

2011 年度後期

確率モデル論 (情報科学研究科)
応用解析学 (工学研究科)
確率モデル論 (国際高等研究教育院)

- 授業科目の目的・概要及び達成目標等

自然科学・生命科学をはじめ人文社会科学に至るまで、ノイズ・ゆらぎ・乱雑さ・不確定さから逃れられない現象には枚挙にいとまがなく、そのようなランダム現象の数理解析はますます重要になってきている。本講義では、確率論の基本的な考え方になじみながら、確率モデルの構成と解析手法を学ぶ。特に、時間発展を含むランダム現象を記述する確率過程としてマルコフ連鎖の基本的事項を学び、その幅広い応用を概観する。

- 目次

1. 序論
2. 確率変数と確率分布
3. ベルヌイ試行列
4. 大数の法則と中心極限定理
5. ランダム・ウォーク
6. マルコフ連鎖
7. ゴルトン-ワトソン分枝過程
8. 出生死亡過程
9. ポアソン過程
10. 待ち行列

— 後半は、やや流動的。トピックスを選んで講義する予定。

- 成績

授業中に出题する問題（レポート問題として通し番号がつく。単なる「問」ではない）から数題を選択してレポートを作成し、1 月後半に提出してもらう予定。詳細は後日発表する（コピーレポートは零点）。

- この講義は拙著「確率モデル序説 (仮題)」(出版遅れ) にもとづく。原稿の一部はダウンロード可能:

www.math.is.tohoku.ac.jp/~obata/ から「大学院講義」のページへ飛ぶ

— ウェブページには、関連する講義録・その他の資料も掲載されている。

— 講義では、簡単なレジュメだけ配布する。

- 担当者連絡先

尾畑伸明 (情報科学研究科・システム情報科学専攻) obata@math.is.tohoku.ac.jp

応用数学連携フォーラム <http://www.dais.is.tohoku.ac.jp/~amf/>

数学と諸分野との連携, ワークショップの開催 (来聴歓迎! 第 22 回 10.6; 第 23 回 10.20; 続く)

- 休講 10.13 (木) 11.10 (木)

● 参考書として、たとえば、

1. 国沢清典：確率論とその応用 (岩波全書), 1982.

本講義はこのレベルをめざす。少し古いので入手困難かも。

2. W. Feller: An Introduction to Probability Theory and Its Applications, Vol. 1, Wiley, 1957.

名著の誉れ高い。この本は講義内容をカバーし、さらに詳しいことがたくさん書かれている (Vol. 2 もある!)。邦訳もある。

W. フェラー (河田龍夫他訳)：確率論とその応用 (紀伊国屋)。こちらは4分冊。

3. B. V. Gnedenko: The Theory of Probability and the Elements of Statistics, AMS Chelsea Publishing Co., 6th ed. 1989.

4. R. Durrett: Probability: Theory and Examples, Duxbury Press, 1996.

この2冊も講義内容をカバーし、さらに詳しいことがたくさん書かれている。この程度の知識があれば、確率モデルを本格的に研究に生かせるだろう。

5. 佐藤坦：はじめての確率論 測度から確率へ (共立), 1994.

コルモゴロフの公理から出発して確率論の基礎を数学的に厳密に展開している。特に、ルベグ積分が気になる人は一読してみよ。ただし、確率モデルという観点は希薄 (この本を読んだから、適当な本を見てねというスタンス)。

6. 舟木直久：確率論, 朝倉書店, 2004.

7. 西尾真喜子：確率論, 実教出版, 1978.

この2冊はさらに高度なところまで数学理論として展開している。

8. 志賀徳造：ルベグ積分から確率論 (共立), 2000.

前半はルベグ積分を展開しているが、後半でランダムウォークを取り扱って確率モデルへの入門をはたす。

9. P. ブレモー (釜江哲朗監修, 向井久訳)：モデルで学ぶ確率入門 (新装版), シュプリンガー東京, 2004.

実用の場面を想定したさまざまな確率モデルが取り上げられている。例題を通して数学的な枠組を学ぶ形式で書かれている。個々の事例は興味深いが、初学者が理論を学ぶには重い。

10. 鳥脇純一郎：工学のための確率論 (オーム社), 2002.

11. 拙著：確率統計要論 (牧野書店), 2007.

この2冊は入門的かつ初等的。このレベルを超えて [1,2] のレベルを目指してほしい。

第1章 序論

1.1 ランダム現象の確率モデル

偶然現象 (ランダム現象) では, 偶然の効果によって結果に揺らぎ (ばらつき, あるいは曖昧さ) を伴うため, 確定的な予測は困難である. その揺らぎの統計的性質をうまく利用して確率モデルを構築し, 偶然現象の諸性質を理論的に導くことが重要である. 揺らぎの統計的性質は, 同一条件下で繰り返して実験または観測して多数のデータを収集することで初めて明らかになる. したがって, それが原理的にできない 1 回限りの予測不可能な現象は, 残念ながら確率論の対象ではない.

たとえば, 日常の動作としてコイン投げを考えてみよう. 空中に投げ上げたコインを受け止めたとき, 表が出るか裏が出るかは一定しないだろう. ところが, コイン投げを多数回繰り返すと, 表の出る相対頻度と裏の出る相対頻度は, とともに $1/2$ に近くなるという統計性が確認される.¹ ここで, 日常の動作としてのコイン投げは, 同一の条件下での繰り返し実験と言えるのか, 厳格に同一条件を設定すれば一定の結果が得られるのではないかと, という批判が出るだろう. 至極もつともである. しかし, 真に厳格な同一条件の設定ができないのであれば, 当然, 結果にゆらぎが生ずる. そうすれば, 表が出る相対頻度が 98%, 裏の出る相対頻度が 2% といった統計性がやはり観察されるだろう. 一般に, 実験や観測には制御できない誤差がつきものであり, そのため観測結果にはゆらぎがともなう. このような事例も偶然現象として確率モデルの対象になる.

本来は偶然現象ではないが, 偶然現象として扱うことで有用な情報を得るという方向性もある. 容器に閉じ込められた気体分子は弾性衝突を繰り返しながら運動している. 容器の壁に衝突する分子が圧力を発生させている. この圧力を計算のために, 分子の運動方程式を解くことは現実的ではない. 容器に閉じ込められた分子の個数は非常に多いため, 平均値とそのまわりの揺らぎという形で偶然現象のように扱うのが極めて有効なのである. (まさに, 統計物理学は確率モデルの実験場なのである.)

さらに, 興味深い研究対象は, 時間とともに変化する偶然現象である. いわゆる時系列データには複雑な挙動を示すものがいくらかでも見つかるだろう. 過去の経過があって今の状態にあるのだが, 次の瞬間にどのように動くか確定的なルールはなく, 多数の可能性の中から確率的に選ばれていると想定されるものは確率過程によってモデル化される. 拡散現象・感染・生物個体数・太陽黒点・気象・株価や為替などは時間とともに複雑に変動するが, その原因は多種多様であり, それらの連関も明らかではない. このような場合は, 原因の詳細を考慮した解析はしばしば現実的ではなく, 変動の統計性に着目した確率モデルが数理解析の鍵となるのである.

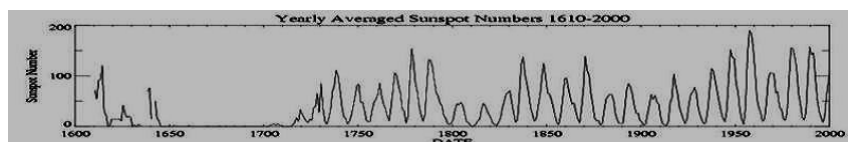


図 1.1: 太陽黒点

¹たとえば, 数理統計学の創設に大きな貢献のあったピアソン (K. Pearson) のような著名な数学者もコイン投げ実験をしたことが伝えられている.

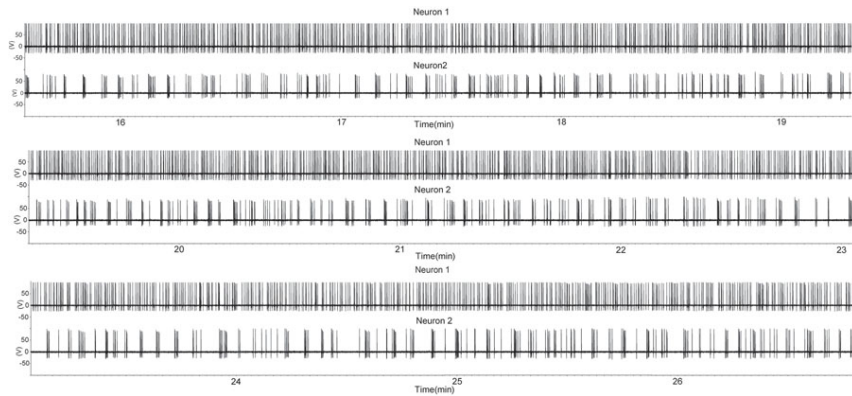


図 1.2: ニューロンの発火

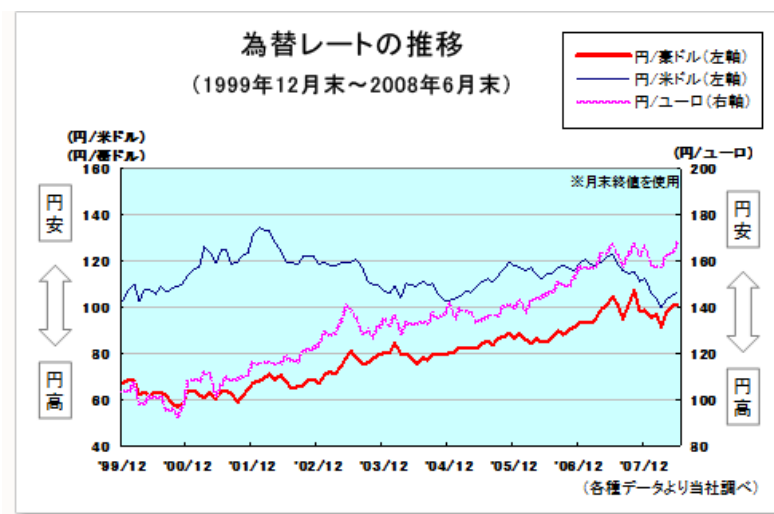


図 1.3: 為替変動

1.2 目標となる課題

例 1.2.1 (ギャンブラーの破産問題) A, B の 2 人が公平なコインによるゲームをする。それぞれの持ち点を A, B として、1 回のコイン投げ毎に、勝者は相手から 1 点を受け取るものとする。一方の持ち点が 0 点になった段階でゲームを終了し、勝者は $A + B$ 点を獲得する。 A, B がそれぞれ勝利する確率を求めよ。また、ゲーム終了までに要するコイン投げの平均回数はどのくらいか？

例 1.2.2 (ランダム・ウォーク) 1 本道を行き来する酔っ払いをモデル化したものが (1 次元) ランダム・ウォークである。原点から出発したランダム・ウォーカーは、再び原点に戻るだろうか？ 原点に戻る確率を求めて答えよ。

例 1.2.3 (つき) A, B の 2 人が公平なコインによるゲームをする。1 回のコイン投げ毎に、勝者は相手から 1 点を受け取るものとする。 A, B とも初めの持ち点は 0 点とする。一晩中、このゲームをするとき、持ち点がつねに正になる (浮きっぱなし) のと、持ち点が正負半々 (浮いたり沈んだり) になるのとどちらが起きやすいか？ 確率を計算して答えよ。

例 1.2.4 (家系の存続) ある家系では、生まれた男子のみが姓を継承する。どのような状況の下に、家系が存続するだろうか？ 各世代の男子が生み出す後継男子の人数はランダムであるとして、この家系が消滅する確率を求めよ。

例 1.2.5 (待ち時間) 多くのメールを受信するが、送信元は気まぐれに互いに無関係に送信しているものと思われる。1回のメール着信の後、次のメールを着信するまでの待ち時間はどのようになるか？

1.3 確率変数

ある範囲を動くことが想定されている実数のことを数学用語で (実) 変数という。文字の使い方は自由ではあるが、変数には $x, y, z, t, \dots$ などがよく用いられる。たとえば、 $0 \leq x \leq 1$ とあれば、 x は 0 以上 1 以下の範囲にある実数を代表する変数と理解される。偶然現象を観測して得られる観測値は、偶然の影響のために定数ではなく、ある範囲を動く変数となる。しかも、どの値が出やすく、どの値が出にくいといった確率的な傾向が付加されている。このような変数を確率変数という。習慣によって、確率変数には $X, Y, Z, T, \dots$ のように大文字を用いる。たとえば、サイコロ投げで得られる観測値を X とすると、 X は $\{1, 2, 3, 4, 5, 6\}$ の範囲を動く変数であるが、さらに確率が付加されていることはすぐに納得されよう。ふつつ

$$P(X = 1) = P(X = 2) = \dots = P(X = 6) = \frac{1}{6}$$

のように書くことで、値の出やすさの確率的な傾向を明らかにする。一方、偶然現象をひとたび観測すれば確定した 1 つの観測値 x が得られる。これを確率変数 X の実現値といい、確率変数そのものとは区別する。つまり、確率変数 X は偶然現象における観測値のゆらぎ方の情報をすべて持ち合わせた変数であり、実現値とは X の取りうる個々の値のことをさす。

一般には、観測データは数値 (実数) ばかりではなく、複数の数値の組 (ベクトル)、あるいは音声や画像などと様々であるから、想定される観測データのなす集合 S をあらかじめ決めておいて、 S -値確率変数を考えるのが好都合である。この S を状態空間という。つまり、偶然現象では複数の状態が可能であり、それぞれの状態が出現する確率が与えられている。

1.4 確率過程

時間とともに変化するランダム現象において、ある時刻 t において観測された値を確率変数の実現値ととらえるのが我々の出発点である。そうすれば、時間とともに変動するランダム現象に対しては、時間パラメータ t をもつ確率変数の系列 $\{X(t)\}$ 、つまり確率過程を考えることになる。

時間パラメータのとり方に 2 通りある。1 ステップごとに進行するランダム現象は確率変数の列 $\{X_n; n = 1, 2, \dots\}$ でモデル化される。これを離散時間の確率過程という。連続的に流れる時間経過とともに変化するランダム現象であっても、それを一定時間ごとに観測する場合は、離散時間の確率過程でモデル化される。一方、時間を連続的に流れるものととらえれば、連続時間の確率過程 $\{X(t); 0 \leq t \leq T\}$ が考察の対象になる。現実世界では、時間間隔 Δt がいかほどに小さいにせよ、離散時間的な観測しかできないので、離散時間の確率過程を考えておけば十分であるとの考え方もあるが、連続時間の確率過程に対しては、確率解析と呼ばれる無限変数の微積分が適用できるため、その応用範囲は非常に広い。連続量を扱う微積分が強力な計算手段であることは経験済みであろう。

例 1.4.1 (ベルヌイ試行列) 同一のコインを繰り返し投げる試行において、 n 回目のコイン投げにあたる確率変数を Z_n とおく。その確率法則は一般性をもたせて、 $P(Z_n = 1) = p, P(Z_n = 0) = q$ とする。ただ

し, $0 \leq p, q \leq 1$ は $p + q = 1$ を満たす定数である. さらに, コイン投げは互いに影響を及ぼさないことを反映させて, 確率変数列 $\{Z_1, Z_2, \dots\}$ は独立であると仮定する. つまり,

$$P(Z_1 = \xi_1, Z_2 = \xi_2, \dots, Z_n = \xi_n) = \prod_{k=1}^n P(Z_k = \xi_k)$$

がすべての $\xi_1, \xi_2, \dots, \xi_n \in \{0, 1\}$ について成り立つ. このような確率変数列 (= 離散時間確率過程) を成功確率 p のベルヌイ試行列と呼ぶ.

例 1.4.2 (ランダムウォーク) コイン投げをしながら直線上を運動する動点 (ランダムウォーカーと呼ぼう) を考えよう. このランダムウォーカーは, 時刻 $t = 0$ において原点 0 から出発し, 各時刻でコインを投げて, 表 (H) が出たら正の方向に 1 だけ移動し, 裏 (T) が出たら負の方向に 1 だけ移動するものとする. 図 1.4 は, そのような動きの一例である.

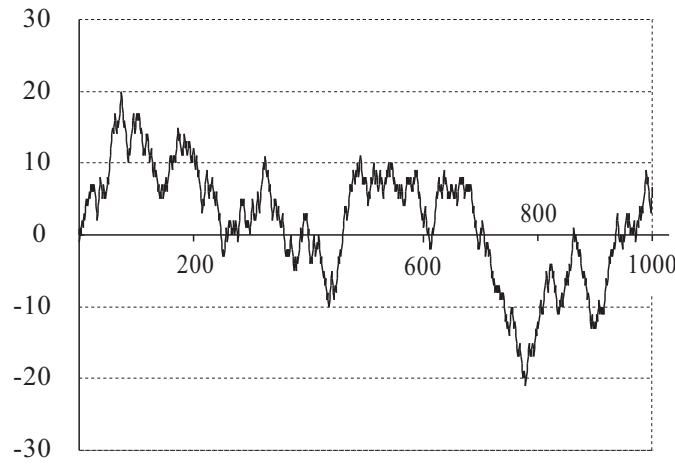


図 1.4: ランダム・ウォークのシミュレーション ($p = q = 1/2$)

時刻 $t = n$ におけるランダムウォーカーの位置は, n 回のコイン投げの結果として定まるので確率変数である. これを X_n と書こう. X_n を簡潔に記述するために, ベルヌイ試行列 $\{Z_n\}$ を用いるのがよい. ただし, Z_n は値として ± 1 を取るものとし,

$$P(Z_n = 1) = p, \quad P(Z_n = -1) = q = 1 - p,$$

とおく. そうすれば,

$$X_n = \sum_{k=1}^n Z_k \tag{1.1}$$

のように表される. こうして得られた離散時間確率過程 $\{X_n\}$ を 1 次元ランダム・ウォークという. X_n の分布は,

$$P(X_n = 2k - n) = \binom{n}{k} p^k q^{n-k}, \quad k = 0, 1, 2, \dots, n, \tag{1.2}$$

で与えられる. つまり, X_n の分布は本質的に二項分布である.

第2章 確率変数と確率分布

2.1 確率変数と状態空間

ある範囲を動くことが想定されている実数のことを数学用語で (実) 変数という。文字の使い方は自由ではあるが、変数には $x, y, z, t, \dots$ などがよく用いられる。たとえば、 $0 \leq x \leq 1$ とあれば、 x は 0 以上 1 以下の範囲にある実数を代表する変数と理解される。偶然現象を観測して得られる観測値は、偶然の影響のために定数ではなく、ある範囲を動く変数となる。しかも、どの値が出やすく、どの値が出にくいといった確率的な傾向が付加されている。このような変数を確率変数という。習慣によって、確率変数には $X, Y, Z, T, \dots$ のように大文字を用いる。たとえば、サイコロ投げで得られる観測値を X とすると、 X は $\{1, 2, 3, 4, 5, 6\}$ の範囲を動く変数であるが、さらに確率が付加されていることはすぐに納得されよう。ふ

$$P(X = 1) = P(X = 2) = \dots = P(X = 6) = \frac{1}{6}$$

のように書くことで、値の出やすさの確率的な傾向を明らかにする。一方、偶然現象をひとたび観測すれば確定した 1 つの観測値 x が得られる。これを確率変数 X の実現値といい、確率変数そのものとは区別する。つまり、確率変数 X は偶然現象における観測値のゆらぎ方の情報をすべて持ち合わせた変数であり、実現値とは X の取りうる個々の値のことをさす。

一般には、観測データは数値 (実数) ばかりではなく、複数の数値の組 (ベクトル)、あるいは音声や画像などと様々であるから、想定される観測データのなす集合 S をあらかじめ決めておいて、 S -値確率変数を考えるのが好都合である。この S を状態空間という。つまり、偶然現象では複数の状態が可能であり、それぞれの状態が出現する確率が与えられている。

2.2 確率変数の分布

2.2.1 離散型確率変数

確率変数 X の取りうる値が高々可算個¹であるとき、離散型確率変数という。 X のとり得る値の集合を $\{a_1, a_2, \dots\}$ として、 X が a_i をとる確率

$$P(X = a_i) = p_i$$

を与えることで確率変数 X が決まる。ここで、

$$\sum_i p_i = 1, \quad p_i \geq 0 \quad (2.1)$$

が成り立つ。もし $p_i = 0$ であれば、 X の取りうる値から a_i を除外してよい。しかし、理論上は $p_i = 0$ も許しておいた方が、例外に言及する手間が省けて便利である。

¹有限個または可算個。番号付け可能な無限集合 (自然数 $\mathbb{N}$ と同様な集合) を可算集合といい、その要素の個数が可算個。

離散型確率変数 X の統計的な情報は、実数列 $a_1, a_2, \dots$ とそこに乗っている加重 $p_1, p_2, \dots$ の組で尽きる。これを X の分布という。逆に、実数の有限または無限列 $a_1, a_2, \dots$ とそこに乗っている加重 $p_1, p_2, \dots$ で (2.1) を満たすものを与えれば、離散型確率変数を 1 つ定めたことになる。このとき、

$$\mu = \sum_i p_i \delta_{a_i}, \quad \text{または} \quad \mu(dx) = \sum_i p_i \delta(x - a_i) dx,$$

のように表記して、 $\mathbf{R}$ 上の離散分布という。

2.2.2 連続型確率変数

次に、確率変数 X の取りうる値が連続的に変化する場合を考えよう。一般に、確率変数 X が連続的な値をとる場合は、ある特定の値 a をとる確率 $P(X = a)$ は 0 になる。したがって、統計的な情報を取り出すためには、ある範囲に値をとる確率、たとえば $P(a \leq X \leq b)$ を決めなければならない。これが

$$P(a \leq X \leq b) = \int_a^b f(x) dx$$

のように、ある関数 $f(x)$ で表わされるとき、 $f(x)$ を確率変数 X の (確率) 密度関数という。明らかに、

$$\int_{-\infty}^{+\infty} f(x) dx = 1, \quad f(x) \geq 0, \quad (2.2)$$

が成り立つ。このとき、 X がある特定の値をとる確率は 0 であるから、左辺は $P(a < X < b)$ などに置き換えても同じである。この積分の一部を取り出して、

$$\mu(dx) = f(x) dx \quad (2.3)$$

と書いて確率変数 X の分布ということもある。逆に、関数 $f(x)$ で (2.2) をみたすものを与えれば、連続型確率変数 X を決めたことになる。

次節で触れるが、連続型確率変数は必ずしも密度関数をもつとは限らないが、応用上重要なものは密度関数であらわされる。

例 2.2.1 半径 R の円板からランダムに 1 点 A を選んだとき、中心 O からの距離 $X = \overline{OA}$ は $[0, R]$ に値をとる確率変数になる。

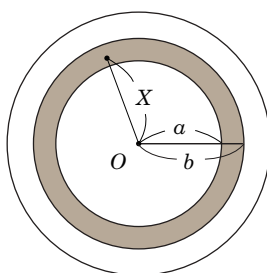


図 2.1: 単位円板からランダムに 1 点を選ぶ

この X の取り得る値は連続的に変化する。 $0 < a < R$ として $P(X = a)$ を考えてみよう。 $X = a$ となるのは、 A が中心 O 、半径 a の円周 $C(a)$ から選ばれたことを意味する。題意から、単位円板のどの点も同等に選ぶことになるから、その確率は面積比で与えるのが適当である。したがって、

$$P(X = a) = \frac{|C(a)|}{\pi R^2} = \frac{0}{\pi R^2} = 0$$

となる。つまり、円周 $C(a)$ には広がりがないので、円板からランダムに選ばれた 1 点が、たまたま $C(a)$ 上に載っている確率は 0 にならざるをえない。このことから、 $P(X = a)$ を与えることでは、確率変数 X の統計的な情報が何も得られないことがわかる。

そこで、少し幅を考えて $P(a \leq X \leq b)$ を与えることを考えてみる。 $0 \leq a < b \leq 1$ として、

$$P(a \leq X \leq b) = \frac{\pi b^2 - \pi a^2}{\pi R^2} = \frac{b^2 - a^2}{R^2}$$

となる。さらに、積分表示

$$P(a \leq X \leq b) = \frac{1}{R} \int_a^b 2x dx, \quad 0 < a < b < R,$$

が得られる。したがって、密度関数は、

$$f(x) = \begin{cases} 0, & x \leq 0, \\ 2x, & 0 \leq x \leq 1, \\ 0, & x > 1, \end{cases}$$

となる。密度関数は積分を通して確率を与えることが本質的な役割であるので、不連続点 $x = 1$ の値はどのように定めても (定めなくても) かまわない。

2.2.3 分布関数

確率変数は離散型または連続型の 2 つに分類されるわけではなく、連続型であっても密度関数をもつとは限らない。統一的な扱いのためには、分布関数が必要である。

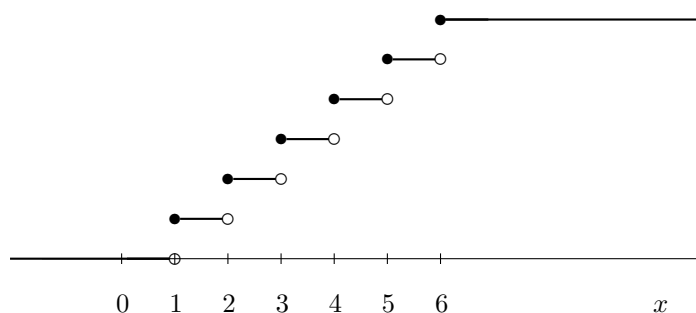
定義 2.2.2 X を確率変数とすると、

$$F(x) = F_X(x) = P(X \leq x), \quad x \in \mathbf{R},$$

を X の (確率) 分布関数 ((probability) distribution function, pdf) という。 (分布関数は全実数 $\mathbf{R}$ を定義域とする実数値の関数になる。)

例 2.2.3 サイコロを 1 個振り、出た目を X とする。 X の分布関数は、

$$F(x) = \begin{cases} 0 & x < 1 \\ 1/6 & 1 \leq x < 2 \\ \vdots & \vdots \\ 5/6 & 5 \leq x < 6 \\ 1 & 6 \leq x \end{cases}$$



確率変数が有限個の値, または離散的な値をとる場合でも, 分布関数 $F(x)$ の x は実数を連続的に動くことに注意しよう.

例 2.2.1 では,

$$F_X(x) = \begin{cases} 0, & x < 0, \\ x^2, & 0 \leq x \leq 1, \\ 1, & x > 1 \end{cases}$$

がわかる. 実際, $x < 0$ と $x \geq 1$ では明らかであり, $0 \leq x \leq 1$ では,

$$F_X(x) = P(X \leq x) = \frac{\pi x^2}{\pi} = x^2$$

となっている.

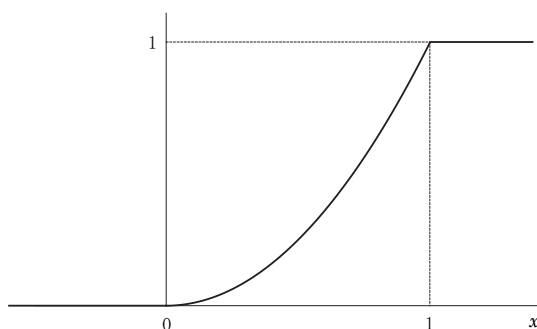


図 2.2: 例 2.2.1 の分布関数

分布関数の例を 2 つ見たが, そこに見られる共通の性質によって, 分布関数を直接的に (確率変数とは無関係に) 定義する.

定義 2.2.4 実変数の実数値関数 $F(x)$ で次の性質をもつものを分布関数という.

- (1) (単調増加) $x_1 \leq x_2$ ならば $F(x_1) \leq F(x_2)$.
- (2) $\lim_{x \rightarrow -\infty} F(x) = 0, \quad \lim_{x \rightarrow +\infty} F(x) = 1$.
- (3) (右半連続) $\lim_{\epsilon \downarrow 0} F(x + \epsilon) = F(x)$.

確率変数 X の分布関数 $F_X(x)$ は定義 2.2.4 の分布関数である. つまり, そこに述べた性質 (i)–(iii) を満たす. 逆に, 定義 2.2.4 の意味の分布関数 $F(x)$ によって確率変数が 1 つ定義されることになる.

注意 2.2.5 分布関数は実験や統計調査で得られたデータから作った累積頻度分布に相当する.

定義 2.2.6 分布関数 $F(x)$ は跳躍 (ジャンプ) のみで増加するとき離散的であるといい, 分布関数が連続であるときは連続的であるという. そして, 対応する確率変数 X をそれぞれ離散型確率変数または連続型確率変数という.

定義 2.2.7 分布関数 $F(x)$ が区分的に微分可能であるとき, その導関数

$$f(x) = F'(x)$$

を $F(x)$ の (確率) 密度関数という.

密度関数は次の性質で特徴づけられる.

$$f(x) \geq 0, \quad \int_{-\infty}^{+\infty} f(x)dx = 1.$$

このとき,

$$F(x) = \int_{-\infty}^x f(t)dt$$

が成り立ち, 確率変数 X が区間 $[a, b]$ に値をとる確率は

$$P(a \leq X \leq b) = F(b) - F(a) = \int_a^b f(t)dt, \quad a < b$$

で与えられる.

離散的な分布 $F(x)$ に対して密度関数を考えることはできない. 実際,

$$P(X = a) = F(a) - F(a - 0)$$

が成り立ち, 分布関数の不連続点でのジャンプ量が確率を表す.

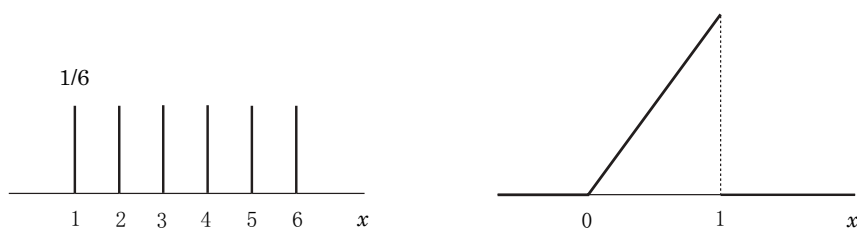


図 2.3: 密度関数 (左図は離散型の場合のイメージ図)

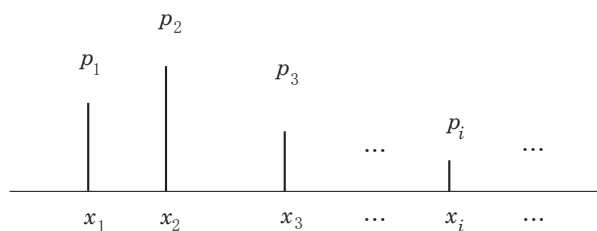
2.3 確率変数の特性量

一般に, 確率変数 X の確率分布を $\mu(dx)$ のような記号で表そう.

- X が離散型の場合は,

$$\mu_X = \sum_n p_n \delta_{x_n} \quad \text{あるいは} \quad \mu_X(dx) = \sum_n p_n \delta(x - x_n)dx$$

を意味する. 実軸上の各点 $x_1, x_2, \dots$ に荷重 $p_1, p_2, \dots$ が載っているイメージである.



- X が連続型で密度関数 $f(x)$ をもつ場合は,

$$\mu(dx) = f(x)dx$$

を意味する.

定 義 2.3.1 X の平均値 (mean value, expectation) を

$$\mathbf{E}[X] = \int_{-\infty}^{+\infty} x\mu(dx)$$

で定義する. (確率変数 X は様々な値をある確率分布をもってとるのだが, その「平均」あるいは「重心」にあたるのが平均値である.)

X が離散的な確率変数ならば,

$$\mathbf{E}[X] = \sum_i x_i p_i$$

となる. また, X が連続型確率変数で密度関数 $f(x)$ をもてば,

$$\mathbf{E}[X] = \int_{-\infty}^{+\infty} x f(x) dx$$

として平均値が計算される.

さらに, $\varphi(x)$ を実関数とすると,

$$\mathbf{E}[\varphi(X)] = \int_{-\infty}^{+\infty} \varphi(x)\mu(dx)$$

となる. たとえば,

$$\mathbf{E}[X^m] = \int_{-\infty}^{+\infty} x^m \mu(dx)$$

を m 次のモーメントという.

定 義 2.3.2 確率変数 X の分散 (variance) を

$$\mathbf{V}[X] = \mathbf{E}[(X - \mathbf{E}[X])^2] = \mathbf{E}[X^2] - \mathbf{E}[X]^2$$

で定義する. 同じことであるが,

$$\mathbf{V}[X] = \int_{-\infty}^{+\infty} (x - \mathbf{E}[X])^2 \mu(dx) = \int_{-\infty}^{+\infty} x^2 \mu(dx) - \left(\int_{-\infty}^{+\infty} x \mu(dx) \right)^2$$

としてもよい.

2.4 重要な離散分布

例 2.4.1 (ベルヌイ分布) 表の出る確率が p であるようなコインを投げて, 表が出たら 1, 裏が出たら 0 とする確率変数の分布は,

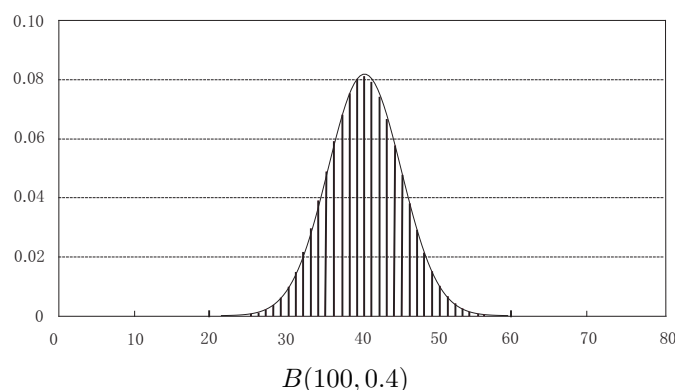
$$(1-p)\delta_0 + p\delta_1$$

で与えられる. これを成功確率 p のベルヌイ分布 (Bernoulli distribution) という. 平均値は p , 分散は $p(1-p)$ となることを示せ.

例 2.4.2 (二項分布) $0 \leq p \leq 1, n \geq 1$ とするとき,

$$\sum_{k=0}^n \binom{n}{k} p^k (1-p)^{n-k} \delta_k$$

で与えられる分布を二項分布 (binomial distribution) といい, $B(n, p)$ で表わす.



表の出る確率が p であるようなコインを n 振ったとき, 表の出る回数を X とすれば, X の分布が $B(n, p)$ となる. 二項分布 $B(n, p)$ の平均値は np , 分散は $np(1-p)$ である.

例 2.4.3 (幾何分布) $0 \leq p \leq 1$ を定数として,

$$\sum_{k=1}^{\infty} p(1-p)^{k-1} \delta_k$$

で定まる離散分布をパラメータ p の幾何分布 (geometric distribution) という.² 表の出る確率が p であるようなコインを投げるとき, 初めて表が出るまでに要する試行回数 (表が出た試行も含める) の分布である. パラメータ p の幾何分布の平均値は $\frac{1}{p}$, 分散は $\frac{1}{p^2}$ である.

例 2.4.4 (ポアソン分布) $\lambda > 0$ とするとき,

$$\sum_{k=0}^{\infty} e^{-\lambda} \frac{\lambda^k}{k!} \delta_k$$

で与えられる分布をパラメータ λ のポアソン分布 (Poisson distribution) という. パラメータ λ のポアソン分布の平均値は λ , 分散も λ である.

2.5 重要な連続分布

例 2.5.1 (一様分布) $a < b$ とする.

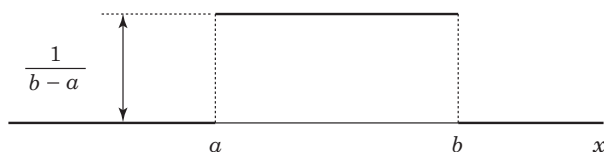
$$f(x) = \begin{cases} \frac{1}{b-a}, & a \leq x \leq b, \\ 0, & \text{その他} \end{cases}$$

²幾何分布の定義は, 本によっては, k の変域を $\{0, 1, 2, \dots\}$ にとって,

$$\sum_{k=0}^{\infty} p(1-p)^k \delta_k$$

を採用している. これは, 表の出る確率が p であるようなコインを投げるとき, 初めて表が出るまでに表の出た回数の分布である. この流儀では, 負の二項分布との関係などが見やすくなる.

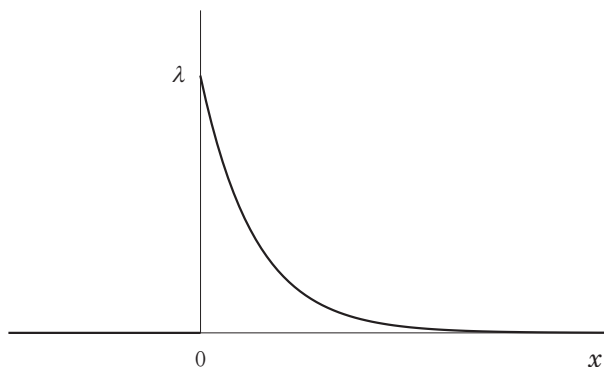
を密度関数とする確率分布を区間 $[a, b]$ 上の一様分布 (uniform distribution) という. 区間 $[a, b]$ 上の一様分布の平均値は $\frac{a+b}{2}$, 分散は $\frac{(b-a)^2}{12}$ となる.



例 2.5.2 (指数分布) $\lambda > 0$ とする.

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0, \\ 0, & \text{その他} \end{cases}$$

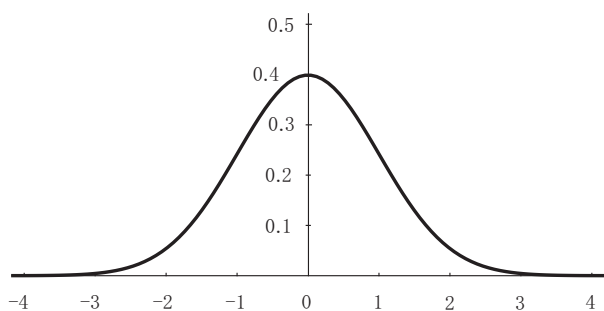
を密度関数とする確率分布をパラメータ λ の指数分布という. パラメータ λ の指数分布の平均値は λ^{-1} , 分散は λ^{-2} となる.



例 2.5.3 (正規分布) $\sigma > 0, m \in \mathbf{R}$ とするとき,

$$\frac{1}{\sqrt{2\pi}\sigma^2} \exp \left\{ -\frac{(x-m)^2}{2\sigma^2} \right\} \quad (2.4)$$

を密度関数とする確率分布を, 平均 m , 分散 σ^2 の正規分布 (normal distribution) またはガウス分布 (Gaussian distribution) といい, $N(m, \sigma^2)$ で表わす. 特に, $N(0, 1)$ を標準正規分布という.



議論としては、まず、(2.4) が密度関数になっていることを確認する。次に、積分を計算して、

$$m = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{+\infty} x \exp \left\{ -\frac{(x-m)^2}{2\sigma^2} \right\} dx,$$

$$\sigma^2 = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{+\infty} (x-m)^2 \exp \left\{ -\frac{(x-m)^2}{2\sigma^2} \right\} dx$$

を示し、初めて m が平均値で σ^2 が分散であることが分かる。

レポート問題 1 半径 1 の円盤からランダムに 1 点を選び、その点を中心とする内接円の半径を X とする。

- (1) X の分布関数, 確率密度関数, 平均値, 分散を求めよ。
- (2) 内接円の面積 $S = \pi X^2$ の平均値, 分散を求めよ。

レポート問題 2 長さ L の棒をランダムに折って、2 本の断片を作る。このとき、長い方の断片の長さを X とするとき、 X の分布関数, 密度関数, 平均値, 分散を求めよ。

レポート問題 3 a, b を異なる実数, $0 \leq p \leq 1$, $q = 1 - p$, に対して 2 点分布が

$$p\delta_a + q\delta_b$$

で定義される。平均 0, 分散 1 となるような 2 点分布を求めよ。

レポート問題 4 (確率母関数) $\{0, 1, 2, \dots\}$ にのみ荷重が乗っている確率分布 $\mu = \sum_{k=0}^{\infty} p_k \delta_k$ に対して、

$$G(z) = \sum_{k=0}^{\infty} p_k z^k$$

を μ の確率母関数という。

- (1) μ の平均値 m と分散 σ^2 は確率母関数を用いて

$$m = G'(1), \quad \sigma^2 = G''(1) + G'(1) - \{G'(1)\}^2$$

で与えられることを示せ。

- (2) 二項分布の確率母関数を求めよ。
- (3) 二項分布の平均値と分散を確率母関数を用いて導出せよ。

レポート問題 5 (1) 公式

$$\int_0^{+\infty} e^{-tx^2} dx = \frac{\sqrt{\pi}}{2\sqrt{t}}, \quad t > 0$$

の両辺を t で微分することによって、

$$\int_0^{+\infty} x^{2n} e^{-x^2} dx$$

を求めよ。ただし、 $n = 0, 1, 2, \dots$ とする。

- (2) $f(x) = ax^2 e^{-x^2}$ が確率密度関数になるように定数 a を定め、その確率分布の平均と分散を求めよ。

第3章 ベルヌイ試行列

3.1 コイン投げの確率モデル

コイン投げは最も基本的な確率モデルを提供する。コイン投げの結果は表裏の2通りであるが、慣例によって数値1,0に対応させることにする。 n 回目のコイン投げの結果を表す確率変数を Z_n で表せば、

$$P(Z_n = 1) = p, \quad P(Z_n = 0) = q = 1 - p, \quad 0 < p < 1, \quad (3.1)$$

となる。ここで、 p はこのコインの表の出る確率である。同じコインを同じ状況で投げ続けることを想定して、 p は n によらない定数とする。言い換えると、 Z_n は成功確率 p のベルヌイ型確率変数ということである。

コインを2回投げる実験を繰り返すことで、

$$\begin{aligned} P(Z_1 = 1, Z_2 = 1) &= p^2 = P(Z_1 = 1)P(Z_2 = 1), \\ P(Z_1 = 1, Z_2 = 0) &= pq = P(Z_1 = 1)P(Z_2 = 0), \\ P(Z_1 = 0, Z_2 = 1) &= qp = P(Z_1 = 0)P(Z_2 = 1), \\ P(Z_1 = 0, Z_2 = 0) &= q^2 = P(Z_1 = 0)P(Z_2 = 0), \end{aligned}$$

が高精度で確認できる。コイン投げの回数を増やしても同様であり、

$$P(Z_1 = \xi_1, Z_2 = \xi_2, \dots, Z_n = \xi_n) = \prod_{k=1}^n P(Z_k = \xi_k) \quad (3.2)$$

がわかる。したがって、コイン投げの確率モデルとして単に成功確率 p のベルヌイ型確率変数列 $Z_1, Z_2, \dots$ を考えるだけでは不十分であり、(3.2) を保証する関係性を規定しなければならない。以下で詳しく見てゆくが、この関係性が「確率変数の独立性」と呼ばれるものである。

こうして、コイン投げは、独立で同分布をもつベルヌイ型確率変数列 $Z_1, Z_2, \dots$ でモデル化されることになる。つまり、確率変数列として独立であり、すべての Z_n は同じ分布をもち、その分布は(3.1)で与えられる。これを成功確率 p のベルヌイ試行列という。結果が2通りであり、毎回の結果が他の回に影響を受けないような偶然現象はコイン投げに模され、ベルヌイ試行列でモデル化される。

3.2 事象の独立性

独立性は確率論で最も基本的な概念であるから、きちんと定義しておきたい。

定義 3.2.1 事象の(有限または無限)列 $A_1, A_2, \dots$ が互いに独立(pairwise independent)であるとは、そこから選んだ任意の2つの事象 A_{i_1}, A_{i_2} ($i_1 \neq i_2$) が

$$P(A_{i_1} \cap A_{i_2}) = P(A_{i_1})P(A_{i_2})$$

を満たすときにいう。

定義 3.2.2 事象の (有限または無限) 列 $A_1, A_2, \dots$ が独立 (independent) であるとは, そこから選んだ任意有限個の事象 $A_{i_1}, \dots, A_{i_n}$ ($i_1 < i_2 < \dots < i_n$) が

$$P(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_n}) = P(A_{i_1})P(A_{i_2}) \cdots P(A_{i_n})$$

を満たすときにいう.

例 3.2.3 52 枚のカードから 1 枚を無作為に抽出するとき, 抜き出したカードがエースである事象 A とスペードである事象 B は独立である.

例 3.2.4 6 つの数字 1, 2, 3, 4, 5, 6 を一列に並べる試行に対して, どの並べ方も等確率で現れる確率モデルを考える. このとき, 1 が 2 に先行する事象 A , 3 が 4 に先行する事象 B , 5 が 6 に先行する事象 C は独立である.

例 3.2.5 壺の中に 112, 121, 211, 222 という番号の着いた 4 個の球が入っている. この壺から 1 個の球を取り出して番号を読むとき, 100 位の数字が 1 である事象を A_1 , 10 位の数字が 1 である事象を A_2 , 1 位の数字が 1 である事象を A_3 とする. このとき, A_1, A_2, A_3 は互いに独立であるが独立でない.

注意 3.2.6 同一の事象 A を 2 つ並べてできる事象の列 A, A が独立であれば, 定義によって,

$$P(A \cap A) = P(A)P(A)$$

が成り立つ. したがって, $P(A) = 0$ または $P(A) = 1$ である. つまり, 自分自身と独立な事象は, 本質的に空事象 $\emptyset$ と全事象 Ω といえる. ($P(A) = 0$ だからといって $A = \emptyset$ とは限らないし, $P(A) = 1$ だからといって $A = \Omega$ であるとは限らない.)

補題 3.2.7 集合 A に対して, $A^\#$ で A または A^c を表わすものとする. 次の主張を証明せよ.

- (1) 2 つの事象 A と B が独立であれば, $A^\#$ と $B^\#$ も独立であることを示せ.
- (2) 事象列 $A_1, A_2, \dots$ が独立であれば, $A_1^\#, A_2^\#, \dots$ も独立であることを示せ.

証明 (1) は容易. (2) は帰納法で証明できる. ■

3.3 独立な確率変数

定義 3.3.1 確率変数の (有限または無限) 列 $X_1, X_2, \dots$ が独立であるとは, そこから選んだ任意有限個の確率変数 $X_{i_1}, \dots, X_{i_n}$ ($i_1 < i_2 < \dots < i_n$) と任意定数 $a_1, \dots, a_n$ に対して

$$P(X_{i_1} \leq a_1, X_{i_2} \leq a_2, \dots, X_{i_n} \leq a_n) = P(X_{i_1} \leq a_1)P(X_{i_2} \leq a_2) \cdots P(X_{i_n} \leq a_n) \quad (3.3)$$

が成り立つときにいう. 言い換えれば, 確率変数 $X_1, X_2, \dots$ が独立であるとは, 任意定数 $a_1, a_2, \dots$ を与えて得られる事象列

$$A_1 = \{X_1 \leq a_1\}, A_2 = \{X_2 \leq a_2\}, \dots, A_n = \{X_n \leq a_n\}, \dots$$

が独立になる (定義 3.2.2) ときにいう.

確率変数 $X_1, X_2, \dots$ が離散型であるときは, (3.3) に代えて,

$$P(X_{i_1} = a_1, X_{i_2} = a_2, \dots, X_{i_n} = a_n) = P(X_{i_1} = a_1)P(X_{i_2} = a_2) \cdots P(X_{i_n} = a_n)$$

としても同じことである.

定義 3.3.2 確率変数の (有限または無限) 列 $X_1, X_2, \dots$ が互いに独立であることも定義 3.3.1 に準じて定義する.

例 3.3.3 長方形 $\Omega = \{(x, y); a \leq x \leq b, c \leq y \leq d\}$ からランダムに 1 点を選び出す確率モデルを考える. 選ばれた点の x 座標を X , y 座標を Y とするとき, X, Y は独立な確率変数である.

3.4 多次元確率変数

n 個の確率変数 $X_1, X_2, \dots, X_n$ を組にして定義される

$$\mathbf{X} = (X_1, X_2, \dots, X_n)$$

を n 次元確率変数 (確率ベクトル) という. これに対して,

$$F_{\mathbf{X}}(x_1, x_2, \dots, x_n) = P(X_1 \leq x_1, X_2 \leq x_2, \dots, X_n \leq x_n)$$

で定義される n 変数関数 $F_{\mathbf{X}}$ を $\mathbf{X} = (X_1, \dots, X_n)$ の同時分布関数または結合分布関数という. この文脈で, 各 X_i の分布関数 $F_{X_i}(x) = P(X_i \leq x)$ を X_i の周辺分布関数という.

同時分布関数で 1 つの変数を除いて他の変数を ∞ に近づければ, 周辺分布関数が得られる. たとえば,

$$\lim_{\substack{x_2 \rightarrow +\infty \\ \vdots \\ x_n \rightarrow +\infty}} F_{\mathbf{X}}(x_1, x_2, \dots, x_n) = F_{X_1}(x_1)$$

が成り立つ. 周辺分布関数は個別に X_i の確率法則を記述するもので, それらをすべて知ったからといって, 確率ベクトル $\mathbf{X}$ の確率法則は定まらない.

n 次元確率変数 $\mathbf{X} = (X_1, X_2, \dots, X_n)$ において, $X_1, \dots, X_n$ がすべて離散型であれば, 同時分布関数 $F_{\mathbf{X}}$ に代えて,

$$P(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n)$$

を考えれば十分である. また, 連続型の場合で,

$$F_{\mathbf{X}}(x_1, \dots, x_n) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} f_{\mathbf{X}}(s_1, \dots, s_n) ds_1 \cdots ds_n$$

のような表示ができるとき, n 変数関数 $f_{\mathbf{X}}$ を $\mathbf{X} = (X_1, \dots, X_n)$ の同時密度関数という. 同時密度関数を, 1 つの変数を除いて他の変数について $\mathbb{R}$ 上で積分すれば, 周辺密度関数が得られる. たとえば,

$$\int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} f_{\mathbf{X}}(x_1, x_2, \dots, x_n) dx_2 \cdots dx_n = f_{X_1}(x_1)$$

が成り立つ.

定理 3.4.1 n 次元確率変数 $(X_1, \dots, X_n)$ に対して, その同時分布関数を $F(x_1, \dots, x_n)$, X_i の周辺分布関数を $F_i(x_i)$ とする. このとき, $X_1, \dots, X_n$ が独立であることと,

$$F(x_1, \dots, x_n) = \prod_{i=1}^n F_i(x_i)$$

が成り立つことは同値である.

定 理 3.4.2 n 次元確率変数 $(X_1, \dots, X_n)$ が同時密度関数 $f(x_1, \dots, x_n)$ をもつとし, X_i の周辺密度関数を $f_i(x_i)$ とする. このとき, $X_1, \dots, X_n$ が独立であることと,

$$f(x_1, \dots, x_n) = \prod_{i=1}^n f_i(x_i)$$

が成り立つことは同値である.

一般に, n 次元確率変数 $(X_1, \dots, X_n)$ の同時分布は周辺分布からだけでは決定されない. しかし, $X_1, \dots, X_n$ が独立であれば, $(X_1, \dots, X_n)$ の同時分布が周辺分布によって完全に決定され (定理 3.4.1), $X_1, \dots, X_n$ から得られる統計量は周辺分布だけを用いて計算できることになる.

3.5 平均値・分散・共分散・相関係数

確率変数 X の平均値は

$$m_X = \mathbf{E}(X) = \int_{-\infty}^{+\infty} x \mu_X(dx)$$

で定義される.

定 理 3.5.1 (平均値の線形性) 確率変数 X, Y と定数 a, b に対して,

$$\mathbf{E}(aX + bY) = a\mathbf{E}(X) + b\mathbf{E}(Y).$$

定 理 3.5.2 (平均値の乗法性) 独立な確率変数列 $X_1, X_2, \dots, X_n$ に対して,

$$\mathbf{E}[X_1 X_2 \cdots X_n] = \mathbf{E}[X_1] \cdots \mathbf{E}[X_n]. \quad (3.4)$$

証 明 まず, $X_k = 1_{A_k}$ の形 (指示確率変数) の場合について証明する. 定義から $X_1, \dots, X_n$ の独立性と $A_1, \dots, A_n$ の独立性は同値であることがわかる. よって,

$$\begin{aligned} \mathbf{E}[X_1 \cdots X_n] &= \mathbf{E}[1_{A_1 \cap \cdots \cap A_n}] = P(A_1 \cap \cdots \cap A_n) \\ &= P(A_1) \cdots P(A_n) = \mathbf{E}[X_1] \cdots \mathbf{E}[X_n]. \end{aligned}$$

これは, (3.4) の成立を示している. 次いで, 平均値の線形性によって, 各 X_k が有限個の値を取る場合 (指示確率変数の線形結合) も成立する. 一般の X_k については, ルベーク積分の定義に戻って, 指示確率変数の線形結合による近似の議論によって証明される. ■

確率変数 X の分散を

$$\sigma_X^2 = \mathbf{V}(X) = \mathbf{E}[(X - m_X)^2] = \mathbf{E}[X^2] - \mathbf{E}[X]^2$$

で定義する. X の分布を $\mu(dx)$ とすれば,

$$\mathbf{V}(X) = \int_{-\infty}^{+\infty} (x - m_X)^2 \mu(dx) = \int_{-\infty}^{+\infty} x^2 \mu(dx) - \left(\int_{-\infty}^{+\infty} x \mu(dx) \right)^2$$

が成り立つ.

定 義 3.5.3 X, Y を確率変数とする.

$$\sigma_{XY} = \mathbf{E}[(X - \mathbf{E}(X))(Y - \mathbf{E}(Y))] = \mathbf{E}[XY] - \mathbf{E}[X]\mathbf{E}[Y]$$

を X, Y の共分散という. (なお, $\sigma_{XX} = \sigma_X^2$ は X の分散である.) $\sigma_{XY} = 0$ のとき X と Y は無相関であるという. また, $\sigma_{XY} > 0, \sigma_{XY} < 0$ にしたがって, X と Y は正の相関または負の相関をもつという.

定 理 3.5.4 確率変数 X, Y が互いに独立ならば、それらは無相関である。

証 明 X, Y が互いに独立であれば、 $\mathbf{E}[XY] = \mathbf{E}[X]\mathbf{E}[Y]$ が成り立つので、

$$\sigma_{XY} = \mathbf{E}[XY] - \mathbf{E}[X]\mathbf{E}[Y] = 0$$

である。 ■

注意 3.5.5 確率変数 X を

$$P(X = -1) = P(X = 1) = \frac{1}{4}, \quad P(X = 0) = \frac{1}{2}$$

で定め、 $Y = X^2$ とおく。このとき、 X, Y は独立ではないが、 $\sigma_{XY} = 0$ であることを示せ。つまり、定理 3.5.4 の逆は一般には成り立たない。

定 理 3.5.6 (分散の加法性) 確率変数列 $X_1, X_2, \dots, X_n$ において、どの 2 つも無相関であれば、

$$\mathbf{V}\left[\sum_{k=1}^n X_k\right] = \sum_{k=1}^n \mathbf{V}[X_k]$$

が成り立つ。

定 義 3.5.7 確率変数 X, Y が $\sigma_X > 0, \sigma_Y > 0$ を満たすとき、

$$\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}$$

を X, Y の相関係数という。

定 理 3.5.8 確率変数 X, Y に対して、

$$-1 \leq \rho_{XY} \leq 1$$

が成り立つ。ただし、 $\sigma_X > 0, \sigma_Y > 0$ とする。

証 明 確率変数 X, Y から平均値を引いて、

$$\bar{X} = X - \mathbf{E}[X], \quad \bar{Y} = Y - \mathbf{E}[Y],$$

とおく。

$$\mathbf{E}[\bar{X}^2] = \mathbf{V}[X] = \sigma_X^2, \quad \mathbf{E}[\bar{Y}^2] = \mathbf{V}[Y] = \sigma_Y^2, \quad (3.5)$$

に注意しておく。さて、任意の実数 $t \in \mathbf{R}$ に対して $(t\bar{X} + \bar{Y})^2 \geq 0$ であるから、

$$0 \leq \mathbf{E}[(t\bar{X} + \bar{Y})^2] = t^2 \mathbf{E}[\bar{X}^2] + 2t \mathbf{E}[\bar{X}\bar{Y}] + \mathbf{E}[\bar{Y}^2]$$

が成り立つ。したがって、2 次式の判別式から

$$\mathbf{E}[\bar{X}\bar{Y}]^2 - \mathbf{E}[\bar{X}^2]\mathbf{E}[\bar{Y}^2] \leq 0.$$

(3.5) に注意して、

$$\mathbf{E}[\bar{X}\bar{Y}]^2 \leq \mathbf{E}[\bar{X}^2]\mathbf{E}[\bar{Y}^2] = \sigma_X^2 \sigma_Y^2$$

が得られる。これを用いて、

$$\rho_{XY}^2 = \frac{\sigma_{XY}^2}{\sigma_X^2 \sigma_Y^2} = \frac{\mathbf{E}[(X - \mathbf{E}(X))(Y - \mathbf{E}(Y))]^2}{\sigma_X^2 \sigma_Y^2} = \frac{\mathbf{E}[\bar{X}\bar{Y}]^2}{\sigma_X^2 \sigma_Y^2} \leq 1.$$

したがって, $-1 \leq \rho_{XY} \leq 1$. ■

相関係数は2つの確率変数 X, Y の間の線形的な関係性の強さを -1 から 1 の実数を指標として表したものといえる. たとえば, $Y = aX + b$ という線形関係があれば, $a > 0$, $a < 0$ にしたがって, $\rho_{XY} = 1$ または $\rho_{XY} = -1$ となる.

注意 3.5.9 確率変数 X のもっとも基本的な統計量は平均値 $E[X]$ と分散 $V[X]$ であった. n 次元確率変数 $\mathbf{X} = (X_1, X_2, \dots, X_n)$ の基本的な統計量で, これらに対応するものは, 平均ベクトル

$$m_{\mathbf{X}} = (E[X_1], E[X_2], \dots, E[X_n])$$

と共分散行列

$$\Sigma_{\mathbf{X}} = \begin{bmatrix} \sigma_{X_1 X_1} & \sigma_{X_1 X_2} & \cdots & \sigma_{X_1 X_n} \\ \sigma_{X_2 X_1} & \sigma_{X_2 X_2} & \cdots & \sigma_{X_2 X_n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{X_n X_1} & \sigma_{X_n X_2} & \cdots & \sigma_{X_n X_n} \end{bmatrix}$$

である. 共分散行列の対角成分 $\sigma_{X_i X_i} = \sigma_{X_i}^2 = V[X_i]$ は X_i の分散である. また, $X_1, X_2, \dots, X_n$ が互いに無相関であることと共分散行列が対角行列になることは同値である.

レポート問題 6 半径 ρ の円板からランダムに1点を選び出す確率モデルを考える. この円板内の点は極座標 (r, θ) で表示することとする ($0 \leq r \leq \rho$, $0 \leq \theta < 2\pi$). さて, 選ばれた点の中心からの距離を R , 回転角成分を Θ とするとき, R, Θ は独立な確率変数であることを示せ.

レポート問題 7 サイコロを2個投げたときに出る目のうち大きい方を L , 小さい方を S とする. (同じ目の場合は, $L = S$ である.) このとき,

- (1) 共分散 σ_{LS} と相関係数 r_{LS} を求めよ.
- (2) L, S は独立かどうかを判定せよ.

レポート問題 8 $a > 0, b > 0$ を定数とする. 三角形 $\Omega = \{(x, y); 0 \leq bx + ay \leq ab, x \geq 0, y \geq 0\}$ から, どの点も同等に選ばれるような確率モデルを考える. 選ばれた点の x 座標を X , y 座標を Y とするとき,

- (1) (X, Y) の同時分布関数と X, Y の周辺分布関数を求め, X, Y は独立ではないことを示せ.
- (2) X, Y の相関係数を求めよ.

第4章 大数の法則と中心極限定理

4.1 大数の法則を観察する

コインを投げ続ける試行を考える. n 回目のコイン投げの結果を X_n と表わすこととして,

$$X_n = \begin{cases} 1, & \text{表が出る} \\ 0, & \text{裏が出る} \end{cases} \quad (4.1)$$

とおく. このとき,

$$S_n = \sum_{k=1}^n X_k$$

は, 初めの n 回のコイン投げのうちで表が出た回数になる. したがって,

$$\frac{S_n}{n} = \frac{1}{n} \sum_{k=1}^n X_k$$

は, 初めの n 回のコイン投げのうちで表が出た相対頻度になる.

計算機を用いたシミュレーションをしてみよう. 下にあげたものはほんの一例であり, 相対頻度の挙動は様々である. しかし, 共通の性質として, 試行回数 n を増やすほどに相対頻度 S_n/n は $1/2$ に近づくように見える. このことを数学的にきちんと述べて証明したい.

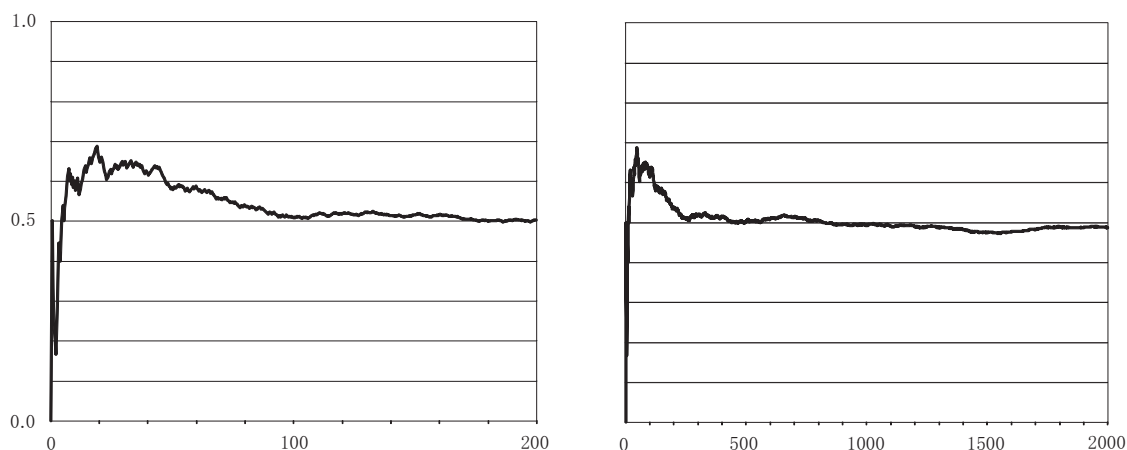


図 4.1: コイン投げで表の出る回数の相対頻度

まず, 定式化であるが, 素朴に,

$$\lim_{n \rightarrow \infty} \frac{S_n}{n} = \frac{1}{2} \quad (4.2)$$

としたのでは不十分である. そもそも, S_n/n も確率変数であり, 表が連続して出れば 1 を, 裏が連続して出れば 0 を, 一般には, $\{0, 1/n, 2/n, \dots, 1\}$ の値をランダムにとる. この状況は n を大きくしても同じで

あるから、とても定数 $1/2$ に収束するとは思えない。実際、コイン投げの結果 $\omega = (\omega_1, \omega_2, \dots)$ から得られる相対頻度の列

$$S_1(\omega), \frac{S_2(\omega)}{2}, \frac{S_3(\omega)}{3}, \dots, \frac{S_n(\omega)}{n}, \dots$$

には様々な挙動が可能である。たとえば、 $\omega = (1, 1, 1, \dots)$ に対しては、 S_n/n は 1 に収束し、 $\omega = (0, 0, 0, \dots)$ に対しては、 S_n/n は 0 に収束する。さらに、任意の $0 \leq t \leq 1$ に対して、 S_n/n が t に収束するような ω が存在し、また、 S_n/n が収束しないような ω も存在する。つまり、すべての標本 ω に対して (4.2) を期待することはできない。したがって、経験的事実の説明のためには、(4.2) になんらかの確率的な定式化が必要なのである。

4.2 大数の弱法則

定理 4.2.1 (大数の弱法則) $X_1, X_2, \dots$ を同分布をもつ確率変数列とし、その平均を m 、分散を σ^2 とする (したがって、有限な分散をもつことが仮定されている)。もし、 $X_1, X_2, \dots$ が互いに無相関であれば、任意の $\epsilon > 0$ に対して、

$$\lim_{n \rightarrow \infty} P\left(\left|\frac{1}{n} \sum_{k=1}^n X_k - m\right| \geq \epsilon\right) = 0$$

が成り立つ。このことを、 $\frac{1}{n} \sum_{k=1}^n X_k$ は m に確率収束するという。

注意 4.2.2 多くの本では、大数の弱法則を確率変数列 $X_1, X_2, \dots$ に「独立」性を仮定して述べているが、より弱い「無相関」の仮定の下で成立する。

定理 4.2.3 (チェビシェフの不等式) 確率変数 X の平均を m 、分散を σ^2 とすれば、任意の $\epsilon > 0$ に対して、

$$P(|X - m| \geq \epsilon) \leq \frac{\sigma^2}{\epsilon^2}$$

が成り立つ。

証明 定義によって、

$$\begin{aligned} m &= \mathbf{E}[X] = \int_{\Omega} X(\omega) P(d\omega), \\ \sigma^2 &= \mathbf{E}[(X - m)^2] = \int_{\Omega} (X(\omega) - m)^2 P(d\omega). \end{aligned}$$

さて、分散を表す積分を 2 つに分ければよい。

$$\begin{aligned} \sigma^2 &= \int_{\Omega} (X(\omega) - m)^2 P(d\omega) \\ &= \int_{|X-m| \geq \epsilon} (X(\omega) - m)^2 P(d\omega) + \int_{|X-m| < \epsilon} (X(\omega) - m)^2 P(d\omega) \\ &\geq \int_{|X-m| \geq \epsilon} (X(\omega) - m)^2 P(d\omega) \\ &\geq \int_{|X-m| \geq \epsilon} \epsilon^2 P(d\omega) \\ &= \epsilon^2 P(|X - m| \geq \epsilon). \end{aligned}$$

よって、示された。 ■

証明 (大数の弱収束) 簡単のため,

$$Y = Y_n = \frac{1}{n} \sum_{k=1}^n X_k$$

とおく. 平均値は,

$$\mathbf{E}[Y] = \frac{1}{n} \sum_{k=1}^n \mathbf{E}[X_k] = m.$$

次に分散を計算する. 無相関の仮定から $\mathbf{E}[X_k X_l] = \mathbf{E}[X_k] \mathbf{E}[X_l]$ ($k \neq l$) に注意すれば,

$$\begin{aligned} \mathbf{E}[Y^2] &= \frac{1}{n^2} \sum_{k,l=1}^n \mathbf{E}[X_k X_l] \\ &= \frac{1}{n^2} \left\{ \sum_{k=1}^n \mathbf{E}[X_k^2] + \sum_{k \neq l} \mathbf{E}[X_k X_l] \right\} \\ &= \frac{1}{n^2} \left\{ \sum_{k=1}^n (\mathbf{V}[X_k] + \mathbf{E}[X_k]^2) + \sum_{k \neq l} \mathbf{E}[X_k] \mathbf{E}[X_l] \right\} \\ &= \frac{1}{n^2} \{ n\sigma^2 + nm^2 + (n^2 - n)m^2 \} \\ &= \frac{\sigma^2}{n} + m^2. \end{aligned}$$

したがって,

$$\mathbf{V}[Y] = \mathbf{E}[Y^2] - \mathbf{E}[Y]^2 = \frac{\sigma^2}{n}.$$

チェビシェフの不等式によって,

$$P(|Y - m| \geq \epsilon) \leq \frac{\mathbf{V}[Y]}{\epsilon^2} = \frac{\sigma^2}{n\epsilon^2}.$$

したがって,

$$\lim_{n \rightarrow \infty} P(|Y_n - m| \geq \epsilon) = 0.$$

これが示したかったことである. ■

例 4.2.4 コインを投げ続ける試行を考えよう. 表の相対頻度はコイン投げの回数が増えるに従って次第に $1/2$ に近づくことを大数の弱法則で説明しよう. まず,

$$X_n = \begin{cases} 1, & n \text{ 回目が表,} \\ 0, & n \text{ 回目が裏,} \end{cases}$$

とおく. $X_1, X_2, \dots$ は独立同分布の確率変数列であり, 平均値と分散は,

$$m = \frac{1}{2}, \quad \sigma^2 = \frac{1}{4}$$

である. 大数の弱法則によれば, 任意の $\epsilon > 0$ に対して,

$$\lim_{n \rightarrow \infty} P\left(\left|\frac{1}{n} \sum_{k=1}^n X_k - \frac{1}{2}\right| \geq \epsilon\right) = 0$$

が成り立つ. ここで, $\frac{1}{n} \sum_{k=1}^n X_k$ は初めの n 回のコイン投げの結果, 表の出る相対頻度である. その相対頻度が $1/2$ から ϵ 以上ずれる確率は $n \rightarrow \infty$ で 0 に近づく. つまり, コイン投げにおける経験的事実「表の出る相対頻度はコイン投げの回数が増えるに従って次第に $1/2$ に近づく」を確率モデルで証明したことになる.

4.3 大数の強法則

大数の弱法則において、 $\epsilon > 0$ は任意であるので、いつそのこと $\epsilon \rightarrow 0$ とできないかという疑問がわくだろう。これに答えるのが次の主張である。

定理 4.3.1 (大数の強法則) $X_1, X_2, \dots$ を同分布の確率変数列として、その平均値を m とする (平均値の存在は仮定するが、分散の存在は仮定しない)。もし、 $X_1, X_2, \dots$ が互いに独立であれば、

$$P\left(\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n X_k = m\right) = 1$$

が成り立つ。言い換えれば、

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n X_k = m \quad \text{a.s.}$$

が成り立つ。

注意 4.3.2 コルモゴロフは大数の強法則を「独立」な同分布確率変数列について証明し、多くの本ではそれが紹介されている。その証明では、いわゆる「コルモゴロフの概収束定理」と呼ばれる一般的な結果を適用するため、「独立」の仮定を弱めることはできない。上で述べた大数の強法則は、N. Etemadi (1981) による。仮定を「互いに独立」に弱めたばかりではなく、その証明も著しく初等化された。たとえば、佐藤坦や Durrett の教科書ではそれが紹介されている。

4.4 ド・モアブル—ラプラスの定理

数値計算によって、二項分布のヒストグラムは同じ平均と分散をもつ正規分布の曲線に近いことが見て取れる。二項分布 $B(n, p)$ の平均は $m = np$ 、分散は $\sigma^2 = np(1-p)$ であるから、このことを

$$B(n, p) \approx N(np, np(1-p)) \quad (4.3)$$

と表そう。さらに、 n が大きくなればなるほど近似が良くなることもわかる。

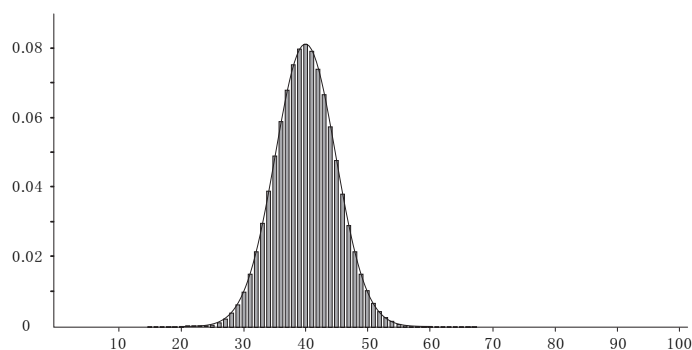


図 4.2: $B(100, 0.4)$ と同じ平均・分散をもつ正規分布

近似 (4.3) の意味を考えてみよう。ヒストグラムと密度関数のグラフが近いとすれば、それらの分布関数が近いということである。つまり、二項分布 $B(n, p)$ に従う確率変数 S に対して、

$$P(S \leq x) \approx \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^x e^{-(t-m)^2/2\sigma^2} dt, \quad m = np, \quad \sigma^2 = np(1-p)$$

が成り立つ。右辺の積分で変数変換を行えば,

$$P(S \leq x) \approx \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{(x-m)/\sigma} e^{-t^2/2} dt$$

となる。明らかな等式

$$P(S \leq x) = P\left(\frac{S-m}{\sigma} \leq \frac{x-m}{\sigma}\right)$$

を用いて, $(x-m)/\sigma$ を x と書き換えれば,

$$P\left(\frac{S-np}{\sqrt{np(1-p)}} \leq x\right) \approx \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-t^2/2} dt \quad (4.4)$$

が得られる。右辺は標準正規分布 $N(0, 1)$ に関する積分であり, n に依存しない。(4.3)において n が大きくなればなるほど近似が良くなるということは, $n \rightarrow \infty$ において (4.4) が等式になることで説明できる。実際, このことを証明することができる (証明は参考書を参照)。

定理 4.4.1 (ド・モアブル–ラプラスの定理) $0 < p < 1$ を定数とする。二項分布 $B(n, p)$ に従う確率変数 S_n に対して,

$$\lim_{n \rightarrow \infty} P\left(\frac{S_n - np}{\sqrt{np(1-p)}} \leq x\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-t^2/2} dt \quad (4.5)$$

が成り立つ。簡単に言えば, 二項分布 $B(n, p)$ は n が大きいとき, 同じ平均と分散をもつ正規分布 $N(np, np(1-p))$ に漸近する。

4.5 中心極限定理

さて, 定理 4.4.1 を少し書き換えてみよう。成功確率 p のベルヌイ試行列を $Z_1, Z_2, \dots$ とすれば,

$$S_n = \sum_{k=1}^n Z_k$$

が二項分布 $B(n, p)$ に従う確率変数となる。ここで, Z_k の正規化を

$$\bar{Z}_k = \frac{Z_k - p}{\sqrt{p(1-p)}}$$

で定義すると, $\bar{Z}_1, \bar{Z}_2, \dots$ は平均 0, 分散 1 に正規化された独立同分布の確率変数列となる。さらに,

$$\frac{S_n - np}{\sqrt{np(1-p)}} = \frac{1}{\sqrt{n}} \sum_{k=1}^n \frac{Z_k - p}{\sqrt{p(1-p)}} = \frac{1}{\sqrt{n}} \sum_{k=1}^n \bar{Z}_k$$

に注意しよう。そうすると, (4.5) は,

$$\lim_{n \rightarrow \infty} P\left(\frac{1}{\sqrt{n}} \sum_{k=1}^n \bar{Z}_k \leq x\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-t^2/2} dt$$

となる。実は, この極限公式はベルヌイ試行列でなくても成り立つのである。

定理 4.5.1 (中心極限定理) $X_1, X_2, \dots$ を独立同分布の確率変数列で, 平均 0, 分散 1 に正規化されているものとする。このとき, すべての $x \in \mathbf{R}$ に対して,

$$\lim_{n \rightarrow \infty} P\left(\frac{1}{\sqrt{n}} \sum_{k=1}^n X_k \leq x\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-t^2/2} dt$$

が成り立つ。言い換えると, $\frac{1}{\sqrt{n}} \sum_{k=1}^n X_k$ の分布は $n \rightarrow \infty$ で標準正規分布 $N(0, 1)$ に弱収束する。

証明のために、分布の特性関数が必要である。

定 義 4.5.2 確率変数 X の特性関数とは、

$$\varphi(z) = \mathbf{E}[e^{izX}], \quad z \in \mathbf{R}$$

によって定義される実数変数で複素数値をとる関数 $\varphi(z)$ のことである。確率変数 X の分布を $\mu(dx)$ とすれば、

$$\varphi(z) = \int_{-\infty}^{+\infty} e^{izx} \mu(dx), \quad z \in \mathbf{R}$$

となる。これを分布 $\mu(dx)$ の特性関数という。文脈によってはフーリエ変換ともいう。

分布の特性関数に関する理論は、中心極限定理の研究とともに深化してきたといえる。ここで必要なことは、次の基本的な結果である。

定 理 4.5.3 (グリヴェンコの定理) 1次元分布 $\mu_1, \mu_2, \dots, \mu$ の特性関数を $\varphi_1, \varphi_2, \dots, \varphi$ とする。すべての $z \in \mathbf{C}$ に対して $\lim_{n \rightarrow \infty} \varphi_n(z) = \varphi(z)$ ならば、 μ_n は μ に弱収束する。つまり、 $F_1, F_2, \dots, F$ を $\mu_1, \mu_2, \dots, \mu$ の分布関数とすると、 F のすべての連続点 x において、

$$\lim_{n \rightarrow \infty} F_n(x) = F(x)$$

が成り立つ。

また、指数関数のテーラー展開に関する簡単な評価に注意しておく。

補 題 4.5.4 すべての $x \in \mathbf{R}$ に対して、

$$\left| e^{ix} - \left(1 + ix + \frac{(ix)^2}{2!} \right) \right| \leq \min \left\{ \frac{|x|^3}{6}, |x|^2 \right\}$$

が成り立つ。

証 明 (中心極限定理) $\frac{1}{\sqrt{n}} \sum_{k=0}^n X_k$ の特性関数を

$$\varphi_n(z) = \mathbf{E} \left[\exp \left\{ \frac{iz}{\sqrt{n}} \sum_{k=0}^n X_k \right\} \right] \quad (4.6)$$

とおく。一方、標準正規分布 $N(0, 1)$ の特性関数は $e^{-z^2/2}$ である (レポート問題??)。したがって、グリヴェンコの定理によって、各 $z \in \mathbf{R}$ に対して、

$$\lim_{n \rightarrow \infty} \varphi_n(z) = e^{-z^2/2} \quad (4.7)$$

を示せば結論が従うことがわかる。

確率変数列 $X_1, X_2, \dots$ は同分布であるから同一の特性関数

$$\varphi(z) = \mathbf{E}[e^{izX_1}]$$

をもつ。また、それらは独立であるから平均値の乗法性によって、(4.6) は

$$\varphi_n(z) = \prod_{k=1}^n \mathbf{E} \left[\exp \left\{ \frac{iz}{\sqrt{n}} X_k \right\} \right] = \varphi \left(\frac{z}{\sqrt{n}} \right)^n \quad (4.8)$$

となる。さて、

$$e^{i \frac{z}{\sqrt{n}} X_1} = 1 + i \frac{z}{\sqrt{n}} X_1 - \frac{z^2}{2n} X_1^2 + R_n(z)$$

とにおいて、 $\mathbf{E}[X_1] = 0$, $\mathbf{V}[X_1] = 1$ に注意して平均値をとれば、

$$\varphi\left(\frac{z}{\sqrt{n}}\right) = \mathbf{E}\left[e^{i \frac{z}{\sqrt{n}} X_1}\right] = 1 - \frac{z^2}{2n} + \mathbf{E}[R_n(z)]$$

を得る。したがって、(4.8) は、

$$\varphi_n(z) = \left(1 - \frac{z^2}{2n} + \mathbf{E}[R_n(z)]\right)^n \quad (4.9)$$

となる。ここで、

$$\lim_{n \rightarrow \infty} n \mathbf{E}[R_n(z)] = 0 \quad (4.10)$$

を示せば、よく知られた指数関数の公式 (レポート問題 34) によって、

$$\lim_{n \rightarrow \infty} \varphi_n(z) = \lim_{n \rightarrow \infty} \left(1 - \frac{z^2}{2n} + \mathbf{E}[R_n(z)]\right)^n = e^{-z^2/2}$$

が得られ、(4.7) が導かれる。

あとは、(4.10) を示すだけである。補題 4.5.4 を用いれば、

$$|R_n(z)| \leq \min \left\{ \frac{1}{6} \left| \frac{z}{\sqrt{n}} X_1 \right|^3, \left| \frac{z}{\sqrt{n}} X_1 \right|^2 \right\}.$$

両辺に n をかけて平均値をとれば、

$$|n \mathbf{E}[R_n(z)]| \leq \mathbf{E}[n |R_n(z)|] \leq |z|^2 \mathbf{E} \left[\min \left\{ \frac{|z|}{6\sqrt{n}} |X_1|^3, |X_1|^2 \right\} \right]. \quad (4.11)$$

ここで、

$$\min \left\{ \frac{|z|}{6\sqrt{n}} |X_1|^3, |X_1|^2 \right\} \leq |X_1|^2$$

は明らかで、仮定より $\mathbf{E}[|X_1|^2] < \infty$ であるから、ルベークの収束定理によって、

$$\lim_{n \rightarrow \infty} \mathbf{E} \left[\min \left\{ \frac{|z|}{6\sqrt{n}} |X_1|^3, |X_1|^2 \right\} \right] = \mathbf{E} \left[\lim_{n \rightarrow \infty} \min \left\{ \frac{|z|}{6\sqrt{n}} |X_1|^3, |X_1|^2 \right\} \right] = 0.$$

したがって、(4.11) より、

$$\lim_{n \rightarrow \infty} |n \mathbf{E}[R_n(z)]| = 0.$$

こうして、(4.10) が示された。この証明では、 $\mathbf{E}[|X_1|^3] < \infty$ を仮定していないところに注意せよ。もし $\mathbf{E}[|X_1|^3] < \infty$ が成り立てば、ルベークの収束定理によらず、簡単な評価で (4.10) がわかる。 ■

レポート問題 9 成功確率 p のベルヌイ試行列を $\{Z_k\}$ 、その正規化を $\{\bar{Z}_k\}$ とする。このとき、

$$\frac{S_n - np}{\sqrt{np(1-p)}} = \frac{1}{\sqrt{n}} \sum_{k=1}^n \bar{Z}_k$$

の特性関数を計算せよ。次に、 $n \rightarrow \infty$ の極限を計算して、ドモアブル-ラプラスの定理を (グリヴェンコの定理を用いて) 証明せよ。

レポート問題 10 (1) すべての $x \in \mathbf{R}$ に対して,

$$e^{ix} - \sum_{k=0}^n \frac{(ix)^k}{k!} = \frac{i^{n+1}}{n!} \int_0^x (x-t)^n e^{it} dt$$

が成り立つことを示せ.

(2) すべての $x \in \mathbf{R}$ に対して,

$$\left| e^{ix} - \sum_{k=0}^n \frac{(ix)^k}{k!} \right| \leq \frac{|x|^{n+1}}{(n+1)!}$$

が成り立つことを示せ.

(3) (1) の右辺に部分積分を施して, すべての $x \in \mathbf{R}$ に対して,

$$\left| e^{ix} - \sum_{k=0}^n \frac{(ix)^k}{k!} \right| \leq \frac{2|x|^n}{n!}$$

が成り立つことを示せ.

(補題 4.5.4 の証明が (もっと一般的な状況で) できた.)

レポート問題 11 $f(x)$ を区間 $[0, 1]$ 上の連続関数とする. $x_1, x_2, \dots$ を $[0, 1]$ から選ばれた乱数列とするとき, 平均値

$$\frac{1}{n} \sum_{k=1}^n f(x_k)$$

は積分

$$\int_0^1 f(x) dx$$

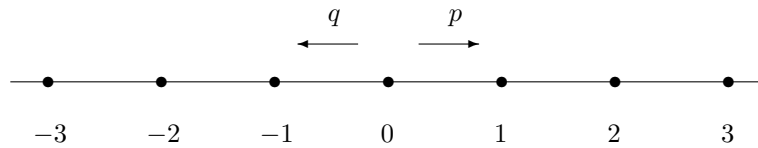
の近似値として用いられる (モンテカルロ法の原理). このことを大数の法則と中心極限定理を用いて説明し, 計算機による実例を示せ.

第5章 ランダム・ウォーク

ランダム・ウォークに「酔歩」とか「乱歩」といった訳語があてられることがある。それは、酔っ払いの動きをモデル化したものである。酔っ払いは、なぜ今この地点にいるのかを覚えておらず、次の一步をいくつかある可能な方向から1つランダムに選んでいるように見える。過去の経緯を覚えておらず、今の状態をもとにして次の状態への遷移が確率的に起こるといった動きをモデル化したものがマルコフ連鎖である。ランダム・ウォークはマルコフ連鎖の中で最も基本的なものである。

5.1 1次元ランダム・ウォーク

一本道を行き来する酔っ払い(ランダム・ウォーカーと呼ぶ)の動きをモデル化しよう。ランダム・ウォーカーの動きの規則は、各時点で右か左に1だけ移動するものとし、右に移動する確率を p 、左に移動する確率を q とおく ($p > 0, q > 0, p + q = 1$)。ただし、 p, q は場所や時間によらず常に一定であるとする。



ランダム・ウォーカーの時刻 n における位置を X_n と表わそう。時刻 0 ではランダム・ウォーカーは原点にいるものとする。つまり、 $X_0 = 0$ である。時刻 n における位置 X_n は $\{-n, -n+2, \dots, n-2, n\}$ に値をとる確率変数である。その確率分布を調べるのに、ベルヌイ試行列の和を用いるのが重要である。

$\{Z_n\}$ を独立同分布 (iid) の確率変数列で

$$P(Z_n = 1) = p, \quad P(Z_n = -1) = q$$

を満たすものとする。これもベルヌイ試行列と呼ぶ(通常のベルヌイ試行列は $\{0, 1\}$ に値をとる)。そうすれば、ランダム・ウォーカーの時刻 n における位置は、

$$X_n = Z_1 + Z_2 + \dots + Z_n, \quad X_0 = 0,$$

のようにベルヌイ試行列の和として表わされる。こうして得られる離散時間確率過程 $\{X_n\}$ を1次元ランダム・ウォークという。

定 理 5.1.1 X_n は $\{-n, -n+2, \dots, n-2, n\}$ に値をとる確率変数であり、その分布は、

$$P(X_n = n - 2k) = \binom{n}{k} p^{n-k} q^k, \quad k = 0, 1, 2, \dots, n.$$

証 明 $k = 0, 1, 2, \dots, n$ とする。

$$X_n = Z_1 + Z_2 + \dots + Z_n = n - 2k = (n - k) - k$$

となるのは, $Z_i = -1$ となる i が k 個, $Z_i = 1$ となる i が $n - k$ 個あるときに限る. そのような $Z_1, Z_2, \dots, Z_n$ の特定の配列が起こる確率は $p^{n-k}q^k$ である. よって,

$$P(X_n = n - 2k) = \binom{n}{k} p^{n-k} q^k$$

が得られる. ■

定 理 5.1.2 ランダム・ウォークの平均と分散は

$$\mathbf{E}[X_n] = (p - q)n, \quad \mathbf{V}[X_n] = 4pqn.$$

証 明 まず,

$$\mathbf{E}[Z_k] = p - q, \quad \mathbf{V}[Z_k] = 4pq$$

は容易である (確認せよ). そうすれば, 平均値の線形性によって,

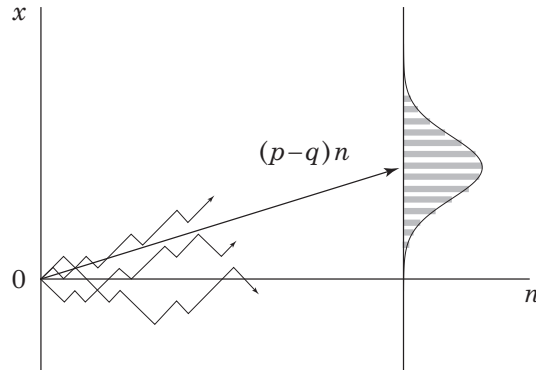
$$\mathbf{E}[X_n] = \sum_{k=1}^n \mathbf{E}[Z_k] = (p - q)n.$$

また, $\{Z_n\}$ は独立であるから, 分散の加法性によって,

$$\mathbf{V}[X_n] = \sum_{k=1}^n \mathbf{V}[Z_k] = 4pqn$$

が得られる. ■

時刻 n でどのあたりにランダム・ウォーカーが見つかりやすいかを示すのが X_n の確率分布である. 平均の位置 $(p - q)n$ の周りに揺らぐことになるが, n が大きくなれば, X_n の到達する範囲も広くなり, 分散も大きくなってゆく. さらに, n が大きいときは, X_n の確率分布は正規分布で近似できる (ド・モアブル-ラプラスの定理).



5.2 再帰性

原点を出発したランダム・ウォーカーが再び原点に戻ってくるかを考察しよう. 特に, 再帰確率に興味がある. 再帰確率とは, いつとはわからないが有限時間内に原点に戻ってくる確率のことである.

前節にしたがって, 各時点ごとに右に移動する確率を p , 左に移動する確率を q とし, 時刻 n におけるランダム・ウォーカーの位置を $\{X_n\}$ とする. 原点に戻ってくるのは偶数ステップ後に限ることに注意しておく. まず, $2n$ ステップ後に原点に (戻って) いる確率 p_{2n} は,

$$p_{2n} = P(X_{2n} = 0) = \binom{2n}{n} p^n q^n = \frac{(2n)!}{n!n!} p^n q^n, \quad n = 1, 2, \dots, \quad (5.1)$$

となることは明らかであろう. 便宜上, $p_0 = 1$ とおく. 求めたい再帰確率 R は

$$R = P\left(\bigcup_{n=1}^{\infty} \{X_{2n} = 0\}\right)$$

である. ここで, 右辺の事象は互いに排反になっていないので, R を p_{2n} の和として求めるわけにはゆかない.

そこで, $2n$ ステップ後に初めて 0 に戻る確率

$$q_{2n} = P(X_2 \neq 0, X_4 \neq 0, \dots, X_{2n-2} \neq 0, X_{2n} = 0) \quad n = 1, 2, \dots,$$

を考えるのが定石である. p_{2n} との違いに注意されたい. あるいは, 原点に初めて戻ってくる時間を

$$T = \inf\{n \geq 1; X_n = 0\} \quad (5.2)$$

とおけば,

$$q_{2n} = P(T = 2n) \quad (5.3)$$

となる.

注意 5.2.1 (5.2) において $\inf\{n \geq 1; X_n = 0\}$ は, 大雑把には, $X_n = 0$ となる最小の n を表すと理解してよい. しかし, $\inf$ を $\min$ で置き換えることはできない. $\{n \geq 1; X_n = 0\}$ が空集合になる場合もあるからで, そのときは, 定義によって $\inf\{n \geq 1; X_n = 0\} = +\infty$ とする. したがって, T は当初の意味で (実数に値をとるとしたので) 確率変数とは呼べないが, $+\infty$ を値として認めれば確率変数と呼んで差し支えない.

そうすれば, 再帰確率は

$$R = \sum_{n=1}^{\infty} q_{2n} \quad (5.4)$$

で与えられる. これを計算するのが目標である. この節では, q_{2n} を組合せ数などを用いて直接計算せずに, 母関数を用いて R を計算する (次節で別の方法を紹介する).

計算の鍵は, $\{p_{2n}\}$ と $\{q_{2n}\}$ の関係式

$$p_{2n} = \sum_{k=1}^n q_{2k} p_{2n-2k}, \quad n = 1, 2, \dots, \quad (5.5)$$

である. これは, $2n$ ステップ後に原点にいるという事象を, 何ステップ目で初めて原点に戻ったかで分類して確率を計算したものである. 2 つの数列 $\{p_{2n}\}$ と $\{q_{2n}\}$ の母関数をそれぞれ

$$g(z) = \sum_{n=0}^{\infty} p_{2n} z^{2n}, \quad h(z) = \sum_{n=1}^{\infty} q_{2n} z^{2n} \quad (5.6)$$

で定義する. これらは $|z| \leq 1$ で収束する. (5.5) の両辺に z^{2n} をかけて, n について和をとると,

$$\begin{aligned} g(z) - 1 &= \sum_{n=1}^{\infty} \sum_{k=1}^n q_{2k} z^{2k} p_{2n-2k} z^{2n-2k} \\ &= \sum_{k=1}^{\infty} \sum_{n=0}^{\infty} q_{2k} z^{2k} p_{2n} z^{2n} \\ &= h(z)g(z). \end{aligned}$$

したがって,

$$h(z) = 1 - \frac{1}{g(z)}. \quad (5.7)$$

一方, 二項級数の公式によって $g(z)$ が具体的に計算できる (レポート問題 12). 実際,

$$g(z) = \sum_{n=0}^{\infty} p_{2n} z^{2n} = \sum_{n=0}^{\infty} \binom{2n}{n} p^n q^n z^{2n} = \frac{1}{\sqrt{1-4pqz^2}}$$

がわかるから, (5.7) は

$$h(z) = 1 - \frac{1}{g(z)} = 1 - \sqrt{1-4pqz^2}. \quad (5.8)$$

特に, $z \rightarrow 1-0$ として,

$$R = h(1) = \sum_{n=1}^{\infty} q_{2n} = 1 - \sqrt{1-4pq} = 1 - |p-q|$$

が得られる (レポート問題 13). 以上をまとめると,

定 理 5.2.2 原点から出発する 1 次元ランダムウォークの原点への再帰確率 R は,

$$R = 1 - |p-q|$$

で与えられる.

原点から原点への再帰確率を考えたが, 任意の点から出発しその点に戻る再帰確率も同じであることは明らかであろう. ランダム・ウォークは, 再帰確率 $R = 1$ のとき再帰的であるといい, $R < 1$ のとき, 非再帰的または過渡的という. 非再帰的というのは, 決して原点に戻れないということではなく, いつまでたっても原点に戻ってこない確率が正であることを意味する.

定 理 5.2.3 1 次元ランダム・ウォークについて, 再帰的であるための必要十分条件は $p = q = 1/2$ (等方的) であること. 非再帰的であるための必要十分条件は $p \neq q$ であること.

ランダム・ウォークが再帰的であれば, 再帰時間の平均値を考えることに意味がある.

定 理 5.2.4 等方的な 1 次元ランダム・ウォークの平均再帰時間は無限大である.

証 明 ランダムウォーカーが初めて原点に戻るのに要する時間を T とする. 平均再帰時間とは $E(T)$ のことであるから,

$$E(T) = \sum_{n=1}^{\infty} 2nq_{2n}. \quad (5.9)$$

一方, $p = q = 1/2$ の場合の母関数は, (5.6) と (5.8) から,

$$h(z) = \sum_{n=1}^{\infty} q_{2n} z^{2n} = 1 - \sqrt{1-z^2}.$$

微分して,

$$h'(z) = \sum_{n=1}^{\infty} 2nq_{2n} z^{2n-1} = \frac{z}{\sqrt{1-z^2}}.$$

ここで, $z \rightarrow 1-0$ とすれば,

$$E(T) = \sum_{n=1}^{\infty} 2nq_{2n} = \lim_{z \rightarrow 1-0} h'(z) = \lim_{z \rightarrow 1-0} \frac{z}{\sqrt{1-z^2}} = +\infty.$$

よって, 再帰時間の期待値は無限大. ■

注意 5.2.5 ランダム・ウォークの再帰性については、マルコフ連鎖の枠組みで再び論じる。

レポート問題 12 α を実数の定数とする。二項級数

$$(1+x)^\alpha = \sum_{n=0}^{\infty} \binom{\alpha}{n} x^n$$

を用いて、

$$\sum_{n=0}^{\infty} \binom{2n}{n} z^n = \frac{1}{\sqrt{1-4z}}, \quad |z| < \frac{1}{4},$$

を証明せよ。

レポート問題 13 $a_n \geq 0$ として、べき級数

$$f(x) = \sum_{n=0}^{\infty} a_n x^n$$

を考える。このべき級数の収束半径が ≥ 1 であれば、

$$\lim_{x \rightarrow 1-0} f(x) = \sum_{n=0}^{\infty} a_n$$

が、 $\infty = \infty$ の場合も含めて成り立つことを示せ。

5.3 カタラン数

カタラン数というのは組合せ数の一種であり、数え上げなどに関連する組合せ数学の分野では常識である。(カタランは19世紀のフランスの数学者である。) 二項係数と同じくらい需要があり、スタンレーという有名な数学者はカタラン数の現れる事例を何百と収集して公開している。ところが、組合せ論など一部の専門家を除いては、一般には(日本では?) あまり知られていないのが不思議である。ランダムウォークの関連でも、カタラン数があちらこちらに顔を出すのだが、それを意識して積極的に用いるという方針で書かれた本は見当たらない。

まず、味もそっけもない数式による定義を与えておこう。 $n \geq 1$ として、 $(\epsilon_1, \epsilon_2, \dots, \epsilon_n)$ を ± 1 からなる長さ n の数列とする。つまり、 $\{-1, 1\}^n$ の要素である。この数列が、

$$\begin{aligned} \epsilon_1 &\geq 0 \\ \epsilon_1 + \epsilon_2 &\geq 0 \\ &\dots \\ \epsilon_1 + \epsilon_2 + \dots + \epsilon_{n-1} &\geq 0 \\ \epsilon_1 + \epsilon_2 + \dots + \epsilon_{n-1} + \epsilon_n &= 0 \end{aligned}$$

をみたすとき、長さ n のカタラン道という。明らかに、長さが奇数のカタラン道は存在しない。

定義 5.3.1 長さ $2n$ のカタラン道の個数をカタラン数といい C_n で表す。便宜上、 $C_0 = 1$ とする。

初めの数項は、

$$C_0 = C_1 = 1, \quad C_2 = 2, \quad C_3 = 5, \quad C_4 = 14, \quad C_5 = 42, \quad C_6 = 132, \dots$$

カタラン数を図形的に表現して, カタラン数を表す簡単な表式を求めよう. $n \times n$ の正方格子に, 左隅の角を原点 $(0, 0)$ とする xy -座標系を考える. 一般に, ± 1 からなる数列 $(\epsilon_1, \epsilon_2, \dots, \epsilon_n)$ に対して,

$$\epsilon_k = +1 \leftrightarrow u_k = (1, 0) \quad \epsilon_k = -1 \leftrightarrow u_k = (0, 1)$$

を対応させ, 頂点

$$(0, 0), u_1, u_1 + u_2, \dots, u_1 + u_2 + \dots + u_{n-1}, u_1 + u_2 + \dots + u_{n-1} + u_n$$

を順に線分で結ぶ経路を考える. このとき, $\epsilon_1 + \epsilon_2 + \dots + \epsilon_{n-1} + \epsilon_n = 0$ であれば, 最後の頂点は

$$u_1 + u_2 + \dots + u_{n-1} + u_n = (n, n)$$

となり, 得られた経路は, $(0, 0)$ と (n, n) を結ぶ最短経路の 1 つになる.

補題 5.3.2 カタラン道と $(0, 0)$ と (n, n) を結ぶ最短経路で対角線 $y = x$ の上方を通らないものは 1 対 1 に対応する.

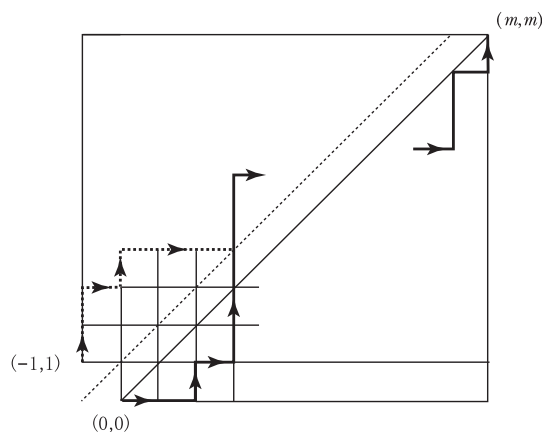
定理 5.3.3 (カタラン数)

$$C_n = \frac{(2n)!}{(n+1)!n!}, \quad n = 0, 1, 2, \dots,$$

証明 $n = 0$ のときは, 定義 $0! = 1$ から明らか. $n \geq 1$ のときは, 下図を参考にして,

$$C_n = \binom{2n}{n} - \binom{2n}{n+1} = \frac{(2n)!}{n!(n+1)!}$$

がわかる. ■



補題 5.3.4 カタラン数 C_n の母関数は,

$$f(z) = \sum_{n=0}^{\infty} C_n z^n = \frac{1 - \sqrt{1 - 4z}}{2z} \quad (5.10)$$

で与えられる.

証 明 レポート問題 14 を参照. ■

カタラン数に対して, 別のルールで経路を対応させよう. xy -座標系を考える. 頂点

$$(0, 0), (1, \epsilon_1), (2, \epsilon_1 + \epsilon_2), \dots, (n-1, \epsilon_1 + \epsilon_2 + \dots + \epsilon_{n-1}), (n, \epsilon_1 + \epsilon_2 + \dots + \epsilon_{n-1} + \epsilon_n)$$

を順に線分で結んでできる経路を考えると, ランダム・ウォークのサンプル・パスで時刻 0 で 0 を出発して時刻 $2n$ で 0 に戻り, 常に $y \geq 0$ を満たすものとカタラン道とが一対一対応する. したがって,

補 題 5.3.5 $n \geq 1$ とする. ランダム・ウォークのサンプル・パスで時刻 0 で 0 を出発して時刻 $2n$ で 0 に戻り, 常に $y \geq 0$ を満たすものの個数はカタラン数 C_n である.

定 理 5.3.6 右移動の確率が p , 左移動の確率が q であるような 1 次元ランダムウォーク $\{X_n\}$ に対して,

$$q_{2n} = P(X_2 \neq 0, X_4 \neq 0, \dots, X_{2n-2} \neq 0, X_{2n} = 0) = 2C_{n-1}(pq)^n, \quad n = 1, 2, \dots$$

証 明 証明は簡単であろう. ■

では, 定理 5.2.2 の別証明を与えよう.

定 理 5.3.7 原点から出発する 1 次元ランダムウォークの再帰確率 R は,

$$R = 1 - |p - q|$$

で与えられる.

証 明 定理 5.3.6 によって,

$$R = \sum_{n=1}^{\infty} q_{2n} = \sum_{n=1}^{\infty} 2C_{n-1}(pq)^n$$

であるが, これはカタラン数の母関数を用いれば,

$$R = 2pqf(pq) = 2pq \frac{1 - \sqrt{1 - 4pq}}{2pq} = 1 - \sqrt{1 - 4pq} = 1 - |p - q|$$

となって, 証明を終える. ■

レポート問題 14 カタラン数 C_n を次の手順で求めよ.

(1) $C_n = \sum_{k=1}^n C_{k-1}C_{n-k}$ が成り立つことを図形的に説明せよ.

(2) (1) を用いて, カタラン数の母関数 $f(z) = \sum_{n=0}^{\infty} C_n z^n$ が $f(z) - 1 = z\{f(z)\}^2$ を満たすことを示せ.

(3) $f(z)$ を求めよ.

(4) (3) をテーラー展開して C_n を求めよ.

レポート問題 15 縦の長さが m 横の長さが $m+n$ であるような格子において, $(0, 0)$ と (m, m) を結ぶ直線の上方を通らずに, $(0, 0)$ から $(m+n, m)$ に至る最短経路の個数は,

$$\frac{(m+n)!}{(m+1)!n!}$$

で与えられることを示せ.

レポート問題 16 $\{X_n\}$ を右移動の確率が p , 左移動の確率が $q = 1 - p$ であるようなランダムウォークで, $X_0 = 0$ とする. このとき, $n = 1, 2, \dots$ に対して,

$$\begin{aligned} P(X_1 \geq 0, X_2 \geq 0, \dots, X_{2n-1} \geq 0) \\ = P(X_1 \geq 0, X_2 \geq 0, \dots, X_{2n} \geq 0) = 1 - q \sum_{k=0}^{n-1} C_k (pq)^k \end{aligned}$$

が成り立つことを示せ. ただし, C_k はカタラン数である. 次に, これを用いて,

$$P(\text{すべての } n \geq 1 \text{ に対して } X_n \geq 0 \text{ となる}) = \begin{cases} 1 - \frac{q}{p}, & p > q, \\ 0, & p \leq q, \end{cases}$$

となることを示せ.

確率過程, 特にマルコフ連鎖とランダム・ウォークを学ぶために

1. 国沢清典: 確率論とその応用 (岩波全書), 1982.

すでに紹介したとおり.

2. W. Feller: An Introduction to Probability Theory and Its Applications, Vol. 1, Vol. 2, Wiley, 1957.

これもすでに紹介したとおり. ランダムウォークに関連する統計量の計算では, 事実上, カタラン数が現れているが, それをカタラン数と呼んではいけない. 邦訳もある.

W. フェラー (河田龍夫他訳): 確率論とその応用 (紀伊国屋). こちらは4分冊.

3. B. V. Gnedenko: The Theory of Probability and the Elements of Statistics, AMS Chelsea Publishing Co., 6th ed. 1989.

これもすでに紹介済み. 邦訳もある.

B. V. グネジェンコ (鳥居一雄訳): 確率論教程 I, II, 森北出版, 1971, 1972.

4. R. Durrett: Probability: Theory and Examples, Duxbury Press, 1996.

これもすでに紹介済み.

5. 志賀徳造: ルベーグ積分から確率論 (共立), 2000.

最後の章である種の平面グラフ上のランダム・ウォークを議論している.

6. R. B. シナジ (今野紀雄・林俊一訳): マルコフ連鎖から格子確率モデルへ, シュプリンガー東京, 2001.

マルコフ連鎖の基本的性質を速習するのに十分であろう.

7. Z. Brzeźniak and T. Zastawniak: Basic Stochastic Processes, Springer, 1999.

確率過程入門として易しく書かれている.

8. F. Spitzer: Principles of Random Walk, Springer, 2nd Ed., 1976.

ランダムウォークに関するプロ向きの本.

9. K. L. Chung: Markov Chains, Springer, 1960.

マルコフ連鎖に関するプロ向きの本.

5.4 逆正弦則

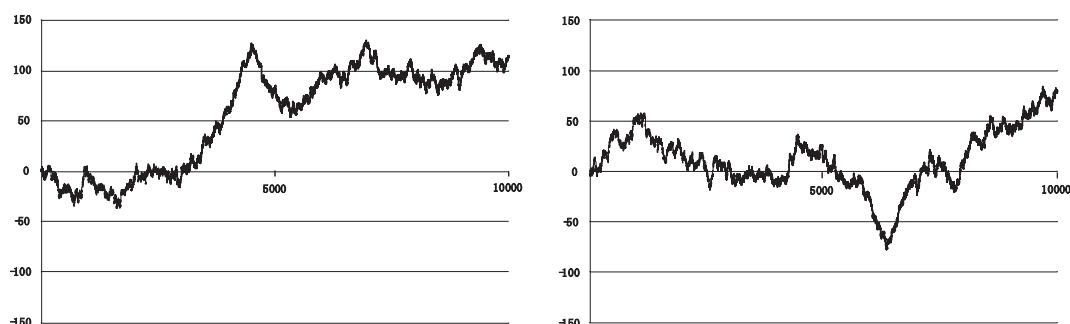
等方的なランダム・ウォーク $\{X_n\}$ を考えよう. つまり, $\{Z_n\}$ を

$$P(Z_n = +1) = P(Z_n = -1) = \frac{1}{2}$$

を満たすベルヌイ試行列として,

$$X_0 = 0, \quad X_n = \sum_{k=1}^n Z_k.$$

下は, シミュレーションによって, $X_0, X_1, X_2, \dots, X_{10000}$ の軌跡を示したものである. 様々なパターンが見られるが, ほんの 2 例だけを示している.



大数の法則によって, ± 1 はほぼ 5000 回ずつ出ているといえる. (X_{10000} の値から, ± 1 はほぼ 5000 ± 50 回である.) つまり, 折れ線のアップダウンはほぼ半々で起こっているのだが, 折れ線が存在範囲が横軸の上または下に偏っているとは思えないだろうか?

$X_0, X_1, \dots, X_{2n}$ を順次結んでできる折れ線を考えよう. $X_i \geq 0$ かつ $X_{i+1} \geq 0$ のとき, 時間区間 $[i, i+1]$ において正の部分にあるといい, $X_i \leq 0$ かつ $X_{i+1} \leq 0$ のとき, 時間区間 $[i, i+1]$ において負の部分にあるということにする. 時間区間 $[0, 2n]$ のうち, 正の部分にある時間区間の長さの和が $2k$ となる確率

$$W(2k|2n), \quad k = 0, 1, \dots, n,$$

を計算してみよう.

補題 5.4.1 $n = 1, 2, \dots$ に対して,

$$p_{2n} \equiv P(X_{2n} = 0) = \binom{2n}{n} \left(\frac{1}{2}\right)^{2n}.$$

証明 既知. ■

補題 5.4.2 $n = 1, 2, \dots$ に対して,

$$f_{2n} \equiv P(X_1 > 0, X_2 > 0, \dots, X_{2n-1} > 0, X_{2n} = 0) = C_{n-1} \left(\frac{1}{2}\right)^{2n} = \frac{1}{4n} p_{2n-2}.$$

証明 既知. ■

補題 5.4.3 $n = 1, 2, \dots$ に対して,

$$P(X_1 \neq 0, X_2 \neq 0, \dots, X_{2n-1} \neq 0, X_{2n} \neq 0) = p_{2n}.$$

証 明 余事象を考え, 等式

$$p_{2n-2} - p_{2n} = 2f_{2n} = \frac{1}{2n} p_{2n-2}$$

を用いる.

補 題 5.4.4 $n = 1, 2, \dots$ に対して,

$$P(X_1 \geq 0, X_2 \geq 0, \dots, X_{2n-1} \geq 0, X_{2n} \geq 0) = p_{2n}.$$

証 明 余事象を考え, 再び等式

$$p_{2n-2} - p_{2n} = 2f_{2n} = \frac{1}{2n} p_{2n-2}$$

を用いる.

定 理 5.4.5 $n = 1, 2, \dots$ に対して,

$$W(2k|2n) = p_{2k}p_{2n-2k} = \binom{2k}{k} \binom{2n-2k}{n-k} \left(\frac{1}{2}\right)^{2n}, \quad k = 0, 1, \dots, n.$$

証 明 まず,

$$W(2k|2n) = \sum_{r=1}^k f_{2r} W(2k-2r|2n-2r) + \sum_{r=1}^{n-k} f_{2r} W(2k|2n-2r)$$

がわかる. これを用いて, k, n に関する帰納法によって証明ができる (レポート問題 17).

$n \rightarrow \infty$ で近似を求めてみよう. $0 < a < b < 1$ として,

$$\begin{aligned} P(a < \text{正の部分にある時間の割合} < b) &= \sum_{k=an}^{bn} W(2k|2n) \\ &= \sum_{k=0}^n \chi_{[an, bn]}(k) W(2k|2n) \\ &= \sum_{k=0}^n \chi_{[a, b]} \left(\frac{k}{n} \right) \binom{k}{n} \binom{2k}{k} \binom{2n-2k}{n-k} \left(\frac{1}{2} \right)^{2n}. \end{aligned}$$

ここで $\chi_I(x)$ は区間 I の指示関数であり, $x \in I$ のとき 1, それ以外では 0 をとる.

スターリングの公式

$$n! \sim \sqrt{2\pi n} \left(\frac{n}{e} \right)^n \quad n \rightarrow \infty,$$

によって,

$$\binom{2k}{k} \left(\frac{1}{2} \right)^{2k} \sim \frac{1}{\sqrt{\pi k}}$$

よって,

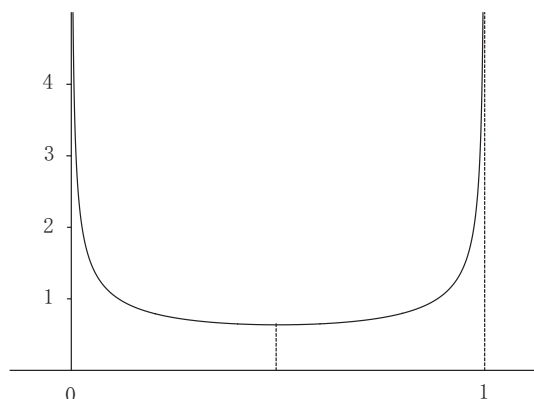
$$\begin{aligned} P(a < \text{正の部分にある時間の割合} < b) &\sim \sum_{k=0}^n \chi_{[a, b]} \left(\frac{k}{n} \right) \frac{1}{\pi \sqrt{k(n-k)}} \\ &= \sum_{k=0}^n \chi_{[a, b]} \left(\frac{k}{n} \right) \frac{1}{\pi \sqrt{\frac{k}{n} (1 - \frac{k}{n})}} \frac{1}{n} \\ &\rightarrow \int_0^1 \chi_{[a, b]}(x) \frac{dx}{\pi \sqrt{x(1-x)}}. \end{aligned}$$

定義 5.4.6 密度関数

$$\frac{dx}{\pi\sqrt{x(1-x)}}. \quad 0 < x < 1,$$

によって定義される確率分布を逆正弦則 (arcsine law) という. 分布関数は

$$F(x) = \int_0^x \frac{dt}{\pi\sqrt{t(1-t)}} = \frac{2}{\pi} \arcsin \sqrt{x} = \frac{1}{2} + \frac{1}{\pi} \arcsin(2x-1).$$



たとえば,

$$F(0.9) = \frac{2}{\pi} \arcsin \sqrt{0.9} \approx 0.795$$

つまり, 長時間ゲームをしたとき, その時間の 90% 以上勝っている確率は $1 - F(0.9) \approx 0.205$ もある!

レポート問題 17 定理 5.4.5 をきちんと証明せよ.

レポート問題 18 スターリングの公式の意味を正しく説明し, その証明の概略を述べよ. 次に, スターリングの公式を用いて,

$$\binom{2n}{n} \left(\frac{1}{2}\right)^{2n} \sim \frac{1}{\sqrt{\pi n}}$$

を導出せよ.

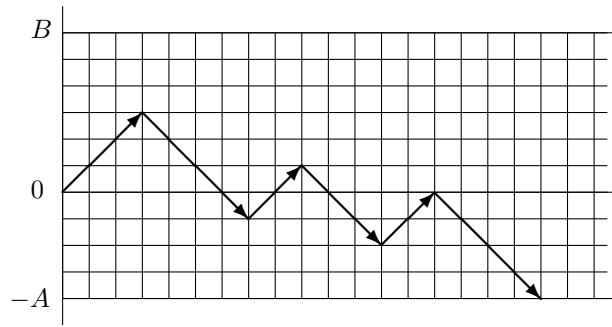
5.5 ギャンブラーの破産問題

時刻 $n = 0$ で原点 0 から出発する $\mathbb{Z}$ 上のランダム・ウォーカーを考えよう. ただし, $-A$ と B ($A \geq 1$, $B \geq 1$) に壁があり, いずれかの壁に触れたらそれ以降, その位置にとどまり続ける (あるいは, その動きを停止する) という条件をつける. このような壁を吸収壁という.

成功確率 $0 < p < 1$ のベルヌイ試行列を $Z_1, Z_2, \dots$ とし, 確率変数列 $X_0, X_1, X_2, \dots$ を次のように定義する.

$$X_0 = 0, \quad X_n = \begin{cases} X_{n-1} + Z_n, & -A < X_{n-1} < B, \\ -A, & X_{n-1} = -A, \\ B, & X_{n-1} = B. \end{cases} \quad (5.11)$$

このような確率変数列 $\{X_n\}$ を吸収壁をもつランダム・ウォークという.



壁に吸収される確率に興味がある。つまり,

$$R = P(\text{ある時刻 } n \text{ で } X_n = -A) = P\left(\bigcup_{n=1}^{\infty} \{X_n = -A\}\right),$$

$$S = P(\text{ある時刻 } n \text{ で } X_n = B) = P\left(\bigcup_{n=1}^{\infty} \{X_n = B\}\right)$$

を計算したい。右辺の和事象は互いに排反になっていないことに注意しておく。

漸化式による巧妙な方法で R, S を求めてみよう。今考えているランダム・ウォークの初期条件は $X_0 = 0$ である。この初期条件を一般化して、 $-A \leq k \leq B$ を満たす整数 k に対して、初期条件を $X_0 = k$ で置き換えた吸収壁をもつランダム・ウォークを $X_n^{(k)}$ で表そう。従来のものは $X_n = X_n^{(0)}$ ということになる。ランダム・ウォーク $X_n^{(k)}$ が壁 $-A, B$ に吸収される確率をそれぞれ R_k, S_k とおく。 $R = R_0, S = S_0$ となる。

補題 5.5.1 $\{R_k; -A \leq k \leq B\}$ は次の漸化式と境界条件をみたす。

$$R_k = pR_{k+1} + qR_{k-1}, \quad R_{-A} = 1, \quad R_B = 0. \quad (5.12)$$

同様に、 $\{S_k; -A \leq k \leq B\}$ は次の漸化式と境界条件をみたす。

$$S_k = pS_{k+1} + qS_{k-1}, \quad S_{-A} = 0, \quad S_B = 1. \quad (5.13)$$

定理 5.5.2 $A \geq 1, B \geq 1$ を整数とする。 $\{X_n\}$ を $-A$ と B に吸収壁をもつランダム・ウォークで (5.11) で定義されるものとする。このランダム・ウォークが壁に吸収される確率は、

$$P(\text{ある時刻 } n \text{ で } X_n = -A) = \begin{cases} \frac{(q/p)^A - (q/p)^{A+B}}{1 - (q/p)^{A+B}}, & p \neq q, \\ \frac{B}{A+B}, & p = q = \frac{1}{2}, \end{cases}$$

$$P(\text{ある時刻 } n \text{ で } X_n = B) = \begin{cases} \frac{1 - (q/p)^A}{1 - (q/p)^{A+B}}, & p \neq q, \\ \frac{A}{A+B}, & p = q = \frac{1}{2}, \end{cases}$$

で与えられる。特に、このランダム・ウォークは確率 1 で壁に吸収される。

吸収壁をもつランダム・ウォークは、ギャンブラーの破産問題 (the gambler's ruin problem) に明快な視点と解法をもたらした。 A, B の 2 人が公正なコインによる賭けをする。それぞれの持ち点を A, B として、コイン投げの勝負によって 1 点ずつやり取りするものとする。つまり、コイン投げに勝てば相手から 1

点を受け取り、負ければ相手に 1 点を譲り渡す。このゲームは、一方の持ち点が 0 になった段階で終了し、勝者は $A + B$ 点を獲得するものとする。A, B がそれぞれ勝利する確率に関心がある。 n 回目のゲームが終了した時点での A の持ち点を $A + X_n$ として、得点の収支を X_n とおくと、 $\{X_n\}$ は吸収壁をもつランダム・ウォークである。定理 5.5.2 によって、A, B がそれぞれが勝つ確率 $P(A), P(B)$ は、

$$P(A) = \frac{A}{A+B}, \quad P(B) = \frac{B}{A+B} \quad (5.14)$$

となり、ゲーム開始時の持ち点に比例することがわかる。たとえば、 $A = 1, B = 100$ なら $P(A) = 1/101$, $P(B) = 100/101$ となり、A に勝ち目はないだろう。

公正なコインによる賭けでは、定理 5.2.2 によって原点再帰性が保障されている。どんなに負けが込んでも賭けを継続しさえすれば、将来必ず収支が 0 に戻ってくるのである。しかしながら、現実には、資金力という壁が存在する。(5.14) はその壁の効果を教えている。資金が少ないギャンブラーは膨大な資金力を有する胴元には勝てないということである。

次に、ゲームが終了するまでに要するコイン投げの平均回数に興味がある。

定理 5.5.3 $\{X_n\}$ を定理 5.5.2 と同じ吸収壁をもつランダム・ウォークとする。このランダム・ウォークが壁に吸収されるまでに要する時間の平均値は次のようになる。

$$\begin{cases} \frac{A}{q-p} - \frac{A+B}{q-p} \frac{1-(q/p)^A}{1-(q/p)^{A+B}}, & p \neq q, \\ AB, & p = q = \frac{1}{2}. \end{cases}$$

証明 時刻 0 で位置 k ($-A \leq k \leq B$) から出発して、壁に吸収されるまでに要する時間を Y_k とする。つまり、

$$Y_k = \min\{j \geq 0; X_j^{(k)} = -A \text{ または } X_j^{(k)} = B\}.$$

求めたいものは $E(Y_0)$ である。定義から明らかに、

$$E(Y_{-A}) = E(Y_B) = 0 \quad (5.15)$$

である。 $-A < k < B$ のとき、

$$E(Y_k) = \sum_{j=1}^{\infty} j P(Y_k = j) \quad (5.16)$$

を 2 つに分けて計算する。アイデアは定理 5.5.2 の証明と同じであり、

$$P(Y_k = j) = pP(Y_{k+1} = j-1) + qP(Y_{k-1} = j-1)$$

を (5.16) に代入すれば、

$$\begin{aligned} E(Y_k) &= p \sum_{j=1}^{\infty} j P(Y_{k+1} = j-1) + q \sum_{j=1}^{\infty} j P(Y_{k-1} = j-1) \\ &= pE(Y_{k+1}) + qE(Y_{k-1}) + 1. \end{aligned} \quad (5.17)$$

こうして、 $E(Y_k)$ は漸化式 (5.17) の解で境界条件 (5.15) を満たすものとして一意的に定まる。この漸化式は標準的な方法で解くことができ、

$$E(Y_k) = \begin{cases} \frac{A+k}{q-p} - \frac{A+B}{q-p} \frac{1-(q/p)^{A+k}}{1-(q/p)^{A+B}}, & p \neq q, \\ (A+k)(B-k), & p = q = \frac{1}{2}. \end{cases}$$

が得られる。ここで $k = 0$ とおけば、求めたい平均値が得られる。 ■

たとえば, $p = q = 1/2$ で, $A = 1, B = 100$ なら吸収されるまでの平均時間は $AB = 100$ である。ギャンブラー A は資金量で圧倒的に劣るのであるが (既にみたように, A が勝つ確率は $1/101$ である), いずれかが破産するまでの平均時間は 100 となり, 意外と長時間にわたりゲームを楽しむことができる。その間, 夢を見ることができるので, ギャンブルはやめられないのであろう。

注意 5.5.4 吸収壁に対して反射壁というものがある。上で考えたのと同様に, $A \geq 1, B \geq 1$ とともに整数の定数として, 壁の位置を $-A$ と B とする。正負両側に壁があるため無限の領域を自由に行き来できないのは同じであるが, いずれかの壁に触れたら, 次の時刻には壁を離れるような動きをするランダムウォークを反射壁をもつランダム・ウォークという。式で表そう。成功確率 $0 < p < 1$ のベルヌイ試行列を $Z_1, Z_2, \dots$ とし, 確率変数列 $X_0, X_1, X_2, \dots$ を次のように定義する。

$$X_0 = 0, \quad X_n = \begin{cases} X_{n-1} + Z_n, & -A < X_{n-1} < B, \\ -A + 1, & X_{n-1} = -A, \\ B - 1, & X_{n-1} = B. \end{cases} \quad (5.18)$$

レポート問題 19 時刻 $t = 0$ で原点 0 を出発し, 半直線 $\{0, 1, 2, \dots\}$ 上を運動する等方的なランダム・ウォークを考える。ただし, 原点 0 は反射壁とする。このランダム・ウォーカーが時刻 $t = 2n$ で原点にいる確率 P_n を求めよ。

第6章 マルコフ連鎖

直線や格子上のランダム・ウォークは、過去の履歴によらず、現時点の位置をもとに次の時点に移動する位置が確率的に定まる。ランダム・ウォーカーは各時点でどこにいるかという「状態」を変化させながら移動してゆく。「状態」は直線や格子の点（座標で表わされる）である必要はなく、より一般の「状態」を考えることができる。過去の履歴によらず、現時点の状態から次の状態に遷移する確率が定まっているような離散時間確率過程がマルコフ連鎖である。たとえば、将棋の対戦では盤面の配置（と持ち駒）が変化してゆく。このとき、盤面の配置が状態ということになり、次の一手で状態に遷移が起こる。次の一手は何通りかある駒の動かし方から1つを選んだものである。下手な人は次の一手を可能なパターンから等確率（に近い確率）で選んでいるだろうし、上級者は特定の一手が選ばれる確率が他に比べて高くなっているだろう。将棋では次の一手が現在の盤面で決められるのでマルコフ連鎖でモデル化される。（囲碁のルールを知っている方：囲碁はマルコフ的ではないがなぜでしょう?）

6.1 条件付き確率

2つの事象 A, B に対して、 $P(B) > 0$ のとき

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

を、条件 B のもとの A の条件付き確率という。これは、事象 B が起こったときに A が起こる確率を与えるものと解釈される。

定 理 6.1.1 2つの事象 A, B に対して次の5つの性質は同値である。

- (i) A と B が独立である。
- (ii) $P(B|A) = P(B)$.
- (ii') $P(A|B) = P(A)$.
- (iii) $P(B|A) = P(B|A^c)$.
- (iii') $P(A|B) = P(A|B^c)$.

ただし、条件付確率を含む性質では、それが定義されることが仮定されているものとする。

証 明 A と B の役割を入れ替えればよいので、(i) $\iff$ (ii) $\iff$ (iii) を示せばよい。(i) $\iff$ (ii) は省略する（試みよ）。

(ii) $\implies$ (iii) を示そう。まず、条件付確率の定義から

$$P(B|A^c) = \frac{P(A^c \cap B)}{P(A^c)} = \frac{P(B) - P(A \cap B)}{1 - P(A)}.$$

仮定から $P(A \cap B) = P(A)P(B)$ が成り立つので、

$$= \frac{P(B) - P(A)P(B)}{1 - P(A)} = \frac{P(B)(1 - P(A))}{1 - P(A)} = P(B).$$

よって, $P(B|A^c) = P(B) = P(B|A)$ が示された.

(iii) $\Rightarrow$ (ii). 定義によって,

$$P(B|A) = \frac{P(A \cap B)}{P(A)}, \quad P(B|A^c) = \frac{P(A^c \cap B)}{P(A^c)} = \frac{P(B) - P(A \cap B)}{1 - P(A)}.$$

仮定からこれらは等しいので,

$$\frac{P(A \cap B)}{P(A)} = \frac{P(B) - P(A \cap B)}{1 - P(A)}.$$

整理すれば, $P(A \cap B) = P(A)P(B)$ が得られ, A, B は独立である. ■

$P(A|B) = P(A)$ であれば, 事象 B の生起に無関係に A の確率が決まることを意味する. これが, 事象の独立性の意味である.

定 理 6.1.2 (乗法法則) 事象 $A_1, A_2, \dots, A_n$ に対して,

$$\begin{aligned} & P(A_1 \cap A_2 \cap \dots \cap A_n) \\ &= P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A_2) \dots P(A_n|A_1 \cap A_2 \cap \dots \cap A_{n-1}). \end{aligned} \quad (6.1)$$

証明は数学的帰納法によればよい. (6.1) がいわゆる樹形図の考え方の根拠になる. たとえば, 2 つの事象 A, B に対して次の図を考えよう. 左から右に矢印を追うことで, $P(A \cap B)$ などの確率を求めることが

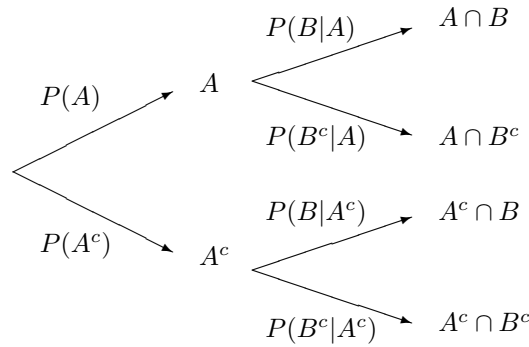


図 6.1: 樹形図

できる. 樹形図は複数の事象 (上の例では A と B) の生起を分類した図であり, 一般には A と B に時間順序や因果関係があるわけではない. もちろん, 矢印の向きに時間が経過するような状況 (ゲームの 1 回戦, 2 回戦, $\dots$ など時間経過を伴う偶然現象) の解析において樹形図がとても有用であることは言うまでもない.

レポート問題 20 (全確率の公式) $A_1, \dots, A_n$ を Ω の事象であって,

$$\Omega = \bigcup_{k=1}^n A_k, \quad A_j \cap A_k = \emptyset \quad (j \neq k)$$

を満たしているものとする. このとき, 任意の事象 B に対して

$$P(B) = \sum_{k=1}^n P(A_k)P(B|A_k)$$

が成り立つ. このことを証明せよ.

6.2 マルコフ連鎖

高々可算集合 (有限集合または可算集合) S に値をとる離散時間の確率過程 $\{X_n; n = 0, 1, 2, \dots\}$ を考えよう. $S \subset \mathbb{R}$ とは限らないほうが便利であるが, 以下では $S = \{1, 2, \dots, N\}$ や $S = \{0, 1, 2, \dots\}$ などを考える. S を状態空間という. この章では, 状態空間 S は常に高々可算集合であるものとする.

定義 6.2.1 S を状態空間とする離散時間の確率過程 $\{X_n; n = 0, 1, 2, \dots\}$ が, 任意の $0 \leq i_1 < i_2 < \dots < i_k < n$ と $a_1, a_2, \dots, a_k, a \in S$ に対して,

$$P(X_n = a | X_{i_1} = a_1, X_{i_2} = a_2, \dots, X_{i_k} = a_k) = P(X_n = a | X_{i_k} = a_k)$$

を満たすとき, $\{X_n\}$ を S 上のマルコフ連鎖という.

マルコフ連鎖は直近の過去にだけ依存して定まる確率過程である.

定義 6.2.2 S 上のマルコフ連鎖 $\{X_n\}$ に対して,

$$P(X_{n+1} = j | X_n = i)$$

を時刻 n における (i から j への) 推移確率という. これが, n によらず一定であるマルコフ連鎖は時間的に一様 (または時間的に斉次) であるという. このとき,

$$p_{ij} = p(i, j) = P(X_{n+1} = j | X_n = i)$$

と書いて, 推移確率という. また, 行列

$$P = [p_{ij}]$$

を推移確率行列という. (行列といっても正方形に配列する必要はない. そうしたければ, S の要素に適当に番号付けをする必要がある.)

明らかに, 任意の $i \in S$ に対して

$$\sum_{j \in S} p(i, j) = \sum_{j \in S} P(X_{n+1} = j | X_n = i) = 1.$$

これを踏まえて, 次の定義を与える.

定義 6.2.3 S を添字とする行列 $P = [p_{ij}]$ は, 次の 2 条件をみたすとき確率行列と呼ばれる.

- (i) $p_{ij} \geq 0$.
- (ii) $\sum_{j \in S} p_{ij} = 1$.

定理 6.2.4 マルコフ連鎖の推移確率行列は (定義 6.2.3 の意味で) 確率行列である. 逆に, 与えられた確率行列を推移確率行列とするマルコフ連鎖が構成できる.

マルコフ連鎖をグラフ表示するのが便利である. 各状態を点で表わし, $p(i, j) > 0$ のときに i と j をその向きの矢印で結んでできるグラフを状態遷移図という.

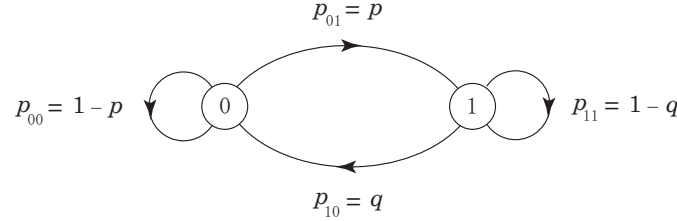
例 6.2.5 (2 状態のマルコフ連鎖) 2 状態 $\{0, 1\}$ 上のマルコフ連鎖は, 推移確率

$$p(0, 1) = p, \quad p(0, 0) = 1 - p, \quad p(1, 0) = q, \quad p(1, 1) = 1 - q$$

で決まる。推移確率行列は,

$$\begin{bmatrix} 1-p & p \\ q & 1-q \end{bmatrix}$$

となる。このマルコフ連鎖の状態遷移図は次のようになる。



例 6.2.6 ($\mathbb{Z}^1$ 上のランダム・ウォーク) 推移確率は,

$$p(i, j) = \begin{cases} p, & j = i + 1 \text{ のとき,} \\ q = 1 - p, & j = i - 1 \text{ のとき,} \\ 0, & \text{その他.} \end{cases}$$

推移確率行列は, 両側に無限に伸びた行列となる:

$$\begin{bmatrix} \ddots & \ddots & \ddots & \ddots & & \\ & \ddots & q & 0 & p & 0 \\ & & 0 & q & 0 & p & 0 \\ & & & 0 & q & 0 & p & 0 \\ & & & & 0 & q & 0 & p & \ddots \\ & & & & & \ddots & \ddots & \ddots & \ddots \end{bmatrix}$$

例 6.2.7 (吸収壁のあるランダム・ウォーク) $A > 0, B > 0$ とする. $-A$ と B に吸収壁のあるランダムウォークの状態空間は $S = \{-A, -A+1, \dots, B-1, B\}$ である. 推移確率は次のようである. $-A < i < B$ のときは,

$$p(i, j) = \begin{cases} p, & j = i + 1 \text{ のとき,} \\ q = 1 - p, & j = i - 1 \text{ のとき,} \\ 0, & \text{その他.} \end{cases}$$

$i = -A$ または $i = B$ のときは,

$$p(-A, j) = \begin{cases} 1, & j = -A \text{ のとき,} \\ 0, & \text{その他,} \end{cases} \quad p(B, j) = \begin{cases} 1, & j = B \text{ のとき,} \\ 0, & \text{その他.} \end{cases}$$

行列で書けば,

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & \cdots & 0 \\ q & 0 & p & 0 & 0 & \cdots & 0 \\ 0 & q & 0 & p & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & q & 0 & p & 0 \\ 0 & 0 & \cdots & 0 & q & 0 & p \\ 0 & 0 & \cdots & 0 & 0 & 0 & 1 \end{bmatrix}$$

例 6.2.8 (反射壁のあるランダム・ウォーク) $A > 0, B > 0$ とする. $-A$ と B に吸収壁のあるランダムウォークの状態空間は $S = \{-A, -A+1, \dots, B-1, B\}$ である. 推移確率は次のようである. $-A < i < B$ のときは,

$$p(i, j) = \begin{cases} p, & j = i + 1 \text{ のとき,} \\ q = 1 - p, & j = i - 1 \text{ のとき,} \\ 0, & \text{その他.} \end{cases}$$

$i = -A$ または $i = B$ のときは,

$$p(-A, j) = \begin{cases} 1, & j = -A + 1 \text{ のとき,} \\ 0, & \text{その他,} \end{cases} \quad p(B, j) = \begin{cases} 1, & j = B - 1 \text{ のとき,} \\ 0, & \text{その他.} \end{cases}$$

行列で書けば,

$$\begin{bmatrix} 0 & 1 & 0 & 0 & 0 & \cdots & 0 \\ q & 0 & p & 0 & 0 & \cdots & 0 \\ 0 & q & 0 & p & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & q & 0 & p & 0 \\ 0 & 0 & \cdots & 0 & q & 0 & p \\ 0 & 0 & \cdots & 0 & 0 & 1 & 0 \end{bmatrix}$$

定 理 6.2.9 n ステップ推移確率が,

$$p_n(i, j) = P(X_{m+n} = j | X_m = i), \quad i, j \in S,$$

で定義される. (時間的に一様との仮定から, 右辺は m の取り方によらない.) このとき, 任意の $0 \leq r \leq n$ に対して,

$$p_n(i, j) = \sum_{k \in S} p_r(i, k) p_{n-r}(k, j) \quad (6.2)$$

が成り立つ. これをチャップマン-コルモゴロフ方程式という.

証 明

$$p_n(i, j) = P(X_{m+n} = j | X_m = i) = \sum_{k \in S} P(X_{m+n} = j, X_{m+r} = k | X_m = i)$$

は明らかであろう. さらに,

$$\begin{aligned} P(X_{m+n} = j, X_{m+r} = k | X_m = i) &= \frac{P(X_{m+n} = j, X_{m+r} = k, X_m = i)}{P(X_{m+r} = k, X_m = i)} \times \frac{P(X_{m+r} = k, X_m = i)}{P(X_m = i)} \\ &= P(X_{m+n} = j | X_{m+r} = k, X_m = i) P(X_{m+r} = k | X_m = i). \end{aligned}$$

マルコフ性によって,

$$P(X_{m+n} = j | X_{m+r} = k, X_m = i) = P(X_{m+n} = j | X_{m+r} = k)$$

であるから,

$$P(X_{m+n} = j, X_{m+r} = k | X_m = i) = P(X_{m+n} = j | X_{m+r} = k) P(X_{m+r} = k | X_m = i).$$

最後に, 時間的に一様であることから,

$$P(X_{m+n} = j, X_{m+r} = k | X_m = i) = p_{n-r}(k, j) p_r(i, k).$$

こうして, (6.2) が得られた. ■

(6.2) を繰り返し, 1 ステップ推移確率 $p_1(i, j) = p(i, j)$ を用いれば,

$$p_n(i, j) = \sum_{k_1, \dots, k_{n-1} \in S} p(i, k_1)p(k_1, k_2) \cdots p(k_{n-1}, j). \quad (6.3)$$

右辺が行列の積であることに注意すれば, n ステップ推移確率 $p_n(i, j)$ は推移確率行列の n 乗の (i, j) 成分に他ならない.

定 理 6.2.10

$$P(X_{m+n} = j | X_m = i) = p_n(i, j) = (P^n)_{ij}$$

注意 6.2.11 そうすると, チャップマン-コルモゴロフ方程式は, 行列の積法則

$$P^n = P^r P^{n-r}$$

を成分で表示したに過ぎないことがわかる. (いつもどおり, $P^0 = E$ (単位行列) とする.)

マルコフ連鎖 $\{X_n\}$ に対して, 時刻 n で状態 $i \in S$ にある確率を

$$\pi_i(n) = P(X_n = i)$$

とおく. これを横ベクトル

$$\pi(n) = [\cdots \pi_i(n) \cdots]$$

で表わし, X_n の分布という. 特に, X_0 の分布 $\pi(0)$ を初期分布という.

定 理 6.2.12

$$\pi(n) = \pi(0)P^n$$

成分で表示すれば,

$$\pi_i(n) = \sum_{j_1, j_2, \dots, j_n} \pi_{j_1}(0)p(j_1, j_2)p(j_2, j_3) \cdots p(j_{n-1}, j_n)p(j_n, i).$$

証 明 まず,

$$\pi_i(n) = P(X_n = i) = \sum_{k \in S} P(X_n = i | X_0 = k)P(X_0 = k) = \sum_{k \in S} \pi_k(0)p_n(k, i)$$

に注意する. $p_n(k, i) = (P^n)_{ki}$ であり, 右辺はベクトルと行列の積になっている. したがって,

$$\pi_i(n) = (\pi(0)P^n)_i.$$

つまり,

$$\pi(n) = \pi(0)P^n$$

が得られた. ■

レポート問題 21 A 党と B 党の支持者は選挙のたびに, 過去の支持政党によらず一定の割合で入れ替わるという. 選挙のたびに, A 党の支持者のうち 30% が B 党に鞍替えし, B 党の支持者のうち 20% が A 党に鞍替えするとする. 当初, A 党の支持者は有権者は 80%, B 党の支持者は有権者は 20% であったという. 何回の選挙ののち, B 党が過半数を制するか? また, この状況が続くとき, 最終的な支持者の割合を求めよ.

6.3 状態の再帰性

引き続き、高々可算集合 S を状態空間とする (時間的に一様な) マルコフ連鎖 $\{X_n\}$ を考える.

定義 6.3.1 2つの状態 $i, j \in S$ に対して, $p_n(i, j) > 0$ となる $n \geq 0$ が存在するとき, i は j に推移可能であるという. 定義によって, i はそれ自身に推移可能である. 2つの状態 $i, j \in S$ に対して, i が j に推移可能かつ j が i に推移可能であるとき, つまり, $p_n(i, j) > 0$ かつ $p_m(j, i) > 0$ を満たす $m \geq 0$ と $n \geq 0$ が存在するとき, i, j は互いに推移可能であるという.

補題 6.3.2 2つの状態 $i, j \in S$ に対して, それらが互いに推移可能であるときに, $i \sim j$ と記すことにする. このとき, $\sim$ は S 上の同値関係になる. すなわち,

- (i) $i \sim i$;
- (ii) $i \sim j$ ならば $j \sim i$;
- (iii) $i \sim j$ かつ $j \sim k$ ならば $i \sim k$.

このことを示せ.

証明 (i), (ii) は定義から明らか. (iii) はチャップマン・コルモゴロフの方程式からわかる (確かめよ). ■

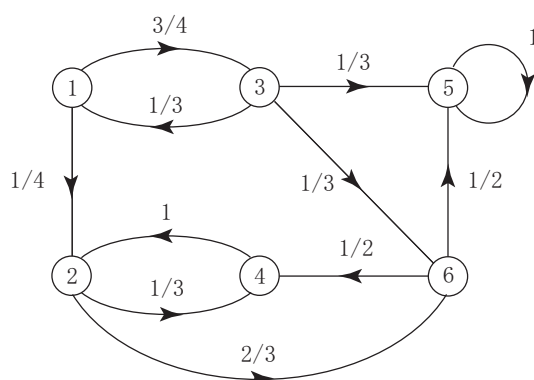
定義 6.3.3 i が吸収状態であるとは,

$$p(i, j) = \begin{cases} 1, & j = i \text{ のとき} \\ 0, & \text{その他,} \end{cases}$$

となるものをいう. 吸収状態は, それ単独でひとつの同値類をなす.

定義 6.3.4 すべての状態が同値であるようなマルコフ連鎖は既約であるという.

例 6.3.5 次の状態遷移図で表わされるマルコフ連鎖について, 状態の同値類別はどうなるだろうか.



定義 6.3.6 状態 $i \in S$ に対して, 初めて i に達するのに要する時間 (初到達時間, first hitting time) が

$$T_i = \inf\{n \geq 1; X_n = i\}$$

で定義される。もし、 $X_n = i$ となる $n \geq 1$ が存在しないときは $T_i = \infty$ とおく。 $P(T_i < \infty | X_0 = i) = 1$ のとき、状態 i は再帰的であるという。そうでないとき、つまり $P(T_i = \infty | X_0 = i) > 0$ のとき、非再帰的であるという。

つまり、状態 i が再帰的であるとは、ある時刻に状態 i を観測すれば、いつとは言えないが有限時間のうちに再び状態 i が観測される確率が 1 であることを意味する。再帰性の判定に関しては、次の強力な結果が知られている。

定 理 6.3.7 状態 $i \in S$ が再帰的であるための必要十分条件は、

$$\sum_{n=0}^{\infty} p_n(i, i) = \infty$$

である。

証 明 (1 次元ランダムウォークのときに用いた方法と基本的に同じである。) まず、

$$\begin{aligned} p_n(i, j) &= P(X_n = j | X_0 = i), \quad n = 0, 1, 2, \dots, \\ f_n(i, j) &= P(T_j = n | X_0 = i) \\ &= P(X_1 \neq j, \dots, X_{n-1} \neq j, X_n = j | X_0 = i), \quad n = 1, 2, \dots, \end{aligned}$$

とおく。 $p_n(i, j)$ はすでに導入した n ステップ推移確率である。一方、 $f_n(i, j)$ は i を出発して n ステップ後に初めて j に達する確率である。このとき、 i から出発して n ステップで j に至る経路を、初めて j に至るまでのステップ数で分類して和をとれば、

$$p_n(i, j) = \sum_{r=1}^n f_r(i, j) p_{n-r}(j, j), \quad i, j \in S, \quad n = 1, 2, \dots \quad (6.4)$$

が成り立つ。

次に、母関数を導入する。

$$\begin{aligned} G_{ij}(z) &= \sum_{n=0}^{\infty} p_n(i, j) z^n, \\ F_{ij}(z) &= \sum_{n=1}^{\infty} f_n(i, j) z^n, \end{aligned}$$

とおく。(6.4) を用いて、少し計算すれば、

$$G_{ij}(z) = p_0(i, j) + F_{ij}(z) G_{jj}(z) \quad (6.5)$$

がわかる。(6.5) で $i = j$ とおけば、

$$G_{ii}(z) = 1 + F_{ii}(z) G_{ii}(z)$$

が成り立つ。したがって、

$$G_{ii}(z) = \frac{1}{1 - F_{ii}(z)}.$$

一方、

$$G_{ii}(1) = \sum_{n=0}^{\infty} p_n(i, i), \quad F_{ii}(1) = \sum_{n=1}^{\infty} f_n(i, i) = P(T_i < \infty | X_0 = i)$$

であるから、 $F_{ii}(1) = 1$ と $G_{ii}(1) = \infty$ が同値になる。 ■

上の証明の中で、次のことも示されている。

定理 6.3.8 状態 i が非再帰的であるとき,

$$\sum_{n=0}^{\infty} p_n(i, i) < \infty$$

であり,

$$\sum_{n=0}^{\infty} p_n(i, i) = \frac{1}{1 - P(T_i < \infty | X_0 = i)}$$

を満たす.

例 6.3.9 ($\mathbb{Z}$ 上のランダムウォーク) 原点から出発したランダムウォークが原点に戻るのは偶数ステップ後に限る. したがって, $p_{2n}(0, 0)$ を求めて, その総和を計算すればよい. $2n$ ステップ後に原点に戻るということは, 右移動の回数と左移動の回数が等しいことを意味し, 右移動と左移動が現れる順序は任意である. したがって,

$$p_{2n}(0, 0) = \frac{(2n)!}{n!n!} p^n q^n, \quad p + q = 1.$$

(これは, すでに第 6 章 (5.1.1) で見ている.) スターリングの公式

$$n! \sim \sqrt{2\pi n} \left(\frac{n}{e}\right)^n \quad (6.6)$$

によって,

$$p_{2n}(0, 0) \sim \frac{1}{\sqrt{\pi n}} (4pq)^n$$

がわかる. したがって, $p \neq q$ ならば,

$$\sum_{n=0}^{\infty} p_{2n}(0, 0) < \infty.$$

$p = q = \frac{1}{2}$ ならば,

$$\sum_{n=0}^{\infty} p_{2n}(0, 0) = \infty.$$

こうして, 1 次元ランダムウォークは, $p \neq q$ ならば非再帰的, $p = q = \frac{1}{2}$ ならば再帰的である.

レポート問題 22 (1) a_n, b_n を正数列とする. $\lim_{n \rightarrow \infty} \frac{a_n}{b_n} = 1$ のときに, $a_n \sim b_n$ と記す (スターリングの公式はこの意味である). このとき, 定数 $c_1 > 0, c_2 > 0$ で

$$c_1 a_n \leq b_n \leq c_2 a_n$$

を満たすものが存在することを示せ.

(2) (1) の条件の下で, $\sum_{n=1}^{\infty} a_n$ と $\sum_{n=1}^{\infty} b_n$ は同時に収束または発散することを示せ.

(3) $\mathbb{Z}$ 上のランダムウォークの再帰性について, 例 6.3.9 の説明を省略せずに詳しく議論せよ.

例 6.3.10 ($\mathbb{Z}^2$ 上のランダムウォーク) 原点から出発した (等方的) ランダムウォークが原点に戻るのは偶数ステップ後に限る. したがって, $p_{2n}(0, 0)$ を求めて, その総和を計算すればよい. 2 次元ランダムウォー

クでは, x 軸方向と y 軸方向の 2 つの移動方向があるから, $2n$ ステップ後に原点に戻る経路を x 軸方向の移動回数 $2i$ と y 軸方向 $2j$ として数え上げよう. まず,

$$\begin{aligned} p_{2n}(0,0) &= \sum_{i+j=n} \frac{(2n)!}{i!j!j!} \left(\frac{1}{4}\right)^{2n} \\ &= \frac{(2n)!}{n!n!} \left(\frac{1}{4}\right)^{2n} \sum_{i+j=n} \frac{n!n!}{i!i!j!j!} \\ &= \binom{2n}{n} \left(\frac{1}{4}\right)^{2n} \sum_{i=0}^n \binom{n}{i}^2 \end{aligned}$$

がわかる. ここで, 二項係数に関する公式 (確かめよ)

$$\sum_{i=0}^n \binom{n}{i}^2 = \binom{2n}{n} \quad (6.7)$$

を用いれば,

$$p_{2n}(0,0) = \binom{2n}{n}^2 \left(\frac{1}{4}\right)^{2n}$$

が得られる. スターリングの公式によって,

$$p_{2n}(0,0) \sim \frac{1}{\pi n}$$

がわかる (計算で確認せよ) から,

$$\sum_{n=1}^{\infty} p_{2n}(0,0) = \infty$$

となる (詳しく述べよ). よって, 2 次元ランダムウォークは再帰的である.

例 6.3.11 ($\mathbb{Z}^3$ 上のランダムウォーク) ここでも等方的なランダムウォークを考えよう. 原点から出発したランダムウォーカーが原点に戻るのは偶数ステップ後に限る. 3 次元ランダムウォークでは, x 軸, y 軸, z 軸の 3 つの移動方向があるから,

$$\begin{aligned} p_{2n}(0,0) &= \sum_{i+j+k=n} \frac{(2n)!}{i!i!j!j!k!k!} \left(\frac{1}{6}\right)^{2n} \\ &= \frac{(2n)!}{n!n!} \left(\frac{1}{6}\right)^{2n} \sum_{i+j+k=n} \frac{n!n!}{i!i!j!j!k!k!} \\ &= \binom{2n}{n} \left(\frac{1}{6}\right)^{2n} \sum_{i+j+k=n} \left(\frac{n!}{i!j!k!}\right)^2 \end{aligned}$$

がわかる. ここで, 和を評価するために 2 つの事実に注意する. まず, 二項展開 (多項展開) によって,

$$\sum_{i+j+k=n} \frac{n!}{i!j!k!} = 3^n. \quad (6.8)$$

また,

$$M_n = \max_{i+j+k=n} \frac{n!}{i!j!k!}$$

は, $\frac{n}{3} - 1 \leq i, j, k \leq \frac{n}{3} + 1$ のときに与えられる. スターリングの公式で

$$M_n \sim \frac{3\sqrt{3}}{2\pi n} 3^n.$$

そうすれば,

$$p_{2n}(0, 0) \leq \binom{2n}{n} \left(\frac{1}{6}\right)^{2n} 3^n M_n \sim \frac{3\sqrt{3}}{2\pi\sqrt{\pi}} n^{-3/2}$$

が得られる. したがって,

$$\sum_{n=1}^{\infty} p_{2n}(0, 0) < \infty$$

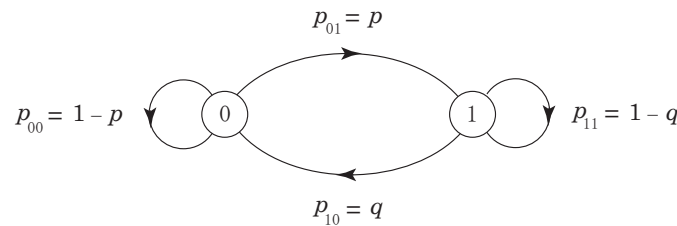
となり, 3次元ランダムウォークは非再帰的である. 酔っ払った人は家に帰れるが, 酔っ払った鳥は家に帰れない (角谷静夫 (1911–2004) の言葉らしい).

状態 i は $P(T_i < \infty | X_0 = i) = 1$ のとき再帰的と呼ばれることはすでに述べた. このときは, 平均再帰時間 $E(T_i | X_0 = i)$ が意味をもつ. 等方的な 1 次元ランダムウォークは再帰的であったが, 平均再帰時間は ∞ であった. このような状態は零再帰的であると呼ばれる. 平均再帰時間 $E(T_i | X_0 = i) < \infty$ となるとき, 状態 i は正再帰的であると呼ばれる.

定理 6.3.12 同一の同値類に属する状態は, すべて同時に正再帰的であるか, 零再帰的であるか, 非再帰的かのいずれかである. 特に, 既約なマルコフ連鎖の状態は, すべて同時に正再帰的であるか, 零再帰的であるか, 非再帰的かのいずれかである.

定理 6.3.13 有限な状態空間 S 上の既約マルコフ連鎖の状態はすべて正再帰的である.

レポート問題 23 次の状態遷移図で表わされる 2 状態のマルコフ連鎖 $\{X_n\}$ を考える.



このとき,

$$P(T_0 = 1 | X_0 = 0), \quad P(T_0 = 2 | X_0 = 0), \quad P(T_0 = 3 | X_0 = 0), \quad P(T_0 = 4 | X_0 = 0)$$

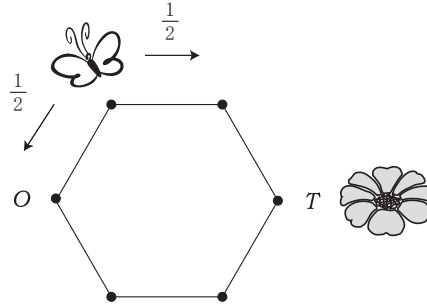
を計算せよ. 次に, $P(T_0 = n | X_0 = 0)$ を求めて,

$$\sum_{n=1}^{\infty} P(T_0 = n | X_0 = 0), \quad \sum_{n=1}^{\infty} n P(T_0 = n | X_0 = 0)$$

を計算せよ.

レポート問題 24 図のような, O と T を定点とする 6 角形の頂点上を蝶が飛び回っている. 時刻 $t = 0$ において 原点 O において, 各時点ごとに確率 $1/2$ で右隣または左隣の頂点に移動する. ただし, T は食虫花であり, T に達した瞬間に蝶は死んでしまう.

- (1) 推移確率行列を求めよ.
- (2) このマルコフ連鎖は既約であるか?
- (3) この蝶が時刻 $t = 12$ においてまだ生きている確率を求めよ.
- (4) 時刻 $t \rightarrow \infty$ のときの様子を考察せよ.



6.4 定常分布

定義 6.4.1 高々可算集合 S を状態空間とするマルコフ連鎖の推移行列を P とする. 横ベクトル $\pi = [\cdots \pi_i \cdots]$ で

$$\pi_i \geq 0, \quad \sum_{i \in S} \pi_i = 1$$

を満たすものを S 上の分布という. S 上の分布 π は

$$\pi = \pi P \quad (6.9)$$

をみたすとき, 定常分布 (不変分布とも) という. 定常分布が存在しないこともあるし, あっても一意とは限らない.

例 6.4.2 (2 状態のマルコフ連鎖) 推移確率行列

$$P = \begin{bmatrix} 1-p & p \\ q & 1-q \end{bmatrix}$$

をもつマルコフ連鎖の定常分布を求めよう. $\pi = [\pi_0 \ \pi_1]$ とおけば, (6.9) は

$$[\pi_0 \ \pi_1] \begin{bmatrix} 1-p & p \\ q & 1-q \end{bmatrix} = [(1-p)\pi_0 + q\pi_1 \quad p\pi_0 + (1-q)\pi_1] = [\pi_0 \ \pi_1]$$

となる. これは,

$$p\pi_0 - q\pi_1 = 0$$

と同値である. これと

$$\pi_0 + \pi_1 = 1$$

を連立させて解く. $p+q > 0$ であれば,

$$\pi_0 = \frac{q}{p+q}, \quad \pi_1 = \frac{p}{p+q},$$

が得られる. $p=q=0$ のときは, π は $\pi_0 + \pi_1 = 1$ を満たす限り任意である.

定常分布の存在と一意性に関する基本定理を述べる. 証明は省略する (国沢「確率論とその応用」, シナジ「マルコフ連鎖から格子確率モデルへ」などを参照).

定理 6.4.3 既約なマルコフ連鎖に対して, 次の 2 条件は同値である:

- (i) 定常分布をもつ.

(ii) すべての状態が正再帰的である.

この条件が成り立つとき, 定常分布 π は一意的であり,

$$\pi_i = \frac{1}{\mathbf{E}(T_i | X_0 = i)}, \quad i \in S,$$

が成り立つ.

有限な状態空間 S 上の既約マルコフ連鎖の状態はすべて正再帰的である (定理 6.3.13) から,

定理 6.4.4 有限な状態空間 S 上の既約マルコフ連鎖はただ 1 つの定常状態 π をもつ. さらに, すべての $i \in S$ に対して $\pi_i > 0$ である.

例 6.4.5 (2 状態のマルコフ連鎖) 推移確率行列

$$P = \begin{bmatrix} 1-p & p \\ q & 1-q \end{bmatrix}$$

に対して, P^n を計算してみよう. ここでは, $p+q > 0$ を仮定する. 固有値を求めると, $1, 1-p-q$ となる. $p+q > 0$ を仮定しているから, この 2 つの固有値は異なる. 固有ベクトルを求めて,

$$T = \begin{bmatrix} 1 & p \\ 1 & -q \end{bmatrix}$$

とくと,

$$PT = T \begin{bmatrix} 1 & 0 \\ 0 & 1-p-q \end{bmatrix}$$

となる. よって,

$$P^n = T \begin{bmatrix} 1 & 0 \\ 0 & (1-p-q)^n \end{bmatrix} T^{-1}$$

ここで, $n \rightarrow \infty$ として,

$$\lim_{n \rightarrow \infty} P^n = T \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} T^{-1} = \frac{1}{p+q} \begin{bmatrix} q & p \\ q & p \end{bmatrix}.$$

一方, $\pi(0) = [\pi_0(0) \ \pi_1(0)]$ を初期分布とするマルコフ連鎖の時刻 n における分布 $\pi(n)$ は, 定理 6.2.12 によって,

$$\pi(n) = \pi(0)P^n$$

で与えられる. ここで $n \rightarrow \infty$ とすれば,

$$\lim_{n \rightarrow \infty} \pi(n) = \frac{1}{p+q} [\pi_0(0) \ \pi_1(0)] \begin{bmatrix} q & p \\ q & p \end{bmatrix} = \frac{1}{p+q} [q \ p] = \left[\frac{q}{p+q} \ \frac{p}{p+q} \right].$$

これは, マルコフ連鎖の定常分布 π である. すなわち, 2 状態のマルコフ連鎖で $p+q > 0$ であれば, 初期分布によらずに, 時間が十分に経過した後は, X_n の分布は定常分布 π に漸近するのである.

どのような仮定のもとで定常分布への収束が言えるかは, 応用上も重要であり, さまざまな結果が知られている. 残念ながら, 定常分布が一意的に存在する (たとえば, 定理 6.4.4 の条件下) だけでは, 収束は言えない.

例 6.4.6 2 状態のマルコフ連鎖で推移行列が

$$P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

で与えられるものを考えよう。まず、定常分布が一意的に存在する (確認せよ)。しかし、任意の初期分布 $\pi(0)$ を与えると、 $\lim_{n \rightarrow \infty} \pi(n)$ が収束するとは限らない。

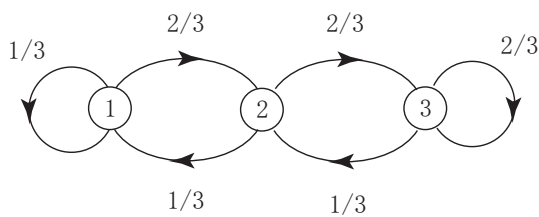
定理 6.4.7 有限な状態空間 S 上の既約マルコフ連鎖の定常状態を π とする (ただ 1 つ存在することは既知)。さらに、 $\{X_n\}$ が非周期的であれば、すべての $j \in S$ に対して

$$\lim_{n \rightarrow \infty} P(X_n = j) = \pi_j$$

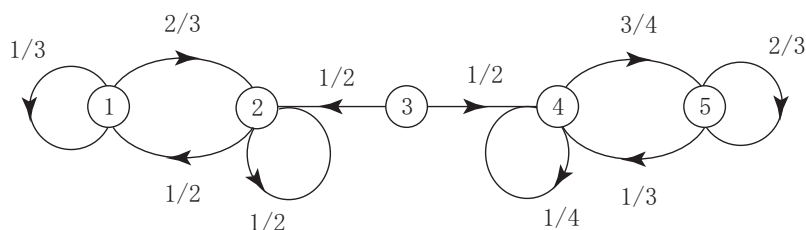
が成り立つ。

ここで、非周期性の条件によって、例 6.4.6 のような状況を排除するのである。詳しくは、国沢「確率論とその応用」、シナジ「マルコフ連鎖から格子確率モデルへ」などを参照されたい。

レポート問題 25 下図で表される状態遷移図をもつマルコフ連鎖は既約であるか調べよ。次に、定常分布について考察せよ。



レポート問題 26 下図で表される状態遷移図をもつマルコフ連鎖は既約であるか調べよ。次に、定常分布について考察せよ。



レポート提出要領

1. 講義中に出題したレポート問題のうち

1 番から 15 番のうち 2 題

16 番から最終番のうち 2 題

の合わせて 4 題を選択して解答せよ (1 題 25 点で採点)。コピーレポートは 0 点。

2. 提出先：情報科学研究科 1 F 事務室前のメールボックスに設置するレポート提出専用のボックス

3. 提出期間：2 月 6 日 (月) - 2 月 10 日 (金)

第7章 ポアソン過程

$T \subset \mathbf{R}$ を区間とすると、確率変数の族 $\{X(t); t \in T\}$ を連続時間の確率過程という。 $T = [0, 1]$ や $T = [0, \infty)$ などを考えることが多い。各 $t \in T$ に対して、 $X(t)$ は確率変数であるから、もうひとつの変数 $\omega \in \Omega$ をもっている。これを明示するときは、 $X(t, \omega)$ や $X_t(\omega)$ のように記す。一つの $\omega \in \Omega$ を固定して、 $X(t, \omega)$ を時間の関数と見るとき、

$$t \mapsto X(t, \omega)$$

を $X(t)$ のサンプルパスという。ランダム現象の時間発展を確率過程でモデル化するとき、現実には多数のサンプルパスのうち一つが実現していると理解するのである。

連続時間の確率過程で最も基本的なものは、ポアソン過程とブラウン運動 (ウィナー過程ともいう) である。講義では触れることができないが、これらを混合したレヴィ過程 (あるいは加法過程) は数理物理や数理ファイナンスなどに多くの応用があり、最近盛んに研究されている。

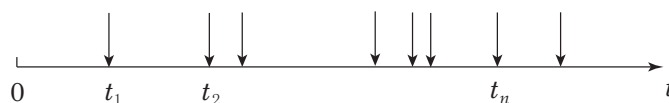
7.1 直感的な導入

時間とともに何らかの個数がランダムに変化する様子を思い浮かべよう。時刻 t におけるその個数を確率過程 $\{X_t\}$ としてモデル化することが課題となる。このとき、 $\{X_t\}$ は $\{0, 1, 2, \dots\}$ に値をとることになるが、このような確率過程を一般に計数過程 (counting process) という。

ランダムに変化する仕方はさまざまであり、計数過程にもさまざまな種類がある。ここでは、時間の経過とともにある特定の事象 E が互いに無関係にランダムに繰り返し起こる状況を考えよう。たとえば、つぎのようなものである：

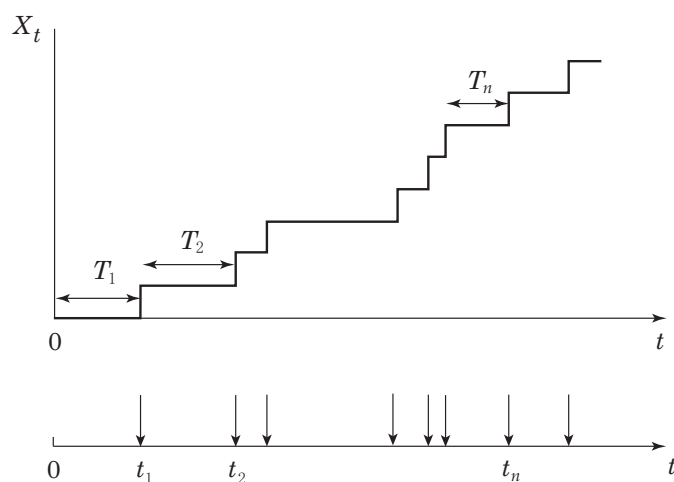
- (i) 携帯電話のメール着信
- (ii) コールセンターにかかってくる電話
- (iii) バス停に並ぶ乗客
- (iv) 商店を訪れる買い物客
- (v) ときどき不良品が現れる製造ライン
- (vi) ある交差点における交通事故
- (vii) 放射線を出して崩壊する原子核

時刻の始点を $t = 0$ として、時刻 t までに起こった E の回数 X_t が興味の対象である。まず、事象 E が起きた時刻 $t_1, t_2, \dots$ を記録すれば、下図のようになる。



このような現象を記述するために注目すべき量は2つある。

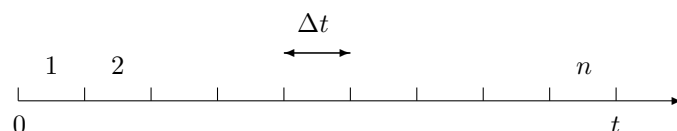
- (i) 時刻 t までに起きた E の回数 X_t . そうすると, $\{X_t, t \geq 0\}$ は計数過程となる.
- (ii) $n-1$ 回目が起きてから n 回目が起きるまでの待ち時間 T_n (ただし, T_1 は初めて E が起こるまでの待ち時間とする). このとき, $\{T_n; n = 1, 2, \dots\}$ は, $[0, \infty)$ に値をとる確率変数列 (離散時間確率過程といってもよい) となる.



ここでは (i) の視点から, いささか直感的に確率過程 (確率モデル) を構成しよう. 最初から連続時間では考えにくいので, $t > 0$ を固定して, 時間区間 $[0, t]$ を n 等分して微小区間に分割する. 微小区間の幅を

$$\Delta t = \frac{t}{n}$$

とおき, 得られた微小区間には順に番号をつけておく.



問題の事象 E の生起に関して次の仮定をおく:

- (1) 各微小時間区間において E の起こる回数の分布は, ある定数 $\lambda > 0$ によって,

$$P(E \text{ が } 1 \text{ 回起こる}) = \lambda \Delta t + o(\Delta t),$$

$$P(E \text{ が起こらない}) = 1 - \lambda \Delta t + o(\Delta t),$$

$$P(E \text{ が } 2 \text{ 回以上起こる}) = o(\Delta t),$$

で与えられる.

- (2) 各微小時間区間において E が起こるかどうかは微小区間ごとに独立である.

現実のランダム現象 (乗客が 1 人ずつ, ぽつりぽつりやってきてバス停に列をなす様子を思い浮かべよう) に即して説明しよう. (1) の $P(E \text{ が } 2 \text{ 回以上起こる}) = o(\Delta t)$ は, Δt が小さければ, 考えている事象 E が 2 つ以上同時には起こらないことを意味する. つまり, E が続けて起こる時の時間差はどんなに短くてもよいが一定時間以上離れているということである. また, $P(E \text{ が } 1 \text{ 回起こる}) = \lambda \Delta t + o(\Delta t)$ は, 時間幅 Δt が小さければ, それに比例して E が起こる確率が増えることを意味する. (2) は, E が起こるかど

うかは、それがすでにいつ何回起こったかによらないことを意味する。先の例では、このような状況が想定されるのである。

さて、第 i 番目の微小区間で E の起こる回数を Z_i で表せば、 $Z_1, Z_2, \dots, Z_n$ は独立同分布の確率変数列であり、

$$P(Z_i = 0) = 1 - \lambda \Delta t + o(\Delta t), \quad P(Z_i = 1) = \lambda \Delta t + o(\Delta t), \quad P(Z_i \geq 2) = o(\Delta t),$$

となる。時間区間 $[0, t]$ において事象 E が起こる回数は、

$$\sum_{i=1}^n Z_i$$

で与えられる。区間幅 Δt を小さくするほど近似の精度が上がると考えられるので、 $\Delta t \rightarrow 0$ あるいは $n \rightarrow \infty$ を考えて、

$$X_t = \lim_{\Delta t \rightarrow 0} \sum_{i=1}^n Z_i \quad (7.1)$$

として X_t を定めれば、連続時間確率過程 $\{X_t\}$ が得られる。これが、時刻 t までに起こった事象の回数を与える確率過程ということになる。このように得られた確率過程をパラメータ λ のポアソン過程という。

(7.1) の Z_i は Δt に (したがって n にも) 依存していることに注意しよう。 Δt が小さくなれば、 n が大きくなり、和をとるべき確率変数の個数が増えるが、各々の Z_i は高い確率で 0 であることから、うまくバランスして極限が定まるのである。実際、

$$P\left(\sum_{i=1}^n Z_i = k\right) = \binom{n}{k} (\lambda \Delta t)^k (1 - \lambda \Delta t)^{n-k} + o(\Delta t),$$

であるから、

$$\Delta t = \frac{t}{n}.$$

に注意して、極限移行すると、

$$P(X_t = k) = \lim_{\Delta t \rightarrow 0} \frac{(\lambda t)^k}{k!} \frac{n(n-1)\dots(n-k+1)}{n^k} \left(1 - \frac{\lambda t}{n}\right)^{n-k} = \frac{(\lambda t)^k}{k!} e^{-\lambda t}.$$

つまり、 X_t の分布はパラメータ λt のポアソン分布である。

注意 7.1.1 上で用いた議論の本質は、「二項分布 $B(n, p)$ は、 n が大きく p が小さいときは、パラメータ np のポアソン分布に近い」というポアソンの少数の法則である (定理 10.0.6)。

定理 7.1.2 ポアソン過程 $\{X_t; t \geq 0\}$ は次の性質をもつ。

- (1) (計数過程) X_t は $\{0, 1, 2, \dots\}$ に値をとる確率変数である。
- (2) $X_0 = 0$.
- (3) (単調増加) $0 \leq s \leq t$ ならば $X_s \leq X_t$
- (4) (独立増分) $0 \leq t_1 < t_2 < \dots < t_k$ ならば、

$$X_{t_2} - X_{t_1}, \quad X_{t_3} - X_{t_2}, \quad \dots, \quad X_{t_k} - X_{t_{k-1}},$$

は独立である。

- (5) (定常性) $0 \leq s < t, h \geq 0$ とするとき、 $X_{t+h} - X_{s+h}$ と $X_t - X_s$ の分布は一致する。

(6) ある定数 $\lambda > 0$ が存在して,

$$P(X_h = 1) = \lambda h + o(h), \quad P(X_h \geq 2) = o(h).$$

証 明 (1) X_t はパラメータ λt のポアソン分布に従う確率変数である. したがって, 非負整数値をとる.

(2) 定義から明らか.

(3) $s = m\Delta t, t = n\Delta t, m < n$ とすれば,

$$X_s = \lim_{\Delta t \rightarrow 0} \sum_{i=1}^m Z_i \leq \lim_{\Delta t \rightarrow 0} \sum_{i=1}^n Z_i = X_t$$

は明らかである.

(4) $t_1 = n_1\Delta t, \dots, t_k = n_k\Delta t, n_1 < \dots < n_k$ とすれば,

$$X_{t_2} - X_{t_1} = \lim_{\Delta t \rightarrow 0} \sum_{i=1}^{n_2} Z_i - \lim_{\Delta t \rightarrow 0} \sum_{i=1}^{n_1} Z_i = \lim_{\Delta t \rightarrow 0} \sum_{i=n_1+1}^{n_2} Z_i$$

となる. つまり, $X_{t_2} - X_{t_1}$ は時間区間 $[t_1, t_2)$ に含まれる微小区間に対応する Z_i の和である. したがって, $X_{t_2} - X_{t_1}, \dots, X_{t_k} - X_{t_{k-1}}$ はそれぞれ共通部分のない時間区間に含まれる微小区間に対応する Z_i の和である. $\{Z_i\}$ は独立な確率変数列であるから, $X_{t_2} - X_{t_1}, \dots, X_{t_k} - X_{t_{k-1}}$ は独立である.

(5) $X_{t+h} - X_{s+h}$ と $X_t - X_s$ を定義するため, 微小区間に対応する Z_i の和で表せば, それらは同数の Z_i から成るので, 分布は一致する.

(6) X_h はパラメータ λh のポアソン分布に従う確率変数である. よって,

$$\begin{aligned} P(X_h = 0) &= e^{-\lambda h} = 1 - \lambda h + \dots = 1 - \lambda h + o(h), \\ P(X_h = 1) &= \lambda h e^{-\lambda h} = \lambda h(1 - \lambda h + \dots) = \lambda h + o(h). \end{aligned}$$

したがって,

$$P(X_h \geq 2) = 1 - P(X_h = 0) - P(X_h = 1) = o(h).$$

■

例 7.1.3 窓口に来る訪問客は 1 分間に平均 2 人である. 訪問客の人数はポアソン過程でモデル化されるものとする. 次の確率を求めよう.

- (1) 窓口を開いてから, 最初の 2 分間に訪問客が全く来ない確率.
- (2) 窓口を開いてから, ある時刻から 2 分間に訪問客が全く来ない確率.
- (3) 最初の 2 分間に訪問客が来ず, 次の 1 分間に 2 人の訪問客がある確率.

時刻 t までに窓口を訪れた訪問客の人数を X_t とする. これがパラメータ $\lambda = 2$ のポアソン過程になる.

(1) $P(X_2 = 0)$ を求めればよい. X_2 はパラメータ $2\lambda = 4$ のポアソン分布に従うから,

$$P(X_2 = 0) = \frac{4^0}{0!} e^{-4} \approx 0.018$$

(2) ある時刻を t_0 とすれば, 求める確率は, $P(X_{t_0+2} - X_{t_0} = 0)$ である. 定常性から,

$$P(X_{t_0+2} - X_{t_0} = 0) = P(X_2 - X_0 = 0) = P(X_2 = 0)$$

なので, (1) と同じである.

(3) 求める確率は $P(X_2 = 0, X_3 - X_2 = 2)$ である. X_2 と $X_3 - X_2$ は独立であるから,

$$P(X_2 = 0, X_3 - X_2 = 2) = P(X_2 = 0)P(X_3 - X_2 = 2)$$

さらに, 定常性から

$$= P(X_2 = 0)P(X_1 = 2) = \frac{4^0}{0!} e^{-4} \times \frac{2^2}{2!} e^{-2} \approx 0.00496$$

レポート問題 27 ある窓口に来る訪問客は 1 時間に平均 20 人である. 訪問客の人数をポアソン過程でモデル化して, 次の量を求めよ.

- (1) 最初の 2 分間に 1 人の訪問客がある確率.
- (2) ある時刻から最初の 2 分間に 1 人の訪問客があり, 次の 3 分間に 2 人の訪問客がある確率.
- (3) 窓口を開けてから最初の訪問客が来るまで 10 分以上待つ確率.

レポート問題 28 $\{X_t\}$ をポアソン過程とする. $0 < s < t$ のとき,

$$P(X_s = k | X_t = n) = \binom{n}{k} \left(\frac{s}{t}\right)^k \left(1 - \frac{s}{t}\right)^{n-k}, \quad k = 0, 1, \dots, n,$$

を示せ. このことを直感的に説明せよ.

7.2 待ち時間

パラメータ λ のポアソン過程 $\{X_t; t \geq 0\}$ を考えよう. ポアソン過程は, 定義によって $X_0 = 0$ であり, 時間とともに 1 ずつ増加する. ポアソン過程は, ランダムに起こる事象が時刻 t までに起きる回数をモデル化したものであった. まず,

$$T_1 = \inf\{t \geq 0; X_t \geq 1\} \quad (7.2)$$

とおく. これは, 初めて E が起こるまでの待ち時間ということになる. T_2 を E が初めて起きてから 2 度目に起こるまでの待ち時間とする. すなわち,

$$T_2 = \inf\{t \geq 0; X_t \geq 2\} - T_1$$

である. 以下同様に,

$$T_n = \inf\{t \geq 0; X_t \geq n\} - T_{n-1}, \quad n = 2, 3, \dots, \quad (7.3)$$

とおけば, T_n は $n-1$ 回目の E が起きてから次の E が起こるまでの待ち時間となる.

定理 7.2.1 パラメータ λ のポアソン過程 $\{X_t\}$ に対して, 待ち時間 T_n を (7.2) と (7.3) で定義する. このとき, $\{T_n; n = 1, 2, \dots\}$ はパラメータ λ の指数分布に従う独立同分布の確率変数列である.

証明 $t = n\Delta t$ として, 微小区間 Δt による近似から導こう. 微小時間区間には Z_i が対応している.

$$\begin{aligned} P(T_1 > t) &= \lim_{\Delta t \rightarrow 0} P(Z_1 = \dots = Z_n = 0) \\ &= \lim_{\Delta t \rightarrow 0} (1 - \lambda\Delta t)^n \\ &= \lim_{\Delta t \rightarrow 0} \left(1 - \frac{\lambda t}{n}\right)^n \\ &= e^{-\lambda t} \end{aligned}$$

したがって、

$$P(T_1 \leq t) = 1 - e^{-\lambda t}.$$

これは、 T_1 がパラメータ λ の指数分布に従うことを示す。

T_2 については、 E が1度起きてから次が起こるまでの待ち時間であるが、微小時間区間による近似の議論から、 T_1 と独立同分布であることは明らかであろう。以下、 $T_3, \dots$ も同様である。 ■

注意 7.2.2 パラメータ λ のポアソン過程 $\{X_t\}$ について $E(X_1) = \lambda$ である。これは、単位時間あたりに起こる事象の平均回数である。したがって、事象の起こる時間間隔の平均は $1/\lambda$ と推測される。定理 7.2.1 によって、待ち時間はパラメータ λ の指数分布に従うが、その平均値は $1/\lambda$ である。したがって、上の推測は結果的に正しい。

レポート問題 29 (これで終わり) パラメータ λ のポアソン過程において、事象 E が n 回起きるまでの待ち時間は $S_n = T_1 + T_2 + \dots + T_n$ で与えられる。ここで T_n は定理 7.2.1 で与えられている。 $P(S_2 \leq t)$ を計算して、 S_2 の密度関数を求めよ。[一般に、 S_n はガンマ分布に従う。]

7.3 ポアソン過程の厳密な定義

(7.1) によるポアソン過程の「定義」はランダム現象のモデル化という観点から直接的で、かつ多くの応用を示唆するものである。しかしながら、数学的には、このままでは確率過程 $\{X_t\}$ を定義したことにはならない。たとえば、確率過程が定義されている確率空間 $(\Omega, \mathcal{F}, P)$ は明らかであろうか?

厳密な定義は待ち時間 $\{T_n\}$ から始めるのがよい。まず、パラメータ $\lambda > 0$ の指数分布に従う独立同分布な確率変数列 $\{T_n; n = 1, 2, \dots\}$ を準備する (もちろん、確率空間 $(\Omega, \mathcal{F}, P)$ をきちんと明示することができますね)。さらに、

$$S_0 = 0, \quad S_n = T_1 + \dots + T_n, \quad n = 1, 2, \dots,$$

としておいて、 $t \geq 0$ に対して、

$$X_t = \max\{n \geq 0; S_n \leq t\}$$

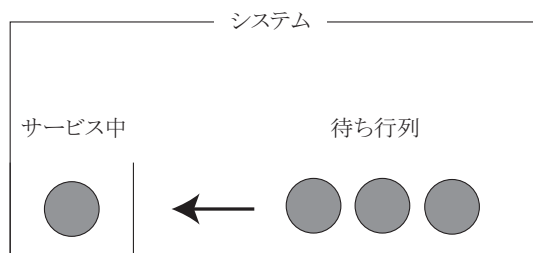
とおく。このとき、各 $t \geq 0$ に対して X_t は $(\{T_n\})$ と同じ確率空間で定義された) 確率変数になる。よって、 $\{X_t; t \geq 0\}$ は連続時間確率過程になる。これをパラメータ λ のポアソン過程というのである。

この定義から定理 7.1.2 が証明される。

7.4 $M/M/1/\infty$ 待ち行列

待ち行列のモデルはケンドールによる記号によって区別される。一般には $A/B/c/K/m/Z$ の形をもつ。 A : 到着時間間隔分布, B : サービス時間分布, c : サービス窓口の個数, K : 最大容量, m : 客の数, Z : サービスの順序 (先着順, ランダム, 優先権ありなど)。

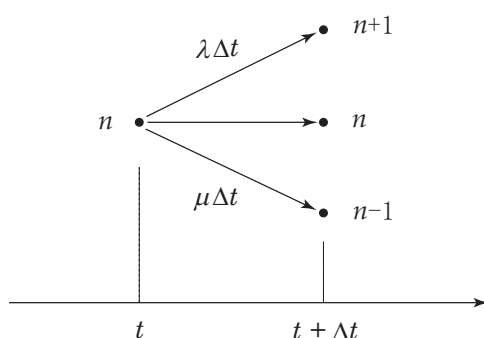
ここで取り上げる $M/M/1/\infty$ はポアソン到着, 指数サービス, 単一窓口, 容量無制限を表す。これは、最も単純な待ち行列モデルといえる。



客はサービスを受けるためにシステムと呼ばれる箱に来る。そこには1つのサービス窓口があり、それが空いていれば直ちにサービスを受けることができるが、それがふさがっている場合は空くまで待たねばならない。すでに待っている客があれば、待ち行列の最後尾に並ぶ。客がポアソン過程に従って到着するとき、つまり、一人の客が来て次の客が来るまでの時間間隔が指数分布に従うとき、ポアソン到着という。また、一人の客に要するサービス時間の分布は指数分布に従うとき、指数サービスという。客の到着とサービスに要する時間は独立であるものとする。時刻 t においてシステム内にいる客（サービス中と待ち行列合わせて）の総数を $X(t)$ とする。この $X(t)$ を確率過程として構成し、その性質を調べることで、待ち行列の様子を解析することができる。

待ち行列を特徴づけるパラメータは2つある。つまり、客の到着を表すポアソン過程のパラメータ $\lambda > 0$ 、とサービス時間を表す指数分布のパラメータ $\mu > 0$ である。客は平均時間間隔 $1/\lambda$ で到着し、平均サービス時間 $1/\mu$ のサービスを受けることになる。待ち行列理論では、 λ を到着率、 μ をサービス率という。

さて、微小時間の経過 $t \rightarrow t + \Delta t$ においてシステムがどのように変化するかを考えよう。微小時間 Δt では事象は高々1つしか起こらないとする。つまり、新たな客が到着する、サービス中の客がシステムを離れる、あるいは何も起こらないものとする。それぞれが起こる確率は、 $\lambda\Delta t$ 、 $\mu\Delta t$ 、 $1 - \lambda\Delta t - \mu\Delta t$ となる。



したがって、 $P(X(t) = n)$ は次式を満たす：

$$\begin{aligned}
 P(X(t + \Delta t) = n) &= P(X(t + \Delta t) = n | X(t) = n - 1)P(X(t) = n - 1) \\
 &\quad + P(X(t + \Delta t) = n | X(t) = n)P(X(t) = n) \\
 &\quad + P(X(t + \Delta t) = n | X(t) = n + 1)P(X(t) = n + 1) \\
 &= \lambda\Delta t P(X(t) = n - 1) \\
 &\quad + (1 - \lambda\Delta t - \mu\Delta t)P(X(t) = n) \\
 &\quad + \mu\Delta t P(X(t) = n + 1), \\
 P(X(t + \Delta t) = 0) &= (1 - \lambda\Delta t)P(X(t) = 0) + \mu\Delta t P(X(t) = 1).
 \end{aligned}$$

$p_n(t) = P(X(t) = n)$ とおけば、

$$\begin{aligned}
 p'_n(t) &= \lambda p_{n-1}(t) - (\lambda + \mu)p_n(t) + \mu p_{n+1}(t), \quad n = 1, 2, \dots, \\
 p'_0(t) &= -\lambda p_0(t) + \mu p_1(t).
 \end{aligned}$$

この連立微分方程式を解析することで、待ち行列の性質がわかるのである。

ここでは $t \rightarrow \infty$ における均衡解（極限推移確率） p_n を求めてみよう。 $t \rightarrow \infty$ では $p_n(t) \rightarrow p_n$ は定数

になると考えるのである。したがって、左辺の微分を 0 として、

$$\begin{aligned}\lambda p_{n-1} - (\lambda + \mu)p_n + \mu p_{n+1} &= 0 \quad n = 1, 2, \dots, \\ -\lambda p_0 + \mu p_1 &= 0.\end{aligned}$$

これは容易に解けて、 $\lambda \neq \mu$ ならば、

$$p_n = A \left(\frac{\lambda}{\mu} \right)^n \quad (A \text{ は任意定数})$$

の形であることがわかる。また、 $\lambda = \mu$ のときは、 $p_n = A$ (定数) となる。 p_n が確率分布を表すためには、 $\sum_{n=0}^{\infty} p_n = 1$ であるから、これが可能なのは $\lambda < \mu$ のときに限り、

$$p_n = \left(1 - \frac{\lambda}{\mu} \right) \left(\frac{\lambda}{\mu} \right)^n, \quad n = 0, 1, 2, \dots$$

が得られる。これは、十分時間が経過したときのシステム内にいる客の人数を表す確率分布であり、幾何分布になっていることがわかる。その平均値は、容易に計算できて

$$\sum_{n=0}^{\infty} n p_n = \frac{\lambda/\mu}{1 - \lambda/\mu} = \frac{\lambda}{\mu - \lambda}$$

が得られる。

待ち行列理論では、到着率 λ とサービス率 μ の比

$$\rho = \frac{\lambda}{\mu}$$

をトラフィック密度という。 $\rho \geq 1$ ではサービスしきれない。 $\rho < 1$ であれば、平衡状態においてシステム内の客の人数は幾何分布

$$(1 - \rho)\rho^n, \quad n = 0, 1, 2, \dots,$$

に従う。これから、窓口が空いている確率は $1 - \rho$ であり、窓口がふさがっていてすぐにサービスが受けられない確率は ρ となる。この意味で ρ は利用率とも呼ばれる。サービス中の客も合わせたシステム中の客の平均人数は $\rho/(1 - \rho)$ で与えられる。

例 7.4.1 ATM で一人当たりの平均所要時間は 3 分であり、客は平均 5 分間隔でやってくるものとする。待ち行列でモデル化すると、

$$\lambda = \frac{1}{5}, \quad \mu = \frac{1}{3}, \quad \rho = \frac{3}{5}$$

となる。したがって、ATM を訪れたとき、利用客がおらずすぐに利用できる確率は $1 - \rho = \frac{2}{5}$ である。また、ATM を利用している客はいるが、ほかに待っている人のない確率は、

$$\frac{2}{5} \times \frac{3}{5} = \frac{6}{25}$$

である。よって、ATM を訪れたとき、すでに待っている人がいて行列につかねばならない状況が起こる確率は、

$$1 - \frac{2}{5} - \frac{6}{25} = \frac{9}{25} = 0.36.$$

平均して 3 回に 1 回は行列に並ぶことになる。

注意 7.4.2 待ち行列で現れた $X(t)$ は、より一般には出生死滅過程と呼ばれる計数過程の中で議論される。

第8章 ブラウン運動

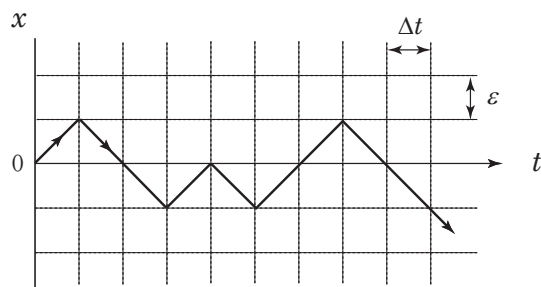
ブラウン運動は、離散時間の確率過程であるランダムウォークにおいて、時間幅を極限まで小さくすることで得られる連続時間の確率過程である。その名前の由来は、植物学者ブラウン (Robert Brown) が花粉から生じた微粒子の不規則な動きを顕微鏡下で観察した (1827) ことに因む。長い間原因が不明であったが、1905 年 (物理学の奇跡の年として知られる) にアインシュタイン (Albert Einstein) が、ブラウン運動は水分子の不規則な衝突によって引き起こされる現象であるとし、熱伝導方程式とブラウン運動との関係を明らかにした。こうして、ブラウン運動の最初の数学的定式化はアインシュタインによると広く認知されているが、それとは独立にポーランドの物理学者スモルコフスキ (Marian Smoluchowski, 1872–1917) もブラウン運動の数学的記述を与えている。スモルコフスキの論文発表は 1906 年であるが、実は、アインシュタインに先行する研究であるとも言われている (クラコフ大学に関連する展示がある)。このような物理学的研究をうけて、ブラウン運動あるいはより一般の確率過程の数学的研究が始まった。ブラウン運動の数学理論はウィナー (Norbert Wiener, 1894–1964) とレヴィ (Paul Lévy, 1886–1971) によるところが大きい。ブラウン運動はウィナー過程とも呼ばれる。1940 年代に入ると、伊藤清 (Kiyoshi Itô, 1915–2008) がブラウン運動に関する微積分や確率微分方程式の理論 (伊藤解析あるいは伊藤理論と呼ばれる) を創始した。この理論は、過去半世紀にわたり確率解析の名のもとで大発展し、ランダム現象の数理解析に飛躍的な進展をもたらして今日に至っているのである。伊藤の公式なしには金融工学は成り立たず、伊藤清はウォール・ストリートでもっとも有名な日本人となったことはその一端に過ぎない。

8.1 ランダムウォーク (再考)

直線上を運動するランダムウォーカーを考えよう。このランダムウォーカーは、時刻 $t = 0$ において原点 0 から出発し、時間が Δt 経過するごとに公平なコインを投げて、表 (H) が出たら正の方向に ϵ だけ移動し、裏 (T) が出たら負の方向に ϵ だけ移動するものとする。成功確率 $1/2$ のベルヌイ確率変数列 $\{Z_n\}$ を用いれば、時刻 $t = n\Delta t$ におけるランダムウォーカーの位置 X_n は、

$$X_n = \sum_{k=1}^n \epsilon Z_k$$

となる。



ここで, $P(Z_k = 1) = P(Z_k = -1) = \frac{1}{2}$ と $Z_1, Z_2, \dots$ の独立性を用いれば, X_n の平均値 (位置)・分散・共分散が求められる:

$$\mathbf{E}(X_n) = \epsilon \sum_{k=1}^n \mathbf{E}(Z_k) = 0, \quad (8.1)$$

$$\mathbf{V}(X_n) = \epsilon^2 n, \quad (8.2)$$

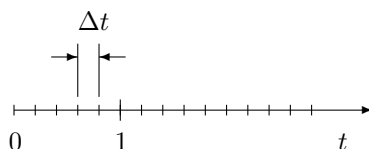
$$\text{Cov}(X_m, X_n) = \epsilon^2 \sum_{j=1}^m \sum_{k=1}^n \text{Cov}(Z_j, Z_k) = \min\{\epsilon^2 m, \epsilon^2 n\}. \quad (8.3)$$

問 8.1 (8.1)–(8.3) を確認せよ.

8.2 ブラウン運動

単位時間当たりのコイン投げの回数を n とすれば, 時刻 t におけるランダムウォーカーの位置は nt 回のコイン投げの積算として与えられる. このとき, Δt は $n\Delta t = 1$ を満たし, 時刻 t までのコイン投げの回数は, Δt のとり方による. そのことを明示して, 時刻 t におけるランダムウォーカーの位置を $X_{\Delta t}(t)$ と書こう. つまり,

$$X_{\Delta t}(t) = \sum_{k=1}^{nt} \epsilon Z_k. \quad (8.4)$$



興味の対象は, 関係 $n\Delta t = 1$ を保ったまま, $\Delta t \rightarrow 0$, $n \rightarrow \infty$ としたとき, (8.4) はどのような運動を表すかである. ここで, $X_{\Delta t}(t)$ の分散

$$\mathbf{V}(X_{\Delta t}(t)) = \epsilon^2 nt = \frac{\epsilon^2}{\Delta t} t$$

が極限において意味をもつためには,

$$\frac{\epsilon^2}{\Delta t} \rightarrow \alpha \text{ (定数)} \quad (8.5)$$

が成り立つ必要がある. 言い換えれば, コイン投げの時間間隔 Δt を狭くすると同時に, 1 回当たりのランダムウォーカーの移動距離 ϵ も小さくするのである. 時間と空間の軸をどのような比率で縮小すればよいかを示しているのが (8.5) である.

以下では, 簡単のため (8.5) で $\alpha = 1$ とする. コイン投げの時間間隔を Δt として得られるランダムウォークを $X_{\Delta t}(t)$ とおく. 時間間隔を小さくして ($\Delta t \rightarrow 0$) 得られる極限のランダムウォーク

$$B(t) = \lim X_{\Delta t}(t) = \lim \sum_{k=1}^{nt} \epsilon Z_k \quad (8.6)$$

に興味がある. ただし, 極限をとるときには,

$$n\Delta t = 1, \quad \frac{\epsilon^2}{\Delta t} = 1$$

をみたしたまま,

$$n \rightarrow \infty, \quad \Delta t \rightarrow 0, \quad \epsilon \rightarrow 0$$

とする. このとき, 平均・分散・共分散の極限は, (8.1)–(8.3) からすぐわかる:

$$\mathbf{E}(B(t)) = 0, \quad (8.7)$$

$$\mathbf{V}(B(t)) = \mathbf{E}(B(t)^2) = t, \quad (8.8)$$

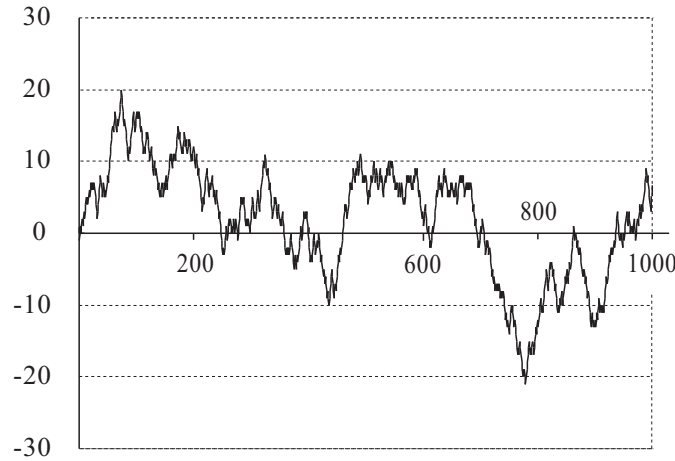
$$\text{Cov}(B(s), B(t)) = \min\{s, t\}. \quad (8.9)$$

また, (8.6) から

$$B(t) = \lim \sqrt{\epsilon^2 n t} \times \frac{1}{\sqrt{n t}} \sum_{k=1}^{n t} Z_k = \sqrt{t} \lim \frac{1}{\sqrt{n t}} \sum_{k=1}^{n t} Z_k$$

となる. したがって, ド・モアブル-ラプラスの定理 (より一般には中心極限定理) によって, $B(t)$ は平均 0, 分散 t の正規分布 $N(0, t)$ に従うことがわかる. このように構成された確率過程 $\{B(t)\}$ をブラウン運動またはウィナー過程と呼ぶ.

問 8.2 (8.7)–(8.9) を確かめよ.



ブラウン運動のシミュレーション

定 理 8.2.1 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

$$a_1 B_{t_1} + \cdots + a_n B_{t_n}, \quad a_i \in \mathbf{R}, \quad t_i \geq 0,$$

は正規分布に従う確率変数である.

- (2) (平均) $\mathbf{E}[B_t] = 0$
(3) (共分散) $\text{Cov}(B_s, B_t) = \mathbf{E}[B_s B_t] = \min\{s, t\}$
(4) (初期値) $B_0 = 0$
(5) (サンプルパスの連続性) $t \mapsto B_t$ は確率 1 で連続.
(6) (独立増分) $0 \leq t_1 < t_2 < \cdots < t_n$ のとき,

$$B_{t_1}, B_{t_2} - B_{t_1}, B_{t_3} - B_{t_2}, \dots, B_{t_n} - B_{t_{n-1}}$$

は独立である.

(2)–(6) はすでに述べた直感的構成から明らかであろう. 各 t ごとに B_t が正規分布 $N(0, t)$ に従う確率変数であることもすでに述べた. このことと (6) を組み合わせれば (1) もわかる (独立なガウス型確率変数の和もまたガウス型になるから).

しかしながら, ポアソン過程の構成の場合と同様に, ブラウン運動なる確率過程がきちんと存在しているかは, 上の構成的導入では明らかにされていないのである. ブラウン運動を定義すべき確率空間 $(\Omega, \mathcal{F}, P)$ は何でしょうか? 理論を一貫させるためには, まず, ガウス系の一般論を構築して, (1–3) を満たす $\{B_t; t \geq 0\}$ をきちんと構成する. 次に, それを詳しく調べて (4)–(6) を証明する.

そのような背景を従えて, (1)–(5) を満たす確率過程をブラウン運動 (ウィナー過程ともいう) と定義する. 本によってはサンプルパスの連続性 (5) を仮定しないが, (1)–(3) からサンプルパスの連続変形ができることがわかるから実質的には同じことになる.

定 理 8.2.2 ブラウン運動はマルコフ過程である.

連続時間かつ状態空間が連続な確率過程に対するマルコフ性の定義のためには, 条件付き確率についてやや高度な知識を要するので省略する.

8.3 確率微分方程式

$x = x(t)$ を未知関数とする常微分方程式

$$\frac{dx}{dt} = f(t, x), \quad x(0) = x_0 \text{ (初期条件)} \quad (8.10)$$

を考えよう. これは時間とともに変化する量 $x = x(t)$ の変化の仕方を規定している. たとえば, α を定数として,

$$\frac{dx}{dt} = \alpha x, \quad x(0) = x_0 \quad (8.11)$$

で与えられれば, 時刻 t における x の変化率が $x(t)$ に比例することを表す. 無制限に生物が増殖する場合を単純化すれば, (8.11) が想定されよう.

記号の方便であるが, (8.10) を

$$dx(t) = f(t, x)dt, \quad x(0) = x_0 \quad (8.12)$$

と記すこともある. これは時間が t から $t + \Delta t$ と経過したときに, $x(t)$ の変化量が

$$x(t + \Delta t) - x(t) = f(t, x(t))\Delta t + o(\Delta t) \quad (8.13)$$

のように与えられることを示す. (8.12) の両辺を積分することで, 微分方程式 (8.10) あるいは (8.12) は積分方程式

$$x(t) = x_0 + \int_0^t f(s, x(s))ds \quad (8.14)$$

と同値である.

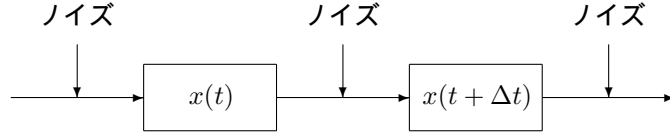
微分方程式の一般論によって, 2 変数関数 $f(t, x)$ に対する適当な条件 (たとえばリプシッツ条件) のもとで, 微分方程式 (8.10) は一意的な解をもつ. つまり, 微分方程式によって, $t = 0$ における初期状態から時間が経過するときに, $x = x(t)$ がどのように変化するかが一意的に定まるのである.



次に、各時刻で偶然の効果で値が揺らぐ場合を考えよう。つまり、(8.13) に「ノイズ」を加えることを考える。

$$x(t + \Delta t) - x(t) = f(t, x(t))\Delta t + (\text{ノイズ})$$

このとき、 $x(t)$ をノイズに対比して信号 (シグナル) と呼ぶことも多い。ノイズは、各時刻で独立に加わることを想定しよう。



ブラウン運動 $\{B_t\}$ は独立増分をもつので、

$$x(t + \Delta t) - x(t) = f(t, x(t))\Delta t + (B(t + \Delta t) - B(t))$$

とすると、各時刻で独立なノイズを加えたことになる。少し一般化して、ノイズの大きさは、信号の大きさと時刻によって変化してもよいことにすれば、

$$x(t + \Delta t) - x(t) = f(t, x(t))\Delta t + g(t, x(t))(B(t + \Delta t) - B(t))$$

を考えることになる。 $\{B(t)\}$ は確率過程だから $\{x(t)\}$ も確率過程になる。習慣に合わせて、今後は、 $x(t)$ を大文字 $X(t)$ で記すことにしよう。 $\Delta t \rightarrow 0$ とすれば、

$$dX(t) = f(t, X(t))dt + g(t, X(t))dB_t, \quad X(0) = x_0 \text{ (初期条件)} \quad (8.15)$$

という表式が想定される。積分すれば、

$$X(t) = x_0 + \int_0^t f(s, X(s))ds + \int_0^t g(s, X(s))dB_s \quad (8.16)$$

という表式が想定される。

(8.15) は確率過程 $\{X(t)\}$ に対する方程式である。これを確率微分方程式という。すでに説明してきた通り、各時点でブラウン運動でモデル化されるノイズが独立に加わってゆくことを直感的に示しており、有用な表式であるといえる。しかしながら、この程度の形式的計算では、 dB_t の意味は全く不明である。 dB_t に数学的意味を与え、確率微分方程式論を構築したのが伊藤清 (1915–2008) である。その理論体系を伊藤解析と称する。

8.4 確率積分と伊藤の公式

伊藤解析においては、微分形の方程式 (8.15) は積分形の方程式 (8.16) を簡略化した表現であるとし、積分方程式 (8.16) をまず定式化する。右辺の第 2 項はふつうの積分であるが、第 3 項は確率過程を dB_t に関して積分するものである。これを確率積分またはより明確に伊藤型確率積分と称する (今日では、本来の伊藤積分がさまざまに拡張されて論じられているから)。

一般に、ブラウン運動 $\{B_t\}$ に適合した (適合の意味を説明するのは、目下の知識では困難なので省略) 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 (8.15) はブラウン運動 $\{B_t\}$ に適合した確率過程 $\{X(t)\}$ で (8.16) を満たすものを求めよという問題なのである。

さて、確率過程 $\{X(t)\}$ が

$$X(t) = x_0 + \int_0^t F(s)ds + \int_0^t G(s)dB_s$$

の形をもつとき、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、

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

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 8.4.1 (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、

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

したがって、

$$\begin{aligned} F_1 &= \frac{\partial f}{\partial t}(t, X(t)) + \frac{\partial f}{\partial x}(t, X(t)) \cdot F(t) + \frac{1}{2} \frac{\partial^2 f}{\partial x^2}(t, X(t)) \cdot G(t)^2, \\ F_2 &= \frac{\partial f}{\partial x}(t, X(t)) \cdot G(t) \end{aligned}$$

が得られる。なお、伊藤の公式の本質は、次の積公式にある：

$$(dt)(dt) = (dB)(dt) = (dt)(dB) = 0, \quad (dB)(dB) = dt.$$

問 8.3 上の F_1, F_2 を確かめよ。

8.5 簡単な確率微分方程式

8.5.1 ランジュヴァン方程式

$\gamma > 0, \sigma > 0$ を定数とする。確率微分方程式

$$dX = -\gamma Xdt + \sigma dB, \quad X_0 = x_0 \tag{8.17}$$

をランジュヴァン方程式 (Langevin equation) という。ノイズ項がない微分方程式

$$dx = -\gamma xdt, \quad x(0) = x_0,$$

は摩擦係数 γ の摩擦力を受けながら運動する質点の速度 (初速度を x_0 としている) を表す. したがって, ランジュヴァン方程式は, そのような質点が外界から揺らぎ (揺動力) を受けながら運動するときの速度を記述している.

では, (8.17) を伊藤の公式を応用して解いてみよう. まず,

$$Y(t) = e^{\gamma t} X(t)$$

とおいて微分してみよう. ($f(t, x) = e^{\gamma t} x$ である.)

$$\begin{aligned} dY &= \gamma e^{\gamma t} X dt + e^{\gamma t} dX \\ &= \gamma e^{\gamma t} X dt + e^{\gamma t} (-\gamma X dt + \sigma dB) = e^{\gamma t} \sigma dB \end{aligned}$$

積分して,

$$e^{\gamma t} X(t) = x_0 + \int_0^t e^{\gamma s} \sigma dB_s$$

したがって,

$$X(t) = x_0 e^{-\gamma t} + \sigma \int_0^t e^{-\gamma(t-s)} dB_s$$

が得られる. ノイズがなければ ($\sigma = 0$) ならば,

$$x(t) = x_0 e^{-\gamma t}$$

である. よく知られているように, 摩擦力を受けながら運動する質点の速度は指数的に減衰して 0 になることに一致する. 一方, 外界からランダムな揺動力を受けるときは, $t \rightarrow \infty$ でも静止することはない. 実際,

$$\lim_{t \rightarrow \infty} \mathbf{E}[X_t] = 0, \quad \lim_{t \rightarrow \infty} \mathbf{E}[X_t^2] = \frac{\sigma^2}{2\gamma}$$

がわかるので, 長時間経過したときの $X(t)$ は正規分布 $N\left(0, \frac{\sigma^2}{2\gamma}\right)$ に従う.

8.5.2 ブラック・ショールズ・モデル

$\alpha \in \mathbf{R}, \sigma > 0$ を定数とする. 確率微分方程式

$$dX = \alpha X dt + \sigma X dB, \quad X_0 = x_0 \quad (8.18)$$

を考えよう. ノイズがなければ (つまり, $\sigma = 0$), 常微分方程式

$$dx = \alpha x dt, \quad x(0) = x_0$$

に帰着する. これは, $x(t)$ が成長率 α で成長することを意味し, $\alpha > 0$ なら指数的增长, $\alpha < 0$ なら指数的減少であることは既知であろう. α を (連続) 利率と考えれば, $x(t)$ は銀行に預金を預けっぱなしにしたときの預金残高 (連続複利による運用) となる. さて, (8.18) は,

$$dX = (\alpha dt + \sigma dB)X,$$

と書き換えられるから, 成長率 α が時々刻々, 外界からノイズの影響を受けて変化する場合は記述している.

伊藤の公式を応用して, (8.18) を解いてみよう. まず,

$$X(t) = Y(t) e^{\sigma B_t}$$

と想定しよう. 微分して,

$$dX = dY e^{\sigma B_t} + Y d(e^{\sigma B_t}) = dY e^{\sigma B_t} + Y \left(\sigma e^{\sigma B_t} dB + \frac{\sigma^2}{2} e^{\sigma B_t} dt \right).$$

(8.18) に代入すると,

$$dY e^{\sigma B_t} + Y \left(\sigma e^{\sigma B_t} dB + \frac{\sigma^2}{2} e^{\sigma B_t} dt \right) = \alpha Y e^{\sigma B_t} dt + \sigma Y e^{\sigma B_t} dB.$$

整理して,

$$dY = \left(\alpha - \frac{\sigma^2}{2} \right) Y dt$$

これは常微分方程式であり, 一般解

$$Y = K \exp \left\{ \left(\alpha - \frac{\sigma^2}{2} \right) t \right\}, \quad K: \text{任意定数},$$

が得られる. したがって, (8.18) の解は, 初期条件も考慮して

$$X(t) = x_0 \exp \left\{ \left(\alpha - \frac{\sigma^2}{2} \right) t + \sigma B_t \right\} \quad (8.19)$$

となる. 一般に, a, b を定数として, $\exp(at + bB_t)$ の形の確率過程を幾何ブラウン運動という.

幾何ブラウン運動 (8.19) は確率微分方程式 (8.18) の解として与えられた. この $\{X(t)\}$ は株価の変動のモデルとしても重要である ($X(t)$: 株価, α : 安全連続利子率, σ : 株価変動率). そのような文脈では, ブラック-ショールズ・モデル (Black-Scholes model) という. 有名なブラック-ショールズのオプション価格 (1973 年ノーベル経済学賞) は, このモデルから導出される.

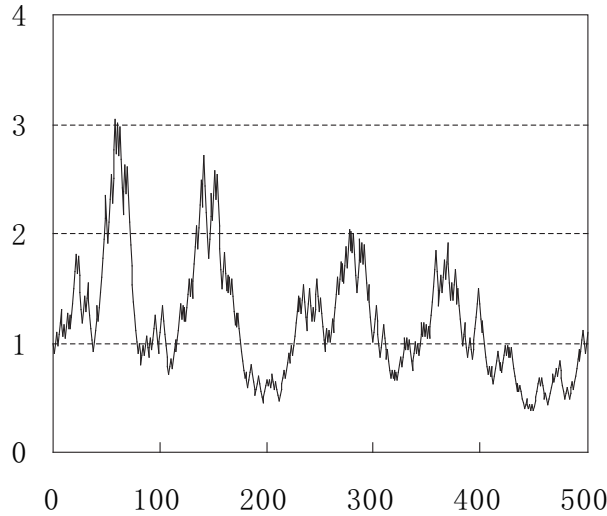


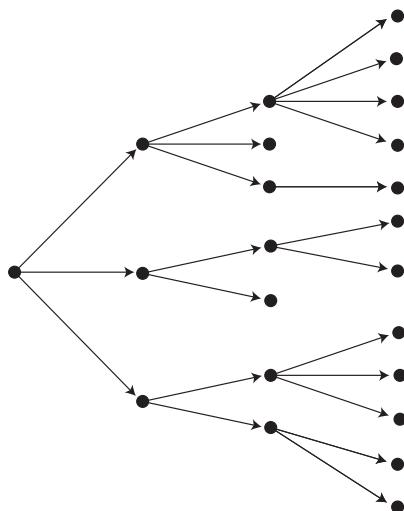
図 8.1: 株価変動 (幾何ブラウン運動) のシミュレーション

伊藤解析の参考書は多い. 入門的なところで,

1. B. エクセンドール (谷口説男訳): 確率微分方程式, シュプリンガー東京, 1999.
2. 小川重義: 確率解析と伊藤過程, 朝倉書店, 2005.
3. Z.Brzeźniak and T. Zastawniak: Basic Stochastic Processes, Springer, 1999.

第9章 ゴルトン・ワトソン分枝過程

家系を観察すれば、各世代の個体は何個かの個体を生んで死んでゆくことになる。この状況を単純化して得られる確率モデルがゴルトン・ワトソン分枝過程 (Galton-Watson branching process) であり、家系の存続について確率的に答えることができる。生物学的な家系に限らず、たとえば、遺伝子の存続 (集団遺伝学) や粒子の生成消滅 (物理学) など家系のようにふるまうものに適用できる。さらに、感染症モデルや拡散現象などのより複雑な確率モデルの解析で、ゴルトン・ワトソン分枝過程が基本的な道具として多用される。ゴルトン・ワトソン分枝過程という名前の由来は、F. Galton が 1873 年に提示し、翌年に Galton と Watson が共著論文を発表して、その主要な性質を明らかにしたことに因む。しかしながら、Irénée-Jules Bienaymé (1796–1878) が、すでに 1845 年に、同様のモデルについて臨界現象を見出していたので、著者によっては Bienaymé の名を冠して BGW 過程と呼んでいる。



9.1 定義

各 $n = 0, 1, 2, \dots$ に対して、第 n 世代の個体数を X_n で表わす。この個体数の変化を規定することでマルコフ連鎖を構成する。各世代のそれぞれの個体は、何個かの個体を生んでそれ自身は死ぬものとする。このとき、各個体が生み出す個体数は個体および世代によらず同じ確率分布に従い、しかも、個体ごとに独立であるものとする。ここで、ある世代のある個体が生み出す個体数を Y とおき、その確率分布を

$$P(Y = k) = p_k, \quad k = 0, 1, 2, \dots$$

とおくことにする。 Y と同じ分布をもつ、独立な確率変数列 $Y_1, Y_2, \dots$ を準備する。そうすれば、第 n 世代の個体数が $i \geq 1$ であるとき、次の第 $n+1$ 世代の個体数が j に変化する確率 (推移確率) は

$$p(i, j) = P(X_{n+1} = j | X_n = i) = P\left(\sum_{k=1}^i Y_k = j\right), \quad i \geq 1, \quad j \geq 0,$$

で与えられる. ある世代で個体数が 0 になってしまうと, もはや新しい個体は生まれないから,

$$p(0, j) = \begin{cases} 0, & j \geq 1, \\ 1, & j = 0 \end{cases}$$

とおく. つまり, 0 は吸収状態である. 上のように定義された推移確率 $p(i, j)$ をもつ $\{0, 1, 2, \dots\}$ 上のマルコフ連鎖 $\{X_n\}$ をゴルトン・ワトソン分枝過程という.

構成の仕方からわかるように, ゴルトン・ワトソン分枝過程 $\{X_n\}$ の本質は Y の確率分布 $\{p_k; k = 0, 1, 2, \dots\}$ によって決定される. これを, ゴルトン・ワトソン分枝過程 $\{X_n\}$ の子孫分布という. 多くの場合, 初期条件を $X_0 = 1$ とおいて考察すれば十分である (家系は 1 個の個体から始まる).

問 9.1 $p_0 + p_1 = 1$ のとき, ゴルトン・ワトソン分枝過程 ($X_0 = 1$ とする) はどのようなものか? 特に,

$$q_1 = P(X_1 = 0), \quad q_2 = P(X_1 \neq 0, X_2 = 0), \quad \dots, q_n = P(X_1 \neq 0, \dots, X_{n-1} \neq 0, X_n = 0), \quad \dots$$

を計算し, 家系が消滅する確率

$$P = \left(\bigcup_{n=1}^{\infty} \{X_n = 0\} \right) = P(\text{ある } n \geq 1 \text{ で } X_n = 0 \text{ となる})$$

を求めよ.

注意 9.1.1 ゴルトン・ワトソン分枝過程 $\{X_n\}$ において興味深いのは, その子孫分布が

$$p_0 + p_1 < 1, \quad p_2 < 1, \quad \dots, \quad p_k < 1, \quad \dots \quad (9.1)$$

のときである. 以下では, このことを仮定する.

9.2 母関数

子孫分布 $\{p_k; k = 0, 1, 2, \dots\}$ をもつゴルトン・ワトソン分枝過程 $\{X_n\}$ を考える. また, $X_0 = 1$ とする. 子孫分布の母関数を

$$f(s) = \sum_{k=0}^{\infty} p_k s^k \quad (9.2)$$

とおく. 右辺の級数は $|s| \leq 1$ において絶対収束する. また,

$$f_0(s) = s, \quad f_1(s) = f(s), \quad f_n(s) = f(f_{n-1}(s)),$$

と定義する.

補題 9.2.1 ゴルトン・ワトソン分枝過程の推移確率 $p(i, j)$ について

$$\sum_{j=0}^{\infty} p(i, j) s^j = [f(s)]^i, \quad i = 1, 2, \dots \quad (9.3)$$

証明 定義によって,

$$p(i, j) = P(Y_1 + \dots + Y_i = j) = \sum_{\substack{k_1 + \dots + k_i = j \\ k_1 \geq 0, \dots, k_i \geq 0}} P(Y_1 = k_1, \dots, Y_i = k_i)$$

であるが, $Y_1, \dots, Y_i$ は独立であるから,

$$p(i, j) = \sum_{\substack{k_1 + \dots + k_i = j \\ k_1 \geq 0, \dots, k_i \geq 0}} P(Y_1 = k_1) \cdots P(Y_i = k_i) = \sum_{\substack{k_1 + \dots + k_i = j \\ k_1 \geq 0, \dots, k_i \geq 0}} p_{k_1} \cdots p_{k_i}$$

となる. したがって,

$$\begin{aligned} \sum_{j=0}^{\infty} p(i, j) s^j &= \sum_{j=0}^{\infty} \sum_{\substack{k_1 + \dots + k_i = j \\ k_1 \geq 0, \dots, k_i \geq 0}} p_{k_1} \cdots p_{k_i} s^j \\ &= \sum_{k_1=0}^{\infty} p_{k_1} s^{k_1} \cdots \sum_{k_i=0}^{\infty} p_{k_i} s^{k_i} \\ &= [f(s)]^i \end{aligned}$$

よって示された. ■

補題 9.2.2 ゴルトン・ワトソン分枝過程の n ステップ推移確率 $p_n(i, j)$ について

$$\sum_{j=0}^{\infty} p_n(i, j) s^j = [f_n(s)]^i, \quad i = 1, 2, \dots \quad (9.4)$$

証明 定義によって, $p_1(i, j) = p(i, j)$, $f_1(s) = f(s)$ であるから, $n = 1$ のとき示すべき式は,

$$\sum_{j=0}^{\infty} p(i, j) s^j = [f(s)]^i, \quad i = 1, 2, \dots \quad (9.5)$$

である. これは, 補題 9.2.1 によって示した. $n \geq 1$ として, (9.4) が n まで正しいと仮定する. 推移確率に関するチャップマン・コルモゴロフの方程式を用いて,

$$\sum_{j=0}^{\infty} p_{n+1}(i, j) s^j = \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} p(i, k) p_n(k, j) s^j$$

を得る. 帰納法の仮定から,

$$\sum_{j=0}^{\infty} p_n(k, j) s^j = [f_n(s)]^k$$

が成り立つから,

$$\sum_{j=0}^{\infty} p_{n+1}(i, j) s^j = \sum_{k=0}^{\infty} p(i, k) [f_n(s)]^k$$

右辺は, (9.5) で s を $f_n(s)$ でおきなおした式であるから,

$$\sum_{j=0}^{\infty} p_{n+1}(i, j) s^j = [f(f_n(s))]^i = [f_{n+1}(s)]^i$$

となる. これが示したかったことである. ■

$X_0 = 1$ を仮定しているので,

$$P(X_n = j) = P(X_n = j | X_0 = 1) = p_n(1, j),$$

特に,

$$P(X_1 = j) = P(X_1 = j | X_0 = 1) = p_1(1, j) = p(1, j) = p_j$$

であることに注意しよう.

定 理 9.2.3 子孫分布が有限の平均値

$$m = \sum_{k=0}^{\infty} k p_k < \infty$$

をもてば,

$$\mathbf{E}[X_n] = m^n.$$

証 明 まず, (9.2) を微分して,

$$f'(s) = \sum_{k=0}^{\infty} k p_k s^{k-1}, \quad |s| < 1, \quad (9.6)$$

を得る. $s \rightarrow 1-0$ とするとき, 仮定によって右辺は m に収束する. よって,

$$\lim_{s \rightarrow 1-0} f'(s) = m.$$

一方, (9.4) で $i = 1$ とおけば,

$$\sum_{j=0}^{\infty} p_n(1, j) s^j = f_n(s) = f_{n-1}(f(s)). \quad (9.7)$$

微分して,

$$f'_n(s) = \sum_{j=0}^{\infty} j p_n(1, j) s^{j-1} = f'_{n-1}(f(s)) f'(s). \quad (9.8)$$

$s \rightarrow 1-0$ として,

$$\lim_{s \rightarrow 1-0} f'_n(s) = \sum_{j=0}^{\infty} j p_n(1, j) = \lim_{s \rightarrow 1-0} f'_{n-1}(f(s)) \lim_{s \rightarrow 1-0} f'(s) = m \lim_{s \rightarrow 1-0} f'_{n-1}(s).$$

したがって,

$$\lim_{s \rightarrow 1-0} f'_n(s) = m^n$$

がわかる. これは,

$$\mathbf{E}(X_n) = \sum_{j=0}^{\infty} j P(X_n = j) = \sum_{j=0}^{\infty} j p_n(1, j) = m^n$$

を意味する. ■

こうして, 第 n 世代の個体数の平均値 $\mathbf{E}(X_n)$ は n とともに, $m < 1$ であれば減少して 0 に収束し, $m > 1$ であれば増加して ∞ に発散し, また $m = 1$ ならば平均値は一定ということになる. このことから, $m = 1$ を境目として家系の消滅が起こるのではないかと示唆される.

レポート問題 30 子孫分布の分散が有限であると仮定して $\mathbf{V}[Y] = \sigma^2$ とおく. 定理 9.2.3 にならって,

$$\mathbf{V}[X_n] = \begin{cases} \frac{\sigma^2 m^{n-1} (m^n - 1)}{m - 1}, & m \neq 1, \\ n \sigma^2, & m = 1 \text{ のとき,} \end{cases}$$

を示せ.

9.3 消滅確率

ゴルトン・ワトソン分枝過程 $\{X_n\}$ が第 n 世代の個体数を表すとすれば, 事象 $\{X_n = 0\}$ はこの家系が第 n 世代に至るまでに消滅していることを表す. したがって, この家系が消滅する確率 q は,

$$q = P\left(\bigcup_{n=1}^{\infty} \{X_n = 0\}\right)$$

で与えられる. 右辺の和事象はたがいに排反ではなく,

$$\{X_1 = 0\} \subset \{X_2 = 0\} \subset \cdots \subset \{X_n = 0\} \subset \cdots$$

となっている. したがって,

$$q = \lim_{n \rightarrow \infty} P(X_n = 0) \quad (9.9)$$

が成り立つ. $q = 1$ ならば, この家系は (有限時間で) 必ず消滅し, $q < 1$ であれば, この家系が消滅せず永遠に続く確率が $1 - q > 0$ ということになる. したがって, $q = 1$ かどうかは興味のある問題となる.

補題 9.3.1 子孫分布の母関数を $f(s)$, その n 回合成関数を $f_n(s) = f(f_{n-1}(s))$ とするとき,

$$q = \lim_{n \rightarrow \infty} f_n(0)$$

が成り立つ. したがって, q は,

$$q = f(q) \quad (9.10)$$

を満たす.

証明 補題 9.2.2 から,

$$f_n(s) = \sum_{j=0}^{\infty} p_n(1, j) s^j$$

よって,

$$f_n(0) = p_n(1, 0) = P(X_n = 0 | X_0 = 1) = P(X_n = 0).$$

最後の等号は, 仮定 $X_0 = 1$ による. (9.9) と合わせて, 結論を得る. 後半は, $f(s)$ は $[0, 1]$ 上の連続関数であることからわかる. ■

補題 9.3.2 子孫分布の母関数 $f(t)$ は次の性質をもつ. (ただし, 子孫分布は (9.1) を満たすものと仮定されている.)

- (1) $f(s)$ は増加関数である. つまり, $0 \leq s_1 \leq s_2 \leq 1$ ならば, $f(s_1) \leq f(s_2)$.
- (2) $f(s)$ は狭義凸関数, つまり, $0 \leq s_1 < s_2 \leq 1$, $0 < \theta < 1$ ならば,

$$f(\theta s_1 + (1 - \theta)s_2) < \theta f(s_1) + (1 - \theta)f(s_2).$$

証明 (1) はべき級数 $f(s)$ の係数がすべて ≥ 0 であるから明らか. (2) は $f''(s) > 0$ からわかる. ■

補題 9.3.3 (1) $m \leq 1$ のとき, $0 \leq s < 1$ で $f(s) > s$.

- (2) $m > 1$ のとき, $0 \leq s < 1$ の中に $f(s) = s$ をみたす s がただ 1 個存在する.

補題 9.3.4 $f_1(0) \leq f_2(0) \leq \cdots \rightarrow q$.

定理 9.3.5 ゴルトン・ワトソン分枝過程が消滅する確率 q は,

$$s = f(s), \quad 0 \leq s \leq 1,$$

を満たす s のうちで最小のものに一致する. さらに, $m \leq 1$ なら $q = 1$ であり, $m > 1$ なら $q < 1$ である.

ゴルトン・ワトソン分枝過程は, $m < 1$, $m = 1$, $m > 1$ にしたがって, 劣臨界的 (subcritical), 臨界的 (critical), 優臨界的 (supercritical) と呼ばれる. 家系の存続を決めるパラメータは子孫分布の平均値 m だけであり, $m = 1$ を境目に存続が劇的に変化しているといえる. これを統計物理学の言葉を借用して相転移現象という.

レポート問題 31 p を $0 < p < 1$ を満たす定数とする. 子孫分布が,

$$p_0 = 1 - p, \quad p_1 = 0, \quad p_2 = p, \quad p_3 = p_4 = \cdots = 0$$

で与えられているゴルトン・ワトソン分枝過程について考察せよ. たとえば, 確率

$$P(X_1 \neq 0, \dots, X_{n-1} \neq 0, X_n = 0), \quad P(X_n = 0), \quad \lim_{n \rightarrow \infty} P(X_n = 0)$$

などの計算や, 計算機シミュレーション, その他.

レポート問題 32 b, p を $b > 0$, $0 < p < 1$, $b + p < 1$ を満たす定数とする. 子孫分布が

$$p_k = bp^{k-1}, \quad k = 1, 2, \dots, \\ p_0 = 1 - \sum_{k=1}^{\infty} p_k$$

で与えられているとき, $f(s)$ を計算せよ. また, $m = 1$ のとき, $f_n(s)$ を計算せよ.

レポート問題 33 異なる家系が 100 ある孤島を考える. 各家系の消滅確率が $q = 9/10$ であるとき, この島の全人口は増加するだろうか減少するだろうか? 考察せよ.

ゴルトン・ワトソン分枝過程について詳しく知りたければ,

1. R. B. シナジ (今野紀雄・林俊一訳): マルコフ連鎖から格子確率モデルへ, シュプリンガー東京, 2001.
2. K. B. Athreya and P. E. Ney: Branching Processes, Dover 2004 (original version, Springer 1972)