



Title	確率算について : 若い学徒の為に
Author(s)	園, 信太郎
Citation	経済學研究, 74(2), 39-45
Issue Date	2024-12-05
Doc URL	https://hdl.handle.net/2115/93760
Type	departmental bulletin paper
File Information	ES_74(2)_039.pdf



確率算について

—若い学徒の為に—

園 信 太 郎

未知の記号や術語に出会った場合には、数学辞典、テキスト、想像力及び common sense に相談してもらいたい。

1. A. N. Kolmogorov による迂回

確率算の加法及び乗法法則の証明が迂回されていることの確認である。少なくとも科学においては、その原則がいかに流布していようと、人には、その証明を要求する生来の資格がある。この態度は事実上既に、ドイツの数学者 Richard Dedekind が採っている。

1.1 加法法則の迂回

Ω を集合とし $\mathcal{F}$ をその上の σ 集合体とする。つまり、

$$\begin{aligned}\mathcal{F} &\subset \mathcal{P}(\Omega); \\ \Omega &\in \mathcal{F}; \\ A \in \mathcal{F} &\Rightarrow A^c = \Omega - A \in \mathcal{F}; \\ \forall n \in \mathbb{N} (A_n \in \mathcal{F}) &\Rightarrow \bigcup \{A_n | n \in \mathbb{N}\} \in \mathcal{F};\end{aligned}$$

である。ここに $\mathbb{N}$ は 1 から始まる自然数系列。

P が、

$$\begin{aligned}P: \mathcal{F} &\rightarrow \mathbb{R}; \\ P(A) &\geq 0 \forall A \in \mathcal{F} \text{ 非負性}; \\ A_n \in \mathcal{F} \forall n \in \mathbb{N}, A_i \cap A_j &= \emptyset (i \neq j) \Rightarrow \\ P(\bigcup \{A_n | n \in \mathbb{N}\}) &= \sum_{n=1}^{\infty} P(A_n) \text{ 完全加法性}; \\ P(\Omega) &= 1, [1];\end{aligned}$$

をみたす場合、 P は可測空間 $(\Omega, \mathcal{F})$ 上の確率とよばれる。ここに $\mathbb{R}$ は実数体だが、実数連続体（実数直線）と同一視する。従って点集合。数学的には、 $(\Omega, \mathcal{F})$ 上の完全加法的測度で全空間の測度が 1 であるものを確率測度とよぶ。 $(\Omega, \mathcal{F}, P)$ は確率空間とよばれる。

P は実数値をとるので完全加法性より $P(\emptyset) = 0$ となる。再び完全加法性より、 $\mathcal{F}$ に属する A, B が $A \cap B = \emptyset$ ならば、 $P(A \cup B) = P(A) + P(B)$ が従う。この有限加法性は確率算の加法法則である。しかし、加法法則が合理的説得力を持っていることが示されたわけではない。完全加法性という強力な前提から有限加法性を出しただけである。なお、 $0 \neq 1$ より $\Omega \neq \emptyset$ となる。これは前提ではなく帰結である。空集合の記号 $\emptyset$ はウーと読むことにしておく。

1.2 乗法法則の迂回

A, B を $\mathcal{F}$ の要素とし、 $P(A) \neq 0$ とし、 $P(B|A) = P(A \cap B) / P(A)$ とおく。 $P(B|A)$ は、 A が与えられている場合の、 B の条件つき確率とよばれる。これより、

$$P(A \cap B) = P(A) \times P(B|A)$$

となる。 $P(A) = 0$ の場合 $P(B|A)$ を任意の実数値 x_A に固定しておけば、 $0 = 0 \times x_A$ よりこの等式は常に成立する。これは乗法法則だが、条件つき確率の定義の合理的根拠が示されていないので、合理的説得力を持つ導出とは言い難い。

1.3 本来の事象

Ω の各要素 ω は根元事象とよばれる。根元事象の集まりを事象とよぶ。 $\mathcal{F}$ に属する事象は可測とよばれる。 $\emptyset, \Omega$ は可測な事象であり、各空事象、全事象とよばれる。 ω の単元集合 (singleton) $\{\omega\}$ は事象だが、可測とは限らない。

本来の事象とは「できごと」だが、それは一回性のものである。「雨がふる」という現象はくりかえすが、「その雨」は一回限りでありくりかえさない。くりかえすのは現象である。Bruno de Finetti は 1920 年代に、事象の一回性について鋭敏な分析を行っている。Kolmogorov は集合概念を援用して、「本来の事象」を巧みに迂回している。

A.N.Kolmogorov は、確率の解析学を展開するために彼の公理系を提示し、確率に関する難題を迂回しているのである。その際、集合概念が基本的な役割を演じている。確率は終わっていない。(なお、事象の一回性という場合の「一」だが、一期一会という場合の「一」であり、対応する英語は unique とか uniqueness である。)

1.4 G. Pólya (ポーヤ) の壺

a, b, c を非負整数とし、 $a+b>0$ とする。壺の中に同形同質の、 a 個の黒玉と b 個の白玉が入っており、この $a+b$ 個の玉から無差別に 1 個を取り出し色を観察し、それを壺に戻すとする。但し、その際同色の玉 c 個を追加する。この試行をくりかえす。はじめが黒ならば、次の抽出は、 $a+c$ 個の黒と b 個の白との内訳で行われる。 n 回目の抽出で黒となる確率は $a/(a+b)$ となる。

確率を直観的に扱って Prob とすれば、加法及び乗法法則と数学的帰納法とにより、 $a/(a+b)$ が容易にでる。しかし、Kolmogorov の P をつかって出そうとすると、うまくは行かない。例えば、暗黙のうちに乗法法則を使うようなことになりかねない。ぜひとも試みてもらいたい。

1.5 包除原理

$\mathcal{A}$ を Ω 上の集合体とする。つまり、

$$\begin{aligned}\mathcal{A} &\subset \mathcal{P}(\Omega); \\ \Omega &\in \mathcal{A}; \\ A \in \mathcal{A} &\Rightarrow A^c = \Omega - A \in \mathcal{A}; \\ A, B \in \mathcal{A} &\Rightarrow A \cup B \in \mathcal{A}.\end{aligned}$$

また P を、

$$\begin{aligned}P: \mathcal{A} &\rightarrow \mathbb{R}; \\ P(A) &\geq 0 \forall A \in \mathcal{A}; \\ A, B \in \mathcal{A}, A \cap B = \emptyset &\Rightarrow P(A \cup B) = P(A) + P(B); \\ P(\Omega) &= 1;\end{aligned}$$

とする。 $\mathcal{A}$ の要素を事象とよぶ。

包除原理の最も基本的な場合は、事象 A, B に対して、

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

である。一般に、事象列 $A_i, i=1, \dots, n, n>1$, に対して、

$$\begin{aligned}P(A_1 \cup \dots \cup A_n) &= \sum_{i=1}^n P(A_i) - \sum_{i(1)<i(2)} P(A_{i(1)} A_{i(2)}) \\ &\quad + \sum_{i(1)<i(2)<i(3)} P(A_{i(1)} A_{i(2)} A_{i(3)}) - \dots \\ &\quad + (-1)^{r+1} \sum_{i(1)<\dots<i(r)} P(A_{i(1)} \dots A_{i(r)}) + \dots \\ &\quad + (-1)^{n+1} P(A_1 \dots A_n)\end{aligned}$$

である。

直接には数学的帰納法だが、 P が確率ならば期待値作用素 E が使えて、 $P(A) = E(I_A)$ に注意して、 $\prod_{i=1}^n (1 - x_i)$ の展開式が既知 (証明は数学的帰納法) ならば、 $I_{A^c} = 1 - I_A$, $I_{AB} = I_A I_B$ を使えばほとんど自明。ここに、 AB は $A \cap B$ の略記で、 I_A は A の指示関数、つまり、 ω が A に属すれば 1、属しなければ 0、となる関数。いわゆる、 A の定義関数である。(なお、包除原理に対応する英語は、principle of inclusion and exclusion である。)

2. 順序統計量

$$J_j = \sum_{i=1}^n I\langle X_i < X_j \rangle + \sum_{i < j} I\langle X_i = X_j \rangle + 1$$

普通は略式に扱われていることを正式に規定しておく。

とおいて,

$$X_{(i)} = \sum_{j=1}^n I\langle X_j = i \rangle X_j$$

2.1 標本

確率空間 $(\Omega, \mathcal{F}, P)$ が与えられているとする。 n を非負の整数とし、実数値をとる確率変数 (random variable) からなる列 $(X_i; i=1, \dots, n)$ を考える。この列が独立同一分布 (i.i.d., independent and identically distributed) であれば、大きさ n の標本とよばれる。 n が 0 ならば空列 $\emptyset$ となり、考察者の資質と経験とがためされる。以下では n を正としておく。標本概念は可測空間を使えば一般化できるが、ここではこれで充分である。

とおけばよい。但し、 Ω 上の命題 Q に対して $I\langle Q \rangle = I\langle \{\omega \in \Omega | Q(\omega)\} \rangle$ である。これによって順序統計量 $(X_{(i)}; i=1, \dots, n)$ が正式に導入される。なお可測性は、任意の $i, j \in \{1, \dots, n\}$ に対して,

$$\Delta(i, j) = \{x \in R^n | x_i < x_j\}$$

とおくと、 $\Delta(i, j)$ (特に $\Delta(i, i) = \emptyset$) が R^n の Borel 集合となることなどより従う。

2.2 統計量

標本の関数は統計量 (statistic) とよばれる。例えば、正の整数 m に対して R^m を考え、 R^n から R^m への Borel 可測な関数 t をとり、合成関数 $t(X_1, \dots, X_n)$ を考えれば、これが一つの統計量となる。ここでは、 R^m を m 次元 Euclid 空間と同一視しておく。つまり点集合となる。

2.3 順序を考える

任意の $\omega \in \Omega$ に対して、実数列 $(X_i(\omega); i=1, \dots, n)$ を小から大へと順に並べて,

$$X_{(1)}(\omega) \leq \dots \leq X_{(i)}(\omega) \leq \dots \leq X_{(n)}(\omega)$$

とする。ただし値が等しい場合 (いわゆるタイ) には、添数が若いほうを先に並べる。ここで、各 ω に $X_{(i)}(\omega)$ を対応させる関数 $X_{(i)}$ を、第 i 番目の順序統計量とよび、また列 $(X_{(i)}; i=1, \dots, n)$ をも順序統計量 (order statistic) とよぶ。

しかし、 $X_{(i)} = t_{(i)}(X_1, \dots, X_n)$ となる可測な $t_{(i)}$ を指定しなければ正式の統計量とはならない。この $t_{(i)}$ を構成する。これは、

3. データによる変数の置き換え

n 次元 Euclid 空間上の Lebesgue (ルベーグ) 測度を λ^n とする。 $(X_1, \dots, X_n)$ を大きさ n の標本とし、 $t(X_1, \dots, X_n)$ を実数値をとる統計量とする。母数 $\theta \in \Theta \neq \emptyset$ に対する推定を問題とする。簡単のため Θ は R の部分集合としておく。 θ に対応する確率を $P(\cdot; \theta)$ 、この確率に基づく期待値作用素を $E(\cdot; \theta)$ とする。 X_1 の分布は λ^1 に関して絶対連続とする。次の $(*)$ を仮定する。

$$(*) \quad E(t(X_1, \dots, X_n); \theta) = \theta, \quad \forall \theta \in \Theta.$$

一方、計算可能な実数の全体を、Alan Mathison Turing の頭文字をとって AMT とし、任意に固定された実数を α とする。 t を変形する。 $x \in \text{AMT}^n$ ならば α , $x \in R^n - \text{AMT}^n$ ならば $t(x)$ 、によって定義される関数を $t(\cdot; \alpha)$ とする。よく知られているように AMT は可算無限集合であるので、 $\lambda^n(\text{AMT}^n) = 0$ 。 $t = t(\cdot; \alpha)$ a.e. (λ^n) 。従って、 $t(\cdot; \alpha)$ も $(*)$ をみたす。しかし、データ $(x(1), \dots, x(n)) (\in \text{AMT}^n)$ に対

しては, $t(x(1), \dots, x(n); \alpha) = \alpha_0$.

つまり (*) だけに基づいて, データがもたらす値を「真の母数」に結びつけるのには無理がある。

実際には common sense によって t を「えらぶ」ので, 破局は避けられる。しかし, この common sense の一角を鋭敏に定式化すれば立派な業績である。AIC を視よ。なお Jerzy Neyman は, Lebesgue 積分に通じていたので, データによる置き換えがもたらす一個の区間と, 確率値がもたらす信頼係数とを結びつけることを峻拒している。もっとも若いころの話だが。

4. 発見的接近の利点と限界と

発見的接近は特に自然科学者には歓迎されることであろう。ここでは, Γ 及び Poisson (ポアソン) 分布との関連で, その利点と限界とを確認する。

4.1 Γ 分布

この分布の実数直線上の密度関数は, $(-\infty, 0]$ 上で 0, $(0, \infty)$ に属する各 x に対して

$$\beta^{-\alpha} \Gamma^{-1}(\alpha) x^{\alpha-1} \exp(-x/\beta)$$

をとる。ただし α 及び β は正の実数であり, 各形状母数 (shape parameter) 及び尺度母数 (scale parameter) とよばれる。

4.2 ある過程と Prob と

$[0, \infty)$ 上で生起する事象たちを考える。一つ事象が起こればこの半直線上に一つ印をつける。またこの半直線を時間軸ともみなす。事象たちは独立に起こり, 有界区間 ($\emptyset$ を含む) I で起こる事象の総数を $N(I)$ とする。ここで確率 Prob を直観的に導入し, 以下を仮定する。

$$\text{Prob}(N(I)=0)=1-\lambda\Delta+\varepsilon_0(\Delta),$$

$$\text{Prob}(N(I)=1)=\lambda\Delta+\varepsilon_1(\Delta),$$

$$\text{Prob}(N(I)>1)=\varepsilon_2(\Delta).$$

ここに λ は正の実数で定数。 Δ は I の長さ。 $\varepsilon_i(\Delta)/\Delta \rightarrow 0, \Delta \downarrow 0, \varepsilon_i(0)=0, i=0,1,2$. さらに, n を任意の正の整数として, 区間列 $I_i, i=1, \dots, n$, が disjoint (交わりを持たない) ならば, 変数列 $N(I_i), i=1, \dots, n$, は Prob に関して独立とする。

すると, t を正の実数とすると, 非負の整数 m に対して,

$$\text{Prob}(N([0,t])=m)=e^{-\lambda t}(\lambda t)^m/m!$$

となる。実際, 任意の正の整数 n に対して, 区間 $[0,t]$ を n 等分して, 但し独立性を使って, $n \uparrow \infty$ とすると,

$$N([0,t]) \sim \text{Poisson}(\lambda t)$$

である。「剰余」は有限的な劣加法性でみつもればよい。

さらに正の整数 n に対して, n 番目の事象が起こるまでの時間を T_n とすれば, t を正の実数として,

$$\begin{aligned} (d/dt)\text{Prob}(T_n \leq t) &= -(d/dt)\text{Prob}(T_n > t) \\ &= (\lambda^n t^{n-1} / (n-1)!) e^{-\lambda t} \end{aligned}$$

となり,

$$T_n \sim \text{Gamma}(n, \lambda^{-1})$$

が従う。これは形状及び尺度母数が各 n 及び λ^{-1} の Γ 分布だが, 形状母数が正の整数である場合, Erlang (アーラン) 分布とよばれることがある。なお, (α, β) が $(1, \beta)$ ならば指数分布, $(v/2, 2)$ ならば $\chi^2(v)$, v は正の実数で自由度, となる。

4.3 利点と限界と

自然科学者ならば発見的接近 (heuristic approach) を好むであろう。上のように先に分布が、つまりは答えが、分かってしまうのである。しかし大きな問題がある。それは果たして Prob や $N(\cdot)$ が存在するのか、また存在するとしていかなる意味で一意か、ということである。最初にこれらの存在を基礎づけたのは、おそらく伊藤清である。

5. 正規分布への短い注意

非正規性は偏在している。測定誤差ですら非正規であることが、古くから知られている。しかし正規分布は、中心極限定理 (CLT, Central Limit Theorem) との連関でしばしば言及される。古典的な CLT は Lindeberg (リンドバーク)-Feller の CLT に帰着する。但し、ここで古典的とは、無限次元 Euclid 空間 $\mathbb{R}^\infty$ 上の確率測度を持ち出さなくとも、古典的な解析学の範囲内で、但し特性関数 (結局 Fourier 変換) を巧妙に用いて、導出できるということである。

分布関数 (右連続としておく) の列が分布関数へと各点収束する場合、収束先の分布関数が実数軸上で連続ならば、その収束は実数軸上で一様であるという、G.Pólya による指摘はよく知られている。つまり、CLT での収束は一様である。関数列の収束の場合、各点収束のみでなく、どの範囲で一様なのかも突き止めておく必要がある。正の実数 v を自由度とする t 分布の密度が、 $v \uparrow \infty$ とすると、標準正規分布の密度へと各点収束することはよく知られているが、また Stirling の公式を使えば容易に示せるが、実軸上のどの範囲で一様となるのかも、やはり明記すべきである。

6. 未知固定確率

しばしばコインの投げ上げとの連関で、未知固定確率 p が持ち出される。確率を定義して

いないのに未知固定 (unknown but fixed) と断じるのは異様である。さらには、 p に関する二項分布も導入されたりする。ここでは裏及び表に対して各 0 及び 1 を対応させて数値化しておく。

独立試行という言葉が持ち出されて、 $\{0,1\}^N$ 上に確率測度が導入されるのが普通だが、この測度を $\mu(p)$ としておく。すると、 n 回までの表の相対度数 $\bar{X}$ が、 $n \rightarrow \infty$ とすると、 $\mu(p)$ に関して確率 1 で、 p へと収束することが示せる。 p は無定義なのだから、この命題が何を意味するのかは不明である。しかし $\mu(p)$ は、 p を使っているのだから、「 $\bar{X}$ の p への非収束」に対するマチゴマであろう。この $\mu(p)$ の存在を導くのに Kolmogorov の拡張定理を持ち出すのは大げさである。無限直積なので、角谷静夫 (かくたに・しずお) の (あるいは John von Neumann の) 拡張定理で充分である。

未知固定確率については、Bruno de Finetti が 1920 年代に、今日の交換可能性を用いて鋭敏な分析をしており、de Finetti (デ・フィネッティ) の表現定理を導出している。

7. 文献のことなど

筆者は既に次の冊子四冊を出している。

- [1] 『サヴェジ基礎論覚書』, 岩波出版サービスセンター, 東京, 2001 年 12 月 20 日。
- [2] 『サヴェジ氏の思索』, 岩波出版サービスセンター, 東京, 2007 年 8 月 31 日。
- [3] 『確率概念の近傍—ベイズ統計学の基礎をなす確率概念—』, 内田老鶴圃, 東京, 2014 年 5 月 15 日。
- [4] 『サヴェジ・システム試論—統計的決定理論の公理化と期待効用の最大化—』, 内田老鶴圃, 東京, 2017 年 5 月 10 日。

諸文献については [3] の付録 C を、未知固定確率については [1] の第 4 章を一瞥して頂

ければ幸いである。[3] [4] を視れば了解できるように、個人的価値判断の合理性が確率に先行しているとすれば、加法及び乗法法則へのかなり説得力のある証明が得られる。しかし、個人的損得勘定を確率に先行させ、しかもそれを損得勘定の統治下に置くことは、生粋の自然科学者には我慢ならない遣り口であろう。Fisher, Ronald Aylmer, Sir, でなくとも不快に思うに違いない。Fisher は「基準確率, fiducial probability」なる言葉を持ち出しているが、実に苦しい。確率の基礎はまだまだ甘くない。

鈴木雪夫教授（故人、理学博士）は、漸近論的接近は確率論ではあっても、本来の統計学ではないという、生涯変わらぬ信念をお持ちであった。論点は二つある。第一は、確率の定義がないこと。第二は、データが与えられれば、その標本の大きさ N は、そのデータに固有の定数となること。例えば、 N が 5279 と定まれば、これを勝手に 5281 に変更することは許されない。いかなる有限の限界をも越えて増大する変量 N の出る幕はない。思うに、漸近論的接近は統計学固有 (statistics proper) ではない。ただし、この統計学固有とは、岡部純一博士（故人、経済統計学会）が行っている作業を念頭に置く。

筆者が数学科の学生であった頃、実に幸運にも、伊藤清三先生の授業を受けることとなった。伊藤先生は、ルベーク積分入門と確率論の基礎概念とを、懇切丁寧に講義して下さった。繰り返すが実に幸運であった。ゼミでは、岡部靖憲先生に、確率過程の基礎を、じっくりと骨の髄に徹するまで、御指導頂いた。また中学生の頃、担任の浅田日出夫先生（藤沢市立明治中学校教諭、当時）に、的確な集合論の入門書を御教示頂いた。これにより、公理的集合論の存在を知った。

なお数理統計学への入門書として、簡潔かつ卓越した、

[5] 鍋谷清治（なべや・せいじ）、『数理統計学』、共立出版、東京、1978 年 2 月 5 日、

がある。この第 4.4 節 C(153 頁) で、可測な選択との関係で（集合論における）P.S. Novikov の定理（1931 年）が引用されているが、入門書ではあっても初等的ではない証である。

伊藤清の確率論（新版）だが、旧版も同様だが、条件つき期待値の導入の仕方が良くない。K.Itô-P.Lévy 式の確率過程論への優れた入門書なので残念である。ここは定評ある、

[6] Breiman, Leo, *Probability*, Addison-Wesley, Reading, Massachusetts, 1968, 但し、SIAM より paper 版が出ている、

の第 4 章を学ぶのが良い。その際第 4.3 節 80 頁の中央の段落の注意と、81 頁の NOTES とは、読み落とさないで欲しい。なお確率論（旧版）の付録での測度の拡張定理の取り扱いも良くない。

Exchangeability（交換可能性）と Lindeberg-Feller の CLT との簡潔な説明が次の [7] にある。

[7] Loève, Michel, *Probability Theory, Third Edition*, Dover, New York, 2017. もとは 1963 年に Van Nostrand から出ている。第一版は同社より 1955 年に出ている。第四版は GTM の二巻本である。

なお A.N.Kolmogorov の「基礎概念」については、次の労訳がある。

[8] 坂本實（さかもと・みのる）訳、A.N. コルモゴロフ著、『確率論の基礎概念』、ちくま学芸文庫、筑摩書房、東京、2010 年 7 月 10 日。

この [8] には、A.N.Shiryaev の解説がつい

ている。但し、そこでの Leonard Jimmie Savage
の業績への言及は不十分。

2024 年 4 月 22 日（月） 札幌の寓居にて

註. 北大経済学部同窓会の塚田久美子氏より、編集上
の助力を得た。ここに謝意を記す。

末尾の一言。学徒の本務とは真理の探究であ
る。有用性は言わば副産物だ。