現実の現象を描く確率・統計－Pattern Theory

田中紀子　松原望

D.Mumford は実世界（real world）の信号（signal）に見る偶然 stochastic 現象を分析する広汎かつ高度な数学理論として、確率論・統計学の今後の重要性と役割発展を強調し Pattern Theory を提唱した。世界的に高名な（フィールズ賞受賞）著者の一見分野外の成果だけにその衝撃は大きい。今回の発表はその最初の一部として P.Lévy の無限分解可能性（ID）の深遠な結果に触れ、それを理解するための基本確率論の要諦および厳密数学として K.Itô ,H.McKean の定評ある成書の関連部分を紹介する。

1．確率論･統計学の第 3 千年紀*新時代の夜明け

D.Mumford 'Pattern Theory'の見通しはこうである。 2000 年以上もの間、西洋での考え方はアリストテレスの論理に基づいてきた。すべての精密な理論、科学的モデル、また思考過程のモデルそのものすら原則として論理の縛りに従ってきた。しかしながら、ギャンブルの戦略や死体の数え方を考案した中世ロンドンの暗い時代から、確率論や統計的推論が現在、科学的モデル、特に思考過程のモデルのよりよい基盤や、理論数学、さらには数学の基礎そのものの本質要素となってきている。我々の総体的な物事の見方に関するこの大いなる変革が次の千年紀の数学すべてに実質的な影響を与えるだろう。

Mumford によれば:

―私の全体的結論は、確率的方法は第 3-千年紀の最初には、純粋数学及び応用数学を変えると信じているということである。確率と統計学は科学的モデリングだけでなく数学においても使われるべき自然の道具として見られるようになるであろう。知的世界は全体に、論理学を美しい<u>優雅な理想像</u>としてみなすが、統計学を私たちが推論し考える<u>標準的方法</u>とみなすようになるであろうー

＊（以下引用者注）日本語には *century*(=100 年)に「世紀」の訳語があるが、*millennium*(=1000 年)に対する適訳はないので、「千年紀（期）」とした。現在は第 3 千年紀（紀元 2001～3000 年）内にある。

2．D.Mumford 'Pattern Theory'の基礎と意義

現実の現象の世界は変動し、一回限りのものである。変動性を扱うのが確率論およびそれを測る統計学である。この考え方にたって、現象からの信号（自然科学の信号）を分析するアプローチが D.Mumford の Pattern Theory である。

とりわけ今回紹介するのはその重要な一例で、どんな現象もこれより小さい同様の独立な確率的変動の和からできている、いいかれば、どんな現象もどこまでも無限に分解しうる（無限分解可能 infinitely divisible)という基本的性質であり、この概念を見出したのはフランスの P.Lévy である。これは一例に過ぎない。Mumford は確率イコール測度（確率

測度）という今日の優勢な立場に疑問を呈し、いわば「P.Lévy の自然の本流へ還ろう」とP.Lévy を賞揚する。

しかしながら、日本の他の数学分野にくらべ受容が遅れている確率論でも、原典読み（本論文第 9 章）からも Lévy が当初忘却されていたわけではなく、むしろ伊藤清にあっては Lévy こそ指導原理であったことは明白である（たとえば Brown 運動)。ただ、数学領域での確率論の位置づけをかんがえると、数学的理論整備にも意を用いざるを得ななかった事情から、以後さまざまないわば「伊藤解釈」が生まれたことは否定できない。今回も昨年度に引き続き Lévy や Mumford の視点から自然現象を考察するが、それ自体の意義のほかに、日本の確率論史が辿った道すがらを顧みることができ、であればこそ、これからの新しい千年が「確率論と統計的推論の時代」と（分野以外の著名数学者から）力づけられるのも、ありがたく面はゆい感を禁じ得ない。

3.「実世界」の Pattern Theory による定式化(総論)

Mumford 視野はこれに限られずさらに広い。彼によれば、この論文は非常に本質的な点を議論する論争となるだろう。すなわち、確率的モデルと統計的推論は、

i)　現実社会へ

ii）科学や数学の多くの部分へ

iii)特に我々の心の中での計算を理解することへ

強く関連している。

確率における研究の基本的対象は偶然変数*(Stochastic variable)であり、それは空間、群、関数といった基本的構成物として[並んで：引用者]扱われるべきもので、それを測度論を用いて定義するのは人為的で不自然である。確率や偶然変数も数学の基礎になり得、それらはより本質的で強力な体系をなすことを提案している。

このため、確率の基本理論を展開するには二つのアプローチがある。一つは、可能な限り、確率の[本来的]言葉を削り<u>測度理論へと単純化する</u>ことである。そこで確率空間Ωはなくなり[抽象的に定義されれば済み具体性は問われない]、集合 X には、Ω上の確率測度の写像 x の下での像で得られる p(x)あるいは p(x)dx が与えられる[写像Ω－>X によって実体空間での操作が可能になる]。もう一つは<u>「偶然変数」の概念を中心に持ってきて、可能な限り偶然変数を操作するものである。</u>単純家の方法では、偶然変数は測度で定義される。その測度は、実数の理論で定義され、その実数の理論は集合論で定義され、その集合論は述語計算のうえで定義される。私(Munford)は代わりに、偶然変数を論理学と数学の両方の基礎に入れ、確率的視点においてより完璧でより透明度の高い公式に辿り着くはずだといいたいのである。

＊通常は「確率変数」random variable であるが、stochastic は「確率的」というより「偶然的」という実体的な語感がある。たしかに、「確率」を定義することが主目的ならそのための偶然量は、日本でそうであるように、常にさいころなどに決まっている。これは非現実的である。われわれには偶然量自体がまず関心対象であり、それがいくらの確率かは二義的であることは多い。「確率論」とはある偶然量の確率

の計算論である。

これら二つのスタイルを対比させる例がある。実数値偶然変数 x の'infinite divisibility'(ID:無限分解可能)のコンセプトを取り上げよう。現象的には二次元、三次元であるが、ほとんどすべての場合のように、数学的には一次元で定式化される。

一次元 X の確率密度関数を p(x), x ϵ R と表し、任意の n について

$$p=q_n * \cdot\cdot\cdot * q_n \quad (n \text{ 個の因数 } q_n)$$

と分解される確率密度関数 $q_n(x)$ があれば、X は「無限分解可能」(Infinitely divisible, ID)と言われる。＊は「たたみこみ」(convolution)あるいは「合成積」で、ここでは密度関数 f(x),g(x), x ϵ R に対

$$(f*g)(x) = \int_{-\infty}^{\infty} f(x-u)g(u)\,du$$

で定義される。あきらかに＊は可換、また結合律を満たす。

いま、X が f(x)に、Y が g (x)に従う（確率分布の記号〜を用いるなら、X〜f,Y〜 g ）とき、X、Y が独立なら

$$X+Y \sim f*g$$

である。なお、〜の記号は同一分布に従うとき、X〜Y のように転用される。

したがって、むしろつぎのようにいいかえられる。任意の n に対し、独立で同分布をもつ Y i(i=1,2, ・・・,n)が存在して

$$X \sim Y_1 + Y_2 + \cdot\cdot\cdot Y_n$$

のように、分布において分解される。

これは、表記の単純な変化に過ぎないが、この二通りの方法で Lévy-Khintchine 理論について述べるときに何が生じるか考えてほしい。最初の方法では、この理論は p(x)のフーリエ変換(後述する特性関数)が次のように表される場合のみ、X はＩＤ（無限分解可能)であることを述べている。

$$\varphi(t) = \exp \left\{ i\mu t - \frac{\sigma^2}{2} t^2 + \int_{-\infty}^{\infty} \left(e^{itu} - 1 - \frac{itu}{1+u^2} \right) d\lambda(u) \right\}$$

(Lévy - Khintchine)

二番目の方法では、同じ条件が偶然変数 x を使って直接以下のように表される。

$$X \sim a + bX_{normal} + c\Sigma(X_i - \text{convergence factor } c_i)$$

ここで X_{normal} は標準正規変数、Xi はジャンプ u（－もあり）の強度 intensity λ のポワソン過程である。

さて、私（Mumford）にはこれらは実に違うものに見える！二番目の方法で Lévy-Khintchine 理論を説明するほうがはるかに明確である。その偶然変数を明確にすることに結果の真の確率的意味が明らかになる。

＊ここで「Lévy-Khintchine 理論」とは上述$\varphi(t)$表現が無限分解可能の十分条件とうことである。$\varphi(t)$はいささか奇妙な形をしているが、よくみれば（後述５.2）正規分布、ポアソン分布を含んでいることが一目瞭然で、それを確率変数で表現した式 X〜・・・はたしかに理解しやすさで優れている。ただし、以上の議論も５.2 の特性関数の理解がないと真意は徹底しないであろう。

4. Pattern Theory―実世界の信号の確率的分析―

David Mumford と Angnes Desolneux の「Pattern Theory」を紹介し、確率・統計的視点で現実世界をみることの有用性について述べたい **Pattern Theory とは**偶然的な（stochastic）すべての形の実世界の信号（all forms of *real* world signals)を分析することで、背後の隠れた確率構造をより直接的に把握するモデルである。

本来「 Pattern Theory 」はスウェーデンの数学者・統計学者 Ulf Grenander（1923－2016）により始められたに対するアプローチである。その中心にはまず広い範囲の確率論モデルがあり、本書はその統計的サンプル*になっていて、実信号、そのパターンおよびその変動の様子（real signals, patterns　and their variability)を感覚することができる。

＊確率から生じた一つの代表的実現例をいう。さいころを１０回投げれば 6^{10} 通りの結果が確率 $1/6^{10}$ で生じるが、もちろん実際にはそのただ１通りが代表的データとして観察される。

また、ベイズ統計学による推論にこれらのモデルを当てはめれば、さらに新しい信号の分析も可能になる。ここでは、まず数理的ツール、次にモデルそのもの、さらに６段階の信号の複雑さに応じた分類(six representative classes of signals of increasing complexity)に対し統計量を分析する計算アルゴリズム(computational　algorithm for applying statistics)を扱う。６段階は以下の通りである。

1. 英語テクストとマルコフ連鎖
2. 音楽と区分的ガウス性
3. 文字認識とシンタクス的分類
4. 画像構造のセグメント化とギッブスモデル
5. 顔と弾力性鋳型
6. 自然科学とマルチスケール型分析

このうち、６．の２実例を紹介しよう。

＜例：自然界における高い尖度＞　「尖度」(*kurtosis*, coefficient of excess、記号 κ ）とは分布の中心付近の尖り具合の指標を指し,確率分布の４次のモーメント（後述 4.2）から定義される。高い κ は中心付近が tall and slim であることを示す。(Cramér, p.184)（ただし、最近のテキストでは経験上いわゆる分布の裾の長さ、重さを指す指標として説明される）。 Mumford は、自然界のある写真をピクセル化（格子で区切り小細分化）して離散データ$I(i,j)$にした結果、x 方向の段差のデータ

$$I(i,j)-I(i-1,j)$$

から κ ＝10～20 と正規分布の $\kappa=3$ に比べて極端に高い尖度を得た。これは自然界イメージ（画像）は事物の明確な「輪郭」によって区画化されていることを示す。

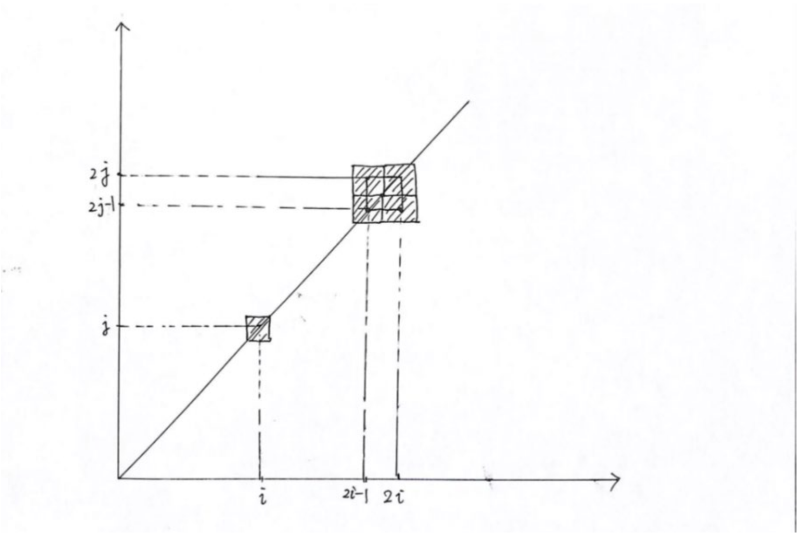
＜例：イメージ・ピラミッドと尺度不変＞　同じ設定で、もとのイメージを $2^n\times 2^n$ のサイズに分割し、「イメージ・ピラミッド」として、まず

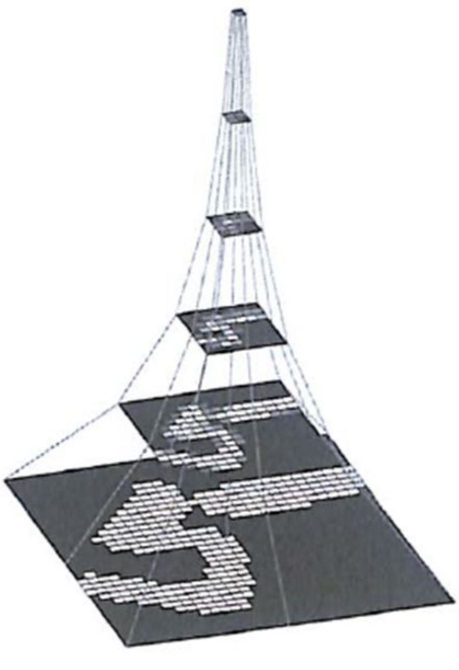
$$I^{(1)}(i,j)=\frac{1}{4}\ \{I(2i-1,\ 2j-1)+I(2i,\ 2j-1)+I(2i-1,\ 2j)+I(2i,2j)\}$$

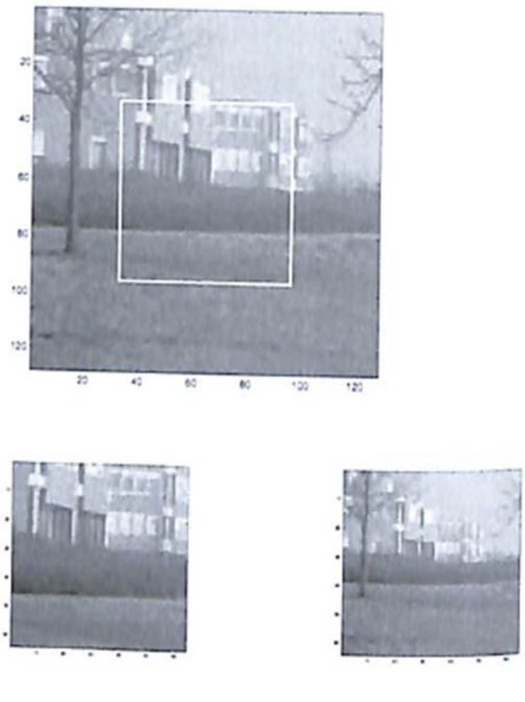
のように縮小し、以後 k（≤n）に対し''blown down''イメージ（サイズ$2^{n-k}\times 2^{n-k}$）

$$I^{(k)}(i,j) = \frac{1}{4^k} \sum_{r=0}^{r=2^n-1} \sum_{s=0}^{s=2^n-1} I(2^k i - r,\ 2^k j - s)$$

を作成する。‘’blown down’’とは風で飛んでいく、という意味である（図参照）。このように、視覚イメージはスケール（尺度）を変えてもそのものの原イメージの表現であることは変らない。これに対して、聴覚データは時間の尺度を変えれば音の周波数が変り音自体が別物に変容する。







＜例：無限分解可能分布＞　すでに述べたが、理論的表現は次章以降とする。

５. 偶然自然現象の確率･統計論的描写の基礎（各論 I）

５.1 確率の測度アプローチ対偶然変数アプローチ

Kolmogorov 流の測度論的確率の定義は、まず確率を定義するべき集合の族（完全加法族、σ-代数、ボレル集合族 etc.)を

(i)　$\Omega \in \mathcal{F}$　(7.1.1)

(i)　$A \in \mathcal{F}$　なら　$A^c \in \mathcal{F}$　(7.1.2)

(ii)　$A_1, A_2, \cdots \in F$ なら $\bigcup_{i=1}^{\infty} A_i \in F$　(7.1.3)

のように定めて置き、しかるのちに $\mathcal{F}$ 上に確率測度を公理として

(I)　$A \in \mathcal{F}$ に対し $P(A) \geqq 0$　(7.2.1)

(II) $P(\Omega) = 1$　(7.2.2)

(III) $A_1, A_2, \cdots \in \mathcal{F}$ で $A_i \cap A_j = \phi \ (i \neq j)$ ならば

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i) \qquad \text{（完全加法性）}$$

を満たすように定義するものであるが、確率論の数学的厳密化ではきわめて大きな役割を果たす一方（ことに日本に於いて）、確率の内在的な意味内容を把握しているようには考えられないとする消極的評価が一定程度根強くある。

＊Mumford の評価もこれに属する。評者としてはおおむね同調しつつも、提唱者 Kolmogorov もこれが確率であると云い切ったわけではなく、（確率は数学的には）このようにも表現できる、と述べたにとどまると理解したい。戦後まもなくこの公理的確率が日本に知られたとき、数学界からは驚きと戸惑いを以って迎えられたという（東京帝国大学理学部数学科卒業当時の松下嘉米男（後に統計数理研究所第一研究部長）の述懐。）

５.2 数理的準備：モーメント母関数と特性関数

５.2.1 確率分布のモーメント母関数　mgf（moment generating function）

＜定義＞　実数値関数 f(x), $x \in R$ は

$$f(x) \geqq 0, \qquad \int_{-\infty}^{\infty} f(x)dx = 1$$

のとき、（ルベーグ測度 d x に関する）確率密度関数と言われる。

注）　f (x)は連続と限らないがルベーグ可測でなければならない。

＜定義:モーメント＞　確率分布 f(x)に対し

$$E(X^k) = \int_{-\infty}^{\infty} x^k f(x)dx$$

を X あるいは f(x)の k 次のモーメントという。

よく知られているのは、一次のモーメント

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

で「期待値」Expectation ,Expected value と呼ばれる。

一連の次数のモーメントはスカラーであるが、分布の重要なパターン特性を通して確率

象の実体を表わす。4 次までは名称がある。

i)　E(X)：期待値、分布の大略の位置、X の大略を表す

ii)　E(X^2) :V(X)＝E(X－E(X))^2=E(X^2) − (E(X))^{2}で分散(ばらつき）を表わす

大方の不確実現象もおおむねこれら 2 次のモーメント（期待値と分散）で片がつくが、さらに関心がある場合、あるいは詳細を必要とする場合は以下となる。

iii)　E(X^3)：E(X)＝０, V(X)=1 とする標準化後、「歪度」(非対称性）を表わす

iv)　$E(X^4)$ ：E(X)＝０, V(X)=1 とする標準化後、「尖度」(中心付近の尖り方）を表わす。

*歪度は skewness，尖度は kurtosis（後述）

i)〜iv)で f (x)の概形がほぼ定まるが、特に命名のないさらに高次のすべての次数のモーメントにより f(x)は一意に決定されると考えられる。全ての次数のモーメントを知るのは次の定義、定理による：

＜定義：モーメント母関数＞　X の確率分布を$f(x)$とするとき、t を助変数として

$$M(t)=\int_{-\infty}^{\infty} e^{tx} f(x)dx$$

を X の（あるいはfの）モーメント母関数（moment generating function）という。

＜定理＞　正則条件のもとでマクローリン展開から、第 k 階導関数

$$M^{(k)}(0)=E(X^k),$$

M_X (t)は積分計算を迂回して微分演算により全ての次数のモーメントを生成する。この利便性は大きい。モーメントに限らず、「母関数」generating　function の方法は歴史的には S.Laplace に遡ると言われるが、ただ利便性だけではなく、以下が成立する。

<定理：対応>　　　　M:　f(x)　－＞　M_X(t) は 1 対 1

確率分布(x)は現象実体に近いので扱いは数学的には容易でないが、その形状パラメータであるモーメントを通して、モーメント母関数にすべての情報が集約される。また次のように、高い演算性を持つ：

<定理：和の mgf>　X,Y は独立で X の mgf を$M_X(t)$，Y の mgf を$M_Y(t)$とするとき、X+Y の mgf は$M_X(t)\ M_Y(t)$で与えられる。

確率変数の和 X+Y のたたみこみ演算＊は、モーメント母関数に移せば単なる積になる。これもモーメント母関数の優れた利便性である。結局は、確率分布の便利な「身代わり」といって差し支えない。

＜例＞　いくつかの重要例は今後の理論上も頻用される。とりわけ正規とポアソンについては要注意であり、本論文でも無限分解可能分布の例になっている。

正規分布 N(μ，σ^2)　　$M(t) = e^{\mu t+\sigma^2\frac{t^2}{2}}$

ポアソン分布 Po(λ)　　M(t)= λ ($e^t - 1$)

指数分布 Ex(λ)　　　　　　　M(t)=$\frac{\lambda}{\lambda-t}$　(t< λ)

二項分布 Bi(n, p)　　　　　　M(t)=(q+pe$^{t})^{n}$　　(q=1-p)

モーメント母関数の不便な点をいえば、定義で積分が存在しないケースであり、そのいい例が指数分布である。もっとも t=0 での微分のためにはその近傍で存在すればよいから、その限りでは実用上の問題はない。

5 .2.2 確率分布の特性関数　chf　(characteristic function)

モーメント母関数において、助変数 t を it に代えれば特性関数が得られる。

＜定義：特性関数＞　X の確率分布を$f(x)$とするとき、t を助変数として

$$\varphi_X(t)=\int_{-\infty}^{\infty} e^{itx} f(x)dx$$

を X の（あるいはfの）特性関数(characteristic function)という。

<定理：対応>　　　　φ ： f(x)－＞φ_X（t）　は 1 対 1

<定理：和の mgf>　X,Y が独立で、X, Y の chf をそれぞれ$\varphi_X(t)$，$\varphi_Y(t)$とするとき、X+Y の chf は$\varphi_X(t)\ \varphi_Y(t)$で与えられる。

<例＞　　モーメント母関数と並行的である。

正規分布 N(μ，σ^2)　　　φ (t) = $e^{i\mu t-\sigma^2\frac{t^2}{2}}$

ポアソン分布 Po(λ)　　　φ (t) = λ($e^{it}-1$)

指数分布 Ex(λ)　　　　　φ (t) = $\frac{\lambda}{\lambda-it}$

二項分布 Bi(n, p)　　　　φ (t) =(q+pe$^{it})^{n}$

特性関数 chf はモーメント母関数 mgf と形の上では並行的な実質繰り返しで、存在理由がわからない向きがあるかもしれない。しかし、モーメント母関数は実用には十分であるが、数学的には特性関数は数々の理由で優れている。まず｜e^{itx}｜＝1 から積分の存在の正則条件が保証される。また、この被積分関数が有界のため数々の収束定理が使える。また、複素関数として解析関数となることを追求してもよい。

＜無限分解可能と Lévy‐Khintchine の表現＞　ここで、無限分解可能性の Lévy Khintchine の表現の chf の形が大方理解できる。正規分布（さらにはポアソン分布）の chf の形および和の chf から：

X, Yが独立に正規分布$N(\mu_1, \sigma_1^2), N(\mu_2, \sigma_2^2)$にしたがうとすると、$X+Y$は

$$\mu = \mu_1 + \mu_2, \qquad \sigma^2 = \sigma_1^2 + \sigma_2^2$$

とおいて$N(\mu, \sigma^2)$にしたがうことが、モーメント母関数、あるいは特殊関数から導かれる。すなわち $X+Y$の分布はたたみこみで

$$N(\mu, \sigma^2) = N(\mu_1, \sigma_1^2) * N(\mu_2, \sigma_2^2)$$

であるが、これは正規分布が 2 つの正規分布に任意に分布において分解できることを示す。もちろん、さらに任意個数 n の多重分割もできることは明らかである

同様にポアソン分布 Po(λ)おいても、任意に$\lambda = \lambda_1 + \lambda_2$ と分解すれば、

$$\mathrm{Po}(\lambda) = \mathrm{Po}(\lambda_1)\ \mathrm{Po}(\lambda_2)$$

とこれまた任意に分布において分解できる。しかも、ポアソン分布においては飛躍は正で 1 単位であるが、Lévy - Khintchine 表現ではより実世界に近く$-\infty < u < \infty$（ただし u ごとにλは異なる）としており、より完璧に実世界の描像に近づいていることがわかる。

ここで、あらためて Mumford が無限分解可能性を自然の実世界の確率論イメージを表現するモデルとして重要視したことを次のようにまとめておこう。

最高度の複雑性を持つ信号の例は自然科学の信号であり、スケールが大きいものから小さいものまでひろがっている。このとき、どんな現象もこれより小さい同様の独立な確率的変動の和からできている、いいかれば、どんな現象もどこまでも無限に分解しうる（無限分解可能 infinitely divisible)という著しい基本的性質が普遍的に観察される。

6 確率変数の極限定理

確率論は確率分布の表し方や各論を中心とした基礎の理論でスタートするが、その準備のあとは確率変数の演算が主役になって来る。その始まりは確率変数の和あるいは平均、これら変数の個数 n が多くなってくるどういう振舞になるか、そして確率変数の数列、その極限などと本格化する。その先に、大数の法則、中心極限定理がなどが成立し、まさに「確率論」の真ん中へ入る。これらを徹底的に応用したのが統計学であるが、もう一つの進展はランダムウォーク、Brown 運動などの興味深い領域が広がる。はじめから見て行こう。

6.1 確率変数の和：期待値 E と分散 V

ほとんどの場合、不確実性は期待値 E と分散（１， ２次のモーメント）で見てゆくのだが、まず

$$E(X+Y) = E(X) + E(Y)$$

は無条件である。しかし

$$V(X+Y) = V(X) + V(Y)$$

は無条件ではなく独立の条件下であり、一般には共分散 Cov が入り

$$V(X+Y) = V(X) + V(Y) + 2Cov(X,Y)$$

である。この場合（相関がある場合）はかなり違った展開になるので、ここでは、独立性を仮定し

$$S_n = X_1 + X_2 + \cdots + X_n \ , \qquad \bar{X} = \frac{X_1 + X_2 + \cdots + X_N}{n}$$

につき、もし$X_1, X_2, \cdots X_n$ が独立で同一の分布（independently, identically distributed, *i.i.d.*)なら、

$$E(X_1) = \cdots = E(X_n) = \mu, \qquad V(X_1) = \cdots = V(X_n) = \sigma^2$$

であって、和を繰り返せば n 倍になり、あるまとまった中間的結果

$$E(S_n) = n\mu, \quad E(\bar{X}) = \mu \quad ; \qquad V(S_n) = n\sigma^2,\ V(\bar{X}) = \frac{\sigma^2}{n}$$

となる。$V(\bar{X})$については、分散の性質から分母にｎが残る点が重要である。これよりｎ→ ∞のとき$\bar{X}$はμ一点に集中傾向となり、次の重要２法則がなりたつ。

6.2 大数の法則 Law of Large Numbers

1700 年代には経験的に発見されていたが、J.Bernoulli が *Ars Conjectandi* で最初の古典的証明を与えた。

＜定理：大数の弱法則＞ ｎ→ ∞ のとき、 $\bar{X}$ → μ （確率収束）

＜定理：大数の強法則＞ ｎ→ ∞ のとき、 $\bar{X}$ → μ （概収束）

弱法則がほどよい強さで実用上有用で（例：統計分析）証明も数行で済むが、強法則は「ほとんどいたるところ almost everywhere」（a.e.）あるいは a.s.(almost surely)などと表現され、証明は容易でない（触れていないテキストも多い）。

なお、確率変数の収束の位相には、そのほか、分布の収束（法則収束）、２次の平均収束（L_2 での収束）がある。後者は今後確率積分の近似の位相で登場する。

6 .3 中心極限定理 Central Limit Theorem

i.i.d.確率変数の和 $S_n = X_1 + X_2 + \cdots + X_n$（平均 0 、分散 1 に標準化後）に確率論史上最大の発見として ubiquitous な「中心極限定理」central limit theorem

$$1)\quad \lim_{n\to\infty} P\ \left[a \leqq \frac{X_1 + \cdots + X_n - n\mu}{\sqrt{n}\sigma} < b\right] = \int_a^b \frac{e^{-\frac{y^2}{2}}dy}{\sqrt{2\pi}} \qquad \text{（標準正規分布）}$$

が成立する。大数の法則は n で割るが、ここでは「弱く」$\sqrt{n}$で割っており結果は拡がる（分布になる）のである。定理はすでに 18,19 世紀に「*De Moivre - Laplace* の定理」として知られ、二項分布の精密計算からの古典的証明がある（いくつかのテキストでは紹介）。20 世紀に入ってから一般的になり名称も Central Limit Theorem となり、今日、経済学、工学、医学、心理学をはじめ確率・統計計算のある所多くの領域でこの正規分布仮定の結果の下に理論計算がなされている。特に金融数理ではほぼ普遍的な理論前提（対数正規分布含む）とされている。

古典的証明は二項分布については Stirling の公式で難なくたどり着くが、定理は分布を仮定せずこの証明は部分的である。以下、モーメント母関数が一般の証明の役に立つ。

期待値μ，標準偏差σ（分散σ²）で標準化し（X-μ）/σを再び Xi とすれば、X のモーメント母関数では、あらためてμ=0，σ=１であるはずだから、テイラー展開は

$$M\ (t) = 1 + \frac{t^2}{2} + \cdots$$

であり、（$X_1 + X_2 + \cdots + X_n$）/$\sqrt{n}$のモーメント母関数はｎ重積でｔをｔ/$\sqrt{n}$に置き換えれば

$$\left\{M\ \left(\frac{t}{\sqrt{n}}\right)\right\}^n = \left\{1 + \frac{t^2}{2n} + \cdots\right\} \to e^{\frac{t^2}{2}} \qquad (n \to \infty)$$

この右辺は標準正規分布Ｎ（０，１）のモーメント母関数である。ここで

$$\left(1+\frac{1}{n}\right)^n \xrightarrow{n\to\infty} e,\quad \left(1+\frac{a}{n}\right)^n \xrightarrow{n\to\infty} e^a \qquad (n \to \infty)$$

を用いている。

7. ランダムウォークと Brown 運動

確率現象に興味、関心があるなら、「ランダムウォーク」Random walk と「Brown 運動」Brownian motion を知る人は多いだろう。世界を確定論で割り切ることができないという考え方も急速に力を得ている中で、確率論の「ランダムウォーク」と「Brown 運動」は数理の常識に加えられ、今後はますますその傾向は強まるだろう。つまり、数理の人間には「現代人の常識」である。

7.1 ランダムウォーク

独立で

$X_i=1$ (確率 p で), —1 (確率 q =1- p で) $i = 1,2,\cdots,\ n$

である確率変数列に対し、確率過程

$$S_0 = 0,\quad S_n = X_1 + X_2 + \cdots + X_n, \qquad n = 1,2,\cdots$$

を「(単純) ランダムウォーク」simple *random walk*、とくに p=q=1/2 のとき「対称ランダムウォーク」symmetric *random walk* という($S_0 = 0$ は便宜である)。 容易に

$$E(S_n) = n(p-q),\qquad V(S_n) = 4npq,\qquad n = 1,2,\cdots$$

を得るが、いずれも「時間」n に比例する結果である。対称なら、さらに簡単に

$$E(S_n) = 0 \qquad V(S_n) = n,\quad n = 1,2,\cdots$$

もし関心あるなら、この場合、

・$Cov(S_n, S_m) = n^m$

・S_nはマルチンゲール

などを確認しておくと、今後のために有益である。

＊拙著『入門 確率過程』東京図書

7.2 Brown 運動

ランダムウォークにおいては時間 n = 1,2,・・・は離散的であるが、連続時間 t にすれば Brown 運動 W(t)が構成される。それには[0,t]を微小間隔Δt で区切り n=[t/Δt]とすればよい。Δt —> 0 とするので n->∞となり、中心極限定理が働く。空間刻みも別途 1 から微小Δr に縮小するのが自然である。しかして、対称ランダムウォークS_nを元に

$$W(t) = S_{[t/\Delta t]}\cdot\Delta r$$

と定義すれば、

$$E(S_n) = 0,\quad V(S_n) = n$$

から

$$E\big(W(t)\big) = 0\ ,\qquad V\big(W(t)\big) = (\Delta r)^2\cdot\left(\frac{t}{\Delta t}\right)$$

ここで、ともにΔt, Δr —> 0とするが, 不定とならないようにΔr $=\sqrt{\Delta t}$の関係を保つものとする。E,V は定まったが、W(t) の分布にはΔt—> 0 から中心極限定理が働く。Δt を極限移行の整数パラメータ

N—> ∞によって$\Delta t=1/N$ —> 0 のようにとれば、$\Delta r = 1/\sqrt{\mathrm{N}}$である.これより$W(t)$は

$$\frac{S_{[Nt]}}{\sqrt{N}} = \sqrt{t}\cdot\frac{S_{[Nt]}}{\sqrt{Nt}}$$

であり、N—> ∞のときその分布は中心極限定理から N(0,1)変数の$\sqrt{t}$倍、すなわち正規分布 N（0,t）となる。これを「標準 Brown 運動」standard Brownian motion,またσW(t)+μ t は正規分布 N（μ t，σ^2t）に従い、ドリフト（drift）μ，ボラテリティ(volatility)σの Brown 運動という。

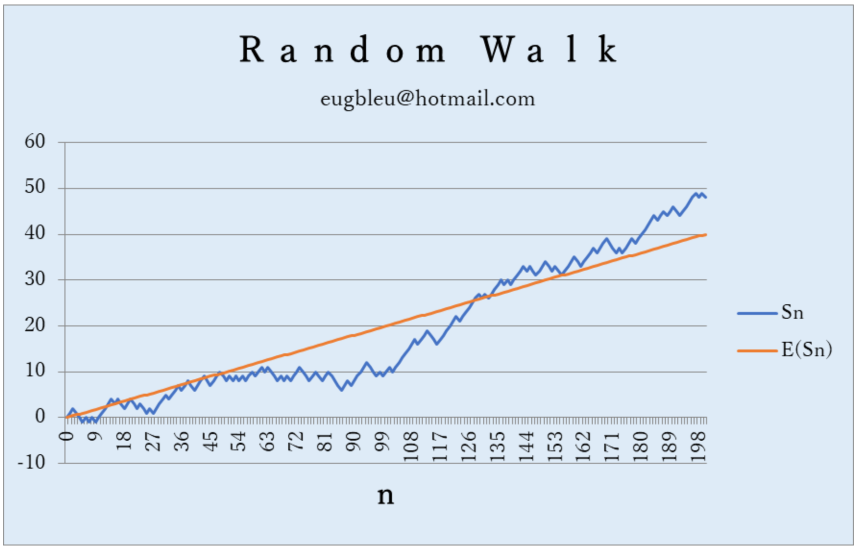


図１　ドリフトのある一次元ブラウン運動の標本径路

以上は構成的定義であるが、これから導かれる以下の性質が多くのテキストで「Brown 運動」の定義とされている。

＜Brown 運動の性質＞　確率過程 W(t), t≧0 が

$$W(0) = 0$$

さらに任意の n に対し、n+1 個の時間の分点

$$0 \le t_0 < t_1 < \cdots < t_{n-1} < t_n$$

で作った変化分（増分）

$$W(t_1) - W(t_0), W(t_2) - W(t_1), \cdots, W(t_n) - W(t_{n-1})$$

が独立であり、かつそれぞれ正規分布

$$N\left(\mu(t_k - t_{k-1}), \sigma^2(t_k - t_{k-1})\right) \ \ (k = 1,2,\cdots,n)$$

に従う場合を「Brown 運動」という。

したがって、W(t)は任意の独立増分の和となる。この性質を満たす確率過程を「加法過程」という。伊藤の『確率論』（旧版）がマルコフ過程に先立って加法過程から出発しているのは重要である。なぜなら、W(t)は本式に見ての通り文字通り無限分解可能の可視的な例となっているからである。

7.3 W(t)の標本径路の性質

W(t)は確率的にそのたびに異なった（それでいて一定の確率法則で）t の関数として生じるが、現に生じたある W(t)を一般に（統計学の用語から）「標本径路」Sample path という。標本径路はそれ

自体ある t の関数である。

標準 Brown 運動の標本径路には３通りの著しい特異な基本性質がある（証明は略）。

＜定理 I：連続性＞　確率 1 で（a.s.）連続 continuous

＜定理 II：微分不可能性＞確率 1 で（a.s.）至る所微分不可能 nowhere　differentiable

＜定理 III：変分＞　[0,t]の分点幅をΔ_k、W(t)の増分をΔW_kとすると、max Δ_k →０の極限で

(i)　絶対変分　$V_1 = \sum_k |\Delta W_k| = \infty$

(ii)　２次変分　$V_2 = \sum_k |\Delta W_k|^2 < \infty$ $\left(2\text{ 次変分有限}\right)$, $V_2 \to t$ (全区間幅)

一般に連続関数は積分可能であるから、I は以下の展開のための重要な前提である。II が主要な論題であるが、難点回避の理論戦略として I に基づき積分（確率積分）＝＞微分という方法を採用する。その際、確率積分の成立に III(ii)が重要役割を果たす。３性質に基づき、以下の Stieltjes 式確率積分、確率微分方程式が定義される。

なお、この際、不連続な標本径路をもたらす典型として、導出を省略して「ポアソン過程」も図２に示しておこう。ポアソン過程は通常ジャンプ高 u＝１であるが、－∞＜u<∞を許せば、これも無限分解可能性の要素の目に見える例になっている。

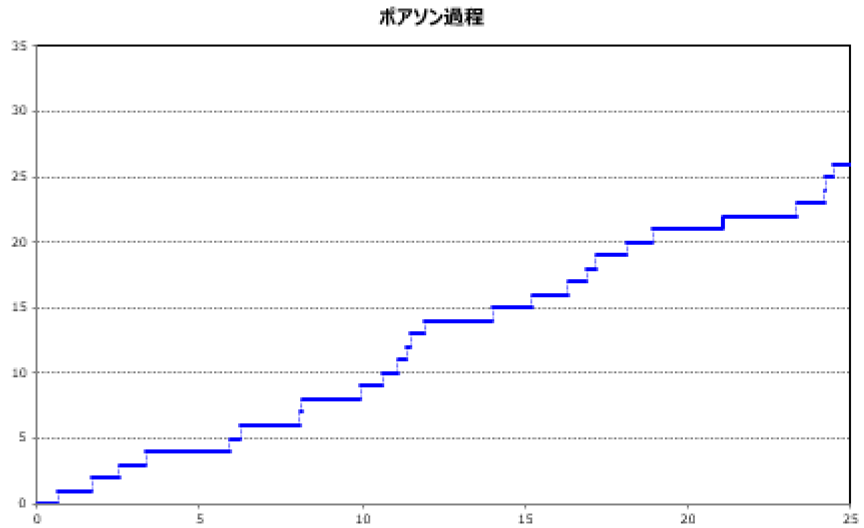


図２　　ポアソン過程の標本径路

8.確率積分 Stochastic integral

8.1　定義

標準 Brown 運動 W(t)を基礎として確率論的な微分積分学が成りたち、「確率積分」、「確率微分方程式」（Stochastic differential equation, SDE）が定義され、 今日さまざまな理論と応用がある。 今後ますますその領域は広まると思われるが、要点だけ述べておこう。

まず、積分区間を分点で

$$\alpha \equiv t_0 < t_1 < t_2 < \cdots t_n \equiv \beta$$

のように分割し（必ずしも等分でなくてもよい）、各区分幅での W(t)の増分

$$W(t_1) - W(t_0),\ \ W(t_2) - W(t_1),\ \ W(t_n) - W(t_{n-1})$$

を

$$\Delta W(t_o), \qquad \Delta W(t_2), \qquad \Delta W(t_{n-1})$$

とし、他方被積分関数ｂ（以下でみるようにこれ自体確率的でもよい）を

$$b_0, b_1, \cdots, b_{n-1}$$

として、それ自体確率的な積和

$$b_0 \Delta W(t_0) + b_1 \Delta W(t_1) + \cdots + b_{n-1} \Delta W(t_{n-1}) = \sum_{k=0}^{n-1} b_k \Delta W(t_k)$$

をつくる。ちなみに E(W(t))＝０からこの期待値＝０だが、分散は

$$\sigma_b^2 = \sum_{k=0}^{n-1} b_k^2 (t_{k+1} - t_k)$$

となる。この積和の、分点を一様に細かくした極限が

$$\sum_{k=0}^{n-1} b_k \Delta W(t_k) \rightarrow I_b \quad (n \rightarrow \infty) \qquad (L_2)$$

なら、これを記号

$$I_b = \int_\alpha^\beta b(t) dW(t)$$

であらわす。これが関数bの「確率積分」である。 なおこの極限は、２次の平均収束

$$E\left(\sum_{k=0}^{n-1} b_k \Delta W(t_k) - I_b\right)^2 \rightarrow \quad 0$$

であって、

$$\sum_{k=0}^{n-1} b_k \Delta W(t_k) \fallingdotseq \int_\alpha^\beta b(t) dW(t)$$

とあらわすときは注意する。

8.2 計算実例

確率積分はある程度までのところなら定義通りで計算できるが、そこから先は全微分であらわして「伊藤の補題」による。たとえば、

$$\int_0^s W(t) dW(t) = \frac{1}{2}\big(W(s)\big)^2 - \frac{1}{2} s$$

を導く。定義から

$$\sum_{k=0}^{n-1} W(t_k) \Delta \mathrm{W}(t_k)$$

求めるが、中途を省略すると、

$$W(s)^2 = \sum_{k=0}^{n-1}\left(\Delta W(t_k)\right)^2 + 2\sum_{k=0}^{n-1} W(t_k)\Delta W(t_k)$$

を得るのは煩雑だがむずかしくない。ここで、分点を細かくすると、右辺第１項は２次変分であるから、

$$\left(W(s)\right)^2 = s + 2\int_0^s W(t)dW(t)$$

となる。s がなければ、通常の積分

$$\int_0^y F(s)d(x) = \frac{1}{2}\left(F(y)\right)^2$$

となるが、そこが確率積分の大きな特質である。同様に

$$\int_0^s \left(W(t)\right)^2 dW(t) = \frac{1}{3}((W(s))^3 - \int_0^s W(t)dt$$

も言えるが

$$\int_0^y \left(F(x)\right)^2 dF(x) = \frac{1}{3}\left(F(y)\right)^3$$

とは似ているが異なる。なお

$$\int_0^s tdW(t) = sW(s) - \int_0^s W(t) \qquad \text{(部分積分)}$$

も得られる。

以上を、見方を変えて

$$\left(W(s)\right)^2 = \int_0^s 1dt + \int_0^s 2W(t)dW(t)$$

$$\left(W(s)\right)^3 = \int_0^s 3W(t)dt + \int_0^s 3W(t)dW(t)$$

$$sW(s) = \int_0^s W(t)dt + \int_0^s tdW(t)$$

とすると、右辺にdt ，$dW(t)$という微小増分（全微分という）があらわれる。これらの増分の関係を。例えば３番目については、右辺を全微分にして、全微分の関係式として記号的に、

$$\boldsymbol{d}\left(sW(s)\right) = W(s)\boldsymbol{ds} + s\boldsymbol{dW}(\boldsymbol{s})$$

とあらわす一般的約束が行われているが、あくまで記号であって（それ自体の意味はない）元は確率積分である。 そこで、W(t)の一般的な変換公式が望まれるところである。

＜定理：伊藤の補題＞ ds も入れて

$$dX(s) = A(s)ds + B(s)dW(s) \qquad \text{(伊藤過程)}$$

としておこう。$A(s) \equiv$ ０，$B(s) \equiv$ １ならで X(s)≡W(s)で問題ない。変換gが（ t が入ってもよい）

$$Y(t) = g\left(t, X(t)\right)$$

ならば、$dY(s)$は、

$$dY(s) = \frac{\partial g}{\partial s}\left(s, X(s)\right)ds + \frac{\partial g}{\partial x}\left(s, X(s)\right)dX(s) + \frac{1}{2}\frac{\partial^2 g}{\partial x^2}\left(s, X(s)\right)\left(dX(s)\right)^2$$

ただし、 dX(s)を代入する際

$$(ds)^2 = 0, \qquad dsdW(s),\ dW(s)ds = 0, \qquad \left(dW(s)\right)^2 = ds$$

と約束する。よって、X(s)≡W(s)の場合は、簡潔に

$$dY(s) = \ \{ \frac{\partial g}{\partial s}\left(s, X(s)\right) + \frac{1}{2}\frac{\partial^2 g}{\partial x^2}\left(s, X(s)\right)\}\ ds + \frac{\partial g}{\partial x}\left(s, X(s)\right) dW(s)$$

を得る。

本式は２次のテイラー展開であるが、見ての通り２次の３項のうち$\left(dW(s)\right)^2$だけが残る。この項を ds とするのは奇異だが、これも Brown 運動の２次変分から来ている。公式の適用例は割愛しよう。

＊多くの場合そうであるように、人名からの「○○の定理（補題）」も本人によるものではない。本定理も『確率論』（旧版）の定理 65.1 とされているだけである。

9. Ornstein-Uhrenbeck の Brown 運動モデル（伊藤、国澤）

ここまでの Brown 運動の発展的応用として、有名な Ornstein-Uhrenbeck モデルがある。これは、Brown 運動に速度に比例する摩擦抵抗力が働いているモデルで、従来から Langevin 方程式として知られていた。しかし、標本径路は微分不可能で速度は定義できず、その解き方が大変有用なものである。モデルは以前から知られ、伊藤のテキストでも最終に近い第 69 節に簡単に触れている部分がある。最近は金融モデル（Vasicek モデル）にも応用されることになって多くの確率論テキストが取り上げている。そこで重要部分を紹介することにしよう。

ここで、$U(t)$は速度として、第１項にマイナスがついていることが抵抗を表している：

$$dU(t) = \ -\alpha U(t)dt + \ \sigma dW(\tau)$$

これを解くために、伊藤の補題で

$$\mathrm{g}\ (t, x) = \ e^{\alpha t}x$$

とおくと

$$d\left(e^{\alpha t}U(t)\right) = 0 \cdot dt + \ \sigma e^{\alpha t}dW(t)$$

となる。したがって、これを$[s, t]$で積分して

$$e^{\alpha t}U(t) - \ e^{\alpha s}U(s) = \ \sigma \int_S^t e^{\alpha \tau}dW(\tau)$$

両辺を$e^{\alpha t}$で割れば

$$(*) \quad U(t) - \ e^{\alpha(s-t)}U(s) = \ \sigma \int_S^t e^{-\alpha(t-\tau)}dW(\tau)$$

ここで

$$\int_{-\infty}^S \left|e^{-\alpha(t-\tau)}\right|^2 d\tau \ < \ \infty$$

なので、（＊）は$s \to -\infty$のとき、２次の平均収束で

$$U(t) - \ e^{-\alpha t}V = \ \sigma \int_{-\infty}^t e^{-\alpha(t-\tau)}d\tau$$

のように書き換えることができる。さらにもし、$E(U(s)^2)$が s について有界なら V=0 となるから、結局

$$U(t) = \sigma \int_{-\infty}^{t} e^{-\alpha(t-\tau)} dW(\tau)$$

が得られる。これが Ornstein-Uhrenbeck の Brown 運動の速度の表現式である。

次の性質がある。

① $U(t)$は Brown 運動である。

② $E\big(U(t)\big) = 0$

$$\mathrm{Cov}\big(U(t), U(s)\big) = \sigma^2 \int_{-\infty}^{t\wedge s} e^{-\alpha(t-\tau)} e^{-\alpha(s-\tau)} d\tau = \frac{\sigma^2}{2\alpha} e^{-\alpha|t-s|} \quad (t \neq s)$$

すなわち、自己共分散関数は

$$\gamma(\tau) = \frac{\sigma^2}{2\alpha} e^{-\alpha|\tau|}, \quad (\tau \neq 0)$$

したがって、ここでの範囲外だが、スペクトル密度はそのフーリエ変換

$$f(\lambda) = \int_{-\infty}^{\infty} \gamma(\tau) e^{-i\lambda\tau} d\tau = \frac{\sigma^2}{\pi} \frac{1}{\lambda^2 + \alpha^2}$$

と求められる。

このように OU 過程は伊藤、国澤、 $Breiman, Feller$, $en.wikipedia$などで扱われているが、この確率微分方程式をマルコフ過程の推移確率として解くための「前向き方程式」(Forward equation) は

$$\frac{\partial}{\partial t} f(y; t \mid x; s) = -\frac{\partial}{\partial y}\{A(t,y) f(y; t \mid x; s)\} + \frac{\partial^2}{\partial y^2}\{B^2(t,y) f(y; t \mid x; s)\}$$

である。ここで瞬間的推移先の期待値、分散をそれぞれ

$$A(t,y) = -\alpha y, \qquad B^2(t,y) \equiv \sigma^2/2$$

とすると、$f(y) = f(y; t \mid x_0; 0)$ $(t > 0)$ が満たすべき偏微分方程式は

$$\frac{\partial f}{\partial t} = \alpha \frac{\partial (yf)}{\partial y} + D \frac{\partial^2 f}{\partial y^2} \qquad (D = \sigma^2/2)$$

となる。（Fokker-Plank の方程式）この方程式を解くのは易しくないが、解法の解説はある（国澤）。それによると、$f(y)$は正規分布

$$N\left(x_0 e^{-\alpha t}, \quad \frac{\sigma^2}{2\alpha}(1 - e^{-2\alpha t})\right)$$

となる＊。すなわち、時間の経過とともに出発点x_0へ戻る傾向（0 なら原点復帰傾向 mean reverting property）があるが、ばらつきは時間とともに漸増する。

＊ただし、国澤 p.227 の$f(x: t \mid x_0, 0)$の式には軽微なミスがある）

10. Ito-McKean『拡散過程と標本径路 』原著（第 1 章）から Lévy を理解する

ここまでの内容はヒューリスティクスであって、数学的厳密性は最優先していない。ここからは数学的に構成された理論内容を紹介しよう。そのエッセンスは標本経路から出発して数学的な関数空間論へ発展するもので、Ito-McKean の研究書は歴史的な著作としてこの方面の研究者が引用するが、やはり相当に高度で読みごたえがある。Mumford の今後の発展解説の準備として、今回ここまでの内容に関

連する事項のみをかい摘んで要約しよう。以下、節番号は伊藤『確率論』に対応させるものとする。

10.1　1次元標準ランダムウォーク

$\boldsymbol{\mathcal{W}}$ は標本径路 Sample path（以下「径路」）

$$\omega:\quad n = 0,1,2,\ldots. \quad \rightarrow\ s_n = s(n) = 0\,,\pm1,\pm2,\ldots.$$

の空間であり、$\mathcal{F}$は径路ωの集合

$$\{\omega : s(n) = l\ \}\ n \geqq 0,\quad l = 0,\pm1,\pm2,\ldots.$$

の全体で生成される完全加法族（σ-加法族、σ-代数、ボレル代数）とする。いま、
径路 s(n), n≧0 が出発点 s(0)= ℓ =0,±1,±2, ・・・から出発して、 s(n)=k となる確率を

$$P_l\Big[\ s\,(n) = k\Big] = \binom{n}{n_+}(1/2)^n$$

ただし

$$n_+ = \frac{n}{2} + \frac{1}{2}(k-l),\quad \binom{n}{n_+} = \frac{n!}{(n-n_+)!}n_+!\quad \text{for}\ n_+ = 0,1,2.\ldots,n$$

と定義し、$\mathcal{F}$上に確率P_lを導入する。そこで、$\boldsymbol{\mathcal{D}}$= ($\mathbf{W}$,$\mathcal{F}$ P_l,： $l = 0,\pm1,\pm2,\ldots.$ ）

を**標準ランダムウォーク** standard random walk という。

さらに、これに基づき、m≧0 に対し、

$$P_{.}\ [s(m+1) = l\,|\,s(n){:}\,n \leqq m] = P_{S(m)}[s(1) = l]\,;\quad l = 0,\pm1,\pm2,\ldots.$$

が成り立つ。すなわち、$\boldsymbol{\mathcal{D}}$ は（一重）マルコフ過程である。

なお、ここでは述べないが続いて「マルコフ時刻(間)」Markov time が定義されている。これは応用確率論で「停止時間」stopping time といわれるものと同一意味である（この方がマルチンゲールを論じる場合などをはじめ理解しやすい。）

10.2　通過時間（Passage time）　略

「通過時間」は直訳であるが、実際には（最初の）到達時間である。なぜなら、到達しなければ通過はないからである。実際、Feller は *first passage* といっているが、passage times という語は見あたらない。

10.3　拡散方程式から de Moivre-Laplace 型中心極限定理を導く（Khintchine, Feller）

径路 s(n)に、確率論史上最大の発見として ubiquitous な「中心極限定理」central limit theorem

$$2)\quad \lim_{n\to+\infty} P_0\left[a \leqq \frac{S(n)}{\sqrt{n}} < b\right] = \int_a^b \frac{e^{-\frac{y^2}{2}}dy}{\sqrt{2\pi}}\qquad（正規分布）$$

が成立することはすでに 18,19 世紀に「*De Moivre - Laplace* の定理」として知られ（原型は J.Bernoulli の *Ars Conjectandi*)、二項分布の精密計算からの古典的証明がある。今日ではモーメント母関数、特性関数、あるいは Sterling の公式による。s(n)が n 回の試行の結果径路であるにもかかわらず、除数が n でなく$\sqrt{n}$であるところが意味深淵で

ある。

ところで、Khintchine[11]は径路の性質から拡散方程式を導き、その解として目的のガウス分布（正規分布）を導出する巧妙で天才的証明を与えており、径路の内在的性質を元にする Ito-McKean の接近法にふさわしい結果である。

まず、　u_nを、位置 a,b は$1/\sqrt{n}$単位、時間 t は1/n単位で測った上で、ℓ ＝[$(\sqrt{n})$a]を出発点とするランダムウォークの確率で期待値演算から

$$f \in C(R^1) \quad \text{に対し，} u_n(t,a) = E_{[\sqrt{n},a]}[f(s_{[nt]}/\sqrt{n})] \quad t \geqq 0, a \in R^1$$

と定義する。 f に対する作用（演算） E (f)から見てゆく関数解析の方法をとっている。

u_nは、中途を省略して

$$u_n\left(t+\tfrac{1}{n},a\right) = \tfrac{1}{2}u_n\left(t,a-\tfrac{1}{\sqrt{n}}\right) + \tfrac{1}{2}u_n\left(t,a+\tfrac{1}{\sqrt{n}}\right), \qquad \text{[原著 p.10]}$$

さらに工夫して、$1/\sqrt{n}$単位の 2 階差分方程式

$$n \times [u_n\left(t+\tfrac{1}{n},a\right) - u_n(t,a)]$$

$$= \tfrac{1}{2}(\sqrt{n})^2 \times [u_n\left(t,a+\tfrac{1}{\sqrt{n}}\right) - 2u_n(t,a) + u_n\left(t,a-\tfrac{1}{\sqrt{n}}\right)]$$

$$u_n(0,b) = f([\sqrt{n}b]/\sqrt{n})$$

を与える。これより、所期の解

$$u = u(t,a) = \int_{R^1} \frac{e^{-\frac{(b-a)^2}{2t}}}{\sqrt{2\pi t}} f(b)db$$

が、n→∞の 2 階偏微分方程式移行から拡散方程式

$$\frac{\partial u}{\partial t} = \frac{1}{2}\frac{\partial^2 u}{\partial a^2} \qquad \text{（拡散方程式）}$$

$$u\left(0+,\ \cdot\right) = f$$

を解いて得られる。すなわち

$$\lim_{n\to+\infty} u_n = u$$

である(証明略)。このようにして（離散的な）径路の空間に正規分布が与えられる。

10.4　連続関数の空間上に標準 Brown 運動を構成する

Brown 運動の標本径路は連続である。そのため、[0, ∞]上の連続関数 x_t

$$\omega : t \in \left[0,\ \infty\right] \to x_t = x(t) \in R^1$$

の関数空間 $\boldsymbol{w}$ に正規分布を導入し Brown 運動を構成する。連続確率過程の標準的構成法により、t に任意 n 個の分点$t_1, t_2, \dots t_n$を入れ、各分点毎に確率を指定する。まず、$\boldsymbol{w}$ の集合

$$C = x_t{}^{-1}(B) = x_{t_1 t_2 \dots t_n}{}^{-1}(B) \qquad t = (t_1, t_2, \dots t_n)\ ,$$

$$0 < t_1 < t_2 < \dots < t_n\ ,\ \mathrm{B} \in \boldsymbol{\mathcal{F}}[R^{\,n}],\ n \geqq 1$$

の族$\boldsymbol{\mathcal{C}}$を導入する。ここで$x_t{}^{-1}$は、各分点$t_1, t_2, \dots t_n$におけるx_tの射影

$$x_t: \quad \omega \to (x_{t_1}(\omega), \qquad x_{t_2}(\omega), \dots x_{t_n}(\omega)) \in R^n$$

の逆写像である。したがって、$C = x_t{}^{-1}(B)$はx_tの$t_1, t_2, \dots t_n$における振舞である。

つぎに正規分布（ガウス核）

$$\mathrm{g}(\mathrm{t}, a, b) = \frac{e^{-\frac{(b-a)^2}{2t}}}{\sqrt{2\pi t}} \qquad \mathrm{t} > 0,\ \ a, b \in R^1$$

を用意する。ここで$b - a$は正規変量（よって g は差の関数）、t は分散であることに注意する。x_tが各分点幅内で幅 $\mathrm{t_i}$-$\mathrm{t_{i\text{-}1}}$ に比例する分散で正規変量bに従い変動するとし

$$P_t(C) = \int_{\mathrm{B}} \dots \int \mathrm{g}(\mathrm{t}_1, 0, b_1)\mathrm{g}(\mathrm{t}_2 - \mathrm{t}_1, b_1, b_2)\ \mathrm{g}(\mathrm{t}_n - \mathrm{t}_{n-1}, b_{n-1}, b_n) db_1 db_2 \dots db_n$$

と定義すると、$\boldsymbol{\mathcal{C}}$が完全加法族であること、P_tが$\boldsymbol{\mathcal{C}}$上の確率測度となることが証明され。さらに Brown 運動の重要条件である各分点幅内での独立増分性は 5)の構成（増分の分布 g が積になっている）ことから定義そのものである。この（$\boldsymbol{w}, \boldsymbol{\mathcal{C}},\ P_t$)を**標準 Brown 運動**という（Ito-McKean では標準ランダムウォーク$\boldsymbol{\mathcal{D}}$と同一記号）

ここで、構成において分点$t_1, t_2, \dots t_n$の数 n はいくらでも大きくとれるから、Brown 運動はいくらでも小さい独立な正規増分の和に分割でき、Lévy のいう「無限分解可能」infinitely divisible の性質を持っている。

10.5　Haar 関数基底付きの Hilbert 空間に Brown 運動を構成（Lé vy）

関心ある話題であり Mumford も注目するところであって、池田信行『偶然の輝き』も触れている。フーリエ係数、wavelet など広い関連があるので、次の機会とする。

10.6　狭義マルコフ性　　略

10.7 標準 Brown 運動の通過時間

前述のように、実際には到達時間

$$m_a = \min(t: x_t = a) \quad a \in R^1$$

である。本来は min は inf であるが、x_tは連続な標本径路であるからこの定義となる。このモーメント母関数は、詳細は省くが、いわゆる「鏡映原理」を用いて

$$E_0(e^{-\alpha m_a}) = 2\,\alpha \int_0^{+\infty} e^{-\alpha t} dt \int_\alpha^{+\infty} \frac{e^{-\frac{b^2}{2t}}}{\sqrt{2\pi t}} db \ = \ e^{-\sqrt{2\alpha}\alpha} \qquad \text{[原著 p.26]}$$

と得られる。この結果の意義は、標本径路としては

$$m_a = \int_0^{+\infty} lp([0, a) \times dl) \qquad \text{（Poisson 測度による積分）}$$

であり、これが Lévy- Khintchine の無限分解可能分布を構成する正規分布、Poisson 分布以外の剰余項として出現することが重要である。実際、m_aは「片側安定過程」one-sided stable process であり、無限分解可能分布で重要な役割を果たすが、Mumford ではここまで

は配慮していない。次の機会とする。

10.8 Kolmogorov による重複対数の法則　略

10.9　L é vy-Hölder 条件　略

10.1　標準 Brown 運動によるランダムウォークの近似　略

参考文献

20 世紀より前の確率論史は割愛した。

[1] 伊藤清,「確率論」(旧版)，岩波書店，1953

[2] 国沢清典,「確率論 とその応用」 岩波全書, 1982

[3] 吉田耕作、加藤敏夫「応用数学 I」, 1962

[4] 佐藤健一,「加法過程」，紀伊国屋書店，1990

[5] 飛田武幸,「ホワイトノイズ」，丸善，2014

[6] 池田信行,「偶然の輝きーBrown 運動を巡る 2000 年」，岩波書店，2018

[7] 松原望,「入門確率過程」，東京図書，2003

[8] 藤田岳彦「大学生の確率・統計」, 東京図書, 2010

[9] 岩沢宏和,「分布からはじめる確率・統計入門」, 2016

[10] A.Kolmogorov *Grundbegriffe der Wahrscheinlichkeitrechnung*, Ergebn.Math. **2** No3, Berlin 1933

[11] A.Khintchine *Asymtotische Gesetz der Wahrscheinlichkeitrechnung*, Ergebn. Math. **2** No4, Berlin 1933

[12] H.Cramér, *Mathematical Methods of Statistics*, Princeton, 1945

[13] P. Lévy, *Processus stochastiques et movement brownien,* Paris, 1948

[14] Itô ,H.McKean, *Diffusion Processes and their Sample Paths,* Springer, 1964

[15] W.Feller, *An Introduction to Probability Theory and its Applications* Vo.II John Wiley, 1966

[16] L.Breiman, *Probability,* Addison-Wesley, 1968

[17] A.N.Shiryaev, *Probability,* Springer, 1989,

[18] D. Mumford, *The Dawning of the Age of Stochasticity*, 2000

[19] D. Mumford, Angnès Desolneux, *Pattern Theory*, CSC Press, 2010