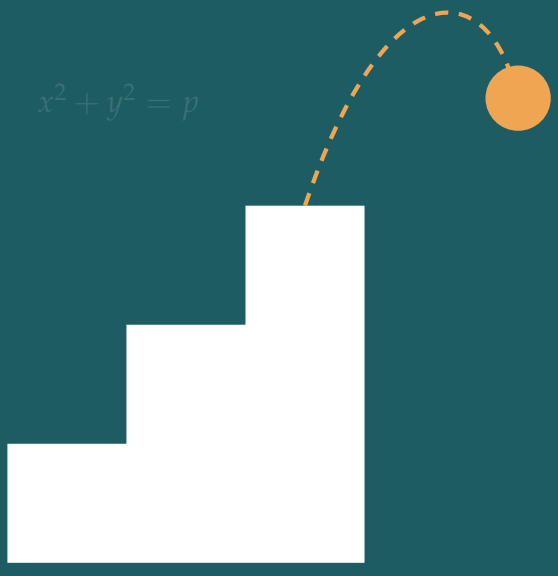
STEP 対策 · 大学の数学

Beyond STEP

第 12 章

確率



KEITH McNULTY

Beyond STEP

大学数学への入り口としての
STEP 過去問

第 12 章

確率

Keith McNulty

Beyond STEP：大学数学への入り口としての STEP 過去問

第 12 章：確率

これは *Beyond STEP* 全 13 章のうちの第 12 章である。本書全体、各章単独の PDF、各節の印刷用ワークシート、そしてすべての STEP の問題の詳しい解答は、`beyondstep.keithmcnulty.org` で無料で閲覧・ダウンロードできる。

2026 年 11 月、オンラインにて初版発行。

本 PDF のうち著者自身の著作物（本文、解答、図、演習問題とその答えを含む）はすべて、クリエイティブ・コモンズ CC0 1.0 全世界（Creative Commons CC0 1.0 Universal）ライセンスのもとでパブリックドメインに提供されている。許可を求めることなく、いかなる目的でも複製、改変、共有してよい。詳しくは `creativecommons.org/publicdomain/zero/1.0` を参照のこと。

ここで扱う STEP の問題の著作権は Cambridge University Press & Assessment に帰属する。問題そのものは転載しておらず、各問題には公式の出典である STEP Support Programme の問題データベース、または OCR の過去問へのリンクを付している。

誤りの指摘は、`github.com/keithmcnulty/step_book_mcnulty` への issue として歓迎する。

章

12

確率

不確かさについても厳密に推論することはできる。ところが直観のほうは、そうはいかないのである。

この章の内容

12.1	正確な条件付け	STEP I 2007
12.2	基準率の誤謬	STEP II 1998
12.3	「ランダム」とは何を意味するか？	STEP III 2022
12.4	二つずつなら独立、全体では独立でない	STEP III 2021
12.5	最後の乗客の座席	STEP II 2021
12.6	もう 1 枚の硬貨	STEP I 2016
12.7	勝ち抜き戦で誰と誰が当たるか	STEP II 試験見本
12.8	ペニーのゲーム	STEP III 2001
12.9	偏った硬貨から公平な判定を	STEP I 2019
12.10	ひもの端を結ぶ	STEP III 2008
12.11	切手、カップ、そして e	STEP I 1996
12.12	いつ探すのをやめるか	STEP I 2001
12.13	三つ巴の決闘	STEP II 1995
12.14	混合戦略	STEP I 1989
12.15	絶滅か爆発か	STEP II 2003
12.16	マルコフ連鎖を順向きにも逆向きにも	STEP III 1996
12.17	バナッハのマッチ箱	STEP II 1992
12.18	無記憶性	STEP III 2021
12.19	平面上のランダムな点	STEP II 2012
12.20	ビュフォンの針	STEP I 1993
12.21	棒を折る	STEP III 1990
12.22	平均をもたない分布	STEP III 1995
12.23	確率から導くスターリングの公式	STEP II 2016
12.24	中心極限定理	STEP III 2008
12.25	多数の中の最大値	STEP III 2025
A	演習問題の答え	p. 161

偶然について推論する

> 確率論とは、計算に還元された常識にほかならない。
>
> — ピエール＝シモン・ラプラス（Pierre-Simon Laplace）、*Théorie analytique des probabilités*（『確率の解析的理論』、1812 年）

ラプラスのこの有名な言葉は、半分だけ正しい。確率（probability）とは常識を精密にしたものである。しかし本章で繰り返し見るように、常識はそれだけに任せておくと誤る。人はこう考えがちである。まれな病気の検査で陽性が出たら、おそらくその病気にかかっている。飛行機に最初に乗り込んだ乗客が違う席に座ってしまったら、最後の乗客が自分の席に座れる見込みはほとんどない。コイン投げのある並びが、同じ長さの別の並びに勝ちやすいということはありえない。多数の測定値の平均はいつでも一定の値に落ち着く。これらはすべて誤りであり、どの場合も、短い計算によって真実が明らかになる。さらによいことに、その理由まで明らかになるのである。

確率論は 17 世紀、ブレーズ・パスカル（Blaise Pascal）とピエール・ド・フェルマー（Pierre de Fermat）の間で交わされた書簡の中で、賭け事に関する問いから始まった。それが厳密な数学の一分野となったのは、ようやく 1933 年、アンドレイ・コルモゴロフ（Andrey Kolmogorov）がその公理（axiom）を定めたときである。その間にも、そしてそれ以後も、確率論は数学の中で最も広く使われる分野の一つへと成長してきた。

確率はなぜ重要なのか

不確かさを扱う分野はどれも確率を用いる。医学（診断検査や臨床試験）、法学（証拠の重み）、金融（リスクと保険）、工学（信頼性や洪水）、計算機科学（乱択アルゴリズム（randomised algorithm）、暗号、機械学習）、物理学（量子力学や統計力学）、生物学（遺伝学や感染症の流行）などである。現代の人工知能は、その大部分が応用確率論であるといってよい。言語モデルは次に来る単語の確率を予測しているのである。そして、リスク、偶然、証拠について正しく推論する能力は、教養ある人なら誰にとっても最も価値ある技能の一つである。

大学で出会うこと

大学 1 年の確率論では、公理、条件付き確率（conditional probability）とベイズの定理（Bayes′ theorem）、独立性（independence）、離散型および連続型の確率変数（random

variable)、期待値（expectation）と分散（variance)、母関数（generating function)、そして二つの偉大な極限定理である大数の法則（law of large numbers）と中心極限定理（central limit theorem）を学ぶ。その後の講義では、確率過程（random process）（マルコフ連鎖（Markov chain)、ポアソン過程（Poisson process)、ランダムウォーク（random walk)、分枝過程（branching process))、測度論的確率論（measure-theoretic probability）（極限や連続分布を確固とした基礎の上に置くもの)、マルチンゲール（martingale）を学び、確率論を統計学、解析学、組合せ論、物理学と結びつけていく。

本章を貫く問い

三つの大きな問い

1. **問われていることは正確には何か？**　ランダムな試行とは何か、何を条件としているのか、そして「無作為に」とはどういう意味か。
2. **うまい切り口は何か？**　対称性（symmetry)、指示変数（indicator)、条件付け、あるいはうまく選んだ状態を用いることで、長い計算を省けるのはどのようなときか。
3. **長い目で見ると何が起こるか？**　確率過程は時間とともにどのように振る舞うのか。また、多数のランダムな量の和、平均、最大値はどのような姿をしているのか。

本章は五つの部分からなる。**明晰に考える**（Section 12.1〜Section 12.4）では落とし穴を扱う。条件付け、基準率（base rate)、「ランダム」の意味、独立性の意味である。**対称性と巧妙な議論**（Section 12.5〜Section 12.9）には、美しく短い解法をもつ問題を集めた。**期待値、ゲーム、停止**（Section 12.10〜Section 12.14）では平均と意思決定に目を向ける。いつ止めるべきか、誰を撃つべきか、どのようにランダム化すべきか、である。**確率過程**（Section 12.15〜Section 12.19）では、時間と空間の中でのランダムさを追いかける。そして**極限と連続型の確率**（Section 12.20〜Section 12.25）は、幾何的確率（geometric probability）と偉大な極限定理で締めくくる。

本章で扱う考え方

§	主題	考え方
12.1	条件付け	「特定の一つ」と「少なくとも一つ」の違い
12.2	ベイズの定理	基準率と自然頻度（natural frequency）
12.3	ベルトラン（Bertrand）の逆説	「ランダム」には定義が必要である
12.4	独立性	対ごとに独立（pairwise）であっても互いに独立（mutual）とは限らない
12.5	飛行機の座席	二つの座席の間の対称性
12.6	もう 1 枚のコイン	互いに補い合う対称な事象
12.7	トーナメント	試合数と組の数を数える
12.8	Penney のゲーム	第一歩解析（first-step analysis）、非推移性（non-transitivity）
12.9	公平なコイン	偏りから公平さを取り出す
12.10	ランダムな輪	無記憶な指示変数の和
12.11	カップと e	階乗と指数関数の級数
12.12	最適停止（optimal stopping）	$\frac{1}{e}$ の法則
12.13	三者決闘（truel）	戦略上の弱さ
12.14	ゲーム理論（game theory）	ランダム化された戦略、ミニマックス（minimax）
12.15	分枝過程	不動点（fixed point）としての絶滅
12.16	マルコフ連鎖	定常分布（stationary distribution）、時間反転
12.17	バナッハ（Banach）のマッチ箱	負の二項分布（negative binomial distribution）
12.18	指数分布（exponential distribution）	無記憶性（memorylessness）と独立性
12.19	空間のランダムさ	最近接点、レイリー（Rayleigh）分布
12.20	ビュフォン（Buffon）の針	幾何的確率と π
12.21	ランダムな三角形	面積としての確率
12.22	コーシー（Cauchy）分布	平均をもたない分布
12.23	スターリング（Stirling）の公式	正規曲線から得られる階乗
12.24	中心極限定理	積率母関数（moment generating function）
12.25	極値（extreme value）	ガンベル（Gumbel）分布

Section 12.24 の一節では積率母関数を用いる。これは 2019 年に STEP の出題範囲から外されたが、この節では必要なことを説明している。

本章の読み方

確率の問題がわかりにくく思えるときは、小さな場合について標本空間（sample space）を書き出し、数えてみるとよい。答えが怪しいほど簡単に出てきたときは、そ

れを説明する対称性を探すとよい。そして直観と計算が食い違うときは、計算を信じ、そのうえでなぜ直観が誤っていたのかを突き止めることである。学びが生まれるのはまさにそこなのである。

12.1 正確な条件付け

問題

STEP I 2007、問 13

先に進む前に、まず問題を読んでおこう。問題は STEP Support Programme の問題データベースにある：

`step.maths.org/questions/07-s1-q13`

導入

「〜という条件のもとで」は、確率論で最も危険な言い回しである。この問題の最後の四つの小問は、ほとんど同じことを言っているように聞こえる。番号 3 の円盤が最初に取り出される。番号 3 の円盤が取り出される。番号つきの円盤が最初に取り出される。番号つきの円盤が取り出される。このうち三つは答えが $\frac{1}{2}$ になるが、残る一つはそうならない。

この違いは、有名な「男の子か女の子か」のパズル（ある家庭に子どもが二人いて、少なくとも一人は男の子である。二人とも男の子である確率はいくらか）の背後にあるものとまったく同じである。このパズルは何世代もの学生を、そして少なからぬ数のプロの数学者をもつまずかせてきた。ここから学ぶべきことは、条件とは「どの結果がまだ起こりうるか」についての精密な言明であり、情報そのものと同じくらい、その情報がどのようにして得られたかが重要だということである。

取りかかる前に

? **考えてみよう** 11 枚の円盤から 4 枚を取り出す。事象「4 枚すべてに番号がある」にとって、取り出す順序は関係するだろうか。同様に確からしい 4 枚の組は何通りあるだろうか。

4 枚の円盤のどの組も、取り出される組となる確率は等しいので、順序は無視してよい。同様に確からしい組は $\binom{11}{4} = 330$ 通りある。

順を追って解く

段階 1　条件なし

番号つきの円盤 4 枚からなる組は $\binom{5}{4} = 5$ 通りあるので、

$$\mathrm{P}(\text{4 枚すべてに番号がある}) = \frac{5}{330} = \frac{1}{66}.$$

段階 2　3 が最初に出る

> **考えてみよう**　番号 3 の円盤が最初に取り出されたとき、残りの 3 回の取り出しはどうなっているだろうか。

残りの 3 枚は、残る 10 枚（番号つき 4 枚、無地 6 枚）から無作為に選ばれた 3 枚の組であり、$\binom{10}{3} = 120$ 通りの組はすべて同様に確からしい。よって

$$\text{(ii)}\quad \frac{\binom{4}{3}}{120} = \frac{4}{120} = \frac{1}{30}, \qquad \text{(iii)}\quad \frac{\binom{4}{1}\binom{6}{2}}{120} = \frac{4 \times 15}{120} = \frac{1}{2},$$

である。(iii) では、残りの 3 枚のうちちょうど 1 枚に番号がなければならないからである。

段階 3　3 が（いずれかの時点で）取り出される

> **考えてみよう**　番号 3 の円盤が**いつ**取り出されたかは関係するだろうか。4 枚の中に 3 が含まれるという条件のもとで、残りの 3 枚について何が言えるだろうか。

対称性（symmetry）により、4 枚の組が 3 を含むという条件のもとで、残りの 3 枚は残る 10 枚から無作為に選ばれた 3 枚の組であり、段階 2 とまったく同じである。よって (iv) も $\frac{1}{2}$ である。

段階 4　番号つきの円盤が最初に出る

最初の円盤に番号がある（その番号が何であれ）という条件のもとで、残りの 3 枚はやはり残る 10 枚から無作為に選ばれた 3 枚の組であり、その 10 枚のうち 4 枚が番号つき、6 枚が無地である。よって (v) も $\frac{1}{2}$ である。

段階 5　番号つきの円盤が取り出される

考えてみよう 今度の条件は「4 枚のうち少なくとも 1 枚に番号がある」である。これはこれまでの条件のどれかと同じだろうか。330 通りの組のうち、この条件を満たすものはいくつあるだろうか。

この条件は異なる。特定の円盤を指定していないからである。番号つきの円盤を 1 枚も含まない組は $\binom{6}{4} = 15$ 通りしかないので、条件を満たす組は 315 通りある。このうち、番号つきの円盤をちょうど 2 枚含む組は $\binom{5}{2}\binom{6}{2} = 10 \times 15 = 150$ 通りである。よって

$$\text{(vi)} \quad \frac{150}{315} = \frac{10}{21}. \qquad \blacksquare$$

手を止めて、やってみよう なぜ (vi) は (v) と異なるのだろうか。(v) では、**どの**円盤に番号があるか（最初の円盤）がわかっていて、残りの 3 枚には何の制約もない。(vi) では、**どれか**の円盤に番号があることしかわからない。番号つきの円盤を 3 枚または 4 枚含む組は、「どれかの円盤に番号がある」をより多くの仕方で満たすので、(v) のときのように過小に数えられることがない。

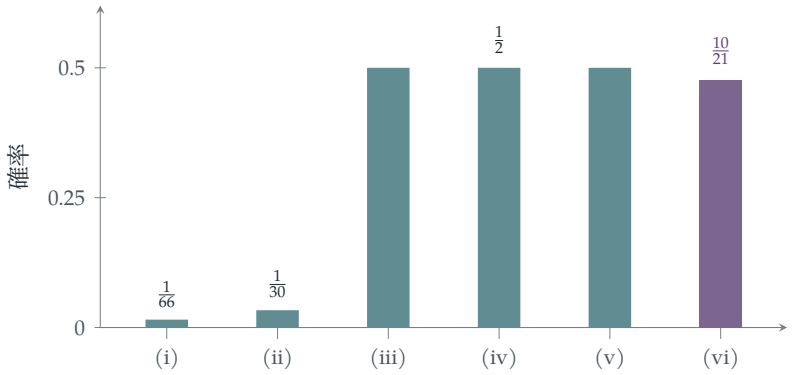


図 12.1.1 六つの答え。小問 (iii)、(iv)、(v) はいずれも**特定の**円盤に番号があることを条件としており、答えは $\frac{1}{2}$ となる。小問 (vi) は**どれか**の円盤に番号があることだけを条件としており、答えは $\frac{10}{21}$ となる。

振り返り

- **条件は標本空間（sample space）を制限する。** $\mathrm{P}(A \mid B)$ は、B に属する結果のうち A にも属するものの割合である。
- **対称性は手間を省く。** 3 が最初に取り出されたか、いずれかの時点で取り出されたかは関係なかった。すべての位置が対等だからである。

- **「特定の一つ」は「少なくとも一つ」ではない。**似たように聞こえる条件が、まったく異なる結果の集合を選び出すことがある。

背後にある数学

男の子か女の子かのパズル

マーティン・ガードナー（Martin Gardner）は 1959 年に次の問題を出した。スミス氏には子どもが二人いて、少なくとも一人は男の子である。二人とも男の子である確率はいくらか。同様に確からしい家族構成 BB、BG、GB、GG を列挙し、GG を除くと三つが残り、そのうち一つが BB であるから、答えは $\frac{1}{3}$ である。しかし、上の子が男の子であるとわかっている場合には、答えは $\frac{1}{2}$ になる。これはまさに、この問題の小問 (vi) と (v) の縮小版である。ガードナーは後に、最初に出した形の問題はあいまいであったと認めた。答えは、男の子がいることをどのようにして知ったかによるのである。子どもの一人に無作為に出会い、その子が男の子だった場合、答えは $\frac{1}{2}$ である。「息子さんが少なくとも一人いますか」と尋ねて「はい」と答えられた場合、答えは $\frac{1}{3}$ である。確率は情報だけでなく、その情報を生み出した過程にも依存するのである。

モンティ・ホール問題

同じ教訓が、確率論で最も激しく論争されてきたパズル、モンティ・ホール問題（Monty Hall problem）の根底にもある。あるゲーム番組で、三つの扉のうち一つの後ろに車がある。あなたは扉 1 を選ぶ。車の位置を知っている司会者は、必ず残り二つの扉のうち一つを開けてヤギを見せ、必ず選択を変える機会を与える。車があなたの扉の後ろにあって司会者に選択の余地がある場合には、司会者は残り二つの扉から無作為に選ぶ。司会者は扉 3 を開け、扉 2 に変えてもよいと言う。変えるべきだろうか。答えは「変えるべき」である。変えれば確率 $\frac{2}{3}$ で勝つ。司会者の扉の選び方は無作為ではない。車の位置に依存しており、それが情報をもたらすのである。正確なルールが重要である。司会者が可能な限りいつも扉 3 を開けるのであれば、扉 3 が開いたのを見たとき扉 1 と扉 2 は同様に確からしくなり、変えて勝つ確率は $\frac{1}{2}$ にすぎない。それでも、あらかじめ必ず変えると決めている参加者は、やはり $\frac{2}{3}$ の割合で勝つ。変えて勝つのは、最初の選択が外れていたときにちょうど一致するからである。マリリン・ボス・サバント（Marilyn vos Savant）が 1990 年に雑誌のコラムでこの答えを示したとき、彼女のもとには、間違っていると指摘する手紙が何千通も届き、その中には数学者からのものもあった。コンピュータによるシミュレーションか、同様に確からしい場合を注意深く列挙すれば、決着がつく。

ベイズとラプラス

条件付き確率（conditional probability）は、トマス・ベイズ（Thomas Bayes）と、ピエール＝シモン・ラプラス（Pierre-Simon Laplace）によって体系的な基礎の上に置かれた。ベイズの論文は彼の死後の 1763 年に出版され、ラプラスは 1774 年に、今日ベイズの定理（Bayes' theorem）と呼ばれる法則を一般的な形で述べた。今日あらゆる大学の講義の出発点となっている、アンドレイ・コルモゴロフ（Andrey Kolmogorov）の 1933 年の公理（axioms）は $\mathrm{P}(A \mid B) = \mathrm{P}(A \cap B)/\mathrm{P}(B)$ と定義し、確率論の残りの部分はこの定義の上に築かれている。次の節 Section 12.2 では、ベイズの定理が実際に何をするかを示す。

思考のスキル

条件について正確であれ

条件付き確率を計算する前に、どの結果が条件を満たすのかを正確に書き出すこと。その情報がどのようにして得られたかを問うこと。

大学では、条件付き確率と条件付き期待値（conditional expectation）は確率論と統計学における最も重要な道具の一つであり、同時に最もよくある誤りの源でもある。少なくとも小さな場合について標本空間を書き出すことが、最良の防御である。

次の一歩

1. **一般化する。** 無地の円盤が b 枚、番号つきの円盤が n 枚あり、k 枚を取り出すとき、(iv) と (v) の答えが常に等しいことを示せ。また、(vi) がそれらと等しくなるのはどのようなときかを求めよ。

2. **火曜日生まれの男の子。** ある家庭に子どもが二人いて、少なくとも一人は火曜日生まれの男の子である。二人とも男の子である確率が $\frac{13}{27}$ であることを示せ。(14×14 通りの同様に確からしい場合を列挙するか、うまく数えよ。)

3. **モンティ・ホール問題。** 結果を列挙して $\frac{2}{3}$ を確かめよ。司会者が車の位置を知らず、残りの扉のうち一つを無作為に開け、たまたまヤギが現れた場合はどうなるか。

4 **2 枚のエース。** 13 枚からなるブリッジの手札に、少なくとも 1 枚のエースが含まれている。少なくとも 2 枚含まれている確率はいくらか。スペードのエースが含まれているという条件のもとでの確率と比較せよ。

さらに学ぶために

- **Jason Rosenhouse, *The Monty Hall Problem*** (Oxford University Press)。このパズル、その歴史、そして数多くの変種を扱う。
- **Geoffrey Grimmett & Dominic Welsh, *Probability: An Introduction*** (Oxford University Press)。条件付き確率の章。
- **Joseph Blitzstein & Jessica Hwang, *Introduction to Probability*** (CRC Press；オンラインで無料公開)。条件付けについての解説が優れており、多くのパラドックスを扱っている。

12.2 基準率の誤謬

問題

STEP II 1998、問 12

先に進む前に、まず問題を読んでおこう。問題は STEP Support Programme の問題データベースにある：

`step.maths.org/questions/98-s2-q12`

導入

ある病気の検査は、その病気にかかっている人に対して 90% の精度をもち、誤報もほとんど出さない。Frankie の検査結果は陽性であった。Frankie はどれほど心配すべきだろうか。多くの人は、そして多くの研究によれば大半の医師さえも、「約 90%」と答える。この問題における正しい答えは $\frac{1}{11}$、すなわち約 9% である。10 分の 1 も小さいのである。

誤りの原因は**基準率**（base rate）を無視することにある。この病気はまれであるため、多数の健康な人々の中での偽陽性（false positive）の割合がわずかであっても、実際の患者数をはるかに上回る誤報が生じるのである。この**基準率の誤謬**（base-rate fallacy）は、医療、法律（そこでは検察官の誤謬と呼ばれる）、保安検査、機械学習において重要な問題となる。その処方箋がベイズの定理（Bayes' theorem）である。

取りかかる前に

考えてみよう 1000 人を想像してみよう。そのうち何人が病気にかかっているか。そのうち何人が陽性と判定されるか。健康な人のうち何人が陽性と判定されるか。

1000 人のうち 10 人が病気にかかっており、そのうち 9 人が陽性と判定される。健康な 990 人のうち $\frac{1}{11}$ が陽性と判定されるので、偽陽性は 90 人である。したがって陽性と判定されるのは 99 人であり、そのうち実際に病気なのはわずか 9 人である。

順を追って解く

段階 1　ベイズの定理

考えてみよう　「Frankie が病気にかかっている」を D、「AL が陽性である」を $+$ と書く。$\mathrm{P}(+ \mid D)$ はわかっていて、求めたいのは $\mathrm{P}(D \mid +)$ である。この二つはどのような関係にあるか。

条件付き確率（conditional probability）の定義により $\mathrm{P}(D \mid +) = \frac{\mathrm{P}(D \cap +)}{\mathrm{P}(+)}$ であり、$\mathrm{P}(+)$ は Frankie が病気にかかっているかどうかで場合分けできる：

$$\begin{aligned}\mathrm{P}(D \mid +) &= \frac{\mathrm{P}(+ \mid D)\mathrm{P}(D)}{\mathrm{P}(+ \mid D)\mathrm{P}(D) + \mathrm{P}(+ \mid \bar{D})\mathrm{P}(\bar{D})} \\ &= \frac{0.9 \times 0.01}{0.9 \times 0.01 + \frac{1}{11} \times 0.99} = \frac{0.009}{0.009 + 0.09} = \frac{1}{11}.\end{aligned}$$

■

これは 1000 人の例における $\frac{9}{99}$ にほかならない。

段階 2　二つ目の検査

考えてみよう　二つ目の検査は AL で陽性となった人だけに行われる。その集団における患者の割合はいくらか。それを新たな「基準率」として用いよ。

AL で陽性となった人のうち、$\frac{1}{11}$ が患者であり、$\frac{10}{11}$ は患者でない。これらを事前確率（prior probability）としてベイズの定理を再び適用すると、

$$\mathrm{P}(D \mid \text{両方とも陽性}) = \frac{0.9 \times \frac{1}{11}}{0.9 \times \frac{1}{11} + p \times \frac{10}{11}} = \frac{0.9}{0.9 + 10p}.$$

これを $\frac{1}{2}$ に等しいとおくと $10p = 0.9$ となり、$p = 0.09$ である。 ■

（ここでは、問題が暗に仮定しているとおり、Frankie が病気かどうかを与えたときに二つの検査の誤りは独立（independent）であると仮定している。）

手を止めて、やってみよう　$p = 0.09$ のとき、STEP の偽陽性率は AL のそれとほぼ同じであるにもかかわらず、二度目の陽性結果によって確率は $\frac{1}{11}$ から $\frac{1}{2}$ へと上がる。同じ品質の検査が、二度目にはなぜはるかに大きな効果をもつのだろうか。



- 病気にかかっている
- 健康だが陽性と判定される
- 健康で陰性と判定される

（典型的な 100 人）

図 12.2.1　典型的な 100 人。1 人が病気にかかっており（紫）、10 回中 9 回は陽性と判定される。健康な 99 人のうち約 9 人が誤って陽性と判定される（琥珀色）。陽性と判定された人の中では、健康な人が病気の人をおよそ 10 対 1 で上回る。

振り返り

- **自然頻度。** 1000 人を想像することで、答えが明白になり、確かめるのも容易になった。
- **事前確率と尤度の積。**事後確率（posterior probability）は、検査の精度と基準率の両方に依存する。
- **段階的な更新。**昨日の事後確率は今日の事前確率である。二つ目の検査は $\frac{1}{100}$ ではなく $\frac{1}{11}$ から出発した。

背後にある数学

オッズと証拠

ベイズの定理は**オッズ**（odds）を用いると最もわかりやすい。D のオッズとは $\mathrm{P}(D) : \mathrm{P}(\bar{D})$ のことである。一つの証拠 E は、オッズに**尤度比**（likelihood ratio）$\frac{\mathrm{P}(E|D)}{\mathrm{P}(E|\bar{D})}$ を掛ける：

$$\text{事後オッズ} = \text{事前オッズ} \times \text{尤度比}.$$

ここで事前オッズは $1 : 99$ であり、AL の尤度比は $0.9 \div \frac{1}{11} = 9.9$ であるから、事後オッズは $9.9 : 99 = 1 : 10$、すなわち確率 $\frac{1}{11}$ となる。尤度比 10 をもつ独立な二つ目の検査（$p = 0.09$ のときの STEP がそうである）によって、オッズは $1 : 1$ になる。対数をとると、証拠は**足し算**になる。すなわち、独立な検査はそれぞれ自分の尤度比の対数を寄与する。チューリング（Turing）はまさにこれを用い、証拠を「バン」と「デシバン」で測って、ブレッチリー・パーク（Bletchley Park）でエニグマの暗号文を解読した。

検察官の誤謬

1999 年、事務弁護士の Sally Clark は二人の幼い息子を殺害したとして有罪判決を受けた。その根拠の一部は、一つの家族で乳幼児突然死が二度起こる確率は「7300 万分の

1」であるという証拠であった。その数値自体の誤り（二つの死が独立であると仮定し、遺伝的・環境的要因を無視していた）は別としても、この推論は P(証拠 | 無実) と P(無実 | 証拠) を混同していた。母親による二重殺人もまたきわめてまれであり、比較すべきは二つの説明どうしなのである。英国王立統計学会（Royal Statistical Society）はこの誤りを指摘する公開書簡を出し、2003 年に彼女の有罪判決は覆された。その数学は、まさにこの問題のものと同じである。

Gigerenzer と自然頻度

1990 年代、心理学者の Gerd Gigerenzer は、検査の統計を条件付き確率として与えられた医師たちは、この問題のような問いにほとんどの場合ひどく間違えるが、同じ情報を自然頻度（natural frequencies）として提示されると（「1000 人の女性のうち 10 人が乳がんにかかっている。そのうち 8 人が陽性と判定される…」）大半が正しく答えることを示した。彼の研究は、多くの国で医療リスクを患者に伝える方法を変えた。

思考のスキル

ベイズの定理で条件を逆転させる

$\mathrm{P}(A \mid B)$ と $\mathrm{P}(B \mid A)$ は異なる数である。一方から他方を得るには、基準率 $\mathrm{P}(A)$ が必要である。

ベイズ的推論、すなわち証拠に照らして信念を更新することは、スパムフィルタから医療診断に至るまで、現代の統計学と機械学習の基礎である。大学では、単一の確率の代わりに事前分布と事後分布（prior and posterior distributions）を用いるベイズ推論（Bayesian inference）としてこれに出会うことになる。

次の一歩

1. **冷蔵庫の中の死体。** STEP II 1987、問 16（`step.maths.org/questions/87-s2-q16`）に挑戦してみよう。信頼性の異なる二人の証人についてのベイズの定理の問題であり、STEP の中でも特に印象的な設定の一つである。

2. **どれほどまれか。** 感度（sensitivity）0.9、偽陽性率 $\frac{1}{11}$ の検査について、陽性結果が「病気である確率が五分五分を上回る」ことを意味するためには、病気はどれほどありふれていなければならないか。

3 **検査を繰り返す。** AL そのものを繰り返し、その人の真の状態を与えたときに結果が独立であるとする。AL で二回陽性となった後に病気である確率はいくらか。三回の場合はどうか。

4 **スクリーニング。**あるスクリーニング計画で、有病率（prevalence）が 10 000 人に 1 人の病態について 100 万人を検査する。用いる検査は感度 99%、偽陽性率 1% である。その病態の可能性があると告げられるのは何人で、そのうち実際にその病態をもつのは何人か。

さらに学ぶために

- **Gerd Gigerenzer, *Reckoning with Risk*** (Penguin)。自然頻度、医療検査、そして法廷。
- **David Spiegelhalter, *The Art of Statistics*** (Pelican)。医療スクリーニングと Sally Clark 事件を含む。
- **Sharon Bertsch McGrayne, *The Theory That Would Not Die*** (Yale University Press)。チューリングによる利用を含む、ベイズの定理の歴史。

12.3 「ランダム」とは何を意味するか？

問題

STEP III 2022、問 12

先に進む前に、まず問題を読んでおこう。問題は、OCR が配布している 2022 年の STEP 過去問一式（zip ファイル）のうち、STEP III の問題用紙にある：

`www.ocr.org.uk/Images/696587-2022-past-papers.zip`

導入

「円の弦（chord）を無作為に 1 本選ぶ。その長さの期待値（expected value）はいくらか。」この問いには答えがありそうに聞こえる。しかし「無作為に」が何を意味するかを決めるまでは、答えは存在しない。円周上に無作為に 2 点をとって結べば、長さの期待値は $\frac{4a}{\pi} \approx 1.27a$ である。円の内部に無作為に 1 点をとり、その点を通る方向を無作為に選べば、答えは $\frac{16a}{3\pi} \approx 1.70a$ となる。どちらも「無作為な弦」の選び方としてまったく妥当である。

これが**ベルトランの逆説**（Bertrand's paradox）であり、ジョゼフ・ベルトラン（Joseph Bertrand）が 1889 年に警告として発表したものである。確率は、確率的な試行が明確に指定されて初めて定まる。この逆説は、20 世紀に築かれた確率論の厳密な基礎へと数学者たちを向かわせる一因となった。その基礎においては、あらゆる確率は厳密に定義された結果の空間の上で考えられる。

取りかかる前に

考えてみよう 角 AOB が θ のとき、弦 AB の長さはいくらか。（O から垂線を下ろしてみよ。）

O からの垂線は弦と角をともに二等分するので、$AB = 2a\sin\frac{\theta}{2}$ である。

順を追って解く

段階 1　円周上の 2 点

$$\mathrm{E}(AB) = \frac{1}{2\pi}\int_0^{2\pi} 2a\sin\frac{\theta}{2}\,\mathrm{d}\theta = \frac{a}{\pi}\Big[-2\cos\frac{\theta}{2}\Big]_0^{2\pi} = \frac{a}{\pi}\cdot 4 = \frac{4a}{\pi}. \qquad \blacksquare$$

段階 2　内部の無作為な点

> **考えてみよう**　$0 \le r \le a$ に対して $\mathrm{P}(R \le r)$ はいくらか。微分して確率密度を求めよ。

$\mathrm{P}(R \le r)$ は半径 r の円板の面積を円板全体の面積で割ったもの、すなわち $\frac{r^2}{a^2}$ である。したがって R は $[0,a]$ 上で確率密度（density）$\frac{2r}{a^2}$ をもつ。（中心から遠い点ほど選ばれやすい。そこにより多くの面積があるからである。）

> **考えてみよう**　C を通る弦が OC と角 t をなすとする。この弦は O からどれだけ離れているか。

O からこの直線までの距離は $R\sin t$ なので、弦の長さは $2\sqrt{a^2 - R^2\sin^2 t}$ である。

段階 3　C について平均する

$s = \sin t$ とおくと、

$$L(t) = \int_0^{a} 2\sqrt{a^2 - r^2 s^2}\cdot\frac{2r}{a^2}\,\mathrm{d}r = \frac{4}{a^2}\Big[-\frac{(a^2 - r^2s^2)^{3/2}}{3s^2}\Big]_0^{a}$$
$$= \frac{4}{a^2}\cdot\frac{a^3 - a^3(1-s^2)^{3/2}}{3s^2} = \frac{4a(1-\cos^3 t)}{3\sin^2 t}. \qquad \blacksquare$$

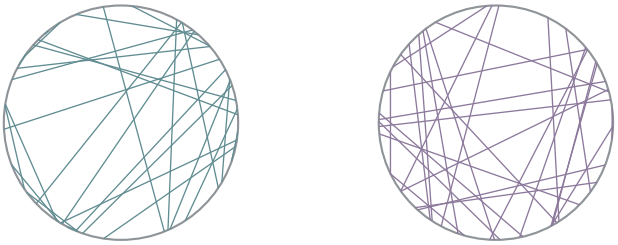
> **考えてみよう**　$1-\cos^3 t$ と $\sin^2 t = 1-\cos^2 t$ を因数分解せよ。何が約分できるか。

$c = \cos t$ とおくと $\frac{1-c^3}{1-c^2} = \frac{(1-c)(1+c+c^2)}{(1-c)(1+c)} = \frac{1+c+c^2}{1+c} = c + \frac{1}{1+c}$ であり、また $1+\cos t = 2\cos^2\frac{t}{2}$ である。よって

$$L(t) = \frac{4a}{3}\Big(\cos t + \tfrac{1}{2}\sec^2\tfrac{t}{2}\Big). \qquad \blacksquare$$

段階 4　方向も無作為に

$$\begin{aligned}\mathrm{E}\big(L(T)\big) &= \frac{2}{\pi}\int_0^{\pi/2} \frac{4a}{3}\Big(\cos t + \tfrac{1}{2}\sec^2\tfrac{t}{2}\Big)\,\mathrm{d}t \\ &= \frac{8a}{3\pi}\Big[\sin t + \tan\tfrac{t}{2}\Big]_0^{\pi/2} = \frac{8a}{3\pi}(1+1) = \frac{16a}{3\pi}. \qquad \blacksquare\end{aligned}$$



無作為な 2 端点
平均の長さ $\frac{4a}{\pi} \approx 1.27a$

円板内の無作為な中点
平均の長さ $\frac{4a}{3} \approx 1.33a$

図 12.3.1　2 通りの方法で選んだ 28 本の「無作為な弦」。左：円周上の独立な一様（uniform）な 2 点。右：円板内の一様に無作為な点を弦の中点とする。右の方法では、縁の近くの短い弦が少なくなる。方法が異なれば分布が異なり、答えも異なる。

振り返り

- **試行を指定する。**「無作為な弦」は各小問で明確な手順となり、手順ごとに異なる答えが得られた。
- **面積から確率密度を得る。** $\mathrm{P}(R \le r)$ は幾何から直接得られ、それを微分すると確率密度が得られた。
- **段階的に平均する。**まず（t を固定して）C の位置について平均し、次に t について平均した。

背後にある数学

ベルトランの 3 つの答え

ベルトラン自身の問いは、無作為な弦が内接正三角形（inscribed equilateral triangle）の 1 辺 $\sqrt{3}a$ より長くなる確率はいくらか、というものであった。彼は 3 つの方法と 3 つの答えを示した。

- 無作為な端点：一方の端点を固定する。もう一方の端点が円周の中央の 3 分の 1 にあれば、弦は長い。確率は $\frac{1}{3}$。

- 無作為な半径とその上の無作為な点を中点とする：O からの距離が $\frac{a}{2}$ より小さければ、弦は長い。確率は $\frac{1}{2}$。
- 円板内のどこかに無作為な中点をとる：中点が O から $\frac{a}{2}$ 以内にあれば、弦は長い。確率は $\frac{1}{4}$。

それぞれの方法は弦の集合の上の異なる確率分布であり、矛盾はない。これらは異なる問いに答えているのである。教訓は、「一様に無作為に」という言葉は選んだパラメータ付け（parametrisation）に対してのみ意味をもち、連続的な集合に対しては、自然なパラメータ付けがただ 1 つに定まることはふつうない、ということである。

正しい答えはあるのか

1973 年、物理学者のエドウィン・ジェインズ（Edwin Jaynes）は、2 番目の方法こそが「正しい」と論じた。その根拠は、円を少し動かしたり拡大縮小したりしても答えは変わるべきではない、というものである。無作為な弦とは、床に描いた円の上にストローを投げたようなものであるべきだ、というわけである。この不変性（invariance）をもつのは、中心からの距離が一様な分布だけであり、実際にストローを投げる実験でも $\frac{1}{2}$ が確かめられている。正しい分布とはその状況の対称性を尊重するものである、という同じ考え方が、無作為な直線の数学である**積分幾何学**（integral geometry）の根底にあり、これは Section 12.20 のビュフォン（Buffon）の針も説明する。

逆説から公理へ

ベルトランは弦の逆説を『*Calcul des probabilités*』（1889 年）で発表した。その目的の一部は、彼が「幾何学的確率」の軽率な使い方とみなしたものを皮肉ることであった。アンリ・ポアンカレ（Henri Poincaré）もこれを論じている。1900 年、ダフィット・ヒルベルト（David Hilbert）は新世紀に向けた有名な問題の中に確率の公理化（axiomatisation）を挙げ、1933 年にアンドレイ・コルモゴロフ（Andrey Kolmogorov）がこれを解決した。確率とは結果の空間の上の測度であり、「無作為な」対象についての問いはすべて、まずどの空間とどの測度を用いるかを述べなければならない。ベルトランの逆説は、なぜそれが必要なのかを示す標準的な例である。

思考のスキル

確率的な試行を定義する

確率を計算する前に、無作為な対象がどのように生成されるかを正確に述べよ。記述が異なれば答えも異なりうる。

大学では、あらゆる確率の計算は標本空間（sample space）と確率測度（probability measure）から始まる。統計学や物理学では、しばしば対称性によって正しい無作為性のモデルを選ぶことが、仕事の大きな部分を占める。

次の一歩

1. **ベルトランの確率。** 上の 3 つの確率 $\frac{1}{3}$、$\frac{1}{2}$、$\frac{1}{4}$ を確かめよ。
2. **もう 1 つの平均。** 弦の中点を円板内で一様に選ぶとき（右の図）、長さの期待値が $\frac{4a}{3}$ であることを示せ。（ヒント：これは $L(\frac{\pi}{2})$ である。なぜか。）
3. **無作為な距離。** O から弦までの距離が $[0, a]$ 上で一様であるとき、長さの期待値が $\frac{\pi a}{2}$ であることを示せ。
4. **3 つの方法、3 つの分布。** 3 つの方法それぞれについて、$[0, 2a]$ 上での弦の長さ ℓ の確率密度を求めよ。

さらに学ぶために

- **Edwin Jaynes, “The well-posed problem”**, *Foundations of Physics* 3 (1973). 不変性による議論。読みやすく、刺激的である。
- **Gábor Székely, *Paradoxes in Probability Theory and Mathematical Statistics*** (Reidel). ベルトランの逆説をはじめ、多くの逆説を扱う。
- **Daniel Klain & Gian-Carlo Rota, *Introduction to Geometric Probability*** (Cambridge University Press). 積分幾何学。いずれ読むときのために。

12.4 二つずつなら独立、全体では独立でない

問題

STEP III 2021、問 12

先に進む前に、まず問題を読んでおこう。問題は、OCR が配布している 2021 年の STEP 過去問一式（zip ファイル）のうち、STEP III の問題用紙にある：

`www.ocr.org.uk/Images/696588-2021-past-papers.zip`

導入

プレイヤー 1 と 2 が同じ目を出したとする。このことから、プレイヤー 2 と 3 が同じ目を出すかどうかについて何かわかるだろうか。意外なことに、何もわからない。これらの事象は独立（independent）なのである。プレイヤー 1 と 3 についても同様である。ところが、1 が 2 と一致し、2 が 3 と一致するならば、1 は必ず 3 と一致する。3 つの事象のどの 2 つも独立であるのに、3 つ全体としては独立ではないのである。

対ごとの独立性（pairwise independence）は完全な独立性よりも弱く、その違いは重要である。分散（variance）に関しては対ごとの独立性さえあれば十分であり、この問題はその事実を利用している。しかし、それ以外の多くの目的には十分ではない。計算機科学者は対ごとに独立な乱数を意図的に用いる。完全に独立な乱数よりもはるかに安価に生成でき、しかも多くのアルゴリズムにとってはそれで十分だからである。

取りかかる前に

考えてみよう $\mathrm{P}(X_{12}=1)$ はいくらか。$\mathrm{P}(X_{12}=1 \text{ かつ } X_{23}=1)$ はいくらか。

プレイヤー 1 がどの目を出しても、プレイヤー 2 がそれと一致する確率は $\frac{1}{6}$ である。$X_{12}=1$ と $X_{23}=1$ がともに成り立つのは 3 人全員が同じ目を出すときであり、その確率は $\frac{6}{216}=\frac{1}{36}$ である。

順を追って解く

段階 1　対ごとの独立性

$\mathrm{P}(X_{12}=1, X_{23}=1) = \frac{1}{36} = \frac{1}{6}\cdot\frac{1}{6} = \mathrm{P}(X_{12}=1)\mathrm{P}(X_{23}=1)$ であり、各変数は 2 つの値しかとらないので、X_{12} と X_{23} は独立である（2 値をとる変数では、同時分布表（joint table）の 1 つのマスが残りのマスをすべて決定する。演習問題 1 でこのことを正確に扱う）。X_{12} と X_{34} のように共通のプレイヤーをもたない対は明らかに独立である。したがって、すべての X_{ij} は対ごとに独立である。

段階 2　得点の平均と分散

得点は $S = \sum_{i<j} X_{ij}$ であり、これは $\binom{n}{2}$ 個の変数の和で、各変数の平均は $\frac{1}{6}$、分散は $\frac{1}{6}\cdot\frac{5}{6} = \frac{5}{36}$ である。期待値（expectation）はつねに和に分けられるので、$\mathrm{E}(S) = \frac{1}{6}\binom{n}{2}$ である。

> **考えてみよう**　和の分散が分散の和に等しくなるのはどのようなときか。

和の分散は、分散の和に共分散（covariance）の和の 2 倍を加えたものであり、独立な対の共分散は 0 である。したがって対ごとの独立性で十分である：

$$\mathrm{E}(S) = \frac{n(n-1)}{12}, \qquad \mathrm{Var}(S) = \frac{5}{36}\binom{n}{2} = \frac{5n(n-1)}{72}.$$

段階 3　分散の公式

Y_i の平均がすべて 0 ならば $\mathrm{Var}(\sum Y_i) = \mathrm{E}\big((\sum Y_i)^2\big)$ であり、平方を展開すると各 Y_i^2 が 1 回ずつ、各積 Y_iY_j（$i<j$）が 2 回ずつ現れる。期待値をとれば公式が得られる。 ■

段階 4　奇数・偶数のゲーム

> **考えてみよう**　$\mathrm{E}(Z_{12})$ はいくらか。$Z_{12}Z_{23}$ はどのような値をとりうるか。また、それが 0 でないのはどのようなときか。

$Z_{12}=1$ となる確率は $\frac{3}{36} = \frac{1}{12}$（2 人とも 2、2 人とも 4、または 2 人とも 6 を出す場合）であり、-1 となる確率も $\frac{1}{12}$ である。よって $\mathrm{E}(Z_{12}) = 0$、$\mathrm{E}(Z_{12}^2) = \frac{1}{6}$ である。

$Z_{12}Z_{23} \neq 0$ となるのは 3 人全員が同じ目を出すときに限られ、そのとき $Z_{12} = Z_{23}$（ともに $+1$ またはともに -1）であるから、積は 1 である。したがって $\mathrm{E}(Z_{12}Z_{23}) = \frac{1}{36}$ であり、これは $\mathrm{E}(Z_{12})\mathrm{E}(Z_{23}) = 0$ と等しくない。よってこれらは独立ではない。 ■

（さらに直接的には、$\mathrm{P}(Z_{12}=1, Z_{23}=-1) = 0$ であるのに対し、各事象の確率は単独

ではそれぞれ $\frac{1}{12}$ である。）

段階 5 **分散**

総得点 $T = \sum Z_{ij}$ の平均は 0 である。小問 (ii) により、$\operatorname{Var}(T)$ は $\binom{n}{2} \cdot \frac{1}{6}$ に、対の対すべてにわたる $\mathrm{E}(Z_{ij}Z_{kl})$ の和の 2 倍を加えたものである。共通のプレイヤーをもたない対の対の寄与は 0 である（それらは独立で、それぞれ平均が 0 だからである）。ちょうど 1 人のプレイヤーを共有する対の対はそれぞれ $\frac{1}{36}$ の寄与をもち、その個数は $n\binom{n-1}{2}$ である（共有するプレイヤーを選び、次に他の 2 人を選ぶ）。したがって

$$\begin{aligned}\operatorname{Var}(T) &= \frac{n(n-1)}{12} + 2 \cdot \frac{1}{36} \cdot \frac{n(n-1)(n-2)}{2} \\ &= \frac{n(n-1)}{36}\big(3 + (n-2)\big) = \frac{n(n^2-1)}{36}. \qquad \blacksquare\end{aligned}$$

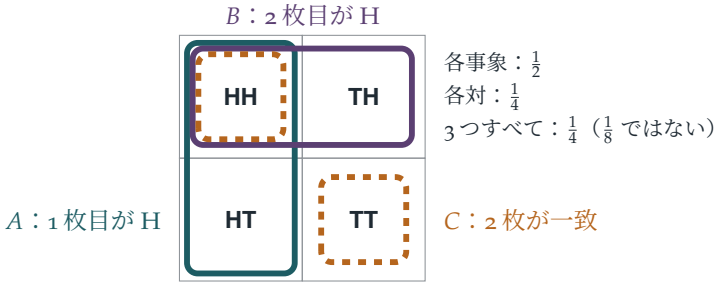


図 12.4.1 ベルンシュタイン（Bernstein）の例：公平な 2 枚の硬貨。事象「1 枚目が表」（ティール）、「2 枚目が表」（プラム）、「2 枚が一致する」（アンバー、破線）の確率はいずれも $\frac{1}{2}$ であり、どの 2 つも確率 $\frac{1}{4}$ の 1 つのマスで交わるので、これらは対ごとに独立である。しかし 3 つすべてが交わるのは HH であり、その確率は $\frac{1}{8}$ ではなく $\frac{1}{4}$ である。どの 2 つの事象も残りの 1 つを決定してしまうのである。

振り返り

- **指示変数。**得点を指示変数（indicator variable）の和として書くことで、すべてがプレイヤーの対の問題に帰着した。
- **分散に必要なのは対だけ。**和の分散は対ごとの共分散だけで決まるので、分散を足し合わせるには対ごとの独立性で十分である。
- **独立性は仮定せず確かめよ。**見た目の似た 2 つのゲームが、異なる依存構造をもっていた。

背後にある数学

複数の事象の独立性

事象 $A_1, \ldots, A_n$ が**独立**であるとは、そのどの部分集まりについても、共通部分の確率が各確率の積に等しいことをいう。3 つの事象の場合、これは 4 つの条件を意味する。対についての 3 つの条件と、$\mathrm{P}(A \cap B \cap C) = \mathrm{P}(A)\mathrm{P}(B)\mathrm{P}(C)$ である。セルゲイ・ベルンシュタイン（Sergei Bernstein）は 1920 年代に図 12.4.1 の例を示し、対についての条件から 3 つについての条件は導かれないことを示した（逆に、3 つについての条件だけからも対についての条件は導かれない）。問題における一致の事象 X_{12}、X_{23}、X_{13} は、ベルンシュタインの硬貨とまったく同じ構造をしている。

無相関は独立ではない

共分散が 0、すなわち $\mathrm{E}(XY) = \mathrm{E}(X)\mathrm{E}(Y)$ であること（無相関（uncorrelated））は、独立性よりさらに弱い。STEP III 2000 Q14（演習問題にある）は、2 つの値しかとらない変数では 2 つの概念が一致するが、3 つの値をとる場合には一致しないことを示している。簡単な例を挙げよう。X を $\{-1, 0, 1\}$ 上の一様分布（uniform distribution）に従う変数とし、$Y = X^2$ とする。すると $\mathrm{E}(XY) = \mathrm{E}(X^3) = 0 = \mathrm{E}(X)\mathrm{E}(Y)$ であるが、Y は X によって完全に決まる。共分散が測るのは線形な関連だけなのである。

安価な乱数

乱数を生成することは、計算機にとって高くつく作業である。1979 年、Larry Carter と Mark Wegman は、データを擬似乱数的な位置に格納する**ハッシュ法**（hashing）をはじめとする多くのアルゴリズムでは、対ごとに独立なランダムな選択でも完全に独立なものと同じくらいうまくいくことを示した。対ごとに独立な値は非常に安価に生成できる。たとえば、a と b をランダムに選んで $k = 0, 1, \ldots, p-1$ に対して $a + bk \pmod{p}$ をつくると、どの 2 つも独立な p 個の値が得られる。これがうまくいく理由は、この問題の分散の計算にある。チェビシェフ（Chebyshev）の不等式に必要なのは分散だけであり、分散に必要なのは対だけなのである。

思考のスキル

指示変数に分解する

ある個数の平均と分散を求めるには、数えられうる対象それぞれに *1* つずつ指示変数を対応させ、その和として個数を書け。平均には個々の指示変数が、分散には指示変数の対が必要である。

これは確率論で最も有用な手法の一つであり、すでに Section 9.8 にも登場している。大学では、ランダムグラフ（random graph）、アルゴリズム、統計的検定の解析を支える手法となる。

次の一歩

1 **2 つの値。** STEP III 2000、問 14（`step.maths.org/questions/00-s3-q14`）に挑戦せよ。この問題は、それぞれ 2 つの値しかとらない確率変数では $\mathrm{E}(XY) = \mathrm{E}(X)\mathrm{E}(Y)$ から独立性が従うことを示し、3 つの値をとる場合の反例（counterexample）を求めるものである。

2 **より大きなチーム。** 小問 (i) で $n = 7$ のとき、得点の平均と標準偏差（standard deviation）を求め、例として挙げられた得点 7 と比較せよ。

3 **再び誕生日。** 指示変数を用いて、k 人の中で誕生日が一致する対の個数の平均と分散を求めよ（Section 9.9）。それらの指示変数は対ごとに独立か。

4 **3 つについては成り立つが対ごとには成り立たない。** 確率がいずれも $\frac{1}{2}$ で、$\mathrm{P}(A \cap B \cap C) = \frac{1}{8}$ を満たすが、対ごとに独立ではない 3 つの事象を見つけよ。

さらに学ぶために

- **Geoffrey Grimmett & David Stirzaker, *Probability and Random Processes***（Oxford University Press）。独立性を扱った最初のいくつかの章。
- **Michael Mitzenmacher & Eli Upfal, *Probability and Computing***（Cambridge University Press）。対ごとの独立性とそのアルゴリズムへの応用。
- **Joseph Blitzstein & Jessica Hwang, *Introduction to Probability***（CRC Press、オンラインで無料公開）。指示変数と共分散について。

12.5 最後の乗客の座席

問題

STEP II 2021、問 11

先に進む前に、まず問題を読んでおこう。問題は、OCR が配布している 2021 年の STEP 過去問一式（zip ファイル）のうち、STEP II の問題用紙にある：

`www.ocr.org.uk/Images/696588-2021-past-papers.zip`

導入

満席の飛行機に 100 人の乗客が乗り込む。最初の乗客は搭乗券をなくしてしまい、適当な席に座る。それ以降の乗客は、自分の席が空いていればそこに座り、空いていなければ空いている席の中から無作為に 1 つを選んで座る。最後の乗客が自分の席に座れる確率（probability）はいくらだろうか。混乱が飛行機全体に連鎖していくかもしれないので、確率は小さいと予想する人もいるだろう。あるいは、ほとんどの乗客は影響を受けないので、確率は 1 に近いはずだと考える人もいるだろう。

答えは、乗客の人数によらず、ちょうど $\frac{1}{2}$ である。この問題では、それを漸化式（recurrence）と帰納法（induction）によって証明する。しかし、対称性（symmetry）を用いた美しい 1 行の議論もあり、それによって答えがなぜ $\frac{1}{2}$ になるのかが説明でき、さらにどの乗客についても答えをただちに書き下すことができる。

取りかかる前に

手を止めて、やってみよう 起こりうる場合を列挙して、P_2 と P_3 を求めよ。

座席が 2 つのとき、T_1 は S_1 に座る（このとき T_2 は S_2 に座れる）か S_2 に座るかであり、それぞれの確率は $\frac{1}{2}$ である。よって $P_2 = \frac{1}{2}$ である。座席が 3 つのときを考える。T_1 が S_1 に座れば（確率 $\frac{1}{3}$）、全員が自分の席に座れる。T_1 が S_3 に座れば、T_3 は自分の席に座れない。T_1 が S_2 に座れば、T_2 は S_1 と S_3 のうちから無作為に選び、T_3 が自分の席に座れるのは T_2 が S_1 を選んだときに限る。したがって $P_3 = \frac{1}{3} + \frac{1}{3} \cdot \frac{1}{2} = \frac{1}{2}$ で

ある。

順を追って解く

段階 1　漸化式

> **考えてみよう**　T_1 が S_k に座ったとする。ただし $2 \le k \le n-1$ とする。乗客 $T_2, \dots, T_{k-1}$ はどうなるか。T_k が来たとき、どの席が空いているか。また、その状況はより小さな列車の場合とどう比べられるか。

乗客 $T_2, \dots, T_{k-1}$ は自分の席が空いているので、そこに座る。T_k が来たとき、空いている席は S_1 と $S_{k+1}, \dots, S_n$ であり、その数は $n-k+1$ である。T_k はその中から無作為に 1 つを選び、以後の各乗客は、自分の席が空いていればそこに座り、空いていなければ無作為に選んだ席に座る。これは座席数が $n-k+1$ の元の問題そのものであり、T_k が最初の乗客の役割を、S_1 が T_k の「自分の」席（そこに座れば全てが収まる席）の役割を果たしている。したがって、T_n が S_n に座れる条件付き確率（conditional probability）は P_{n-k+1} である。

次に、T_1 の座る席で場合分けする。各席の確率は $\frac{1}{n}$ である。S_1 ならば必ず成功し、S_n ならば必ず失敗し、S_k ならば成功の確率は P_{n-k+1} である。よって

$$P_n = \frac{1}{n}\Bigl(1 + 0 + \sum_{k=2}^{n-1} P_{n-k+1}\Bigr) = \frac{1}{n}\Bigl(1 + \sum_{r=2}^{n-1} P_r\Bigr). \qquad \blacksquare$$

段階 2　答え

> **考えてみよう**　$P_2 = P_3 = \frac{1}{2}$ である。P_n を予想し、強い帰納法（strong induction）で証明せよ。

主張：すべての $n \ge 2$ に対して $P_n = \frac{1}{2}$ である。$n = 2$ のときは成り立つ。$2 \le r \le n-1$ に対して $P_r = \frac{1}{2}$ が成り立つと仮定すると、

$$P_n = \frac{1}{n}\Bigl(1 + \frac{n-2}{2}\Bigr) = \frac{1}{n} \cdot \frac{n}{2} = \frac{1}{2}. \qquad \blacksquare$$

段階 3　最後から 2 番目の乗客

> **考えてみよう**　もっと鮮やかな議論を探してみよ。最後に T_n に残される席はどれか。それが S_2 になることはありうるか。

T_n に最後に残される席は S_1 か S_n のいずれかである。$2 \le k \le n-1$ について、それが S_k であることはありえない。なぜなら、その席には T_k が座っているはずだからである。そして、それまでに行われる無作為な選択はすべて、S_1 と S_n を対称に扱う。つまり、席を奪われた乗客が選ぶときにはいつでも、S_1 と S_n が選ばれる確率は等しい。この 2 つの席のうち先に選ばれた方が結果を決めるので、どちらも等しい確率で残される席となる。よって $P_n = \frac{1}{2}$ である。

同じ議論は、任意の乗客 T_k、$k \ge 2$ についても通用する。T_k が自分の席に座れないのは、それより前の乗客がその席に座った場合に限る。席を奪われる連鎖は、誰かが S_1 を選ぶまで続く。$S_1, S_k, S_{k+1}, \ldots, S_n$ の各席は、T_k が来るまでの選択者たちによってすべて同じように扱われるので、これらのうち最初に選ばれる席は、この $n-k+2$ 個の席のどれである確率も等しい。そして T_k が自分の席に座れないのは、それが S_k である場合に限る。したがって

$$\mathrm{P}(T_k \text{ が座る席が} S_k) = \frac{n-k+1}{n-k+2}.$$

$k = n-1$ のとき（$n \ge 3$ として）、これより $Q_n = \frac{2}{3}$ を得る。また $Q_2 = \frac{1}{2}$ である。このとき $T_{n-1} = T_1$ であり、この乗客が S_1 に座る確率は $\frac{1}{2}$ だからである。 ■

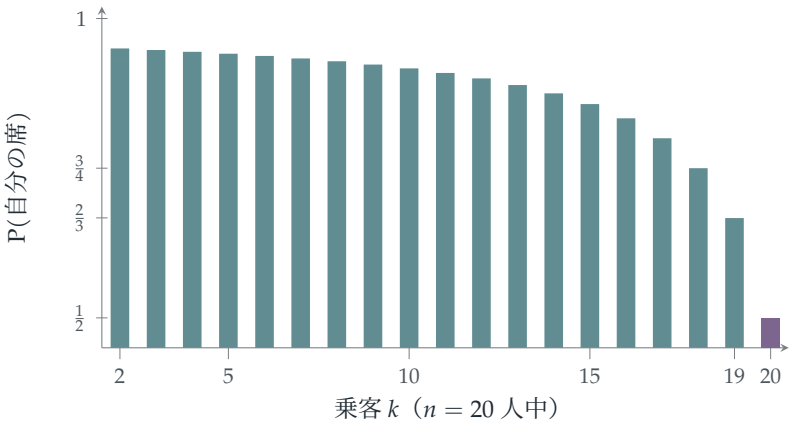


図 12.5.1　$n = 20$ の列車において、乗客 k が自分の席に座れる確率。$2 \le k \le n$ に対して、その値は $\frac{n-k+1}{n-k+2}$ であり、T_2 の $\frac{19}{20}$ から T_{19} の $\frac{2}{3}$ まで減少し、最後の乗客では $\frac{1}{2}$ となる（プラム色）。

振り返り

- **より小さな同じ問題に帰着させる。** T_1 が S_k に座った後、残りの問題は座席数の少ない同じ問題であった。
- **強い帰納法。** 漸化式にはより小さなすべての場合が現れるので、帰納法ではそれら

すべてについて結果が成り立つことを仮定した。

- **対称性を見つける。** 席 S_1 と S_n は終始入れ替え可能であり、そのことが計算を一切せずに $\frac{1}{2}$ を説明する。

背後にある数学

対称性による議論

2 つ目の証明は、確率論における強力な考え方の一例である。すなわち、2 つの結果が確率過程の対称性によって互いに対応しているならば、それらの確率は等しい。ここでの過程は、どちらかが選ばれるまで S_1 と S_n を区別しないので、両者が最後まで残る確率は等しくなければならない。対称性による議論は、長い計算をたった 1 つの観察で置き換えることが多い。この先のいくつかの節にはそうした議論がたくさん登場する。Section 12.6 での 1 回余分な硬貨投げ、Section 12.7 でのトーナメントの組み合わせ抽選などである。

席を奪われる乗客は何人か

席を奪われる乗客（自分の席に座れない乗客）の人数も、扱いやすい振る舞いをする。T_1 は S_1 を選ばない限り席を奪われた状態になり、T_k（$k \geq 2$）が席を奪われる確率は $\frac{1}{n-k+2}$ である。これらを足し合わせると、席を奪われる乗客の人数の期待値（expected value）は

$$\left(1 - \frac{1}{n}\right) + \left(\frac{1}{n} + \frac{1}{n-1} + \cdots + \frac{1}{2}\right) = H_n - \frac{1}{n},$$

となり、およそ $\ln n$ である。乗客が 300 人の飛行機であっても、混乱の影響を受けるのは平均して約 6 人にすぎない。ここでもまた調和数（harmonic numbers）が現れる（Section 9.11）。

思考のスキル

対称性を探せ

確率過程が 2 つの結果を区別できないならば、それらは等しい確率で起こる。両者が初めて区別できるようになる瞬間を見つけよ。

大学では、対称性による議論は確率論の至るところに現れる。鏡像原理（reflection principle）（Section 9.10）から交換可能性（exchangeability）まで、

その例は幅広い。交換可能性とは、確率変数（random variables）の列がどのような順序に並べ替えても同じ分布（distribution）をもつという性質であり、現代統計学の多くの部分の基礎となっている。

次の一歩

1. **忘れっぽい乗客が2人。**最初の2人の乗客がともに無作為に席を選ぶとき、最後の乗客が自分の席に座れる確率はいくらか。

2. **席を奪われる乗客。**$n = 3$ の場合について、起こりうる場合を列挙して、席を奪われる乗客の人数の期待値を確かめよ。

3. **遠回りの道。**漸化式を用いて、$\sum_{r=2}^{n-1} P_r = \frac{n-2}{2}$ が対称性による議論と矛盾しないことを示せ。

4. **最後の席。**最後の乗客が自分の席に座れないという条件のもとで、最後の乗客は確率1で S_1 に座ることを示せ。それはなぜか。

さらに学ぶために

- **Peter Winkler, *Mathematical Puzzles: A Connoisseur's Collection***（A K Peters）。『数学パズル：愛好家のためのコレクション』。飛行機の座席パズルをはじめ、対称性によって解ける多くのパズルを収録している。
- **Frederick Mosteller, *Fifty Challenging Problems in Probability***（Dover）。『確率の50の難問』。古典的な問題集である。
- **Joseph Blitzstein & Jessica Hwang, *Introduction to Probability***（CRC Press；オンラインで無料公開）。『確率論入門』。

12.6 もう 1 枚の硬貨

問題

STEP I 2016、問 12

先に進む前に、まず問題を読んでおこう。問題は STEP Support Programme の問題データベースにある：

`step.maths.org/questions/16-s1-q12`

導入

ボブはアリスより 1 枚多くコインを投げる。ボブの表の回数がアリスより多くなる確率（probability）はいくらか。ボブには有利な点があるが、表が出る保証はなく、引き分けはボブの負けとなる。最初の二つの場合を力ずくで計算すると、答えはどちらも $\frac{1}{2}$ になる。これは偶然だろうか。

偶然ではない。ボブは表の回数がアリスより多いか、裏の回数がアリスより多いかのどちらか一方であり、両方が同時に起こることはない。このことに気づけば、一文で済む証明（proof）がある。この問題は、計算がパターンを示唆し、巧妙な対称性（symmetry）がそれを説明するという流れの手本である。

取りかかる前に

考えてみよう アリスが 2 回投げるとき、表の回数の分布（distribution）はどうなるか。ボブが 3 回投げるときはどうか。

アリスの表の回数が 0, 1, 2 となる確率はそれぞれ $\frac{1}{4}, \frac{2}{4}, \frac{1}{4}$ であり、ボブの表の回数が 0, 1, 2, 3 となる確率はそれぞれ $\frac{1}{8}, \frac{3}{8}, \frac{3}{8}, \frac{1}{8}$ である。

順を追って解く

段階 1 力ずくで計算する

小問 (i) では、アリスの表の回数について和をとる。

$$\begin{aligned}\mathrm{P}(B > A) &= \tfrac{1}{4}\mathrm{P}(B \geq 1) + \tfrac{2}{4}\mathrm{P}(B \geq 2) + \tfrac{1}{4}\mathrm{P}(B = 3) \\ &= \tfrac{1}{4}\cdot\tfrac{7}{8} + \tfrac{1}{2}\cdot\tfrac{4}{8} + \tfrac{1}{4}\cdot\tfrac{1}{8} = \frac{7+8+1}{32} = \frac{1}{2}.\end{aligned}$$

小問 (ii) は項が増えるだけで同じ計算である。同様に確からしい $8 \times 16 = 128$ 通りの結果を数えると、ボブがリードするのは 64 通りであり（図 12.6.1）、やはり答えは $\frac{1}{2}$ となる。

段階 2　ボブの試行を分ける

考えてみよう　ボブの $n+1$ 回の試行を、n 回の試行とそれに続く 1 回の追加の試行とみなそう。最初の n 回を終えた時点で、ボブはアリスより多いか、同数か、少ないかのいずれかである。それぞれの場合に何が起こるか。

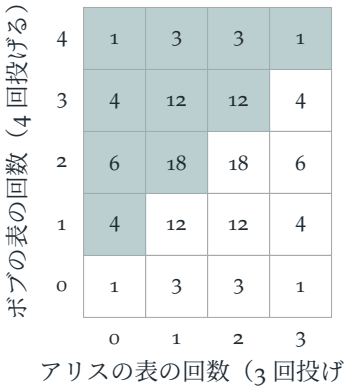
互いに n 回ずつ投げた時点で、ボブがリードしている確率を p_2、同数である確率を p_1 とすると、（アリスとボブの対称性により）ボブが負けている確率も p_2 である。リードしていれば、ボブはそのままリードを保つ。同数であれば、追加の 1 回で確率 $\frac{1}{2}$ でリードする。負けていれば、追加の 1 回で表が出てもせいぜい同数になるだけである。したがって

$$\mathrm{P}(\text{ボブの表が多い}) = p_2 + \tfrac{1}{2}p_1.$$

三つの場合ですべてを尽くすので $2p_2 + p_1 = 1$ であり、よって $p_2 + \frac{1}{2}p_1 = \frac{1}{2}$ である。コインが 1 枚多ければ、どの n に対しても、ボブの表の回数が多くなる確率はちょうど $\frac{1}{2}$ である。 ■

手を止めて、やってみよう　一文で済む証明を示そう。ボブは $n+1$ 回投げるので、表と裏の回数の和は $n+1$ であり、アリスより 1 多い。ボブの表の回数がアリスより多いことと、ボブの裏の回数がアリスより多くないことが同値であることを示せ。表と裏は対称なので、この二つの事象の確率は等しく、したがってそれぞれ $\frac{1}{2}$ である。

実際、$H_A + T_A = n$、$H_B + T_B = n+1$ と書くと、$H_B > H_A$ ならば $H_B \geq H_A + 1$ なので $T_B = n + 1 - H_B \leq n - H_A = T_A$ となる。逆に $T_B \leq T_A$ ならば $H_B \geq H_A + 1$ である。よって「表が多い」と「裏が多い」は互いに余事象（complementary event）であり、表と裏を入れ替えると（コインに偏りがないので確率は変わらない）この二つが入れ替わる。



ボブがリード：
$15 + 33 + 15 + 1 = 64$
（128 通り中）

図 12.6.1 小問 (ii)：同様に確からしい 128 通りの結果のうち、表の回数の各組合せとなるものの数。ボブがリードするのは網掛けのマスであり、そこにはちょうど半分の結果が含まれる。

振り返り

- **小さな場合を計算する。**二つの計算から、答えが $\frac{1}{2}$ であると予想できた。
- **自然な分け方で条件付ける。**ボブの追加の 1 回を切り離すことで、問題は互角の勝負とコイン 1 枚に帰着された。
- **互いに余事象となる組を見つける。**「表が多い」と「裏が多い」は互いに余事象であり、対称性によって入れ替わる。

背後にある数学

対称性と余事象

この鮮やかな証明は二つの事実を組み合わせている。対称性によって入れ替わる二つの事象は確率が等しいこと、そして互いに余事象である二つの事象は確率の和が 1 になることである。両方が成り立てば、それぞれの確率は $\frac{1}{2}$ である。同じ論法は確率論のいたるところに現れる。飛行機の座席の問題（Section 12.5）では、S_1 と S_n が対称性によって入れ替わり、最後にはそのちょうど一方が残るのであった。ヘックス（Hex）というゲームでは、盤面の対称性と、必ずどちらか一方のプレイヤーが勝つという事実から、先手が有利であることが示される。ただし、大きな盤面での必勝法は誰も知らない。

引き分けはどのくらい起こりやすいか

投げる回数が等しい場合には、引き分けの確率 p_1 が重要になる。アリスとボブが引き分けるのはアリスの表の回数とボブの表の回数が等しいときであり、これは（対称性によりボブの表の回数を裏の回数に置き換えると）アリスの表の回数とボブの表の回数の和が n になる確率、すなわち $2n$ 回の試行で表が合計 n 回出る確率に等しい。よって、スターリングの近似（Stirling's approximation）（Section 12.23）を用いると

$$p_1 = \binom{2n}{n}\frac{1}{4^n} \approx \frac{1}{\sqrt{\pi n}},$$

となる。引き分けはまれになっていくが、その減り方はゆっくりである。互いに 100 回ずつ投げても、引き分けの確率はまだ約 5.6% ある。

思考のスキル

> **計算し、予想し、そして説明する**
>
> 計算の結果が怪しいほど単純な答えになったら、その理由を探そう。単純な答えには、たいてい単純な説明がある。
>
> 大学では、この習慣が「理解すること」と「計算すること」を分ける。ここでの一つ目の証明は答えを確かめるものであり、二つ目の証明は答えを説明するもので、すぐに一般化できる（演習 2 を試してみよう）。

次の一歩

1. **コインが 2 枚多い場合。** ボブが $n+2$ 枚、アリスが n 枚のコインを投げるとき、$\mathrm{P}(\text{ボブの表が多い}) = \frac{1}{2} + \frac{1}{2}\mathrm{P}(H' = H_A)$ を証明せよ。ただし H' はボブの最初の $n+1$ 回の試行における表の回数である。

2. **偏りのあるコイン。** 二人とも表の出る確率が p のコインを使うとき、1 回多く投げる場合の答えはやはり $\frac{1}{2}$ だろうか。鮮やかな証明のどの段階が成り立たなくなるか。

3. **引き分け。** 本文の工夫を用いて $p_1 = \binom{2n}{n}/4^n$ を示し、$n = 1, 2, 3$ に対してその値を求めよ。

4. **サイコロ。** アリスが n 個、ボブが $n+1$ 個のサイコロを振る。ボブの 6 の目の個数が多くなる確率は $\frac{1}{2}$ に等しいだろうか。

さらに学ぶために

- **Peter Winkler, *Mathematical Mind-Benders***（A K Peters）。対称性を用いた解法をもつ確率のパズルが多数収められている。
- **Gabor Székely, *Paradoxes in Probability Theory and Mathematical Statistics***（Reidel）。確率論と数理統計学におけるパラドックス集。
- **Frederick Mosteller, *Fifty Challenging Problems in Probability***（Dover）。確率の難問 50 題。

12.7 勝ち抜き戦で誰と誰が当たるか

問題

STEP II Specimen paper、問 16

先に進む前に、まず問題を読んでおこう。問題は STEP Support Programme の問題データベースにある：

`step.maths.org/questions/spec-s2-q16`

導入

ウィンブルドン（Wimbledon）では、128 人の選手がドローに名を連ねる。無作為に 2 人を選んだとき、その 2 人がどこかで対戦する確率（probability）はいくらだろうか。各ラウンドで対戦する確率を足し合わせようとするかもしれない。それぞれは、勝ち残る確率と組み合わせで当たる確率の積である。この方法でも求められるが、はるかによい方法がある。試合の数を数えるのである。どの試合も 2 人の選手の組の間で行われ、試合はちょうど 127 試合ある。対称性（symmetry）により、どの選手の組も、それぞれの試合の対戦カードになる確率は等しい。したがって、求める確率は 127 を組の総数で割ったものになる。

この問題は、「正しいものを数える」ことで確率の問題を解く格好の例である。答え $\frac{1}{2^{n-1}}$ はたった 1 行で得られる。

取りかかる前に

考えてみよう 2^n 人の選手によるノックアウト方式のトーナメント（knockout tournament）では、何試合が行われるか。（各試合でちょうど 1 人の選手が敗退する。）

各試合でちょうど 1 人が敗退し、1 人を除く全員が敗退するのだから、試合数は $2^n - 1$ である。

順を追って解く

段階 1　1 回戦

選んだ 2 人の選手を X と Y とする。1 回戦では、X の対戦相手は他の $2^n - 1$ 人のうち誰である確率も等しいので、

$$\mathrm{P}(\text{出会うのがラウンド1}) = \frac{1}{2^n - 1}.$$

段階 2　鍵となる対称性

> **考えてみよう**　各ラウンドの組み合わせは無作為に決まり、どの試合も公平な硬貨投げで勝敗が決まる。この過程は、ある選手の組を他の組と区別して扱うだろうか。

区別しない。組み合わせは無作為で、どの試合も 50–50 なので、選手の番号を付け替えてもどの確率も変わらない。したがって、特定の試合（たとえば決勝）は、$\binom{2^n}{2}$ 通りの組のどれの対戦になる確率も等しい。

段階 3　決勝

決勝は特定の 1 試合なので、それが X と Y の対戦である確率は

$$\frac{1}{\binom{2^n}{2}} = \frac{2}{2^n(2^n - 1)}.$$

である。（検算：2 人とも決勝に進む必要があり、そのためには 2 人が異なる半分の山に入り、それぞれ $n - 1$ 試合に勝たなければならない。）

段階 4　いずれかの段階で

> **考えてみよう**　X と Y が対戦するのは高々 1 回である。このことから、試合ごとの確率を足し合わせてよいのはなぜか。

試合 m が X と Y の対戦であるとき 1、そうでないとき 0 をとるものを I_m とする。$2^n - 1$ 試合のそれぞれについて $\mathrm{P}(I_m = 1) = \binom{2^n}{2}^{-1}$ であり、これらの事象は互いに排反（mutually exclusive）である（2 人が 2 回対戦することはない）から、

$$\mathrm{P}(\text{出会う}) = \frac{2^n - 1}{\binom{2^n}{2}} = \frac{2(2^n - 1)}{2^n(2^n - 1)} = \frac{1}{2^{n-1}}. \qquad \blacksquare$$

ウィンブルドンの場合は $n = 7$ であり、特定の 2 人が対戦する確率は $\frac{1}{64}$ である。検算

として、1 回戦は $2^n - 1$ 試合のうち 2^{n-1} 試合を占めるので、$\frac{2^{n-1}}{\binom{2^n}{2}} = \frac{1}{2^n-1}$ となり、段階 1 の結果と一致する。

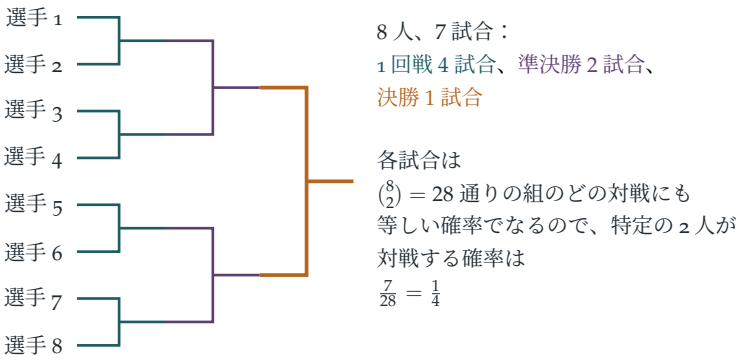


図 12.7.1　$2^3 = 8$ 人の選手によるノックアウト方式のトーナメント（問題では各ラウンドで無作為に組み合わせをやり直すが、ここでは固定した組み合わせ表として描いている）。試合は 7 試合ある。対称性により、各試合は 28 通りの組のどれの対戦にも等しい確率でなるので、選んだ 2 人が対戦する確率は $\frac{7}{28} = \frac{1}{4} = \frac{1}{2^{n-1}}$ である。

振り返り

- **試合を数える。**試合数 $2^n - 1$ が鍵となる事実であった。
- **組の間の対称性。**どの選手の組も、特定の試合で対戦する確率は等しい。
- **試合についての線形性。**2 人の選手が対戦するのは高々 1 回なので、対戦する確率は対戦回数の期待値に等しく、試合についての和として表される。

背後にある数学

個数の期待値

段階 4 の議論は、実は期待値（expectation）の計算である。X と Y が対戦する回数は、試合ごとに 1 つずつの指示変数（indicator）の和であり、その期待値（値は 0 か 1 なので、対戦する確率に等しい）は、各指示変数の期待値の和に等しい。これが期待値の線形性（linearity of expectation）であり、Section 9.11 や Section 12.4 でも用いた。この方法の大きな利点は、ラウンド間の複雑な従属関係をすべて無視できることである。

シード付きのトーナメント

現実のトーナメントはシード（seeded）制である。最も強い選手たちは、終盤のラウンドまで対戦しないように配置される。また、現実の試合は 50–50 ではない。どちらも対称性を壊すので、特定の 2 人が対戦する確率を求めるのは本格的な計算になる。スポーツをモデル化する統計学者は、各試合の確率を選手のレーティングから推定したうえで、まさにこうした計算を用いてトーナメントの結果を予測する。対称な場合は、現実の効果を測るための基準となるのである。

思考のスキル

> **正しいものを数える**
>
> 確率は多くの場合、数える対象を正しく選びさえすれば、（条件を満たす対象の個数）÷（同様に確からしい対象の個数）として求められる。ここでは、それが試合と組であった。
>
> 何を数えるかを選ぶことこそ、組合せ確率論（combinatorial probability）の技である。大学では、同じ考え方が**二重数え上げ**（double counting、Section 7.9）として、また確率的手法（probabilistic method）として現れる。確率的手法では、ある構造の個数の期待値を用いて、その構造が存在することを証明する。

次の一歩

1. **ラウンドごとに。** X と Y がラウンド k で対戦する確率が $\frac{2^{n-k}}{\binom{2^n}{2}}$ であることを直接示し、これらを足すと $\frac{1}{2^{n-1}}$ になることを確かめよ。

2. **固定された組み合わせ表。** 代わりに、組み合わせ表が（図のように）あらかじめ固定されていて、すべての選手がそこに無作為に配置されるとする。答えは変わるか。

3. **強い選手。** 選手 X がすべての試合に勝つとする。このとき、X が Y と対戦する確率はいくらか。

4. **リーグ戦。** $2n$ チームによるリーグ戦（league）で、どの 2 チームも 1 回ずつ対戦するとき、試合数はいくらか。これに対応する「対戦する確率」は何か。

さらに学ぶために

- **Frederick Mosteller, *Fifty Challenging Problems in Probability***（Dover）。トーナメントの問題も収められている。
- **David Spiegelhalter, *The Art of Statistics***（Pelican）。モデル化と予測について。
- **Joseph Blitzstein & Jessica Hwang, *Introduction to Probability***（CRC Press、オンラインで無料公開）。指示変数と対称性について。

12.8 ペニーのゲーム

問題

STEP III 2001、問 13

先に進む前に、まず問題を読んでおこう。問題は STEP Support Programme の問題データベースにある：

`step.maths.org/questions/01-s3-q13`

導入

あなたと友人がそれぞれ硬貨投げ 3 回分の列を 1 つずつ選び、どちらかの列が現れるまで硬貨を投げ続ける。3 回分の列はどれも、どの時点においても等しい確率（probability）で現れるのだから、どれを選んでも同じくらい有利なはずではないか。まったくそうではない。友人がどの列を選んでも、あなたはそれに勝つ列を選ぶことができ、ときには 7 対 1 という差をつけることさえできる。

これは *Penney* のゲーム（Penney's game）であり、1969 年に Walter Penney が考案したものである。この問題では、A が B に勝ち、B が C に勝ち、C が D に勝ち、D が A に勝つことを示す。じゃんけんのような循環である。「勝つ」という関係は推移的（transitive）ではない。後から選ぶ者は、長い目で見れば常に勝つことができるのである。

取りかかる前に

考えてみよう ここまでの投げの結果が TT で終わっていて、まだどちらも勝っていないとする。TTH より先に THH が現れることはありうるか。

TT の状態から T が続く限り、最後の 2 回は TT のままであり、最初の H が出た時点で TTH が完成する。THH が現れるには TH が必要だが、TH は T の後に H が出ることでしか生じず、その H によってすでに TTH が完成してしまう。したがって、いったん TT が現れれば A の勝ちである。

順を追って解く

段階 1　状態

> **考えてみよう**　どちらの列も 2 回の投げでは完成しない。2 回投げた後、それぞれの状態から次の 1 回はどのような効果をもつか。

ゲームの行方は最後の 2 回の結果だけで決まるので、状態（state）（最後の 2 回の結果）を追跡すればよい。

- TT から：A が勝つ（上で示した）。よって $p_{\mathrm{TT}} = 1$ である。
- HT から：次の 1 回で TH か TT になる。よって $p_{\mathrm{HT}} = \frac{1}{2}p_{\mathrm{TH}} + \frac{1}{2}p_{\mathrm{TT}}$ である。
- TH から：H が出れば THH が完成し（B の勝ち）、T が出れば HT になる。よって $p_{\mathrm{TH}} = \frac{1}{2}\cdot 0 + \frac{1}{2}p_{\mathrm{HT}}$ である。
- HH から：H が出れば再び HH、T が出れば HT になる。よって $p_{\mathrm{HH}} = \frac{1}{2}p_{\mathrm{HH}} + \frac{1}{2}p_{\mathrm{HT}}$ であり、$p_{\mathrm{HH}} = p_{\mathrm{HT}}$ を得る。

段階 2　解く

$p_{\mathrm{TH}} = \frac{1}{2}p_{\mathrm{HT}}$ を HT の式に代入すると $p_{\mathrm{HT}} = \frac{1}{4}p_{\mathrm{HT}} + \frac{1}{2}$ となるので、$p_{\mathrm{HT}} = \frac{2}{3}$、$p_{\mathrm{HH}} = \frac{2}{3}$、$p_{\mathrm{TH}} = \frac{1}{3}$ である。4 つの初期状態はそれぞれ確率 $\frac{1}{4}$ で生じるから、

$$\mathrm{P}(A \text{ が勝つ}) = \tfrac{1}{4}\left(\tfrac{2}{3} + \tfrac{2}{3} + \tfrac{1}{3} + 1\right) = \tfrac{2}{3}. \qquad \blacksquare$$

段階 3　B 対 C

> **考えてみよう**　C は HHT を、B は THH を望んでいる。C が勝つのはどのようなときか。（最初の HH の直前に何がなければならないかを見よ。）

最初の 2 回が HH であれば、その後に出る最初の T によって、THH が現れる前に HHT が完成する。C の勝ちである。そうでなければ、HH が初めて現れるより前のどこかに T があるので、最初の HH の直前は必ず T である。その瞬間に THH が現れ、これは（HH の後に T を必要とする）HHT より先である。したがって C が勝つのはゲームが HH で始まるときに限られ、

$$\mathrm{P}(B \text{ が勝つ}) = \tfrac{3}{4}. \qquad \blacksquare$$

段階 4　循環

同じ方法（あるいは「背後にある数学」の節で述べる近道）により、C（HHT）は確率 $\frac{2}{3}$ で D（HTT）に勝ち、D（HTT）は確率 $\frac{3}{4}$ で A（TTH）に勝つことがわかる。後者は、ステージ 3 で表と裏を入れ替えた鏡像である。よって

A が勝つ相手はB, B が勝つ相手はC, C が勝つ相手はD, D が勝つ相手はA. ■

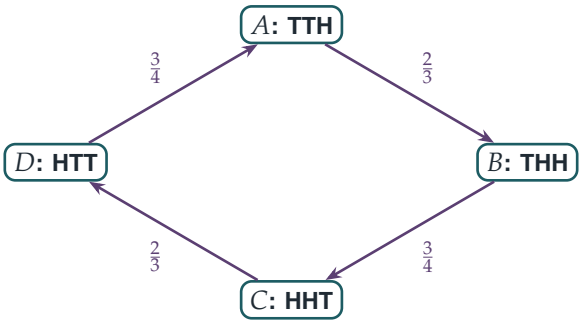


図 12.8.1 Penney の循環。X から Y への矢印は、X の列が Y の列に、示された確率で勝つことを表す。どの列にもそれに勝つ列があるので、後から選ぶ者は常に勝つ列を選ぶことができる。

振り返り

- **適切な状態を選ぶ。**重要なのは最後の 2 回だけだったので、4 つの状態と 4 つの方程式でゲーム全体を捉えることができた。
- **第一ステップ解析。**各方程式は次の 1 回の結果で条件づけたものである。ある状態からの確率は、次の 1 回で移る先についての平均である。
- **決定的な事象を探す。**ステージ 3 では、「最初の HH の直前は T である」という事実によって、方程式を使わずに勝敗が決まった。

背後にある数学

なぜ有利な列があるのか

3 回分の列はどれも、どの位置においても等しい確率で現れ、それぞれの期待（expected）待ち時間（waiting time）は容易に計算できる。しかし待ち時間は列によって異なる。TTH は平均 8 回で現れるが、TTT は平均 14 回かかる。自分自身と重なりをもつ列（TTT のように、惜しくも TTH となると 2 回分の進捗が無駄になるもの）は現れるのが遅いのである。そして 2 つの列の競争で重要なのは、それぞれが**相手の列**とどう重なるかである。THH が HHT に勝つのは、HHT のどの出現も（一番最初を除けば）直

前に T があり、つまりその時点で THH がちょうど現れているからである。後から選ぶ者の秘訣は常に同じである。相手の列の最初の 2 つの記号をとり、その前に相手の列の真ん中の記号の反対を置けばよい。

コンウェイの公式

ジョン・コンウェイ（John Conway）は、この勝率について驚くべき公式を発見した。長さ n の 2 つの列 X, Y に対して、**相関**（correlation）XY を n 桁の 2 進数として次のように定める。左から読んで第 i 桁は、X の第 i 記号から始まる部分が、それと同じ長さの Y の先頭部分に一致すれば 1 とする。このとき、X より先に Y が現れる勝率の比は

$$(XX - XY) : (YY - YX).$$

である。$X = \text{HHT}$, $Y = \text{THH}$ の場合、$XX = 100_2 = 4$（HHT 全体だけが自分自身と一致する）、$XY = 001_2 = 1$（HHT の最後の T だけが THH の先頭と一致する）、$YY = 100_2 = 4$、$YX = 011_2 = 3$（THH の末尾の HH と H がともに HHT の先頭と一致する）である。よって THH が先に現れる勝率の比は $(4-1) : (4-3) = 3 : 1$ であり、これはステージ 3 の $\frac{3}{4}$ である。コンウェイの公式は任意の長さ、任意の文字集合の列に対して成り立ち、その証明には**マルチンゲール**（martingale）の理論が用いられる。列が現れることに賭けるギャンブラーたちがいるカジノを想像し、公平なゲームの期待利益は 0 であるという事実を使うのである。

Penney とガードナー

Walter Penney は 1969 年、*Journal of Recreational Mathematics* 誌の短い記事でこのゲームを紹介した。マーティン・ガードナー（Martin Gardner）は 1974 年、*Scientific American* 誌の連載コラムでこれを有名にし、そこにはコンウェイが発見したものの発表していなかったコンウェイの公式も含まれていた。この種の非推移的な関係（エフロン（Efron）のサイコロもその一例で、サイコロ A が B に勝ち、B が C に勝ち、C が D に勝ち、D が A に勝つ）は繰り返し驚きを生む源であり、多数派の選好が循環しうる投票理論においても重要である。これがコンドルセ（Condorcet）のパラドックス（Condorcet's paradox）である。

思考のスキル

第一ステップ解析

状態の間を移り変わる過程について確率や期待値（*expectation*）を求めるには、最初の 1 ステップで条件づけよ。そうすれば、状態ごとに 1 つの一次方程式が得られる。

これはマルコフ連鎖（Markov chain）の基本的な手法であり、Section 12.16 や、大学の確率論の講義全体を通じて出会うことになる。同じ方法で、ある状態に別の状態より先に到達する確率、吸収（absorption）までの期待時間、賭けの問題における破産（ruin）の確率を求めることができる。

次の一歩

1. **4 人同時に。** STEP II 2015、問 12（`step.maths.org/questions/15-s2-q12`）に挑戦せよ。4 人が同時に行う Penney のゲームである。
2. **待ち時間。** HH が初めて現れるまでの投げの回数の期待値は 6、HT の場合は 4 であることを示せ。（第一ステップ解析を用いよ。）
3. **最善の応手。** 長さ 3 の 8 つの列それぞれについて、それに最大の確率で勝つ列を求めよ。最も簡単に負かされる列はどれか。
4. **コンウェイの公式。** コンウェイの公式を用いて、HHT が確率 $\frac{2}{3}$ で HTT に勝つことを確かめよ。

さらに学ぶために

- **Martin Gardner, *Time Travel and Other Mathematical Bewilderments***（W. H. Freeman）。Penney のゲームとコンウェイの公式を収録している。
- **Ronald Graham, Donald Knuth & Oren Patashnik, *Concrete Mathematics***（Addison-Wesley）、8.4 節。母関数（generating function）を用いた Penney のゲームの扱い。
- **David Williams, *Probability with Martingales***（Cambridge University Press）。マルチンゲールによる証明。将来のために。

12.9 偏った硬貨から公平な判定を

問題

STEP I 2019、問 11

先に進む前に、まず問題を読んでおこう。問題は、OCR が配布している 2019 年の STEP 過去問一式（zip ファイル）のうち、STEP I の問題用紙にある：

`www.ocr.org.uk/Images/696590-2019-past-papers.zip`

導入

公平な決定を下さなければならないのに、手元にある硬貨は曲がっていて、どれほど偏っているのかもわからない。それでも公平な選択はできるだろうか。1951 年にフォン・ノイマン（John von Neumann）が与えた答えはこうである：硬貨を 2 回投げよ。HT が出ればホテル、TH が出れば喫茶店を選び、HH か TT が出たらもう 2 回投げる。HT と TH はどちらも確率が pq であるから、p がどんな値であっても、この決定は完全に公平である。

この問題の手順はその逆で、HH と TT で決定する。これは速いが公平ではない（ホテルが選ばれる確率は $\frac{p^2}{p^2+q^2}$）。この問題は、どんな硬貨であってもこの手順が長くはかからないことを示す。2 つの手順を比べると、乱数生成における基本的なトレードオフ、すなわち速さと公平さの兼ね合いが見えてくる。

取りかかる前に

考えてみよう 1 回のラウンドで決定が下される確率はいくらか。決定が下されない確率はどうか。$2pq$ がとりうる最大の値はいくらか。

硬貨の結果が一致すれば決定が下され、その確率は $p^2 + q^2$ である。そうでなければ、確率 $2pq$ で続行となる。$pq = p(1-p) \leq \frac{1}{4}$（等号は $p = \frac{1}{2}$ のとき）であるから、$2pq \leq \frac{1}{2}$ が成り立つ。

順を追って解く

段階 1　第 n ラウンドでの決定

各ラウンドは独立（independent）である。第 n ラウンドで決定が下されるには、$n-1$ 回続けて結果が一致せず、その次に一致しなければならない：

$$\mathrm{P}(\text{決定するラウンドが}n) = (2pq)^{n-1}(p^2+q^2). \qquad \blacksquare$$

段階 2　第 n ラウンドまでの決定

> **考えてみよう**　最初の n ラウンドで決定が下されない確率はいくらか。

n ラウンドで決定が下されない確率は $(2pq)^n \le \left(\frac{1}{2}\right)^n$ であるから、p がどんな値であっても、第 n ラウンドまでに決定が下される確率は少なくとも $1-\frac{1}{2^n}$ である。 $\blacksquare$

段階 3　3 人の場合

> **考えてみよう**　第 1 ラウンドで表が 2 枚、裏が 1 枚出たあと、3 人目は裏の硬貨をひっくり返して表にする。第 2 ラウンドで決定が下されるには何が起こらなければならないか。裏が 2 枚、表が 1 枚の場合はどうか。

第 1 ラウンドで決定が下される確率は p^3+q^3 である。そうでなければ、2 枚の硬貨が一致し、1 枚だけが異なる。多数派が表であった場合（確率 $3p^2q$）、異なる硬貨は表に返されるので、第 2 ラウンドで決定するには残りの 2 人が表を出す必要があり、その確率は p^2 である（全員が裏になることはもはやありえない）。同様に、多数派が裏であった場合（確率 $3pq^2$）、決定には裏が 2 枚出る必要があり、その確率は q^2 である。したがって

$$\mathrm{P}(\text{決定までに要するラウンド数が高々2}) = p^3+q^3+3p^2q\cdot p^2+3pq^2\cdot q^2.$$

> **考えてみよう**　すべてを $s=pq$ を用いて表せ。（$p+q=1$ を用いよ。）

$p^3+q^3=(p+q)^3-3pq(p+q)=1-3s$ であり、$3p^4q+3pq^4=3s(p^3+q^3)=3s(1-3s)$ である。よって、$s\le\frac{1}{4}$ より

$$\mathrm{P} = 1-3s+3s-9s^2 = 1-9s^2 \ge 1-\frac{9}{16} = \frac{7}{16},$$

となる。 $\blacksquare$

振り返り

- **余事象（complement）。**「第 n ラウンドまでに少なくとも 1 回」は、1−（「n ラウンドで一度も起こらない」）として考えるのがいちばん簡単であった。
- **最悪の場合を評価する。**未知の p は、不等式 $pq \leq \frac{1}{4}$ によって処理した。
- **対称関数（symmetric function）。**p と q について対称な式は、$p+q=1$ と pq を用いて表すと簡単になる。

背後にある数学

フォン・ノイマンのトリック

この問題の手順は速いが偏り（bias）がある。フォン・ノイマンの手順は公平である：2 回投げ、HT か TH なら決定し、HH か TT なら繰り返す。各試行が成功する確率は $2pq$ なので、平均して $\frac{1}{2pq}$ 回の試行を要し、硬貨が大きく偏っているときには時間がかかる（図 12.9.1）。この問題の手順は 1 ラウンドあたり確率 $p^2+q^2 \geq \frac{1}{2}$ で決定を下すが、$p > \frac{1}{2}$ のときにはその答えがホテルに偏る。

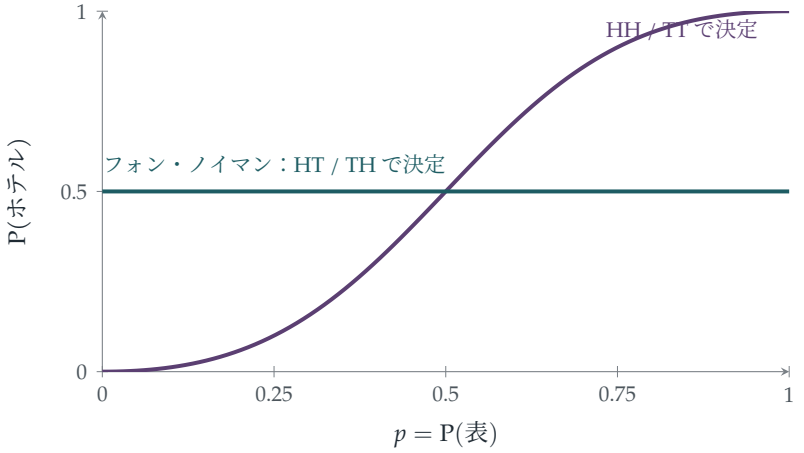


図 12.9.1　硬貨の偏りに対する、ホテルが選ばれる確率。硬貨の結果が一致したときに決定する方法（プラム色）は速いが、出やすい面に有利になる。フォン・ノイマンの手順（ティール色）はすべての p に対して厳密に公平であるが、その代償として、硬貨が大きく偏っているときには時間がかかる。

ランダムネスの抽出

フォン・ノイマンのトリックは、最初の**乱数抽出器**（randomness extractor）である。これは不完全なランダムネスを完全なランダムネスに変える手順である。この方法は無駄が多い（2 回投げるごとに得られる公平なビットは高々 1 つで、1 つも得られないことも多い）。より優れた方法では、捨てられた HH や TT の組を再利用して、硬貨がもつランダムネスのほぼすべてを抽出する。このランダムネスの量は、1 回の投げあたり $-p\log_2 p - q\log_2 q$ ビットという**エントロピー**（entropy）で測られる。抽出器は、コンピュータのハードウェア乱数生成器で用いられている。ハードウェア乱数生成器は、予測不可能だが偏りのある物理的なノイズを集め、それを暗号のために偏りのないビットに変換しなければならないのである。

フォン・ノイマンと乱数

フォン・ノイマンは、1951 年の論文 “Various techniques used in connection with random digits”（乱数に関連して用いられるさまざまな技法）でこのトリックを述べた。この論文が書かれたのは、彼と同僚たちが核兵器の設計のために、最初期のコンピュータ・シミュレーションのいくつか（「モンテカルロ法（Monte Carlo method）」）を実行していた時期である。同じ論文の中で、彼は算術によって乱数を生成することについて、有名な言葉を残している：「算術的な方法で乱数を作り出そうと考える者は、言うまでもなく罪の状態にある。」

思考のスキル

対称性を使って未知のものを消す

パラメータが未知であるときは、その値によらず確率が等しくなる事象を探し、それにもとづいて決定せよ。

HT と TH は同じ確率 pq をもつので、両者を比べれば p は完全に消える。大学では、同じ考え方が統計学の**枢軸量**（pivotal quantity）や実験計画（design of experiments）に現れる。実験計画では、無作為化（randomisation）と対称性によって未知の効果が打ち消される。

次の一歩

1 **ラウンド数の期待値。**小問 (i) の手順とフォン・ノイマンの手順のそれぞれについて、ラウンド数の期待値（expected value）を求めよ。どのような p に対して、フォン・ノイマンの手順のほうが遅くなるか。

2 **公平な三者択一。**偏りのある硬貨を用いて、3 つの選択肢のうち 1 つを、それぞれちょうど確率 $\frac{1}{3}$ で選ぶ手順を考案せよ。

3 **公平な硬貨から偏ったさいころを。**公平な硬貨を用いて、表が出る確率がちょうど $\frac{1}{3}$ である硬貨を模擬する手順を考案せよ。投げる回数の期待値はいくらか。

4 **再び 3 人の場合。**小問 (ii) において、3 人がホテルに行く確率を p の関数として求めよ。$p = \frac{1}{2}$ のとき、この手順は公平か。

さらに学ぶために

- **John von Neumann, “Various techniques used in connection with random digits”**, *National Bureau of Standards Applied Mathematics Series* 12 (1951). 短く、有名な論文である。
- **Donald Knuth, *The Art of Computer Programming*, Vol. 2** (Addison-Wesley)。第 3 章で乱数を扱っている。
- **Michael Mitzenmacher & Eli Upfal, *Probability and Computing*** (Cambridge University Press)。ランダムネスの抽出とエントロピーを扱っている。

12.10 ひもの端を結ぶ

問題

STEP III 2008、問 13

先に進む前に、まず問題を読んでおこう。問題は STEP Support Programme の問題データベースにある：

`step.maths.org/questions/08-s3-q13`

導入

100 本の紐を袋に入れ、端だけが外に出るようにしておく。無作為に端を 2 つ選んで結び、ほどけた端がなくなるまでこれを繰り返す。最後には輪がいくつか絡み合った状態になる。輪はいくつできるだろうか。100 本の紐からは、原理的には小さな輪が 100 個できることもあれば、巨大な輪が 1 つだけできることもある。驚くべきことに、輪の個数の期待値（expected value）はわずか 3.3 程度であり、紐が 4 万本あってもまだ 6 程度にすぎない。

鍵となる観察をひとつ押さえると、この問題はほとんど自明になる。何が起こったとしても、結び目を 1 つ作るたびに、箱の中の紐はちょうど 1 本減るのである。この過程はそれまでの履歴を「忘れて」しまい、輪の個数は、確率が変化していく独立なコイン投げの和になる。

取りかかる前に

考えてみよう 2 つの端を結んだあと、紐は何本残っているか（輪になったものはもう「紐」としては数えない）。それは輪ができたかどうかに依存するだろうか。

2 つの端が同じ紐のものであれば、その紐は輪になり、$n-1$ 本の紐が残る。異なる紐のものであれば、その 2 本はつながって 1 本の長い紐になり、やはり $n-1$ 本の紐が残る。いずれの場合も、それぞれ 2 つの端をもつ $n-1$ 本の紐が残るのである。

順を追って解く

段階 1　最初の結び目

端は $2n$ 個ある。そのうち 1 つを選ぶと、結ばれる相手は残りの $2n-1$ 個の端のどれもが等しい確率で選ばれ、そのうちちょうど 1 つが同じ紐の端である。したがって、最初の結び目が輪を作る確率は $\frac{1}{2n-1}$ であり、最初の段階で作られる輪の個数の期待値は $\frac{1}{2n-1}$ である。

段階 2　すべての結び目

> **考えてみよう**　上の観察から、2 回目の結び目を作る前の状況はどうなっているか。3 回目の前はどうか。紐が k 本残っているときに作る結び目で輪ができたかどうかを表す指示変数（indicator）を I_k とする。$\mathrm{E}(I_k)$ はそれぞれいくらか。

$k = n, n-1, \dots, 1$ に対して、各結び目を作る前には箱の中に k 本の紐があり、その結び目は確率 $\frac{1}{2k-1}$ で輪を作る（紐が 1 本だけ残っているときの最後の結び目は、必ず輪を作る）。輪の総数は $R = I_n + I_{n-1} + \cdots + I_1$ であるから、

$$\mathrm{E}(R) = \sum_{k=1}^{n} \frac{1}{2k-1} = 1 + \frac{1}{3} + \frac{1}{5} + \cdots + \frac{1}{2n-1}. \qquad \blacksquare$$

段階 3　分散

> **考えてみよう**　指示変数 I_k は独立か（ある段階で輪ができる確率は、それ以前に起こったことに依存するか）。

各段階で輪ができる確率は、それ以前に何が起こったかによらず $\frac{1}{2k-1}$ である。状態はつねに「k 本の紐」だからである。よって I_k は独立（independent）であり、分散（variance）は足し合わせることができる：

$$\mathrm{Var}(R) = \sum_{k=1}^{n} \frac{1}{2k-1}\left(1 - \frac{1}{2k-1}\right) = \sum_{k=1}^{n} \frac{2k-2}{(2k-1)^2}. \qquad \blacksquare$$

段階 4　4 万本の紐

> **考えてみよう**　$2n-1$ までの奇数の逆数の和を、調和数（harmonic number）を用いて表せ。

$2n$ までのすべての数の逆数の和から、偶数の逆数の和を引くと

$$\sum_{k=1}^{n} \frac{1}{2k-1} = H_{2n} - \tfrac{1}{2}H_n \approx \ln(2n) - \tfrac{1}{2}\ln n = \tfrac{1}{2}\ln(4n).$$

$n = 40\,000$ のとき $4n = 160\,000 = 20^4$ であるから、$\mathrm{E}(R) \approx \frac{1}{2} \cdot 4 \ln 20 \approx 6$ となる。 ■

（正確な値は 6.28 である。この差は、$H_n \approx \ln n + 0.577$ であることからわかるように、オイラー（Euler）定数（Euler's constant）に由来する。）

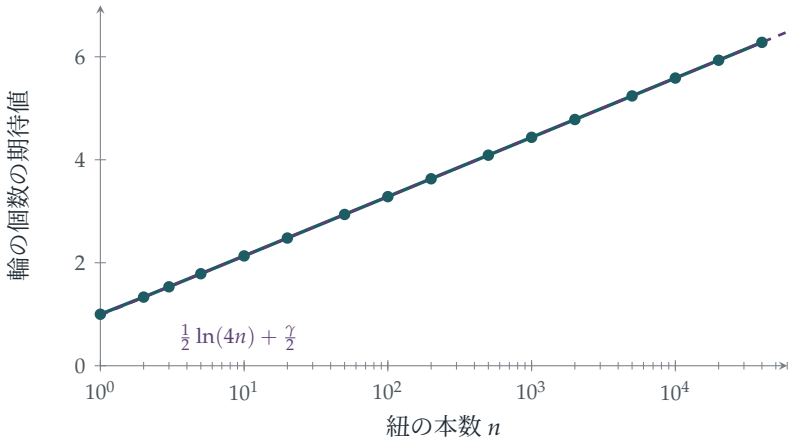


図 12.10.1　紐の本数に対する輪の個数の期待値（対数目盛（logarithmic scale））。正確な値（ティール）は直線 $\frac{1}{2}\ln(4n) + \frac{\gamma}{2}$（プラム、破線）の上にのっている。ここで $\gamma \approx 0.577$ はオイラー定数である。紐の本数を 10 倍にしても、輪は約 1.15 個しか増えない。

振り返り

- **過程が何を忘れるかを見つける。**各結び目のあと、箱の中は履歴によらず「k 本の紐」であった。
- **ここでも指示変数。**輪の個数は、結び目 1 つにつき 1 つの指示変数の和であった。
- **独立性から分散が得られる。**各段階の確率が過去に依存しないので、指示変数は独立であった。

背後にある数学

対数的な増加

輪の個数の期待値は $\frac{1}{2}\ln n$ のように増加し、分散は平均に近い（どちらも小さな確率の和である）。中心極限定理（central limit theorem）（Section 12.24）により、n が大き

いとき輪の個数は近似的に正規分布（normal distribution）に従い、平均と分散はいずれも約 $\frac{1}{2}\ln(4n)$ である。紐が 40 000 本あるとき、輪はたいてい 2 個から 11 個の間になる。

ランダムな置換の巡回

同じ対数は、n 個のものの無作為な置換（permutation）の巡回（cycle）の個数にも現れる。その平均は $H_n \approx \ln n$ である（紐の問題はこれの近い親戚であり、そちらは $\frac{1}{2}\ln n$ となる）。ランダムな置換の巡回は少ない。たいてい 1 つの巨大な巡回が要素の大部分を含み、最大の巡回に含まれる要素の割合の平均は約 0.62 である。これがゴロム–ディックマン定数（Golomb–Dickman constant）である。紐を無作為に結んでいく場合も同様にふるまい、紐の大部分は 1 つか 2 つの長い輪に入る。これらの事実は**ランダムな組合せ構造**（random combinatorial structures）の研究の一部であり、整数論とも関係がある（無作為に選んだ整数の素因数は、ランダムな置換の巡回のようにふるまう）。

ランダムな結び目と高分子

紐を結ぶこの問題は、物理系の簡単なモデルになっている。溶液中の長い高分子（polymer）は端どうしでつながって輪や鎖を作り、輪の形成の統計は似た規則に従う。結び目理論（knot theory）では、数学者はもっと難しい問い、すなわち無作為な閉じた輪が結ばれている（自明でない結び目になっている）確率はどれくらいか、という問いを考える。長いランダムな輪については、その答えは 1 に近づく。これは 1988 年に De Witt Sumners と Stuart Whittington によって証明された結果である。

思考のスキル

記憶をもたない構造を見つける

各段階のあとの系の状態が、それまでに行った段階の回数だけで決まるならば、各段階は独立にふるまい、期待値と分散は単純に足し合わせられる。

これはクーポン収集問題（coupon collector problem）（Section 9.11）の背後にある考え方であり、そこでは各段階が新たな幾何分布の待ち時間となる。この考え方はマルコフ連鎖（Markov chain）の理論にも再び現れる。そこでは現在の状態だけが意味をもつ。

次の一歩

1 **小さな場合。** $n = 2$ のとき、起こりうる場合をすべて書き出し、輪の個数の期待値が $\frac{4}{3}$ であることを確かめよ。

2 **1つの大きな輪。**過程が、すべての紐を含む1つの輪で終わる確率は $\prod_{k=2}^{n}\left(1 - \frac{1}{2k-1}\right)$ であり、n が大きいときこれは近似的に $\sqrt{\frac{\pi}{4n}}$ であることを示せ。（**ヒント：**この積を $\binom{2n}{n}$ と関連づけよ。）

3 **置換の巡回。** 置換を1要素ずつ組み立てていくこと（「中華料理店」構成（Chinese restaurant construction））により、n 個のものの無作為な置換の巡回の個数の期待値が H_n であることを示せ。

4 **分散。** $\mathrm{Var}(R) = \mathrm{E}(R) - \sum_{k=1}^{n} \frac{1}{(2k-1)^2}$ であること、および引かれている和が $\frac{\pi^2}{8}$ に収束することを示せ。

さらに学ぶために

- **Peter Winkler, *Mathematical Puzzles: A Connoisseur's Collection***（A K Peters）。紐を結ぶパズルが収録されている。
- **Richard Arratia, Andrew Barbour & Simon Tavaré, *Logarithmic Combinatorial Structures***（European Mathematical Society）。巡回、素数、ランダムな構造について。もう少し先で読むとよい。
- **Colin Adams, *The Knot Book***（American Mathematical Society）。結び目理論の入門書で、ランダムな結び目も扱っている。

12.11 切手、カップ、そして e

問題

STEP I 1996、問 13

先に進む前に、まず問題を読んでおこう。問題は STEP Support Programme の問題データベースにある：

`step.maths.org/questions/96-s1-q13`

導入

切手の一風変わった売り方である。買い手はどのカップの下に切手が隠れているかを当て続け、カップは 1 回ごとに 1 個ずつ増えていく。そして初めて外したとき、テーブルの上にあるカップ 1 個につき 1 ポンドを支払う。平均すると支払額は £e、すなわち約 £2.72 になる。

数 e は確率の中で不思議なほど頻繁に現れる。誰も自分のシャツを受け取らない確率（Section 9.8）、秘書問題（secretary problem、Section 12.12）、ポアソン分布（Poisson distribution）などである。この問題は e が現れる例の中でおそらく最も単純なものであり、しかも別の有名なパズルとまったく同じ分布をもつことがわかる。そのパズルとは、「0 と 1 の間の乱数をいくつ足せば合計が 1 を超えるか」というものである。平均すると、e 個である。

取りかかる前に

考えてみよう $N = 4$ となるには何が起こらなければならないか。その確率はいくらか。

ジムはカップが 2 個のとき（確率 $\frac{1}{2}$）と 3 個のとき（$\frac{1}{3}$）に正しく当て、カップが 4 個のときに外さなければならない（$\frac{3}{4}$）。よって $\mathrm{P}(N = 4) = \frac{1}{2} \cdot \frac{1}{3} \cdot \frac{3}{4} = \frac{3}{4!} = \frac{1}{8}$ である。

順を追って解く

段階 1　分布

$N = k$（ただし $k \geq 2$）となるのは、ジムがカップ $2, 3, \ldots, k-1$ 個のときに当て、k 個のときに外す場合である。

$$\mathrm{P}(N = k) = \frac{1}{2} \cdot \frac{1}{3} \cdots \frac{1}{k-1} \cdot \frac{k-1}{k} = \frac{k-1}{k!}.$$

（確認：$\frac{1}{2!} + \frac{2}{3!} + \frac{3}{4!} + \cdots = \sum_{k \geq 2} \left(\frac{1}{(k-1)!} - \frac{1}{k!} \right) = 1$ であり、これは隣り合う項が打ち消し合う和（telescoping sum）である。）

段階 2　期待値

考えてみよう　$k \cdot \frac{k-1}{k!}$ を簡単にせよ。どのような級数が得られるか。

$$\mathrm{E}(N) = \sum_{k \geq 2} \frac{k(k-1)}{k!} = \sum_{k \geq 2} \frac{1}{(k-2)!} = \frac{1}{0!} + \frac{1}{1!} + \cdots = e. \qquad \blacksquare$$

段階 3　分散

考えてみよう　$k^2 = k(k-1) + k$ と書くか、$\frac{k^2(k-1)}{k!}$ を直接簡単にすることで、$\mathrm{E}(N^2)$ を求めよ。

$$\mathrm{E}(N^2) = \sum_{k \geq 2} \frac{k^2(k-1)}{k!} = \sum_{k \geq 2} \frac{k}{(k-2)!} = \sum_{j \geq 0} \frac{j+2}{j!} = \sum_{j \geq 1} \frac{1}{(j-1)!} + 2 \sum_{j \geq 0} \frac{1}{j!} = e + 2e = 3e.$$

よって $\mathrm{Var}(N) = 3e - e^2 = e(3 - e) \approx 0.766$ である。 $\blacksquare$

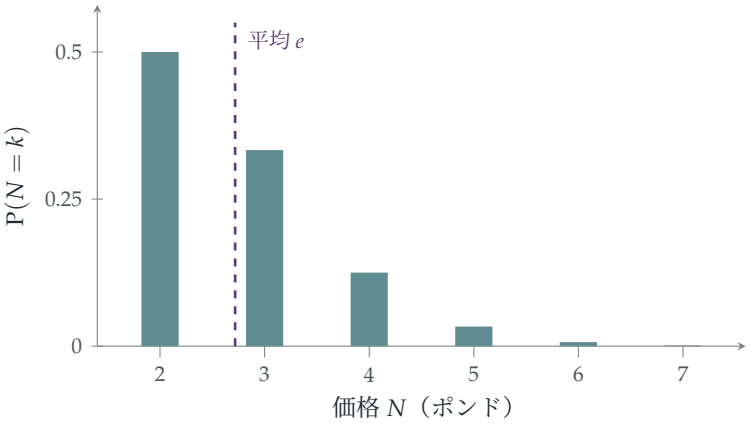


図 12.11.1　価格の分布（distribution）$\mathrm{P}(N=k)=\frac{k-1}{k!}$。ジムは半分の確率で £2 を支払う。長いが細い裾の部分が、平均を $£e \approx £2.72$ まで引き上げている。

振り返り

- **確率の積。**独立な成功が続いた後に失敗が 1 回起こる、という事象の確率は積で表され、それが $\frac{k-1}{k!}$ という簡単な形にまとまった。
- **階乗を約分する。**k や k^2 を掛けて約分することで、和が指数級数（exponential series）に変わった。
- **添字をずらして $\mathrm{E}(N^2)$ を求める。**$j=k-2$ と添字を付け替えることで、和が e 二つ分に分かれた。

背後にある数学

姿を変えた同じ分布

乱数 $U_1, U_2, \ldots$ を $[0,1]$ から一様に（uniformly）選び、それらの和が 1 を超えるのに必要な個数を M とする。最初の k 個の数の和が 1 未満となる確率は、領域 $\{u_1+\cdots+u_k<1, u_i\geq 0\}$ の体積に等しく、それは $\frac{1}{k!}$ である（面積 $\frac{1}{2}$ の三角形、体積 $\frac{1}{6}$ の四面体、以下同様）。したがって

$$\mathrm{P}(M>k)=\frac{1}{k!}, \qquad \mathrm{P}(M=k)=\frac{1}{(k-1)!}-\frac{1}{k!}=\frac{k-1}{k!},$$

となり、これは N の分布とまったく同じである。どちらの問題でも $\mathrm{P}(N>k)=\frac{1}{k!}$ が成り立つ。カップの問題では、これは $k-1$ 回続けて正しく当てる確率である。$\mathrm{E}(N)=\sum_{k\geq 0}\mathrm{P}(N>k)$（Section 9.9 の公式）を用いれば、1 行で $\mathrm{E}(N)=\sum\frac{1}{k!}=e$ が

得られる。

なぜ e は繰り返し現れるのか

確率に $\frac{1}{k!}$ が含まれるときはいつでも、それが k 個のものが特定の順序に並ぶことから来るのであれ、確率 $\frac{1}{2}, \frac{1}{3}, \ldots$ の k 個の事象が続けて起こることから来るのであれ、単体（simplex）の体積から来るのであれ、k についての和は指数級数を生み出す。完全順列（derangements、Section 9.8）では $\sum \frac{(-1)^k}{k!} = e^{-1}$ が得られた。ここでは $\sum \frac{1}{k!} = e$ である。そして、起こりにくい独立な事象が多数あることから生じる $\left(1 - \frac{1}{n}\right)^n \to e^{-1}$ という形の極限が、秘書問題やポアソン分布に e が現れる理由を説明している。

モンテカルロ法による e

「和が 1 を超える」問題を使えば、実験によって e を推定することができる。乱数を生成し、1 を超えるのに何個必要かを数え、これを何度も繰り返して平均をとればよい。これは π を推定するビュフォンの針（Buffon's needle、Section 12.20）の親戚である。このような**モンテカルロ法**（Monte Carlo methods）は、カジノの名にちなんで名付けられたもので、直接計算するのが難しい量を無作為抽出によって推定する。現在では素粒子物理学から金融に至るまで科学のあらゆる分野で使われており、その多くは多次元の積分の計算に用いられている。

思考のスキル

級数を見抜く

期待値が階乗を含む和で表されているときは、標準的な級数（指数級数、等比級数、二項級数）が現れるまで各項を簡単にせよ。

大学では、母関数（generating functions）がこの手法を体系化する。確率母関数（probability generating function）$\mathrm{E}(t^N)$ は分布全体をひとまとめにしたものであり、$t = 1$ におけるその導関数から平均と分散が得られる。

次の一歩

1 **母関数。** $\mathrm{E}(t^N) = 1 + (t-1)e^t$ を示し、それを用いて $\mathrm{E}(N)$ と $\mathrm{E}(N(N-1))$ を求めよ。

2 **裾の公式。**非負整数値をとる確率変数（random variable）について $\mathrm{E}(N) = \sum_{k\geq 0} \mathrm{P}(N > k)$ が成り立つことを証明し、それを用いて 1 行で $\mathrm{E}(N)$ を求めよ。

3 **カップ 1 個から始める。**ゲームがカップ 1 個（ジムが外すことはありえない）から始まるとすると、分布はどう変わるか。価格の期待値はいくらか。

4 **単体。**$\{u_1 + u_2 + u_3 < 1, u_i \geq 0\}$ の体積が $\frac{1}{6}$ であることを示し、数学的帰納法（induction）を用いて k 次元の場合の体積が $\frac{1}{k!}$ であることを示せ。

さらに学ぶために

- **Eli Maor, *e: The Story of a Number***（Princeton University Press）。e の歴史を、確率における登場例も含めて解説している。
- **Geoffrey Grimmett & Dominic Welsh, *Probability: An Introduction***（Oxford University Press）。母関数について。
- **Paul Nahin, *Digital Dice***（Princeton University Press）。モンテカルロ・シミュレーションによって解かれる確率パズル集。

12.12 いつ探すのをやめるか

問題

STEP I 2001、問 14

先に進む前に、まず問題を読んでおこう。問題は STEP Support Programme の問題データベースにある：

`step.maths.org/questions/01-s1-q14`

導入

あなたは候補者を 1 人ずつ面接し、その場で決断しなければならない。この人を採用して終わりにするか、永久に不採用として次の人に会うかである。あなたが望むのは最も優れた人物である。早く採用しすぎれば、より優れた人を逃す可能性が高い。待ちすぎれば、最良の人はすでに去ってしまっている可能性が高い。どうすればよいだろうか。

確率論で最も有名な結果の一つである答えは、次のとおりである。最初の 37% の候補者を見送り、その後、見送った全員より優れた最初の候補者を採用せよ。この方法は約 37% の確率で成功し、37% とは $\frac{1}{e}$ のことである。この問題では基本となる確率を計算し、姉妹問題である STEP II 2006 の問題（演習に収録）では、この戦略の最も単純な形、すなわち最初の 1 人だけを見送る場合を分析する。

取りかかる前に

考えてみよう 無作為な順序で面接するとき、全体で最良の候補者が最初の n 人の中にいる確率はいくらか。

最良の候補者は N 個の位置のどこにも等しい確率でいるので、その確率は $\frac{n}{N}$ である。

順を追って解く

段階 1　最初の n 人の中の最良が全体の最良である

最初の n 人の中の最良が全体の最良であるのは、全体の最良が最初の n 人の中にいるときであり、かつそのときに限る。よって確率は $\frac{n}{N}$ である。 ■

段階 2　最良または 2 番目

考えてみよう　最初の n 人の中の最良が全体の 2 番目であるのは、2 番目の候補者が最初の n 人の中にいて、最良の候補者がいないときであり、かつそのときに限る。その確率はいくらか。

2 番目の候補者は確率 $\frac{n}{N}$ で最初の n 人の中にいる。このとき、最良の候補者は残りの $N-1$ 個の位置のいずれかにあり、そのうち $N-n$ 個が最初の n 人より後ろである。よって

$$\mathrm{P}(\text{最良または 2 番目}) = \frac{n}{N} + \frac{n}{N}\cdot\frac{N-n}{N-1}. \qquad \blacksquare$$

$N=4, n=2$ のとき、$\frac{1}{2}+\frac{1}{2}\cdot\frac{2}{3}=\frac{5}{6}$ である。上位 2 人の候補者が 4 つの位置のどこに入るかを列挙すると（順序を区別した 12 通りの配置はすべて同様に確からしい）、最初の 2 人の中の最良が上位 2 人のいずれかになるのは、上位 2 人がともに位置 3 と 4 に来る 2 通りを除くすべての場合である。よって $\frac{10}{12}=\frac{5}{6}$ となる。

段階 3　最適戦略

考えてみよう　戦略：最初の r 人の候補者を見送り、その後、見送った全員より優れた最初の候補者を採用する。最良の候補者が位置 $k>r$ にいるとき、この戦略がその人を選ぶのはどのような場合か。

この戦略が位置 k にいる最良の候補者を選ぶのは、位置 $r+1,\dots,k-1$ のどの候補者も最初の r 人全員を上回らないとき、すなわち最初の $k-1$ 人の中の最良が最初の r 人の中にいるときであり、かつそのときに限る。段階 1（N を $k-1$ に置き換えたもの）により、この確率は $\frac{r}{k-1}$ である。各位置 k の確率は $\frac{1}{N}$ であるから

$$\mathrm{P}(\text{成功}) = \frac{1}{N}\sum_{k=r+1}^{N}\frac{r}{k-1} = \frac{r}{N}\Bigl(\frac{1}{r}+\frac{1}{r+1}+\cdots+\frac{1}{N-1}\Bigr).$$

考えてみよう　N が大きいとき、$x=\frac{r}{N}$ とおく。和はおよそ $\ln\frac{N}{r}$ である。$x\ln\frac{1}{x}$ を最大にする x の値は何か。

成功確率はおよそ $x \ln \frac{1}{x}$ である。微分すると、$x = \frac{1}{e}$ のとき $\ln \frac{1}{x} - 1 = 0$ となり、最大値は $\frac{1}{e} \approx 0.368$ である（図 12.12.1）。

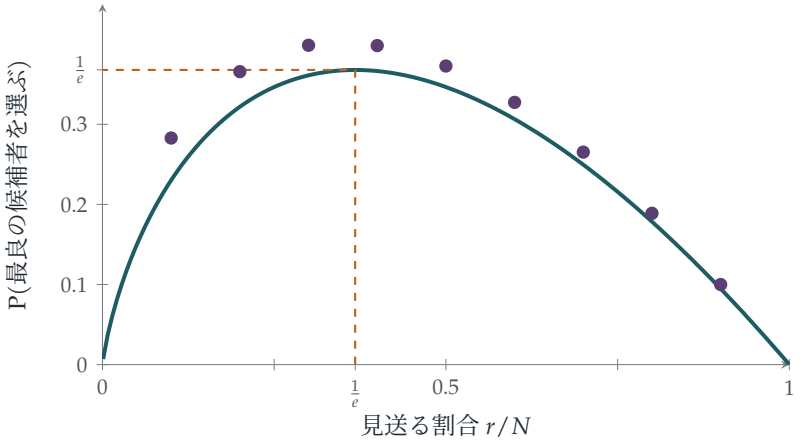


図 12.12.1　最初の割合 r/N を見送るときに最良の候補者を選ぶ確率。極限の曲線 $x \ln \frac{1}{x}$（青緑）と、$N = 10$ のときの正確な値（紫）。最良の戦略は候補者の約 $\frac{1}{e} \approx 37\%$ を見送るもので、N がどれほど大きくても約 $\frac{1}{e}$ の確率で成功する。

振り返り

- **位置の対称性（symmetry）。**無作為な順序では、各候補者はどの位置にも等しい確率でいる。このことから直ちに $\frac{n}{N}$ が得られた。
- **最良の候補者の位置で条件付ける（condition）。**戦略の成功確率は、最良の候補者の位置で条件付けることによって計算した。
- **和は積分になる。** N が大きいとき、調和和（harmonic sum）は対数となり、最適化によって $\frac{1}{e}$ が得られた。

背後にある数学

最適停止

秘書問題（secretary problem）は**最適停止**（optimal stopping）の原点となる例である。最適停止とは、時間とともに届く情報にもとづいて、期待される報酬を最大にするためにいつ行動を起こすかを決める問題である。最後の候補者から逆向きにたどる議論によって、（「r 人を見送り、その後それまでで最良の人を採る」という戦略に限らず）**あらゆる**戦略の中で、この種の閾値規則（threshold rule）が最適であることが証明で

きる。同じ数学が、株をいつ売るか、金融オプションをいつ行使するか、駐車スペースをどれだけ長く探すかといった問題を支配している。注目すべき特徴は、極限では答えが N によらないことである。候補者が百人でも百万人でも、成功確率は $\frac{1}{e}$ である。

変種

上位 2 人のどちらでも満足するなら、最良の戦略は約 57% の確率で成功する。最良の人を選ぶ確率ではなく、選んだ候補者の順位の期待値（expected rank）を最適にしたいなら、最適戦略はより複雑になり、順位の期待値は約 3.87 に近づく（Chow, Moriguti, Robbins and Samuels, 1964）。見送った候補者をある確率で呼び戻せる場合や、候補者の得点の分布がわかっている場合（「完全情報」問題）には、答えはまた変わる。演習にあるアイスクリームの問題は最も単純な場合であり、1 つだけ見送り、その後で最初に上回ったものを採る。

誰が解いたのか？

この問題は 1950 年代に数学者の間で広まり、1960 年にマーティン・ガードナー（Martin Gardner）が「グーゴル・ゲーム」として広く紹介した（紙片に数を書く。相手はそれを 1 枚ずつ裏返し、最大の数で止まらなければならない）。Thomas Ferguson の論文 “Who solved the secretary problem?”（秘書問題を解いたのは誰か、1989 年）は、その入り組んだ歴史をたどっている。一説によれば、天文学者ケプラー（Johannes Kepler）は 1611 年に再婚相手を選ぶにあたって 11 人の候補と面会し、結局 5 人目に求婚したという。

思考のスキル

戦略をパラメータで表して最適化する

最良の戦略がわからないときは、パラメータをもつ自然な戦略の族を考え、成功確率をそのパラメータの関数として計算し、最適化せよ。

大学では、この手法は決定理論、ゲーム理論、統計学で用いられる。族の中で最良の戦略が全体でも最良であることを示すのは、より難しく、より興味深い段階であり、多くの場合後ろ向き帰納法（backward induction）によって行われる。

次の一歩

1 **アイスクリーム。** STEP II 2006、問 13（`step.maths.org/questions/06-s2-q13`）に挑戦せよ。単純な規則による秘書問題の小型版である。最初の 1 つを見送り、その後それを上回る最初のものを採る。

2 **小さい N。** $N = 4$ のとき、最良の r の値と成功確率を求めよ。列挙して確かめよ。

3 **最良または 2 番目。** 小問 (ii) を用いて、「最初の n 人の中の最良が上位 2 人のいずれかである」ことは「最初の n 人の中の最良が全体の最良である」ことよりはるかに起こりやすいことを論ぜよ。このことは、上位 2 人のいずれかを得るための戦略について何を示唆するか。

4 **正確な最大値。** $N = 10$ のとき、図から $r = 3$ が最適であることを確かめよ。$N = 100$ のとき、$r = 37$ が最適であることを示せ（小数第 4 位までで、$r = 36, 37, 38$ に対する成功確率は 0.3710, 0.3710, 0.3708 である）。

さらに学ぶために

- **Thomas Ferguson, “Who solved the secretary problem?”**, *Statistical Science* 4 (1989). 歴史と変種。オンラインで無料で読める。
- **Brian Christian & Tom Griffiths, *Algorithms to Live By*** (William Collins). 最適停止をはじめとするアルゴリズムを日常の意思決定に応用する。
- **Martin Gardner, *New Mathematical Diversions*** (Mathematical Association of America). グーゴル・ゲームを収録。

12.13 三つ巴の決闘

問題

STEP II 1995、問 12

先に進む前に、まず問題を読んでおこう。問題は STEP Support Programme の問題データベースにある：

`step.maths.org/questions/95-s2-q12`

導入

3 人がロールパンを投げ合って決闘をする。カスバートは決して外さず、バーティーは 5 回に 3 回命中させ、いちばん下手なアルジャーノンは 5 回に 2 回しか当てられない。最初に投げるのはアルジャーノンである。彼は誰を狙うべきだろうか。意外な答えは「誰も狙わない」である。彼はわざとロールパンを大きく外して投げるべきなのである。

理由は戦略（strategy）にある。3 人全員が残っている間、危険な投げ手であるバーティーとカスバートは互いを狙い合う。もしアルジャーノンがどちらか一方を倒せば、生き残ったほうが彼に向かってくる。わざと外すことで、彼は 2 人に決着をつけさせ、そのうえで生き残ったほうに対して先に投げる権利を得る。実際、最善の作戦をとれば、いちばん下手な投げ手が**最も**勝つ可能性の高い者になる。この「三者決闘（truel）」は、弱さが有利に働きうることを示す、ゲーム理論（game theory）でよく知られた例である。

取りかかる前に

考えてみよう 2 人の決闘で、パーシーが先に投げる。彼がすぐに命中させるか、あるいは 2 人とも外して状況がまったく同じに戻るかのどちらかである。パーシーが勝つ確率（probability）についての方程式を立てよ。

パーシーが勝つ確率を P とする。彼はすぐに確率 p で命中させる。そうでない場合（確率 $1-p$）はロドニーが投げ、確率 $1-r$ で外して最初の状態に戻る。したがって

$P = p + (1-p)(1-r)P$ である。

順を追って解く

段階 1　2 人の決闘

これを解くと $P(1-(1-p)(1-r)) = p$ であり、$1-(1-p)(1-r) = p+r-pr$ であるから、

$$P = \frac{p}{p+r-pr}. \qquad \blacksquare$$

アルジャーノン（$p = \frac{2}{5}$）とバーティー（$r = \frac{3}{5}$）の決闘でアルジャーノンが先に投げる場合、これは $\frac{2/5}{2/5+3/5-6/25} = \frac{10}{19}$ となる。バーティーが先に投げる場合、アルジャーノンが勝つ確率は $\left(1-\frac{3}{5}\right)\cdot\frac{10}{19} = \frac{4}{19}$ である。

段階 2　アルジャーノンの 3 つの選択肢

考えてみよう　それぞれの場合に、誰が残り、次に誰が投げるか。

(i) **アルジャーノンがカスバートに命中させる場合。** 次にバーティーがアルジャーノンに向かって投げる。アルジャーノンはバーティーが先に投げる決闘に勝たねばならず、その確率は $\frac{4}{19}$ である。

(ii) **アルジャーノンがバーティーに命中させる場合。** 次にカスバートがアルジャーノンに向かって投げ、決して外さない。アルジャーノンが勝つ見込みは 0 である。

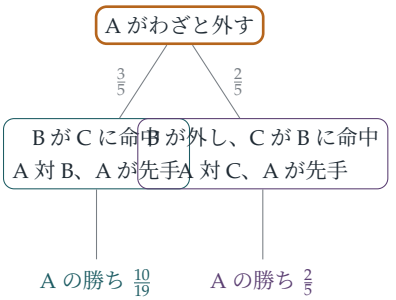
(iii) **アルジャーノンが外す場合。** バーティーがカスバートに向かって投げる。確率 $\frac{3}{5}$ で命中し、アルジャーノンとバーティーが残ってアルジャーノンが投げる番になる。このときアルジャーノンが勝つ確率は $\frac{10}{19}$ である。確率 $\frac{2}{5}$ でバーティーは外し、カスバートがバーティーに命中させ、アルジャーノンは 1 回だけ投げてカスバートと対決する。彼が勝つのは命中させたときだけで、その確率は $\frac{2}{5}$ である。合わせて、

$$\frac{3}{5}\cdot\frac{10}{19} + \frac{2}{5}\cdot\frac{2}{5} = \frac{6}{19} + \frac{4}{25} = \frac{150+76}{475} = \frac{226}{475} \approx 0.476.$$

段階 3　助言

考えてみよう　アルジャーノンがカスバートを狙うと、命中する（確率 $\frac{2}{5}$）こともあれば外れることもある。全体として彼が勝つ確率はいくらか。わざと外す場合と比較せよ。

カスバートを狙うと $\frac{2}{5}\cdot\frac{4}{19}+\frac{3}{5}\cdot\frac{226}{475}=\frac{200+678}{2375}=\frac{878}{2375}\approx 0.370$ となる。バーティーを狙うのはさらに悪い。わざと外すと $\frac{226}{475}\approx 0.476$ となり、これが最善である。したがってアルジャーノンは大きく外して投げるべきであり、そのとき彼が勝つ確率は $\frac{226}{475}$ である。 ■



この作戦のとき：
A の勝ち $\frac{226}{475}\approx 0.48$
B の勝ち $\frac{135}{475}\approx 0.28$
C の勝ち $\frac{114}{475}\approx 0.24$

図 12.13.1　アルジャーノンの最善の作戦。わざと外すことで、彼はより強い 2 人の相手に先に戦わせ、生き残ったほうに対して必ず先に投げられる。いちばん下手な投げ手が最も勝つ可能性が高くなり、完璧な投げ手が最も勝つ可能性が低くなる。

振り返り

- **繰り返される状況からは方程式が得られる。**決闘は 2 回外れると「最初からやり直し」になるので、勝つ確率についての 1 次方程式（linear equation）が得られる。
- **相手がどう応じるかを考える。**アルジャーノンの選択肢はどれも、バーティーとカスバートが次にとる行動を変える。
- **何もしないことも含めて、すべての選択肢を比較する。**最善の手は、いちばん悪そうに見えた手であった。

背後にある数学

戦略とゲーム理論

三者決闘は**ゲーム理論**の中心的な考え方を示している。すなわち、最善の行動は他のプレイヤーがどう応じるかに依存する。最も危険な相手であるカスバートに命中させるのは自然な手に見えるが、それはアルジャーノンに不利な形でゲームを変えてしまう。この問題ではバーティーとカスバートの戦略は固定されている（2 人は互いを狙う）。完全なゲーム理論的分析では、各プレイヤーは他のプレイヤーも同じように考えていることを承知のうえで、自分の戦略を選ぶことになる。このような分析

は、1944 年のフォン・ノイマン（John von Neumann）とモルゲンシュテルン（Oskar Morgenstern）の研究、そして 1950 年のナッシュ（John Nash）の研究に始まり、今では経済学、進化生物学、政治学の基礎となっている。

弱点としての強さ

三者決闘に勝つのは誰か。アルジャーノンが最善の作戦をとるとき、勝つ確率はアルジャーノンが $\frac{226}{475}$、バーティーが $\frac{135}{475}$、カスバートが $\frac{114}{475}$ である。最も腕のよい投げ手は、選択権をもつ者全員から狙われるため、生き残る可能性が最も低い。同じパターンは、複数の候補者がいる選挙、企業間の競争、そして最強の競争者が最初の標的になりうる自然界にも現れる。1950 年代に三者決闘をゲーム理論に導入した Martin Shubik は、これを用いて「最も適した者が必ず生き残るのか」という問いを立てた。

思考のスキル

一手先を読む

行動を選ぶ前に、それが他の全員にとって状況をどう変えるか、そして彼らがそれにどう応じるかを考えよ。

大学では、これはゲーム理論における後ろ向き帰納法（backward induction）になる。まずゲームの終盤を分析し、そこから遡って最善の初手を決めるのである。

次の一歩

1. **バーティーとカスバート。** アルジャーノンがわざと外すとき、バーティーが勝つ確率は $\frac{135}{475}$、カスバートが勝つ確率は $\frac{114}{475}$ であることを確かめよ。
2. **より腕のよい投げ手。** アルジャーノンが命中させる確率が $\frac{2}{5}$ ではなく a であるとき、わざと外すことが依然として最善の初手となるのは a がどのような値のときか。
3. **バーティーはカスバートを狙うべきか。** アルジャーノンがわざと外すとき、バーティーにとってアルジャーノンではなくカスバートを狙うほうが得策か。（カスバートの行動は変わらないものとする。）

4 **投げる回数が決まっている決闘。**各プレイヤーが 2 回しか投げられない決闘において、パーシーが勝つ確率を求め、回数に制限のない決闘と比較せよ。

さらに学ぶために

- **D. Marc Kilgour & Steven Brams, "The truel"**, *Mathematics Magazine* 70 (1997). 三者決闘についての読みやすい概説。
- **Avinash Dixit & Barry Nalebuff, *Thinking Strategically*** (Norton). 一般読者向けのゲーム理論の本で、三者決闘も扱っている。
- **Ken Binmore, *Game Theory: A Very Short Introduction*** (Oxford University Press). ゲーム理論の簡潔な入門書。

12.14 混合戦略

問題

STEP I 1989、問 16

先に進む前に、まず問題を読んでおこう。問題は STEP Support Programme の問題データベースにある：

`step.maths.org/questions/89-s1-q16`

導入

じゃんけん（rock–paper–scissors）では、決まった手を出し続けると必ずつけ込まれる。毎回グーを出していれば、相手はやがてパーを出すことを覚えるだろう。安全な戦略はただ一つ、手をランダムに選ぶことである。この問題はその込み入った版であり、利得が均等ではなく、どのようにランダムに選べばよいかは自明ではない。実は、1, 2, 3 をそれぞれ確率 $\frac{1}{4}, \frac{1}{2}, \frac{1}{4}$ で言うことにすれば、相手がどうしようとも、長い目で見て負けないことを保証できるのである。

これは、1928 年のジョン・フォン・ノイマン（John von Neumann）による**ミニマックス定理**（minimax theorem）を初めて垣間見る例である。この定理はゲーム理論（game theory）の礎となった結果であり、一方の利得が他方の損失となるあらゆる二人ゲームにおいて、各プレイヤーには、完璧にプレイする相手に対して可能な限り最良の結果を保証するランダム化された戦略が存在する、と主張する。

取りかかる前に

考えてみよう 二人の言う数の各組について、B の獲得額の表を書いてみよ（B の数を縦に、A の数を横に並べる）。

規則を一つずつ当てはめていくと（たとえば B が 1、A が 3 と言えば、差は 2 であり、小さい方の B が £2 を支払う）、次の表が得られる。

B \ A	1	2	3
1	0	+1	−2
2	−1	0	+1
3	+2	−1	0

この表は**交代的**（antisymmetric）である（プレイヤーを入れ替えると符号が変わる）。したがって、どちらのプレイヤーにも生まれつきの有利さはないという意味で、このゲームは公平である。

順を追って解く

段階 1　決まった数に対して

A がいつも 1 と言うなら、B は 3 と言えば毎回 £2 を得る。A がいつも 2 と言うなら、B は 1 と言って £1 を得る。A がいつも 3 と言うなら、B は 2 と言って £1 を得る。どんな決まった数もつけ込まれてしまう。

段階 2　一様にランダムな選択に対して

考えてみよう　A が一様にランダムに数を選ぶとき、B が j と言った場合の B の期待利得（expected winnings）はいくらか。（表の第 j 行の平均をとればよい。）

各行の平均は、$j=1$ のとき $\frac{1}{3}(0+1-2)=-\frac{1}{3}$、$j=2$ のとき 0、$j=3$ のとき $\frac{1}{3}(2-1+0)=\frac{1}{3}$ である。よって、1 回あたりの B の期待利得は

$$-\tfrac{1}{3}b_1+0\cdot b_2+\tfrac{1}{3}b_3,$$

であり、これは $b_3=1$ のとき最大となる。すなわち、いつも 3 と言えば、1 回あたり平均で $\frac{1}{3}$ ポンドを得る。一様なランダム化では A は身を守れない。

段階 3　一般の場合

考えてみよう　今度は A が確率 a_j で j を言うとする。B の期待利得を A の数ごとにまとめよ。a_1 の係数（第 1 列の重み b_j による平均）は何か。

表を列に沿って読むと、

$$\text{B の期待利得}=(-b_2+2b_3)a_1+(b_1-b_3)a_2+(-2b_1+b_2)a_3,$$

となるので、$X=-b_2+2b_3$, $Y=b_1-b_3$, $Z=-2b_1+b_2$ である。

考えてみよう 三つすべてを 0 にできるか。（方程式は三つあるが、独立ではない。なぜか。）

これらは独立ではない。恒等的に $X + 2Y + Z = 0$ が成り立つからである（確かめてみよ）。したがって、$Y = 0$ から $b_1 = b_3$、$X = 0$ から $b_2 = 2b_3$ が得られれば、自動的に $Z = 0$ も成り立つ。$b_1 + b_2 + b_3 = 1$ と合わせて

$$b_1 = \tfrac{1}{4}, \qquad b_2 = \tfrac{1}{2}, \qquad b_3 = \tfrac{1}{4},$$

を得て、$X = Y = Z = 0$ となる。このとき、a_j が何であっても B の期待利得は 0 であり、B は長い目で見て何も失わない。 ■

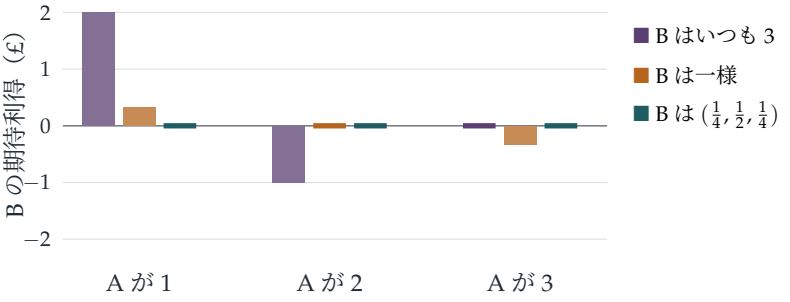


図 12.14.1 B の三つの戦略それぞれについて、A の各選択に対する B の期待利得。いつも 3 と言う戦略は A の 1 には勝つが、A の 2 には負ける。一様な選択は A の 3 に負ける。戦略 $(\frac{1}{4}, \frac{1}{2}, \frac{1}{4})$ は A が何をしても期待利得がちょうど 0（細い棒で表示）となるので、決してつけ込まれることがない。

振り返り

- **利得表（payoff table）を書く。** 推論はすべて 3×3 行列の行と列についてのものであった。
- **純粋戦略（pure strategy）はつけ込まれる。** どんな決まった数に対しても、さらには一様な選択に対してさえ、相手には得をする応じ方がある。
- **均等にする。** 安全な戦略とは、相手にとってどの選択も無差別になるようにするものである。すなわち、A のどの選択も同じ期待値を与える。

背後にある数学

ミニマックス定理

利得行列（payoff matrix）M で与えられる二人**ゼロ和**（zero-sum）ゲームにおいて、B が確率 $\mathbf{b}$ を、A が確率 $\mathbf{a}$ を選ぶとき、B の期待利得は $\mathbf{b}^{\mathsf{T}} M \mathbf{a}$ である。フォン・ノイマンのミニマックス定理は

$$\max_{\mathbf{b}} \min_{\mathbf{a}} \mathbf{b}^{\mathsf{T}} M \mathbf{a} = \min_{\mathbf{a}} \max_{\mathbf{b}} \mathbf{b}^{\mathsf{T}} M \mathbf{a}.$$

が成り立つことを主張する。言葉で言えば、B がランダム化された戦略を前もって宣言することで保証できる額は、A が自分の戦略を宣言することで B を抑え込める額に等しい。この共通の値をゲームの**値**（value）という。ここでは交代性から値は 0 であり、戦略 $\left(\frac{1}{4}, \frac{1}{2}, \frac{1}{4}\right)$ は両プレイヤーにとって最適である。この定理は線形代数と凸性（convexity）を用いて証明される（線形計画法（linear programming）の双対定理（duality theorem）と同値である）。また、その多人数への一般化である**ナッシュ均衡**（Nash equilibrium）により、ジョン・ナッシュ（John Nash）はノーベル経済学賞を受賞した。

なぜランダムにするのか

ここでのランダム性は運の問題ではなく、情報の問題である。B の選択が予測できれば、A はそれにつけ込める。適切な確率でランダムに選ぶことで、B は自分の選択を予測不能にするとともに、どんな予測も役に立たないようにするのである。実例としては、サッカーのペナルティーキックがある。プロ選手の研究では、キッカーもゴールキーパーも左右をランダムに選んでおり、その比率はゲーム理論上の最適値に近いことがわかっている。テニスのサーブも同様の例である。

フォン・ノイマンの 1928 年の論文

ジョン・フォン・ノイマンは 1928 年、論文「On the theory of games of strategy」（戦略ゲームの理論について）においてミニマックス定理を証明した。動機となったのはポーカーであり、そこではブラフがランダム化の一形態となっている。1944 年、彼は経済学者オスカー・モルゲンシュテルン（Oskar Morgenstern）とともに *Theory of Games and Economic Behavior*（『ゲームの理論と経済行動』）を出版し、ゲーム理論を一つの学問分野として確立した。フォン・ノイマンは後年、ミニマックス定理こそが鍵となる結果であり、それなしには出版に値するゲーム理論は存在しえなかったと考えていた、と回想している。

思考のスキル

ランダムに選んで相手の有利さを消す

どんな決まった選択もつけ込まれるときは、相手にとってどの選択肢も無差別になるような確率で、ランダムに選べ。

大学では、ゲーム理論、線形計画法、アルゴリズムの理論のいずれもがこの考え方を用いる。たとえば乱択アルゴリズム（randomised algorithm）は、どんな敵対者も予測できないランダムな選択を行うことで、最悪の入力から身を守っている。

次の一歩

1 **A の最適な応じ方。** B が $(\frac{1}{4}, \frac{1}{2}, \frac{1}{4})$ を用いるとき、A のどの戦略も期待利得 0 を与えることを示せ。また、B が代わりに $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ を用いるとき、A の最適な応じ方を求めよ。

2 **じゃんけん。** じゃんけんの利得行列を書き、最適戦略を求めよ。ここでは一様になるのに、問題ではそうならないのはなぜか。

3 **一意性。** 負けないことを保証する B の戦略は $(\frac{1}{4}, \frac{1}{2}, \frac{1}{4})$ **ただ一つ**であることを示せ。（ヒント：その戦略は X, Y, Z をすべて非負にしなければならない。$X + 2Y + Z = 0$ を用いよ。）

4 **修正されたゲーム。** 差が 2 のときの支払いが £2 ではなく £3 であるとき、最適戦略を求めよ。

さらに学ぶために

- **Ken Binmore, *Playing for Real*** (Oxford University Press)。ゲーム理論への、本格的でありながら親しみやすい入門書。
- **Avinash Dixit & Barry Nalebuff, *The Art of Strategy*** (Norton)。スポーツやビジネスの例を用いたゲーム理論の本。
- **William Poundstone, *Prisoner's Dilemma*** (Anchor)。フォン・ノイマン、ゲーム理論、そして冷戦。

12.15 絶滅か爆発か

問題

STEP II 2003、問 12

先に進む前に、まず問題を読んでおこう。問題は STEP Support Programme の問題データベースにある：

`step.maths.org/questions/03-s2-q12`

導入

1873 年、ゴルトン（Galton）は *Educational Times* 誌上で一つの問いを投げかけた。貴族の家名はなぜこれほど頻繁に途絶えてしまうのか。男性はそれぞれ無作為な人数の息子をもち、その息子たちがそれぞれ家名を受け継ぐ。たとえ一人あたりの息子の数の平均が 1 人を超えていたとしても、家名は永遠に存続するのだろうか。この問いから**分枝過程**（branching process）の理論が生まれた。この理論は現在、核の連鎖反応、感染症の流行、細胞集団の増殖、突然変異の存続などのモデルとして用いられている。

本問は最も単純なモデルを設定している。各粒子は 0 個、1 個、または 2 個の粒子に分裂する。集団の大きさの期待値（expectation）は等比数列（geometric sequence）的に増加または減少する。しかし、期待値がすべてを語るわけではない。期待値が爆発的に増大する場合でさえ、系統が完全に途絶える確率は正であり、それは美しい不動点（fixed point）の議論によって求められる。

取りかかる前に

考えてみよう ある段階で粒子が 2 個あるとき、次の段階の粒子の個数はどのような分布に従うか。

2 個の粒子はそれぞれ独立に $\mathrm{B}(2, p)$ に従う個数の子を生むので、合計は $\mathrm{B}(4, p)$ に従う。すなわち、それぞれ確率 p で成功する独立な 4 回の「機会」である。

順を追って解く

段階 1　2 個、そして 2 個

$\mathrm{P}(X_1 = 2) = p^2$ であり、さらに $\mathrm{P}(X_2 = 2 \mid X_1 = 2) = \binom{4}{2} p^2 q^2 = 6p^2q^2$ である。よって

$$\mathrm{P}(X_1 = 2, X_2 = 2) = 6p^4q^2. \qquad \blacksquare$$

段階 2　さかのぼって考える

> **考えてみよう**　$X_2 = 2$ となるのはどのような場合か。（X_1 はどのような値をとりうるか。）

$X_2 = 2$ となるのは、$X_1 = 1$（確率 $2pq$）の後に子が 2 個生まれる（確率 p^2）場合か、$X_1 = 2$ の後に $\mathrm{B}(4, p)$ から 2 が出る場合である。（$X_1 = 0$ のときは何も残らない。）よって $\mathrm{P}(X_2 = 2) = 2p^3q + 6p^4q^2$ であり、

$$\mathrm{P}(X_1 = 2 \mid X_2 = 2) = \frac{6p^4q^2}{2p^3q + 6p^4q^2} = \frac{3pq}{1 + 3pq}.$$

これを $\frac{9}{25}$ とおくと $75pq = 9 + 27pq$ となるので $pq = \frac{3}{16}$、すなわち $p(1-p) = \frac{3}{16}$ であり、$p = \frac{1}{4}$ または $p = \frac{3}{4}$ を得る。 $\blacksquare$

段階 3　集団の大きさの期待値

> **考えてみよう**　第 $n-1$ 段階で粒子が k 個あるとき、第 n 段階の個数の期待値はいくらか。そのうえで k について平均をとれ。

各粒子は平均 $\mathrm{E}(X) = 2p$ 個の子をもつので、k 個の粒子の子の数の平均は $2pk$ である。すなわち $\mathrm{E}(X_n \mid X_{n-1} = k) = 2pk$。$k$ について平均をとると $\mathrm{E}(X_n) = 2p\,\mathrm{E}(X_{n-1})$ となる。$\mathrm{E}(X_0) = 1$ より

$$\mathrm{E}(X_n) = (2p)^n.$$

$p = \frac{1}{4}$ のとき $(2p)^n = 2^{-n} \to 0$、$p = \frac{3}{4}$ のとき $(2p)^n = \left(\frac{3}{2}\right)^n \to \infty$ である。 $\blacksquare$

振り返り

- **独立な二項分布の和。** それぞれ $\mathrm{B}(2, p)$ に従う子を生む k 個の粒子からは、合計で $\mathrm{B}(2k, p)$ に従う個数の子が生まれる。二項分布（binomial distribution）の再生性である。

- **時間をさかのぼるベイズの定理。** $\mathrm{P}(X_1 = 2 \mid X_2 = 2)$ は、後の観測から前の状態へとさかのぼって推論するベイズの定理（Bayes' theorem）の応用である。
- **条件付き期待値。** 前の世代で条件づけることにより、期待値が等比数列になった。条件付き期待値（conditional expectation）の考え方である。

背後にある数学

絶滅確率

$p = \frac{3}{4}$ のとき、集団の大きさの期待値は限りなく増大する。では、集団は必ず存続するのだろうか。そうとは限らない。最初の粒子が子をもたなければ（確率 $q^2 = \frac{1}{16}$）、系統はただちに途絶える。最終的に絶滅する確率を e とする。最初の粒子は k 個の子をもち、それぞれの子の系統は独立に確率 e で途絶えるので、

$$e = q^2 + 2pq\,e + p^2e^2 = (q + pe)^2 = \mathrm{G}(e),$$

となる。ここで $\mathrm{G}(s) = (q + ps)^2$ は子の数の**確率母関数**（probability generating function）である。絶滅確率（extinction probability）は G の不動点であり、実際にはそのうち**最小の**非負のものである。$p = \frac{3}{4}$ のとき、$(1 + 3e)^2 = 16e$ から $9e^2 - 10e + 1 = 0$ となるので $e = \frac{1}{9}$ または 1 であり、絶滅確率は $\frac{1}{9}$ である。$p = \frac{1}{4}$ のとき、$[0, 1]$ にある不動点は $e = 1$ のみであり、絶滅は確実である（図 12.15.1）。

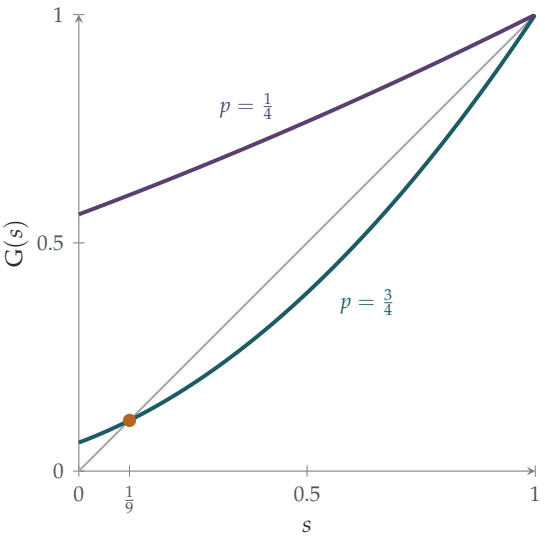


図 12.15.1　子の数の母関数 $\mathrm{G}(s) = (q + ps)^2$ と直線 $y = s$。絶滅確率は両者が交わる最小の点である。$p = \frac{1}{4}$（紫、平均 $\frac{1}{2}$）のとき、それは $s = 1$ であり、絶滅は確実である。$p = \frac{3}{4}$（青緑、平均 $\frac{3}{2}$）のとき、それは $s = \frac{1}{9}$（琥珀色）であり、集団は確率 $\frac{8}{9}$ で永遠に存続する。

臨界値

一般に、分枝過程は子の数の平均が 1 より小さければ確実に絶滅する。平均がちょうど 1 の場合も、すべての個体がちょうど 1 個の子をもつという自明な場合（このときは系統が 1 個ずつ永遠に続くだけである）を除けば、確実に絶滅する。平均が 1 より大きければ、正の確率で存続する。子の数の平均は、疫学における**基本再生産数**（reproduction number）R にあたる。$R > 1$ の感染症は爆発的に広がりうるが、その場合でも初期の流行が偶然に収まることがあり、その確率はまさに上の不動点方程式で与えられる。原子核物理学では、各核分裂が平均してちょうど 1 回の核分裂を引き起こすとき、連鎖反応は**臨界**（critical）にあるという。原子炉の運転とは、この状態を保つ技術にほかならない。

ゴルトン、ワトソン、ビエネメ

ゴルトンの問いには、1874 年に Henry William Watson 牧師が答えたが、その答えは部分的に誤っていた。Watson は不動点方程式を見いだしたものの、2 つ目の解を見落とし、あらゆる家名はいずれ途絶えると結論したのである。実は正しい解析は、1845 年にフランスの統計学者ビエネメ（Irénée-Jules Bienaymé）によってすでに与えられていたが、その業績は 1970 年代まで忘れられていた。このモデルは現在**ゴルトン–ワトソン過程**（Galton–Watson process）（ビエネメ–ゴルトン–ワトソン過程とよばれることもある）とよばれ、確率過程のあらゆる講義で中心的な例となっている。

思考のスキル

第 1 世代で条件づける

自己再生的な過程（各子孫が過程全体の独立な複製を新たに始める過程）は、最初の 1 ステップで条件づけることで得られる方程式を満たす。

方程式 $e = \mathrm{G}(e)$ は不動点問題であり、Section 2.2 や Section 8.7 で出会ったものと同じ種類である。大学では、母関数と不動点が、分枝過程、ランダム木、ランダムネットワークを扱うための主要な道具となる。

次の一歩

1 **一般の場合の絶滅確率。** 子の数の分布が $\mathrm{B}(2, p)$ のとき、絶滅確率は $p \le \frac{1}{2}$ ならば 1、$p > \frac{1}{2}$ ならば $\left(\frac{q}{p}\right)^2$ であることを示せ。

2 **分散。**分散（variance）$\mathrm{Var}(X_n)$ が $\mathrm{Var}(X_n) = (2p)^2\mathrm{Var}(X_{n-1}) + 2pq\,\mathrm{E}(X_{n-1})$ を満たすことを示し、$p = \frac{1}{2}$ のときの値を求めよ。

3 **第 n 世代までの絶滅。** $e_n = \mathrm{P}(X_n = 0)$ とする。$e_n = \mathrm{G}(e_{n-1})$ を示し、$p = \frac{3}{4}$ のときの e_1、e_2、e_3 を計算せよ。これらは $\frac{1}{9}$ に近づくか。

4 **姓。**男性がそれぞれ 0、1、2、3 人の息子をもつ確率がいずれも $\frac{1}{4}$ であるとき、その姓が最終的に途絶える確率はいくらか。

さらに学ぶために

- **Geoffrey Grimmett & David Stirzaker, *Probability and Random Processes***（Oxford University Press）。分枝過程の節。
- **Adam Kucharski, *The Rules of Contagion***（Profile）。一般読者向けに、再生産数と感染の流行を解説している。
- **David Kendall,「Branching processes since 1873」**, *Journal of the London Mathematical Society* 41 (1966)。ビエネメを含む歴史。

12.16 マルコフ連鎖を順向きにも逆向きにも

問題

STEP III 1996、問 12

先に進む前に、まず問題を読んでおこう。問題は STEP Support Programme の問題データベースにある：

`step.maths.org/questions/96-s3-q12`

導入

エックス教授の定理には一定の規則がある。次に何を証明するかは、直前に何を証明したかだけで決まるのである。未来が現在には依存するが過去には依存しない過程を**マルコフ連鎖**（Markov chain）という。マルコフ連鎖はいたるところに現れる。気象モデル、すごろく遊びの「蛇と梯子」（Snakes and Ladders）、DNA のモデル、Google の初期の PageRank アルゴリズム、そして統計学者が複雑な確率分布をシミュレーションするために用いる手法などである。

この問題では、定理の種類が長期的にどのような割合で現れるか、すなわち**定常分布**（stationary distribution）を求める。長期的には、彼女の定理の 6 つに 1 つは新しく正しいものであり、半分は誤りである。さらに、この連鎖を時間について逆向きにたどり、「今日の定理がわかっているとき、昨日の定理は何だったか」を問う。

取りかかる前に

考えてみよう 「各行の和が 1 である」という条件は何を意味するか。各列の和は 1 になるか。

第 i 行には、現在の種類が i であるときの、次の種類の確率が並んでいる。三つの種類のいずれかは必ず起こるので、各行の和は 1 である。一方、列についてはその必要はない。第 2 列の和はちょうど 1 であるが、第 1 列の和は 0.6、第 3 列の和は 1.4 である。

順を追って解く

段階 1　1 ステップ

現在の定理の種類 i で条件づけて、全確率の法則（law of total probability）を用いると、

$$b_j = \sum_i \mathrm{P}(\text{現在が}i)\,\mathrm{P}(\text{次が}j \mid \text{現在が}i) = a_1 p_{1j} + a_2 p_{2j} + a_3 p_{3j}. \qquad \blacksquare$$

行列で表せば、行ベクトルについて $\mathbf{b} = \mathbf{a}P$ である。

段階 2　定常分布

> **考えてみよう**　$a_1 + a_2 + a_3 = 1$ のもとで $\mathbf{a} = \mathbf{a}P$ となるものを求めたい。方程式を書き出してみよ。どれが最も簡単か。

方程式は

$$\begin{aligned} a_1 &= 0.3a_1 + 0.2a_2 + 0.1a_3, \\ a_2 &= 0.3a_1 + 0.4a_2 + 0.3a_3, \\ a_3 &= 0.4a_1 + 0.4a_2 + 0.6a_3. \end{aligned}$$

である（これらは独立ではない。三つを足し合わせると両辺とも $a_1 + a_2 + a_3$ になる）。第 2 式から $0.6a_2 = 0.3(a_1 + a_3) = 0.3(1 - a_2)$ となるので、$a_2 = \frac{1}{3}$ である。第 1 式から $7a_1 = 2a_2 + a_3 = \frac{2}{3} + a_3$ であり、また $a_1 + a_3 = \frac{2}{3}$ であるから、$8a_1 = \frac{4}{3}$ となり、$a_1 = \frac{1}{6}$、$a_3 = \frac{1}{2}$ を得る。すなわち

$$(a_1, a_2, a_3) = \left(\tfrac{1}{6}, \tfrac{1}{3}, \tfrac{1}{2}\right). \qquad \blacksquare$$

段階 3　連鎖を逆向きにたどる

> **考えてみよう**　$q_{ij} = \mathrm{P}(\text{直前が}i \mid \text{現在が}j)$ である。直前の定理と現在の定理の両方について定常分布を用いて、ベイズの定理（Bayes' theorem）を適用せよ。

ベイズの定理により、

$$q_{ij} = \frac{\mathrm{P}(\text{直前が}i)\,\mathrm{P}(\text{現在が}j \mid \text{直前が}i)}{\mathrm{P}(\text{現在が}j)} = \frac{a_i p_{ij}}{a_j}.$$

たとえば $q_{11} = \frac{\frac{1}{6}\times 0.3}{\frac{1}{6}} = 0.3$、$q_{21} = \frac{\frac{1}{3}\times 0.2}{\frac{1}{6}} = 0.4$ である。すべてまとめると（i が行、j

が列を表す)、

$$(q_{ij}) = \begin{pmatrix} 0.3 & 0.15 & \frac{2}{15} \\ 0.4 & 0.4 & \frac{4}{15} \\ 0.3 & 0.45 & 0.6 \end{pmatrix}. \qquad \blacksquare$$

今度は各**列**の和が 1 になる。第 j 列には直前の定理としてありうる種類が並んでいるからである。

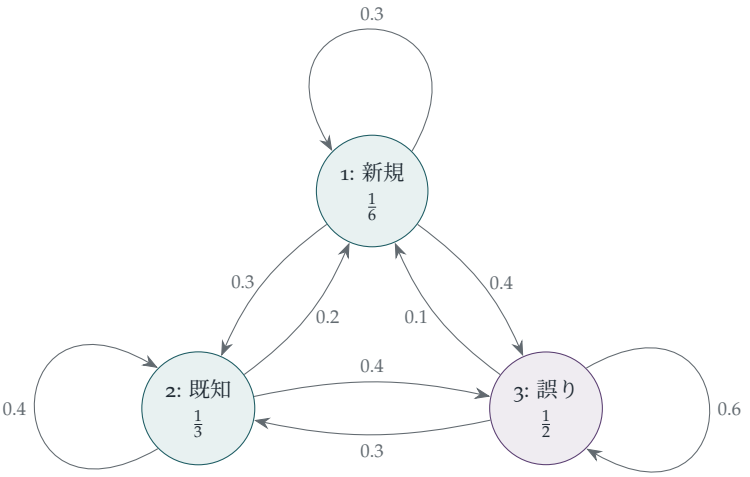


図 12.16.1　マルコフ連鎖として見たエックス教授の定理。矢印は、次の定理がそれぞれの種類になる確率を表す。円の中の分数は定常分布であり、長期的には定理の 6 つに 1 つは新しく正しいもの、3 つに 1 つは既知のもの、半分は誤りである。

振り返り

- **行列としての全確率の法則。**連鎖の 1 ステップは、行ベクトルに行列 P を掛けることである。
- **定常性とは固有ベクトルである。**$\mathbf{a} = \mathbf{a}P$ は、$\mathbf{a}$ が P の固有値（eigenvalue）1 に対する左固有ベクトル（left eigenvector）であることを意味する。
- **ベイズの定理で逆向きにする。**逆向きの確率もまた確率行列（stochastic matrix）をなし、それは順向きの行列と定常分布から作られる。

背後にある数学

平衡への収束

エックス教授が任意の定理から始めたとしよう。そこから n 個の定理の後、種類の分布は $\mathbf{a}_0 P^n$ である。この連鎖のように、どの種類の後にもどの種類も続きうる連鎖では、出発点によらず、P^n はすべての行が定常分布 $(\frac{1}{6}, \frac{1}{3}, \frac{1}{2})$ に等しい行列に収束（convergence）する。収束の速さは P の残りの固有値（ここでは 0.2 と 0.1）によって決まり、それらの寄与は等比数列的に小さくなる。わずか数個の定理の後には、彼女の過去は忘れ去られてしまうのである。マルコフ連鎖に関するこの**エルゴード定理**（ergodic theorem）は、大学 2 年次の確率論の講義の中心的な結果であり、マルコフ連鎖モンテカルロ法（Markov chain Monte Carlo）が機能する根拠でもある。複雑な分布からサンプルを得るには、その分布を定常分布にもつマルコフ連鎖を設計し、長時間走らせればよい。

可逆性

逆向きの連鎖 (q_{ij}) は順向きの連鎖 (p_{ij}) とは異なる。つまり、彼女の定理の長い記録を見れば、時間の向きを見分けることができる。逆向きと順向きが一致する連鎖、すなわちすべての i、j について $a_i p_{ij} = a_j p_{ji}$ が成り立つ連鎖を**可逆**（reversible）であるという。ネットワーク上のランダムウォーク（random walk）や、熱平衡（thermal equilibrium）にある物理系のモデルがその例である。可逆性は強力な性質である。定常分布を容易に求められるようになるからである。

マルコフとプーシキン

アンドレイ・マルコフ（Andrey Markov）は 1906 年に自らの連鎖を導入した。その動機の一つは、大数の法則（law of large numbers）が成り立つには独立性（independence）が必要だと主張していた数学者パーヴェル・ネクラーソフ（Pavel Nekrasov）との論争において、独立性は必要ないことを示すことであった。1913 年、マルコフはこの考えをプーシキン（Pushkin）の韻文小説**エヴゲーニイ・オネーギン**の最初の 20 000 文字に適用した。各文字を母音か子音かに分類し、各文字の種類が直前の文字の種類に依存し、その依存の仕方が 2 状態の連鎖でよく記述できることを示したのである。これは文章に対する最初の統計的分析であり、あらゆる予測変換システムの祖先である。

思考のスキル

確率をベクトルとして、推移を行列として捉える

系が一定の確率で状態の間を移り変わるときは、分布をベクトルとして、推移を行列として書け。すると、長期的な振る舞いは固有ベクトルについての問題になる。

大学では、これによって第 7 章の線形代数が確率論に持ち込まれる。ペロン–フロベニウスの定理（Perron–Frobenius theorem）が定常分布の存在を保証し、残りの固有値がそこへ到達する速さを測るのである。

次の一歩

1 **ウサギの巣穴。** STEP III 2003、問 13（`step.maths.org/questions/03-s3-q13`）に挑戦せよ。各頂点に出口のある正方形の頂点上のランダムウォークを、最初の 1 ステップに関する方程式によって解く問題である。

2 **2 ステップ。** エックス教授の直前の定理が新しく正しいものであったとき、その次の次の定理が三つの種類それぞれである確率を求めよ。

3 **固有値。** P の固有値が 1、0.2、0.1 であることを示せ。

4 **可逆か？** エックス教授の連鎖について $a_i p_{ij} = a_j p_{ji}$ が成り立つかどうか確かめよ。その答えから何がわかるか。

さらに学ぶために

- **James Norris, *Markov Chains***（Cambridge University Press）。学部生向けの標準的な教科書である。
- **Olle Häggström, *Finite Markov Chains and Algorithmic Applications***（Cambridge University Press）。簡潔で、マルコフ連鎖モンテカルロ法の解説に優れる。
- **Brian Hayes, “First links in the Markov chain”,** *American Scientist* 101 (2013)。マルコフによるプーシキンの分析を改めて紹介する記事である。

12.17 バナッハのマッチ箱

問題

STEP II 1992、問 16

先に進む前に、まず問題を読んでおこう。問題は STEP Support Programme の問題データベースにある：

`step.maths.org/questions/92-s2-q16`

導入

数学者ステファン・バナッハ（Stefan Banach）は大変なヘビースモーカーで、左右のポケットにマッチ箱を一つずつ入れていた、という話が伝わっている。マッチが欲しくなるたびに、彼はどちらかのポケットを無作為に選んだ。ある日、彼は箱の一つが空であることに気づいた。もう一方の箱には何本のマッチが残っていたか。これが**バナッハのマッチ箱問題**（Banach's matchbox problem）であり、ヒューゴ・シュタインハウス（Hugo Steinhaus）がバナッハにちなんで考案したものである。本問は、その問題をタクシー運転手とお菓子に置き換えたものである。

微妙な点は「気づく」という言葉にある。運転手は、ある袋の最後のお菓子を取った時点では袋が空になったことに気づかず、次にその袋に手を伸ばしたときに初めて気づく。したがって考えるべき事象は、「トフィーに手を伸ばしたとき、n 個のトフィーがすでにすべてなくなっている」というものである。これが N 回目に手を伸ばしたときに起こる場合の数を数えることは、二項分布（binomial distribution）の格好の演習になる。そしてバナッハの問題の答えには意外な特徴がある。もう一方の箱に残っているマッチの本数として最も起こりやすい値は 0 なのである（1 と同率である）。

取りかかる前に

考えてみよう　運転手が N 回目に手を伸ばしたときにトフィーがなくなっていることに気づくためには、N 回目はどちらの袋に手を伸ばしたのでなければならないか。それより前の $N-1$ 回には何が含まれていなければならないか。

N 回目はトフィーに手を伸ばしたのでなければならず、それより前の $N-1$ 回では、ちょうど n 個のトフィー（すべて）と $N-1-n$ 個のミントを取っていなければならない。さらに、それより前にミントの袋が空であることに気づいていてはならない。

順を追って解く

段階 1　トフィーがなくなっていることに気づく

考えてみよう　最初の $N-1$ 回はどのような順序で起こりうるか。$N-1-n$ にどのような条件があれば、ミントの袋が空であることに一度も気づかずに済むか。

トフィーに手を伸ばした n 回は、最初の $N-1$ 回のうちのどの n 回であってもよい。ミントがなくなった後にミントに手を伸ばすことが一度もないのは、取ったミントが高々 m 個、すなわち $N-1-n \le m$ のときである。よって $n+1 \le N \le n+m+1$ に対して

$$\mathrm{P}(T_N) = \binom{N-1}{n} p^{n+1} q^{N-1-n},$$

が成り立つ。ここで T_N は事象「N 回目に手を伸ばしたときにトフィーがなくなっていることに気づく」である（因子 p^{n+1} には N 回目そのものも含まれている）。

段階 2　ミントがなくなっていることに気づく

役割を入れ替えて同じ議論をすれば、$m+1 \le N \le n+m+1$ に対して

$$\mathrm{P}(M_N) = \binom{N-1}{m} q^{m+1} p^{N-1-m}.$$

段階 3　条件付き確率

考えてみよう　N 回目に手を伸ばしたときに空の袋に気づいたことがわかっているとする。それがトフィーであった確率はいくらか。

事象 T_N と M_N は排反（mutually exclusive）であるから、条件付き確率（conditional probability）は

$$\mathrm{P}(\text{トフィー} \mid \text{発見時刻} N) = \frac{\binom{N-1}{n} p^{n+1} q^{N-1-n}}{\binom{N-1}{n} p^{n+1} q^{N-1-n} + \binom{N-1}{m} q^{m+1} p^{N-1-m}},$$

となる。ただし、二項係数（binomial coefficient）は上で述べた範囲の外では 0 とみなす。（$N \leq m$ ならば、なくなりうるのはトフィーだけであり、確率は 1 である。同様に、$N \leq n$ かつ $N > m$ ならば確率は 0 である。） ■

> **手を止めて、やってみよう** $n = m$ かつ $p = q = \frac{1}{2}$ のときに、この公式を確かめよ。対称性（symmetry）から、ありうるすべての N に対して答えは $\frac{1}{2}$ になるはずである。実際にそうなるか。

振り返り

- **事象を正確に読む。**「気づく」とは、最後のお菓子を取ることではなく、空の袋に手を伸ばすことであった。
- **最後の一歩を固定し、残りを数える。** N 回目を固定すると、最初の $N-1$ 回は任意の順序でよく、二項係数が現れた。
- **境界を確かめる。**制約 $N-1-n \leq m$ と $N-1-m \leq n$ によって、起こりえない経過が除外された。

背後にある数学

バナッハのマッチ箱

バナッハの設定では、どちらの箱にも最初 n 本のマッチが入っており、どちらのポケットも等確率で選ばれる。初めて箱が空であることに気づいたとき、もう一方の箱に入っているマッチの本数を K とする。$K = k$ となるためには、空の箱からは n 本、もう一方の箱からは $n-k$ 本を任意の順序で取り、その後もう一度空の箱に手を伸ばさなければならない。また、空になるのはどちらの箱でもよい。よって

$$\mathrm{P}(K = k) = 2\binom{2n-k}{n}\left(\frac{1}{2}\right)^{2n-k+1} = \binom{2n-k}{n}\left(\frac{1}{2}\right)^{2n-k}.$$

これらの確率は k について減少するが、例外として $\mathrm{P}(K = 0) = \mathrm{P}(K = 1)$ である（図 12.17.1）。残っている本数の期待値（expected value）はおよそ $2\sqrt{n/\pi} - 1$ であり、各箱に 50 本ずつ入っている場合はおよそ 7 本となる。$\sqrt{n}$ は、$2n$ 回の無作為な選択の

後に二つの箱の間に生じる偏りの典型的な大きさであり、誕生日問題（Section 9.9）やランダムウォーク（random walk）に現れるのと同じ平方根の法則である。

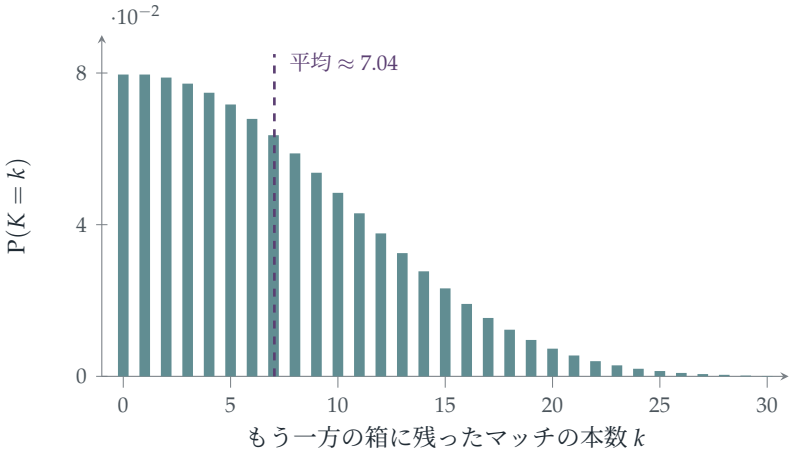


図 12.17.1　各箱に 50 本ずつマッチが入っている場合のバナッハのマッチ箱問題：一方の箱が空であることに初めて気づいたとき、もう一方の箱に残っている本数の分布。最も起こりやすい値は 0 と 1 で同率であるが、平均はおよそ 7 である。

負の二項分布

各回が確率 p で成功する独立な試行において、r 回目の成功を得るまでに必要な試行回数は**負の二項分布**（negative binomial distribution）$\mathrm{P}(N = k) = \binom{k-1}{r-1} p^r q^{k-r}$ に従う。タクシー運転手の問題は、二つの負の二項過程（トフィーとミント）の競争であり、どちらも袋のお菓子の個数より一つ多い「成功」を必要とする。負の二項分布は保険（請求件数）、生態学（群れをなす動物の個体数）、遺伝学などに現れ、ポアソン分布（Poisson distribution）より柔軟な代替としてしばしば用いられる。

バナッハとスコティッシュ・カフェ

関数解析（functional analysis）の創始者の一人であるステファン・バナッハは、ルヴフ（現在のリヴィウ）のスコティッシュ・カフェで数学の多くを行った。彼と同僚たちはウェイターが預かるノート、すなわち**スコティッシュ・ブック**（Scottish Book）に問題を書き込み、その懸賞はビール一杯から生きたガチョウまでさまざまであった。バナッハの師であるヒューゴ・シュタインハウスは、バナッハの喫煙癖を称えるスピーチのためにマッチ箱問題を考案した。この問題を初めて活字で解析したのは、ウィリアム・フェラー（William Feller）による 1950 年の名高い確率論の教科書である。

思考のスキル

最後の事象を固定し、そこに至る経路を数える

ある過程が特定の時刻に特定の仕方で終わる確率を求めるには、最後の一歩で起こることを固定し、そこに至る経過を数えよ。

大学では、これは負の二項分布、ランダムウォークの初到達時刻（first-passage time）、そして Section 9.10 の鏡像原理（reflection principle）の基礎となる。

次の一歩

1 **小さな場合。** $n = m = 1$ かつ $p = \frac{1}{2}$ のとき、起こりうる列をすべて書き出し、N の分布を求めよ。

2 **バナッハの公式。** $n = 1$ と $n = 2$ に対して $\sum_{k=0}^{n} \binom{2n-k}{n} 2^{-(2n-k)} = 1$ が成り立つことを確かめよ。

3 **同率の最大値。**バナッハの問題において、すべての n に対して $\mathrm{P}(K = 0) = \mathrm{P}(K = 1)$ であることを示せ。

4 **先になくなるのはどちらか。** タクシー運転手の問題で、$n = m = 2$ かつ $p = \frac{2}{3}$ のとき、（時刻を問わず）最初に空であると気づかれる袋がトフィーである確率を求めよ。

さらに学ぶために

- **William Feller, *An Introduction to Probability Theory and Its Applications*, Vol. 1**（Wiley）の第 VI 章。バナッハの問題を世に紹介した文献である。
- **R. Daniel Mauldin (ed.), *The Scottish Book***（Birkhäuser）。スコティッシュ・カフェの問題集で、解説付き。
- **Geoffrey Grimmett & Dominic Welsh, *Probability: An Introduction***（Oxford University Press）。負の二項分布について。

12.18 無記憶性

問題

STEP III 2021、問 11

先に進む前に、まず問題を読んでおこう。問題は、OCR が配布している 2021 年の STEP 過去問一式（zip ファイル）のうち、STEP III の問題用紙にある：

`www.ocr.org.uk/Images/696588-2021-past-papers.zip`

導入

ある放射性原子が、研究室で 100 万年ものあいだ崩壊せずに残っているとする。この原子はそろそろ崩壊する「はず」だろうか。そうではない。次の 1 秒間に崩壊する確率は、つくられたばかりの原子の場合とまったく同じである。放射性原子は年をとらないのである。崩壊までの待ち時間は**指数分布**（exponential distribution）に従い、その本質的な性質は記憶をもたないことである。

この問題は、その無記憶性（memorylessness）を印象的な形で示している。指数分布に従う待ち時間を、整数単位の部分と小数部分に分ける。すると、この二つは**独立**（independent）であることがわかる。何時間（整数部分）待ったかを知っても、端数の分については何もわからないのである。また演習問題では、バスを待つ時間がバスの運行間隔より長く感じられることがある理由も説明する。

取りかかる前に

考えてみよう $x \geq 0$ に対して $\mathrm{P}(X > x)$ はいくらか。$\mathrm{P}(X > s+t \mid X > s)$ はいくらか。

$\mathrm{P}(X > x) = \int_x^\infty \lambda e^{-\lambda u}\,\mathrm{d}u = e^{-\lambda x}$ である。したがって $\mathrm{P}(X > s+t \mid X > s) = \frac{e^{-\lambda(s+t)}}{e^{-\lambda s}} = e^{-\lambda t} = \mathrm{P}(X > t)$ となる。すでに s だけ待ったとしても、残りの待ち時間は新たに待ち始める場合と同じ分布に従う。これが**無記憶性**である。

順を追って解く

段階 1　整数部分

$Y = n$ となるのは、ちょうど $n \le X < n+1$ のときである。

$$\mathrm{P}(Y = n) = e^{-\lambda n} - e^{-\lambda(n+1)} = (1 - e^{-\lambda})e^{-n\lambda}. \qquad \blacksquare$$

よって Y は幾何分布（geometric distribution）に従う。各単位時間を確率 $e^{-\lambda}$ で「生き延びる」わけである。

段階 2　小数部分

> **考えてみよう**　$Z < z$ とは、X が区間 $[n, n+z)$ のいずれかに入るということである。それらの確率を足し合わせよ。

$$\mathrm{P}(Z < z) = \sum_{n \ge 0} \left(e^{-\lambda n} - e^{-\lambda(n+z)}\right) = (1 - e^{-\lambda z}) \sum_{n \ge 0} e^{-\lambda n} = \frac{1 - e^{-\lambda z}}{1 - e^{-\lambda}}. \qquad \blacksquare$$

段階 3　Z の平均

> **考えてみよう**　$[0,1]$ に値をとる確率変数（random variable）に対しては $\mathrm{E}(Z) = \int_0^1 \mathrm{P}(Z > z)\,\mathrm{d}z$ が成り立つ。これを用いよ。

$c = 1 - e^{-\lambda}$ とおくと $\mathrm{P}(Z > z) = 1 - \frac{1 - e^{-\lambda z}}{c}$ であるから、

$$\mathrm{E}(Z) = 1 - \frac{1}{c}\left(1 - \frac{1 - e^{-\lambda}}{\lambda}\right) = 1 - \frac{1}{c} + \frac{1}{\lambda} = \frac{1}{\lambda} - \frac{e^{-\lambda}}{1 - e^{-\lambda}} = \frac{1}{\lambda} - \frac{1}{e^{\lambda} - 1}. \qquad \blacksquare$$

（λ が小さい、すなわち待ち時間が長いとき、これは $\frac{1}{2}$ に近づく。小数部分はほぼ一様分布になるのである。λ が大きいときは $\frac{1}{\lambda}$ に近い。X がほとんど常に 1 より小さいからである。）

段階 4　独立性

$$\mathrm{P}(Y = n, z_1 < Z < z_2) = \mathrm{P}(n + z_1 < X < n + z_2) = e^{-\lambda n}\left(e^{-\lambda z_1} - e^{-\lambda z_2}\right).$$

> **考えてみよう**　小問 (i) と (ii) を用いて、$\mathrm{P}(Y = n)\,\mathrm{P}(z_1 < Z < z_2)$ と比較せよ。

(i) と (ii) より $\mathrm{P}(Y=n)\mathrm{P}(z_1<Z<z_2)=(1-e^{-\lambda})e^{-n\lambda}\cdot\frac{e^{-\lambda z_1}-e^{-\lambda z_2}}{1-e^{-\lambda}}$ であり、これは上と一致する。よって Y と Z は独立である。 ■

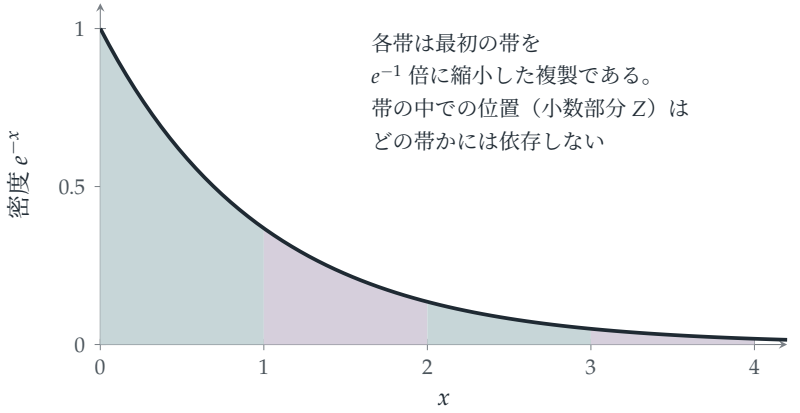


図 12.18.1　$\lambda=1$ の指数分布の密度関数を幅 1 の帯に分割したもの。$e^{-(n+z)}=e^{-n}e^{-z}$ であるから、どの帯も同じ形をしている。X がどの帯に入るか（整数部分 Y）と、帯の中のどこに入るか（小数部分 Z）は独立である。

振り返り

- **区間ごとに足し合わせる。** Z の分布も同時確率も、区間 $[n,n+1)$ ごとの確率を足し合わせることで得られた。
- **裾の公式。** $\mathrm{E}(Z)=\int \mathrm{P}(Z>z)\,\mathrm{d}z$ を使うことで、密度関数を求めずに済んだ。
- **積に分解できることは独立性を意味する。**同時確率が、n の関数と z_1, z_2 の関数の積に分解された。

背後にある数学

指数分布はなぜ特別なのか

Y と Z の独立性は $e^{-\lambda(n+z)}=e^{-\lambda n}e^{-\lambda z}$ から直接導かれるが、これは無記憶性が姿を変えたものにほかならない。実は、記憶をもたない連続分布は指数分布**だけ**である。すべての $s,t\geq 0$ に対して $\mathrm{P}(X>s+t)=\mathrm{P}(X>s)\mathrm{P}(X>t)$ が成り立つならば、$g(t)=\ln \mathrm{P}(X>t)$ は $g(s+t)=g(s)+g(t)$ を満たし、（g が単調ならば）これにより $g(t)=-\lambda t$ となる。だからこそ指数分布は、蓄積なしに「ランダムに」起こる事象の待ち時間のモデルとなるのである。放射性崩壊、客の到着、摩耗しない電子部品の故障などがその例である。

最小値と待ち時間のパラドックス

X_1 と X_2 がそれぞれ率 λ_1, λ_2 の独立な指数分布に従うならば $\mathrm{P}(\min > x) = e^{-\lambda_1 x} e^{-\lambda_2 x}$ であるから、最小値は率 $\lambda_1 + \lambda_2$ の指数分布に従う。二つの独立なランダムな事象のうち最初のものを待つことは、合わせた率で起こる一つの事象を待つことと同じである（演習問題の STEP III 2002 Q13）。無記憶性は**待ち時間のパラドックス**（waiting-time paradox）も説明する。バスがランダムに到着し、平均の間隔が 10 分であるとき、ランダムな時刻に到着した乗客の平均待ち時間は 5 分ではなく 10 分である。残りの待ち時間は記憶をもたないからである。さらに、乗客が到着した時点を含む間隔の長さは平均 20 分である。長い間隔ほど、その中に到着しやすいからである。

ラザフォードとポアソン過程

1910 年、アーネスト・ラザフォード（Ernest Rutherford）とハンス・ガイガー（Hans Geiger）は、放射線源から放出されるアルファ粒子を 7.5 秒ごとの区間で数え、その個数がポアソン分布（Poisson distribution）に従うことを見いだした。放出が一定の率で独立に起こるならばそうなるはずであることを、数学者 Harry Bateman が示していた。このとき放出と放出の間の時間は指数分布に従う。**ポアソン過程**（Poisson process）、すなわち間隔が指数分布に従う事象の列は、今日では時間におけるランダムさの基本モデルであり、電話の呼び出し、保険金の請求、ニューロンの発火などに用いられている。次の節 Section 12.19 では、これを空間へと拡張する。

思考のスキル

積への分解で独立性を確かめる

二つの確率変数は、その同時確率が一方の関数と他方の関数の積に分解されるならば独立である。

大学では、同時密度関数の積への分解（factorisation）が独立性を確かめる標準的な方法であり、多変量の問題で独立な成分を見抜く手段となる。たとえば Section 12.22 における、正規分布に従う点の動径と偏角がその例である。

次の一歩

1 **二つのバス路線。** STEP III 2002、問 13（`step.maths.org/questions/02-s3-q13`）に挑戦せよ。二つの指数分布に従う確率変数の最小値を用いて、次のバスまでの

待ち時間を求める問題である。

2 **離散分布の無記憶性。**幾何分布が記憶をもたないこと、すなわち $\mathrm{P}(Y \geq m+n \mid Y \geq m) = \mathrm{P}(Y \geq n)$ を示せ。

3 **待ち時間のパラドックス。**バスが率 λ のポアソン過程に従って到着する。固定された時刻 t を含む間隔の長さの期待値は（t が大きいとき）$\frac{2}{\lambda}$ であり、平均間隔の 2 倍であることを示せ。

4 **Z の密度関数。** Z の確率密度関数を求め、$\lambda = 1$ と $\lambda = 5$ の場合についてその概形をかけ。

さらに学ぶために

- **Geoffrey Grimmett & Dominic Welsh, *Probability: An Introduction***（Oxford University Press）。指数分布とポアソン過程について。
- **Sheldon Ross, *Introduction to Probability Models***（Academic Press）。指数分布、ポアソン過程、待ち行列。
- **David Spiegelhalter, *The Art of Uncertainty***（Pelican）。日常生活におけるランダムさと待ち時間。

12.19 平面上のランダムな点

問題

STEP II 2012、問 13

先に進む前に、まず問題を読んでおこう。問題は STEP Support Programme の問題データベースにある：

`step.maths.org/questions/12-s2-q13`

導入

スーパーマーケット、木、星、携帯電話の基地局などがある領域に「ランダムに」散らばっているとき、いちばん近いものまでどれだけ行かなければならないだろうか。その答えは密度 k（単位面積あたりの平均個数）だけで決まり、しかも特徴的な形をもつ。それが**レイリー分布**（Rayleigh distribution）である。最も近いものまでの距離は平均すると $\frac{1}{2\sqrt{k}}$ であり、1 平方キロメートルあたりスーパーマーケットが 1 軒なら 500 メートルになる。

このモデルは**空間ポアソン過程**（spatial Poisson process）であり、時間上のランダムな到着（Section 12.18）を 2 次元にしたものである。計算は短く美しいが、本当の面白さは基準としての役割にある。生態学者は実際の最近接距離（nearest-neighbour distance）をこの公式と比べ、木や巣が集中しているのか、ランダムに散らばっているのか、意図的に間隔をあけて配置されているのかを判断するのである。

取りかかる前に

考えてみよう 平均（mean）μ のポアソン確率変数（Poisson random variable）が値 0 をとる確率はいくらか。

$\mathrm{P}(0) = e^{-\mu}$ である。半径 y の円の面積は πy^2 なので、その内部にあるスーパーマーケットの個数の平均は $k\pi y^2$ である。

順を追って解く

段階 1　近くにスーパーマーケットがない

距離 y 以内にスーパーマーケットが 1 軒もない確率は $e^{-\pi k y^2}$ である。

段階 2　最近接距離の分布

> **考えてみよう**　$Y \geq y$ とは、距離 y 以内にスーパーマーケットが 1 軒もないということである。では $\mathrm{P}(Y < y)$ はいくらか。

$\mathrm{P}(Y < y) = 1 - e^{-\pi k y^2}$ であり、これを微分すると確率密度関数（probability density function）

$$\mathrm{f}(y) = 2\pi k y e^{-\pi k y^2}, \qquad y \geq 0. \qquad \blacksquare$$

が得られる。

段階 3　平均

> **考えてみよう**　$y\,\mathrm{f}(y)$ を部分積分（integration by parts）するか、$\mathrm{E}(Y) = \int_0^\infty \mathrm{P}(Y > y)\,\mathrm{d}y$ を用いる。そのあと置換して、与えられたガウス積分（Gaussian integral）が使える形にする。

$$\mathrm{E}(Y) = \int_0^\infty e^{-\pi k y^2}\,\mathrm{d}y.$$

$x = y\sqrt{2\pi k}$ と置換すると $\pi k y^2 = \frac{1}{2}x^2$ となり、$\mathrm{E}(Y) = \frac{1}{\sqrt{2\pi k}}\sqrt{\frac{\pi}{2}} = \frac{1}{2\sqrt{k}}$ を得る。 $\blacksquare$

段階 4　分散

$$\mathrm{E}(Y^2) = \int_0^\infty y^2 \cdot 2\pi k y e^{-\pi k y^2}\,\mathrm{d}y.$$

$u = \pi k y^2$ と置換すると（$\mathrm{d}u = 2\pi k y\,\mathrm{d}y$ である）、これは $\int_0^\infty \frac{u}{\pi k} e^{-u}\,\mathrm{d}u = \frac{1}{\pi k}$ となる。したがって分散（variance）は

$$\mathrm{Var}(Y) = \frac{1}{\pi k} - \frac{1}{4k} = \frac{4 - \pi}{4\pi k}. \qquad \blacksquare$$

である。

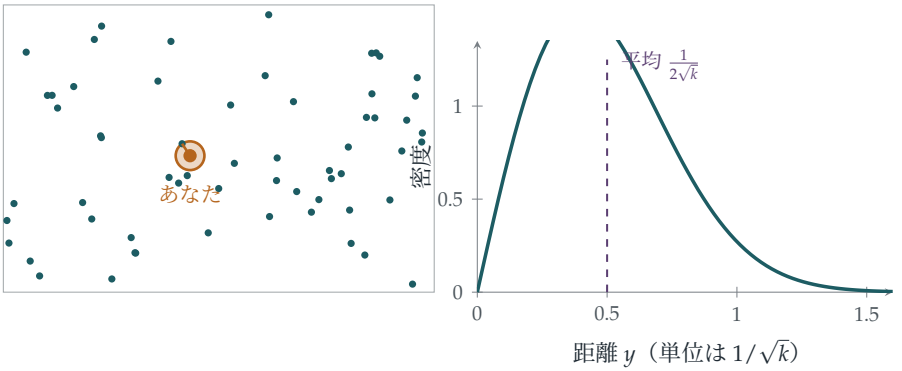


図 12.19.1　左：面積 24 の領域にランダムに散らばった 60 個の点（したがって $k = 2.5$）と、選んだ 1 点からそのうち最も近い点までの距離（琥珀色）。右：その距離のレイリー密度 $2\pi k y e^{-\pi k y^2}$。距離は $1/\sqrt{k}$ を単位として測っている。平均は $\frac{1}{2\sqrt{k}}$ である。

振り返り

- **距離の問題を個数の問題に言い換える。**「最も近いものが y より遠い」とは「y 以内に 1 つもない」ということである。
- **まず累積分布、次に密度。** $\mathrm{P}(Y < y)$ は直接求まり、それを微分して密度が得られた。
- **置換によって既知の積分に帰着させる。**どのモーメントも、ガウス積分か指数関数の積分に帰着した。

背後にある数学

平面上のポアソン過程

強度（intensity）k の空間ポアソン過程とは、点のランダムな散らばりであって、互いに交わらない領域に含まれる点の個数が独立であり、面積 A の領域に含まれる点の個数が平均 kA のポアソン分布に従うものである。これは、膨大な個数の点を膨大な面積の上に独立かつ一様に、適切な平均密度で落としたときに得られるものであり、攪乱順列（derangements、Section 9.8）を説明するのと同じ「稀な事象の法則」（law of rare events）による。これは完全空間ランダム性（complete spatial randomness）の自然なモデルであり、現実のパターンはこれと比べて評価される。2 人の生態学者が 1954 年に考案した**クラーク–エヴァンス検定**（Clark–Evans test）は、観測された最近接距離の平均を $\frac{1}{2\sqrt{k}}$ と比較する。それよりずっと小さければ集中（親木の近くに落ち

る種子）を示唆し、ずっと大きければ規則的な間隔（縄張りをもつ鳥や、水をめぐる植物どうしの競争）を示唆する。

3 次元の場合

同じ計算を 3 次元で行うと、最も近い星、分子、銀河までの距離について $\mathrm{P}(Y > y) = e^{-\frac{4}{3}\pi k y^3}$ が得られる。チャンドラセカール（Subrahmanyan Chandrasekhar）は 1943 年にこれを用いて、ランダムに分布する近隣の星から 1 つの星が受ける重力を研究した。また物理学者は、気体分子の平均自由行程（mean free path）にこれを用いている。

レイリー

密度 $\frac{y}{\sigma^2} e^{-y^2/(2\sigma^2)}$ の名は、1880 年に音響理論の中でこれに出会ったレイリー卿（Lord Rayleigh）にちなむ。これは、位相（phase）がランダムな多数の波を重ね合わせたときの振幅を表す。また、2 つの座標が独立な正規確率変数であるような点の原点からの距離もこの分布に従い、これがここに現れる理由を別の角度から見る方法になっている（演習 3）。今日では、多数の経路をランダムな位相で通って受信機に届く携帯電話の電波のフェージング（fading）のモデルとして使われている。

思考のスキル

まず分布関数を求める

最小値、最大値、最近接距離の分布については、$\mathrm{P}(Y > y)$ を直接書き下すのがたいてい簡単である。それを微分して密度を得る。

この方法は、多数のランダムな寿命のうち最大のものを扱う Section 12.25 でも再び登場する。大学では、あらゆる「順序統計量」（order statistic）を扱う際の標準的な第一歩である。

次の一歩

1. **中央値。**最も近いスーパーマーケットまでの距離の中央値（median）を求め、平均と比較せよ。

2. **2 番目に近いもの。**2 番目に近いスーパーマーケットまでの距離の分布を求めよ。（ヒント：$\mathrm{P}(Y_2 > y)$ は、y 以内にスーパーマーケットが高々 1 軒しかない確率である。）

3 **正規分布に従う座標。** U と V が独立に $\mathrm{N}(0,\sigma^2)$ に従うとき、$\sqrt{U^2+V^2}$ の密度が $\frac{r}{\sigma^2}e^{-r^2/(2\sigma^2)}$ であることを示せ。σ をどの値にすればスーパーマーケットの分布が得られるか。

4 **クラーク–エヴァンス。** 400 本の木がある 1 km^2 の森を調査したところ、木から最も近い隣の木までの距離の平均は 35 m であった。これは完全空間ランダム性と整合するか。それとも集中、あるいは規則的な間隔を示唆するか。

さらに学ぶために

- **Peter Diggle, *Statistical Analysis of Spatial Point Patterns***（Arnold）。空間点パターンの統計解析に関する標準的な文献で、クラーク–エヴァンス検定とその発展形を扱っている。
- **John Kingman, *Poisson Processes***（Oxford University Press）。任意の次元のポアソン過程を扱った、短く美しい本である。
- **Geoffrey Grimmett & David Stirzaker, *Probability and Random Processes***（Oxford University Press）。ポアソン過程の章。

12.20 ビュフォンの針

問題

STEP I 1993、問 16

先に進む前に、まず問題を読んでおこう。問題は STEP Support Programme の問題データベースにある：

`step.maths.org/questions/93-s1-q16`

導入

平行に板が張られた床の上に針を落とす。針が板の継ぎ目をまたぐ確率には、どこにも円が見当たらないにもかかわらず π が現れる。何度も投げて交わった回数を数えれば、実験によって π を推定できるのである。これが**ビュフォンの針**（Buffon's needle）であり、ビュフォン伯ジョルジュ＝ルイ・ルクレール（Georges-Louis Leclerc, Comte de Buffon）が 1733 年に提起し、1777 年に自ら解いたものである。これは**幾何確率**（geometric probability）の最初の問題であり、あらゆるモンテカルロ法（Monte Carlo method）の祖先でもある。

π は角度から生じる。針の向きは半回転の範囲で一様に分布しており、直線に対する針の「影」の長さをすべての向きについて平均すると、因子 $\frac{2}{\pi}$ が得られる。この問題では、その計算を $\cos^{-1}$ の積分によって行う。

取りかかる前に

考えてみよう 針の中心が最も近い直線から距離 x のところにあり、針が直線の垂線と角 y をなしているとする。針は直線に向かってどこまで届くか。

針の半分ずつの長さは a であり、その直線に垂直な成分は $a\cos y$ である。したがって、針が最も近い直線と交わるのは、ちょうど $a\cos y \geq x$ のときである。（$a < b$ であるから、針がほかの直線に届くことはない。）

順を追って解く

段階 1　積分

$t = \cos y$ とおくと、t が 0 から 1 まで動くとき、$y = \cos^{-1} t$ は $\frac{\pi}{2}$ から 0 まで動き、$\mathrm{d}t = -\sin y \,\mathrm{d}y$ である。よって

$$\int_0^1 \cos^{-1} t \,\mathrm{d}t = \int_0^{\pi/2} y \sin y \,\mathrm{d}y = \left[-y\cos y\right]_0^{\pi/2} + \int_0^{\pi/2} \cos y \,\mathrm{d}y = 0 + 1 = 1. \qquad \blacksquare$$

（別解：この積分は $\cos^{-1}$ のグラフと座標軸で囲まれた部分の面積であり、これは $y = \cos x$ の 0 から $\frac{\pi}{2}$ までのグラフの下の面積を直線 $y = x$ に関して折り返したものに等しい。）

段階 2　位置を固定したとき

> **考えてみよう**　x を固定したとき、どのような角 y で交わるか。その確率はいくらか。

$x \geq a$ ならば、針は直線に届かない。$x < a$ ならば、交わるためには $\cos y \geq \frac{x}{a}$、すなわち $y \leq \cos^{-1}\frac{x}{a}$ が必要である。Y は $[0, \frac{\pi}{2}]$ 上の一様分布（uniform distribution）に従うから、

$$\mathrm{P}(\text{交わる} \mid X = x) = \frac{2}{\pi}\cos^{-1}\frac{x}{a} \qquad (0 \leq x < a). \qquad \blacksquare$$

段階 3　位置について平均する

X は $[0, b]$ 上の一様分布に従い、その確率密度（density）は $\frac{1}{b}$ であるから、

$$\mathrm{P}(\text{交わる}) = \frac{1}{b}\int_0^a \frac{2}{\pi}\cos^{-1}\frac{x}{a}\,\mathrm{d}x = \frac{2a}{\pi b}\int_0^1 \cos^{-1} t \,\mathrm{d}t = \frac{2a}{\pi b}, \qquad \blacksquare$$

ここで $x = at$ と置換した。

> **手を止めて、やってみよう**　針の長さがちょうど間隔の半分である場合（$a = \frac{b}{2}$、したがって針の長さ $2a$ は b に等しい）、交わる確率は $\frac{1}{\pi}$ となる。π を小数第 2 位まで推定するには、何回投げる必要があるだろうか。

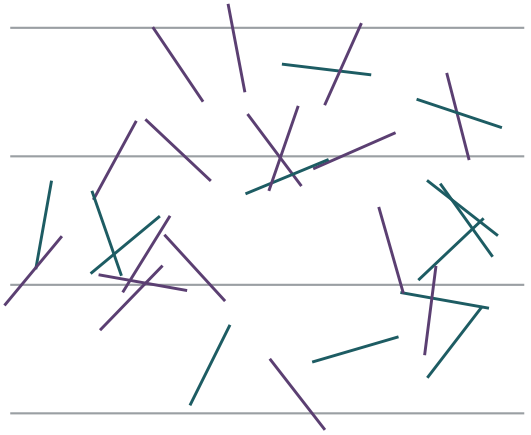


図 12.20.1　長さ $2a = 1.4$ の針 30 本を、間隔 $2b = 2$ の平行線の上に投げたもの。直線と交わる針は紫色で示してある。ここでは期待値 13.4 に対して 17 本が交わっている。標本が小さいと推定は粗く、ビュフォンの針は π の計算法としては遅いことで有名である。

振り返り

- **2 つの独立な座標。**針の位置と角度は独立（independent）かつ一様であったので、確率は 1 変数ずつ計算できる二重積分となった。
- **条件をつけてから平均する。**まず x を固定して角度について平均し、次に x について平均した。
- **置換によって定数が姿を現す。**$x = at$ と尺度を変えることで、答えはただ 1 つの普遍的な積分に帰着した。

背後にある数学

ビュフォンの麺

ほとんど計算なしで答えを見通す方法がある。針を任意の形に曲げて「麺」（noodle）にし、それが直線と交わる回数を C とする。期待値の線形性（linearity of expectation）により、曲線の交わる回数の期待値はその小さな各部分についての和であるから、長さに比例する。すなわち、ある定数 κ に対して $\mathrm{E}(C) = \kappa L$ である。κ を求めるために、麺として直径が直線の間隔 $2b$ に等しい円をとる。この円は常にちょうど 2 回交わり、その長さは $2\pi b$ である。よって $\kappa \cdot 2\pi b = 2$ となり、$\kappa = \frac{1}{\pi b}$ を得る。長さ $2a < 2b$ のまっすぐな針は高々 1 回しか交わらないので、$\mathrm{P}(\text{交わる}) = \mathrm{E}(C) = \frac{2a}{\pi b}$ である。この議論はバルビエ（Joseph-Émile Barbier）によるもので（1860 年）、確率論的推論の珠玉

である。π は、定数を較正するために選んだ円から来ているのである。

π の推定と、してはいけないこと

n 本の針を投げて c 本が交わったとすると、$\frac{c}{n} \approx \frac{2a}{\pi b}$ より $\pi \approx \frac{2an}{bc}$ である。誤差は $\frac{1}{\sqrt{n}}$ 程度でしか小さくならないので、π を小数第 3 位まで求めるには約 100 万回投げる必要がある。1901 年、イタリアの数学者 Mario Lazzarini は、3408 回の試行から $\pi \approx \frac{355}{113} = 3.1415929\ldots$（小数第 6 位まで正しい）を得たと主張した。統計学者たちは、この結果は出来すぎていて信じがたいと指摘している。3408 回の試行と注意深く選んだ針の長さのもとでは、$\frac{355}{113}$ はちょうど 1808 回の交差から生じるのであり、ちょうどよいところで実験を打ち切ればこの値が得られてしまう。これは都合のよい結果だけを報告することへの戒めとして、古典的な教訓話となっている。

ビュフォン

ビュフォンは博物学者であり、パリの王立植物園の園長を務め、全 36 巻の博物誌の著者でもあった。彼は 1733 年に針の問題（および正方形のタイルを用いた版である「フラン・カロー」（franc-carreau）遊び）を科学アカデミー（Académie des Sciences）に提出し、1777 年にその解を発表した。彼のねらいは、算術だけでなく幾何学も確率に用いることができると示すことであった。ここから積分幾何学（integral geometry）、すなわちランダムな直線と、それらが曲線や曲面と交わる様子を扱う理論が生まれた。これは現在、ステレオロジー（stereology）、すなわち 2 次元の断面から 3 次元の構造（組織中の細胞など）の大きさや形を推定する科学において用いられている。

思考のスキル

領域上の積分としての確率

結果がいくつかの独立な連続確率変数に依存するとき、ある事象の確率は、その事象が起こる領域の上で同時確率密度関数を積分したものである。

大学では、これが連続確率論の基礎となる。同時確率密度関数（joint density）、周辺分布（marginal distribution）、変数変換（change of variables）であり、これらはすべて Section 12.21 と Section 12.22 で再び現れる。

次の一歩

1. **長い針。** $a > b$ ならば、針は 2 本以上の直線と交わることがある。それでも交わる回数の期待値は $\frac{2a}{\pi b}$ であることを示せ。
2. **麺。** 長さ $2a$ の針を直角に曲げる（長さ a の 2 本の腕をもつ）。交わる回数の期待値はいくらか。
3. **正方形のタイル。** 半径 r の硬貨を、1 辺 $s > 2r$ の正方形のタイルを敷いた床の上に投げる。硬貨が完全に 1 枚のタイルの内部に収まる確率を求めよ。
4. **Lazzarini の針。** $a = \frac{5}{6}b$（したがって $\frac{2a}{\pi b} = \frac{5}{3\pi}$）のとき、3408 回の試行で 1808 回交わると、推定値 $\pi \approx \frac{355}{113}$ が得られることを示せ。

さらに学ぶために

- **Daniel Klain & Gian-Carlo Rota, *Introduction to Geometric Probability***（Cambridge University Press）。ビュフォンの針から話が始まる。
- **Lee Badger,「Lazzarini's lucky approximation of π」**, *Mathematics Magazine* 67 (1994)。疑わしい実験の統計学。
- **Paul Nahin, *Digital Dice***（Princeton University Press）。確率の問題に対するモンテカルロ法。

12.21 棒を折る

問題

STEP III 1990、問 16

先に進む前に、まず問題を読んでおこう。問題は STEP Support Programme の問題データベースにある：

`step.maths.org/questions/90-s3-q16`

導入

棒をランダムな 2 点で折る。3 本の断片で三角形を作れるだろうか。有名な答えは「確率（probability）$\frac{1}{4}$ で作れる」である。では、作れたとき、その三角形は鋭角三角形（acute triangle）だろうか、それとも鈍角三角形（obtuse triangle）だろうか。たいていは鈍角である。できる三角形のうち鋭角三角形はおよそ 3 分の 1 にすぎない。ランダムな図形に関するこの種の問題は 19 世紀以来、確率論の研究者たちを楽しませてきた。また、連続確率（continuous probability）を扱う格好の練習の場でもある。領域内のランダムな点、部分領域を定める条件、そしてその面積を求める積分がそろっているからである。

この問題では棒の折り方を 2 通り考え、それぞれについて鈍角が現れる確率を問う。答えは $3-2\sqrt{2}$ と $3-4\ln 2$ であり、単純な幾何から意外な数が現れる。

取りかかる前に

考えてみよう 3 つの長さで三角形が作れるのは、どの長さも他の 2 つの和より小さいとき、かつそのときに限る。長さの和が L であるとき、これは最も長いものについて何を意味するか。辺 c の対角が鈍角になるのはどのようなときか。

$a+b+c=L$ のもとで $c<a+b$ は $c<\frac{L}{2}$ を意味する。したがって 3 本の断片で三角形が作れるのは、どの断片も全長の半分より短いときにちょうど一致する。余弦定理（cosine rule）により、c の対角が鈍角になるのは $c^2>a^2+b^2$ のときにちょうど一致す

る。

順を追って解く

段階 1　1 回切ってから半分にする

断片は X, $\frac{1-X}{2}$, $\frac{1-X}{2}$ である。等しい 2 本はつねに $\frac{1}{2}$ より短いので、三角形ができるのは $X < \frac{1}{2}$ のとき、かつそのときに限る。この三角形は二等辺三角形（isosceles triangle）であり、等しい 2 角は鋭角であるから、鈍角三角形になるのは X の対角が鈍角のときに限る：

$$X^2 > 2\Big(\frac{1-X}{2}\Big)^2 = \frac{(1-X)^2}{2}, \qquad \text{すなわち} \qquad \sqrt{2}X > 1-X, \quad X > \frac{1}{1+\sqrt{2}} = \sqrt{2}-1.$$

$X < \frac{1}{2}$ という条件のもとで、X は $[0, \frac{1}{2}]$ 上の一様分布（uniform distribution）に従うので、

$$\mathrm{P}(\text{鈍角} \mid \text{三角形}) = \frac{\frac{1}{2} - (\sqrt{2}-1)}{\frac{1}{2}} = 3 - 2\sqrt{2} \approx 0.172. \qquad \blacksquare$$

段階 2　2 つ目の作り方

> **考えてみよう**　頂点が $(1,0)$, $(1,1)$, $(0,1)$ の三角形内の (x,y) について、辺 x, y, $2-x-y$ に対する 3 つの三角不等式（triangle inequality）を確かめよ。

領域は $\{x \le 1, y \le 1, x+y \ge 1\}$ である。不等式は、$x+y \ge 2-x-y$ すなわち $x+y \ge 1$、$x+(2-x-y) \ge y$ すなわち $y \le 1$、$y+(2-x-y) \ge x$ すなわち $x \le 1$ である。これらはすべて領域内で成り立つ（領域の辺上では等号が成り立ち、そこでは三角形は退化（degenerate）している）。 $\blacksquare$

> **考えてみよう**　B における角は辺 $CA = y$ の対角である。条件 $y^2 > x^2 + (2-x-y)^2$ を y についての不等式に書き直せ。

展開すると $y^2 > x^2 + 4 + x^2 + y^2 - 4x - 4y + 2xy$ となるので $4y - 2xy > 2x^2 - 4x + 4$、すなわち $2y(2-x) > 2(x^2 - 2x + 2)$ である。$2 - x > 0$ であるから、

$$\angle ABC \text{ が鈍角} \iff y > \frac{x^2 - 2x + 2}{2 - x}. \qquad \blacksquare$$

段階 3　面積

考えてみよう x^2-2x+2 を $2-x$ で割れ。次に、領域のうち曲線より上にある部分の面積を求め、領域の面積 $\frac{1}{2}$ で割れ。

$x^2-2x+2=-x(2-x)+2$ であるから、曲線は $y=-x+\frac{2}{2-x}$ である。この曲線は $(0,1)$ から $(1,1)$ まで延び、$0<x<1$ では直線 $y=1-x$ より上、$y=1$ より下にある（図 12.21.1）。鈍角となる領域は曲線と $y=1$ の間にあり、その面積は

$$\int_0^1 \left(1+x-\frac{2}{2-x}\right)\mathrm{d}x = 1+\frac{1}{2}+\left[2\ln(2-x)\right]_0^1 = \frac{3}{2}-2\ln 2.$$

$\frac{1}{2}$ で割って、$\mathrm{P}(\angle ABC \text{ が鈍角}) = 3-4\ln 2 \approx 0.227$ を得る。 ■

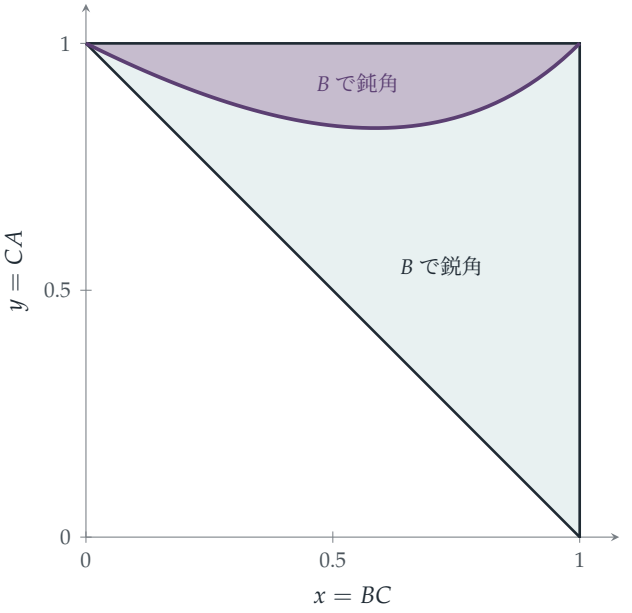


図 12.21.1 小問 (ii) における (x,y) のとりうる範囲を表す三角形。B における角は曲線 $y=\frac{x^2-2x+2}{2-x}$ より上（紫色の領域）で鈍角となり、この領域は三角形全体の $3-4\ln 2 \approx 0.23$ を占める。

振り返り

- **幾何を不等式に翻訳する。**「三角形を作る」も「鈍角である」も、長さの間の不等式になった。
- **連続変数に対する条件付き確率（conditional probability）。**$X<\frac{1}{2}$ という条件のもとで、X はより短い区間上の一様分布に従った。
- **面積の比としての確率。**領域内に一様に分布する点については、確率は面積を全体

の面積で割ったものである。

背後にある数学

古典的な折れた棒の問題

棒を、独立（independent）に一様分布する 2 点 U, V で折る。組 (U, V) は単位正方形上に一様に分布し、3 本の断片で三角形が作れるのは各断片が $\frac{1}{2}$ より短いときにちょうど一致する。この条件は合計面積 $\frac{3}{4}$ の 4 つの小さな三角形を切り取るので、確率は $\frac{1}{4}$ となる。三角形ができたという条件のもとで、それが鋭角三角形である確率は $\frac{1}{4}$ のうちの $3\ln 2 - 2 \approx 0.079$、すなわち約 32% である。これらの断片は**一様ランダム分割**（uniform random partition）の一例であり、その同時分布は**ディリクレ分布**（Dirichlet distribution）である。ディリクレ分布はベイズ統計学（Bayesian statistics）や、資源がランダムに分配される様子のモデルに現れる。

棒の折り方が結果を左右する

問題の 2 つの方法は、よく似て聞こえる問いに対して $3 - 2\sqrt{2}$ と $3 - 4\ln 2$ という異なる答えを与え、しかもどちらも古典的な方法とは異なることに注意しよう。ベルトラン（Bertrand）の弦の問題（Section 12.3）と同様に、手順を指定しない限り「ランダム」という言葉には何の意味もない。

ランダムな三角形

ランダムな三角形に関する問題は、ヴィクトリア朝時代に人気のあったパズルである。ルイス・キャロル（Lewis Carroll）の *Pillow Problems*（1893 年）は、「ランダムに」選んだ三角形が鈍角三角形である確率を問うている。その答えは、本節の答えと同じく、彼が選んだランダムさのモデルに完全に依存する。ランダムな三角形に関する問題はその後も驚きを生み続けてきた。たとえば、正方形内に一様ランダムにとった 3 点が鈍角三角形をなす確率は約 0.725 であり、これが厳密に計算されたのは 20 世紀になってからである。

思考のスキル

標本空間（sample space）の幾何

ランダムな結果を領域内の点として、事象を部分領域として表し、確率を面積の比（あるいは密度（*density*）の積分）として計算する。

大学では、この見方が多変量確率論（multivariate probability）の基礎となり、また座標を取り替えたときに密度がどう変換されるかを教えてくれる変数変換公式（change-of-variables formula）の基礎にもなる。

次の一歩

1. **古典的な問題。**棒を独立に一様分布する 2 点で折るとき、断片で三角形が作れる確率は $\frac{1}{4}$ であることを示せ。

2. **長いほうの断片を折る。**棒を一様分布する 1 点で折り、次に長いほうの断片を一様分布する 1 点で折る。3 本の断片で三角形が作れる確率は $2\ln 2 - 1 \approx 0.386$ であることを示せ。

3. **鋭角か鈍角か。**小問 (ii) において、三角形 ABC が C において（$AB = 2 - x - y$ の対角として）鈍角をもつ確率を求めよ。これを用いて、どこかに鈍角をもつ確率を求めよ。

4. **最長の断片の期待値。**古典的な折れた棒の問題において、最も長い断片の長さの期待値（expected value）は $\frac{11}{18}$ であることを示せ。

さらに学ぶために

- **Martin Gardner, *The Colossal Book of Mathematics***（Norton）。折れた棒の問題を収録している。
- **Frederick Mosteller, *Fifty Challenging Problems in Probability***（Dover）。
- **Lewis Carroll, *Pillow Problems***（Dover、*A Tangled Tale* と合本）。ランダムな三角形を含むヴィクトリア朝時代のパズル集。

12.22 平均をもたない分布

問題

STEP III 1995、問 12

先に進む前に、まず問題を読んでおこう。問題は STEP Support Programme の問題データベースにある：

`step.maths.org/questions/95-s3-q12`

導入

多数の測定値の平均をとれば、偶然誤差は打ち消し合い、平均は真の値に落ち着いていく。この**大数の法則**（law of large numbers）はあらゆる統計学の土台である。しかし、そこには隠れた仮定がある。測定値が有限な平均をもたなければならないのである。この問題では、ごく自然な確率変数、すなわち独立な 2 つの標準正規確率変数の比を構成するが、これは平均をまったくもたない。その平均値は決して落ち着かない。実際、それを 100 万個平均したものは、1 個だけのものとまったく同じ分布をもつのである。

これが**コーシー分布**（Cauchy distribution）である。その確率密度関数は無害な釣鐘形に見えるが、裾が非常に重いため、ときおり現れる巨大な値がどんな平均をも支配してしまう。この問題では、美しい幾何学的議論によってこの分布を導く。座標の比は直線の傾きであり、放射対称性によってその直線の角度は一様に分布するのである。

取りかかる前に

? **考えてみよう** (U, V) が、確率密度が原点からの距離だけに依存するような確率的な点であるとき、原点からその点へ引いた直線が軸となす角 Θ について何がいえるか。

平面を回転しても確率密度は変わらないので、どの方向も同様に確からしい。すなわち角 Θ は $[0, 2\pi)$ 上の一様分布に従う。そして $\frac{U}{V}$ は方向だけに依存し、距離には依存しない。

順を追って解く

段階 1　同時確率密度関数

X と Y は独立であるから、それらの同時確率密度関数（joint density）はそれぞれの確率密度関数の積である。

$$\mathrm{f}(x,y) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \cdot \frac{1}{\sqrt{2\pi}} e^{-y^2/2} = \frac{1}{2\pi} e^{-\frac{1}{2}(x^2+y^2)}.$$

これは x と y に、原点からの距離の 2 乗である x^2+y^2 を通してのみ依存する。つまり、確率密度は原点を中心とする円の上で一定である。これが「放射対称（radially symmetrical）」の意味である。

段階 2　対称性を利用する

> **考えてみよう**　$\frac{U}{V} < k$ となる場合は、V の符号に応じて 2 通りある。180° の回転 $(u,v) \mapsto (-u,-v)$ によって、これらはどのように結びつくか。

$\frac{U}{V} < k$ とは、$V > 0$ かつ $U < kV$ であるか、または $V < 0$ かつ $U > kV$ であることを意味する。回転 $(u,v) \mapsto (-u,-v)$ は 1 つ目の領域を 2 つ目の領域に写し、確率密度を保つ。したがって両者の確率は等しく、

$$\mathrm{P}\left(\frac{U}{V} < k\right) = 2\mathrm{P}(U < kV, V > 0). \qquad \blacksquare$$

段階 3　角度

> **考えてみよう**　u を横軸、v を縦軸とする平面に領域 $\{V > 0, U < kV\}$ を図示せよ。これは原点を頂点とする扇形の領域である。その角の大きさはいくらか。

領域 $V > 0$ は上半平面である。原点を通る直線 $u = kv$ は、正の v 軸と角 $\arctan k$ をなす（$k > 0$ のときは正の u の側に傾く）。上半平面のうちこの直線より左側の部分は、角 $\frac{\pi}{2} + \arctan k$ の扇形の領域である。放射対称性により、扇形の領域の確率はその角を 2π で割ったものであるから、

$$\mathrm{P}\left(\frac{U}{V} < k\right) = 2 \cdot \frac{\frac{\pi}{2} + \arctan k}{2\pi} = \frac{1}{2} + \frac{1}{\pi} \arctan k.$$

これを k で微分すると、確率密度関数

$$\mathrm{g}(k) = \frac{1}{\pi(1+k^2)}. \qquad \blacksquare$$

が得られる。

振り返り

- **独立な正規分布は放射対称である。** 2 つの正規分布の確率密度関数の積は距離だけに依存する。
- **比は方向を表す。** $\frac{U}{V}$ は点の角度だけに依存し、その角度は一様分布に従う。
- **扇形の領域の確率。**放射対称性があれば、扇形の領域の確率はその角に比例する。

背後にある数学

平均が存在しない

コーシー分布の平均は $\int_{-\infty}^{\infty} \frac{k}{\pi(1+k^2)}\,\mathrm{d}k$ となるはずであるが、$\int_0^K \frac{k}{1+k^2}\,\mathrm{d}k = \frac{1}{2}\ln(1+K^2) \to \infty$ である。この積分は収束（converge）せず、平均は存在しない。（対称性から答えは 0 だと思うかもしれないが、正の部分も負の部分もともに無限大であり、$\infty - \infty$ には値がない。）分散も無限大である。

さらに悪いことに、独立な n 個のコーシー確率変数の平均は再びコーシー分布に従い、しかもまったく同じ分布をもつ。（これを確かめる 1 つの方法は、特性関数（characteristic function）$\mathrm{E}(e^{itX}) = e^{-|t|}$ を用いるものである。）したがって平均をとってもまったく役に立たない。100 万個の測定値の平均は、1 個の測定値より少しも信頼できないのである。図 12.22.1 はその様子を示している。正規分布の標本の逐次平均は落ち着いていくが、コーシー分布の標本の逐次平均は、ときおり現れる巨大な値に揺さぶられて、いつまでもふらつき続ける。

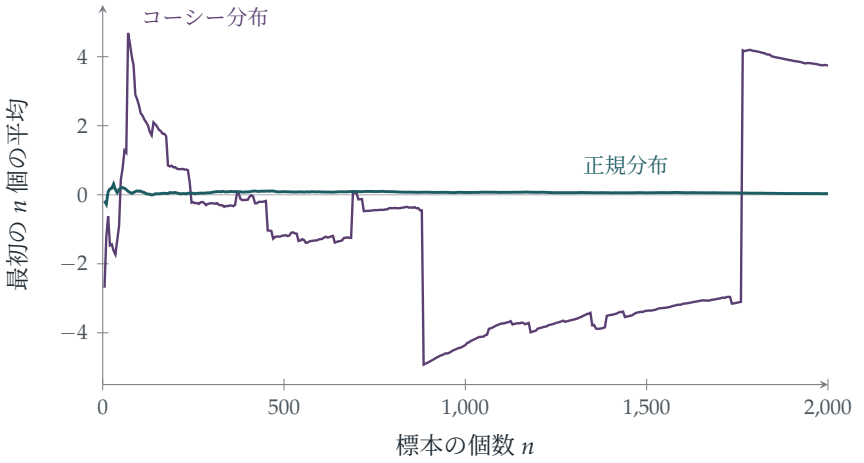


図 12.22.1　独立な 2000 個の標本の逐次平均。標準正規分布の標本（ティール）では、大数の法則が予言するとおり、平均は 0 に落ち着いていく。コーシー分布の標本（プラム）では決して落ち着かない。ときおり現れるたった 1 つの巨大な値が、平均をどこか新しい場所へ放り出してしまうのである。

コーシー確率変数はどこから現れるか

一様に無作為な方向に向けた懐中電灯で壁を照らすと、壁に映る光点の位置はコーシー分布に従う。物理学では、同じ確率密度関数が**ローレンツ関数**（Lorentzian）とよばれ、スペクトル線や共鳴ピークの形を記述する（Section 11.1 と比較せよ）。金融の分野では、正規分布によるモデルが大幅に過小評価してしまうときおりの暴落をモデル化するために、コーシー分布のような裾の重い分布が用いられる。統計学にとっての教訓は深刻である。平均や分散に基づく手法は裾の重いデータに対しては破綻し、中央値（コーシー分布に対してもまったく問題なく振る舞う）を用いるなどの頑健な手法が必要となるのである。

ポアソン、コーシー、そして論争

この分布は 1824 年にポアソン（Siméon Denis Poisson）によって研究され、彼はそれが大数の法則を満たさないことに気づいていた。コーシー（Augustin-Louis Cauchy）は 1853 年、最小二乗法をめぐるビエネメ（Irénée-Jules Bienaymé）との公開論争の中でこの分布に出会った。コーシーはこれを用いて、誤差の裾が重い場合には最小二乗法が劣った方法になりうることを示したのである。名前はコーシーに定着したが、これは Stephen Stigler が「スティグラーの名祖の法則」とよぶものの一例である。すなわち、科学上の発見は最初の発見者の名前ではよばれない、という法則である。

思考のスキル

期待値が存在することを確かめる

平均や分散を用いる前に、それを定義する積分が収束することを確かめよ。対称性があっても、発散する積分が 0 になるわけではない。

大学では、可積分性の条件が確率論のあらゆる定理の一部となっている。大数の法則には有限な平均が、中心極限定理（Section 12.24）には有限な分散が必要である。コーシー分布は、それらの条件が成り立たないときの標準的な反例である。

次の一歩

1. **懐中電灯。** 長くまっすぐな壁から距離 1 の位置にある懐中電灯が、壁への垂線と角 Θ をなす方向を向いている。ただし Θ は $(-\frac{\pi}{2}, \frac{\pi}{2})$ 上の一様分布に従う。光点の位置 $\tan\Theta$ がコーシー分布に従うことを示せ。

2. **逆数。** C がコーシー分布に従うならば、$\frac{1}{C}$ もコーシー分布に従うことを示せ。（ヒント：$\frac{U}{V}$ と $\frac{V}{U}$。）

3. **中央値はうまく働く。** コーシー分布の中央値は 0 であり、四分位数は ± 1 であることを示せ。

4. **半径と角度。** X, Y を独立な標準正規確率変数とするとき、極座標に変換することにより、$R^2 = X^2 + Y^2$ が平均 2 の指数分布に従い、かつ角度と独立であることを示せ。（これは正規乱数を生成するボックス＝ミュラー法の基礎である。）

さらに学ぶために

- **Geoffrey Grimmett & Dominic Welsh, *Probability: An Introduction*** (Oxford University Press)。確率変数の変換について。
- **Nassim Nicholas Taleb, *The Black Swan*** (Penguin)。邦題『ブラック・スワン』。裾の重い分布とそれがなぜ重要なのかを、一般読者向けに解説している。
- **Stephen Stigler, *The History of Statistics*** (Harvard University Press)。コーシーとビエネメの論争も扱っている。

12.23 確率から導くスターリングの公式

問題

STEP II 2016、問 13

先に進む前に、まず問題を読んでおこう。問題は STEP Support Programme の問題データベースにある：

`step.maths.org/questions/16-s2-q13`

導入

100! はどれくらい大きいのだろうか。実際に掛け算をすると時間がかかる。答えは 158 桁である。**スターリングの公式**（Stirling's formula）$n! \approx \sqrt{2\pi n}\,(n/e)^n$ を使えば、たった 1 行で誤差 0.1% 以内の値が得られる。これは数学で最も役に立つ近似の一つであり、確率論、統計物理学、アルゴリズムの解析で絶えず用いられている。

証明は数多くあり、積分を使うもの（Section 4.4）や、オイラー＝マクローリンの公式（Euler–Maclaurin formula）を使うものがある。この問題が与えるのは、その中でも最も意外な証明、すなわち確率による証明である。平均 n のポアソン分布（Poisson distribution）は近似的に正規分布（normal distribution）に従い、両者をピークで比較すると $n!$ の値が決まってしまう。スターリングの公式に現れる $\sqrt{2\pi}$ は、文字どおり正規分布の密度に現れる $\sqrt{2\pi}$ なのである。

取りかかる前に

考えてみよう 平均（mean）μ、分散（variance）σ^2 の離散確率変数 X が近似的に正規分布に従うとき、$\mathrm{P}(X = \mu)$ はおよそいくらか。（μ を中心とする幅 1 の区間の確率を考えよ。）

正規近似（normal approximation）では $\mathrm{P}(X = \mu) \approx \mathrm{P}(\mu - \frac{1}{2} < X < \mu + \frac{1}{2})$ であり、これはおよそ幅 1 にピークでの密度（density）を掛けたもの、すなわち $\frac{1}{\sigma\sqrt{2\pi}}$ である。

順を追って解く

段階 1　16 枚の硬貨

$X \sim \mathrm{B}(16, \tfrac{1}{2})$ の平均は 8、分散は $16 \cdot \tfrac{1}{2} \cdot \tfrac{1}{2} = 4$ であるから、$\sigma = 2$ である。正規近似により、

$$\mathrm{P}(X = 8) \approx \frac{1}{2\sqrt{2\pi}} \approx 0.1995.$$

正確な値は $\binom{16}{8}/2^{16} = \tfrac{12870}{65536} \approx 0.1964$ であり、誤差は 2% 以内である。

段階 2　中心二項係数

> **考えてみよう**　$\mathrm{B}(2n, \tfrac{1}{2})$ の平均と標準偏差（standard deviation）は何か。$\mathrm{P}(X = n)$ を正確な形と近似の形で書け。

平均は n、分散は $\tfrac{n}{2}$ であるから、$\sigma = \sqrt{n/2}$ である。正確には $\mathrm{P}(X = n) = \binom{2n}{n}\frac{1}{2^{2n}}$ であり、近似的には $\frac{1}{\sqrt{n/2}\sqrt{2\pi}} = \frac{1}{\sqrt{\pi n}}$ である。両者を等しいとおくと、

$$\frac{(2n)!}{(n!)^2\, 2^{2n}} \approx \frac{1}{\sqrt{\pi n}}, \qquad (2n)! \approx \frac{2^{2n}(n!)^2}{\sqrt{n\pi}}. \qquad \blacksquare$$

段階 3　スターリングの公式

> **考えてみよう**　パラメータ n のポアソン分布は、平均と分散がともに n に等しい。$\mathrm{P}(X = n)$ を正確な形と近似の形で書け。

正確には $\mathrm{P}(X = n) = e^{-n}\frac{n^n}{n!}$ である。n が大きいとき、ポアソン分布は近似的に正規分布に従い（これは n 個の独立なポアソン (1) 確率変数の和である）、$\sigma = \sqrt{n}$ であるから、$\mathrm{P}(X = n) \approx \frac{1}{\sqrt{2\pi n}}$ となる。両者を等しいとおくと、

$$n! \approx \sqrt{2\pi n}\, e^{-n} n^n. \qquad \blacksquare$$

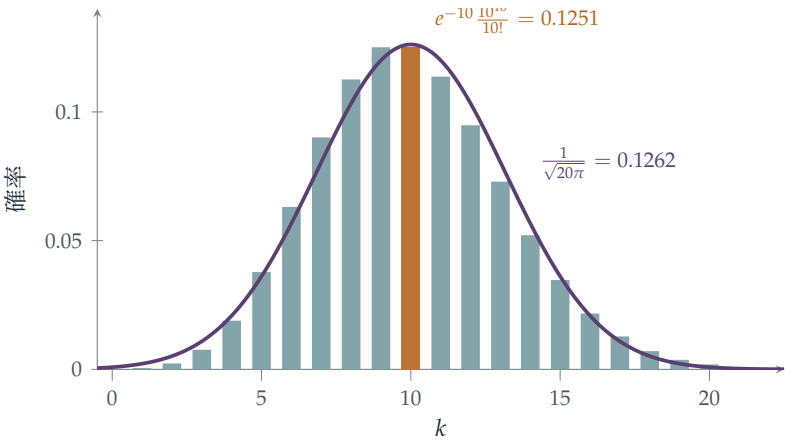


図 12.23.1　平均 10 のポアソン分布（棒）と、同じ平均・分散をもつ正規分布の密度（紫）。ピークにおいて、ポアソン分布の確率 $e^{-10}10^{10}/10!$（琥珀色）は正規分布の密度 $\frac{1}{\sqrt{2\pi\cdot 10}}$ に近い。このほぼ等しいという関係を変形したものがスターリングの公式である。

振り返り

- **1 点での確率 ≈ 密度 × 幅。** 格子分布（lattice distribution）では、ある点における正規分布の密度がその点の確率を近似する。
- **正確な式と近似が出会う。** 二項分布とポアソン分布の確率は正確な値である。それらを正規近似と等しいとおくことで、階乗の評価が得られた。
- **式に合わせて分布を選ぶ。** $\binom{2n}{n}$ からは $\mathrm{B}(2n, \frac{1}{2})$ が、$\frac{n^n}{n!}$ からはポアソン (n) が思い浮かぶ。

背後にある数学

近似の精度はどれくらいか

この議論はヒューリスティック（heuristic）である。正規近似を *1 点*で用いているが、通常の中心極限定理（central limit theorem）はそれを正当化しない。その厳密版である**局所中心極限定理**（local central limit theorem）がまさにそれを保証し、この議論を本物の証明にする。実際、スターリングの公式は驚くほど精度が高い。比 $\frac{n!}{\sqrt{2\pi n}(n/e)^n}$ はおよそ $1 + \frac{1}{12n}$ であり、$n = 1$ のときでさえ 1 の代わりに 0.92 を与え、$n = 10$ のときには 0.8% 小さいだけである。完全な級数 $1 + \frac{1}{12n} + \frac{1}{288n^2} - \cdots$ は**漸近級数**（asymptotic series）の一例である。n が大きいほどよい近似を与えるが、n を固定すると最終的に

は発散する。

階乗と正規分布はなぜ結びつくのか

スターリングの公式に現れる $\sqrt{2\pi}$ と正規分布の密度に現れるものは同じ定数である。$\int_{-\infty}^{\infty} e^{-x^2/2}\,\mathrm{d}x = \sqrt{2\pi}$ だからである。アブラーム・ド・モアブル（Abraham de Moivre）は 1733 年、まさにこの方法で二項分布の正規近似を発見した。そのためには n が大きいときの $\binom{n}{k}$ の近似が必要だったのである。彼は $n! \approx Cn^{n+1/2}e^{-n}$ という形を見いだしたが、定数までは求められなかった。ジェームズ・スターリング（James Stirling）が $C = \sqrt{2\pi}$ であることを示した。正規分布は階乗の漸近挙動から生まれたのである。

ド・モアブルとスターリング

アブラーム・ド・モアブルはフランスのプロテスタントで、ロンドンに亡命し、家庭教師や賭博師への助言で生計を立てていた。彼は 1718 年に『*The Doctrine of Chances*』（偶然の理論）を出版し、1733 年の小冊子で二項分布の正規近似を発見した。これが正規曲線の最初の登場である。スコットランドの数学者ジェームズ・スターリングは、著書『*Methodus Differentialis*』（1730 年）で定数の値を発表し、ド・モアブルは寛大にもその功績を彼に帰した。階乗の近似は今も活発な研究分野であり、巨大な n に対する $n!$ の計算は組合せ論や暗号理論で必要とされている。

思考のスキル

確率とは関係のないことを確率で証明する

正確な確率を近似と比べると、式に現れる数についての情報が得られることがある。

この「確率論的方法」（probabilistic method）は両方向に働く。ここでは確率が階乗の評価を与えた。エルデシュ（Erdős）はこの方法を用いてグラフの存在を証明した（Section 9.2）。大学では、解析学、数論、組合せ論の定理を証明するのに確率が使われる場面に出会うだろう。

次の一歩

1 **公式を確かめる。** スターリングの公式を用いて 10! と 20! を見積もり、真の値 3 628 800 および 2 432 902 008 176 640 000 と比較せよ。

2 **(iii) から (ii) へ。** 小問 (iii) のスターリングの公式から小問 (ii) の近似が導かれることを示せ。

3 **100! の桁数。** $\log_{10} n! \approx n \log_{10} \frac{n}{e} + \frac{1}{2} \log_{10}(2\pi n)$ を用いて、100! が 158 桁であることを示せ。

4 **硬貨投げの引き分け。** 小問 (ii) を用いて、公平な硬貨を 1000 回投げたとき表がちょうど 500 回出る確率がおよそ 0.025 であることを示せ。

さらに学ぶために

- **William Feller, *An Introduction to Probability Theory and Its Applications*, Vol. 1**（Wiley）、第 II 章。スターリングの公式を証明している。
- **Ronald Graham, Donald Knuth & Oren Patashnik, *Concrete Mathematics***（Addison-Wesley）、第 9 章。漸近解析を扱う。
- **Stephen Stigler, *The History of Statistics***（Harvard University Press）。ド・モアブルと正規曲線の誕生について。

12.24 中心極限定理

問題

STEP III 2008、問 12

先に進む前に、まず問題を読んでおこう。問題は STEP Support Programme の問題データベースにある：

`step.maths.org/questions/08-s3-q12`

導入

互いに独立な小さな確率的効果を数多く足し合わせると、個々の分布が何であれ、その合計は釣鐘型の正規分布（normal distribution）の曲線に従う。これが**中心極限定理**（central limit theorem）であり、確率論でおそらく最も重要な定理である。身長や測定誤差や試験の点数がおおよそ正規分布に従う理由も、統計学者がもとの分布について何も知らなくても、その分散（variance）が有限でありさえすれば平均値がどれほど信頼できるかを述べられる理由も、この定理で説明される（Section 12.22 と比較せよ）。

この問題では、特別な場合にこの定理を証明する。1 匹のカエルが両側指数分布（two-sided exponential distribution）に従ってランダムに跳ぶ。その密度は鋭く尖ったピークをもち、釣鐘型とは似ても似つかない。それでも何度も跳んだあとのカエルの変位の合計は正規分布に従う。道具となるのは**積率母関数**（moment generating function）である。積率母関数は独立な確率変数の和を積に変え、さらに対数をとれば積は和に変わる。

出題範囲についての注意

積率母関数は 2019 年に STEP の出題範囲から外された。確率変数 X の積率母関数は $\mathrm{M}_X(\theta) = \mathrm{E}(e^{\theta X})$ である。その 0 における導関数から積率（モーメント）が得られる：$\mathrm{M}'(0) = \mathrm{E}(X)$、$\mathrm{M}''(0) = \mathrm{E}(X^2)$。独立な X と Y に対しては $\mathrm{M}_{X+Y} = \mathrm{M}_X \mathrm{M}_Y$ であり、また $\mathrm{M}_{cX}(\theta) = \mathrm{M}_X(c\theta)$ である。積率母関数が収束すれば、分布も（適切な意味で）収束する。

取りかかる前に

考えてみよう　密度 $\frac{1}{2}e^{-|x|}$ は対称である。そのグラフはどのような形か。平均（mean）はいくらか。

これはテントのような形の曲線で、0 に高さ $\frac{1}{2}$ の鋭いピークをもち、両側に指数関数的に減衰する（図 12.24.1）。対称性より平均は 0 である。

順を追って解く

段階 1　積率母関数

$|\theta| < 1$ のとき、積分を 0 で分けると

$$\mathrm{M}_X(\theta) = \frac{1}{2}\int_{-\infty}^{0} e^{(1+\theta)x}\,\mathrm{d}x + \frac{1}{2}\int_{0}^{\infty} e^{-(1-\theta)x}\,\mathrm{d}x = \frac{1}{2}\Big(\frac{1}{1+\theta} + \frac{1}{1-\theta}\Big) = \frac{1}{1-\theta^2}. \qquad \blacksquare$$

展開すると $\mathrm{M}_X(\theta) = 1 + \theta^2 + \theta^4 + \cdots$ となるので、$\mathrm{E}(X) = 0$ であり、$\mathrm{E}(X^2) = 2$（$\frac{\theta^2}{2!}$ の係数）である。分散は 2 である。

段階 2　カエルの変位

考えてみよう　$Y = X_1 + \cdots + X_n$ の積率母関数は何か。$Y/\sqrt{2n}$ の積率母関数はどうか。

独立性より $\mathrm{M}_Y(\theta) = (1-\theta^2)^{-n}$ であり、尺度を変えると

$$\mathrm{M}_{Y/\sqrt{2n}}(\theta) = \mathrm{M}_Y\Big(\frac{\theta}{\sqrt{2n}}\Big) = \Big(1 - \frac{\theta^2}{2n}\Big)^{-n}.$$

となる。（Y の標準偏差（standard deviation）である $\sqrt{2n}$ で割ることで、分散が 1 になる。）

段階 3　極限

考えてみよう　対数をとり、小さな u に対して $\ln(1-u) \approx -u$ を用いよ。

$\ln \mathrm{M}_{Y/\sqrt{2n}}(\theta) = -n\ln\big(1 - \frac{\theta^2}{2n}\big) = -n\big(-\frac{\theta^2}{2n} - \frac{\theta^4}{8n^2} - \cdots\big) \to \frac{\theta^2}{2}$ であるから

$$\mathrm{M}_{Y/\sqrt{2n}}(\theta) \to e^{\theta^2/2}, \qquad \blacksquare$$

となり、これは標準正規分布の積率母関数である。

段階 4 カエルの確率

考えてみよう n が大きいとき、Y は近似的に $\mathrm{N}(0, 2n)$ に従う。$\mathrm{P}(|Y| \geq 25) = 0.05$ となるのはいつか。

標準正規分布に従う Z に対して $\mathrm{P}(|Z| \geq 1.96) = 0.05$ である。よって $\frac{25}{\sqrt{2n}} = 1.96$ が必要であり、$2n = \left(\frac{25}{1.96}\right)^2 \approx 162.7$、すなわち $n \approx 81.3$ を得る。遠くに着地する確率は n とともに増えるので、少なくとも 5% の確率を与える最小の跳躍回数はおよそ 82 である。 ■

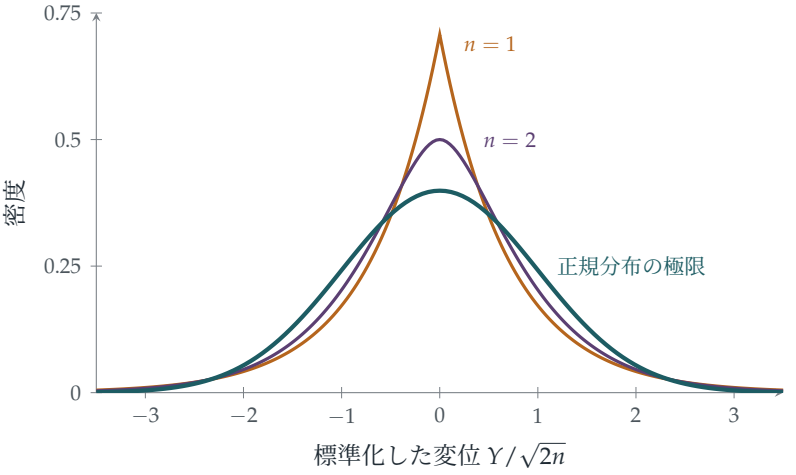


図 12.24.1 $n = 1$ 回跳んだあと（琥珀色：鋭いラプラス分布のピーク）と $n = 2$ 回跳んだあと（紫色）のカエルの標準化した変位の密度を、標準正規分布の密度（青緑色）と比べたもの。2 回跳んだだけで、すでにピークは丸みを帯びている。中心極限定理によれば、$n \to \infty$ のときこれらの曲線は正規分布の曲線に収束する。

振り返り

- **変換は和を積に変える。**独立な確率変数の和の積率母関数は、それぞれの積率母関数の積である。
- **対数は積を和に変える。**$\ln\left(1 - \frac{\theta^2}{2n}\right)^{-n}$ は、極限の求めやすい和になった。
- **標準化せよ。**標準偏差 $\sqrt{2n}$ で割ることで分散が 1 に保たれ、極限が 0 にも無限大にもならなかった。

背後にある数学

一般の定理

同じ議論は、平均が 0 で有限の分散 σ^2 をもつ任意の分布に対して成り立つ（積率母関数の存在も仮定するが、この仮定は外すことができる）。このとき $\mathrm{M}(\theta) = 1 + \tfrac{1}{2}\sigma^2\theta^2 + o(\theta^2)$ であり、

$$\left(\mathrm{M}\left(\frac{\theta}{\sigma\sqrt{n}}\right)\right)^n = \left(1 + \frac{\theta^2}{2n} + o\left(\frac{1}{n}\right)\right)^n \to e^{\theta^2/2}.$$

となる。極限で生き残るのは分散だけであり、もとの分布の形は忘れ去られる。正規分布が普遍的である理由はここにある。大学では、この定理は常に存在する**特性関数**（characteristic function）$\mathrm{E}(e^{i\theta X})$ を用いて証明され、分布の収束はレヴィ（Lévy）の連続性定理によって厳密に述べられる。

収束の速さは？

中心極限定理は極限で何が起こるかを述べるが、どれだけ速く収束するかは述べない。**ベリー・エッセーンの定理**（Berry–Esseen theorem）は、誤差を定数と $\frac{\mathrm{E}|X|^3}{\sigma^3\sqrt{n}}$ の積で上から評価する。カエルの分布のような対称な分布では、数十回跳べば近似は良好になる。歪んだ分布や、分布の裾の極端な部分では、n が大きくても近似が悪いことがある。正規近似は非常にまれな事象の確率を過小評価するのであり、この事実は保険や金融に影響を及ぼす。

ド・モアブルからリンデベルグまで

ド・モアブル（de Moivre）は 1733 年に二項分布の正規近似を発見し（Section 12.23）、ラプラス（Laplace）は 1810 年にそれをより一般の確率変数の和へと拡張して、測定誤差は多くの小さな効果の和であるという考え方を打ち出した。厳密な証明が得られたのはずっと後のことである。1901 年にアレクサンドル・リャプノフ（Aleksandr Lyapunov）が、そして 1922 年にヤール・リンデベルグ（Jarl Lindeberg）が証明を与え、リンデベルグは定理が成り立つための本質的に最も弱い条件を見いだした。「中心極限定理」という名前は 1920 年にジョージ・ポリア（George Pólya）によって付けられたもので、この定理が確率論の中心にあるという意味である。

思考のスキル

変換し、簡単にし、変換して戻す

難しい演算（独立な確率変数を足すこと。これは密度の畳み込みにあたる）は、変換を施すと易しい演算（掛け算）になる。変換した先で計算を行い、その結果を解釈せよ。

同じ戦略は、母関数（Section 9.4）、解析学や工学におけるフーリエ（Fourier）変換やラプラス変換、そして掛け算を足し算に変える対数の根底にもある。

次の一歩

1 **2 回跳んだ場合の厳密な計算。** $Y = X_1 + X_2$ の密度が $\frac{1}{4}(1 + |y|)e^{-|y|}$ であることを示し、その分散が 4 であることを確かめよ。

2 **二項分布から正規分布へ。** $X \sim \mathrm{B}(1, \frac{1}{2})$ を平均が 0 になるように平行移動したもの（値 $\pm\frac{1}{2}$ をとる）について、$\mathrm{M}_X(\theta) = \cosh\frac{\theta}{2}$ を求め、そのような確率変数 n 個の和を標準化したものの積率母関数が $e^{\theta^2/2}$ に近づくことを示せ。

3 **カエル、厳密に。** コンピュータを使って（あるいは検算として $n = 2$ のときの厳密な密度を使って）、$n = 82$ のとき正規近似がどれほど正確かを調べよ。

4 **なぜコーシー分布ではだめなのか。** コーシー分布（Section 12.22）は積率母関数をもたない。議論のどの段階が破綻するか。また、n 個のコーシー分布に従う確率変数の平均はどうなるか。

さらに学ぶために

- **Geoffrey Grimmett & Dominic Welsh, *Probability: An Introduction***（Oxford University Press）。母関数と極限定理の章。
- **Hans Fischer, *A History of the Central Limit Theorem***（Springer）。中心極限定理の歴史。ラプラスからリンデベルグまで。
- **David Spiegelhalter, *The Art of Statistics***（Pelican）。統計学の技法。正規分布がいたるところに現れる理由について。

12.25 多数の中の最大値

問題

STEP III 2025、問 11

先に進む前に、まず問題を読んでおこう。問題は、OCR が配布している 2025 年の STEP 過去問一式（zip ファイル）のうち、STEP III の問題用紙にある：

`www.ocr.org.uk/Images/750486-2025-past-papers.zip`

導入

次の百年で最大の洪水はどれほどの高さになるだろうか。千個の電球のうち最も長もちするものは、どれだけの時間もつだろうか。n 人が出場するマラソンで最も速い走者は、どれほど速いだろうか。多数の確率的な値のうち最大のものや最小のものといった**極値**（extremes）に関する問いは、工学、保険、気候科学できわめて重要であり、その振る舞いは平均とはまったく異なる。

この問題では、n 個の指数分布（exponential distribution）に従う寿命の最大値の分布を求める。その典型的な大きさは $\ln n$ のように、ゆっくりとではあるが限りなく増大する。しかし、そのばらつきはまったく増大しない。千個の寿命の最大値は、おおよそ十個の寿命の最大値と同じ程度に予測可能なのである。極限において、その分布の形は**ガンベル分布**（Gumbel distribution）となる。これは正規分布の曲線に対応する、極値版の分布である。

取りかかる前に

? **考えてみよう** $X_1, \ldots, X_n$ の最大値が y 以下になるのは、ちょうどどのようなときか。

$Y \leq y$ となるのは、ちょうどすべての $X_i \leq y$ となるときである。独立性（independence）により、これは積に変わる。

順を追って解く

段階 1　最大値の分布

各 X_i は $y \ge 0$ に対して $\mathrm{F}(y) = 1 - e^{-\lambda y}$ をみたし、独立性より

$$\mathrm{G}(y) = \mathrm{P}(X_1 \le y, \dots, X_n \le y) = (1 - e^{-\lambda y})^n. \qquad \blacksquare$$

段階 2　分位点

> ? **考えてみよう**　$\mathrm{G}(L) = \alpha$ を L について解け。次に $\mathrm{G}(U) = 1 - \alpha$ を解け。

$(1 - e^{-\lambda L})^n = \alpha$ から $e^{-\lambda L} = 1 - \alpha^{1/n}$ が得られるので、

$$L(\alpha) = -\frac{1}{\lambda} \ln\bigl(1 - \alpha^{1/n}\bigr), \qquad U(\alpha) = -\frac{1}{\lambda} \ln\bigl(1 - (1-\alpha)^{1/n}\bigr). \qquad \blacksquare$$

段階 3　n が大きい場合

> ? **考えてみよう**　$\alpha^{1/n} = e^{(\ln \alpha)/n}$ と書け。n が大きいとき、指数は小さい。$1 - \alpha^{1/n}$ はおよそいくらか。

$\alpha^{1/n} = e^{(\ln\alpha)/n} \approx 1 + \frac{\ln \alpha}{n}$ であるから、$1 - \alpha^{1/n} \approx \frac{-\ln\alpha}{n} = \frac{\ln(1/\alpha)}{n}$ となる。したがって

$$\lambda L(\alpha) \approx -\ln \frac{\ln(1/\alpha)}{n} = \ln n - \ln\Bigl(\ln \frac{1}{\alpha}\Bigr). \qquad \blacksquare$$

同様に、$\lambda U(\alpha) \approx \ln n - \ln\bigl(\ln \frac{1}{1-\alpha}\bigr)$ である。

段階 4　増大とばらつき

中央値（median）は $L(\frac{1}{2})$ であり、$\lambda L(\frac{1}{2}) \approx \ln n - \ln\ln 2 \to \infty$ となる。しかし

$$\lambda\bigl(U(\alpha) - L(\alpha)\bigr) \approx \ln\Bigl(\ln \frac{1}{\alpha}\Bigr) - \ln\Bigl(\ln \frac{1}{1-\alpha}\Bigr),$$

となり、$\ln n$ の項は打ち消し合っている。つまり、幅は n によらない。 $\blacksquare$

段階 5　90% 区間

> ? **考えてみよう**　$\alpha = 0.05$ とせよ。$\ln \frac{1}{\alpha} = \ln 20$ と $\ln \frac{1}{0.95}$ はおよそいくらか。

$\alpha = 0.05$ のとき $\ln 20 \approx 3$ であるから、$\ln\ln \frac{1}{\alpha} \approx \ln 3$ である。また $\ln \frac{1}{0.95} = -\ln(1 - 0.05) \approx 0.05 = \frac{1}{20}$ であるから、$\ln\ln \frac{1}{1-\alpha} \approx \ln \frac{1}{20} \approx -3$ である。したがって幅はおよそ

$\frac{1}{\lambda}(\ln 3 + 3) \approx \frac{4.1}{\lambda}$ であり、Y は確率 0.9 で、幅およそ $4\lambda^{-1}$ の区間 $[L(0.05), U(0.05)]$ に入る。 ■

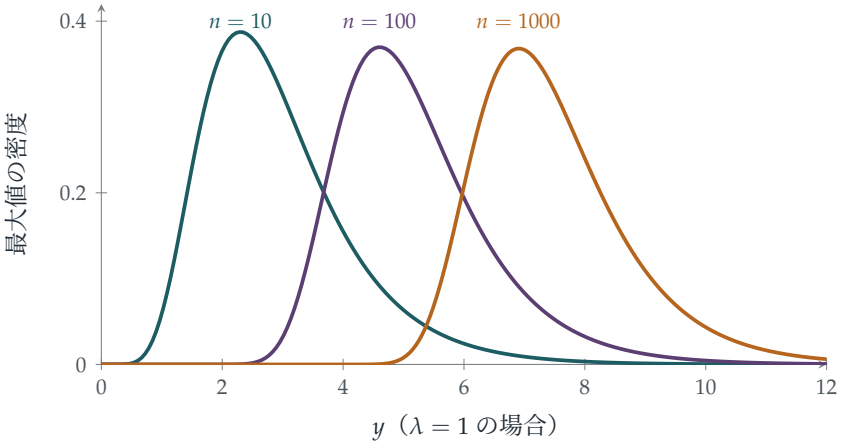


図 12.25.1　$n = 10$、100、1000 の場合の、n 個の指数分布に従う寿命（平均 1）の最大値の密度。n が 10 倍になるごとに曲線は右へ $\ln 10 \approx 2.3$ だけずれるが、その形と幅は本質的に変わらない。これは $\ln n$ だけずらしたガンベル分布である。

振り返り

- **最大値が y 以下であるのは、すべてがそうであるとき。**独立性によって、最大値の分布はべき乗になった。
- **逆に解いて分位点を求める。**$G(y) = \alpha$ を解くことで、区間の端点の厳密な式が得られた。
- **近似してから比較する。**L と U の近似はどちらも $\ln n$ を含み、それが幅では打ち消し合った。

背後にある数学

ガンベル分布

$P(\lambda Y - \ln n \le x) = \left(1 - \frac{e^{-x}}{n}\right)^n \to \exp(-e^{-x})$ であるから、ずらした最大値 $\lambda Y - \ln n$ は、極限において分布関数（distribution function）$\exp(-e^{-x})$ をもつ**ガンベル分布**に従う。その平均はオイラー（Euler）の定数 $\gamma \approx 0.577$ である。この分布にはすでに出会っている。n 種類すべてのクーポンを集めるまでの時間（Section 9.11）は、n 個の（ほぼ独立な）待ち時間の最大値であり、それがおよそ $n \ln n + \gamma n$ となり、n のオーダ

ーのガンベル型のゆらぎをもつのはこのためである。

極値理論

フィッシャー（Fisher）、ティペット（Tippett）、グネジェンコ（Gnedenko）による注目すべき定理（1928–1943）によれば、独立同分布な多数の確率変数の最大値は、適切にずらし尺度を変えると、極限において 3 通りの形しかとりえない。すなわち、ガンベル型（指数分布や正規分布のような軽い裾（tail）をもつ場合）、フレシェ（Fréchet）型（コーシー（Cauchy）分布のような重い裾をもつ場合）、ワイブル（Weibull）型（明確な上限をもつ変数の場合）である。これは中心極限定理（central limit theorem）（Section 12.24）の極値版にあたる。技術者はこれを用いて「百年に一度の波」に耐える防潮施設や海洋プラットフォームを設計し、保険会社は巨大災害補償の価格を決め、気候科学者は熱波がどのように変化しているかを推定する。

ガンベルと洪水

エミール・ユリウス・ガンベル（Emil Julius Gumbel）はドイツの統計学者であり、ナチスに公然と反対したために 1932 年に大学の職を追われ、最終的にアメリカ合衆国へ亡命した人物である。彼は 1930 年代から 1940 年代にかけて、極値理論を河川の氾濫に応用した。その著書 *Statistics of Extremes*（『極値統計学』、1958 年）によって、この分野は実用的な道具となった。1953 年の北海大洪水の後にオランダを守るために建設されたデルタ計画の施設は、極値統計を用いて、およそ 10 000 年に一度と見込まれる高潮に耐えるよう設計された。

思考のスキル

最大値と最小値の分布

独立な確率変数の最大値については $\mathrm{P}(\max \le y) = \prod \mathrm{P}(X_i \le y)$、最小値については $\mathrm{P}(\min > y) = \prod \mathrm{P}(X_i > y)$ である。

この一つの考え方で、最寄りのスーパーマーケット（Section 12.19）、最初に来るバス（Section 12.18）、最も長もちする電球を扱うことができる。大学では、これが順序統計量（order statistics）と極値理論（extreme value theory）の出発点となる。

次の一歩

1 **最大値の期待値。** $\mathrm{E}(Y) = \frac{1}{\lambda}\left(1 + \frac{1}{2} + \cdots + \frac{1}{n}\right)$ を示せ。（**ヒント：**無記憶性（memorylessness）により、n 個の寿命のうち最初の一つが尽きた後、残りの $n-1$ 個は新たに始まったのと同じである。）

2 **最小値。** $X_1, \ldots, X_n$ の最小値が、パラメータ $n\lambda$ の指数分布に従うことを示せ。

3 **中央値を厳密に。** Y の厳密な中央値を求め、$n = 10$ の場合について近似値 $\frac{1}{\lambda}(\ln n - \ln\ln 2)$ と比較せよ。

4 **一様分布の最大値。** $X_1, \ldots, X_n$ が $[0,1]$ 上の一様分布（uniform distribution）に従うとき、その最大値 M の分布を求め、n が大きいとき $n(1-M)$ が近似的に指数分布に従うことを示せ。これは極値の 3 つの形のうちどれか。

さらに学ぶために

- **Stuart Coles, *An Introduction to Statistical Modeling of Extreme Values***（Springer）。洪水や海面水位の実データを用いた標準的な入門書。
- **Emil Julius Gumbel, *Statistics of Extremes***（Dover）。古典的名著。
- **David Spiegelhalter, *The Art of Uncertainty***（Pelican）。まれな事象と、それについてどう考えるべきかを論じた本。

学んだこと

本章の冒頭で、不確実性は正確に推論できるが、直感はそうはいかないと述べた。そして三つの問いを立てた。問題は正確には何を問うているのか、うまい切り口は何か、長い目で見ると何が起こるのか、である。本章の 25 問が示したことを以下にまとめる。

1. 問題は正確には何を問うているのか

最初の四つの節は、問題を正しく捉えることを扱った。「番号の付いた円板が取り出されたとき」と「3 番の円板が取り出されたとき」とでは答えが異なった（Section 12.1）。まれな病気に対する精度 90% の検査では、陽性と出た患者が実際に病気である確率はわずか 9% であった。基準率（base rate）が重要だからである（Section 12.2）。「無作為な弦」は、どのように選ぶかを定めるまで長さの期待値（expected value）をもたなかった（Section 12.3）。折れた棒の問題（Section 12.21）も同じことを示した。さらに「独立（independent）」にはいくつかの意味があることがわかった。事象は二つずつでは独立でも、全体としては独立でないことがある（Section 12.4）。また、無相関（uncorrelated）な確率変数が独立であるとは限らない。いずれの場合も、処方箋は、どのような結果が起こりうるか、それらがどのように生じるかを正確に書き出すことであった。

2. うまい切り口は何か

多くの問題には、長い計算と短い洞察の両方があった。

- **対称性（symmetry）。**過程が区別できない二つの座席（Section 12.5）、互いに補い合う鏡像としての「表が多い」と「裏が多い」（Section 12.6）、どの二人の選手もどの試合で対戦する可能性も等しいこと（Section 12.7）、どんな硬貨でも HT と TH が等確率であること（Section 12.9）、方向が一様に分布する回転対称な点（Section 12.22）。
- **指示変数（indicator）と線形性（linearity）。**組（Section 12.4）、結び目（Section 12.10）、試合（Section 12.7）を指示変数の和として数えることで、分布全体を求めずに平均と分散（variance）が得られた。
- **状態と第一段階解析（first-step analysis）。**Penney のゲーム（Penney's game）（Section 12.8）、ウサギの巣穴と Ecks 教授（Section 12.16）、家名の断絶（Section 12.15）は、いずれも次の一歩で条件付ける（conditioning）ことで解けた。

- **戦略。**三つ巴の決闘（truel）では、最も腕の悪い者はわざと外すべきである（Section 12.13）。ランダム化された戦略は相手に付け込まれない（Section 12.14）。候補者の最初の $\frac{1}{e}$ を見送るのが最適である（Section 12.12）。

3. 長い目で見ると何が起こるのか

本章の後半では、時間の経過に沿って、また多数回の試行にわたって、ランダムネスを追った。分枝過程（branching process）は絶滅するか爆発的に増えるかのいずれかであり、爆発的に増えうる過程でさえ偶然によって絶滅することがある（Section 12.15）。マルコフ（Markov）連鎖（Markov chain）は定常分布（stationary distribution）に落ち着き、過去を忘れる（Section 12.16）。指数分布に従う待ち時間は記憶をまったくもたない（Section 12.18）。平均はふつう一定の値に落ち着くが、コーシー分布（Cauchy distribution）ではそうならない。コーシー分布には平均が存在しないからである（Section 12.22）。多数の独立な効果の和は、中心極限定理（central limit theorem）により正規分布に近づく（Section 12.24）。そしてその正規性の中には、$n!$ に関するスターリング（Stirling）の公式が潜んでいる（Section 12.23）。最大値はまた違った振る舞いをする。最大値は $\ln n$ のように増大するが、そのばらつきはまったく増大しない（Section 12.25）。そして定数 e、π、γ が、カップや切手、針や輪の問題に繰り返し現れた。

4. 確率の道具箱

手法	使った場面
標本空間（sample space）を書き出す	条件付け（12.1）、ベルトラン（Bertrand）（12.3）
ベイズ（Bayes）の定理とオッズ（odds）	基準率（12.2）、マルコフ連鎖の逆転（12.16）
対称性	12.5、12.6、12.7、12.22
指示変数と線形性	12.4、12.7、12.10
第一段階解析	Penney（12.8）、巣穴（12.16）、分枝過程（12.15）
余事象（complement）と評価	公平な決定（12.9）
期待値のための級数	切手と e（12.11）
パラメータ表示と最適化	秘書問題（12.12）
ゲーム理論的な推論	三つ巴の決闘（12.13）、混合戦略（mixed strategies）（12.14）
終わりに至る経路を数える	マッチ箱（12.17）
まず分布関数（distribution function）を求める	12.18、12.19、12.25
面積や積分としての確率	ビュフォン（Buffon）（12.20）、折れた棒（12.21）
正規近似（normal approximation）	スターリング（12.23）、中心極限定理（12.24）

この先へ

- 第 13 章では確率から統計へと移り、データから未知の量を推定することを扱う。まずは Section 13.1 から始める。
- Section 9.8、Section 9.9、Section 9.11 の数え上げの問題は、本章の組合せ論的な側面にあたる。
- マルコフ連鎖では第 7 章の線形代数を用い、分枝過程の不動点では Section 8.7 の力学系を用いる。

大学入学前にさらに先へ進みたい人へ

- **Joseph Blitzstein & Jessica Hwang, *Introduction to Probability***（確率論入門）：すばらしい最初の一冊。オンラインで無料公開されており、講義動画もある。
- **Geoffrey Grimmett & Dominic Welsh, *Probability: An Introduction***（確率論入門）：簡潔かつ厳密で、大学 1 年次の水準である。

- **Frederick Mosteller, *Fifty Challenging Problems in Probability***（確率の難問 50 題）：古典的なパズルを完全な解答付きで収録している。
- **David Spiegelhalter, *The Art of Statistics***（統計の技法）および ***The Art of Uncertainty***（不確実性の技法）：現実世界における確率とリスクを扱う。

ラプラス（Laplace）は、確率とは計算に帰着された常識であると述べた。本章が示したのは、まさに常識がうまく働かないからこそ計算が必要になることが多いということ、そして優れた計算は単に数値を出すだけでなく、物事を説明してくれるということである。一つの対称性、一つの指示変数、うまく選んだ状態、あるいは一つの極限定理が、パズルを洞察へと変える。そしてそうした洞察こそが、大学の確率論の講義の土台となるのである。

「次の一歩」の解答

ここに載せるのは簡単な答えとヒントであり、完全な解答ではない。結果と鍵となる一歩を示すので、自分の解答の確認に使ってほしい。答えを見る前に、まずそれぞれの問題にきちんと取り組むこと。

12.1 正確な条件付け

1. (iv) と (v) のどちらでも、残りの $k-1$ 枚の円盤は、残っている $b+n-1$ 枚（そのうち $n-1$ 枚に番号がついている）から無作為に選ばれたものである。したがって、どちらも $\frac{(n-1)\binom{b}{k-2}}{\binom{b+n-1}{k-1}}$ に等しい。(vi) の答えは $\frac{\binom{n}{2}\binom{b}{k-2}}{\binom{b+n}{k}-\binom{b}{k}}$ である。$N=b+n$ とおくと、両者が一致するのは $(2N-nk)\binom{N-1}{k-1}=2k\binom{b}{k}$ のときに限る。これが起こるのは、$(b,n,k)=(2,4,3)$、$(3,3,4)$、$(6,11,3)$ のような特殊な場合か、自明な場合として $n=2$ ですべての円盤を取り出すときだけである。問題で与えられた値ではこれは成り立たない。

2. 2 人の子どもについて、同様に確からしい（性別、曜日）の組は $14^2=196$ 通りあり、そのうち少なくとも 1 人が火曜日生まれの男の子であるものは $196-13^2=27$ 通りである。そのうち 2 人とも男の子であるものは $7^2-6^2=13$ 通りである。よって、求める確率は $\frac{13}{27}$ である。

3. ドア 1 を選んだとする。車がドア 1 の後ろにあれば（確率 $\frac{1}{3}$）、選び直すと負ける。車がドア 2 または 3 の後ろにあれば、司会者はもう一方のヤギのドアを開けるしかなく、選び直すと勝つ。その確率は $\frac{2}{3}$ である。司会者が無作為にドアを開け、たまたまヤギが出た場合には、見たことを条件とすると、「自分のドアの後ろに車がある」と「もう一方の閉じたドアの後ろに車がある」は同様に確からしく、確率は $\frac{1}{2}$ である。

4. $\mathrm{P}(\geq 2 \mid \geq 1)=\frac{1-\mathrm{P}(0)-\mathrm{P}(1)}{1-\mathrm{P}(0)}\approx 0.370$ であるが、$\mathrm{P}(\geq 2 \mid \text{スペードのエース})=1-\binom{48}{12}/\binom{51}{12}\approx 0.561$ である。**どの**エースであるかを知ると、2 枚目のエースがある可能性が高くなる。

12.2 基準率の誤謬

1. 2 人がともに「遺体がある」と言う確率は、実際にある場合は pq、ない場合は $(1-p)(1-q)$ である。事前オッズ（prior odds）が 1 : 1 のとき、事後確

率（posterior probability）は $\frac{pq}{pq+(1-p)(1-q)} = \frac{pq}{1-p-q+2pq}$ である。後半については、「A がある」は同様に確からしい 3 つの場合のうち 2 つで正しいので、$\mathrm{P}(\text{両方} \mid \text{両者の証言が} A) = \frac{pq}{2pq+(1-p)(1-q)}$ となる。

2. $0.9\pi > \frac{1}{11}(1-\pi)$、すなわち $\pi > \frac{10}{109} \approx 9.2\%$ であればよい。

3. AL の陽性結果が 1 回出るごとに、オッズは $\frac{0.9}{1/11} = 9.9$ 倍になる。事前オッズ $\frac{1}{99}$ から始めると、陽性 1 回の後はオッズ 0.1（確率 $\frac{1}{11}$）、2 回の後はオッズ 0.99（確率は約 0.497）、3 回の後はオッズ 9.8（確率は約 0.907）となる。

4. その病気をもつ 100 人のうち 99 人が陽性となり、もたない 999 900 人のうち約 9999 人が陽性となる。よって、病気の可能性があると告げられるのは約 10 100 人であり、そのうち実際に病気であるのは 99 人（約 1%）にすぎない。

12.3 「ランダム」とは何を意味するか？

1. 弦（chord）が内接正三角形の 1 辺 $\sqrt{3}a$ より長くなるのは、次のときである。遠い側の端点が、近い側の端点の反対側にある円周の 3 分の 1 の弧上にあるとき（$\frac{1}{3}$）。O からの距離が $\frac{a}{2}$ より小さいとき（$\frac{1}{2}$）。中点が半径 $\frac{a}{2}$ の円の内部にあるとき（$\frac{1}{4}$）。

2. 中点が C である弦は OC に垂直なので、C を通り角 $t = \frac{\pi}{2}$ をなす弦であり、$L(\frac{\pi}{2}) = \frac{4a}{3}\left(0 + \frac{1}{2}\sec^2\frac{\pi}{4}\right) = \frac{4a}{3}$ となる。

3. 長さは $2\sqrt{a^2-d^2}$ であり、その平均（mean）は $\frac{2}{a}\int_0^a \sqrt{a^2-d^2}\,\mathrm{d}d = \frac{2}{a}\cdot\frac{\pi a^2}{4} = \frac{\pi a}{2}$ である。

4. 端点を無作為に選ぶ場合：$\frac{2}{\pi\sqrt{4a^2-\ell^2}}$。距離を無作為に選ぶ場合：$\frac{\ell}{2a\sqrt{4a^2-\ell^2}}$。中点を無作為に選ぶ場合：$\frac{\ell}{2a^2}$。

12.4 二つずつなら独立、全体では独立でない

1. 展開すると $\mathrm{E}(XY) - \mathrm{E}(X)\mathrm{E}(Y) = (a-pq)(x_1-x_2)(y_1-y_2)$ となるので、$a = pq$ である。残りの 3 つのセルもこれに従って決まり、独立性（independence）が得られる。例：$(0,0)$ に確率 $\frac{1}{3}$、$(\pm1, \pm1)$ のそれぞれに確率 $\frac{1}{6}$ を置く。周辺分布（marginal distribution）はどちらも一様で $\mathrm{E}(AB) = 0$ であるが、$\mathrm{P}(A=0, B=1) = 0 \neq \frac{1}{9}$ である。

2. 組は 21 個あり、それぞれ確率 $\frac{1}{6}$ で一致し、その指示変数（indicator variable）は対ごとに独立（pairwise independent）である。よって、平均は 3.5、分散（variance）は $21 \cdot \frac{1}{6}\cdot\frac{5}{6} \approx 2.92$、標準偏差（standard deviation）は約 1.71 である。得点 7 は平均より標準偏差の約 2 倍だけ上にある。

3. 平均は $\binom{k}{2}\frac{1}{365}$ である。指示変数はさいころの場合とまったく同じように対ごとに独立なので、分散は $\binom{k}{2}\frac{1}{365}\cdot\frac{364}{365}$ である。（ただし、互いに独立（mutually independent）ではない。1,2 と 2,3 が一致すれば、1,3 も一致する。）

4. 同様に確からしい 8 個の結果をとり、$A = B = \{1,2,3,4\}$、$C = \{1,5,6,7\}$ とする。すると $\mathrm{P}(A \cap B \cap C) = \frac{1}{8}$ であるが、$\mathrm{P}(A \cap B) = \frac{1}{2} \neq \frac{1}{4}$ である。

12.5 最後の乗客の座席

1. すべての $n \geq 3$ について $\frac{1}{3}$ である。最後の席は S_1、S_2、S_n のいずれかとなり、どの無作為な選択もこれら 3 つを対称に扱う。

2. T_1 が S_1 に座れば、誰も自分の席を追われない。T_1 が S_3 に座れば、T_1 と T_3 が追われる。T_1 が S_2 に座れば、T_2 が選ぶことになり、S_1 を選べば追われるのは 2 人、S_3 を選べば 3 人である。期待値（expectation）は $\frac{1}{3}\left(0+2+\frac{5}{2}\right) = \frac{3}{2}$ であり、公式から得られる $\frac{2}{3}+\frac{1}{3}+\frac{1}{2}$ と一致する。

3. $r \geq 2$ に対して $P_r = \frac{1}{2}$ とすると、$\frac{1}{n}\left(1+\frac{n-2}{2}\right) = \frac{1}{2}$ となるので、漸化式（recurrence relation）は $P_n = \frac{1}{2}$ と矛盾しない。

4. 最後に空いている席は必ず S_1 か S_n である（それ以外の S_k は、遅くとも T_k が搭乗する時点までに埋まっている）。よって、それが S_n でなければ S_1 である。

12.6 もう 1 枚の硬貨

1. Bob の最初の $n+1$ 回の投げを H'、最後の 1 回を X とする。Bob が勝つのは、$H' > H_A$ のとき、または $H' = H_A$ かつ X が表のときである。よって確率は $\mathrm{P}(H' > H_A) + \frac{1}{2}\mathrm{P}(H' = H_A)$ であり、硬貨が 1 枚多い場合の結果から $\mathrm{P}(H' > H_A) = \frac{1}{2}$ である。

2. いいえ。$n = 0$ のとき、Bob は確率 p で勝つ。鮮やかな証明は表と裏を入れ替えるが、それが対称性（symmetry）になるのは公平な硬貨の場合だけである。

3. $H_A = H_B$ となるのは、ちょうど $H_A + (\text{Bob の裏の回数}) = n$ のときであり、これは $\mathrm{B}(2n, \frac{1}{2})$ に従う確率変数が n に等しいということである。値は $\frac{1}{2}$、$\frac{3}{8}$、$\frac{5}{16}$ である。

4. いいえ。6 の目は $p = \frac{1}{6}$ の硬貨のようなものである。$n = 0$ のとき確率は $\frac{1}{6}$ であり、一般に $\frac{1}{2}$ より小さい。

12.7 勝ち抜き戦で誰と誰が当たるか

1. 第 k ラウンドには 2^{n-k} 試合があり、対称性から、各試合が $\binom{2^n}{2}$ 組のどの組の対戦になるかは同様に確からしい。大会全体の $2^n - 1$ 試合について足し合わせると、$\frac{2^n-1}{\binom{2^n}{2}} = \frac{2}{2^n}$ となる。
2. いいえ。固定された組み合わせ表に無作為に配置しても対称性は保たれ、同じ数え方が適用できる。
3. 第 k ラウンドでの X の対戦相手は 2^{k-1} 人のグループから出てきて、その各人がグループを勝ち抜く可能性は等しい。よって、Y がその対戦相手である確率は $\frac{2^{k-1}}{2^n-1} \cdot \frac{1}{2^{k-1}} = \frac{1}{2^n-1}$ であり、n ラウンド全体での合計は $\frac{n}{2^n-1}$ である。
4. 試合数は $\binom{2n}{2} = n(2n-1)$ である。すべての組が対戦するので、確率は 1 である。

12.8 ペニーのゲーム

1. HHT が勝つのは、最初の 2 回が HH の場合だけである。そうでなければ HH が揃う前に T が出て、THH が先に現れるしかない。よって、HHT が勝つ確率は $\frac{1}{4}$ である。4 人全員で行う場合：HH の後は HHT が必ず勝ち、TT の後は TTH が必ず勝つ。HT と TH からは、HTT と THH の勝つ確率は $\frac{2}{3}$ と $\frac{1}{3}$（またはその逆）なので、各プレイヤーの勝つ確率は $\frac{1}{4} \cdot 1$ または $\frac{1}{4}\left(\frac{2}{3} + \frac{1}{3}\right)$、すなわち $\frac{1}{4}$ である。TTH は確率 $\frac{2}{3}$ で THH に勝つ（小問 (i)）。
2. HH について：$e_0 = 1 + \frac{1}{2}e_{\mathrm{H}} + \frac{1}{2}e_0$ かつ $e_{\mathrm{H}} = 1 + \frac{1}{2}e_0$ より、$e_0 = 6$ である。HT について：いったん H が出たら T を待てばよく（平均 2 回）、最初の H までにも平均 2 回かかるので、4 である。
3. HHH と TTT は、THH と HTT に確率 $\frac{7}{8}$ で負ける。HHT と TTH は、THH と HTT に確率 $\frac{3}{4}$ で負ける。残りの 4 つ（HTH、HTT、THH、THT）は、それぞれ HHT、HHT、TTH、TTH に確率 $\frac{2}{3}$ で負ける。最もたやすく負かされるのは HHH と TTT である。
4. $A = \mathrm{HTT}$、$B = \mathrm{HHT}$ とすると、相関（correlation）は $AA = 4$、$AB = 0$、$BB = 4$、$BA = 2$ である。B が勝つオッズは $(AA - AB) : (BB - BA) = 4 : 2$、すなわち確率 $\frac{2}{3}$ である。

12.9 偏った硬貨から公平な判定を

1. ラウンド数は幾何分布（geometric distribution）に従い、その平均は小問 (i) では $\frac{1}{p^2+q^2}$、フォン・ノイマン（von Neumann）の方法（HT または TH）では $\frac{1}{2pq}$ である。$p^2 + q^2 - 2pq = (p-q)^2$ なので、$p \neq \frac{1}{2}$ であればつねにフォン・ノイマ

ンの方法のほうが遅い。ただし、公平なのはこちらである。

2. フォン・ノイマンの方法を使って公平なビットを作る。それを 2 つずつ取り、00、01、10 で 3 つの選択肢を選び、11 ならやり直す。

3. 投げた結果を 1 回ずつ、二進展開（binary expansion）$\frac{1}{3} = 0.010101\ldots$ と比べる（H を 1、T を 0 と読む）。初めて異なる回で、投げた結果が桁の数字より**小さい**（1 に対して T）なら「表」、そうでなければ「裏」と宣言する。こうすると、ちょうど確率 $\frac{1}{3}$ で表となり、投げる回数は平均 2 回である。

4. 決着のつかないラウンドの後には、一方の面が 2 枚、もう一方の面が 1 枚出ており、この過程はこの 2 つの状態の上のマルコフ連鎖（Markov chain）になる。これを解くと $\mathrm{P}(\text{ホテル}) = \frac{p^3}{p^3+q^3}$ となり、$p = \frac{1}{2}$ のときは $\frac{1}{2}$ である。そのときは公平であるが、それ以外では公平でない。

12.10 ひもの端を結ぶ

1. 最初の結びで輪ができる確率は $\frac{1}{3}$ であり（他の 3 つの端のうち 1 つが同じひもの端である）、最後の結びでは必ず輪ができる。輪の数は確率 $\frac{1}{3}$ で 2、確率 $\frac{2}{3}$ で 1 であり、平均は $\frac{4}{3}$ である。

2. ひもが k 本残っているとき、結びで輪ができない確率は $1 - \frac{1}{2k-1}$ である。積 $\prod_{k=2}^{n} \frac{2k-2}{2k-1}$ は $\frac{4^n}{2n\binom{2n}{n}}$ に等しく、これはおよそ $\frac{\sqrt{\pi n}}{2n} = \sqrt{\frac{\pi}{4n}}$ である。

3. 要素 $1, 2, \ldots, n$ を順に挿入する。要素 k は新しい巡回（cycle）を始める（確率 $\frac{1}{k}$）か、それより前の $k-1$ 個の要素のどれかの後に置かれる。指示変数を足し合わせると H_n が得られる。

4. 順に行う結びのそれぞれについての輪の指示変数は、確率 $\frac{1}{2k-1}$ の独立な変数なので、分散は $\sum \frac{1}{2k-1}\left(1 - \frac{1}{2k-1}\right)$ である。また、$\sum_{k\geq 1} \frac{1}{(2k-1)^2} = \frac{\pi^2}{6} - \frac{1}{4} \cdot \frac{\pi^2}{6} = \frac{\pi^2}{8}$ である。

12.11 切手、カップ、そして e

1. $\sum_{k\geq 2} \frac{k-1}{k!} t^k = t\mathrm{e}^t - (\mathrm{e}^t - 1) = 1 + (t-1)\mathrm{e}^t$ である。微分すると $\mathrm{E}(N) = \mathrm{G}'(1) = \mathrm{e}$、$\mathrm{E}(N(N-1)) = \mathrm{G}''(1) = 2\mathrm{e}$ となるので、$\mathrm{Var}(N) = 3\mathrm{e} - \mathrm{e}^2$ である。

2. $\mathrm{P}(N = j)$ は $k = 0, \ldots, j-1$ に対する $\mathrm{P}(N > k)$ に現れる、すなわち j 回現れる。ここで $\mathrm{P}(N > k) = \frac{1}{k!}$ なので、$\mathrm{E}(N) = \sum \frac{1}{k!} = \mathrm{e}$ である。

3. まったく変わらない。カップが 1 つなら、Jim は必ず正しく「当てる」。価格の期待値はやはり e である。

4. $\int_0^1 \frac{(1-u)^2}{2}\,\mathrm{d}u = \frac{1}{6}$ である。一般に、$u_k = u$ での切り口は $1-u$ 倍に縮小された $(k-1)$ 次元単体（simplex）なので、$V_k = V_{k-1}\int_0^1 (1-u)^{k-1}\,\mathrm{d}u = \frac{V_{k-1}}{k}$ である。

12.12 いつ探すのをやめるか

1. 一番大きいものが位置 $k \geq 2$ にあるとき、それが手に入るのは、ちょうど最初のアイスクリームが最初の $k-1$ 個の中で一番大きいときであり、その確率は $\frac{1}{k-1}$ である。k について平均すると $\frac{1}{n}\sum_{k=2}^{n}\frac{1}{k-1}$ となり、$n=4$ のときこれは $\frac{11}{24}$ である。24 通りの並びを書き出すと、2 番目、3 番目、4 番目に大きいものが手に入る確率は、それぞれ $\frac{7}{24}$、$\frac{1}{6}$、$\frac{1}{12}$ である。

2. $r = 0,1,2,3$ 人を見送ると、確率はそれぞれ $\frac{1}{4}$、$\frac{11}{24}$、$\frac{5}{12}$、$\frac{1}{4}$ となる。よって、1 人を見送るのがよく、成功確率は $\frac{11}{24}$ である。

3. 最初の n 人の中で最良の者が上位 2 人に入る確率は $\frac{n}{N} + \frac{n}{N}\cdot\frac{N-n}{N-1}$ であり、n が小さいときは $\frac{n}{N}$ のほぼ 2 倍である。上位 2 人のどちらでもよいなら、決断する前に見る候補者の数を少なくしてよい。（この目標に対する最良の戦略の成功確率は約 0.57 である。）

4. $\mathrm{P}(r) = \frac{r}{N}\sum_{k=r}^{N-1}\frac{1}{k}$ とすると、$N=10$ では $r=2,3,4$ に対して 0.366、0.399、0.398 となり、$N=100$ では $r=36,37,38$ に対して 0.3710、0.3710、0.3708 となって、$r=37$ がわずかに上回る。

12.13 三つ巴の決闘

1. Algernon が外した場合、Bertie は Cuthbert を狙って投げる。当たれば（$\frac{3}{5}$）、Algernon と Bertie の対決で Algernon が先に投げることになり、Bertie が勝つ確率は $1 - \frac{2/5}{2/5+3/5-6/25} = \frac{9}{19}$ である。外せば Cuthbert が Bertie に当て、その後は Algernon が先に当てない限り（$\frac{3}{5}$）Cuthbert が勝つ。よって、Bertie が勝つ確率は $\frac{3}{5}\cdot\frac{9}{19} = \frac{135}{475}$、Cuthbert が勝つ確率は $\frac{2}{5}\cdot\frac{3}{5} = \frac{114}{475}$ である。

2. すべての $a<1$ に対してである。わざと外すと $\frac{3}{5}\cdot\frac{a}{a+\frac{3}{5}-\frac{3}{5}a} + \frac{2}{5}a$ となる。Cuthbert に当てると $1 - \frac{3/5}{3/5+a-\frac{3}{5}a}$ となり（その後は Bertie が Cuthbert を狙って先に投げる）、これはつねに小さい。Bertie に当てると 0 である。

3. いいえ。Bertie が Algernon を狙って当てれば、次に Cuthbert が Bertie に当てる。外しても、やはり Cuthbert が Bertie に当てる。どちらにしても Bertie には何も残らず、Cuthbert を狙えば得られる $\frac{27}{95}$ に及ばない。

4. $p + (1-p)(1-r)p$ であり、これは等比級数全体の和 $\frac{p}{p+r-pr}$ より小さい。

12.14 混合戦略

1. $b = (\frac{1}{4}, \frac{1}{2}, \frac{1}{4})$ のとき $X = Y = Z = 0$ なので、B はどの戦略に対しても 0 を得る。$(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ に対しては、A の得点の期待値は予想 1、2、3 に対してそれぞれ $-\frac{1}{3}$、0、$\frac{1}{3}$ なので、A はつねに 3 と予想すべきである。

2. 利得行列（payoff matrix）は対角成分が 0、それ以外が ± 1 であり、3 つの戦略を巡回的に入れ替えても変わらないので、最適戦略は $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ である。問題では、差が 1 の場合と 2 の場合で罰点が異なるので、戦略は入れ替え可能ではない。

3. $X, Y, Z \geq 0$ かつ $X + 2Y + Z = 0$ ならば、3 つとも 0 である。これより $b_1 = b_3$、$b_2 = 2b_3$ が導かれる。

4. 今度は $X = -b_2 + 3b_3$、$Y = b_1 - b_3$、$Z = -3b_1 + b_2$ であり、$X + 3Y + Z = 0$ である。3 つすべてを 0 とおくと $(\frac{1}{5}, \frac{3}{5}, \frac{1}{5})$ が得られる。

12.15 絶滅か爆発か

1. 絶滅確率（extinction probability）は、$[0, 1]$ における $s = (q + ps)^2$ の最小の根である。根は 1 と $\left(\frac{q}{p}\right)^2$ であり、後者が 1 より小さくなるのはちょうど $p > \frac{1}{2}$ のときである。

2. X_{n-1} が与えられたとき、X_n の平均は $2pX_{n-1}$、分散は $2pqX_{n-1}$ なので、$\mathrm{Var}(X_n) = \mathrm{E}(2pqX_{n-1}) + (2p)^2\mathrm{Var}(X_{n-1})$ である。$p = \frac{1}{2}$ のとき、$\mathrm{Var}(X_n) = \mathrm{Var}(X_{n-1}) + \frac{1}{2}$ なので、$\mathrm{Var}(X_n) = \frac{n}{2}$ である。

3. 第 n 世代までに絶滅するとは、第 1 世代の各粒子の系統が $n - 1$ 世代以内に途絶えることである。$p = \frac{3}{4}$ のとき、$e_1 = 0.0625$、$e_2 \approx 0.0881$、$e_3 \approx 0.0999$ であり、$\frac{1}{9} \approx 0.111$ に向かって増加する。

4. $s = \frac{1}{4}(1 + s + s^2 + s^3)$ は $(s - 1)(s^2 + 2s - 1) = 0$ と因数分解できるので、絶滅確率は $\sqrt{2} - 1 \approx 0.414$ である。

12.16 マルコフ連鎖を順向きにも逆向きにも

1. 対称性から $p_B = p_D$ であり、さらに $p_A = \frac{1}{3} + \frac{2}{3}p_B$、$p_B = \frac{1}{3}p_A + \frac{1}{3}p_C$、$p_C = \frac{2}{3}p_B$ である。よって $p_A = \frac{7}{15}$ となる。A から出発して C を訪れずに A で抜け出す確率は、$x_A = \frac{1}{3} + \frac{2}{3}x_B$ かつ $x_B = \frac{1}{3}x_A$ より、$x_A = \frac{3}{7}$ である。条件付き確率は $\frac{3/7}{7/15} = \frac{45}{49}$ である。

2. 第 1 行に P を掛けると $(0.19, 0.33, 0.48)$ となる。

3. トレース（trace）は 1.3、行列式（determinant）は 0.02 であり、1 は固有値（eigenvalue）である（各行の和が 1 だから）。残りの 2 つは和が 0.3、積が 0.02 な

ので、0.2 と 0.1 である。

4. $a_1 p_{12} = \frac{1}{20}$ であるが $a_2 p_{21} = \frac{1}{15}$ なので、この連鎖は可逆（reversible）ではない。時間を逆向きにたどると、推移確率が異なる（小問 (iii) の q_{ij} は p_{ij} と異なる）。

12.17 バナッハのマッチ箱

1. 2 回目に手を伸ばしたのが 1 回目と同じ種類であれば、そこで空であることに気づく（$N = 2$）。そうでなければもう一方のお菓子を取り、3 回目に手を伸ばしたときに空の袋に気づく。よって $N = 2$ または 3 であり、確率はそれぞれ $\frac{1}{2}$ である。

2. $n = 1$ のとき：$\frac{1}{2} + \frac{1}{2} = 1$。$n = 2$ のとき：$\frac{6}{16} + \frac{6}{16} + \frac{4}{16} = 1$。

3. $\binom{2n-1}{n} 2^{-(2n-1)} = \frac{1}{2}\binom{2n}{n} \cdot 2 \cdot 4^{-n} = \binom{2n}{n} 4^{-n}$ である。

4. 残っているタフィーとミントの個数について第 1 段階で場合分けする再帰を立てると、$\frac{64}{81} \approx 0.79$ が得られる。

12.18 無記憶性

1. $\mathrm{P}(\min > t) = \mathrm{e}^{-\lambda_1 t}\mathrm{e}^{-\lambda_2 t}$ である。次のバスまでの待ち時間は、1 分あたりの率が $\frac{1}{15} + \frac{1}{30} = \frac{1}{10}$ の指数分布（exponential distribution）に従うので、その中央値（median）は $10 \ln 2 \approx 6.9$ 分である。

2. $\mathrm{P}(Y \geq k) = (1-p)^k$ なので、比は $(1-p)^n$ である。

3. t を含む間隔は、直前のバスまでさかのぼる時間と次のバスまでの時間からなり、（t が大きいとき）どちらも平均 $\frac{1}{\lambda}$ の指数分布に従うので、その長さの期待値は $\frac{2}{\lambda}$ である。長い間隔ほど、与えられた時刻を含みやすい。

4. $[0,1]$ 上で $\frac{\lambda \mathrm{e}^{-\lambda z}}{1-\mathrm{e}^{-\lambda}}$ である。$\lambda = 1$ のときは約 1.58 から 0.58 へゆるやかに減少し、$\lambda = 5$ のときは 0 の近くに集中して、約 5.03 から 0.03 へ減少する。

12.19 平面上のランダムな点

1. $1 - \mathrm{e}^{-\pi k y^2} = \frac{1}{2}$ より中央値は $\sqrt{\frac{\ln 2}{\pi k}} \approx \frac{0.470}{\sqrt{k}}$ であり、平均 $\frac{1}{2\sqrt{k}}$ より少し小さい。

2. $\mathrm{P}(Y_2 > y) = \mathrm{e}^{-\pi k y^2}(1 + \pi k y^2)$ なので、確率密度関数（probability density function）は $2\pi^2 k^2 y^3 \mathrm{e}^{-\pi k y^2}$ である。

3. 極座標（polar coordinates）では、R の確率密度関数は $\frac{1}{2\pi\sigma^2}\mathrm{e}^{-r^2/(2\sigma^2)} \cdot 2\pi r$ である。これは $\sigma^2 = \frac{1}{2\pi k}$ のときスーパーマーケットの分布と一致する。

4. $k = 4 \times 10^{-4}$ per m^2 のとき、完全に無作為な配置からは平均 $\frac{1}{2\sqrt{k}} = 25\,\mathrm{m}$ が予測される。観測された 35 m はこれより大きいので、木は密集しているのではなく、（競合によって）間隔をあけて並んでいることが示唆される。

12.20 ビュフォンの針

1. 針を短い断片に切り分ける。交差回数の期待値は加法的であり、各断片は長さに比例して寄与するので、a がいくつであっても合計は $\frac{2a}{\pi b}$ である。
2. 同じ加法性（ビュフォン（Buffon）のヌードル）により、やはり $\frac{2a}{\pi b}$ である。
3. 中心はどの辺からも r 以上離れていなければならないので、$\frac{(s-2r)^2}{s^2}$ である。
4. $\frac{1808}{3408} \approx \frac{5}{3\pi}$ より $\pi \approx \frac{5\times 3408}{3\times 1808} = \frac{17\,040}{5424} = \frac{355}{113}$ となる。出来すぎていて怪しい。

12.21 棒を折る

1. 三角形ができるには、どの断片も $\frac{1}{2}$ より短くなければならない。折る点を表す単位正方形の中で、これは面積 $\frac{1}{4}$ の領域を切り取る。
2. 最初に折る点が $x < \frac{1}{2}$ にあり（もう一方の場合も対称である）、長いほうの断片 $1-x$ を一様に折るとする。新しい 2 つの断片がどちらも $\frac{1}{2}$ より短くなければならず、その確率は $\frac{x}{1-x}$ である。よって答えは $2\int_0^{1/2} \frac{x}{1-x}\,\mathrm{d}x = 2\ln 2 - 1$ である。
3. C における角が鈍角になるのは $(2-x-y)^2 > x^2 + y^2$、すなわち $y < \frac{2-2x}{2-x}$ のときである。この曲線と $x+y=1$ の間の面積は $\frac{3}{2} - 2\ln 2$ なので、確率は $3 - 4\ln 2$ であり、B の場合と同じである。これは偶然ではない。3 辺はそれぞれ 1 以下で和が 2 となるように一様に分布しているので、互いに入れ替え可能である。よって $\mathrm{P}(\text{鈍角}) = 9 - 12\ln 2 \approx 0.68$ である。
4. 包除原理（inclusion–exclusion principle）により、最長の断片 M について $\mathrm{P}(M > m) = 3(1-m)^2 - 3(1-2m)_+^2 + (1-3m)_+^2$ が成り立つ。ここで $x_+ = \max(x, 0)$ である。$[0,1]$ で積分すると $\mathrm{E}(M) = 1 - \frac{1}{2} + \frac{1}{9} = \frac{11}{18}$ が得られる。

12.22 平均をもたない分布

1. $\mathrm{P}(\tan\Theta \le k) = \frac{1}{\pi}\left(\tan^{-1} k + \frac{\pi}{2}\right)$ であり、微分すると $\frac{1}{\pi(1+k^2)}$ が得られる。
2. U と V を入れ替えても回転対称な確率密度関数は変わらないので、$\frac{V}{U}$ は $\frac{U}{V}$ と同じ分布に従う。
3. 確率密度関数は対称なので中央値は 0 であり、$\mathrm{P}(C \le 1) = \frac{1}{2} + \frac{1}{\pi}\tan^{-1} 1 = \frac{3}{4}$ である。

4. $\frac{1}{2\pi}\mathrm{e}^{-r^2/2}\,r\,\mathrm{d}r\,\mathrm{d}\theta$ であり、$s = r^2$ とおくとこれは $\frac{1}{2\pi}\,\mathrm{d}\theta \cdot \frac{1}{2}\mathrm{e}^{-s/2}\,\mathrm{d}s$ となる。角は一様分布に従い、R^2 は平均 2 の指数分布に従い、両者は独立である。

12.23 確率から導くスターリングの公式

1. $\sqrt{20\pi}(10/\mathrm{e})^{10} \approx 3\,598\,696$（約 0.8% 小さい）、$\sqrt{40\pi}(20/\mathrm{e})^{20} \approx 2.4228 \times 10^{18}$（約 0.4% 小さい）である。
2. $\frac{(2n)!}{(n!)^2} \approx \frac{\sqrt{4\pi n}(2n/\mathrm{e})^{2n}}{2\pi n(n/\mathrm{e})^{2n}} = \frac{4^n}{\sqrt{\pi n}}$ である。
3. $100\log_{10}\frac{100}{\mathrm{e}} + \frac{1}{2}\log_{10}(200\pi) \approx 157.97$ なので、100! は 158 桁である。
4. $\binom{1000}{500}2^{-1000} \approx \frac{1}{\sqrt{500\pi}} \approx 0.0252$ である。

12.24 中心極限定理

1. 畳み込み（convolution）を計算すると $\frac{1}{4}\int \mathrm{e}^{-|x|-|y-x|}\,\mathrm{d}x = \frac{1}{4}(1+|y|)\mathrm{e}^{-|y|}$ となる。分散は $2+2=4$ である。
2. $\mathrm{E}(\mathrm{e}^{\theta X}) = \frac{1}{2}(\mathrm{e}^{\theta/2} + \mathrm{e}^{-\theta/2}) = \cosh\frac{\theta}{2}$ である。標準化された和 $\frac{2}{\sqrt{n}}\sum X_i$ の積率母関数（moment generating function）は $\cosh^n\frac{\theta}{\sqrt{n}}$ であり、その対数は $n\ln\bigl(1 + \frac{\theta^2}{2n} + \cdots\bigr) \to \frac{\theta^2}{2}$ となる。
3. Y を独立な 2 つの $\Gamma(82,1)$ 確率変数の差として表し、数値積分すると $\mathrm{P}(|Y| \geq 25) \approx 0.051$ が得られる。正規近似（normal approximation）は非常によい。
4. 積率母関数は存在せず、平均さえ存在しない。実際、n 個の独立なコーシー分布（Cauchy distribution）に従う変数の平均は、それ自身がコーシー分布に従う。平均をとっても、まったく集中しないのである。

12.25 多数の中の最大値

1. n 個の寿命のうち最初の 1 つは平均時間 $\frac{1}{n\lambda}$ の後に尽きる。その後は $n-1$ 個が残り、新たに始まったのと同じ状況になる。以下同様である。よって $\mathrm{E}(Y) = \frac{1}{\lambda}\bigl(\frac{1}{n} + \frac{1}{n-1} + \cdots + 1\bigr)$ である。
2. $\mathrm{P}(\min > t) = (\mathrm{e}^{-\lambda t})^n = \mathrm{e}^{-n\lambda t}$ である。
3. $(1-\mathrm{e}^{-\lambda y})^n = \frac{1}{2}$ より $\lambda y = -\ln(1-2^{-1/n})$ である。$n=10$ のときこれは 2.704 であり、近似値 $\ln 10 - \ln\ln 2 \approx 2.669$ と比べられる。
4. $\mathrm{P}(M \leq m) = m^n$ なので、$\mathrm{P}\bigl(n(1-M) > s\bigr) = \bigl(1-\frac{s}{n}\bigr)^n \to \mathrm{e}^{-s}$ である。これはワイブル型（Weibull type）である。変数は上に有界であり、最大値はその上界に押し寄せるように集まる。

Beyond STEP

STEP は、ほとんどの中高生がこれまでに受けるなかで最も難しい数学の試験である。そして最も優れた試験の一つでもある。その問題は短く、見慣れず、奥が深い。なかでも最良の問題は、姿を変えた大学数学の小さな断片なのである。

本書は、1987 年の第 1 回試験に先立つ見本問題（specimen papers）から 2025 年の試験問題までのなかから、特に興味深い STEP の問題を選び、その一つひとつを扉として用いる。どの問題も、優れた学生が途中で抱く疑問に触れながら丁寧に解き、そのうえで扉の向こうへと踏み出す。その考え方はどこから来たのか、大学では何へと発展するのか、次に何を読めばよいのか。

本書は、STEP に向けて準備を進める才能ある数学者の卵のうち、過去問演習以上のものを求める人のために書かれている。数学専攻の学部生のように考え始め、その先に何が待っているのかを知ってほしい。

- 純粋数学（pure mathematics）、力学（mechanics）、確率（probability）、統計（statistics）にわたる 13 の章
- STEP の過去問 136 題、それぞれに道筋を示した解答つき
- すべての問題の背後にある大学数学
- 解答つきの発展演習と、さらに学ぶための文献案内

全 13 章のうちの第 12 章。 本書全体、すべての章、すべての解答は、次のサイトで自由に閲覧・ダウンロード・共有できる：

beyondstep.keithmcnulty.org

本文はパブリックドメイン（public domain）に提供されている（CC0 1.0）。STEP の問題 © Cambridge University Press & Assessment.