数理解析研究所講究録
第 405 巻 1980 年 112-158

確率過程論における新概念導入の歴史

学習院大学理学部　伊藤 清

1. 確率過程という概念

「確率過程論における新概念導入の歴史」という標題をつけたが，実は「確率過程という概念」そのものが新概念であって，解析学における時変数 t の関数に対応する確率論的概念である.

決定論的（因果論的）立場では同一条件（原因）から同一結果がでるということになっている. ある条件の下である量の時間的変動を観察すると，唯一つの観測値関数が得られる. 同じ条件の下で，誰が，何回観測しても同じ関数が得られる. これが決定論的立場であり，こゝで得られる観測値関数は数学的には解析学における 1 変数（時変数）の関数であらわされる.

例えばある容器の中の水（温度は例えば 18°C とする）の表面の 1 点に玉をおくと，玉は水の抵抗を受けながら，重力

によって下向きに落ちて行く．この運動を表わす関数は，誰が何回観測しても同じで，原理的には流体力学の方程式を解いて得られるはずである．所がこの玉が極めて小さいときには（例えばコロイド粒子），この粒子は水の分子と衝突してジクザク運動をする．これが R. Brown が発見した有名な Brown 運動である（1828年）（江沢洋著：だれが原子をみたか（岩波）参照）．この場合には，同じ観測を何回も繰返すと，その度ごとに異なるジクザク運動が観測される．だから同一原因，同一結果という決定論的立場はとれない．

これに対して次のように反論する人があろう．18°C の水といっても，分子論的には同一条件ではない．分子論的同一条件は，各子の位置と運動量をすべて指定して始めて可能である．それをすれば，粒子の運動は唯一つの関数であらわされ，その関数は運動方程式を解いて得られるはずである．あるいはそうかも知れない．しかし 10^{24} のオーダーの数の分子の位置，運動量を指定して，実験を繰返すことは不可能である．我々が制御できるのは温度 18°C と容器の形で，これを巨視状態という．この巨視状態には無数の微視状態（すべての分子の位置と運動量）が対応する．確率論では個々の微視状態を見本（sample）といい，ω であらわし，その全体を見本空間（sample space）といい，Ω であらわす．個々の

微細状態 ω に対しては，運動方程式を解いて粒子の位置の時間的変位 $(X(t,\omega), Y(t,\omega), Z(t,\omega))$ が定まる．与えられた容器の中の 18°C の水の微細状態を何回も観測したと仮定すると，Ω の点の列が得られるはずであるが，Ω のある部分は表われ易く，ある部分は表われ難く，Ω の上の確率分布（確率測度）P が得られると考えられる．勿論実際にはこんな観測は不可能であって，物理的考察から P についていろいろの仮説がある．有名なものは Maxwell の microcanonical distribution と Gibbs の canonical distribution である．粒子の数が極めて多いことを考慮に入れると，両者は本質的にあまり違わない（A. Khinchin：Mathematical Foundations of Statistical Mechanics, 1949 参照）．何れにせよ，Ω と P との対 (Ω, P)（確率空間）が与えられた容器の中の 18°C の水という巨視状態に対応する微細状態の確率的表現である。

個々の見本（微細状態）ω に対して運動方程式を解いて定まった上記の $(X(t,\omega), Y(t,\omega), Z(t,\omega))$ から，つぎの関数（t と ω との関数）が得られる．

$$X(t,\omega),\quad 0 \leq t < \infty,\quad \omega \in \Omega = (\Omega, P)$$
$$Y(t,\omega),\quad \text{〃},\quad \text{〃}\quad \text{〃}$$
$$Z(t,\omega),\quad \text{〃},\quad \text{〃}\quad \text{〃}$$

この各々は Ω の上の確率過程とよばれる．

この際 $X(t,\omega)$, $Y(t,\omega)$, $Z(t,\omega)$ の具体的な関数形には我々は関心を持たない．それは観測し得ないものであるから．では観測できるものは何か．例えば，t を固定したとき，この粒子が E の中に入る確率（E は $\mathbb{R}^3$ の部分集合）：

$$Pr.\{(X(t), Y(t), Z(t)) \in E\} \equiv P\{\omega : (X(t,\omega), Y(t,\omega), Z(t,\omega)) \in E\} \quad (1)$$

である．この値を得るには，何度も観測して，時間 t の後に E の中に粒子が観測される頻度を調べたらよい．何度も観測するのが繁雑であれば，多数の粒子をおいて，時間 t の後に E の中に入るものの割合を見てもよい．後者の場合には粒子相互の影響があるから，正確には前者と一致しないが，近似的には十分である．

上の (1) は E の関数と見て，$\mathbb{R}^3$ 上の確率測度となるが，これを $X(t)$, $Y(t)$, $Z(t)$ の結合分布といって $P_{X(t),Y(t),Z(t)}(E)$ とかくこともある．同様に $X(t)$ のみに着目して，$X(t)$ の確率分布

$$P_{X(t)}(E) = P\{\omega : X(t,\omega) \in E\}, \qquad E \subset \mathbb{R}^1,$$

も考えられる．またいくつかの時点 $t_1, t_2, \cdots, t_n$ で観測して，その結合分布

$$P_{X(t_1), Y(t_1) Z(t_1) \cdots X(t_n) Y(t_n) Z(t_n)}(E), \qquad E \subset \mathbb{R}^{3n},$$

も同様に考えられる．

このような考察を背景として，確率過程は数学的に次のように定義される．$\Omega = (\Omega, P)$ を任意の確率空間とし，時変数 $t \in T$，見本 $\omega \in \Omega$ の関数 $X = X(t,\omega)$ を (Ω, P) 上の確率過程という．こゝで T は実数の区間のこともあり，整数の集合のこともある．これは観測を連続的に行うか，断続的に行うかによる．上記の確率分布 $P_{X(t)}(E)$ や結合確率分布 $P_{X(t_1)\cdots X(t_n)}(E)$ が，少なくとも E が区間または矩形集合のときに定義できるように，写像

$$X_t : \Omega \to R \qquad \omega \longmapsto X(t,\omega)$$

の P-可測性を仮定しておく．これは各時点 t における X の値 $X_t = X(t,\omega)$ が $\omega \in \Omega$ の P-可測関数（確率論の言葉でいえば，Ω 上の確率変数）であることを意味する．このような確率過程の数学的定義は Kolmogorov によって与えられたもので，これによって確率過程をその直観的描像（intuitive image）から独立に純粋論理的に研究することができるようになったのである．これは数学者にとって，理解するにも，討論するにも都合のよい言葉が与えられたことになり，数学的確率過程論の大きい推進力となった．しかし同時に数学者とそれ以外の科学者との間の溝も大きくなった．

2. 見本過程(見本関数)

確率過程が時変数 $t \in T$ と見本 $\omega \in \Omega = (\Omega, P)$ の関数として

$$X(t, \omega), \quad t \in T, \ \omega \in \Omega$$

の形であらわされることは前節に述べた. これをつぎの3通りに眺めることができる.

第一の見方は写像

$$X : T \times \Omega \to \mathbb{R}$$

と見ることで, これは最も単純な見方である.

つぎに個々の $t \in T$ に対して写像

$$X_t : \Omega \to \mathbb{R}, \quad \omega \longmapsto X_t(\omega) \equiv X(t, \omega)$$

を考える. X_t はこの確率過程が時点 t においてとる値をあらわす確率変数である。したがって $X_t \in L^0 \equiv L^0(\Omega)$.

さて t に X_t を対応させる写像

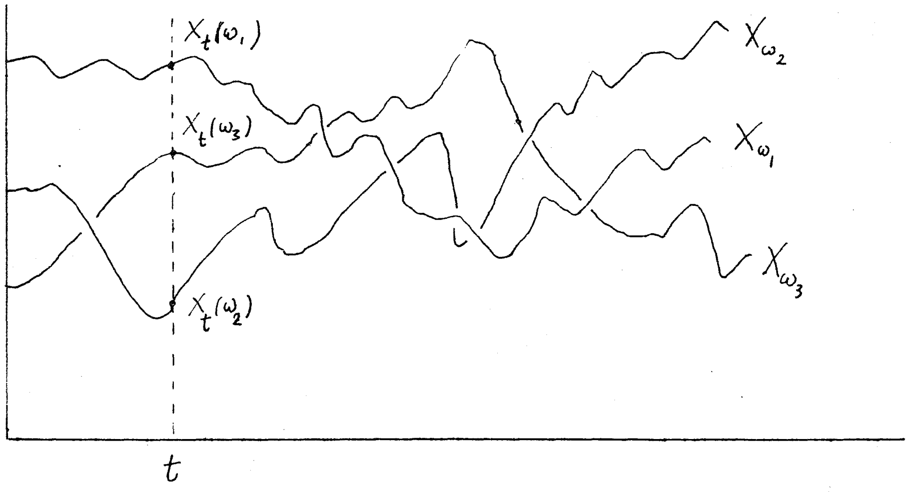
$$\hat{X} : T \longrightarrow L^0, \quad t \longmapsto \hat{X}(t) \equiv X_t$$

として, X を見ることができる. この意味で $\hat{X}$ は L^0 の中の曲線と考えてもよい. この場合には確率過程を時と共に変動する偶然量として眺めるのである。これが第二の見方である。

最後に個々の $\omega \in \Omega$ に対して写像

$$X_\omega : T \to \mathbb{R}, \quad t \longmapsto X_\omega(t) \equiv X(t, \omega)$$

を考えよう．これは見本（前節の例では微視状態）ω が実現されたときに観測される X の値の時間的変化であって，これを確率過程 X の（見本 ω に対する）見本関数（sample function）という．見本関数を見本過程（sample process）とか見本道（sample path）ということもある．下の図は X_ω と X_t との関係を示している．



個々の ω に対して $X_\omega : T \to \mathbb{R}$ は関数空間 $\mathbb{R}^T$ の上の元になっている．それで写像

$$\tilde{X} : \Omega \to \mathbb{R}^T, \quad \omega \longmapsto \tilde{X}(\omega) \equiv X_\omega$$

が考えられる．これが第三の見方である．この見方こそ確率過程の本質に迫ったものというべきものである．

今 Γ を $\mathbb{R}^T$ の部分集合とするとき，見本関数 X_ω が Γ の中に入る確率は

$$P\{\omega : X_\omega \in \Gamma\} = P(\tilde{X}^{-1}(\Gamma))$$

で与えられる．これを実験的に定めるには，何回もの実験を行ない，観測された見本関数の中で関数族 Γ の中に入るものの頻度を調べたらよい．上の式からわかるように，数学的には $\tilde{X}^{-1}(\Gamma)$ が P-可測である場合にのみ意味がある．まず

$$\Gamma = \{\xi \in \mathbb{R}^T \ : \ a_1 < \xi(t_1) < b_1,\ a_2 < \xi(t_2) < b_2,\ \cdots,\ a_n < \xi(t_n) < b_n\}$$

$$= \bigcap_{i=1}^{n} e_{t_i}^{-1}(a_i, b_i)$$

（e_t は $\xi \in \mathbb{R}^T$ を時点 t における値 $\xi(t) \in \mathbb{R}$ に移す写像）

のときには，$X_\omega \in \Gamma$ は

$$a_1 < X_\omega(t_1) < b_1,\ a_2 < X_\omega(t_2) < b_2,\ \cdots,\ a_n < X_\omega(t_n) < b_n$$

即ち，時点 $t_1, t_2, \cdots, t_n$ で，それぞれ見本関数 X_ω が区間 (a_1, b_1), $(a_2, b_2), \cdots, (a_n, b_n)$ の中を通過して行くことを意味する．

$$X_{t_i}(\omega) = X_\omega(t_i),\quad i=1,2,\cdots n$$

に注意すると，

$$P\{\omega : X_\omega \in \Gamma\} = P\{\omega : a_1 < X_{t_1}(\omega) < b_1, \cdots, a_n < X_{t_n}(\omega) < b_n\}$$

となり，$X_t : \Omega \to \mathbb{R}$ の P-可測性により，上の ω 集合も P-可測で $P\{\omega : X_\omega \in \Gamma\}$ は意味をもつ．

上記の形の集合は $\mathbb{R}^T$ の中の矩形集合とよばれる．$P\{\omega ; X_\omega \in \Gamma\}$ 即ち $P(\tilde{X}^{-1}(\Gamma))$ が意味をもつから，すべての矩形集合で生成される $\mathbb{R}^T$ 上の σ-加法族（これを Kolmogorov σ 加法族といい，$\mathcal{B}_K(\mathbb{R}^T)$ であらわす）の元に対して $P\{\omega : X_\omega \in \Gamma\}$ が意味

をもつ．これは $X:\Omega \to \mathbb{R}^T(\mathcal{B}_K(\mathbb{R}^T))$ が P-可測であることを示し，この意味で見本過程 $\tilde{X}$ は $\mathbb{R}^T$-値確率変数と考えることができる．その確率法則は

$$P_{\tilde{X}}(\Gamma) = P(\tilde{X}^{-1}(\Gamma)) = P\{\omega : X_\omega \in \Gamma\}, \quad \Gamma \in \mathcal{B}_K(\mathbb{R}^T)$$

で，これは無限次元空間 $\mathbb{R}^T$ の上の確率測度である．

では無限次元空間 $\mathbb{R}^T$ の上の確率測度は具体的にいかなるものか．このような問題は 1930 年以前の測度論では殆んど問題にならなかった．こういう問題設定そのものが新しい考え方で，これは確率過程論を厳密な数学理論とするために出されたものである．それ以前の確率過程論では，偶然と時点 $t_1, t_2, \cdots, t_n$ における値 $X_{t_1}, X_{t_2}, \cdots, X_{t_n}$ の結合分布 $P_{X_{t_1}\cdots X_{t_n}}$（これは有限次元 $\mathbb{R}^n$ の上の確率測度）を考え，この $\{t_i\}$ の数を増し，かつますます稠密にして行ったときの極限のようなものと考えていたのである．

μ を $\mathbb{R}^T$ の上の確率測度とする．前述したように

$$e_t : \mathbb{R}^T \longrightarrow \mathbb{R}, \quad \xi \longmapsto \xi(t)$$

を $\mathbb{R}^T$ の t 座標への射影とすると，その組 $(e_{t_1}, e_{t_2}, \cdots e_{t_n})$（これを $e_{t_1 t_2 \cdots t_n}$ とかく）は t_i 座標 $(i=1,2,\cdots,n)$ への組への射影である．

即ち

$$e_{t_1\cdots t_n}(\xi) = (e_{t_1}(\xi), \cdots, e_{t_n}(\xi)) = (\xi(t_1), \cdots, \xi(t_n)).$$

さて

$$\mu_{t_1\cdots t_n}(E) = \mu(e_{t_1\cdots t_n}^{-1}(E)), \quad E \in \mathbb{R}^n$$

とすると，これは所謂 μ の周辺分布（marginal distribution）である．確率過程 $X = X(t, \omega)$ の見本過程 $\tilde{X}$ の確率法則 $P_{\tilde{X}}$ が μ に等しいときには，

$$\mu_{t_1\cdots t_n} = P_{X_{t_1}\cdots X_{t_n}}$$

となる．このようにして，無限次空間上の確率測度の系

$$\{\mu_{t_1\cdots t_n} : n=1, 2, \cdots, t_1, t_2, \cdots t_n \in T\}$$

が得られる．これらの確率測度は全く独立なわけでなく，

$$\mu_{t_1\cdots t_n}((a_1, b_1)\times(a_2, b_2)\times\cdots\times(a_n, b_n))$$
$$= \mu_{t_{p(1)}\cdots t_{p(n)}}((a_{p(1)}, b_{p(1)})\times(a_{p(2)}, b_{p(2)})\times\cdots\times(a_{p(n)}, b_{p(n)}))$$
$$((p(1), \cdots, p(n)) \text{ は } (1, 2, \cdots, n) \text{ の順列})$$

$$\mu_{t_1\cdots t_n, t_{n+1}}((a_1, b_1)\times(a_2, b_2)\times\cdots\times(a_n, b_n)\times(-\infty, \infty))$$
$$= \mu_{t_1\cdots t_n}((a_1, b_1)\times(a_2, b_2)\times\cdots\times(a_n, b_n))$$

が当然成り立つ必要がある．これを両立条件（consistency condition）という．では逆にこの両立条件をみたす系 $\{\mu_{t_1 t_2\cdots t_n}\}$ が与えられたとき，$\mathbb{R}^T$ $(\mathcal{B}_K(\mathbb{R}^T))$ 上の確率測度 μ で上述の

$$\mu_{t_1\cdots t_n}(E) = \mu(e_{t_1\cdots t_n}^{-1}(E)), \quad E \in \mathcal{B}(\mathbb{R}^n) = \mathbb{R}^n \text{上のBorel集合族}$$

がなりたつようなものがあるか．またそれは唯一つか．この問題に Yes の解答を与えたのが Kolmogorov の拡張定理（Grundbegriffe der Wahrscheinlichkeitsrechnung, 1933) である.

具体的な確率過程 X では普通 $\mu_{t_1\cdots t_n} = P_{X_{t_1}\cdots X_{t_n}}$ が指定されるが，このような確率過程 X が数学的に存在するかは，この Kolmogorov の拡張定理によって保証されるのである．即ち $\{\mu_{t_1\cdots t_n}\}$ に対して，上の μ を定め，

$$\Omega = \mathbb{R}^T,\quad P = \mu,\qquad X(t,\omega) = e_t(\omega) \equiv \omega(t)$$

とすれば，この (Ω, P) 上の確率過程 X の見本過程 $\tilde{X}$ の確率法則は μ であり，$X_{t_1}\cdots X_{t_n}$ の結合分布 $P_{X_{t_1}\cdots X_{t_n}}$ は $\mu_{t_1\cdots t_n}$ に等しくなる．

たとえば X_t の確率法則 μ_t を任意に与え，$X_t, t\in T$ が独立となるような確率過程 X の存在は

$$\mu_{t_1\cdots t_n} = \mu_{t_1}\times\mu_{t_2}\times\cdots\times\mu_{t_n} \qquad (\times \text{ は測度の直積})$$

に上記の拡張定理を適用して保証される．また初期分布と推移確率を与えて，Markov 過程を構成する場合も同様である．Kolmogorov 以前には，これらのことは直観的に容認されていたのである．

確率過程を確率空間 (Ω, P) 上の点 ω と時変数 t との関数 $X(t,\omega)$ としてあらわすということは重要な認識であるが，これは形式上の問題である．上の Kolmogorov の拡張定理は，これに実質を与えたもので，これによって確率過程が真に数学の対象として研究できるようになったのである．

3. 見本関数の正則性(連続性、可測性)

$\{\mu_{t_1 t_2 \cdots t_n}\}$ を両立条件を満たす系とするとき，Kolmogorov の拡張定理によって，$\mathbb{R}^T(\mathcal{B}_K(\mathbb{R}^T))$ 上の確率測度 μ で

$$\mu(e_{t_1 t_2 \cdots t_n}^{-1}(E)) = \mu_{t_1 t_2 \cdots t_n}(E), \qquad E \in \mathcal{B}(\mathbb{R}^n)$$

を満たすものが唯一つある．したがって

$$P\{\omega : (X_{t_1}(\omega), \cdots, X_{t_n}(\omega)) \in E\} = \mu_{t_1 t_2 \cdots t_n}(E), \quad E \in \mathcal{B}(\mathbb{R}^n)$$

(いいかえれば

$$P\{\omega : X_\omega \in \Gamma\} = \mu(\Gamma), \qquad \Gamma \in \mathcal{B}_K(\mathbb{R}^T)\ )$$

となる確率過程

$$X(t, \omega), \quad t \in T, \quad \omega \in \Omega = (\Omega, P)$$

が存在する．例えば $\Omega = \mathbb{R}^T$, $P = \mu$, $X(t, \omega) = e_t(\omega)$ とすればよい．

T が離散集合(例えば $\{1, 2, \cdots\}$)のときには，この数学的表現は申し分ない．実際，例えば

$$\{\omega ; \sup_n X_n(\omega) \leqq a\} = \bigcap_n \{\omega : X_n(\omega) \leqq a\}$$

により，$Y(\omega) = \sup_n X_n(\omega)$ は ω の P-可測関数(確率変数)となり，その確率法則を考えることができる．同様に

$$\lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} X_n(\omega) = a$$

という事象の確率も考えられる．大数の強法則の意味がはっきりしたのも，Kolmogorov の表現の成果である．

しかし T が連続集合（例えば区間）のときには問題がある．この場合，たとえば見本過程が連続関数となるという事象：

$$X_\omega \in C \equiv C(T) \ (= T\text{上の連続関数族})$$

は極めて重要な事象である．それにもかゝわらず，

$$\{\omega : X_\omega \in C\}$$

は，上述の Kolmogorov の表現では P-可測でないから，その確率は考えられない．これは，Pを完備化しておいても，同様である．また

$$Y(\omega) = \int_0^1 X_t(\omega)\,dt \equiv \int_0^1 X_\omega(t)\,dt$$

も ω の P 可測関数にならない．したがって，これを確率的に論ずることができない．

例えば Wiener の Brown 運動の見本過程は連続関数であってほしいが，その周辺分布

$$\mu_{t_1 t_2 \cdots t_n}(E) = \iint_E\cdots\int N_{t_1}(dx_1)\,N_{t_2-t_1}(dx_2)\cdots N_{t_n-t_{n-1}}(dx_n) \quad (t_1<t_2<\cdots<t_n)$$

$$(N_t(dx) = \text{平均 } 0,\ \text{分散 } t \text{ の Gauss 分布})$$

から出発して，上記の Kolmogorov の表現をつくると，

$$\{\omega : X_\omega \in C(T)\}$$

は P-可測とならない．本当はこの集合の P 測度は 1 となっ

てほしいのである．

実際 N. Wiener は，この特別な場合には，Gauss 分布の特性を用いて，$\Omega = C(T)$ の上に確率測度 P（これは現在では Wiener 測度という）を巧みに構成して，この難点をきりぬけている（1920〜24）．しかしこのような見本過程の連続性，もっと一般に第一種不連続性，さらに一般に可測性について考察したのは J. L. Doob である（1937）．これが見本過程の正則性の問題である．

Doob は Kolmogorov の表現（$\Omega = \mathbb{R}^T$）では，連続関数の空間 $C(T)$，第一種不連続関数の空間 $D(T)$，可測関数の空間 $L^0(T)$ は決して P-可測とならないことを注意し，これに対処するために，始めに

$$\overline{P}(\Omega') = 1 \qquad (\overline{P} \text{ は } P \text{ に対する外測度})$$

となる Ω'（$= C(T), D(T)$ または $L^0(T)$）があれば，P を $\Omega' \cap \mathcal{B}_K(\mathbb{R}^T)$ 上に制限した P' を Ω' にそえて得られる確率空間 (Ω', P') の上で考えればよいことを示した．その後 (Ω, P) はそのまゝにしておいて，$X_t(\omega)$ を個々の t に対して P-測度 0 の集合（t に依存）の上で修正する方法（modification）をとった．両者は本質的には同じであるが，後者の方が便利であるので，現在では修正法のみが用いられている．修正の一般原則として separable modification（可分変形）という概念を導入

した．可分変形はすべての確率過程に対して定義されるが，特に連続（または第一種不連続）な変形をもつ確率過程に対しては，その変形は可分変形で得られる．では如何なる条件の下で連続な変形，第一種不連続な変形をもつか？

(i)（Kolmogorov）　$\exists\, \alpha, \beta, \gamma > 0 \quad \forall s, t \quad E(|X_t - X_s|^\alpha) \leqq \beta |t-s|^{\gamma+1}$ ならば，連続変形をもつ．

(ii)（Doob）　確率連続（$\lim_{t \to s} E(|X_t - X_s| \wedge 1) = 0$）な加法過程，martingale 過程，（極めてゆるい付加条件の下で）Markov 過程は第一種不連続変形をもつ．

4.　加法過程

$X_1, X_2, \cdots$ が独立な確率変数の列であるとき，その逐次和

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

の行動を調べることは，Bernolli, de Moivre, Laplace, Gauss の時代から，大数の法則，中心極限定理の形で研究されてきた．とくに Laplace が有名な「確率の解析的理論（1812）」で導入した生成関数（Laplace 変換）はこの研究を極めて容易にし，20 世紀になって積分論（Lebesgue 積分，Stieltjes 積分）の進歩に応じて，生成関数よりも有力な特性関数（Fourier 変換）が P. Lévy によって導入され，上記の S_n の行動の研究は急速に進歩した．また見本過程という考えも生れ，

S_n の見本関数（見本列というべきか？）の行動にも関心が向けられ始めた．

$\{S_n\}$ を確率過程と見なすと，時変数 n は離散的で，0, 1, 2, … の値をとる．このれを連続時変数 t でおきかえたものが加法過程である．上記の S_n は

$$X_n = S_n - S_{n-1}, \quad n = 1, 2, \cdots$$

が独立であるから，加法過程 S_t, $t \in T$ $= [0, \infty)$ では，象徴的には

$$dS_t = S_{t+dt} - S_t, \quad t \in T$$

が独立というべきであろう．これは S_t が瞬間毎に独立な増分を得て変動して行く確率過程といってもよい．実際 P. Lévy は初期の論文や著書（Théorie de l'addition des variables aléatoires, 1937）では，「独立な random elements の積分」と呼んでいる．このような象徴的定義を厳密な数学的定義にするには，次のようにいえばよい．

任意の $t_0 < t_1 < \cdots < t_n$ に対し $S_{t_i} - S_{t_{i-1}}$, $i = 1, 2, \cdots, n$ が独立であるとき，S_t を加法過程という．（但し $S_0 = 0$ とする．）

加法過程の特別なものとして，Wiener 過程 (Wiener の Brown 運動)，Poisson 過程，複合 Poisson 過程が Lévy が加法過程の一般論を構築する前から，物理学，災害統計学，生物学の問題に関連して研究されていた．後の二者は見本過程が階段関数になるから，研究も容易であったが，Wiener 過程は見本過程が連続関

数となるから，その研究は容易ではなかった．実際 Wiener がこの理論の数学的基礎を考えたのは，有名な論文'differential space'(1920)に始まるが，その後いくつかの論文をかき，1924年に，Daniel 積分（抽象空間上の積分の始め）を応用して，始めて厳密な証明に到達している．その間に見本過程の微分不可能性などを考察して，Wiener 過程のその後の研究の方向を示唆しているが，その背景となっているのは，数理経済学者 L. Bachlier の投機の理論(1900)，物理学者 Einstein の研究(1906)，コロイド学者 Perrin の研究(1908)などいずれも数学以外の領域である．

前節にのべた Kolmogorov，Doob の一般確率過程の理論によれば，Wiener 過程の基礎づけは容易である．まず13頁に与えた周辺分布の系 $\{\mu_{t_1 t_2 \cdots t_n}\}$ から，Kolmogorov の表現をつくり，

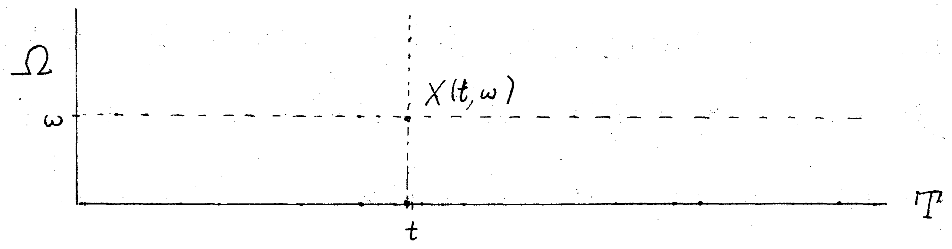
$$E(|X_t - X_s|^4) = 3|t-s|^2$$

であることに着目して，Kolmogorov の連続変形存在定理(15頁)を用いたら，直ちに望むものが得られる．

加法過程の初期の研究（Finetti (1929)，Kolmogorov (1932)）では，加法過程 X の t 時点における値 X_t の確率法則にのみ着目した．既に知られた Wiener 過程，Poisson 過程，複合 Poisson 過程の場合にはそれぞれ Gauss 分布，Poisson

分布，複合Poisson分布となっているが，一般の加法過程では，確率連続の仮定の下で，無限分解可能な確率法則という上記の分布を含む極めて広い分布のクラスが得られる．この無限分解可能な確率法則の標準形を決定するには，Lévyの特性関数を用いるが，それはFourier変換論で，解析学に入り，確率論的な香りは全くない．

P. Lévyは直観的に「確率連続な加法過程の見本関数は第一種不連続であること（これは後にDoobが厳密に証明した，前節末尾(ii)参照）」を見抜き，その立場から，見本関数を連続な部分と飛躍の高さに応ずる階段関数の部分に分けて，その結果として各時点における値X_tの分布を出し，その標準形の真の意味を明らかにしている．このLévyの研究(1937)により，加法過程の確率論的研究が飛躍的に進展し始めたのである．確率過程は(t,ω) 2変数の関数で，模型的には



のようにあらわされる．Lévy以前には個々のtに対して$X_t(\omega)=X(t,\omega)$を眺めていた．Lévyが個々のωに対して，見本関数$X_\omega(t)=X(t,\omega)$を観察して，始めて真の確率過程論

が生れたのである．「縦のものを横に見る」のが独創であるといわれ，Riemann積分からLebesgue積分への進歩がその例であるという．上記のLévyの観点（見本関数）もまたその好例ではなかろうか．

5. Markov過程

確率過程の一般論を厳密に構成しておくことは，数学的な確率過程論を建設するための土台である．土台がしっかりしなくては楼閣はできないが，土台だけにかゝずらわっていることはできない．そこでこの楼閣の大きい柱として，重要な確率過程のクラスを考える必要がある．これを把握するためには，確率過程論の数学的（論理的）基礎である無限次元空間の上の測度をながめていても，何も出てこないので，確率過程論の実体である「確率的現象の時間的変化」を考える必要がある．

前節にのべた加法過程も，その萌芽は，経済現象，物理現象，生物現象である．加法過程が確率過程論の大きな柱であることはいうまでもない．では他に何があるか．第二の柱として考えられるのがMarkov過程である．これは論理的には加法過程を含む大きいクラスである．

加法過程では増分の独立性を考えるから，そのとる値 $X(t,\omega)$ は実数であった．もっと一般に $\mathbb{R}^n$ の点，ベクトル空間の点

としても，定義できるが，全く一般の空間（例えば球面）の上を動く加法過程を考えることはできない．ところがMarkov過程はこういうこととは全く無関係な概念である．

Markov過程は各時点で過去がきまった条件のもとにおける将来の行動の条件付確率法則がその時点における値だけに関係して定まるような確率過程である．この条件付確率法則は一般には過去の行動の履歴全体に関係するのであるが，それが現時点だけできまるのがMarkov性である．たとえば大学入学試験の成績だけで合否をきめるのはMarkov的な態度であり，高校の内申書も考慮に入れるのは非Markov的な態度である．

加法過程はMarkov性をもっていることは定義からすぐ分るが，一般の確率過程はMarkov性をもたないのが普通で，Markov性というのは極めて強い条件である．ある確率現象をMarkov過程として捉えるのは多くの場合近似的にそう考えるのである．明らかにMarkov性を持たないものを無理にMarkov過程として捉え，その推移確率を論じている場合があるので注意を要する．例えば冒頭にのべた水の中の粒子の運動も厳密な意味ではMarkov過程ではない．しかし，この場合には近似的にはMarkov過程と見なしてよいことは直観的に想像される．（厳密には証明できないが．）

Markov性を少し具体的な例で説明しよう．箱の中に$2n$枚の札があり，これに$1, 2, \dots, 2n$の番号がついているとする．これから札を出鱈目に1枚ずつぬきとり，第k回目に出た札の番号をX_kとすると，$X_1, X_2, \dots, X_{2n}$は離散時変数の確率過程となる．明らかに

$$P\{X_1 = i\} = \frac{1}{n}$$

$$P\{X_2 = j \mid X_0 = i\} = \begin{cases} 0 & j = i \text{ のとき} \\ \dfrac{1}{2n-1} & j \neq i \text{ のとき} \end{cases}$$

$$P\{X_3 = k \mid X_0 = i, X_1 = j\} = \begin{cases} 0 & k = i, j \text{ のとき} \\ \dfrac{1}{2n-2} & k \neq i, j \text{ のとき} \end{cases}$$

$$P\{X_4 = l \mid X_0 = i, X_2 = j, X_3 = k\} = \begin{cases} 0 & l = i, j, k \text{ のとき} \\ \dfrac{1}{2n-3} & l \neq i, j, k \text{ のとき} \end{cases}$$

$$\cdots\cdots$$

さて$P\{X_2 = k \mid X_0 = i, X_1 = j\}$は$i, j$両方に関係し，また$P\{X_3 = l \mid X_0 = i, X_1 = j, X_2 = j\}$は$i, j, k$全部に関係する，以下同様である．したがってこの確率過程はMarkov性をもたない．

つぎに同じ試行で，第k回目までにでた奇数札の総枚数をY_kとすると，

$$P\{Y_1 = 0\} = P\{Y_1 = 1\} = \frac{1}{2}，$$

$P\{Y_{k+1}=j \mid Y_1=i_1, Y_2, \cdots, Y_k=i_k\}$ を考えよう．$i_1, i_2, \cdots, i_{k-1}$ が何であっても，上の条件の下では箱の中の総枚数は $2n-k$ で，その組成は奇数札数 $n-i_k$，偶数札数 $n-k+i_k$ である．したがって，上の確率は個々の k, j に対し，$i_1, i_2, \cdots, i_{k-1}$ には関係せず，i_k にのみ関係する．実際

$$P\{Y_{k+1}=j \mid Y_1=i_1, Y_2, \cdots, Y_k=i_k\}$$

$$=\begin{cases} \dfrac{n-k+i_k}{2n-k} & j=i_k \text{ のとき} \\ \dfrac{n-i_k}{2n-k} & j=i_k+1 \text{ のとき} \\ 0 & j \neq i_k, i_k+1 \text{ のとき} \end{cases}$$

$$= \frac{n-k+i_k}{2n-k}\delta_{i_k, j} + \frac{n-i_k}{2n-k}\delta_{i_k+1, j} \qquad (\delta_{ij} = \text{Kronecker の } \delta)$$

となり，個々の k, j に対し，i_k だけで定まる．

さて一般に

$$P\{Y_1=i_1, Y_2=i_2, \cdots, Y_n=i_n\}$$

$$=P\{Y_1=i_1\}P\{Y_2=i_2 \mid Y_1=i_1\}P\{Y_3=i_3 \mid Y_1=i_1, Y_2=i_2\}$$

$$\cdots\cdots P\{Y_n=i_n \mid Y_1=i_1, Y_2=i_2, \cdots, Y_{n-1}=i_{n-1}\}$$

であるが，特に Markov 過程のときには，この第 ν 因子（条件付確率）は最後の $i_{\nu-1}$ にだけ関係する $(\nu=2, 3, \cdots, 4)$ から これをそれぞれ

$p^{(\nu)}_{i_{\nu-1},\, i_\nu}$ とかけば

$$P\{Y_1=i_1,\ Y_2=i_2,\ \cdots,\ Y_n=i_n\}$$
$$= p_{i_1}\, p^{(1)}_{i_1,\, i_2}\, p^{(2)}_{i_2,\, i_3} \cdots p^{(n)}_{i_{n-1},\, i_n}$$

となる．こゝで $p_{i_1}=P\{Y_1=i_1\}$ は初期分布とよばれ，$p^{(\nu)}_{i,j}$ は第ν回から第(ν+1)回への推移確率とよばれる．したがって Markov過程の場合には初期分布と推移確率によって，すべての n に対し，n回目までの結合分布がきまる．あとは Kolmogorov の拡張定理を用いて $\{Y_n\}$ の見本過程（この場合は見本列）の確率法則が考えられる．上の推移確率は ν回から ν+1回へのもので，1回次のものであるが，もっと一般に ν回から ν′回へのものは，

$$p^{(\nu,\nu')}_{i,j} = \sum_{i_1}\sum_{i_2}\cdots\sum_{i_\theta} p^{(\nu)}_{i,i_1}\, p^{(\nu+1)}_{i_1 i_2} \cdots p^{(\nu'-1)}_{i_\theta,\, j} \qquad \theta=\nu'-1-\nu$$

で得られる．　こうすると

$$p^{(\nu,\nu'')}_{i,k} = \sum_{j} p^{(\nu,\nu')}_{i,j}\, p^{(\nu',\nu'')}_{j,k} \qquad (\nu<\nu'<\nu'')$$

となる．これが Chapman-Kolmogorov 方程式である．行列 $(p^{(\nu,\nu')}_{i,k})_{i,k}$ を $P^{(\nu,\nu')}$ であらわすと，上の式は行列の掛算の形で

$$P^{(\nu,\nu'')} = P^{(\nu,\nu')}\, P^{(\nu',\nu'')}.$$

とかける。

以上のことを頭において連続時変数 $t \in T$ (=区間)，一般状態空間 S の場合を考えよう。(普通 Markov 過程といえば，この場合をさし，上述のような離散時変数のものは Markov 連鎖という。) $X(t,\omega)$, $t \in T$, $\omega \in (\Omega, P)$ を Markov 過程とし，その推移確率を $p_{s,t}(x, B)$ であらわす。これは時点 s において x にあったとき $t\,(>s)$ において B に入る条件付確率である。これは s までの時点でどのようであったかには関係しない。Chapman-Kolmogorov の方程式は

$$p_{s,u}(x, B) = \int_S p_{s,t}(x, dy)\, p_{t,u}(y, B)$$

となる。今積分作用素をこの $p_{s,t}(x, B)$ から

$$p_{s,t}(f)(x) = \int_S p_{s,t}(x, dy) f(y)$$

によって導入すると，上の方程式は，作用素の積の形で

$$p_{su} = p_{st} \cdot p_{tu} \qquad s < t < u$$

とかける。$p_{s,t}$ は $s<t$ で定義したが，$s=t$ のときには $p_{s,s}(x, B) = \delta_B(x)$ となるはずであるから，$p_{s,t} = I$ (恒等作用素となる。

さて上の離散時変数のときには $\{p_{ij}^{(\mu,\nu)}\}$ は $\{p_{i,j}^{(\nu)} \equiv p_{i,j}^{(\nu,\nu+1)}\}$ から得られた。連続時変数のときにも，これに相当するもの

を考えよう. $p_{su} = p_{st}\,p_{tu}$ $(s<t<u)$, δ_0がいくら小さくても, $p_{s,s+\delta}$ $(0<\delta<\delta_0)$がすべて与えられたら, すべての$p_{s,t}$ $(s<t)$が定まる. しかし $\delta=0$とおいてしまうと, $p_{s,s}=I$となってしまって, 何等の情報も得られない. それで, 運動の各瞬間における速度を考えるように,

$$\mathcal{G}_s = \lim_{\delta\downarrow 0}\frac{1}{\delta}(p_{s,s+\delta}-I)$$

を考え, これを時点sにおける生成作用素という. $p_{st}f$はS上のすべての有界Borel函数fに対して定義されるが, $\mathcal{G}_s f$はそうではない. しかし$\mathcal{G}_s f$の定義されるfの全体$\mathcal{D}(\mathcal{G}_s)$ ($\mathcal{G}_s$の定義域)は十分広く, $\mathcal{G}_s$から$p_{s,t}$が定まるのである. 実際 $p_{su}=p_{st}\,p_{tu}$から得られる所の

$$\frac{\partial p_{su}}{\partial s} = -\mathcal{G}_s\,p_{su} \qquad s<u$$

を $p_{uu}=I$を初期条件(末期条件というべきか)で解いてp_{su}が得られる. 上記の方程式はKolmogorovの後向き方程式とよばれ, この双対形をKolmogorovの前向き方程式という. 上の方程式は作用素$p_{s,u}$に関するものであるが, これを函数に対する方程式とするには,

$$f(s,u;x) = (p_{su}f)(x)$$

とおいて

$$\frac{\partial f(s,u:x)}{\partial s} = -\mathcal{G}_{s,x} f(s,u:x), \qquad f(u,u:x) = f(x)$$
$$(s < u)$$

とすればよい．$\mathcal{G}_{s,x}$ に x を添えたのは，$f(s,u:x)$ を x の関数と見て（s, u は固定して）$\mathcal{G}_s$ を作用することを示すためである．

$S = R^n$ のときには，$\mathcal{G}_s$ は大体

$$\mathcal{G}_s f(x) = \sum_{i=1}^{n} a_i(s,x)\frac{\partial f}{\partial x_i}(x) + \frac{1}{2}\sum_{i,j=1}^{n} b_{ij}(s,x)\frac{\partial^2 f}{\partial x_i \partial x_j}(x)$$
$$+ \int_{\|\xi\| \neq 0} \left(f(x+\xi) - f(x) - \frac{\sum_i \xi_i \frac{\partial f}{\partial x_i}(x)}{1 + \|\xi\|^2} \right) n_s(d\xi)$$

の形になり，見本過程が連続のときは，最後の積分項は消えて，$\mathcal{G}_s$ は楕円形(偏)微分作用素となる．このようにして見本過程が連続な Markov 過程（これを現在では拡散過程という）を楕円形微分方程式（または放物型微分方程式）と結びつけたのは，Kolmogorov の論文

Über die analytische Methoden in der Wahrscheinlichkeitsrechnung (Math. Ann. 104, 415–458 (1931))

である．Kolmogorov は見本過程が連続な場合（即ち拡散過程）をとり扱ったが，これは，後に(1936)，Feller によって一般化され，精密化された．

Wiener 過程（Wiener の Brown 運動）は拡散過程で，その推移確率 $p_{s,t}(x, E)$ は E の関数と見て平均 x，分散 $t-s$ の

Gauss 分布である．したがって

$$(p_{st}f)(x) = \int_{\mathbb{R}} f(y)\, N_{x,\,t-s}(dy) = \int_{\mathbb{R}} f(y+x)\, N_{0,t-s}(dy)$$

$$(\mathcal{G}_s f)(x) = \frac{1}{2}\frac{d^2}{dx^2}f(x)$$

となり，Kolmogorov の方程式は所謂熱伝導の方程式（伝導率が常数 1/2 のとき）となる．n 次元の Wiener 過程のときには，$\frac{d^2}{dx^2}$ がラプラシアン Δ になる．こゝで $\mathcal{G}_s$ が s に無関係となることに注意すべきである．これは $p_{s,t}(x, E) = N_{x,\,t-s}(E)$ が s, t の関数と見て，$t-s$ にのみ依存することに由来する．このような Markov 過程は時間的に一様（temporally homogeneous）といい，$p_{s,t}$ を p_{t-s} であらわす．時には定常な推移確率をもつ Markov 過程ともいう．後にのべるように定常過程という別のクラスがあり，定常かつ Markov 過程というものも考えられる．これは必ず定常な推移確率をもつが，逆に定常な推移確率をもつ Markov 過程が必ずしも定常 Markov 過程ではない．しかし初期分布 μ が不変性：

$$\mu(E) = \int_{\mathbb{R}} \mu(dx)\, p_t(x, E)$$

をもつときには，$\mathbb{T}$ を $(-\infty, \infty)$ にまで拡張して，定常 Markov 過程にすることができる．

このKolmogorovの論文は数学者とくに解析学者をMarkov過程論にひきつけた最初のものであるが，Kolmogorovは「数学」の中でこの研究の端緒を掴んだのではない．Wiener過程と熱伝導の方程式との関係は物理学者Einsteinによってすでに今世紀初頭に注意されていた．またBachlierも投機の理論に関係して，Wiener過程に注意していた．また係数が2次式の場合は統計学者K. Pearsonが，統計にあらわれる「Gauss分布以外の分布の型」で重要なものを掴むために研究していた．係数が一般の場合は，物理学者Fokker (1914) やPlanck (1917) が分子運動と熱伝導に関係して考察していた．しかしKolmogorov以前には，数学者がこの微分方程式を微分方程式の立場から考察することはあっても，その確率論的意味を考えて興味をもつには至らなかった．Kolmogorovの研究がでて始めて数学者の興味をひく形をとったのである．

Kolmogorov以後のMarkov過程論の発展は

(i) 解析的研究（偏微分方程式論の立場から）

(ii) 関数解析的研究（吉田-Hilleの半群の立場から）

(iii) 確率論的研究（見本過程の立場から）

の3通りがある．これらの研究は別々に出発しているが，見本過程の理論(iii)が進むにつれて，すべてこの見地から統一的に論じられるようになった．(iii)は実は無限次元の

解析学の理論であり，(i), (ii) で得られる結果はすべてその有限次元への射影として理解できるようになった．

生成作用素 $\mathcal{G}_s$ でも確率論的に最も深い考え方は Stroock-Varadhan (1969) による

$$f(X_t) - f(X_0) = \int_0^t \mathcal{G}_\tau f(X_\tau)\,d\tau + \text{“マルチンゲール”}$$

となる $\mathcal{G}_\tau$ として定義することであるが，この見地は見本過程を通じて始めて理解できる．(Stroock-Varadhan: Multidimensional Diffusion Processes, Springer, Grundlehren der math. Wiss. 1979).

n 次元 Brown 運動が $x \in D$ (D は R^n の領域) から出発して ∂D (D の境界) から始めて出るとき，$E (\subset \partial D)$ からでて行く確率 $u(x; E)$ が調和関数で有名な「調和測度」(D で調和，E で 1，$\partial D - E$ で 0 となる関数) である．これは数学的には Lévy や角谷静夫によって発見されたものであるが，物理学者も類似の事実を直観的に知っていた．しかしこの証明を厳密にしようとすると，Wiener 過程の「強 Markov 性」に着目する必要があり，その定義は見本過程の立場において始めて理解できるのである．

Markov 過程論（時間的一様な場合）を半群の立場 (ii) から見た研究は吉田-Hille によって始められたが，それを更に進めて特に一次元拡散過程を完全に決定した W. Feller の 1950 年代

の仕事も，直観的に見本過程を考えることによって始めてなし得たものであり，強 Markov 性を利用している．また有名な Feynman-Kac の公式も見本過程の考えなしには理解できない．これらは現在では見本過程の理論から完全に解明されている（例えば K. Ito - H.P. McKean : Diffusion Processes and Their Sample Paths, Springer 1965）

また Kolmogorov が上述の方程式を導いたときにも，直観的に見本過程の瞬間的変動を捉えて，その平均値，分散という立場から論じているが，これを厳密に表現するには，確率微分方程式という見本過程論的立場が必要である．

1950年後半以降の Markov 過程の数学的研究はすべて見本過程論的立場に立っている．

応用上の見地から導入された出生死亡過程，待ち行列（Poisson arrival のとき）の問題，分枝過程なども Markov 過程となっているが，これを現在の見本過程論的 Markov 過程論の立場から眺めると，今まで繁雑な計算を行って出していた結果も，一望の下に得られることが多い．しかし簡単な問題では牛刀を以て鶏肉を裂くの感もあるせいか，応用確率過程論の本では，古典的直観的方法によっている場合が多い．これは確率過程論に限らず，すべての数学の分野でいえることである．

6. 定常過程とエルゴード性

定常過程は時間的に安定した確率現象をあらわすための確率過程である。時間的安定性は物理現象，生物現象，経済現象などでも極めて重要な概念である。普通の関数 $x(t)$ が時間的に安定しているといえば，$x(t) \equiv const.$ でつまらないものになってしまうが，確率過程では極めて興味ある問題を提供する。時間的安定というと，すぐに考えられるのは，確率過程 $X = X(t, \omega)$，$-\infty < t < \infty$，の時点 t における値 X_t の分布が t に無関係ということであるが，これだけでは相互関係の安定性は無視されている。それで定常過程をつぎのように定義する。

任意の時点 $t_1, t_2, \cdots, t_n$ に対し，$X_{t_1}, X_{t_2}, \cdots, X_{t_n}$ の結合分布が時間のずれに対して不変である，即ち

$$P_{X_{t_1} X_{t_2} \cdots X_{t_n}} = P_{X_{t_1+h}, X_{t_2+h}, \cdots, X_{t_n+h}}$$

がつねに成り立つとき，$X(t, \omega)$ は定常過程であるという。

さて X の見本過程の確率法則を μ とすると，μ は $R^{(-\infty,\infty)}$ 上の確率法則である。さて $R^{(-\infty,\infty)}$ 上のずれの変換を T_t とする，即ち

$$(T_t f)(s) = f(s+t), \qquad f \in R^{(-\infty,\infty)}$$

とすると，$\{T_t\}$ は R^T 上の変換の one parameter group とな

る（$T_t T_s = T_{t+s}$）．上述の定常性の条件は μ が T_t で不変即ち

$$\mu(T_t(\Gamma)) = \mu(\Gamma) \qquad \Gamma \in \mathcal{B}_K(\mathbb{R}^{(-\infty,\infty)})$$

を意味する．見本過程の regularity を考慮すると，数学的に微妙な問題が伴うが，こゝでは立入らない．

このようにして定常過程の理論は実は保測変換群の理論に帰着される．保測変換の理論は力学系の問題に起因するもので，力学における phase space 上の運動による変換について Liouville 測度が不変ということが，その出発点となっている．保測変換群の理論で特に重要なのがエルゴード性である．これはその変換群で不変な集合は測度 0 を除いて空集合か全空間ということである．定常過程の場合には上述の時間のずれによる保測変換群がエルゴード性をもつときに，その定常過程はエルゴード性をもつときに，その定常過程はエルゴード性をもつという．たとえば，確率空間 (Ω, P) 上の保測変換群 $\{T_t\}$ があれば，

$$X(t,\omega) = f(T_t\omega), \qquad f \text{ は任意の } P\text{-可測関数}$$

は定常過程となる．しかも $\{T_t\}$ がエルゴード性をもつならば，$X(t,\omega)$ はエルゴード性をもつ定常過程となる．力学系にあらわれる定常過程はおおむねこのような形であらわれる．

エルゴード性をもつ定常過程では，唯一つの見本関数(確率0の例外を除く)を見れば，定常過程の全貌がわかる．即ち一つの $\omega \in \Omega$ に対し，見本関数 X_ω をとり，

$$T_t X_\omega \ , \ -\infty < t < \infty$$

を考える．これは $\mathbb{R}^{(-\infty,\infty)}$ の中の集合になるが，その頻度は見本関数の確率法則となる．もう少し詳しくいうと，任意の $\Gamma \in \mathcal{B}_K(\mathbb{R}^{(-\infty,\infty)})$ に対し

$$\lim_{a\to\infty} \frac{1}{2a} \left| \{ t \in (-a,a) \ ; \ T_t X_\omega \in \Gamma \} \right| \qquad \left(\begin{array}{l} |\ \ | \text{は Lebesgue} \\ \text{測度} \end{array} \right)$$

が見本過程が Γ の中に入る確率となる．数学的にはこのいい方は不完全であるが，大体の意味はこれでよい．

一つの実験で我々が観測するのは，偶然実現された唯一つの見本関数にすぎないが，エルゴード性を仮定すれば，その見本関数だけを見れば，この確率過程の見本過程を支配する法則が捉えられるので，実験をくりかえす必要はないのである．

エルゴード性より強い条件として，混合性，Kolmogorov 性などがあるが，後者は特に重要である．これについては後節の増大情報系の所でのべる．

7. マルチンゲール

マルチンゲールはもともと各時点において公平な賭けである．たとえば，さいをふって奇数がでたら1円貰い，偶数がでたら1円払うという賭けでは時点 x円もっている人は次の時点 $n+1$ では，自分の持ち金が

$$\frac{1}{2}(x+1)+\frac{1}{2}(x-1)=x(円)$$

となることが期待できる．相手も同様である．だからこの賭けはつづけることができる．もし偶数回目には奇数がでたら1円貰うが，偶数がでたら2円払い，奇数回目にはその逆とすれば，上の期待値は

奇数回の後 x円もっているときには　次の回では　$\frac{1}{2}(x+1)+\frac{1}{2}(x-2)=x-\frac{1}{2}$

偶　〃　〃　〃　$\frac{1}{2}(x+2)+\frac{1}{2}(x-1)=x+\frac{1}{2}$

となり，自分は偶数回目の賭けには応じたくないであろうし，相手は奇数回目の賭けには応じないであろう．これはどの時点でも公平とはいえず，勿論マルチンゲールではない．

余談ではあるが，マルチンゲールは馬具の一種で，馬の首にかけるものらしい．私はパリのレストランの柱に飾りとしてかかっているのを見たことがあり，Doob からそれがマルチンゲールであることを教えてもらったが，それを馬の首にどのようにつけるのかは聞き漏した．実は1964年の「Potential

Theory と Markov Processes に関する国際シンポジュウム (Brelot 主催) で, Brelot が外国人参加者数名をパリのあるレストランに招いてくれた. そのとき Doob が招待された客の代表として謝辞をのべた. その中で柱にかいているマルチンゲールを指して,「このレストランにもマルチンゲールがある. Markov 過程論はマルチンゲール論に含まれ, 今やポテンシャル論もマルチンゲール論となっている. やがて解析や数学もマルチンゲール論となるであろう」という意味のことを述べた. 私も, おそらく同席者全部が, これを Doob 一流の冗談であると思った. しかし今になって見ると, これはまんざら冗談ではなさそうである.

閑話休題(それはさておき), 上述の賭の問題を抽象化して, マルチンゲールを次のような確率過程として導入する. 確率過程 $\{X_t\}$ がマルチンゲールであるとは, 各時点 s でそれまでの X_u, $u \leq s$ が定まったときに, X_t $(t > s)$ の条件付平均値がいつも X_s に等しいことである.

$\{X_t\}$ が加法過程のときには, $E(X_t)$ が一定であれば, マルチンゲールとなる. たとえば1次元 Wiener 過程はマルチンゲールである. もっと一般に $\{X_t\}$ が Markov 過程の場合には,

$$\int_R y\, p_{s,t}(x, dy) = x \qquad (p_{s,t} \text{ は推移確率})$$

ならば，マルチンゲールになる．しかしMarkov過程にならないマルチンゲールはいくらでもある．たとえば$\{B_t\}$をn次元Wiener過程とし，$h: R^n \to R$を調和関数は$h(B_t)$はMarkov過程ではないが，マルチンゲールである．

マルチンゲールは定常過程とは違った意味で安定な確率過程であって，多くのいい性質をもっている．マルチンゲールについては，P. LévyもWiener過程の研究にこれを用い，$\{X_t\}$と$\{X_t^2 - t\}$がマルチンゲールとなる見本連続な確率過程$\{X_t\}$はWiener過程であることを証明している．またマルチンゲールに対する本としてはJ. Ville (1939)のものもあった．しかしマルチンゲールの一般論をつくり，その応用の広さを示したのはJ. L. Doobの研究(1940年代)である．実際マルチンゲールの収束定理，optional stopping や optional sampling (random time change)によるマルチンゲール性の不変など，この理論の基本的事項は殆んどすべてDoobに負う．(Stochastic Processes, J. Wiley 1952)．しかもDoobは確率過程論の中にあらわれる諸種の事項が適当な変換によってマルチンゲールの応用の範囲に入ることを示し，この考え方はその後ますます重要となっている．

さらに最近解析学の諸問題で特殊な技巧，計算によるものも，マルチンゲールの考えで眺めると，見通しよくなる例

がでてきて居り，この間私はあるフランス人から，整数論に応用できるということを聞いた．こうして見ると，Doobのパリのレストランにおける話はあながち冗談ともいえなくなってくる．勿論，19世紀末に数学は不変式論（Invariantentheorie）であるといわれたのに比較すべき程マルチンゲールの理論が大きいかどうかは別としても，公平な賭などという特殊のものとはまるでかけ離れた大きい考え方であるには違いない．

8. 線形理論とGauss過程

確率過程論の形成過程において線形理論があらわれ，いくつかの成果が得られている．この理論は，現在の確率過程から見れば，極めて浅薄なものであるが，それだけにやさしく確率過程論の行くべき道をとらえるのに役立ったといえないことはない．線形理論が容易に受け入れられたのは，それがHilbert空間論の枠内に納めることができ，深い確率論的感覚はなくても，線形関数解析（最近まで関数解析の一般論は線形理論であった）の知識さえあれば，研究することができるからである．

確率過程$\{X_t\}$をながめるのに，その諸時点$\{t_i\}$における一次結合 $\sum_i a_i X_{t_i}$（a_iは定数）だけを考えるとする．平均値$E(X_t)$を引き去っておくことにより，$E(X_t)=0$と仮

定してよい．したがって $E(\sum_i a_i X_{t_i})=0$ となる．$\{X_t\}$ の共分散を $V_{t,s}$ とする．$\left(V_{t,s}=E[(X_t-E(X_t))((X_s-E(X_s))]=E(X_t X_s)\right)$
この結果 $Y=\sum a_i X_{t_i}$，$Z=\sum b_j X_{s_j}$ の共分散は

$$V(Y,Z)=\sum a_i b_j V_{t_i s_j}$$

となるが，$V(Y,Z)$ は Y, Z の関数と見て，内積の性質をもっている．したがって上記一次結合とその極限の全体 H は実 Hilbert 空間をつくる．このように考えると，$\{X_t\}$ は H の中の曲線（または運動）と見なすことができ，その一次結合も H の中で論じ得る．特に

$$(X_t, X_s)\equiv V_{X_t X_s} \text{ が } t-s \text{ にのみ関係する}$$

ときには，$\{X_t\}$ を定常（前述の定常過程と区別するため，弱定常ということもある）という．これが Khinchin の弱定常過程で，Khinchin は $\rho(t)=(X_{s+t}, X_s)$ が Bochner の正定符号関数となることに注意し，それが有限非負測度（スペクトル測度という）の Fourier 変換となることを示し，弱定常過程の研究の基礎を築いた．(1934)

$\{B_t\}$ が Wiener 過程のときには，

$$(B_t-B_s,\ B_v-B_u)=\text{区間 }'(s,t)\cap(u,v)'\text{ の長さ}$$

となり，この量は s, t, u, v を同じだけずらしても変らない．これは $\{B_t\}$ が H の中の螺線（Wiener spiral という）となっていることを示している．このことに着目して，ヒルベクト空間 H の中の

中の一般の螺線の型を決定したのは Kolmogorov である(1940).

また上のいずれの場合も「ずれによる不変性」に注意すると,

$$U_t : Y = \sum a_i X_{t_i} \longrightarrow U_t Y = \sum a_i X_{t_i + t}$$

は H の中のユニタリー作用素の one parameter group を定める. この $\{U_t\}$ に Stone のスペクトル分解定理を適用すると, $\{X_t\}$ そのもののスペクトル分解が得られる. これは屢々 Cramer 分解定理とよばれる.

また H_t を $X_s, s \leq t$ で張られる線形閉部分空間とすると, $\{H_t\}$ は増大する線形閉部分空間の系となる. H_t は線形理論の立場で t までに $X_u, u \leq t$ の与える情報を示していると考えてよい. $s < t$ のとき $H_t \ominus H_s (\equiv H_t \cap H_s^\perp)$ は (s, t) の間に真に新たに加わった情報を示し, この立場から更にすすんで inovation という重要な考えに到達している.

しかしこのような線形理論では $X_t^2 + X_u X_v$ などという非線形量は考察できない. たゞこの線形理論が閉じた理論となるのは, Gauss 過程である. これは任意の有限個の時点の結合分布が Gauss 分布(上の決意により平均 0 としてよい)となる確率過程である. このときには, 直交することは独立と同等になる. しかも X_t (または H の元 Y) を $X_u, u \leq t$ の関数で平均自乗の意味で最良近似するには, 結局一次結合だけ考えたよいことがわかり, したがって H_s への射影を考えたら

よい．　現代的な確率過程論の立場からいえば，線形理論は Gauss 過程に制限した理論といってもよい．実際確率過程論が応用諸分野で生れた頃には，すべての自然な分布は Gauss 分布と考えられて居り，これに伴って線形理論が生じてきたので，Khinchin や Kolmogorov の上述の理論はこれを数学的にまとめたのである．他方応用分野でも Gauss 分布だけに局限することは適当でない（例えば災害統計における Poisson 分布）ことが特に統計学者によって指摘され，それに応ずるものとして，測度論的な確率過程論が生れたのである．

Gauss 過程の見本関数の性質について数学的に興味ある研究がなされているが，これは Wiener 過程の見本過程の性質に関する Wiener, P. Lévy の研究の延長上にあるものであるから，新概念というべきものではない．

9. 確率解析 (stochastic analysis)

解析学の根は微分積分学であり，微分積分学の根は微分方程式である．現在では微分積分学を学んでから，微分方程式論を習うが，これは論理を明快にするためで，物理諸現象の中に微分方程式的な考えは微分積分学が完成する前から存在したのである．微分の考えは非線形的な現象を微小時間では

線形的なものにおきかえられる(高位の無限小を無視する)という，所謂局所的線形化(local linearization)の理論である．

確率解析についても同じで，極めて素朴な確率微分方程式は確率積分や確率微分の考えの出る前からあらわれている．瞬間毎に独立な randomness (または noise) が入ってきて，問題の現象に影響を与えるという形で，物理学や統計学で特殊な確率微分方程式が考えられてきた．例えば

(i) Langevin 方程式 $\dfrac{dX_t}{dt} - \alpha X_t = W_t$

(ii) $a_n \dfrac{d^n X_t}{dt^n} + a_{n-1} \dfrac{d^{n-1} X_t}{dt^{n-1}} + \cdots + a_0 X_t = W_t$

などである．こゝで $\{W_t\}$ は互に独立な量で white noise とよばれる．これは定常過程で平均 0，共分散 $V_{s,t} = \delta_{t-s}$ (δ_a は a に集中した Dirac の δ-関数) の場合である．(これの厳密な意味は超関数に対応する超過程の理論で解明される．) (i) (ii) はいずれも線形理論の範囲で許のすむものである．局所線形化の思想が必要なのは，Kolmogorov の拡散過程の導入においては，彼が dX_t について直観的に考察したことは

$$dX_t = a(t, X_t)\,dt + b(t, X_t)\,W_t\,dt$$

である．W_t は無限大の分散をもつ ($V(W_t) = \delta_0$) ので，Lévy は $W_t\,dt$ を $\xi_t\sqrt{dt}$ とかいて，$V(\xi_t) = 1$ とするようにした．これはむしろ W_t を Wiener 過程の導関数と見て

$$dX_t = a(t, X_t)\,dt + b(t, X_t)\,dB_t$$

とかいた方が意味がはっきりする. この意味は積分形にして

$$X_t = X_0 + \int_0^t a(s, X_s)\,ds + \int_0^t b(s, X_s)\,dB_s \qquad (1)$$

とすればよい. たゞこの第2の積分は今までの積分論ではあつかわれていないもので確率積分を導入する必要がある. これは純粋に確率論的なもので, vector valued integral や超関数的積分としても理解できないものである.

上の式を解いて, Kolmogorov 型の拡散過程をつくるのに, 確率微分や確率積分の意味を確定し, その calculus を定めることが確率解析学の第一歩である. その際最も基本的なものは通常の微分積分学における Leibniz の chain rule に相当する stochastic chain rule:

(Itô) $\quad df(X) = \sum_i f_i\,dX_i + \frac{1}{2}\sum_{ij} f_{ij}\,dX_i\,dX_j$

(Stratonovich)

$$df(X) = \sum_i f_i \circ dX_i$$

である. 後者は

$$Y \circ dX = Y \cdot dX + \frac{1}{2}\,dX \cdot dY$$

を用いて, 前者から出るが, その形が古典的な chain rule と同じであるため便利なこともある.

(1)にでる特殊な形の積分にだけ意味をつけて，(1)を解いたとしても，stochastic calculusができるとはいえない．もっと広い世界で考え，その中の一つの微分方程式として(1)を考察するとき始めて確率解析への道が開けるのである．その意味で上記の stochastic chain rule は最も基本的なもので，これによって従来分点をこまかくして極限をとって長い計算の後やっと得られた結果が，見通しよく数行の確率微分の変形で直ちに得られることになったのである．そのような立場に立って確率制御の問題も定式化できるようになったのである．

確率微分や確率積分は次節にのべる増大情報の立場に立ち，マルチンゲール的定式化をして，さらに透徹した理論となった．このことは Doob が始めて注意し，P. Courrege, D.L. Fisk，渡辺信三，国田寛，P. Meyer などによって完成されたのである．（渡辺：確率微分方程式 1975 産業図書，N. Ikeda－S. Watanabe: *Stochastic Differential Equations and Diffusion Processes*，講談社，J. Wiley 近刊）．

10. 増大情報系

確率変数 X, Y, Z の値がすべてきまると，X, Y, Z の任意の関数 $f(X, Y, Z)$ の値がきまる．X, Y, Z の値がきまったとき，

その値がきまるような確率変数というのは，実は X, Y, Z で生成される σ 加法族 $\sigma[X, Y, Z]$ で可測な確率変数の全体である．このような意味で P 可測集合族（これは σ 加法族）の部分 σ 加法族 $\mathcal{F}$ は $\mathcal{F}$-可測なすべての確率変数 $L^0(\mathcal{F})$ の全体に対応し，これらの確率変数の値で定まる「情報」をあらわすと考えてよい．σ-加法族 $\mathcal{F}$ による条件付確率 $P(A|\mathcal{F})$ は $\mathcal{F}$ 可測な確率変数 Y で

$$P(A \cap B) = \int_B Y(\omega)\, P(d\omega), \quad B \in \mathcal{F}$$

となるものとして定義される．特に $\mathcal{F} = \sigma[X, Y, Z]$ のときには，

$$P(A|\mathcal{F}) = X, Y, Z \text{ が定まったときの条件付確率 } P(A|X,Y,Z)$$

となる．このように条件付確率を σ 加法族に対して定義するのは Doob による．同じ考え方から，σ 加法族 $\mathcal{B}_1, \mathcal{B}_2, \mathcal{B}_3, \cdots, \mathcal{B}_n$ が独立とは

$$P(B_1 \cap B_2 \cap \cdots \cap B_n) = P(B_1)P(B_2)\cdots P(B_n), \quad B_i \in \mathcal{B}_i$$

で定義する．この立場で σ 加法族は「情報」をあらわし，$P(A|\mathcal{F})$ は情報 $\mathcal{F}$ の下における条件付確率であり，σ 加法族 $\mathcal{B}_1, \mathcal{B}_2, \cdots, \mathcal{B}_n$ が独立というのは $\mathcal{B}_1, \mathcal{B}_2, \cdots, \mathcal{B}_n$ の各々から与えられる情報が独立であることを意味する．

このような言葉を用いると，$\{X_t\}$ が加法過程であるとは，

$X_t - X_s$ が，すべての $t > s$ に対し，$\mathcal{F}_s \equiv \sigma[X_u, u \leqq s]$ と独立であるといえばよい．また $\{X_t\}$ が Markov 過程であるとは，

$$P(X_t \in E \mid \mathcal{F}_s) = P(X_t \in E \mid \sigma[X_s]) \qquad t > s$$

といえばよい．そうして右辺はまた $p_{s,t}(X_s, E)$（$p_{s,t}$ は推移確率）に等しいわけである．マルチンゲールも上の $\mathcal{F}_s$ を用いて

$$E(X_t \mid \mathcal{F}_s) = X_s \qquad s < t$$

と定義すればよい．

ところがこの定義では $\{X_t\}$，$\{Y_t\}$ が加法過程でも $X_t + Y_t$ は加法過程とはいえない．マルチンゲールについても同様である．それは $\sigma[X_u, u \leqq s]$ と $\sigma[Y_u, u \leqq s]$ とは一般に違っているからである．それでいくつかの加法過程やマルチンゲールを考えるときには，一つの大きい増大情報系 $\{\mathcal{F}_t, t \in T\}$ を指定しておき，すべての t に対し X_t が $\mathcal{F}_t$ 可測となるもののみを考え，このような $\{X_t\}$ を $\{\mathcal{F}_t\}$ に適合しているという．$\mathcal{F}_t = \sigma[X_u, u \leqq t]$ とすれば，$\{X_t\}$ は勿論 $\{\mathcal{F}_t\}$ に適合するが，もっと大きい $\mathcal{F}_t$ をとっても適合関係はなりたつ．たとえば

$$\mathcal{F}_t = \sigma[X_u, Y_u, Z_u, u \leqq t]$$

とすれば，$\{X_t\}$，$\{Y_t\}$，$\{Z_t\}$ のすべてが $\mathcal{F}_t$ に適合し，また

$$U_t = \varphi_t(X_t, Y_t, Z_t) \qquad \varphi_t : R^3 \to R \ (\text{Borel可測})$$

も$\{\mathcal{F}_t\}$に適合する.

さて $\{X_t\}$が$\{\mathcal{F}_t\}$に適合し, すべての$s<t$に対し$X_t - X_s$が$\mathcal{F}_s$と独立のとき, $\{X_t\}$は$\{\mathcal{F}_t\}$に関して加法過程であるという. Markov過程やマルチンゲールについても同様である. $\{X_t\}$, $\{Y_t\}$が$\{\mathcal{F}_t\}$に関して加法過程であれば, $X_t + Y_t$もそうである. これはマルチンゲールに対してもそうである. (Markov過程にはそうはいえない.)

このようにマルチンゲールをいくつかとり扱うときには, それがある一定の増大情報系$\{\mathcal{F}_t\}$に関連したものを考えた方がよい. この立場でWiener過程に対しては

$$X_t - X_s \text{ が } \mathcal{F}_s \text{ と独立でかつ } N_{0,t-s} \text{ に従う}$$

ものを考え, これを$\{\mathcal{F}_t\}$-Wiener過程という. これは勿論$\{\mathcal{F}_t\}$マルチンゲールであるから, $\{\mathcal{F}_t\}$-Wienerマルチンゲールということもある. 確率微分確率積分も$\{\mathcal{F}_t\}$-マルチンゲールについて考えると, 始めて透明な確率解析の理論ができる.

またエルゴード理論についても, この増大情報系の考えが必要になる. $\{T_t\}$を(Ω, P)上の保測変換群とする. もし増大情報系$\{\mathcal{F}_t\}$が存在して, $T_s\mathcal{F}_t \equiv \{T_sA : A \in \mathcal{F}_t\} = \mathcal{F}_{t+s}$,

$\bigcap_t \mathcal{F}_t = \mathcal{Z} \equiv \{\phi, \Omega\}$ (自明な情報),

$\bigvee_t \mathcal{F}_t$ ($\equiv \mathcal{F}_t$をすべて含む最小のσ加法族) $= \mathcal{F}$ ($\equiv$すべてのP可測集合の族)

がなりたつときには, $\{T_t\}$ は Kolmogorov 型 (Kolmogorov の流れ) という. このときには

$$\mathbb{H}_t = L^2(\Omega, \mathcal{F}_t, P|\mathcal{F}_t), \quad \mathbb{H} = L^2(\Omega, \mathcal{F}, P)$$

とおくと, これが線形理論における H_t, H の非線形的拡張となる.

条件付確率やマルチンゲールに対する情報の利用は Doob の考えであり, エルゴード理論の場合は Kolmogorov による. ただ Kolmogorov は σ-加法族のかわりに, 分割をもって情報を定義している (伊藤: 確率論, 岩波基礎数学講座).

この増大情報系の考えは今後益々重要となると思われる.

結び

以上確率過程論発展に当って導入され, かつその後この理論に新生面を開いた新概念をいくつか示した. 話を簡とわかり易くするために, 所々論理的には厳密でない所や細かい条件を無視した. この他にも確率超過程や一般の空間上の値をとる確率過程などいろいろあろう. また この新概念が有効に作用したのは, 多数の数学者により導入された幾多の巧みな推論技術や計算技術によることはいうまでもない.