 3c) \leq 4E[|X(0)|] + 3E[|X(N)|], N \in Z^+$$

定理 3.17 (Levy の上方定理) 所与の $\xi \in L^1(\Omega, F, P)$ に対して、 $M(n) := E[\xi | F_n], n \in Z^+$ 、とすると $M := \{M(n); n \in Z^+\}$ は、一様可積分なマルチンゲールで

$$M(n) \rightarrow \eta := E[\xi | F_n](n \rightarrow \infty) a.s., \text{ かつ } M(n) \xrightarrow{L^1} \eta(n \rightarrow \infty)$$

定理 3.18 τ を停止時とし、 $X = \{X(t); t \in Z^+\}$ をマルチンゲール (優マルチンゲール) とする。このとき $\{X(t \wedge \tau); t \in Z^+\}$ はマルチンゲール (優マルチンゲール) となる。

X がマルチンゲールの場合では

$$\sum_{k=1}^n \mathbf{1}_{\{k \leq \tau\}} (X(k) - X(k-1)) = X(n \wedge \tau) - X(0)$$

$$\{\mathbf{1}_{\{n \leq \tau\}} = 0\} = \{\tau \leq n-1\} \in F_{n-1}, \{\mathbf{1}_{\{n \leq \tau\}} = 1\} = \{\tau \leq n-1\}^c$$

となることに注意すると、 $X(t \wedge \tau)$ は F_n -可測かつ

$$E[X(n \wedge \tau) - X(n-1 \wedge \tau) | F_{n-1}] = 0$$

$$\therefore E[X(n \wedge \tau) | F_{n-1}] = X(n-1 \wedge \tau), n \in N$$

定理 3.19 τ を停止時とし、 X を優マルチンゲールとする。このとき、次の (1) ~ (3) ;

(1) N があって、 $\tau(\omega) \leq N, \forall \omega \in \Omega$

(2) $\exists K \in R; |X(n)| \leq K, \forall (n, \omega) \in Z^+ \times \Omega$ かつ $\tau < \infty a.s.$

(3) $E[\tau] < \infty$ かつ $\exists K \in R; |X(n) - X(n-1)| \leq K, \forall (n, \omega) \in Z^+ \times \Omega$

のいずれかが成り立つならば、 $X(\tau)$ は可積分であり、

$$E[X(\tau)] \leq E[X(0)]$$

が成立する。また、 X がマルチンゲールの場合には、同条件の下で最後の不等式が等号で成立する。すなわち、

$$E[X_\tau] = E[X_0]$$

となる。

補題 3.3 X を優マルチンゲールとし、 τ をある N に対して、 $\tau \leq N$ となる停止時とする。この時、 $X(\tau) \in L^1(\Omega, F_\tau, P)$,かつ次が成立する。

$$E[X(N)|F_\tau] \leq X(\tau)$$

証明： $E \in F_\tau$ とすると

$$E[X(N)1_E] = \sum_{n=0}^N E[X(N)1_{E \cap \{\tau=n\}}] \leq \sum_{n=0}^N E[X(n)1_{E \cap \{\tau=n\}}] = E[X(\tau)1_E]$$

$$|X(\tau)| \leq \sum_{n=0}^N |X(n)|$$

より、 $E[|X(\tau)|] < \infty$

定理 3.20(任意抽出定理) τ_1, τ_2 を $0 \leq \tau_1 \leq \tau_2 \leq \infty$ となる停止時とする。 X が一樣可積分マルチンゲールならば、次が成立する。

$$E[X(\tau_2)|F_{\tau_1}] \leq X(\tau_1) \text{ a. s.}$$

証明： τ を任意の停止時とする。補題 3.3 と定理 3.13 より、

$$X(\tau \wedge k) = E[X(k)|F_{\tau \wedge k}] = E[E[X(\infty)|F_k]|F_{\tau \wedge k}] = E[X(\infty)|F_{\tau \wedge k}]$$

定理 3.17 より、 $k \rightarrow \infty$ として

$$X(\tau) = E[X(\infty)|G], G := \sigma(\cup_k F_{\tau \wedge k})$$

ここで、 $E \in F_\tau$ とすると、 $E \cap \{\tau \leq k\} \in F_{\tau \wedge k}$ 。したがって、 $E \cap \{\tau < \infty\} \in G$ となり、

$$E[X(\tau)1_{E \cap \{\tau < \infty\}}] = E[X(\infty)1_{E \cap \{\tau < \infty\}}]$$

したがって、

$$E[X(\tau)1_{E \cap \{\tau = \infty\}}] = E[X(\infty)1_{E \cap \{\tau = \infty\}}]$$

以上より、

$$E[X(\tau)1_E] = E[X(\infty)1_E]$$

これより、 $F \in F_{\tau_1}$ とすると

$$E[X(\tau_2) - X(\tau_1)1_F 1_{(\tau_1, \tau_2)}] = 0$$

となるので、証明ができる。

定理 3.21 マルチンゲール中心極限定理

(1) X_1, X_2, \dots がマルチンゲールとは $E(X_{n,i} | F_{n,i-1}) = X_{n,i-1}$ であるのでマルチンゲール差分 $Y_{n,i} = X_{n,i} - X_{n,i-1}$ であり、 $E(Y_{n,i} | F_{n,i-1}) = 0$ 。

(2) 二次モーメント $E(Y_{n,i}^2) < \infty$ 。

(3) 条件付き分散 $\sigma_{n,i}^2 = E[Y_{n,i}^2 | F_{n,i-1}]$ とするならば $\sum_{k=1}^{\infty} \sigma_{n,k}^2 \xrightarrow{p} \sigma^2$ 。

(4) $\sum_{k=1}^{\infty} E\left[Y_{n,k}^2 1_{\{|Y_{n,k}| \geq \varepsilon\}}\right] \xrightarrow{n \rightarrow \infty} 0$ とする。

この時、 $\sum_{k=1}^{\infty} Y_{n,k} \xrightarrow{\text{分布収束}} N(0, \sigma^2)$ 。この条件はリンデベルグ条件と呼ぶ。

1.4 確率過程

1.4.1 ブラウン運動

定義 4.1 B_t は標準ブラウン運動であるとの条件：

- (1) $B_0 = 0$
- (2) 各 ω に対し、 $t \rightarrow B_t(\omega)$ は連続
- (3) $t > s \geq 0$ の時、 $B_t - B_s \sim N(0, t-s) \perp F_s = \sigma\{B_u : u \leq s\}$ 。

定理 4.1 ブラウン運動 $\{B_t, t \geq 0\}$ は次の性質をもつ：

- (1) (ガウス系) $\{B_t\}$ はガウス系である、つまり、任意の線形結合 $a_1 B_{t_1} + \dots + a_n B_{t_n}$, $a_i \in R, t_i \geq 0$ は正規分布に従う確率変数である。
- (2) (平均) $E[B_t] = 0$
- (3) (共分散) $Cov(B_s, B_t) = E[B_s B_t] = \min\{s, t\}$
- (4) (初期値) $B_0 = 0$
- (5) (サンプルパスの連続性) $t \rightarrow B_t$ は確率 1 で連続。
- (6) (独立増分) $0 \leq t_1 < t_2 < \dots < t_n$ のとき、
 $B_{t_0}, B_{t_1}, B_{t_2} - B_{t_1}, B_{t_3} - B_{t_2}, \dots, B_{t_n} - B_{t_{n-1}}$ は独立である。

定理 4.2 ブラウン運動はマルコフ過程である。

任意の $t > s \geq 0$ に対し、 $E(B_t^i | F_s) = B_s^i, i = 1, 2, \dots$

$E\{(B_t^i - B_s^i)(B_t^j - B_s^j) | F_s\} = E\delta_{ij}(t-s), i = 1, 2, \dots$ が成立する。

定理 4.3 Lévy の定理 (1948)

M_t : 連続 $\{F_t\}$ についてマルチンゲール, $M_0 = 0$ が $\langle M \rangle_t = t, (t > 0)$ を満たす時、

M_t は $\{F_t\}$ についてのブラウン運動となる。

($t > s$ である。 $M_t^2 - \langle M \rangle_t$ がマルチンゲールになるもの。

$E(M_t^2 - \langle M \rangle_t | F_s) = M_s^2 - \langle M \rangle_s = E(M_t^2 - M_s^2 | F_s) = E(\langle M \rangle_t - \langle M \rangle_s | F_s)$ 。)

1.4.2 確率積分と伊藤の公式

一般に、ブラウン運動 $\{B_t\}$ に適合した確率過程 $\{G(t)\}$ に対して、伊藤型確率積分が定義される： $I(t) = \int_0^t G(s)dB_s$.

$\{I(t)\}$ もまたブラウン運動 $\{B_t\}$ に適合した確率過程になる。

したがって、確率微分方程式はブラウン運動 $\{B_t\}$ に適合した確率過程 $\{X(t)\}$ が

$$X(t) = x_0 + \int_0^t f(s, X(s))ds + \int_0^t g(s, X(s))dB_s$$

を満たすものを求めよという問題なのである。確率過程 $\{X(t)\}$ が $X(t)$ 形をもつとき、伊藤過程 (Itô process) という。

微分形を用いて $dX = Fds + GdB$ と略記することも多い。伊藤過程 $X(t)$ を関数 $f(t, x)$ と合成してできる確率過程 $Y(t) = f(t, X(t))$ もまた伊藤過程になる。したがって、

$$Y(t) = y_0 + \int_0^t F_1(s)ds + \int_0^t G_1(s)dB_s \text{ または } dY = F_1dt + G_1dB$$

の形をもつ。ここで $F_1(s)$ と $G_1(s)$ の具体形を与えるのが伊藤の公式である：

定理(伊藤の公式) 2 変数関数 $f(t, x)$ が t に関して1回連続的微分可能、 x に関して2回連続的微分可能であれば、

$$dY = \frac{\partial f}{\partial t}(t, X(t))dt + \frac{\partial f}{\partial x}(t, X(t))dx + \frac{1}{2} \frac{\partial^2 f}{\partial x^2}(t, X(t))(dx)^2$$

ここで

$$dX = Fdt + GdB, \quad (dX)^2 = G^2 dt$$

従って

$$F_1 = \frac{\partial f}{\partial t}(t, X(t)) + \frac{\partial f}{\partial x}(t, X(t))F(t) + \frac{1}{2} \frac{\partial^2 f}{\partial x^2}(t, X(t))G(t)^2$$

$$F_2 = \frac{\partial f}{\partial t}(t, X(t))G(t)$$

が得られる。

なお、伊藤の公式の本質は、次の積公式にある(伊藤ルール)：

$$(dt)^2 = (dB)(dt) = (dt)(dB) = 0, (dB)^2 = dt .$$

1.4.3 Bessel 過程

$\beta = (\beta_1, \beta_2, \dots, \beta_n)$ は n 次ブラウン運動であり、 $n \in N$ 。ある確率過程 R は $R_t = \|\beta_t\|$, i. e., $R_t^2 = \sum_{i=1}^n (\beta_i)^2(t)$ と定義される。伊藤の公式より $dR_t^2 = \sum_{i=1}^n 2\beta_i(t) d\beta_i(t) + ndt$ 。

任意 $t > 0$, $P(R_t = 0) = 0$, 確率過程 W が次の式で定義される：

$$dW_t = \frac{1}{R_t} \sum_{i=1}^n \beta_i(t) d\beta_i(t)$$

W は実数値付きのブラウン運動で、確率過程 R は次の式に満たす：

$$d(R_t^2) = 2R_t dW_t + ndt.$$

ここでは $\rho_t = R_t^2$ と設定する。

$$d\rho_t = 2\sqrt{\rho_t}dW_t + ndt$$

伊藤の公式を用いて $dR_t = dW_t + \frac{n-1}{2} \frac{dt}{R_t}$ を得る。従って、 R は Bessel 過程と呼ばれる。 ρ は n 乗 Bessel 過程と呼ばれる。

定義4.2 t 各 $\delta \geq 0$ と $x \geq 0$ について、確立過程 Z_t は初期値 x を持ち、 $BESQ^\delta(x)$ と記入される：

$$Z_t = x + 2 \int_0^t \sqrt{Z_s} d\beta_s + \delta t.$$

δ は Z_t の次元である。

定理 4.4 $\delta > 0$ について、 y の密度を有する $BESQ^\delta$ の半群は

$$q_t^\delta(x, y) = \frac{1}{2} \left(\frac{y}{x}\right)^{\frac{\nu}{2}} \exp\left(-\frac{x+y}{2t}\right) I_\nu\left(\frac{\sqrt{xy}}{t}\right), \quad t > 0, x > 0$$

ν が δ に対応するそして I_ν が ν についての Bessel 関数。
 $x = 0$ において、 y の密度を

$$q_t^\delta(0, y) = (2t)^{-\frac{\delta}{2}} \Gamma(\delta/2)^{-1} y^{\frac{\delta}{2}-1} \exp\left(-\frac{y}{2t}\right).$$

で表す。さらに $x' = x^2, y' = y^2$ 変数変換で $q_t^\delta(x, y)$ は次の式になる：

$$p_t^\delta(x, y) = t^{-1} \left(\frac{y}{x}\right)^\nu \exp\left(-\frac{x^2 + y^2}{2t}\right) I_\nu\left(\frac{xy}{t}\right), \text{ for } t > 0, x > 0$$

そして

$$p_t^\delta(0, y) = 2^{-\nu} t^{-(\nu+1)} \Gamma(\nu+1)^{-1} y^{2\nu+1} \exp\left(-\frac{y^2}{2t}\right).$$

第二章 単一変量分枝過程について

2.1 分枝過程の母関数について

非負整数 Z_n においては分枝過程がマルコフ連鎖である。

$\{Z_n: n=0,1,2,\dots\}$ である時、分枝過程の確率関数は $\{p_k: k=0, 1, 2, \dots\}$, $p_k \geq 0$,

$$\sum p_k = 1$$

で定義される。

$$P(i, j) = P\{Z_{n+1} = j | Z_n = i\} = \begin{cases} p_j^{*i} & \text{if } i \geq 1 \quad j \geq 0, \\ \delta_{0j} & \text{if } i = 0 \quad j \geq 0. \end{cases} \quad (2.1)$$

$$p_k^{*i}: k=0,1,2,\dots$$

分枝過程はある集団の各元素の繁殖過程を表すと思われる。初期値は Z_0 で、各元素が独立で確率関数 $\{p_k\}$ の原則の下でランダムな数で次世代の子孫を作る。このように作られた子孫の和が Z_1 (第一世代) になる。 Z_1 より Z_2 を造って続ける。一つの元素は同じ「お親」より作られた「子孫」とも同一世代の他の元素とも独立である。第 n 世代の元素の数量は確率変数 Z_n 。(2.1) より観ると、 $Z_n=0$ ならば $Z_{n+1}=0$ になる。こうして、任意世代においても数量 0 の意味はこの分枝過程が全滅になる。

分枝過程についての研究が長い歴史のあるものである。研究者たちの予想通り、このモデルは物理や生物などたくさんの分野に応用できる。この問題は 1873 年 Francis Galton によって初めて導入された。最初はイギリス貴族の苗字全滅に関する研究であった。そして Reverend Henry Watson によってうまく解決された。論文 Galton and Watson (1874) に載っているため、分枝過程は Galton and Watson 過程と呼ばれる。分枝過程の歴史については D. Kendall (1966) を参照したほうが良い。初期の文献目録を知るため T. E. Harris (1966) を参照しても良い。

分枝過程の問題に関する研究では、初期値に特に注意を引かれるため、 $\{Z_n^{(i)}, n=0,1,2,\dots\}$ を初期値 i の分枝過程を表示する。多くの場合は $Z_0=1$ とするため、 $Z_n^{(1)} = Z_n$ で書く。

数学において、母関数 (ぼかんすう、英: *generating function*;) は、(自然数で添字付けられた) 数列 $\{a_n\}$ に関する情報を内包した係数を持つ、形式冪級数である。母関数は、一般線型回帰問題の解決のためにド・モアブルによって 1730 年に初めて用いられた。複数の自然数で添字付けられる数の配列 (多重数列) の情報を取り込んだ多変数冪級数を同様に考えることもできる。

母関数には、通常型母関数 (ordinary generating function)、指数型母関数 (exponential generating function)、ランベルト級数 (Lambert series)、ベル級数 (Bell series)、ディリクレ級数 (Dirichlet series) など様々なものがある。これらについては定義と例を後述する。原理的にはあらゆる列についてそれぞれの種類の母関数が存在する (ただし、ランベルト級数とディリクレ型は添字を 1 から始めることが必要) が、扱いやすさについてはそれぞれの種類で相当異なるかもしれない。どの母関数が最も有効かは、その列の性質と解く

べき問題の詳細に依存する。

分枝過程を分析するときの一つ重要な手段である通常型母関数 (Generating Functions) というのは、(2.2)式で定義される。

$$f(s) = \sum_{k=0}^{\infty} p_k s^k, |s| \leq 1 \quad (2.2)$$

(2.2) 式を繰り返してみると

$$f_0(s) = s, f_1(s) = f(s), f_{n+1}(s) = f[f_n(s)]$$

なので

$$\begin{aligned} \sum_j P(1, j) s^j &= f(s); \\ \sum_j P(i, j) s^j &= [f(s)]^i, i \geq 1. \end{aligned} \quad (2.3)$$

理由を簡単に説明すると

$$(2.1) \text{ より } \sum_j P(1, j) s^j = \sum_j p_j s^j = f(s)。$$

(2.3) に対して $l_1 + l_2 + \dots + l_i = j$ とすると

$$\begin{aligned} \sum_j P(i, j) s^j &= \sum_j \sum_{l_1 + l_2 + \dots + l_i} P(l_1) P(l_2) \dots P(l_i) s^{l_1 + l_2 + \dots + l_i} \\ &= \sum_{l_1} P(l_1) s^{l_1} \sum_{l_2} P(l_2) s^{l_2} \dots \sum_{l_i} P(l_i) s^{l_i} \\ &= [f(s)]^i, i \geq 1 \end{aligned}$$

ここでは $P_n(i, j)$ を n 回遷移確率にして $P_n(1, j) = P(Z_n = j | Z_0 = 1)$ 。

コルモゴロフ不等式より

$$\begin{aligned} \sum_j P_{n+1}(1, j) s^j &= \sum_j \sum_k P_n(1, k) P(k, j) s^j = \sum_k P_n(1, k) \sum_j P(k, j) s^j \\ &= \sum_k P_n(1, k) [f(s)]^k \end{aligned}$$

$$\sum_j P_n(1, j) s^j = f_{(n)}(s) \text{ とすると, } f_{(n+1)}(s) = f_{(n)}[f(s)].$$

これでは一つの重要な公式 (2.4) を証明された :

$$f_{(n)}(s) = f_n(s) \quad (2.4)$$

(2.3) と (2.4) 式より

$$\sum_{j=0}^{\infty} P_n(i, j) s^j = [f_n(s)]^i.$$

分枝過程のモーメントは下の式で表す。

$$EZ_1 = \sum P(1, j) j = f'(1) = m$$

さらに

$$EZ_n = \sum P_n(1, j) j = f'_n(1) = f'_{n-1}(1) f'(1) = \dots = [f'(1)]^n = m^n$$

また、次の式を用い、 $f''_{n+1}(1) = f''(1) [f'_n(1)]^2 + f'(1) f''_n(1)$, $f''_n(1)$ を求められる。

$$f''_n(1) = f''(1) [m^{2n-2} + m^{2n-3} + \dots + m^{n-1}]$$

この故に、 $\sigma^2 = \text{variance } Z_1$ とすると、

$$\text{var } Z_n = \begin{cases} \frac{\sigma^2 m^{n-1} (m^n - 1)}{m - 1} & \text{if } m \neq 1 \\ n\sigma^2 & m = 1 \end{cases} \quad (2.5)$$

2.2 パワーシリーズ分布

パワーシリーズ分布はべき級数より構成され、部分集合 N 上の離散分布である。 $a = (a_0, a_1, a_2, \dots)$ が非負実数列と想定する。我々は係数のシーケンスとして持つべき級数に興味を持っている。最初 n 次の部分和を書いておく：

$$g_n(\theta) = \sum_{k=0}^n a_k \theta^k, \theta \in R \quad (2.6)$$

そしてパワーシリーズ g が $g(\theta) = \lim_{n \rightarrow \infty} g_n(\theta)$, $\theta \in R$ と定義される：

$$g(\theta) = \sum_{n=0}^{\infty} a_n \theta^n, \theta \in R \quad (2.7)$$

$\theta = 0$ の時、そのシリーズが収束することを注意する。そして $g(0) = a_0$, この場合のほか、 $r \in [0, \infty)$ で $|\theta| < r$ のときシリーズが絶対収束、そして $|\theta| > r$ のとき発散するような r が必ず存在する。 r が収束の半径であり、 $r > 0$ と仮定する。 $r < \infty$ であればこのシリーズが r で絶対収束あるいは発散する。 $-r$ でこのシリーズが絶対収束あるいは条件収束あるいは発散する。これから $\theta \in [0, r)$ の空間に制限する。この空間がパラメータ空間である。 $r < \infty$ また $\theta = r$ の場合、次の結果も成立するかもしれないが、説明が複雑で不必要である。

確率変数 N は関数 g (パラメータ $\theta \in [0, r)$, シーケンス $a = (a_0, a_1, a_2, \dots)$ を持つ) に関連つけられたパワーシリーズ分布を有する。 N が離散型確率密度関数を持つ：

$$f_{\theta}(n) = \frac{a_n \theta^n}{g(\theta)}, n = 0, 1, 2, \dots \quad (2.8)$$

そして分布関数 $F_{\theta}(n) = \sum_{k=0}^n f_{\theta}(k)$ であるため、 N の分布関数

$$F_\theta(n) = \frac{g_n(\theta)}{g(\theta)}, n=0,1,2,\dots \quad (2.9)$$

$g(\theta)$ を利用して N の確率母関数を簡単に作れる :

$$t \in (0, r/\theta) \text{ に対して } P(t) = \sum_{n=0}^{\infty} t^n f_\theta(n) = \frac{1}{g(\theta)} \sum_{n=0}^{\infty} a_n (t\theta)^n = \frac{g(t\theta)}{g(\theta)} \quad \text{よって}$$

N の確率母関数 P は次の式で表す :

$$P(t) = E(t^N) = \frac{g(t\theta)}{g(\theta)}, \quad t < \frac{r}{\theta} \quad (2.10)$$

従って

$$P'(t) = \frac{g'(t\theta)}{g(\theta)} \cdot \theta, P''(t) = \frac{g''(t\theta)}{g(\theta)} \cdot \theta^2 \quad \text{が明らかである。}$$

確率母関数の性質により $E(N) = P'(1), V(N) = P''(1) + P'(1) - [P'(1)]^2$ 。 よって

$$E(N) = \frac{\theta g'(\theta)}{g(\theta)} \quad (2.11)$$

$$V(N) = \theta^2 \cdot g''(\theta) / g(\theta) + \theta \cdot g'(\theta) / g(\theta) - [\theta \cdot g'(\theta) / g(\theta)]^2 \quad \text{。}$$

また、 $E(N)$ と $V(N)$ が次の関係に満たすことを注意し、 $E(N)$ と $V(N)$ をさらに定義することができる。

$E(N)$ は θ について微分する。

$$\partial E(N) / \partial \theta = \theta g''(\theta) / g(\theta) + g'(\theta) / g(\theta) - \theta (g'(\theta))^2 / g^2(\theta) \quad \text{があるため}$$

$$\theta \cdot \partial E(N) / \partial \theta = V(N) \quad \text{よって}$$

$$V(N) = \left(\frac{\partial}{\partial E(N)} \log \theta \right)^{-1} \quad (2.12)$$

2.3 移民項のある分枝過程の逐次確率比検定の重要な結論

2.3.1 モデルと特性値の推定

$$z_0 = 1, z_n = \sum_{i=1}^{z_{n-1}} \xi_{n,i} + Y_n \quad (2.13)$$

分枝過程では、 z_0 で最初の集団の個体数を表わす。 z_n は第 n 世代の集団の個体数を表わす。集団の各々の個体は他のものとは独立に子を生みだす。たとえば、第 $n-1$ 世代第 i 番目の個体が $\xi_{k,i}$ 個の新しい個体を生みだす。言い方を変えれば、 $\xi_{k,i}$ は Z_{k-1} 世代の各成員の子孫数である。 Z_{k-1} 世代各成員の子孫の全体は $\sum_{i=1}^{z_{k-1}} \xi_{k,i}$ で表示し、子孫項と呼ぶ。 Y_n が移民項であり、第 n 世代の外来の成員である。この Z_{n-1} 世代子孫項と第 n 世代の外来項を合わせて、第 n 世代の集団の全体 Z_n になる。この Z_n 整数値の確率変数の列でマルコフ連鎖をつくる。 ξ_{ni} が子孫数として非負である。

Y_n が移民項として正であれば外から集団の中に来る場合、零であるときに集団の中と外交流がない場合、負であれば集団の成員が集団の外に出ていく場合、それぞれを表わす。

このモデルでは、観測されるもの Z_n, Y_n , 観測されないもの $\sum_{i=1}^{Z_{n-1}} \xi_{ni}$ と仮定する。

図2-1は子孫、移民、集団全体の関係を表わす。黒い点が子孫、白い点が移民、縦線上の全部点を合わせて世代の全体になる。

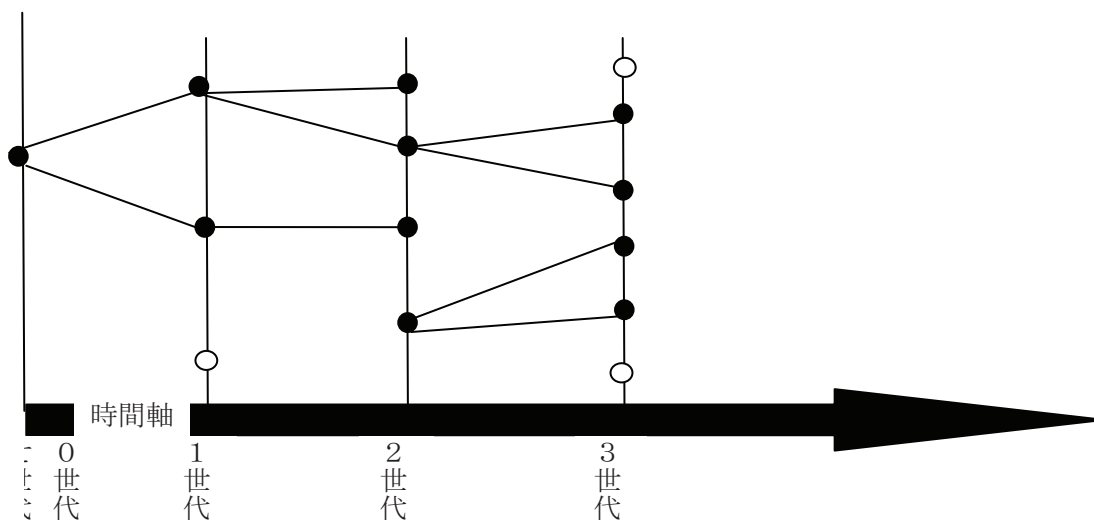


図 2-1

直観的に以下の例で詳しく説明する。

例1 ネズミ講

ねずみ講は、ピラミッド型の階層組織を形成し、上の階層にいる者が下の者から金を吸い上げるという形で、配当のみを目的として活動している組織である。たとえば、100万円で会員になる。新しい会員をどんどん紹介する。新会員が5人増えるごとに100万円の紹介料をもらえる。紹介による新会員の方もその後新しい会員を紹介し、5人ごとに紹介料100万円をもらえる。この方法で講の主催者は黙っていても会員が増えて行き、大儲けできる。初期に会員になった人もかなり儲かれる。美味しい儲け話なので、会員数が爆発的に増える。ちょうど鼠の子孫が爆発的に増えて行く様子に似ているので、ねずみ講と言われる。でも人間の数に限りがある以上、新入会員は必ず頭打ちになって、最後の方に会員になった人たちは大きく損してしまふ。これは法律で禁止されるものである。この例では一人対して新紹介会員数が ξ_{ni} である。第1代の会員 z_1 を新しい人を紹介して、第二代の会員 $z_2 = \sum_i^1 \xi_{ni}$ となる。この例では移民項がない。

例2 新型インフルエンザの流行

日本国内の患者数は Z_n とすれば、医療機関の報告より統計を取ることができる。一人の患者が ξ_{ni} 人に移す。インフルエンザの流行に伴って、新しい患者もどんどん増える。この移された人たちは「子孫」と呼び、 $\sum_{i=1}^{Z_{i-1}} \xi_{ni}$ で表す。この子孫の数は観測できない。国内ではなく、海外よりインフルエンザを感染して日本に来た人は移民(Y_n)である。 Y_n は出入国の記録などのデータより観測可能である。この例と同じように伝染病の流行が移民項のある分枝過程モデルを用いて説明してもよい。

2.3.2 最尤推定のシミュレーションと結果

2.3.2.1 ポアソン分布よりの最尤推定とシミュレーション

ξ_{ni} は期待値 μ のポアソン分布に従う、 $\{\xi_{n,i}\}$ は独立である。

$$\xi_{ni} \sim \text{Po}(\mu) = \frac{e^{-\mu} \mu^{\xi_{ni}}}{\xi_{ni}!}, \quad i=1, 2, \dots, n=1, 2, \dots$$

Y_{ni} は期待値 λ のポアソン分布に従う、 $\{Y_{n,i}\}$ は独立である。

$$Y_{ni} \sim \text{Po}(\lambda) = \frac{e^{-\lambda} \lambda^{y_{ni}}}{y_{ni}!}, \quad i=1, 2, \dots, n=1, 2, \dots \quad \text{さらに}\{\xi_{n,i}\}\text{と}\{Y_{n,i}\}\text{は独立であ}$$

る。

最尤推定量は

$$\hat{\mu} = \frac{\sum_{i=1}^n Z_i - \sum_{i=1}^n Y_i}{\sum_{i=1}^n Z_{i-1}}, \quad \hat{\lambda} = \frac{\sum_{i=1}^n Y_i}{n} = \bar{Y} \quad (2.14)$$

μ の値は三つの可能性がある。 $\mu < 1$ (subcritical) , $\mu = 1$ (critical) と $\mu > 1$ (supercritical)。インフルエンザの流行には、 μ は Z の期待値として、患者は平均的に何人に移すか、その子孫の規模を表わす。 $\mu < 1$ の時に、子孫の規模がちいさくなるから、患者の規模を縮減していることを表示し、流行がいつか終わると考えられる。さらに、 Z_n の期待値を計算すると

$$\begin{aligned} E(Z_n) &= E(E(Z_n | Z_{n-1})) = E(Z_{n-1}\mu + Y_n) \\ &= E(Z_{n-1})\mu + \lambda = \lambda + \mu\lambda + \mu^2\lambda + \dots + \mu^{n-1}\lambda = \lambda \left(\frac{1 - \mu^n}{1 - \mu} \right) \end{aligned} \quad (2.15)$$

これより、 $\mu > 1$ の時に、 Z_n の期待値がかなり大きくなると見える。患者の規模はどんどん拡大し、パンデミック（感染爆発）の可能性が高いと考えられる。

ポアソン分布よりのシミュレーションの結果を見る：

m : 世代 ; n : シミュレーション回数

図 2-2 は $m=1000, n=1000, \mu = 0.7, \lambda = 5$ と設定するとき $\hat{\mu}$ と $\hat{\lambda}$ のヒストグラムである。

図 2-3 は $m=2000, n=1000, \mu = 1.005, \lambda = 6$ と設定する時 $\hat{\mu}$ と $\hat{\lambda}$ のヒストグラムである。

2.3.2.2 幾何分布よりの最尤推定とシミュレーション

ξ_{ni} は成功確率 p の幾何分布に従う, $\{\xi_{n,i}\}$ は独立である。

$$\xi_{ni} \sim Geo(p) = (1-p)^{\xi_{ni}} p, \quad i=1, 2, \dots, n=1, 2, \dots$$

Y_{ni} は成功確率 q の幾何分布に従う, $\{Y_{n,i}\}$ は独立である。

$$Y_{ni} \sim Geo(q) = (1-q)^{Y_{ni}} q, \quad i=1, 2, \dots, n=1, 2, \dots \quad \text{さらに} \{\xi_{n,i}\} \text{ と } \{Y_{n,i}\} \text{ は独立である。}$$

$$Z_m \text{ の期待値 } \mu = \frac{1-p}{p}, \quad Y_n \text{ の期待値 } \lambda = \frac{1-q}{q}$$

最尤推定量は

$$\hat{\mu} = \frac{1 - \hat{p}}{\hat{p}} = \frac{\sum_{i=1}^n Z_i - \sum_{i=1}^n Y_i}{\sum_{i=1}^n Z_{i-1}}, \quad \hat{\lambda} = \frac{1 - \hat{q}}{\hat{q}} = \frac{\sum_{i=1}^n Y_i}{n} = \bar{Y}$$

幾何分布よりのシミュレーションの結果を見る：

m：世代；n：シミュレーション回数

図 2-4 は m=1000, n=1000, p=0.52, q=0.1 と設定するとき $\hat{\mu}$ と $\hat{\lambda}$ のヒストグラムである。この時 μ の理論値は 0.923 λ の理論値は 9 である。

図 2-5 は m=1000, n=500, p=0.49, q=0.1 と設定する時 $\hat{\mu}$ と $\hat{\lambda}$ のヒストグラムである。この時、 μ の理論値は 0.923 で、 λ の理論値は 9 である。

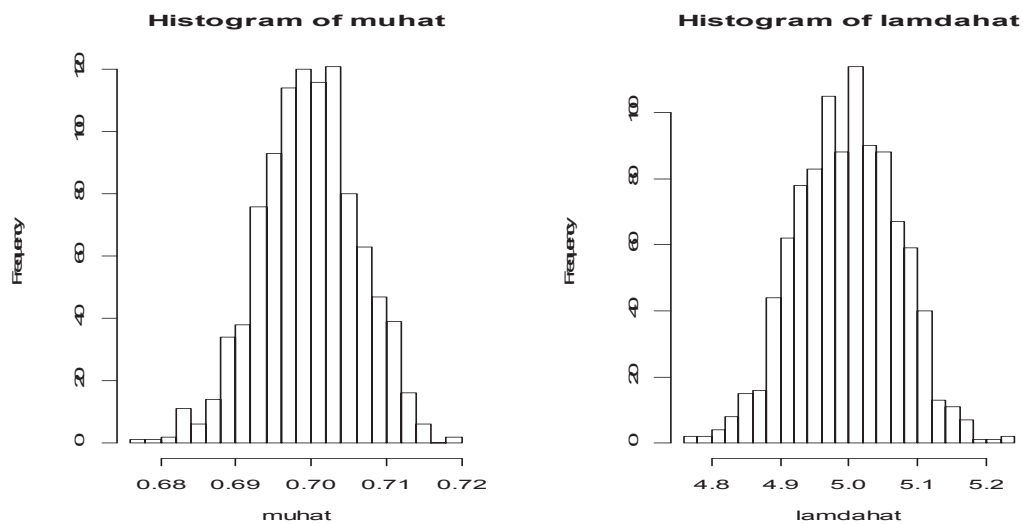
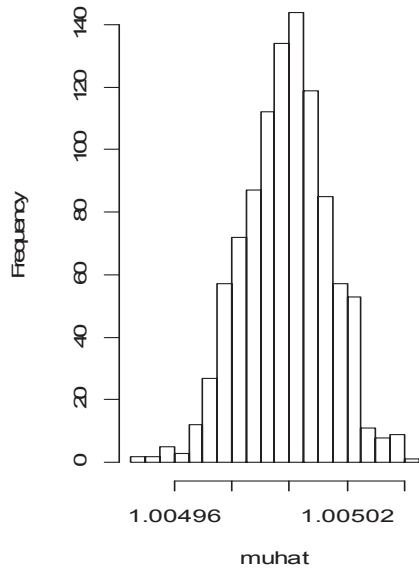
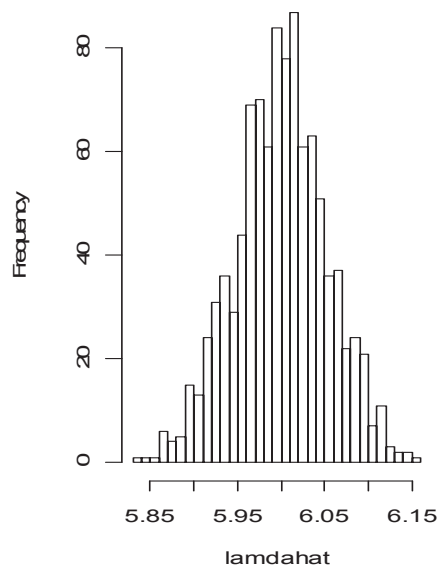


図 2- 1

Histogram of muhat

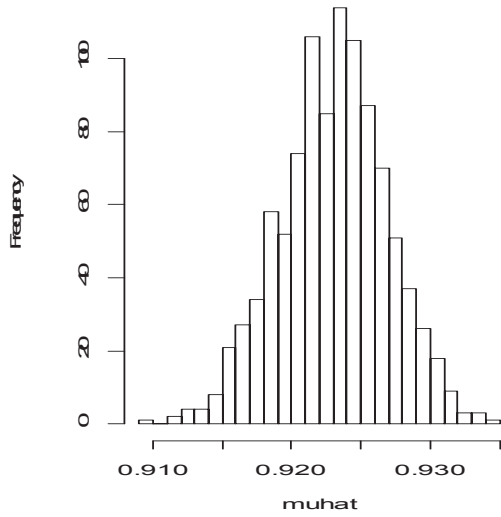


Histogram of lamdahat

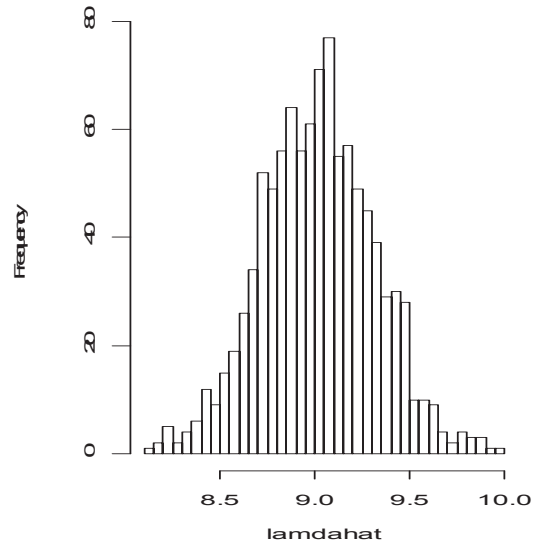


☒ 2-2

Histogram of muhat



Histogram of lamdahat



☒ 2-3

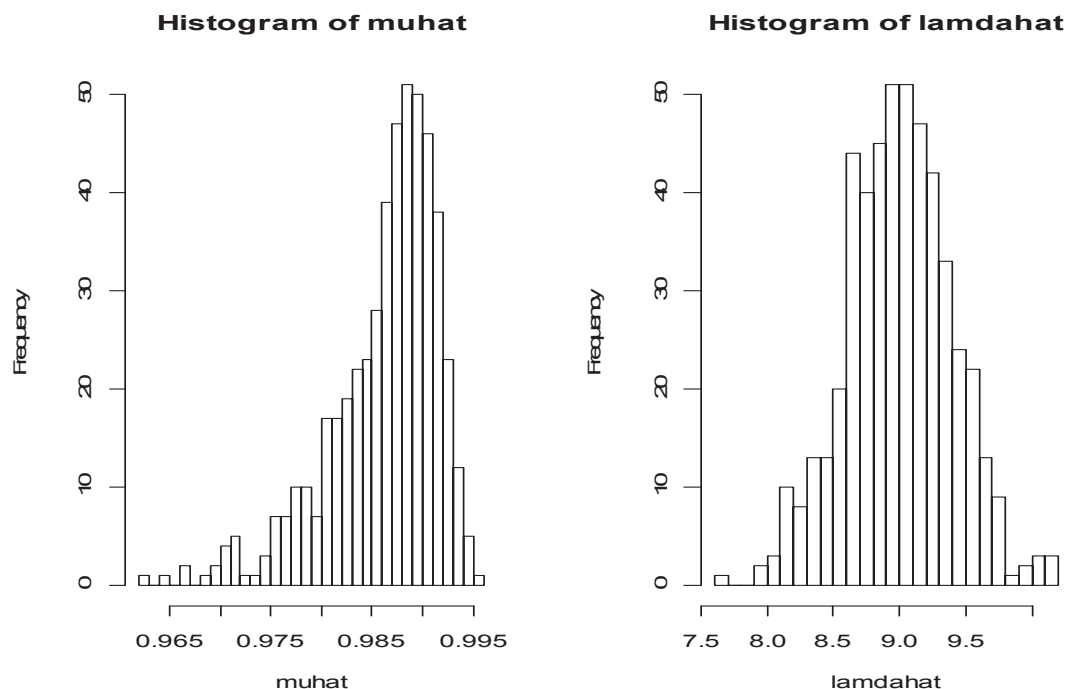


図 2-4

2.3.3 逐次確率比検定

Z_1, Z_2, \dots, Z_n が一次独立同分布に従うならば、逐次確率検定 (Sequential Probability Ratio Test, 略称 SPRT) では、一般の検定のように標本数をあらかじめ固定することをせず、オンラインデータに対応可能である。標本を得るたびに、検定仮説 H_0 と対立仮説 H_1 の尤度を算出し、その比 L がグレイゾーン (採択限界と棄却限界の間) にある限り判断を保留し、サンプリングを続ける。この範囲を超えたときに検定を終了する。本論文では、この SPRT 検定を利用する。

$$\text{停止時刻 } T = \min \left\{ t \geq 1 : \log \frac{f_1(z_1) f_1(z_2) \dots f_1(z_n)}{f_2(z_1) f_2(z_2) \dots f_2(z_n)} \notin [a, b] \right\} \quad (2.14)$$

$H_0 : \mu = \mu_0$ VS $H_1 : \mu = \mu_1$ の検定の中、検定仮説 H_0 が真なのに對立仮説 H_1 と判断される場合は第一種の誤りが発生し、確率が α である。對立仮説 H_1 が真なのに検定仮説 H_0 と判断される場合は第二種の誤りが発生し、確率が $(1 - \beta)$ である。 α は検定のサイズ (size) であるので、0 と近づけば近いほど望ましい。 β は検定力 (power) であるので、1 と接近すればするほど良い。

A. WALD (1945) (参考文献1) の論文より α 、 β と a 、 b の関係は下記の式で計算できる：

$$a = \log \frac{1-\beta}{1-\alpha}, b = \log \frac{\beta}{\alpha} \quad (2.15)$$

逐次確率比検定は幅広い分野で利用されている。実例は地震の被害情報処理、電子情報通信システム適応的話者識別[参6]などである。本論文のモデルでは、逐次確率比検定がインフルエンザの流行と抑制の課題に運用できる。たとえば、インフルエンザ流行中、患者が増えて続ける状況より、データがオンラインである。インフルエンザ流行を抑制するためにワクチンを投入する。そしたらワクチンの効果を検定する。効果があるなら使用し続ける。効果がなければ使用中止、新たな研究をしなければならない。

2.3.3.1 ポアソン逐次確率比検定の過程

$$\begin{cases} Z_0=1 \\ Z_n = \sum_{i=1}^{Z_{n-1}} Z_{ni} + Y_n \\ \xi_{ni} \sim Po(\mu) \quad \xi \text{ は } n, \text{ について独立} \\ Y_n \sim po(\lambda) \quad y \text{ は } n \text{ について独立} \end{cases} \quad \text{とのモデルに従う。}$$

Y の期待値 λ は不変であると仮定する。

$$\begin{cases} H_0 : \mu = \mu_0 \quad \text{ワクチンの効果なし} \\ H_1 : \mu = \mu_1 \quad \text{ワクチンの効果あり} \end{cases}$$

ポアソン分布よりの検定過程についての理論的な計算は具体的に次である。

$$p(z_n, y_n | z_{n-1}) \sim P_o(z_{n-1}, \mu) P_o(\lambda) = \frac{e^{-z_{n-1}\mu} (z_{n-1}\mu)^{z_n - y_n}}{(z_n - y_n)!} \times \frac{e^{-\lambda} \lambda^{y_n}}{y_n!} \text{ より}$$

$$p_0(z_n, y_n | z_{n-1}) = \frac{e^{-z_{n-1}\mu_0} (z_{n-1}\mu_0)^{z_n - y_n}}{(z_n - y_n)!} \times \frac{e^{-\lambda} \lambda^{y_n}}{y_n!}$$

$$p_1(z_n, y_n | z_{n-1}) = \frac{e^{-z_{n-1}\mu_1} (z_{n-1}\mu_1)^{z_n - y_n}}{(z_n - y_n)!} \times \frac{e^{-\lambda} \lambda^{y_n}}{y_n!}$$

$$\text{尤度比 } L = \frac{P_1(z_1, y_1, \dots, z_n, y_n | z_{n-1})}{P_0(z_1, y_1, \dots, z_n, y_n | z_{n-1})} = \prod_{i=1}^n \frac{P_1(z_i, y_i | z_{i-1})}{P_0(z_i, y_i | z_{i-1})}$$

$L < A \Rightarrow$ ワクチン投入前後変化なし $\Rightarrow H_0$ 採択 (効果なし)

$L > B \Rightarrow$ ワクチン投入前後変化あり $\Rightarrow H_1$ 採択 (効果あり)

$$L \text{ の対数尤度 } l_t = \sum_{i=1}^t \log \frac{P_1(z_i, y_i | z_{i-1})}{P_0(z_i, y_i | z_{i-1})}, \quad \log A = a, \quad \log B = b$$

$l_t < a$ の時にワクチン効果なし、 $l_t > b$ のときにワクチン効果ある。SPRT 検定の停

止時刻は $S = \min \{t \geq 1 : l_t > b, \text{ or } l_t < a\}$ と定義する。[効果がある]あるいは[効果がない]どちらか一方、検定される時刻 S でサンプリングを止める。この検定より、ワクチンの効果について早い段階で判断が可能になる。

第一種の誤り $\alpha = P_0(l_s > b)$, 第二種の誤り $1 - \beta = P_1(l_s < a)$ 。ワクチンも普通の薬も、効果があると言ってもある程度の誤差を許す。この検定では α と β の値が与えられた想定数と考えれば良い。たとえば $\alpha = 5\%$ の意味は効果ありのワクチンは 5% の人に効果なくても許す。 β は真実どおりの効果率。たとえば β は 95% であれば、このワクチンが 95% の人に効果がある。ここでは a, b の値は (2.15)

$$\text{式で決められる。つまり } a = \log \frac{1 - \beta}{1 - \alpha}, \quad b = \log \frac{\beta}{\alpha}.$$

2.3.3.2 幾何分布よりの検定過程

2.3.2.2 の推導より

$$P_1(Z_1 = z_1, Y_1 = y_1, \dots, Z_n = z_n, Y_n = y_n) = p_1^{\sum_{i=1}^n z_{i-1}} (1 - p_1)^{\sum_{i=1}^n z_i - \sum_{i=1}^n z_i} q^n (1 - q)^{\sum_{i=1}^n y_i} \cdot \prod_{i=1}^n Q_i$$

$$P_0(Z_1 = z_1, Y_1 = y_1, \dots, Z_n = z_n, Y_n = y_n) = p_0^{\sum_{i=1}^n z_{i-1}} (1 - p_0)^{\sum_{i=1}^n z_i - \sum_{i=1}^n z_i} q^n (1 - q)^{\sum_{i=1}^n y_i} \cdot \prod_{i=1}^n Q_i$$

$$\begin{aligned} L &= \frac{P_1(Z_1 = z_1, Y_1 = y_1, \dots, Z_n = z_n, Y_n = y_n)}{P_0(Z_1 = z_1, Y_1 = y_1, \dots, Z_n = z_n, Y_n = y_n)} \\ &= \left(\frac{1 - p_1}{1 - p_0} \right)^{\sum_{i=1}^n z_i - \sum_{i=1}^n y_i} \cdot \left(\frac{p_1}{p_0} \right)^{\sum_{i=1}^n z_{i-1}} \end{aligned} \quad (2.16)$$

$L < A \Rightarrow$ ワクチン投入前後変化なし $\Rightarrow H_0$ 採択 (効果なし)

$L > B \Rightarrow$ ワクチン投入前後変化あり $\Rightarrow H_1$ 採択 (効果あり)

$$L \text{ の対数尤度 } l_t = \left(\sum_{i=1}^n z_i - \sum_{i=1}^n y_i \right) \log \frac{1-p_1}{1-p_0} + \sum_{i=1}^n z_{i-1} \log \frac{p_1}{p_0}$$

$l_t < a$ の時にワクチン効果なし、 $l_t > b$ のときにワクチン効果ある。

停止時刻 $S = \min\{t \geq 1: l_t > b, \text{ or } l_t < a\}$

第一種の誤り $\alpha = P_0(l_s > b)$ 、第二種の誤り $1 - \beta = P_1(l_s < a)$ 。

検定の判断原則と a、b 決まり方法は分布に限らず、普遍的に適応である。

2.3.4 逐次確率比検定のシミュレーションと結果分析

2.3.4.1 ポアソン分布の場合

m は世代、n はシミュレーションの回数である。 $\alpha = 0.05$, $\beta = 0.75$ と想定する。

	m	n	λ	μ_0	μ_1	α	β
第1回	1000	1000	5	1.005	0.995	0.046	0.778
第2回	1000	2000	5	1.01	0.995	0.041	0.779
第3回	2000	1000	5	1.005	0.999	0.047	0.786

見方：

停止時刻 (StoppingTime) の最大値 = 212

停止時刻の期待値 (MeanTime) = 70.076

停止時刻の見方：> StoppingTime[1:5]

[1] 40 96 70 98 33

シミュレーションを 1000 回行うと、停止時刻が一番多いのが停止時刻 212 個がある。停止時刻の期待値は 70.076 であるから見ると、各回のシミュレーションの停止時刻の平均個数は 70.076 である。

停止時刻の見方のところで StoppingTime の 1 ~ 5 の元素を表示させ、第一回のシミュレーションの停止時刻は第 40 世代に出現、第二回シミュレーションの停止時刻は第 96 世代に出現するという意味を表わす。

シミュレーションの結果より α 、 β は想定値のまわりにあるということがわかる。

第一種誤りの確率が 0.05 と想定する時に、実は 0.046 である。つまり平均停止時刻 (70.076) の中第一種の誤りが 3.2 回 ($70.076 * 0.046$) ぐらい発生する。第二種誤りの確率が $1 - \beta$ であり、実の確率は 0.222 である。つまり平均停止時

刻の中第二種類の誤りが 15.5 回が発生する。β は検定力であり、0.778 の実の値がワクチンの効果検定では有効率が 0.778 という意味である。以下は比較と検証のため、特性値を入れ替え、数回のシミュレーションの結果を載る。

	停止時刻 (StoppingTime) の最大値	停止時刻の期待値 (MeanTime)
第 1 回	212	70.076
第 2 回	181	83.0635
第 3 回	466	164.448

2.3.4.2 幾何分布の場合

$\alpha=0.05$, $\beta=0.75$ と想定する。

	m	n	p	q	α	β
第 1 回	1000	1000	0.49	0.1	0.046	0.769
第 2 回	1000	2000	0.52	0.1	0.042	0.82
第 3 回	2000	1000	0.4	0.1	0.037	0.882

	停止時刻 (StoppingTime) の最大値	停止時刻の期待値 (MeanTime)
第 1 回	38	12.809
第 2 回	53	16.31
第 3 回	229	56.61

以上の二つの分布に対してシミュレーションを行う結果、 α (size) , β (power) 想定値と実際値が大体上一致である。時間の長さは停止時刻の回数に影響を与えることをわかる。 μ の値が 1 と近くにすればするほど、 α と β の値が希望する値と値が近づいてくる。

第3章 臨界値に対する分枝過程の逐次検定 (Sequential Tests for The Criticality of Branching Processes)

3.1 モデルの基本

Z_n を移民項のある分枝過程の第 n 世代数として定義される。

$$Z_n = \sum_{i=1}^{Z_{n-1}} \xi_{n,i} + Y_n, \quad n=1, 2, \dots, \quad i=1, 2, \dots \quad (3.1)$$

$\xi_{n,i}$ が第 $n-1$ 世代第 i 番目個体の子孫である。 Y_n が第 n 世代の移民の数である。

$\{\xi_{n,i}\}$ と $\{Y_n\}$ は独立であり、それぞれ i. i. d. で特性値 (m, σ^2) と (λ, b^2) を持つ。

$\{Z_n\}$ と $\{Y_n\}$ が観測可測されると仮定する。初期値 $Z_0 = 1$ とする。

分枝過程による実際問題の研究例がたくさんある。

生態学分野では、京都大学佐藤葉子・瀬野裕美先生より書かれた本「姓の継承と絶滅の数理生態学—Galton-Watson 分枝過程によるモデル解析」は一つの応用例である。数理生態学の観点を取りいれて、応用数理分野である確率過程のなかの分枝過程を応用した数理モデルの解析によって、姓の存続性や絶滅性、さらに、多様性に関する数理的な考察を試みるための、また、試みた道筋を、できるだけ丁寧にまとめようとした。数理生態学の観点の導入については、とりわけ、姓の存続性の検討に関して、出生率や出生性比を陽に含む数理モデルの解析によって考察している点に顕著である。日本の姓の多様性や存続性、絶滅性に関する数理的な考察が詳しく、計量的な結論が出た。姓の問題は、社会的あるいは文化的な要素を抜きには考察できないことは当然だが、統計的な扱いとして分枝過程によるアプローチは有意義な面がある。

米フェイスブック会員数の増え方は分枝過程に従う。

プリンストン大学博士課程に在籍する Joshua Spechler さんと John Cannarella さんが執筆した論文ではフェイスブックの滅亡が迫っていると主張し、研究論文がインターネット上で物議を醸している。ソーシャルネットワーキングサービス (SNS) 大手のフェイスブックのようなソーシャルメディアは通常、伝染病と同じ成長と衰退のパターンをたどると論じている。2人は世界最大の SNS であるフェイスブックが向こう3年以内に事実上消滅すると結論付けた。

論文には、2008年に最盛期を迎えた後で急速に衰え、11年までに無きに等しい存在になった SNS の米マイスペースの例が挙げられた。また、グーグルの検索クエリーのデータを利用して人気度を測り、フェイスブックは12年に人気が最高潮に達し、既に衰退し始めていると述べた。同年から「検索頻度の下降トレンド」が見られるとしている。

3.1.1 Basawa の研究について

The Annals of Statistics(1991)に発表された「Sequential Estimation Branching Processes with Immigration」では移民項のある分枝過程の逐次推定について重要な結論がある。

m と σ^2 の推定値が 3.2 と 3.3 で定義された。(参考文献[10])

$$\hat{m}_n = \frac{\sum_1^n (Z_i - Y_i)}{\sum_1^n Z_{i-1}}, \quad (3.2)$$

$$\hat{\sigma}_n^2 = \frac{\sum_1^n [(Z_i - \hat{m}_n Z_{i-1} - Y_i)^2 / (1 + Z_{i-1})^2]}{\sum_1^n [Z_{i-1} / (1 + Z_{i-1})^2]}. \quad (3.3)$$

(3.2) は (3.1) より簡単で計算できる。(3.3) は σ の一致推定量である。
この時 $n \rightarrow \infty$ の場合 Basawa(1991) より

$$\left(\sum_{i=1}^n Z_{i-1} / \hat{\sigma}_n^2 \right)^{1/2} (\hat{m}_n - m) \Rightarrow \begin{cases} N(0,1) & \text{if } m < 1 \\ \frac{X(1) - \lambda}{\sqrt{\int_0^1 X(s) ds}} & \text{if } m = 1 \end{cases} \quad (3.4)$$

$n \rightarrow \infty, X_t$ は Diffusion Process (3) 式の $\Delta = 1$ の場合である。

3.2 Diffusion Approximation について

3.2.1. Diffusion Approximation の定理

逐次解析に関する研究には、誤りの確率と停止時間の期待値と分散など特性を計算することが重要である。多くの研究者が、非線形更新理論が重要な役割を果たして、特性値の評価を研究した。しかし非線形再生理論は Wald の研究結果に強く依存し、i. i. d. で観測される時のみ有効である。この問題を克服するために、我々は、local alternatives 仮説を想定した Diffusion Approximation を使用する。

マルコフは、物理学、化学、生物学で処理し、多くの場合一般化拡散と見なされているため、Diffusion Approximation 数学理論も必要である。Kurtz and Protter(1991)では、マルコフ過程の Diffusion Approximation 問題については、次の結論がある。

各 n に対して、 X_k^n は filtration マルコフ過程。

$\mu(x)$: the drift coefficient

$$(1) \quad E \left[X_k^n - X_{k-1}^n \mid F_{k-1}^n \right] = \mu \left(X_{k-1}^n \right) \frac{1}{n} + o \left(\frac{1}{n} \right)$$

$\sigma(x) > 0$: the diffusion coefficient

$$(2) \quad E \left[\left(X_k^n - X_{k-1}^n \right)^2 \mid F_{k-1}^n \right] = \sigma^2 \left(X_{k-1}^n \right) \frac{1}{n} + o \left(\frac{1}{n} \right)$$

diffusion approximation

$$(3) \quad X_{[nt]}^n \Rightarrow X_t = X_0 + \int_0^t \mu(X_s) ds + \int_0^t \sigma(X_s) dW_s$$

3.2.2 分枝過程における Diffusion Approximation

2.2 の内容より、分枝過程がマルコフ過程である。3.2.1 の定理を利用し、分枝過程における Diffusion Approximation を計算できる。この結論は Nagai (2007) [参考文献 4] に載っている。

$X_k^n = Z_k / n$ にして、分枝過程の拡散近似具体的には次三つの式となる。

$$(1)' \quad E \left[X_k^n - X_{k-1}^n \mid F_{k-1}^n \right] = (\lambda - \Delta X_{k-1}^n) \frac{1}{n}$$

$$(2)' \quad E \left[\left(X_k^n - X_{k-1}^n \right)^2 \mid F_{k-1}^n \right] = \sigma^2 X_{k-1}^n \frac{1}{n} + o \left(\frac{1}{n} \right)$$

(3)' この時 $X_{[nt]}^n \Rightarrow X_t$, X_t は CIR (Cox=Ingersoll=Ross) 過程である:

$$dX_t = (\lambda - \Delta X_t) dt + \sigma \sqrt{X_t} dW_t$$

この結論を証明するため local alternative の数学方法を行う必要である。

$H_0: m = 1$ vs $H_1: m = 1 - \Delta/n$ で考える。次の書換を行う:

$$Z_k = \sum_{i=1}^{Z_{k-1}} \xi_{k,i} + Y_k = \lambda + m Z_{k-1} + \varepsilon_k, \quad (3.5)$$

$$\varepsilon_k = \sum_{i=1}^{Z_{k-1}} (\xi_{k,i} - m) + Y_k - \lambda \text{ である。}$$

この時 $E(\varepsilon_k \mid F_{k-1}) = 0$, $E(\varepsilon_k^2 \mid F_{k-1}) = \sigma^2 Z_{k-1} + b^2$.

3.1 より $\{\xi_{n,i}\}$ と $\{Y_n\}$ は独立であり、i. i. d. でそれぞれ特性値 (m, σ^2) と (λ, b^2) を持つ。そして、分枝過程のマルコフ性を用いて、 ε_k の条件付期待値と分散が簡単に計算できる。具体的には

$$\begin{aligned}
E(\varepsilon_k | F_{k-1}) &= E\left(\sum_{i=1}^{Z_{k-1}} (\xi_{k,i} - m) + Y_k - \lambda \mid F_{k-1}\right) \\
&= E(\xi_{k,1} + \xi_{k,2} + \dots + \xi_{k,Z_{k-1}} - Z_{k-1}m + Y_k - \lambda \mid F_{k-1}) \\
&= Z_{k-1}E(\xi_{k,i}) - Z_{k-1}m + E(Y_k) - \lambda \\
&= 0 \\
E(\varepsilon_k^2 | F_{k-1}) &= V(\varepsilon_k | F_{k-1}) + E^2(\varepsilon_k | F_{k-1}) \\
&= E\left[\left(\sum_{i=1}^{Z_{k-1}} (\xi_{k,i} - m) + Y_k - \lambda\right)^2 \mid F_{k-1}\right] \\
&= \sigma^2 Z_{k-1} + b^2 \quad \text{となる。}
\end{aligned}$$

$X_k^n = Z_k/n$ に対し、diffusion approximation を用いて(1)' ~ (3)' の結論が出る。(1)' は(3.5)式と対立仮説を合わせて計算すると自然に証明できる。(2)' はチェビシエフ不等式などの知識が必要である。(3)' 関数替代より diffusion approximation から CIR 過程が生成する。具体的には

$$\begin{aligned}
&(1)' \\
&E[X_k^n - X_{k-1}^n \mid F_{k-1}^n] \\
&= \frac{1}{n} E[Z_k - Z_{k-1} \mid F_{k-1}^n] = \frac{1}{n} \left(E\left(\lambda + mZ_{k-1} + \varepsilon_k - Z_{k-1}F_{k-1}^n\right) \right) \\
&= \frac{1}{n} (\lambda - (1-m)Z_{k-1}) = \frac{1}{n} \left(\lambda - \frac{\Delta}{n} Z_{k-1} \right) \\
&= \frac{1}{n} (\lambda - \Delta X_{k-1}^n)
\end{aligned}$$

$$\begin{aligned}
&= \frac{1}{n^2} \text{Var}(\lambda + mZ_{k-1} + \varepsilon_k - Z_{k-1} | F_{k-1}^n) + E^2(X_k^n - X_{k-1}^n | F_{k-1}^n) \\
&= \frac{Z_{k-1}\sigma^2 + b^2}{n^2} + \frac{(\lambda - \Delta X_{k-1}^n)^2}{n^2} \\
&= \frac{\sigma^2 X_{k-1}^n}{n} + \frac{b^2 + \lambda^2 - 2\lambda\Delta X_{k-1}^n + \Delta^2(X_{k-1}^n)^2}{n^2} \\
&= \frac{\sigma^2 X_{k-1}^n}{n} + o\left(\frac{1}{n}\right) \\
&(\text{注} : \frac{b^2 + \lambda^2 - 2\lambda\Delta X_{k-1}^n + \Delta^2(X_{k-1}^n)^2}{n^2} = o\left(\frac{1}{n}\right))
\end{aligned}$$

注の証明を詳細に書いておく。

$$Z_n = \sum_{i=1}^{Z_{n-1}} \xi_{n,i} + Y_n, \quad Z_0 = 1$$

$$(a) \frac{X_{k-1}^n}{n^2} = o\left(\frac{1}{n}\right) \text{を示す。}$$

$$E[Z_n | F_{n-1}] = E\left[\sum_{i=1}^{Z_{n-1}} \xi_{n,i} + Y_n | F_{n-1}\right] = E\left[\sum_{i=1}^{Z_{n-1}} \xi_{n,i} + Y_n | F_{n-1}\right]$$

この式の両辺同時期待値を取ると

$$E(Z_n) = E(Z_{n-1}) + m\lambda$$

$m=1$ の場合で考える。

$$E(Z_n) = E(Z_{n-1}) + \lambda = E(Z_n) = E(Z_{n-2}) + 2\lambda = \dots = E(Z_0) + n\lambda$$

$$\text{よって } E(Z_n) = 1 + n\lambda$$

$$P\left(\frac{Z_{k-1}}{n^2} > \varepsilon\right) \leq \frac{E(Z_{k-1})}{n^2 \varepsilon}$$

$$\frac{X_{k-1}^n}{n^2} = \frac{Z_{k-1}}{n^3} \text{であるので } \lim_{n \rightarrow \infty} \frac{X_{k-1}^n}{n^2} = \lim_{n \rightarrow \infty} \frac{Z_{k-1}}{n^3} = \lim_{n \rightarrow \infty} \frac{Z_{k-1}}{n} \cdot \frac{1}{n^2} = 0 \text{つまり } \frac{X_{k-1}^n}{n^2} = o\left(\frac{1}{n}\right)$$

(b) $\frac{(X_{k-1}^n)^2}{n^2} = o\left(\frac{1}{n}\right)$ を示す。

$$V(Z_n | F_{n-1}) = E\left[\left(\sum_{i=1}^{Z_{n-1}} (\xi_{n,i} - m) + (Y_i - \lambda)\right)^2 \mid F_{n-1}\right]$$

$$= Z_{n-1}\sigma^2 + b^2$$

$m=1$ の場合を考える。前の証明より $E(Z_n) = E(Z_{n-1}) + m\lambda$ である。

$$E(Z_n^2) = E(Z_n^2 | F_{n-1}) = E(V(Z_n | F_{n-1}) + E^2(Z_n | F_{n-1}))$$

$$= E((Z_{n-1})\sigma^2 + b^2 + (Z_{n-1} + \lambda)^2)$$

$$= E(Z_{n-1})\sigma^2 + b^2 + E(Z_{n-1}^2) + 2\lambda E(Z_{n-1}) + \lambda^2$$

$$E(Z_n^2) - E(Z_{n-1}^2) = (n-1)(2\lambda^2 + \lambda\sigma^2) + \sigma^2 + b^2 + \lambda^2 + 2\lambda$$

$$\text{従って } E(Z_n^2) = \frac{2\lambda^2 + \lambda\sigma^2}{2}n(n-1) + (\sigma^2 + b^2 + \lambda^2 + 2\lambda)n$$

チェビシエフの不等式 $\left\{\frac{P(Z_{k-1}^2)}{n^3} > \varepsilon\right\} \leq \frac{E(Z_{n-1}^2)}{n^3\varepsilon}$ を用いて、

$$\frac{(X_{k-1}^n)^2}{n^2} = \frac{Z_{k-1}^2}{n^4} \text{ なるので } \lim_{n \rightarrow \infty} \frac{\frac{Z_{k-1}^2}{n^4}}{\frac{1}{n}} = 0 \text{ つまり } \frac{(X_{k-1}^n)^2}{n^2} = o\left(\frac{1}{n}\right).$$

(3)'

(1)の $\mu(X_{k-1}^n)$ を(1)'の $(\lambda - \Delta X_{k-1})$ で、

(2)の $\sigma^2(X_{k-1}^n)$ を(2)'の $\sigma^2 X_{k-1}^n$ で、(3)の $\sigma(X_s)$ を $\sigma\sqrt{X_s}$

で入れ替える。

diffusion approximation $X_t = X_0 + \int_0^t \mu(X_s)ds + \int_0^t \sigma(X_s)dW_s$ に対して

両辺微分を取ると、 $X_{[m]}^n$ がCIR過程 X_t に収束する：

$$dX_t = (\lambda - \Delta X_t)dt + \sigma\sqrt{X_t}dW_t \tag{3.6}$$

3.3 尤度関数と fisher 情報量

3.3.1 尤度関数

第2章では Power series function について詳しく述べておいた。ここでは Power series function を用いて尤度関数を作る。

Power series $g(\theta) = \sum_{n=0}^{\infty} a_n \cdot \theta^n$ である時 Power series function が $f(n) = \frac{a_n \theta^n}{g(\theta)}$ で

定義される。

(3.1) に満たす分枝過程に対して Power series function を設定する：

$$P(\xi_{n,j} = j) = \frac{a_j \theta^j}{A(\theta)}, P(Y_n = k) = \frac{b_k \phi^k}{B(\phi)}$$

2.2の内容に従って $\xi_{n,j}$ と Y_n の期待値と分散を表示することができる。

$$E(\xi_{n,j}) = \frac{\theta A'(\theta)}{A(\theta)}, V(\xi_{n,j}) = \left(\frac{\partial}{\partial m} \log \theta \right)^{-1} \quad (3.7)$$

$$E(Y_n) = \frac{k B'(\phi)}{B(\phi)}, V(Y_n) = \left(\frac{\partial}{\partial m} \log \phi \right)^{-1} \quad (3.8)$$

fisher 情報量を得るため、尤度関数 $L = P(Z_0, Y_0, Z_1, Y_1 \cdots Z_n, Y_n)$ を求めなければならぬ。 $m_0 = 1$ VS $m_\Delta^n = 1 - \frac{\Delta}{n}$ 仮定の下で、対数尤度比を作る。

$$P(\xi_{n,j} = j) = \frac{a_j \theta^j}{A(\theta)} \text{ であるので } \sum_{j=1}^{\infty} \frac{a_j \cdot \theta^j}{A(\theta)} = 1$$

左辺の分母を右辺に移動すれば $A(\theta) = \sum_{j=0}^{\infty} a_j \cdot \theta^j$

$P(\xi_{n,j} = j)$ がわかるため λ^ξ の期待値を求めることができる。

$$E[\lambda^\xi] = \sum_{j=1}^{\infty} \lambda^j \cdot \frac{a_j \cdot \theta^j}{A(\theta)} = \frac{A(\lambda \cdot \theta)}{A(\theta)}$$

この式に従って、 ξ_n が *i.i.d* であるため

$$E[\lambda^{\xi_1 + \xi_2 + \dots + \xi_n}] = \prod_{i=1}^n E[\lambda^{\xi_i}] = \frac{A^n(\lambda \cdot \theta)}{A^n(\theta)}$$

$$P(\xi_1 + \xi_2 + \dots + \xi_n = z) = \frac{a_z^{(n)} \cdot \theta^z}{A^n(\theta)}$$

(3.1) $\xi_1 + \xi_2 + \dots + \xi_n = Z_i - Y_i$ であることを利用して、

$$\begin{aligned}
L_n &= P\{Z_0 = z_0, Y_1 = y_1, \dots, Y_n = y_n, Z_n = z_n\} \\
&= P_{n|n-1}(Z_n, Y_n | Z_{n-1}) P_{n-1|n-2}(Z_{n-1}, Y_{n-1} | Z_{n-2}) \dots P_{0|1}(Z_1, Y_1 | Z_0) \\
&= \prod_i^n \frac{\theta^{z_i - y_i} a_{z_i - y_i}^{(z_i - 1)} b_{y_i} \phi^{y_i}}{A^{z_i - 1}(\theta) B(\phi)} \\
l_n = \log L_n &= \sum_{i=1}^n \left\{ \log a_{z_i - y_i}^{(z_i - 1)} + (z_i - y_i) \log \theta - z_{i-1} \log A(\theta) + \log b_{y_i} + y_i \log \phi - \log B(\phi) \right\}
\end{aligned}$$

次のステップは $\begin{cases} H_0 : m = 1 \\ H_1 : m = 1 - \frac{\Delta}{n} \end{cases}$ の基で尤度比関数をもとめる。単純な数学計算で

あるため、読みにくいかもしれない。尤度比を計算するため、次の微分をしておく。

- $\sigma^2 = \left(\frac{\partial}{\partial m} \log \theta \right)^{-1}$ より、 $\sigma^2 = \frac{1}{\theta} \frac{\partial m}{\partial \theta}$
- $m = \frac{\theta \cdot A'(\theta)}{A(\theta)} = \frac{\partial}{\partial (\log \theta)} \log A(\theta)$
- $\frac{\partial}{\partial m} \log A(\theta) = \frac{\partial}{\partial (\log \theta)} \log A(\theta) \cdot \frac{\partial}{\partial m} \log \theta = \frac{m}{\sigma^2}$
- $\frac{\partial^2 \log \theta}{\partial m^2} = \frac{\partial}{\partial m} \left(\frac{\partial \log \theta}{\partial m} \right) = \frac{\partial}{\partial m} \left(\frac{1}{\theta} \frac{\partial \theta}{\partial m} \right) = -\frac{1}{\theta^2} \left(\frac{\partial \theta}{\partial m} \right)^2 + \frac{1}{\theta} \frac{\partial^2 \theta}{\partial m^2}$
- $\frac{\partial^2 \log A(\theta)}{\partial m^2} = \frac{\partial}{\partial m} \left(\frac{\partial \log A(\theta)}{\partial (\log \theta)} \frac{\partial \log \theta}{\partial m} \right) = \frac{\partial}{\partial m} \left(m \cdot \frac{\partial \log \theta}{\partial m} \right) = \frac{\partial \log \theta}{\partial m} + m \frac{\partial^2 \log \theta}{\partial m^2}$

帰無仮説と対立仮説それぞれの尤度関数を取り、差をつくる。

$$l_{H_1} - l_{H_0} = (\log(\theta_1) - \log(\theta_0)) \sum_{i=1}^n (z_i - y_i) - (\log A(\theta_1) - \log A(\theta_0)) \sum_{i=1}^n Z_{i-1}$$

この式を計算するため、 $\sum_{i=1}^n (z_i - y_i)$ と $\sum_{i=1}^n Z_{i-1}$ のそれぞれの係数を求める必要がある。

$\log(\theta_1) - \log(\theta_0)$ を (*) 式、そして $\log A(\theta_1) - \log A(\theta_0)$ を (**) 式で記号する。

テイラー展開公式

$f(\beta) = f(\alpha) + f'(\alpha)(\beta - \alpha) + \frac{f''(\alpha)}{2!}(\beta - \alpha)^2 + o((\beta - \alpha)^2)$ を用い

$f = \log \theta$ を $\alpha=1, \beta=1-\frac{\Delta}{n}$ でテイラー展開する。

$$\log \theta \Big|_{m=1-\frac{\Delta}{n}} = \log \theta \Big|_{m=1} + \frac{\partial \log \theta}{\partial m} \Big|_{m=1} \cdot \left(-\frac{\Delta}{n}\right) + \frac{1}{2} \frac{\partial^2 \log \theta}{\partial m^2} \Big|_{m=1} \cdot \left(-\frac{\Delta}{n}\right)^2 + o\left(\frac{1}{n^2}\right)$$

よって

$$* = \frac{\partial \log \theta}{\partial m} \Big|_{m=1} \cdot \left(-\frac{\Delta}{n}\right) + \frac{1}{2} \frac{\partial^2 \log \theta}{\partial m^2} \Big|_{m=1} \cdot \left(-\frac{\Delta}{n}\right)^2 + o\left(\frac{1}{n^2}\right)$$

$$** = \frac{\partial \log A(\theta)}{\partial m} \Big|_{m=1} \cdot \left(-\frac{\Delta}{n}\right) + \frac{1}{2} \frac{\partial^2 \log A(\theta)}{\partial m^2} \Big|_{m=1} \cdot \left(-\frac{\Delta}{n}\right)^2 + o\left(\frac{1}{n^2}\right)$$

* と ** を $l_{H_1} - l_{H_0}$ に代入し

$$l_{H_1} - l_{H_0} = -\frac{\Delta}{n\sigma^2} \sum_{i=1}^n (Z_i - Y_i - Z_{i-1}) - \frac{\Delta}{2n^2\sigma^2} \sum_{i=1}^n Z_{i-1} + \frac{\Delta^2}{2n^2} \frac{\partial^2 \log \theta}{\partial m^2} \sum_{i=1}^n (Z_i - Y_i - Z_{i-1}) + o\left(\frac{1}{n^2}\right)$$

$$\Rightarrow \log \frac{L_n(m_\Delta^n)}{L_n(m_0)} = -\frac{\Delta}{n\sigma^2} \sum_{i=1}^n (Z_i - Y_i - Z_{i-1}) - \frac{\Delta^2}{2n^2\sigma^2} \sum_{i=1}^n Z_{i-1} + o\left(\frac{1}{n^2}\right) \quad (3.9)$$

逐次検定を行うため (3.9) の対数尤度比を用意する。尤度に関する初期の考察はデンマークの数学者トルバルド・ティエレ (Thorvald Nicolai Thiele) による 1889 年の著書にみられる。尤度についての完全な考察が現れた最初の論文は、ロナルド・フィッシャーによる 1922 年の "On the mathematical foundations of theoretical statistics" である。ここでフィッシャーはまた「最尤法」

(method of maximum likelihood) の語を初めて使われていた。フィッシャーは統計学的推計の基礎として事後確率を用いることに反対し、代わりに尤度関数に基づく推計を提案している。

3.3.2 Fisher 情報量

Fisher 情報量は、パラメータを持つ確率モデルがあったとき、与えられたデータ (情報) を用いて、どこまで真のパラメータを推定できるかを表現したものである。つまり、パラメータを正確に推定できるデータが、大きな情報量を持つと考える。この考え方を提唱したのは、遺伝学者で統計学者の Sir Ronald Aylmer Fisher (1890 年 2 月 17 日-1962 年 7 月 29 日)。20 世紀前半に Fisher によって考案された、統計学的な推定理論において重要な役割を担う Fisher 情報行列は、幾

何学における Riemann 計量であることは、統計学者 C. Rao が 1945 年に「Fisher 情報量を Riemann 計量として考察することは重要ではないか」との指摘から明らかになった。

最尤推定においては、Fisher 情報行列が重要な役割を演じる。一般に、データ Z がパラメータ $\theta_1, \dots, \theta_M$ をもつ密度関数 0 をもつ分布に従うとき、

$$F_{ij} = -E \left(\frac{\partial^2}{\partial \theta_i \partial \theta_j} \log f(z, \theta_1, \dots, \theta_M) \right)$$

を Fisher 情報量と呼び、行列 $F = [F_{ij}]$ を

Fisher 情報行列という。Fisher 情報量は不変推定量の分散と密接に関係している。この理論に従って、尤度関数 L_n が m について二回微分すると、分枝過程モデルの Fisher 情報量を得る。この結論は Nagai (2007) [参考文献 4] に載っている。

$$\begin{aligned} \frac{\partial L_n}{\partial m} &= \frac{\partial \log \theta}{\partial m} \sum_{i=1}^n (z_i - y_i) - \frac{\partial \log A(\theta)}{\partial m} \sum_{i=1}^n z_{i-1} \\ &= \frac{1}{\sigma^2} \sum_{i=1}^n (z_i - y_i) - \frac{m}{\sigma^2} \sum_{i=1}^n z_{i-1} \end{aligned}$$

$$\text{Fisher 情報量は } \frac{\partial^2 L_n}{\partial m^2} = \sum_{i=1}^n z_{i-1} / \sigma^2 \text{ である。} \quad (3.10)$$

3.4 停止時刻 (Stopping Time) と逐次検定 (Sequential Test)

3.4.1 Stopping Time の定義

Sequential Test では Fisher 情報量が十分大きくなる時点でサンプリングが止まる。まず Fisher 情報量より Stopping Time を定義する。

Sriam Basawa and Huggins (1991) [参考文献 10]

$$T_c = \inf \left\{ n \geq 0; \sum_{i=1}^n Z_{i-1} / \hat{\sigma}_n^2 \geq c^2 \right\}, c \rightarrow \infty \quad (3.11)$$

3.4.2 Nagai (2007) 重要な定理

[参考文献 4] にて次二つ重要な定理を主張した。

定理 1

帰無仮説 $H_0 : m=1$ の下で

3.2.1 の diffusion approximation で $\Delta = 0$.

逐次検定統計量 T_c が次の分布に従う：

$$\left(c(\hat{m}_{T_c} - 1) \quad \frac{T_c}{c} \right) \rightarrow \left(B_1 \quad \int_0^1 \frac{1}{\sigma \rho_s} ds \right)$$

$c \rightarrow \infty, B_t$ が標準ブラウン運動 $t \in [0, \infty)$

ρ_t がBessel Process, 初期値 $\rho_0 = 0$.

$$d\rho_t = \frac{\lambda}{\sigma^2 \rho_t} dt + dB_t$$

定理 2

対立仮説 $H_1 : m = 1 - \Delta / c$ の下で

逐次検定統計量 T_c が次の分布に従う :

$$\left(c(\hat{m}_{T_c} - 1) \quad \frac{T_c}{c} \right) \rightarrow \left(-\Delta + B_1 \quad \int_0^1 \frac{1}{\sigma \rho_s^\Delta} ds \right)$$

$c \rightarrow \infty, B_t$ が標準ブラウン運動 $t \in [0, \infty)$

ρ_t がBessel Processであり, $\rho_0 = 0$.

$$d\rho_t^\Delta = \left(-\Delta + \frac{\lambda}{\sigma^2 \rho_t} \right) dt + dB_t$$

3.4.3 定理の証明

$T_c = \inf \left\{ n \geq 0; \sum_{i=1}^n Z_{i-1} / \hat{\sigma}_n^2 \geq c^2 \right\}, c \rightarrow \infty$ である。 $t = [ct]$ で書き換え :

$$T_c = \inf \left\{ ct \geq 0; \frac{1}{c^2} \sum_{i=1}^{[ct]} Z_{i-1} / \hat{\sigma}_n^2 \geq 1 \right\}, c \rightarrow \infty$$

$$= c \inf \left\{ t : \frac{1}{\hat{\sigma}_n^2} \sum_{i=1}^{[ct]} \frac{Z_{i-1}}{c} \cdot \frac{1}{c} \geq 1 \right\}$$

3.2.1 より diffusion approximation では $X_k^n = Z_k / n, X_{[ct]}^c \rightarrow X_t$

(Kurtz and Protter(1991)).

まず $c(\hat{m}_{T_c} - 1) \sim B_1$ を証明する。

(3.2) 式で求めた \hat{m}_{T_c} の値を利用し、

$$c(\hat{m}_{T_c} - 1) = \frac{\frac{1}{c} \sum_{i=1}^{T_c} (Z_i - Y_i - Z_{i-1})}{\frac{1}{c^2} \sum_{i=1}^n Z_{i-1}}. \quad (3.12)$$

この式の分布を求めるため、分子と分母を分けて考えて行こう。

$$\text{分母} : \frac{1}{c^2} \sum_{i=1}^n Z_{i-1} = \frac{1}{c} \sum_{i=1}^{[ct]} X_{i-1}^c = \int_0^{T_c} X_s ds$$

(3.11)の定義より $\int_0^{T_c} X_s ds = \sigma^2$ 。このため(3.12)の分母は σ^2 に収束する。

$$\begin{aligned} \text{分子} &: \frac{1}{c} \sum_{i=1}^{T_c} (Z_i - Y_i - Z_{i-1}) = \frac{1}{c} \sum_{i=1}^{T_c} (Z_i - Z_{i-1}) - \frac{1}{c} \sum_{i=1}^{T_c} Y_i \\ &= \frac{1}{c} (z_1 - z_0 + z_2 - z_1 + \dots + z_{T_c} - z_{T_c-1}) - \frac{1}{c} \sum_{i=1}^{T_c} Y_i \\ &= \frac{1}{c} (z_{T_c} - 1) - \frac{1}{c} \sum_{i=1}^{T_c} Y_i \end{aligned}$$

分子が三つ部分を分けて考える。

- $-\frac{1}{c} : c \rightarrow \infty$ なので $-\frac{1}{c} \rightarrow 0$ 。
- $-\frac{1}{c} \sum_{i=1}^{T_c} Y_i$: 大数法則より $-\frac{1}{c} \sum_{i=1}^{T_c} Y_i = -\frac{1}{c} \cdot T_c \cdot \lambda = -U\lambda$
- $\frac{1}{c} (z_{T_c})$: $\frac{z_{T_c}}{c} = X_{T_c}^c = X_{\left[\frac{T_c}{c}\right]}^c = X_U$ (3.13)

これから (3.6) CIR 過程を便利で利用できる。

$dX_t = (\lambda - \Delta X_t) dt + \sigma \sqrt{X_t} dW_t$ 両辺同時に積分を取ると

$$X_t = \int_0^t (\lambda - \Delta X_s) ds + \sigma \int_0^t \sqrt{X_s} dW_s \quad (3.14)$$

U で書き換え :

$$\begin{aligned} X_U &= \int_0^U (\lambda - \Delta X_s) ds + \sigma \int_0^U \sqrt{X_s} dW_s \\ &= \lambda U - \Delta \int_0^U X_s ds + \sigma \int_0^U \sqrt{X_s} dW_s \\ &= \lambda U - \Delta \sigma^2 + \sigma \int_0^U \sqrt{X_s} dW_s \end{aligned} \quad (3.15)$$

$\sigma \int_0^U \sqrt{X_s} dW_s$ を計算するため、DDS定理を導入する。

DDS 定理 :

M_t が $\{F_t\}$ 上の連続ロカルマルチンゲールで,
 $\langle M \rangle_\infty = \infty$ とする。この時, $U_s \equiv \inf \{t \geq 0 : \langle M \rangle_t = s\}$
 $(= \langle M \rangle_s^{-1})$ に対して $B_s = M_{U_s}$ は標準ブラウン運動である。

ここでは次の三つ式を設定する :

$$M_t = \int_0^t \sqrt{X_s} dW_s \quad (3.16)$$

$$\langle M \rangle_t = \int_0^t X_s ds \quad (3.17)$$

$$\inf \{t \geq 0 : \langle M \rangle_t = \sigma^2\} = U \quad (3.18)$$

よって $\int_0^U \sqrt{X_s} dW_s = B_U \sim N(0, \sigma^2)$

$$X_U = \lambda U - \Delta \sigma^2 + \sigma Z, Z \sim N(0, \sigma^2)$$

以上の計算まとめると

$$\text{分母} = \sigma^2$$

結論 :

$$c(\hat{m}_{\tau_c} - 1) = \frac{\sigma(-\Delta + Z)}{\sigma^2} \sim \begin{cases} N(0, 1) & H_0 : m = 1, \Delta = 0 \\ N(-\Delta, 1) & H_1 : m = 1 - \frac{\Delta}{c}, \Delta \neq 0 \end{cases} \quad (3.19)$$

3.4.4 ベッセル過程 (Bessel Process)

Revuz & Yor[参考文献 14]より δ で定義される初期値 0 の Bessel Process とは

$$d(\rho_s) = dB_s + \frac{\delta - 1}{2} \frac{1}{\rho_s} ds \quad (3.19)$$

ベッセル過程の期待値を求めるため, The transition density を利用する。

$p_t^\delta(x, y) = p(\rho_{s+t} \in dy | \rho_s = x), (t > 0, s \geq 0)$ は次の式で定義される。

$$(1) p_t^\delta(x, y) = t^{-1} (y/x)^v y \exp(-(x^2 + y^2)/2t) I_\nu(xy/t), x > 0$$

$$(2) p_t^\delta(0, y) = 2^{-v} t^{-(v+1)} \Gamma(v+1)^{-1} y^{2v+1} \exp(-y^2/2t)$$

ここでは

$$v = (\delta/2) - 1$$

$$I_\nu(x) = \sum_{k=0}^{\infty} (x/2)^{2k+\nu} / k! \Gamma(\nu+k+1).$$

(3.15) ~ (3.17) の証明より

$$\left. \begin{aligned} M_t &= \int_0^t \sqrt{X_s} dW_s \\ \langle M \rangle_t &= \int_0^t X_s ds \\ U_s &= \inf \{ t : \langle M \rangle_t = s \} \end{aligned} \right\} \text{とする。}$$

CIR Process に注目し、

$$\begin{aligned} X_s &= \int_0^s (\lambda - \Delta X_s) ds + \sigma \int_0^s \sqrt{X_s} dW_s \\ &= \lambda t - \Delta \langle M \rangle_t + \sigma M_t \end{aligned}$$

よって

$$\frac{dU_s}{ds} = \frac{1}{\lambda t - \Delta \langle M \rangle_{U_s} + \sigma M_{U_s}} = \frac{1}{\lambda U_s - \Delta s + \sigma B_s} \quad (3.21)$$

(注意: U_s の定義より $\langle M \rangle_{U_s} = s$ 、DDS定理より $M_{U_s} = B_s$ である)

$$\sigma \rho_s = \lambda U_s - \Delta s + \sigma B_s \text{ とする。従って } \frac{dU_s}{ds} = \frac{1}{\sigma \rho_s}.$$

$$d(\sigma \rho_s) = \lambda dU_s - \Delta ds + \sigma dB_s = \left(-\Delta + \frac{\lambda}{\sigma \rho_s}\right) ds + \sigma dB_s$$

$$\Rightarrow d(\rho_s) = \left(-\frac{\Delta}{\sigma} + \frac{\lambda}{\sigma^2 \rho_s}\right) ds + dB_s, \rho_0 = 0 \quad (3.22)$$

3.5 シミュレーションと結果

3.5.1 帰無仮説下でシミュレーションの理論的な分析

$H_0: m=1$ で検定を行う。

$$\Delta = 0 \text{ なので、 } d(\rho_s) = \frac{\lambda}{\sigma^2 \rho_s} ds + dB_s, \rho_0 = 0$$

$$E_0(\rho_s) = \frac{\lambda}{\sigma} E_0(U_s) \quad \text{つまり} \quad E_0\left(\frac{T_c}{c}\right) = E_0(U_{\sigma^2}) = \frac{\sigma}{\lambda} E_0(\rho_{\sigma^2})$$

$$E_0(\rho_t) = 2^{-v} t^{-(v+1)} \Gamma(v+1)^{-1} \int_0^\infty y \cdot y^{2v+1} \exp(-y^2 / 2t) dy$$

$$2^{-v} t^{-(v+1)} \Gamma(v+1)^{-1} = \Phi_1, \int_0^\infty y \cdot y^{2v+1} \exp(-y^2 / 2t) dy = \Phi_2 \text{ とする。}$$

ガンマ関数の性質が必要である：

$$\Gamma(\alpha) = \int_0^\infty y^{\alpha-1} e^{-y} dy, y = \frac{x^2}{\beta} \text{ の場合は}$$

$$\Gamma(\alpha) = \int_0^\infty \left(\frac{x^2}{\beta}\right)^{\alpha-1} e^{-y} d\left(\frac{x^2}{\beta}\right) = \int_0^\infty \frac{x^{2\alpha-2}}{\beta^{\alpha-1}} e^{-\frac{x^2}{\beta}} \cdot \frac{2x}{\beta} dx$$

Φ_2 はこの式の右辺のような形であるので、 $\alpha = v + 3/2, \beta = 2t$ で書換えると

$$\Phi_2 = \frac{1}{2} \Gamma\left(v + \frac{3}{2}\right) (2t)^{v+\frac{3}{2}}$$

$$E_0(\rho_t) = \Phi_1 \cdot \Phi_2 = 2^{-v-1} t^{-(v+1)} \Gamma(v+1)^{-1} \Gamma\left(v + \frac{3}{2}\right) (2t)^{v+\frac{3}{2}}$$

$t=1$ のとき

$$E_0(\rho_1) = \Phi_1 \cdot \Phi_2 = \frac{\sqrt{2} \Gamma\left(v + \frac{3}{2}\right)}{\Gamma(v+1)}$$

$$E_0(U_1) = \frac{\sigma}{\lambda} E_0(\rho_1) = \frac{\sigma}{\lambda} \frac{\sqrt{2} \Gamma\left(v + \frac{3}{2}\right)}{\Gamma(v+1)}$$

シミュレーションでは λ と σ の値を設定して、 $E_0(U_1)$ を考察する。

例： $\lambda = 1/4$ そして $\sigma = 1$ の場合、

$$\frac{\delta-1}{2} = \frac{\lambda}{\sigma^2} \text{ より } \delta = 3/2$$

$$v = (\delta/2) - 1 = -\frac{1}{4}$$

$$E_0(\rho_1) = \frac{\sqrt{2} \Gamma\left(v + \frac{3}{2}\right)}{\Gamma(v+1)} = \frac{\sqrt{2} \Gamma\left(\frac{5}{4}\right)}{\Gamma\left(\frac{3}{4}\right)} = 1.0461$$

$$E_0(U_1) = 4.1836$$

(ガンマ関数の値が以下のサイトより計算された：
<http://keisan.casio.jp/exec/system/1161228683>)

サンプリングの大きさに関する希望は、 c の値を調整することによって達成できる。停止時刻の条件に満たす上、200個データを取り出したいとき、理論値の計算より $c = T_c/4.1836 = 200/4.1836 = 47.8$ と設定するべきだ。シミュレーションでは H_0 の基で逐次 (σ_n) の手法により乱数に対して計算値とモデル対して理論値は一致するかどうかと観察したい。そして 3.18 より $c(\hat{m}_T - 1)$ をシミュレーションより検証したい。

3.5.2 対立仮説下でシミュレーションの理論的な分析

$H_1 : m = 1 - \Delta/n$ の下で (3.21) より

$d(\rho_s) = \left(-\frac{\Delta}{\sigma} + \frac{\lambda}{\sigma^2 \rho_s}\right) ds + dB_s, \rho_0 = 0$ となるので、 $\lambda = 1/4, \sigma = 1$ と設定したとき、

ρ_s は $\delta = 3/2$ 次元、ドリフト $-\frac{\Delta}{\sigma}$ 、初期値 0 のベッセル過程である。s=1 の時、

$$E_1\left(\frac{T_c}{c}\right) = E_1(U_1) = \frac{\sigma}{\lambda} E_1\left(\rho_1 - \frac{\Delta}{\sigma}\right) = \frac{\sigma}{\lambda} E_1(\rho_1) - \frac{\Delta}{\lambda}.$$

ドリフト付 $E_1(\rho_1)$ の値については、Linetsky (2004) より計算方法が紹介されている。複雑な数学理論で証明されたため、本論文では理論部分の証明を省略する。

シミュレーションの目的は対立仮説の基で c の値を調整することによって停止時刻の条件に満たす上サンプリングの大きさをコントロールする。サンプリングの大きさの規模が決めれば、 H_1 の基で逐次 (σ_n) の手法により乱数に対して計算値とモデル対して理論値は一致するかどうかと観察したい。

3.5.3 シミュレーションの結果

この節ではシミュレーションの結果を載せてある。

帰無仮説を検証するために統計解析ソフト R を用いた。プログラムを書く際、次のサイトをよく参考させていただいた：

<http://cse.naro.affrc.go.jp/takezawa/r-tips/r.html>.

Stopping Time の理論値とシミュレーション計算値がほぼ一致の結論ができている。(表 3-1 を参照)

$c(\hat{m}_T - 1)$ が H_0 の基で標準正規分布に従うことを検証するため正規性に有効な

Kolmogorov-Smirnov 検定 (ks. test) を行っていた。R の中では p 値で表す。p の値は大きいほど、正規分布が確定できる。一般的 0.05 と比べるものである。値とヒストグラムを合わせて分析すると、 $c(\hat{m}_T - 1)$ の正規性をよく確認できた。

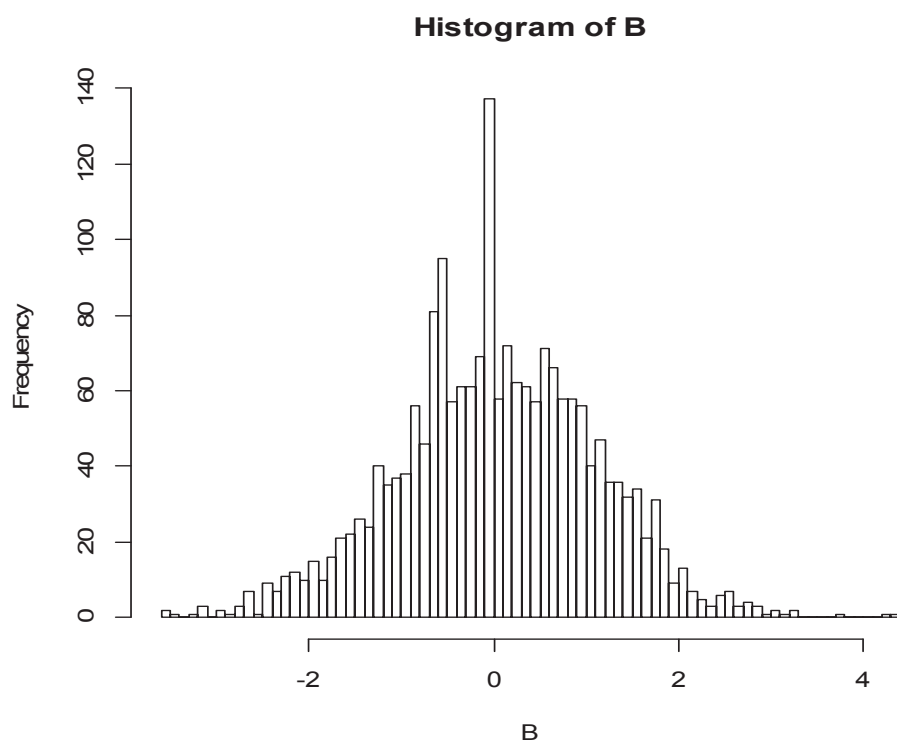
表 3-1
R より帰無仮説が真の時

$m_0=1, \text{真の } \sigma=1$						
世代	2000	10000	2000	10000	2000	10000
シミュレーション回数 n	2000	1000	2000	1000	2000	1000
Tc	50	50	100	100	200	200
$c = Tc/4.1836$	11.591	11.591	23.903	23.903	47.805	47.805
λ	0.25	0.25	0.25	0.25	0.25	0.25
(停止時刻) ST	49.34	49.8	101.134	97.967	204.349	203.779
p-value	0.3006	0.1928	0.4952	0.6728	0.7251	0.6540
$c(\hat{m}_{T_c} - 1)$	Picture1	Picture2	Picture3	Picture4	Picture5	Picture6

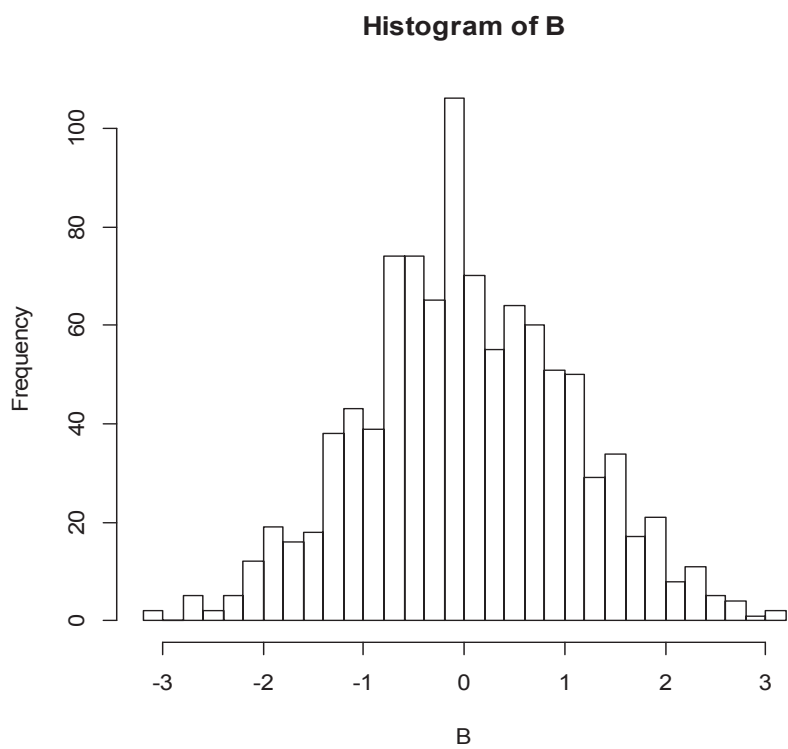
表 3-2

$m_1=0.99, \text{真の } \sigma=1$						
世代	2000	10000	2000	10000	2000	10000
シミュレーション回数 n	2000	1000	2000	1000	2000	1000
Tc	50	50	100	100	200	200
Δ	0.07	0.07	0.09	0.09	0.14	0.14
λ	0.25	0.25	0.25	0.25	0.25	0.25
(停止時刻) ST	51.92	48.22	108.13	109.55	210.33	216.51
p-value	0.7755	0.5224	0.8438	0.1910	0.3406	0.5472
$c(\hat{m}_{T_c} - 1)$	Picture7	Picture8	Picture9	Picture10	Picture11	Picture12

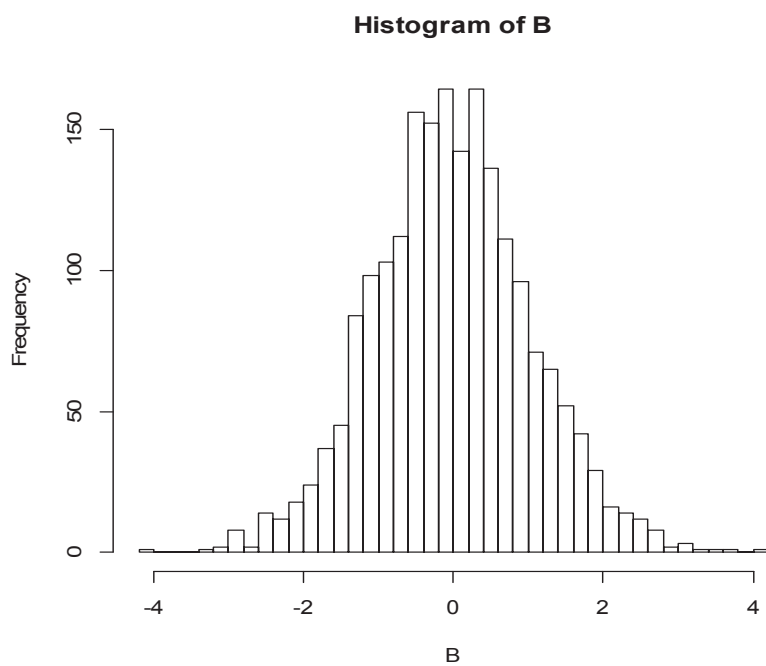
picture1



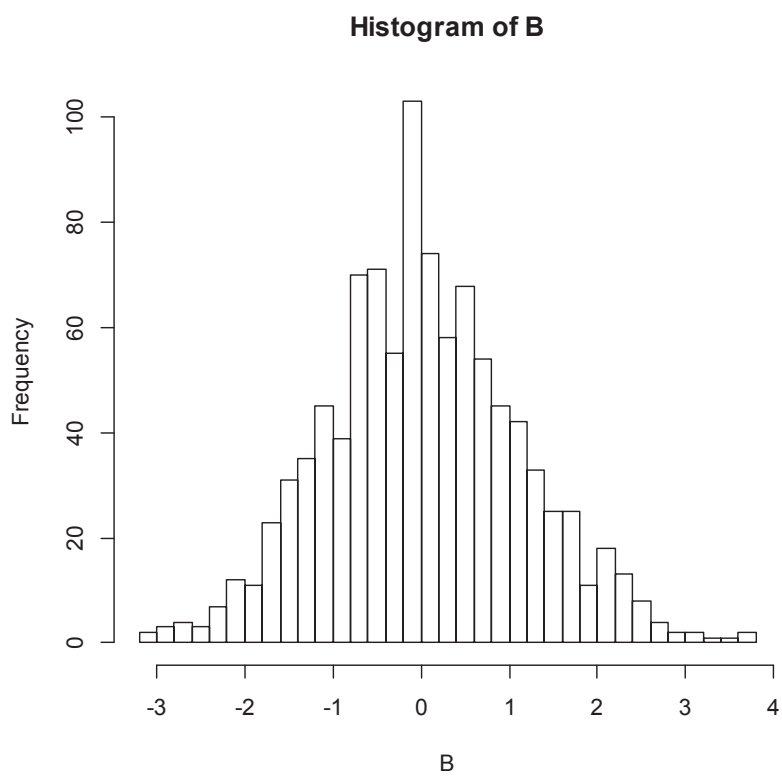
picture2



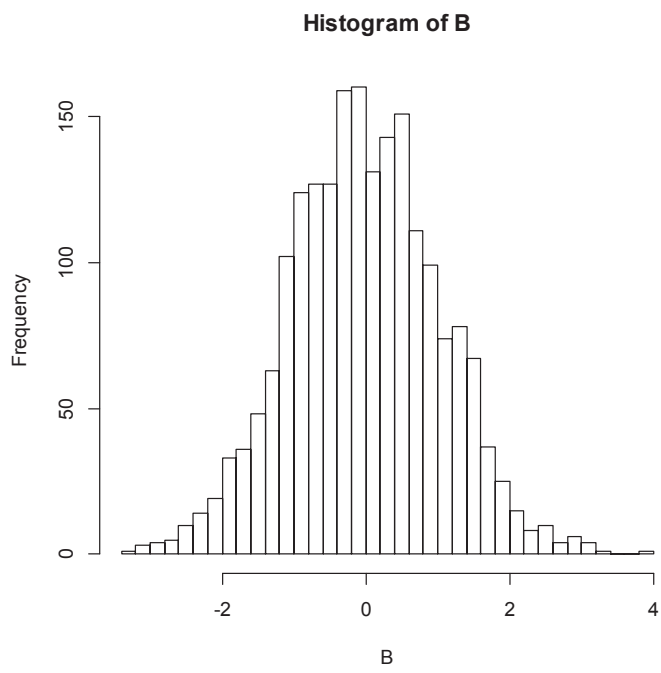
picture3



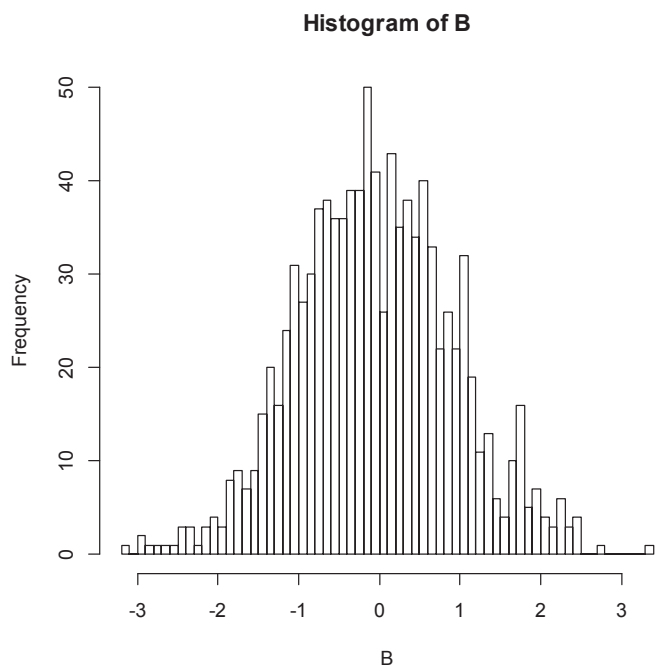
picture4



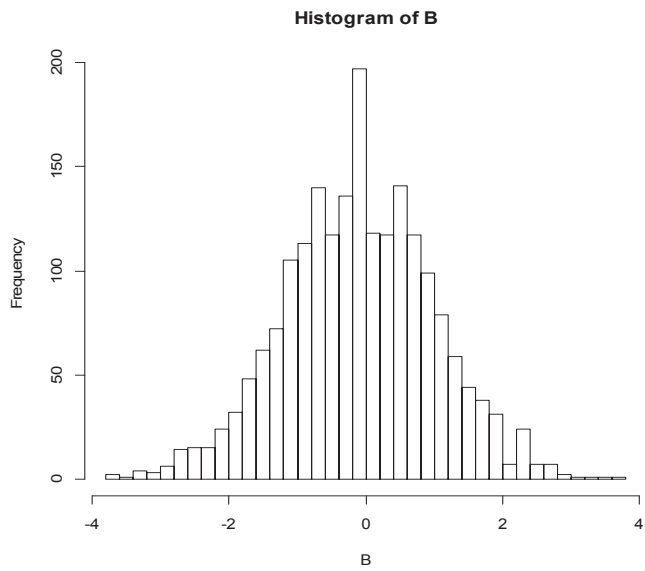
picture5



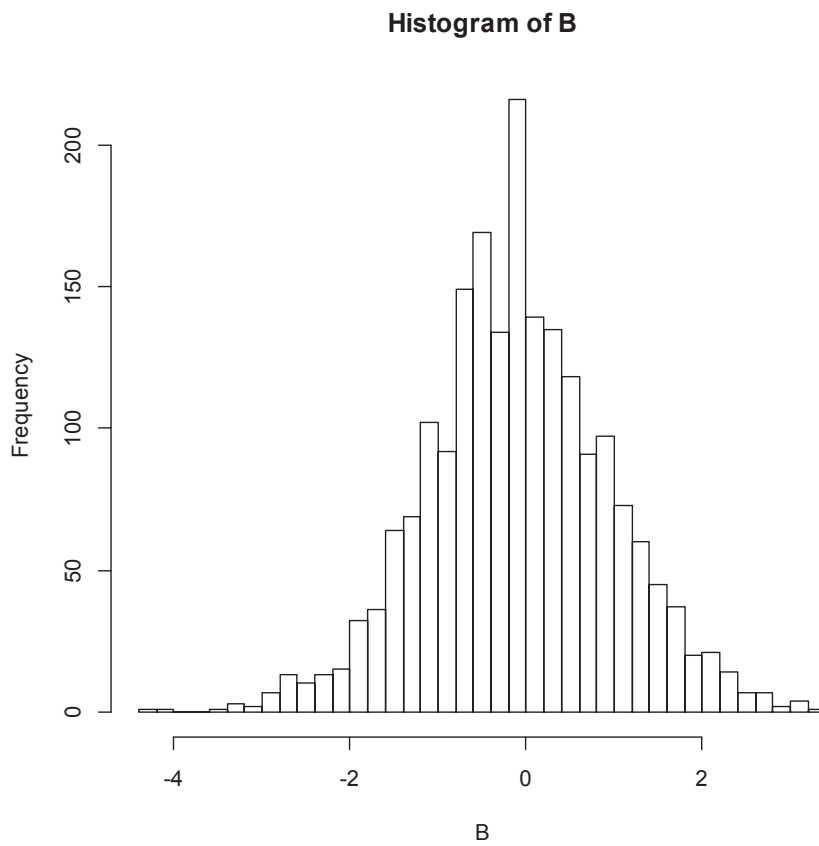
picture6



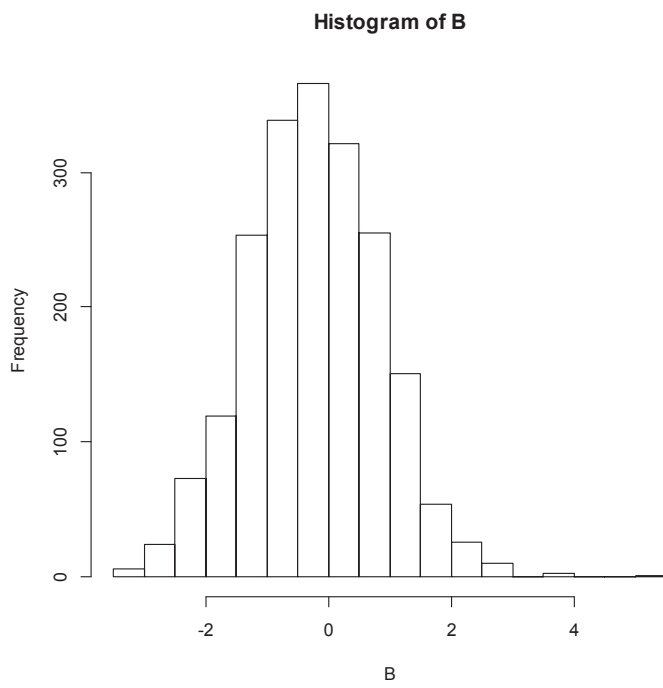
Picture7



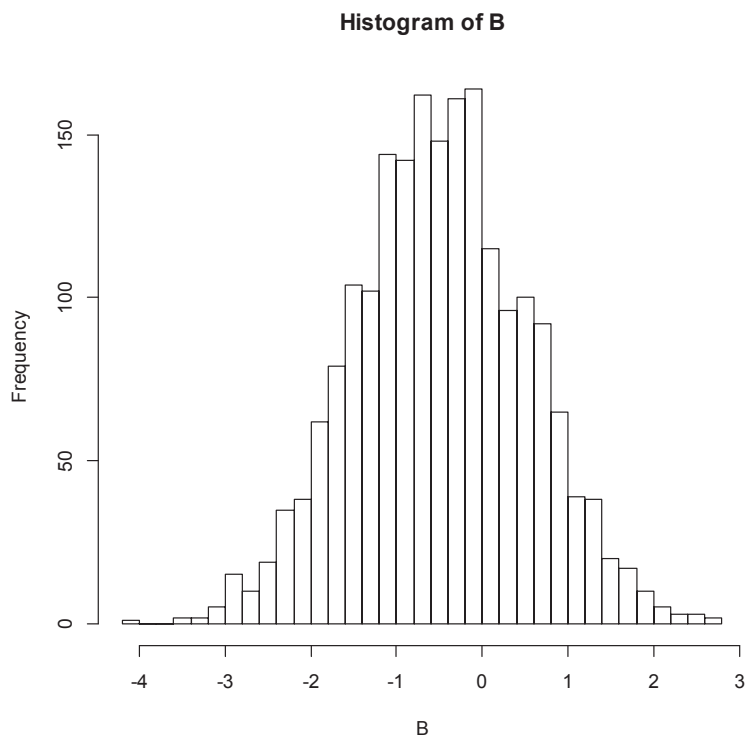
Pictrure8



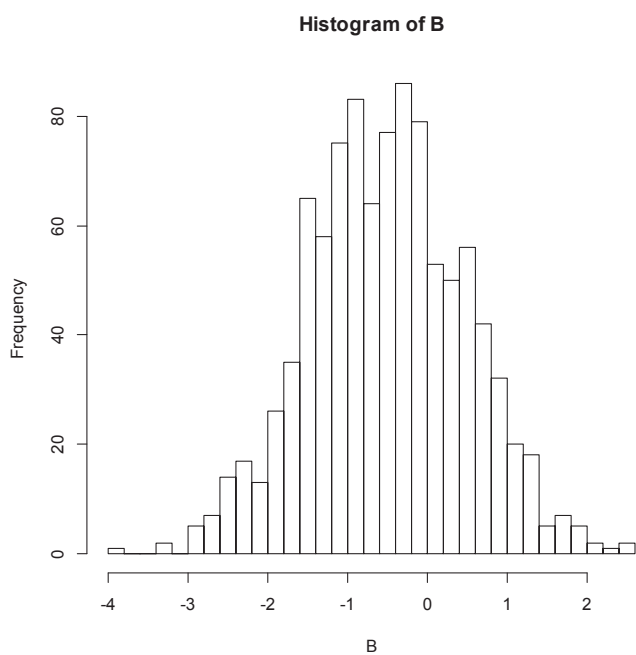
Picture9



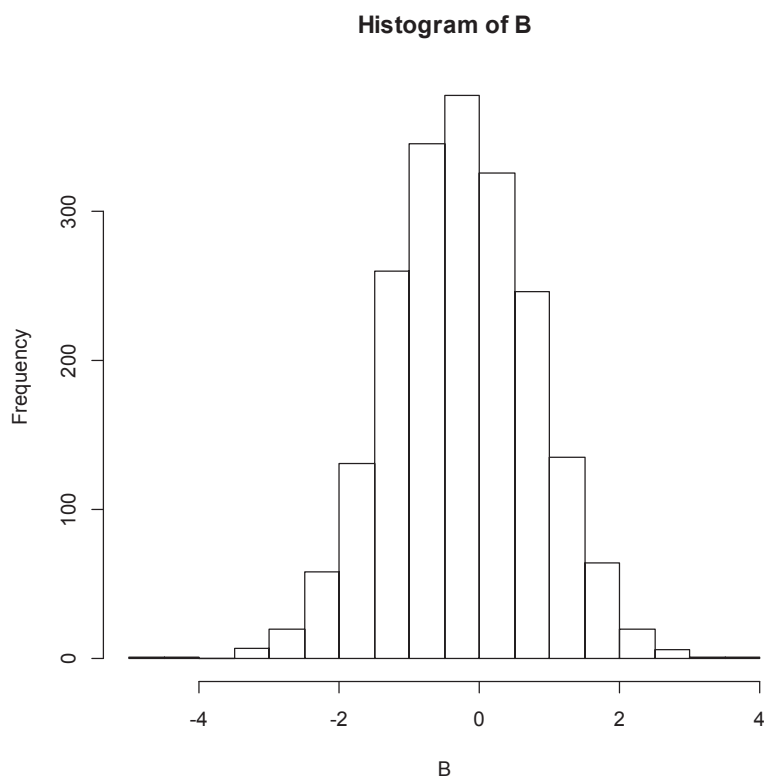
Picture10



Picture11



Picture12



第四章 SPRT 検定 (Sequential Probability Ratio Test)

逐次確率比検定 (SPRT) は Abraham Wald 1947 年よって開発した特定の逐次仮説検定である。Neyman と Pearson の 1933 結果は、逐次解析問題として再定式化した。この結果は Wald に影響を与えた。

もともと製造業の分野での品質管理の研究で使用するために開発されたが、現在様々な分野で使用されている。例えば地震警報システムの研究や医薬品関連のテストなど。停止時刻の設定より、希望する条件に至るまでサンプリングを続け、希望条件に満足させる一番早い時点でサンプリングが止まる。このような特徴を持ち、警報性と予測性が付く実用性の高い解析方法である。

4.1 Wald の SPRT 検定

4.1.1 Wald の SPRT 検定の意味

Wald の逐次確率比検定 (SPRT) は、現在 SISA で唯一のベイズ統計的処理である。ベイズ統計に必要なのは、事前のデータ収集を実行することと、モデルの下で成果の期待は非常に詳細に説明することである。Wald の SPRT において、特定の条件の下、データ収集の決定時に満たされている場合には、データ収集、収集したデータの解釈を継続に関して取られる。

逐次解析は、実験または収集されたサンプリング数は、実験開始時に固定された古典的な仮説検定とは異なります。古典的な仮説検定では、データの収集、データの解析を考慮せずに実行される。すべてのデータが収集された後に分析が行われ、結論が引き出される。しかしながら、シーケンシャル分析ですべてのケースが収集された後に直接分析され、その時点までに収集されたデータは、その後、新たに収集されたケースから得られた新たな情報を組み込んだ特定の数値と比較されている。このアプローチは、データ収集中に結論を出すことを可能にし、古典仮説検定の場合のように、最終的な結論は、おそらく、かなり早い段階で到達することができる。データ収集が少ない例とそれ以前の決定事項の後に終了することができるように、人間の生活と不幸、金融貯蓄の面で節約は、かなりの可能性がある。このような例を考えると逐次解析の利点はわかりやすい。

状況を説明するためにまず基準を設定しなければならない。ある基準の下であれば「少ない」、ある基準の上であれば「多い」と判断が可能である。一つは、データ収集前に、これらの基準値を定義する必要がある。通常であれば正の数のパーセントを与える。一つ目のパーセント以上の場合は改善 (多い) された。二つ目のパーセント以下の割合は劣化 (少ない) したとしての状況を判断する。定義され閾値に至るまでサンプリングを続けて追加する。収集されたサンプルをしきい値を超える時点でデータ収集が終了される。

この二つのパーセント値が必要とされ、統計的では検出力 $1-\beta$ と有意水準 α と呼ばれる。検出力とは、真の中に相違存在する場合、検出力 $1-\beta$ の確率で相

違データを観察する機会である。たとえば検出力を 80% に設定されている。サンプル中るに実際に存在する差異を発見するが 80% である。α は次事象の確率を表す：相違を発見したと思うので検定を終了させたのに、実際にはこの相違が存在していない。α は 5% を設定された時、サンプリングは 5% の確率で誤って停止されていたとの意味である。一般的では検定力と有意水準の設定は広く受け入れられている通常値にする。検定の内容など考えて、変更も可能である。

4.1.2 しきい値の決定

Wald の結果より、a と b は次の関係がある。a, b は正と負の閾値であり、ここで、α は第一種の過誤確率、β は第二種の過誤確率である。

$$a = \log\left(\frac{\beta}{1-\alpha}\right), b = \log\left(\frac{1-\beta}{\alpha}\right).$$

具体的に見ると、SPRT 検定の帰無仮説と対立仮説は次のものとなる：

$$H_0: \beta = \beta_0 \quad VS \quad H_1: \beta = \beta_1$$

次の尤度関数を用いて SPRT 検定を行う。

$$\Lambda_i = \prod_{j=1}^n \frac{f_1(x_j)}{f_0(x_j)}, n = 1, 2, \dots \quad (4.1)$$

SPRT 検定の目的はより早い段階で仮説の正確性を判断するため、 $a < b$ の条件のもとで、b が大きいほど、a が小さいほどサンプリングの停止時刻が遅くなって、より厳密な検定ができる。この段階で所望の検出力、誤報率を提供するために、しきい値を設定しよう。検出率 P_D (検定力 β と対応) そして誤報率 P_{FA} (有意水準 α と対応)。

表記を簡略化するために $x := (x_1, \dots, x_n)$ そして $f_j(x) := \prod_{i=1}^n f_i(x_i)$, ($j = 0, 1$) .

$$\beta = P_D = \int_{R_1} f_1(x) dx \int_{R_1} \frac{f_1(x)}{f_0(x)} f_0(x) dx = \int_{R_1} \Lambda_n(x) f_0(x) dx \quad w = \alpha \Lambda_n(x)$$

$\log \Lambda_n(x) = \log \frac{\beta}{\alpha} \geq b$ との事実を利用して、次の式が導くことができる。

$$1 - \alpha = 1 - P_{FA} = 1 - \int_{R_0} f_0(x) dx = \int_{R_0} \frac{f_0(x)}{f_1(x)} f_1(x) dx = \int_{R_0} \Lambda_n^{-1}(x) f_1(x) dx$$

$$\text{なので } \log \Lambda_n^{-1}(x) = \frac{1-\alpha}{1-\beta} \geq a^{-1}. \quad (4.2)$$

これらの式は、私たちに達成するために必要なしきい値の上の境界を与える。 $P_D(\beta)$ と $P_{FA}(\alpha)$ は次の関係に満たす：

$$a \leq \log \frac{1-\beta}{1-\alpha}, b \geq \log \frac{\beta}{\alpha}$$

$a = \log \frac{1-\beta}{1-\alpha}$ として $b = \log \frac{\beta}{\alpha}$ とする。これらのしきい値は、試験の誤り確率は、 α と β の選択によって指定されたように、少なくとも限り小さくすることを保証するが、結果はあまりにも保守的である可能性がある。これより単純な閾値法で停止規則が表示される：

- ・ $a < \log \Lambda_n < b$: サンプルングを継続
- ・ $\log \Lambda_n \geq b$: H_1 を選択
- ・ $\log \Lambda_n \leq a$: H_0 を選択

つまり、適切にしきい値を設定するためには、 α, β を事前に決定しなければならない。数値は、アプリケーションに依存するであろう。離散的な場合には、サンプル間の閾値を交差することができる。したがって、誤差をコントロールするため、より積極的にしきい値を設定するかもしれない。もちろん、厳密な境界は連続的な場合に使用することができる。

4.2 分枝過程の SPRT 検定

$$\begin{cases} H_0 : m = 1 \\ H_1 : m = 1 - \frac{\Delta}{n} \end{cases} \text{ と設定する。}$$

分枝過程において SPRT 検定とはサンプルングが次の式に満たすとき停止する：

$$T_{a,b}^c = \inf \{ n \geq 1 : \log \Lambda_n^c \notin (a, b) \} \quad (4.3)$$

ここでは Λ_n^c が対数尤度比である。

$$\Lambda_n^c = \frac{P_1(z_1, y_1, \dots, z_n, y_n | z_{n-1})}{P_0(z_1, y_1, \dots, z_n, y_n | z_{n-1})} = \prod_{i=1}^n \frac{P_1(z_i, y_i | z_{i-1})}{P_0(z_i, y_i | z_{i-1})}. \quad (4.4)$$

$\log \Lambda_{T_{a,b}^c}^c \leq a$ のとき H_0 を選択する。

$\log \Lambda_{T_{a,b}^c}^c \geq b$ のとき H_1 を選択する。

a と b がサイズ (size) とパワー (power) であるので、指定水準に従う。(例えば 5% と 70% など)。 a と b の値は検定の検出力を表す重要な数値である。検出力はこの表のように理解しておく。

		実際に成り立っているのは	
		帰無仮説 H_0	対立仮説 H_1
検定結果	H_0	正しい $1 - \alpha$	第二種類誤り β
	H_1	第一種類誤り α	正しい $1 - \beta$

まず分枝過程において対数尤度比を見てみよう。第三章では (3.9) 式で対数

尤度比を与えられた。重要なので、もう一度書いておく。

$l_n = \log L_n$ とする。

$$l_n = \sum_{i=1}^n \left\{ \log a_{z_i - y_i}^{(z_{i-1})} + (z_i - y_i) \log \theta - z_{i-1} \log A(\theta) + \log by_i + y_i \phi - \log B(\phi) \right\} \quad (4.5)$$

$$l_n^c = \log \Lambda_n^c.$$

$n = [ct]$ で書き換えると

$$l_{[ct]}^c = -\frac{\Delta}{c\sigma^2} (Z_{[ct]} - 1) + \frac{\Delta}{c\sigma^2} \sum_{i=1}^{[ct]} Y_i - \frac{\Delta^2}{2c^2\sigma^2} \sum_{i=1}^{[ct]} Z_{i-1} + o\left(\frac{1}{c^2}\right) \quad (4.6)$$

$$\rightarrow -\frac{\Delta}{\sigma^2} X_t + \frac{\Delta\lambda t}{\sigma^2} - \frac{\Delta^2}{2\sigma^2} \int_0^t X_u du \quad (4.7)$$

従って、分枝過程に対する SPRT 検定の停止時刻は次の式で定義される。

$$T_{a,b}^c = \inf \{ n : l_n^c \notin (a, b) \} = \inf \{ ct : l_{[ct]}^c \notin (a, b) \}$$

両辺に t で割ると

$$\begin{aligned} \frac{T_{a,b}^c}{c} &= \inf \{ t : l_{[ct]}^c \notin (a, b) \} \\ &\rightarrow T_{a,b} \equiv \inf \left\{ t : -\frac{\Delta}{\sigma^2} X_t + \frac{\Delta\lambda t}{\sigma^2} - \frac{\Delta^2}{2\sigma^2} \int_0^t X_u du \notin (a, b) \right\} \end{aligned} \quad (4.8)$$

3.2.2 の CIR 過程より $dX_t = (\lambda - \Delta X_t)dt + \sigma\sqrt{X_t}dW$ 。従って X_t は次二つの表現がある。

$$H_0 : m = 1, \Delta = 0$$

$$X_t = \int_0^t \lambda du + \sigma \int_0^t \sqrt{X_u} dW_u \quad (4.9)$$

$$H_1 : m = 1 - \Delta/c, \Delta > 0$$

$$X_t = \int_0^t \lambda du - \int_0^t \Delta X_u du + \sigma \int_0^t \sqrt{X_u} dW_u \quad (4.10)$$

この時仮説ごとに SPRT の表現が異なる。

$$H_0 : m = 1, \Delta = 0$$

$$T_{a,b} = T_{a,b}^+ \equiv \inf \left\{ t : -\frac{\Delta}{\sigma} \left(\frac{\Delta}{2\sigma} \int_0^t X_u du + \int_0^t \sqrt{X_u} dW_u \right) \notin (a, b) \right\} \quad (4.11)$$

$$H_1 : m = 1 - \Delta/c, \Delta > 0$$

$$T_{a,b} = T_{a,b}^- \equiv \inf \left\{ t : -\frac{\Delta}{\sigma} \left(-\frac{\Delta}{2\sigma} \int_0^t X_u du + \int_0^t \sqrt{X_u} dW_u \right) \notin (a, b) \right\} \quad (4.12)$$

ここではもう一度 DDS 定理を使って、時間変更をする。そうすると、SPRT の表現をブラウン運動で表せることが見つけた。DDS 定理の条件は第三章と同様に設定する。

$$\text{DDSによつて} \left\{ \begin{array}{l} M_t = \int_0^t \sqrt{X_s} dW_s \\ \langle M \rangle_t = \int_0^t X_s ds \\ U_s = \langle M \rangle_s^{-1} \end{array} \right\} \text{に対して } B_s = M_{U_s} \text{として,}$$

$$\mu = \frac{\Delta}{2\sigma} > 0 \text{とすると,}$$

$$H_0 : \tau_{a,b}^- \equiv \inf \left\{ s \geq 0 : -\mu s - B_s \notin \left[\frac{a}{2\mu}, \frac{b}{2\mu} \right] \right\} \quad (4.13)$$

$$H_1 : \tau_{a,b}^+ \equiv \inf \left\{ s \geq 0 : +\mu s - B_s \notin \left[\frac{a}{2\mu}, \frac{b}{2\mu} \right] \right\} \quad (4.14)$$

$$\text{第 1 種誤り } \alpha = P_{H_0} \left\{ I_{\tau_{a,b}^-}^c \geq b \right\} \rightarrow P \left\{ -\mu \tau_{a,b}^- - B_{\tau_{a,b}^-} \geq \frac{b}{2\mu} \right\} \quad (4.15)$$

$$\text{第 2 種誤り } \beta = P_{H_1} \left\{ I_{\tau_{a,b}^+}^c \leq a \right\} \rightarrow P \left\{ \mu \tau_{a,b}^+ - B_{\tau_{a,b}^+} \leq \frac{a}{2\mu} \right\} \quad (4.16)$$

以上より *Wald* の方法により閾値 a, b の値を決めることが正当化される。

$$a = \log \left(\frac{\beta}{1-\alpha} \right), b = \log \left(\frac{1-\beta}{\alpha} \right)$$

以上の計算方法を利用して、シミュレーションを実行してみた。事前に α と β の値を設定した。(4.15) と (4.16) で計算された値とくらべた結果は良いと考える。停止時刻の結果は理論値の計算が不十分で、シミュレーションで値がよいと評価できない。

4.3 分枝過程に対する SPRT 検定のシミュレーション

表 4-1

H_0 が真、 $m = 1$	α	0.05	0.05	0.05
	β	0.2	0.25	0.3
	第 1 種類の誤り	0.073	0.065	0.554
		12.04	32.26	52.55

H_1 が真、 $m = 0.95$	α	0.05	0.05	0.05
	β	0.2	0.25	0.3
	第 2 種類の誤り	0.250	0.314	0.369
		70.22	89.95	124.68

第五章 移民項のある 2 変量分枝過程

5.1 モデルについての解説

5.1.1 モデルと直感の意味

各変量は確率空間 (Ω, A, P) で定義される。 Z_+, N, R, R_+ をそれぞれ非負整数、正整数、実数、非負実数を表す。

$k, j \in Z_+$ として $i, l \in \{1, 2\}$ のとき、第 i 種類物の第 k 世代の数は $X_{k,i}$ で表示される。この数は第 $k-1$ 世代の物の各成員の子孫と外から飛び込んだ成員（移民）の和で決める。ここでは、 $i=2$ (2種類物があり、タイプ 1 とタイプ 2 と呼ぶ) の場合を考えてみる。

$$\begin{bmatrix} X_{k,1} \\ X_{k,2} \end{bmatrix} = \sum_{j=1}^{X_{k-1,1}} \begin{bmatrix} \xi_{k,j,1,1} \\ \xi_{k,j,1,2} \end{bmatrix} + \sum_{j=1}^{X_{k-1,2}} \begin{bmatrix} \xi_{k,j,2,1} \\ \xi_{k,j,2,2} \end{bmatrix} + \begin{bmatrix} \varepsilon_{k,1} \\ \varepsilon_{k,2} \end{bmatrix} \quad (5.1)$$

詳しく解説すると、タイプ 1 のもの k 世代の数 $X_{k,1}$ は三つ部分の和で決められる：

① $k-1$ 世代でタイプ 1 のもの $(X_{k-1,1})$ より生まれた k 世代タイプ 1 の子孫

($\xi_{k,j,1,1}$) ; ② k 世代でタイプ 2 のもの $(X_{k-1,2})$ より生まれた k 世代タイプ

1 の子孫 ($\xi_{k,j,2,1}$) ; ③ 第 k 世代で外から飛び込んだタイプ 1 のもの ($\varepsilon_{k,1}$)。

j は各集団の第 j 番目の成員と表示しておく。

いくつかの例で挙げておく。

例 1 : ある花は白 (タイプ 1) と赤 (タイプ 2) 二色がある。一株にお花の色は同一であるが、種ができれば、白と赤どちらでも生まれる。白色の「母親」

($\xi_{k-1,j,1}$) であるが、白の「子供」2 株 ($\xi_{k,j,1,1}=2$) と赤の「子供」1 株 ($\xi_{k,j,1,2}=1$) が造れる。当然赤色のお花もどちらの色の子供が生まれる。生物試験で一株と一株分けて、子孫数を数えられる環境を作る。そして各世代白花と赤花を数える ($X_{k,1}$ と $X_{k,2}$)。偶々色のバランスを考えて、新しいお花を補充することがあり、これは移民の $\varepsilon_{k,1}$ と $\varepsilon_{k,2}$ である。

例 2 . 二つの銀行へ預金する。 k の期間は 1 月間にする。 1 月 1 日銀行 A と銀行 B それぞれある金額の預かりする。 1 月末元金を引き出して、利息のみを再預金にする。ただ利息は今までの銀行におくこともできるし、一部他の銀行に入れることも可能である。ランダムで決めるでよい。このように月末で元金を引き出して、利息より利息を産生することを考えると、2 変量分枝過程に従うではないだろう。どちらかの預金が 0 になると、そちらの銀行へ元金を投入し、さらに利息を出す。この部分は「移民項」である。このモデルで各期末 A 銀行と B 銀行の預金残高を考察できると思う。

現実で、もし利息の再預金する分割りはランダムではなく、お客様の意思で決められると、預金の行き方は利子率や銀行のサービスなどとの関係があるではないかと考える。なので、本モデルを利用して、利子率変動やサービス改善後

の預金残高が試算が可能ではないかと考えておく。

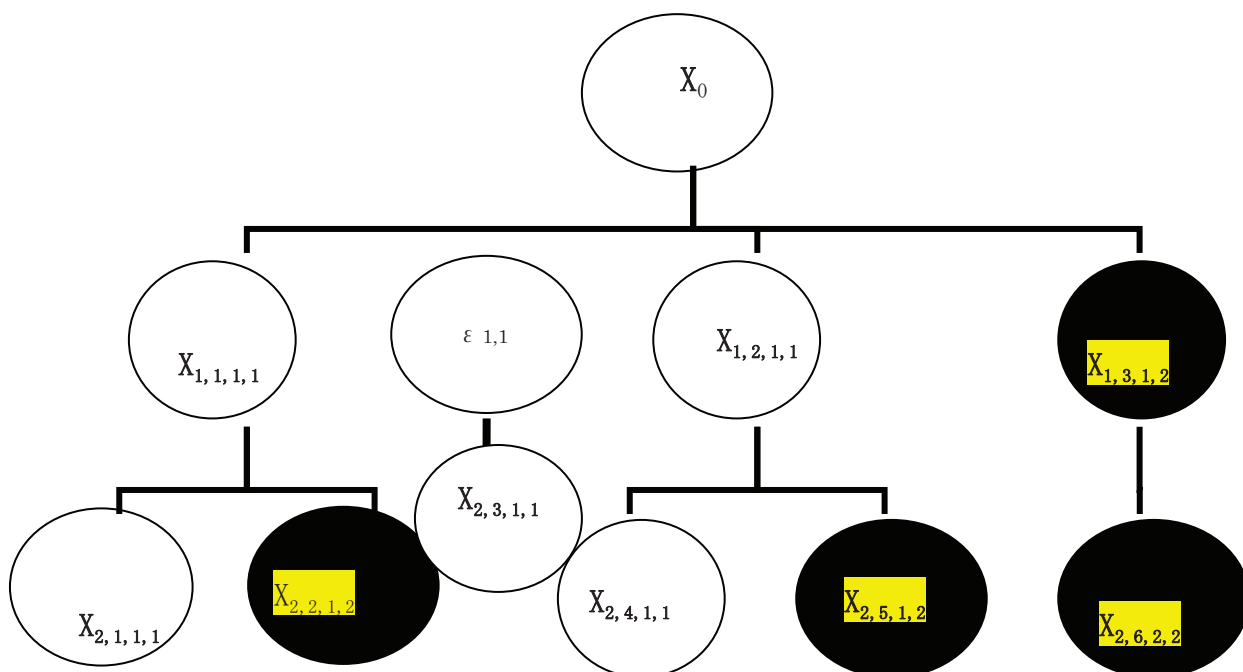


図 5 - 1 モデルの説明

5.1.2 記号の意味

モデルでは $\{X_0, \xi_{k,j,i}, \varepsilon_k : k, j \in N, i=1,2\}$ が独立である :

$$X_k = \begin{bmatrix} X_{k,1} \\ X_{k,2} \end{bmatrix} \quad (5.2)$$

$$\xi_{k,j,i} = \begin{bmatrix} \xi_{k,j,i,1} \\ \xi_{k,j,i,2} \end{bmatrix} \quad (5.3)$$

$$\varepsilon_k = \begin{bmatrix} \varepsilon_{k,1} \\ \varepsilon_{k,2} \end{bmatrix} \quad (5.4)$$

そして、 $E(\|\xi_{1,i}\|^2) < \infty, i=1,2$ $E(\|\varepsilon_1\|^2) < \infty$

下の記号について説明する。

$$m_{\xi_i} = E(\xi_{1,1,i}) \in R_+^2, m_{\xi} = [m_{\xi_1} \quad m_{\xi_2}] \in R_+^{2*2}$$

$$V_{\xi_i} = \text{Var}(\xi_{1,1,i}) \in R^{2*2}, \overline{V_{\xi}} = \frac{1}{2}(V_{\xi_1} + V_{\xi_2}) \in R^{2*2}$$

$$m_{\varepsilon} = E(\varepsilon_1) \in R_+^2, V_{\varepsilon} = \text{var}(\varepsilon_1) \in R^{2*2}$$

$k \in Z_+$ であるとき $F_k = \sigma(X_0, X_1, \dots, X_k)$ とする。3.1 式より

$$E(\mathbf{X}_k | F_{k-1}) = X_{k-1,1} m_{\xi_1} + X_{k-1,2} m_{\xi_2} + m_{\varepsilon} = m_{\xi} \mathbf{X}_{k-1} + m_{\varepsilon}$$

したがって $E(\mathbf{X}_k) = m_{\xi} E(\mathbf{X}_{k-1}) + m_{\varepsilon}$

$$\begin{aligned} E(\mathbf{X}_k) &= m_{\xi} E(\mathbf{X}_{k-1}) + m_{\varepsilon} = m_{\xi} (m_{\xi} E(\mathbf{X}_{k-2}) + m_{\varepsilon}) + m_{\varepsilon} \\ &= m_{\xi}^2 E(\mathbf{X}_{k-2}) + m_{\xi} m_{\varepsilon} + m_{\varepsilon} \\ &= m_{\xi}^2 (m_{\xi} E(\mathbf{X}_{k-3}) + m_{\varepsilon}) + m_{\xi} m_{\varepsilon} + m_{\varepsilon} \\ &= m_{\xi}^3 E(\mathbf{X}_{k-3}) + m_{\xi}^2 m_{\varepsilon} + m_{\xi} m_{\varepsilon} + m_{\varepsilon} \\ &= \text{L L} \end{aligned}$$

$$E(\mathbf{X}_k) = m_{\xi}^k E(\mathbf{X}_0) + \sum_{j=0}^{k-1} m_{\xi}^j m_{\varepsilon} \quad (5.5)$$

5.1.3 criticality parameter について

明らかに、 m_{ξ} はこの連鎖効果の中では重要な役割を担う。

本論文では次のような m_{ξ} を持ち、対称型の移民項のある 2 変量分枝過程を考える：

$$m_{\xi} = \begin{bmatrix} \alpha & \beta \\ \beta & \alpha \end{bmatrix} \quad (5.6)$$

さらに、 $\alpha > 0$ かつ $\beta > 0$ の場合のみ注目する。

この基で臨界パラメータ (criticality parameter) は $\rho = \alpha + \beta$ と定義されると、移民項のある 2 変量分枝過程が次のような三つの状態がある：

臨界未満 (subcritical) ----- $\rho < 1$

臨界 (critical) ----- $\rho = 1$

臨界超過 (supercritical) ----- $\rho > 1$

例 1 を用いて直観的に説明する。白花が同色の白花を生む期待値は α 、異色の赤花を生む期待値は β 、赤花が同色の赤花を生む期待値は α 、異色の白花を生む期待値は β である。 $\alpha + \beta = \rho < 1$ であるとは、各色の花は今の世代より少なく子孫を生むので、移民で植入の花を考えなければ、次世代の花の総数を減っていく傾向である。物种を維持するため、人工で新しいお花の植入が必要であるだろう。 $\alpha + \beta = \rho = 1$ であるとのことは、今の世代の花の規模がキープできるような速度で次世代の子孫を生む。移民で少し増量可能である。 $\alpha + \beta = \rho > 1$ ののであるとき、子孫数は親の数を超えるような増え方で、いずれが爆発的な数になって、異常事態をコントロールするため、事前警報システムが必要

だろう。

初期値については簡単化にするため、移民項のある2変量分枝過程の初期値は $X_0=0$ と定義されることが多い。(言われる0スタート)
本論文では、 $m_\xi \neq 0$, そして $\mathbf{X}_k = \mathbf{0}$ と設定する。

5.2 移民項のある2変量分枝過程の推定

5.2.1 CLS より移民項のある2変量分枝過程の推定

ここでは本論文のモデルに対する CLS 推定を行う。各変量の特性値を求める上、理論的な分布を計算する。

$$E(X_k | F_{k-1}) = m_\xi X_{k-1} + m_\varepsilon \text{ より}$$

$$M_k = X_k - E(X_k | F_{k-1}) = X_k - m_\xi X_{k-1} - m_\varepsilon, k \in \mathbb{N}. \quad (5.7)$$

一つの序列を導入する：

$$U_k = \langle \mathbf{1}, \mathbf{X}_k \rangle = \mathbf{1}^T \mathbf{X}_k = X_{k,1} + X_{k,2} \quad (5.8)$$

$$k \in \mathbb{Z}^+, \mathbf{1} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \in \mathbb{R}^2$$

この式より、次の結果がある：

$$\text{すべての } k \in \mathbb{Z}^+ \text{ に対して } U_k \geq 0 \text{ そして } U_k = U_{k-1} + \langle \mathbf{1}, m_\varepsilon \rangle + \langle \mathbf{1}, M_k \rangle \quad k \in \mathbb{N}.$$

理由は次である：

$\rho = \alpha + \beta = 1$ の意味は、 $\mathbf{1}$ が m_ξ の固有ベクトルであり、固有値が1である。なので

$$\langle \mathbf{1}, m_\xi X_{k,k-1} \rangle = \mathbf{1}^T m_\xi X_{k,k-1} = \mathbf{1}^T X_{k,k-1} = U_{k-1} \quad (5.9)$$

$$\text{さらに、} \mathbf{0} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \in \mathbb{R}^2,$$

$$V_k = \langle \mathbf{0}, \mathbf{X}_k \rangle = \mathbf{0}^T \mathbf{X}_k = X_{k,1} - X_{k,2} \quad k \in \mathbb{N}, \text{ とすると}$$

$$V_k = (\alpha - \beta) V_{k-1} + \langle \mathbf{0}, m_\varepsilon \rangle + \langle \mathbf{0}, M_k \rangle. \quad (5.10)$$

理由は次である：

$\mathbf{0}$ は m_ξ の固有ベクトルであり、固有値が $\alpha - \beta$ である。なので

$$\langle \mathbf{0}, m_\xi X_{k-1} \rangle = \mathbf{0}^T m_\xi X_{k-1} = (\alpha - \beta) \mathbf{0}^T X_{k-1}. \quad (5.11)$$

$$X_{k,1} = (U_k + V_k) / 2, \quad X_{k,2} = (U_k - V_k) / 2, \quad k \in \mathbb{Z}^+. \quad (5.12)$$

5.1 式で表すモデルに対し、CLS 推定を行って、臨界パラメータを見つける。

新しいパラメータ $\delta = \alpha - \beta$ を定義する。そして

$$\alpha = (\rho + \delta) / 2, \quad \beta = (\rho - \delta) / 2. \quad (5.13)$$

5.2 式は次の式に書き換える：

$$M_k = X_k - E(X_k | F_{k-1}) = X_k - m_\varepsilon N_{k-1} - m_\varepsilon, k \in N. \quad (5.14)$$

さらに

$$X_k = \frac{1}{2} \begin{bmatrix} \rho + \delta & \rho - \delta \\ \rho - \delta & \rho + \delta \end{bmatrix} X_{k-1} + M_{k-1} + m_\varepsilon, k \in N. \quad (5.15)$$

CLS 法で推定するとき、まず X_1, \dots, X_n を含む次の式を最小化にする：

$$\sum_{k=1}^n \left\| X_k - \frac{1}{2} \begin{bmatrix} \rho + \delta & \rho - \delta \\ \rho - \delta & \rho + \delta \end{bmatrix} X_{k-1} - m_\varepsilon \right\|^2 \quad (5.16)$$

$$\text{推定値} \quad \begin{cases} \hat{\rho}_n = \frac{\sum_{k=1}^n \langle \mathbf{1}, x_k - m_\varepsilon \rangle \langle \mathbf{1}, x_{k-1} \rangle}{\sum_{k=1}^n \langle \mathbf{1}, x_{k-1} \rangle^2}, \\ \hat{\delta}_n = \frac{\sum_{k=1}^n \langle \mathbb{1}, x_k - m_\varepsilon \rangle \langle \mathbb{1}, x_{k-1} \rangle}{\sum_{k=1}^n \langle \mathbb{1}, x_{k-1} \rangle^2}. \end{cases} \quad (5.17)$$

ここでは

$$\mathbf{1} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \in R^2, \mathbb{1} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \in R^2 \text{ であり、}$$

$$H_n = \left\{ (x_1, \dots, x_n) \in (R^2)^n : \sum_{k=1}^n \langle \mathbf{1}, x_{k-1} \rangle^2 > 0 \right\}$$

$$\mathbb{H}_n = \left\{ (x_1, \dots, x_n) \in (R^2)^n : \sum_{k=1}^n \langle \mathbb{1}, x_{k-1} \rangle^2 > 0 \right\}$$

$X_0 = 0$ は R^2 空間零ベクトル。

$$\rho = \alpha + \beta, \quad \delta = \alpha - \beta \text{ より } \alpha = \frac{1}{2}(\rho + \delta), \quad \beta = \frac{1}{2}(\rho - \delta).$$

$$\begin{bmatrix} \hat{\alpha}_n \\ \hat{\beta}_n \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} \hat{\rho}_n \\ \hat{\delta}_n \end{bmatrix} \quad (5.18)$$

次の論文ではさらに重要な結論がある。

Asymptotic behavior of CLS estimators for 2-type doubly symmetric critical Galton-Watson processes with immigration

By Marton Ispany, Kristof Kormendi, & Gyula Pap

定理 1

$$n(\hat{\rho}_n - 1) \xrightarrow{D} \frac{\int_0^1 Y_t d(Y - \langle 1, m_\varepsilon \rangle t)}{\int_0^1 Y_t^2 dt}, \quad n \rightarrow \infty \quad (5.19)$$

If $\langle \bar{V}_\xi 1 \rangle = 0$ そして

$$n^{3/2}(\hat{\rho}_n - 1) \xrightarrow{D} N\left(0, \frac{3\langle \bar{V}_\xi 1 \rangle}{\langle 1, m_\varepsilon \rangle^2}\right), \quad n \rightarrow \infty \quad (5.20)$$

$f\langle \bar{V}_\xi t \rangle > 0$ そして

$$\begin{bmatrix} n^{1/2}(\hat{\alpha}_n - \alpha) \\ n^{1/2}(\hat{\beta}_n - \beta) \end{bmatrix} \xrightarrow{D} \sqrt{\alpha\beta} \frac{\int_0^1 y_t dW_t^\circ}{\int_0^1 y_t dt} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \quad n \rightarrow \infty \quad (5.21)$$

$(W_t^\circ)_{t \in \mathbb{R}^+}$ が標準ブラウン運動であり、 $(W_t)_{t \in \mathbb{R}^+}$ と独立である。

$$\begin{bmatrix} n^{1/2}(\hat{\alpha}_n - \alpha) \\ n^{1/2}(\hat{\beta}_n - \beta) \end{bmatrix} \xrightarrow{D} N\left(0, \frac{3\langle \bar{V}_\xi t \rangle}{4E(\langle t, \varepsilon_1 \rangle^2)}\right) \begin{bmatrix} 1 \\ -1 \end{bmatrix} \quad n \rightarrow \infty \quad (5.22)$$

5.2.2 CLS より移民項のある2変量分枝過程の推定のシミュレーション
ポアソン分布を利用し、3.2.2の理論をシミュレーションで検証する

$\xi_{k,j,i} = \begin{bmatrix} \xi_{k,j,i,1} \\ \xi_{k,j,i,2} \end{bmatrix}$, $\varepsilon_k = \begin{bmatrix} \varepsilon_{k,1} \\ \varepsilon_{k,2} \end{bmatrix}$ を観測される部分と考える。

$\xi_{k,j,i}$ と ε_k それぞれ期待値 m_ξ と m_ε のポアソン分布に従う。

$$\xi_{k,j,i,1} \sim P_0(m_{\xi_1}) \Rightarrow \xi_{k,j,1,1} \sim P_0(\alpha), \xi_{k,j,2,1} \sim P_0(\beta)$$

$$\varepsilon_{k,1} \sim P_0(\varepsilon_{k1})$$

$$\xi_{k,j,i,2} \sim P_0(m_{\xi_2}) \Rightarrow \xi_{k,j,1,2} \sim P_0(\beta), \xi_{k,j,2,1} \sim P_0(\alpha)$$

$$\varepsilon_{k,2} \sim P_0(\varepsilon_{k2})$$

$X_k = \sum_{j=1}^{X_{k-1,1}} \xi_{k,j,1} + \sum_{j=1}^{X_{k-1,2}} \xi_{k,j,2} + \varepsilon_k$ が期待値 $X_{k-1}m_{\xi_1} + X_{k-2}m_{\xi_2} + m_\varepsilon$ のポアソン分布に従

うことがわかる：

$$X_{k,1} \sim P_0(X_{k-1,1}\alpha + X_{k-1,2}\beta + \varepsilon_{k1})$$

$$X_{k,2} \sim P_0(X_{k-1,1}\beta + X_{k-1,2}\alpha + \varepsilon_{k2})$$

統計ソフト R を利用しシミュレーションを行う。

	m	n	m_{ε_1}	m_{ε_2}	α	β	ρ	δ	$\hat{\alpha}_n$	$\hat{\beta}_n$
第1回	10	10	2	3	0.8	0.2	0.7920	0.4754	0.6337	0.1583
第2回	100	100	2	3	0.75	0.25	0.9152	0.4715	0.6933	0.2218
第3回	100	1000	2	3	0.9	0.1	0.9227	0.7066	0.8096	0.1031

数理モデルは、自然科学においてのみではなく、社会科学や人文科学においても用いられる。数理モデルが用いられている分野を網羅することは難しいと考えられるが、例えば、物理学、工学、生物学、経済学、社会学、心理学、計算機科学、生態学、神経科学、分子生物学、生物統計学、免疫学、地球物理学、天文学、電気回路、機械工学、航空工学、気象学、言語学、計量文献学、伝染病感染予測、オペレーションズ・リサーチなどがある。

近年はコンピュータの性能の向上により、複雑な数理モデルでもそのふるまいをシミュレーションによって見るができるため、様々な分野で用いられるようになってきている。数理モデルは、対象とする現象や、定式化の抽象度などによって様々なものがある。

対象となる現象が大規模で人手による解析が困難、あるいはナビエ-ストークス方程式のようにモデルの解を解析的に得られない場合は、コンピュータによるシミュレーションによって解を求める。代表的なアルゴリズムとして、オイラー法、ルンゲ=クッタ法、有限要素法、モンテカルロ法等がある。コンピュータの性能向上によって、扱える数理モデルの幅が大変広まった。上述したように、数理モデルを構築することによって得られることは、まずは現象の理解があげられる。また、数学的に表現することによって、扱いが容易になり、数学の知見を活用することができる。

終わりに

本論文をまとめるにあたり，横浜国立大学教授永井圭二先生には，研究全般にわたりさまざまな視点よりきめ細かい御指導，御指摘を頂きました。永井圭二先生には，著者が大学院入学時から，研究の進め方，現象の捉え方など数多くの御指導，御示唆を頂いただくとともに，論文作成も大変お世話になりました。心より御礼申し上げます。

本論文の研究内容については同研究室の万ーセイさん、許インさんの御助力をいただきました。まことにありがとうございました。

本研究を遂行するにあたり，以上のお名前を挙げた方以外にも横浜国立大学の先生の皆様から数多くの御指導，御鞭撻を頂きました。心より感謝の意を表します。

研究生生活をサポートして頂いた家族成員たちに、感謝しております。

第六章 参考文献

- [1] A. Wald(1945). Sequential Tests Of Statistical Hypotheses, Annals Of Mathematical Statistics, Vol. 16.
- [2] C. Z. Wei And J. Winnicki(1990). Estimation Of The Means In The Branching Process With Immigration, Annals Of Statistics, vol 18.
- [3] Keiji Nagai, Kohtaro Hitomi And Yoshihiko Nishiyama(2007). Sequential Estimation Test For The Criticality Of Branching Processes.
- [4] K. Nagai, y. Nishiyama And K. Hitomi(2012). Sequential Unit Root Test.
- [5] William Feller(1951). Two Singular Diffusion Problem, Annals Of Mathematics, Vol. 54.
- [6] Kurtz, T. G. and P. Protter(1991). Weak Limit Theorems For Stochastic Integrals And Stochastic Differential Equations. The Annals Of Probability, vol. 19.
- [7] Marton Ispany, Kistof Kormendi And Gyula Pap. Asymptotic Behavior Of CLS Estimators For 2-type Doubly Symmetric Critical Galton-Watson Processes With Immigration.
- [8] Lai, T. L. And D. Siegmund(1983). Fixed Accuracy Estimation Of An Autoregressive Parameter, Annals Of Statistics 11.
- [9] S. カーリン著(1969), 佐藤健一、佐藤由身子訳, 確率過程講義, 産業図書
- [10] T. N. Sriram, I. V. Basawa and R. M. Huggins(1991), Sequential Estimation For Branching Processes With Immigration, Annals of Statistics, Vol 19.
- [11] Linetsky, V. (2004). The Spectral representation of Bessel Processes With Constant drift: Applications in Queueing And Finance, J. Appl. Probab., vol41.
- [12]<http://www.math.is.tohoku.ac.jp/~obata/lecture/2010-graduate/2010-in-09.pdf>
- [13] keisan.casio.jp/exec/system/1161228683

