確率論（数学３年後期選択） probab.tex
服部哲弥

20000420;21;23-30;0503-07;10;15;18-22;25;28;29;0603-08;11-13;15;18-21; 23-0708;10-16;
(Tauber) 20000822;24-30; 0912(文献 Feller 訳); 13(誤植 LDP);15; 17 (Intro 省略箇所);
21;23;1010;1101;08;
15(弱収束同値条件証明 (iv) → (v));29;1206;23;26;27(Kosugi);29;30(大前君);
20010110(集合族の独立性は生成するσ加法族の独立性を意味しない);16; 20020312;0317(正則変動関数 A.2.1,A.2.2, 凸関数 A.3.3 ミスプリ等);
20030119(中心極限定理 cos の評価絶対値忘れ);
20030528; 20030619(theenumi=romanenumi, 独立性のσ [] による表現の命題落丁);
20030626(Fubini の証明概要 χ を X と誤記２カ所);
20031111(熊谷);
20041005(1.1.3cumulant 誤植, 命題 1 注 (ii) 誤植,cvmblefcn3b 誤植, (4) 誤植, 命題 10 証明直後誤植);
20041020(小誤植);21(§1.1 モーメント説明ミス); 10041027 定理 18((v) → (i) 証明「←」を「→」に訂正，他小誤植); 20041103(小誤植);1110(BC2 の証明左辺が抜けていた); 17(Kronecker 補題直前 X_n/n iid →独立); 19(小補正，特性関数の一様収束証明があまりに不親切だったのを改訂);
24(定理 45 の証明コルモゴロフ不等式適用時指数符号間違い, 補遺 C の最後の定理 Hedgeng → Hedging); 1216(概収束→確率収束命題 17 は 優収束定理だけから出る，船野敬君の指摘);
20050112(補遺ブラウン運動の細かい修正と文献追加，RG の方法の追加．ただし，全体の説明は不親切なまま，要改訂);
20050711(大数の法則，増分が L^2 有界のとき実質 martingale の L^2 maximal 不等式); 20051212(小誤植); 20110909(一様可積分の同値条件 F が n-dep でも良いとすべき);

確率論

目 次

0 イントロ.

講義の目標. 確率という概念は，賭博等と関連して，大変古くからあったと思われる．これに統計学（人口統計，品質管理）や遺伝学も加わって，古典的な確率論の数学が発展した．

近代から現代への移行期に重大な研究対象が加わった．19 世紀にブラウン運動という物理現象が発見され，工学で知られていたホワイトノイズとともに，20 世紀の重要な理論的概念が人類の前に姿を現した．(二つの対象のように書いたが，両者とも分子の熱運動に由来するという意味では一つの物理現象である.）連続変数（時間変数）をパラメータとする非可算個の確率変数の列を扱う確率過程論の機運が熟していた．

他方，以上とは独立に 20 世紀初頭にルベーグが測度論を創始して，我々が「大きい（長い，広い)」「小さい（短い，狭い)」と感じる定量（長さ，面積，体積）の概念を σ 加法性を持つ非負値集合関数として理解することが数学的に自然であることを発見した．

コルモゴロフは，当時最先端の数学であったこのルベーグの測度論に立脚して，古典的な「さいころと人口分布と遺伝学」の確率論と分子の熱運動という当時の最先端の理論物理の課題を統一的に扱う数学的枠組み（公理）を完成した．これが 20 世紀の確率論の始まりである．この講義はこの 20 世紀の確率論の基礎を講義する．

解析学は元々は関数に関する性質を極限や不等式を通して調べる学問である．その具体的な手段として微積分が圧倒的に強力である．20 世紀の確率論は積分（期待値）を通して関数（確率変数）の性質を扱うので，解析学の一分野と見ることができる．

測度論に立脚した解析学としては他に関数解析も基礎科目である．強いて違いを言うならば，関数解析は関数の集合を完備な距離空間（関数空間）としてとらえ，近似列によって関数を調べようとするし，確率論は値の分布によって関数を調べると言えなくもないが，差を議論するのはおろかかもしれない．実際，確率過程論の中の一つ拡散過程は拡散方程式と呼ばれる微分方程式と密接な関係にある．20 世紀の確率論の発展は解析学の一分野としての側面が強くなっている．確率論は他の解析学の分野と分けて応用数学に分類することもあるが，20 世紀後半の趨勢からは少々時代遅れかもしれ

ない．もちろん，解析学の他の分野と同様に，数学以外の分野との関わりは確率論自体の発展にとってとてつもなく重要である．統計学，数値解析，統計力学，数理経済学，などとの関わりが指摘できる．

解析学の一分野としての性格が強まった結果として，今日の確率論の講義は,「さいころ」の素朴直感的な確率論に比べると，直感を排した，初学者に抽象的と見えるものになった．上に書いたような高度な現象をも解析できる数学的枠組みが必要だからである．

この講義の到達範囲は（20 世紀後半の解析学としての驚異的な発展ではなく）素朴な（さいころの，高校で習う）確率論の範囲を 20 世紀の（解析計算のしやすい）定式化で表現することである．即ち，この講義では，測度論に基づく確率論の定式化と，その定式化によって自然に導かれる基礎的初等的な事項をできるだけオーソドックスに紹介する．

独立確率変数の和に関する極限定理，具体的には中心極限定理など，を目標とする．時間が余ればランダムウォークあるいは離散時間のマルチンゲールを紹介したい．これらはいずれも自然数をパラメータとする確率変数列に関する基礎的かつ高度な概念であり，20 世紀確率論のもっとも輝かしい成果と考えられる確率過程論の，欠くべからざる準備となる．次のステップの講義で，確率過程論の初歩でありかつ最も重要な典型例であるブラウン運動にすんなり入っていくための直前までの環境整備，ということであり，これがすむことをこの講義としての（おそらく実現不可能な）理想とする．

省略するもの． この講義では，確率論の講義で比較的標準的と思われる内容のうち，特に以下を割愛した．

(i) $\mathbf{R}$ 上の具体的な確率測度関連や統計学関連事項：推定検定，無限可解分布，分布の具体計算，分布関数の諸性質[1]．（アクチュアリー関連では重要なので希望者は各自補充すること．）

(ii) 確率空間の構成に関する少し込み入った事項：Caratheodry–Hopf の拡張定理（有限加法的測度から測度の構成），直積測度の構成，コルモゴロフ拡張定理（無限次元確率測度の構成），Prokhorov の定理（tightness argument，例えば分布の Helly の議論の証明），条件付き期待値．

以上を講義で省略するということは，講義では具体的に背後の空間を構成しなくても言えること（定理の仮定を満たす確率測度が本当に存在するかどうかは検証しないでも言えること）に講義の主題を置くということである．つまり，講義の主題は「もし確率測度や確率変数の定義を満たす対象が与えられたならば何が成り立つか」であって，講義の定理が成り立つような具体例をあげることや，ある対象が確率測度や確率変数になっていることを証明する方法は目標としなかった．

もちろん，このような建前はこだわれば有害無益になる．あくまで学ぶべき内容を精選する[2] ための「割愛」である．受講生には,「講義で言わなかったこと」を各自補充することを期待する．

仮定する知識． 基礎的な初等数学と解析系の基礎数学，特に，初等微積分学，集合と位相，ルベーグ積分論（測度論）．

測度論（ルベーグ積分論）は，2000 年数学科の水準で 1.5 時間 1 年または 3 時間半年のコース程度．特に測度空間の定義や基本性質，および，Hopf の拡張定理は既知として扱い，また，積分の定義と基礎性質も既知とする．

測度論から何を引用するかは「復習」と呼んで明示し，また，随所で定理の「気持ち」を説明することで結果として測度論の「魂」も伝わることを期待するので，測度論からの引用を（公理のように）与えられた正しいものと認めて先に進む限りは一応，測度論と独立に学べる．測度論の講義で出てこなかった題材があっても原理的には心配しなくてよいはずである．

しかし，そうは言っても，測度論の勉強と確率論の勉強を一つの科目でやろうという計画は中身が濃すぎる．解析学の方面によほどの才能がない限り，経験上も勧められない．少なくとも基礎的な測度空間と積分の一般的な定義の「こころ」や基礎公式，および，ルベーグ測度の構成くらいは一度は勉強してあることが必要条件である．

[1] 分布関数を軽く扱ったのは無理がありすぎた．
[2] 日常用語を避けて本質を隠してはいけない．「精選する」とは「内容を減らす」ということだ．

1 確率空間，確率変数，期待値．

確率論とは，**確率変数**の**確率**的性質を**期待値**という視点を中心に調べる現代数学（特に解析学）の一分野である．

1.1 確率空間，確率変数，期待値．

1.1.1 確率空間．

確率空間 とは，下記３つ組 $(\Omega, \mathcal{F}, \mathrm{P}[\ \cdot\ ])$ のことである．

Ω: 集合,

$\mathcal{F}$: Ω の部分集合を要素とする σ 加法族,

$\mathrm{P}[\ \cdot\ ]$: $(\Omega, \mathcal{F})$ 上の測度であって $\mathrm{P}[\,\Omega\,] = 1$ を満たすもの.

$\mathrm{P}[\ \cdot\ ]$ を**確率測度**，$\mathcal{F}$ の要素を**事象**，$A \in \mathcal{F}$ に対して $\mathrm{P}[\,A\,]$ を A の（生じる）**確率**ということ，は周知だろう．即ち，事象とは確率が定義されている集合のことである．

1.1.2 測度（復習）．

可測空間と測度空間． 集合 Ω の部分集合たちを要素とする空でない集合族 $\mathcal{F}$ が σ 加法族であるとは，可算和と補集合という２つの演算に関して閉じていること．このとき $(\Omega, \mathcal{F})$ を可測空間という．

注. 空の集合族（空集合）は σ 加法族とは呼ばない．逆に言えば，σ 加法族であることを示すには，空でないことと可算和と補集合に関して閉じていることを言わないといけない．以下このような注意を省略する． ◇

$\mathcal{F}$ を定義域とし実数値または $+\infty$ を値にとるが恒等的に ∞ とはならない関数（Ω 上の集合関数）$\mathrm{P}[\ \cdot\ ] : \mathcal{F} \to \mathbf{R} \cup \{\infty\}$ が測度であるとは，非負，かつ，σ 加法性を持つこと．

注. あらゆる部分集合に測度が定義可能ならありがたいが，そのような主張はある合理的な公理（選択公理）と矛盾する場合がある（実数上のルベーグ測度）．そこで，定義域 $\mathcal{F}$ は常に明らかにして始めないといけない．

また，定義域の大小は情報の大小（どれくらい複雑な関数を考察の対象とするか）を意味するので，σ 加法族自体に確率論上の積極的な意味がある．σ 加法族はその意味で積極的にとらえるべきであり，抽象的で煩わしいと考えるべきでない． ◇

可測関数． 可測空間 $(\Omega, \mathcal{F})$, $(S, \mathcal{S})$ と関数 $X : \Omega \to S$ があるとき，X が可測であるとは，$X^{-1}(G) := \{\omega \in \Omega \mid X(\omega) \in \mathrm{G}\ \} \in \mathcal{F}$, $G \in \mathcal{S}$, となること．[3]

関数が可測かどうかは σ 加法族による．複数の σ 加法族を同時に考察するときは $\mathcal{F}/\mathcal{S}$–可測と書くこともある．

[3] $\mathcal{S}$ の全ての要素（集合）に対して $X^{-1}(G) \in \mathcal{F}$, という主張をこのように書くことがある．以下でも同様．$(\forall G \in \mathcal{S})\ X^{-1}(G) \in \mathcal{F}$ という書き方との慣用上の使い分けは筆者には分からない．

Borel 集合族（Borel σ 加法族）． 集合 Ω の部分集合たち（全部というわけではない）からなる集合族 $\mathcal{A}$ があるときこれを含む最小の σ 加法族が存在する（復習問題）．これを $\sigma[\mathcal{A}]$ と書く．$\mathcal{A}$ を含む最小の σ 加法族とは，$\mathcal{A}$ の要素は全て要素として持つ σ 加法族であって，そのような性質を持つどんな σ 加法族にも含まれているもののこと．

$\mathbf{R}$ や，一般に $d \in \mathbf{N}$ に対する $\mathbf{R}^d$ など[4]，一般に位相空間において，その開集合を全て含む最小の σ 加法族を Borel 集合族という．これは閉集合を全て含む最小の σ 加法族でもある．

ルベーグ測度． 特に $\mathbf{R}^d$ の Borel 集合族を d 次元 Borel 集合族と呼び，$\mathcal{B}_d$ と書く．$\mathcal{B}_d$ は区間（直方体）を全て含む最小の σ 加法族でもある（復習問題）．

1 次元 $\Omega = \mathbf{R}$ の場合について，区間 $(a,b]$ $(a<b)$ に対して $\mu_1((a,b]) = b-a$ を満たす $(\mathbf{R}, \mathcal{B}_1)$ 上の測度がただ一つ存在する（復習：E. Hopf の拡張定理）．これを**ルベーグ測度**という[5]．

1 次元の場合と全く同様に，一般に「d 次元直方体に対しては全て通常の d 次元体積を与える」という性質を持つ測度 μ_d が $(\mathbf{R}^d, \mathcal{B}_d)$ 上に一意的に存在する．これを d **次元ルベーグ測度**という．

実数値可測関数． 可測関数の値域が実数または $\{\pm\infty\}$ のとき実可測関数ということにしよう．値域が実数ならば実数値可測関数ということにする．測度論では通常は値として $\pm\infty$ を許しておいたほうが便利である．

関数が可測かどうかは値域の σ 加法族によって変わる．この講義（に限らず極めて多くの場合）では実可測関数というときは値域の σ 加法族は

$$\mathcal{B}'_1 = \sigma[\mathcal{B}_1, \{\infty\}, \{-\infty\}]$$

である，と約束する．多くの場合 $\pm\infty$ の付加は便宜的である．以下では，両方とも d **次元 Borel 集合族**ということにし，$\pm\infty$ は重要でない状況では記述を省略することもあるものとする．

可測性は定義域の σ 加法族にもよる．複数の σ 加法族を同時に考察するときは $\mathcal{F}$–可測と書く．

以上をまとめると，

> $X : \Omega \to \mathbf{R} \cup \{\pm\infty\}$ が可測空間 $(\Omega, \mathcal{F})$ 上の**実 ($\mathcal{F}$–) 可測関数**であるとは $A \in \sigma[\mathcal{B}_1, \{\infty\}, \{-\infty\}]$ なる全ての A に対して $X^{-1}(A) = \{\omega \in \Omega \mid X(\omega) \in A\} \in \mathcal{F}$ が成り立つことを言う．
> 値域が d 次元実数空間 $\mathbf{R}^d$ の場合も同様に，σ 加法族は $\mathcal{B}_d$ を考える．また，状況に応じて，各成分について $\pm\infty$ も許すものとする．このとき d 次元実可測関数ということにする．

この講義では（最後のほうを除くと）値域としては（$\pm\infty$ を許した）実数 $\mathbf{R}$ または d 次元実空間 $\mathbf{R}^d$ しか出てこない．無限次元確率空間（例えば $\mathbf{R}^\infty$）も最後のほうで出てくるが，それは追い追い議論する．

1.1.3 確率変数と期待値．

確率変数． $(S, \mathcal{S})$ を可測空間とする．$X : \Omega \to S$ が，確率空間 $(\Omega, \mathcal{F}, \mathrm{P}[\ \cdot\ ])$ 上の**（S 値）確率変数**であるとは，X が $(\Omega, \mathcal{F})$ 上の S 値可測関数であることをいう．可測関数の定義で注意したように，値域 S の可測空間としての構造 $\mathcal{S}$ を指定しないと確率変数かどうか分からないことになるが，この講義では $(S, \mathcal{S})$ は実数と Borel 集合族 $(\mathbf{R}, \mathcal{B}_1)$ およびその多次元版 $(\mathbf{R}^d, \mathcal{B}_d)$ （但し，可測関数の定義で約束したように，$\pm\infty$ を値として許す）しか扱わない．即ち，$X : \Omega \to \mathbf{R} \cup \{\pm\infty\}$ が，確

---

[4] 確率論を含めた数学では n 次元空間 $\mathbf{R}^n$ と書くが，n は確率変数列などで多用するので，物理の流儀に従って $\mathbf{R}^d$ に限り，d を自然数として使う．Dimension の d から来ている．

[5] (a,b) や $[a,b]$ で定義しても同じルベーグ測度 μ_1 を与える．$\mathcal{B}_1$ は比較的標準的な記号だが，μ_1 は標準的な記号がないように思う．この講義では μ_1 と書くことにする．詳しく言うと，このような性質を持つ測度は，$\mathcal{B}_1$ よりももっとたくさんの集合に対しても拡張定義可能で，ある意味でその拡張は一意的である（**測度の完備化**）．この拡張された測度のことをルベーグ測度というのが普通である．しかし，この講義では完備化が本質的になる問題は扱わないので，ここでは基本的に $\mathcal{B}_1$ 上の測度のみ考察する．

率空間 $(\Omega,\mathcal{F},\mathrm{P}[\ \cdot\ ])$ 上の**実確率変数**であるとは， X が $(\Omega,\mathcal{F})$ 上の実可測関数であることをいう．$\mathbf{R}^d$ 値確率変数（d 次元実確率変数）$\vec{X}=(X_1,\cdots,X_d):\Omega\to\mathbf{R}^d$ も同様である．必要に応じて $\mathcal{F}$– と修飾語を入れる．

> 確率変数とは可測関数のことである．

発展：「無限次元空間」に値をとる確率変数． この講義では当分の間 $\mathbf{R}^d$ 値確率変数（d 次元実確率変数）とその極限定理（期待値などをとってから $d\to\infty$ としたもの）しか出てこないが，値域として無限次元 (‘$d=\infty$’) も非常に重要である．

簡単な説明．実確率変数の列 $X_k,\ k=1,2,3,\cdots,$ が与えられたとき，単に $\lim_{n\to\infty}\mathrm{E}[\ \sum_{k=1}^{n}X_k\ ]$ を考えるだけなら，$\mathbf{R}^n$ 値確率変数 $(X_1,\cdots,X_n)$ しか出てこない．しかし，各 $X_k:\ \Omega\to\mathbf{R}$ は関数だから，各 $\omega\in\Omega$ 毎に数列 $X(\omega)=(X_1(\omega),X_2(\omega),\cdots)$ が与えられる．$\omega\in\Omega$ は数列を指定する（区別する）変数であり，その意味で数列の集合と思える[6]．そういう数列の集合上の関数として X をとらえることができると，数列（無限個の実数，即ち無限次元空間）上の確率論（測度論）が展開できることになる．このようなことは実現可能であり，今日の確率論の中心的研究分野である．しかも，パラメータ（添字）n は連続変数にすることもできる．

可算無限個の実確率変数をひとまとめ $X=(X_n)_{n\in\mathbf{Z}_+}$ にしたものが，ある確率空間上の確率変数であるとみることができるとき，$X:\ \Omega\to\mathbf{R}^\infty$ を確率連鎖という．同様に， $I\subset\mathbf{R}$ を実数の区間（無限でもよい）とし，各 $t\in I$ に対して X_t が確率空間 $(\Omega,\mathcal{F},\mathrm{P})$ 上の確率変数のとき，t についてまとめてあつかったもの $X=X_.$ を確率過程と言う．

単にまとめて扱った，だけでなく，通常は各 ω 毎に t について何らかの性質を仮定するときに，これらの概念は極めて興味深い対象になる．例えば各 ω 毎に $X_t(\omega)$ は t の実数値関数であるが，t について連続である確率が 1 ならば，$X=X_.$ は Ω から連続関数の集合 C^0 への関数と思うことができる．このような見方は，最初に確率変数を定義したときの，各 n,t 毎に Ω 上の関数 $X_n:\ \omega\mapsto X_n(\omega)$ とみる見方と視点が変化していることに注意されたい．これらは，今日研究対象としても応用上も非常に重要な視点である．

確率論特有の記号． 実確率変数 X と可測集合 $A\in\mathcal{B}_1$ があると，可測集合 $\{\omega\in\Omega\mid X(\omega)\in A\}\in\mathcal{F}$ が定義できる．

> 確率論では $\{X\in A\}=\{\omega\in\Omega\mid X(\omega)\in A\}$ などと略記することが多い．

$f(X(\omega))$ のような，より複雑な条件式でも同様である．

同様に集合の外でも，関数に関する「各点で成り立つ性質」のときに変数を省略して記述することがある（この講義ではそうする）．$X\le Y$ は $X(\omega)\le Y(\omega),\ \omega\in\Omega,$ を意味する．

確率論以外の数学ではこのような略記を行わないため，最初にこの略記法を使うことを断ることになっているが，Ω を表に出さない視点は，確率論の技術的な強みの由来を表しているように私には思えるので，この講義では積極的に略記法を使う．

> ある命題 $A(\omega)$ が確率 1 で起きるとき，即ち
>
> $$\mathrm{P}[\ A\ ]=\mathrm{P}[\ \{\omega\in\Omega\mid A(\omega)\}\ ]=1$$
>
> とき A, a.e., と書き，**殆ど**至る所 A が成り立つ，などという．

[6] 何も言わずに確率変数列を考える，と書いた瞬間に，実は暗黙のうちにそのような Ω があると仮定していたのだ！だから厳密に言えば，以下全ての定理は「そのような数列の集合があって，σ 加法族や確率測度を含めたしかるべき性質があれば，以下の定理が成り立つ」と言っているだけで，そのような都合のよい確率空間が本当に矛盾なくあるのかどうかは別途議論の必要がある．もちろんこの講義の定理は全てその意味で数学的に空でないが，その証明をとりあえず後回しにしても論理的に困らないところが確率論の定式化の技術上の一つの強みである．

定義通り，確率 1 で A が起きる，と言ってもかまわない．複数の確率測度を同時に考えていて区別したいときは A, P–a.e., などと書き，P に関して殆ど至る所（殆ど常に） A が成り立つ，などという．

確率論の由来に戻って考えれば確率 0 とは「その現象が起きない」という「気持ち」だから，無視すべきである．その意味では「必ず」という言葉は確率論ではあまり出てこないはずで，代わりに「殆ど」という言葉が普通に用いられる[7]．そうはいっても確率論も数学だから，定義としては,「殆ど」という単語は「確率測度 1 」という厳密な定義の言い換えでしかないことは言うまでもない．

確率論特有という意味ではないが,

$$a \wedge b = \min\{a, b\} = \frac{1}{2}(a + b - |a - b|),$$
$$a \vee b = \max\{a, b\} = \frac{1}{2}(a + b + |a - b|),$$

はそれぞれ小さい方と大きい方を表す.

集合 $A \subset \Omega$ に対して A の定義関数を χ_A と書く：

$$\chi_A(x) = \begin{cases} 1, & x \in A, \\ 0, & x \notin A. \end{cases}$$

期待値． 確率空間 $(\Omega, \mathcal{F}, \mathrm{P})$ 上の実確率変数 X を考える．

$$\mathrm{E}[\,X\,] = \int_\Omega X d\mathrm{P} = \int_\Omega X(\omega)\mathrm{P}[\,d\omega\,]$$

が存在するとき，これを X の**期待値**という．即ち，期待値とは積分のことである．
$A \in \mathcal{F}$ における積分を

$$\mathrm{E}[\,X;\ A\,] = \int_A X(\omega)\mathrm{P}[\,d\omega\,] = \mathrm{E}[\,X\,\chi_A\,]$$

と書く．この記法は集合の省略記法とあわせて使うと便利．例えば

$$\mathrm{E}[\,X;\ X \geq a\,] = \mathrm{E}[\,X;\ \{\omega \in \Omega \mid X(\omega) \geq a\}\,].$$

特に

$$\mathrm{E}[\,\chi_A\,] = \mathrm{E}[\,1;\ A\,] = \mathrm{P}[\,A\,].$$

積分は面積である．$\{X \geq 0\}$ （正の部分）と $\{X < 0\}$ （負の部分）でそれぞれ面積を求め，その差で積分が定義されるのであった．確率変数は可測関数なので，積分が存在するかどうかは値として ∞ を許せば，正の部分と負の部分の差が定義できるかどうかだけの問題である．即ち $\infty - \infty$ の形になるときのみ期待値は定義されない．特に,

$\mathrm{E}[\,|X|\,]$ は ∞ を許せば必ず存在し，$\mathrm{E}[\,|X|\,] < \infty$ ならば $\mathrm{E}[\,X\,]$ も存在する．

モーメントと分散，母関数． $f:\ \mathbf{R} \to \mathbf{R}$ が可測関数ならば $f(X) = f \circ X:\ \Omega \to \mathbf{R}$ も確率変数（可測関数）になる（復習：§1.2.2）から,

[7] ある研究集会の講演録を作っていた有能な秘書が「数学なのに「殆ど」という単語が出てくるが，間違いではないか？」と主催者の先生に問い合わせた，というジョークがある．

E[$f(X)$] も同様に定義される．
特に，$n \in \mathbf{N}$ に対して $f(x) = x^n$ をとれば，E[X^n] が定義される．これを X の n 次**モーメント**という．

$Z(t) = \mathrm{E}[\ e^{-tX}\ ]$ を X の母関数という．

（指数部の符号はつけないのが普通のようだが，実際にはつけたほうがみやすい．）例えば $X \geq 0$, a.e.， ならば $\mathbf{Re}t \geq 0$ で $Z(t)$ の右辺が絶対収束して正則になる．(そうでなくても $X \geq -M$, a.e., ならば $X + M$ を考えることで同様．当然もっと緩めることができる．) 従って，広範な X に対して Laplace 逆変換により X の分布が求まる可能性がある．そういう意味で，原理的にはモーメントが全て分かれば

$$\sum_{n=1}^{\infty} \frac{(-t)^n}{n!} \mathrm{E}[\ X^n\ ] = \mathrm{E}[\ e^{-tX}\ ]$$

によって X の分布の母関数が求まり，(Laplace 逆変換ができれば）X の分布が分かる．即ち，

「モーメント E[X^n] ＋ $\displaystyle\sum_{n=1}^{\infty} \frac{(-t)^n}{n!} \mathrm{E}[\ X^n\ ]$ の収束 ＝ X の母関数 ＝ X の分布 ー α」

という標語を得る．母関数が有効な理由は（上記の意味で X の分布の全情報を持っているということ以外に）大きな X の寄与を切って収束性が良くなっているので極限定理などを作りやすい点である．欠点としては，X についてならした量なので分布の詳細な情報を得にくい点である．$t \downarrow 0$ の $Z(t)$ の振る舞いから X の値の下限付近の X の分布の振る舞いを得る定理を **Tauber 型定理**と総称する（§A）．

この他，モーメントは一様可積分性， L^p 収束などを経由する関数解析的な解析にも有効である．

母関数の対数の $(-t)^n$ の係数の $n!$ 倍，即ち

$$\log \mathrm{E}[\ e^{tX}\ ] = \sum_{n=1}^{\infty} \frac{t^n}{n!} u_n$$

の u_n を X の n 次の **cumulant** という．特に，$u_1 = \mathrm{E}[\ X\ ]$，また，

$u_2 = \mathrm{E}[\ (X - \mathrm{E}[\ X\ ])^2\ ] = \mathrm{V}[\ X\ ]$ を**分散**という．

X を確率変数，$c \in \mathbf{R}$ とし，期待値のみが異なる確率変数 $Y = X - c$ を考えると，X の n 次モーメントと Y の n 次モーメントは $n > 1$ で全く異なるが，

$$\log \mathrm{E}[\ e^{tY}\ ] = \log \mathrm{E}[\ e^{tX}\ ] - ct$$

なので，cumulant は全て一致する．即ち，

(i) 複数の確率変数の間で期待値のみが異なる（分布が平行移動しただけ）ことを検証するためには cumulant のほうがモーメントより適当である．
(ii) 期待値から大きくはずれた値をとる確率が小さい（定数関数に近い確率変数である）ことを言いたいときには cumulant，特に分散，を調べるのが自然である．

2 次元以上の実確率変数に対しても同様に期待値などが定義できる．特に，$\mathrm{V}_{12} = C(X_1, X_2) = \mathrm{E}[\ (X_1 - \mathrm{E}[\ X_1\ ])(X_2 - \mathrm{E}[\ X_2\ ])\ ]$ を**共分散**という．2 次元以上に関しては詳しいことは省略する．

1.1.4 積分（復習）.

確率空間 $(\Omega, \mathcal{F}, \mathrm{P})$ 上の実確率変数 X の期待値 $\mathrm{E}[\,X\,]$ とは関数 $X:\ \Omega \to \mathbf{R} \cup \{\pm\infty\}$ の積分のことであった：

$$\mathrm{E}[\,X\,] = \int_\Omega X d\mathrm{P} = \int_\Omega X(\omega)\mathrm{P}[\,d\omega\,].$$

積分の定義を復習しておく.

> $\mathrm{E}[\,X\,] = \displaystyle\int_\Omega X(\omega)\mathrm{P}[\,d\omega\,]$ の定義：
>
> (i) A 上恒等的に 1 他で 0 $(X = \chi_A)$ のとき， A の測度 $\mathrm{P}[\,A\,]$.
>
> (ii) 階段関数 (χ_{A_i} の線形結合) のとき， $\mathrm{P}[\,A_i\,]$ の線形結合．特に，X が定数関数のとき $\mathrm{E}[\,X\,] = X$.
>
> (iii) 非負値可測関数のとき，各点で単調増加な階段関数列で収束するもの $X_n,\ n = 1, 2, \cdots,$ がとれる.（2^{-n} 単位になるように値を切り捨て，かつ n 以上の値は n としてしまって X_n を定義すればよい.）$\displaystyle\int X_n d\mathrm{P}$ は上で定義したから，その単調増加極限で定義．$\mathbf{R} \cup \{\infty\}$ の意味で確定.
>
> (iv) 一般の可測関数 X のとき， $X^{\pm} = (\pm X) \vee 0$ とおく[8] と，$\displaystyle\int X^{\pm} d\mathrm{P}$ は確定する．少なくとも一方が ∞ でないとき $\mathrm{E}[\,X\,] = \displaystyle\int X^{+} d\mathrm{P} - \int X^{-} d\mathrm{P}$ の意味で**積分が確定**する．特に $\mathrm{E}[\,X^{\pm}\,] < \infty$ 即ち $\mathrm{E}[\,|X|\,] < \infty$ のとき**積分可能**，**可積分**等という．明らかにこれは $-\infty < \mathrm{E}[\,X\,] < \infty$ と同値.
>
> (v) 部分集合上の定積分について（可測集合ならば）：
>
> $$\mathrm{E}[\,X; A\,] = \int_A X d\mathrm{P} = \int_\Omega X \chi_A d\mathrm{P}.$$

注. (i) $A \in \mathcal{F}$ でないと $\mathrm{P}[\,A\,]$ が定義されないことに注意．従って階段関数というときも $A_i \in \mathcal{F}$. 可測関数でないと階段関数近似できないのはこのため．それ故積分が定義できるためにはまず可測関数でないといけない.

(ii) いったん可測関数に限れば，非負値関数の積分までは一直線（無条件で OK）．積分が定義できるための残された唯一の条件は $\infty - \infty$ にならないこと.（$\mathrm{E}[\,|X|\,] < \infty$ ならば十分.）

◇

1.2 初等的事項（半分復習）.

1.2.1 確率空間.

$(\Omega, \mathcal{F}, \mathrm{P})$ を確率空間とする.

σ 加法族は $\emptyset$ と Ω を要素に持つ（復習問題）．さらに測度の σ 加法性から $\mathrm{P}[\,\emptyset\,] = 0$ である．なぜなら，測度の定義から $\mathrm{P}[\,A\,] < \infty$ なる $A \in \mathcal{F}$ があるが，$A \cup \emptyset \cup \cdots = A$ で $\emptyset$ は他の集合と共通部分を持たないから $\mathrm{P}[\,\emptyset\,] > 0$ とすると左辺が ∞ で右辺が実数値（有限）となり，矛盾．測度は非負であったから $\mathrm{P}[\,\emptyset\,] = 0$. 他方確率測度の定義から $\mathrm{P}[\,\Omega\,] = 1$ だが，測度の単調性（復習問題）

$$A, B \in \mathcal{F},\ A \subset B,\ \Rightarrow\ \mathrm{P}[\,A\,] \le \mathrm{P}[\,B\,]$$

で $B = \Omega$ とすることにより，確率測度 P の値域が $[0, 1]$ に含まれることが分かる.

[8] 確率論では $\max\{X, Y\} = X \vee Y$, $\min\{X, Y\} = X \wedge Y$ と書くことが多い.

注. $\mathrm{P}[\,A\,]=0$ だからといって $A=\emptyset$ とは限らない. ◇

簡単な性質の中で，特に，**測度（確率測度）の連続性**が極限の考察において基本となる.

命題 1 *(i)* $A_1\subset A_2\subset A_3\subset\cdots$, $A_n\in\mathcal{F}$, $n\in\mathbf{N}$, （増大事象列）ならば $\lim_{n\to\infty}\mathrm{P}[\,A_n\,]=\mathrm{P}[\,\bigcup_{n=1}^{\infty}A_n\,]$.

(ii) $A_1\supset A_2\supset A_3\supset\cdots$, $A_n\in\mathcal{F}$, $n\in\mathbf{N}$, （減少事象列）ならば $\lim_{n\to\infty}\mathrm{P}[\,A_n\,]=\mathrm{P}[\,\bigcap_{n=1}^{\infty}A_n\,]$.

(iii) $\{A_n\}$ が減少事象列でさらに，$\bigcap_{n=1}^{\infty}A_n=\emptyset$ ならば $\lim_{n\to\infty}\mathrm{P}[\,A_n\,]=0$.

証明. (i) $B_1=A_1$, $B_n=A_n\setminus A_{n-1}$, $n=2,3,4,\cdots$, とおくと，$\{B_n\}$ は**排反事象**（互いに共通部分を持たない $\mathcal{F}$ の要素たち）で，$A_n=\bigcup_{k=1}^{n}B_k$, $n\in\mathbf{N}$，かつ，$\bigcup_{k=1}^{\infty}A_k=\bigcup_{k=1}^{\infty}B_k$. （部屋割りの図を書くとわかりやすい.）加法性と σ 加法性から，

$$\lim_{n\to\infty}\mathrm{P}[\,A_n\,]=\lim_{n\to\infty}\sum_{k=1}^{n}\mathrm{P}[\,B_k\,]=\sum_{k=1}^{\infty}\mathrm{P}[\,B_k\,]=\mathrm{P}[\,\bigcup_{k=1}^{\infty}B_k\,]=\mathrm{P}[\,\bigcup_{k=1}^{\infty}A_k\,].$$

(ii) $\{A_n^c\}$ に上記を適用し，$\mathrm{P}[\,A\,]+\mathrm{P}[\,A^c\,]=\mathrm{P}[\,\Omega\,]=1$ を用いる.
(iii) これは上の特別な場合である.

□

注. (i) 最後の性質は非負性と有限加法性と合わせると，確率測度の定義の σ 加法性を置き換えることができる（σ 加法性を導ける）という意味で確率測度の連続性の本質である [9, §2.2 定理 1].
(ii) 最後の性質からその前を導くには，$B=\bigcap_{n=1}^{\infty}A_n$ として $\bigcap_{n=1}^{\infty}(A_n\setminus B)=\emptyset$ を示す.

◇

集合 A, B に対して $A\triangle B=(A\setminus B)\cup(B\setminus A)$ と書く．$A\triangle B$ **は** A **と** B **の違い**を表す集合である.

$$A\triangle B=A^c\triangle B^c \tag{1}$$

は容易に分かる.

$$\left(\bigcup_{n=1}^{N}A_n\right)\triangle C\subset\bigcup_{n=1}^{N}(A_n\triangle B_n)\cup\left(\left(\bigcup_{n=1}^{N}B_n\right)\triangle C\right) \tag{2}$$

もやればできるので練習問題としておく.

定理 2 (近似定理) $\mathcal{A}\subset\mathcal{F}$ が有限加法族のとき $E\in\sigma[\mathcal{A}]$ に対して $\lim_{n\to\infty}\mathrm{P}[\,E\triangle A_n\,]=0$ を満たす集合列 $A_n\in\mathcal{A}$, $n\in\mathbf{N}$, がある.

証明. 定理の結果を満たす $E\in\mathcal{F}$ の集合（集合族）を $\mathcal{C}$ とおく．$A\in\mathcal{A}$ ならば $A_n=A$ とすればよいので $\mathcal{A}\subset\mathcal{C}$.

$C\in\mathcal{C}$ とし，近似列を $A_n\in\mathcal{A}$, $n\in\mathbf{N}$, とする．(1) より $\lim_{n\to\infty}\mathrm{P}[\,C^c\triangle A_n^c\,]=0$ となるから，$C^c\in\mathcal{C}$.

$C_n \in \mathcal{C}$, $n \in \mathbf{N}$, $C = \bigcup_{n\in\mathbf{N}} C_n$, とする．確率の連続性より任意の $\epsilon > 0$ に対して N_0 がとれて，$N \ge N_0$ ならば $\mathrm{P}[\, C \setminus \bigcup_{n=1}^{N} C_n \,] < \epsilon/2$ となる．各 n に対して $C_n \in \mathcal{C}$ だから $\mathrm{P}[\, C_n \triangle A_n \,] < \epsilon 2^{-n-1}$ となる $A_n \in \mathcal{A}$ がある．$\mathcal{A}$ は有限加法族なので $\bigcup_{n=1}^{N} A_n \in \mathcal{A}$ に注意．しかも (2) より，

$$\mathrm{P}[\, C \triangle \left(\bigcup_{n=1}^{N} A_n\right) \,] \le \sum_{n=1}^{N} \mathrm{P}[\, A_n \triangle C_n \,] + \mathrm{P}[\, C \setminus \bigcup_{n=1}^{N} C_n \,] < \epsilon.$$

$\epsilon > 0$ は任意だったから，$C \in \mathcal{C}$ を得る．よって，$\mathcal{C}$ は $\mathcal{A}$ を含む σ 加法族であることが分かるので，$\sigma[\mathcal{A}]$ の最小性から $\sigma[\mathcal{A}] \subset \mathcal{C}$．これは $\sigma[\mathcal{A}]$ の各要素（集合）が定理の結果を満たすことを意味している． □

測度の持つその他の簡単な性質は復習問題とする．

1.2.2 確率変数（可測関数）．

問 1.　$f :\ \Omega \to \mathbf{R}\{\pm\infty\}$ が実可測関数となる必要十分条件として $f^{-1}((a,\infty]) \in \mathcal{F}$, $a \in \mathbf{R}$, をとれることを証明せよ．

（略証：必要性は自明．$\mathcal{G} = \{A \in \mathcal{B}_1' \mid f^{-1}(A) \in \mathcal{F}\}$ とおくと仮定から，区間と $\{\infty\}$, $\{-\infty\}$ は $\mathcal{G}$ の要素であり，かつ，$\mathcal{G}$ は σ 加法族である．従って $\mathcal{G} \supset \mathcal{B}_1'$，即ち，$\mathcal{B}_1'$ の要素は全て $\mathcal{G}$ の性質を持つ.）

問 2.

(i) $(\mathbf{R}^d, \mathcal{B}_d)$ 上の実数値連続関数は実可測関数であることを証明せよ.（略証．連続関数なので，開集合の逆像は開集合だから 問 1 より実可測関数である.）

(ii) 上記を一般の位相空間 Ω 上の実数値連続関数に一般化せよ.（定義域の σ 加法族として何をとればよいか？それを一般に位相空間の**ボレル集合族**という.）

問 3.　$(\Omega_i, \mathcal{F}_i)$, $i = 1,2,3$, と可測関数 $f :\ \Omega_1 \to \Omega_2$, $g :\ \Omega_2 \to \Omega_3$ があるとき，合成関数 $h = g \circ f :\ \Omega_1 \to \Omega_3$ は $\mathcal{F}_1/\mathcal{F}_3$–可測であることを証明せよ．

問 4.　$f :\ \Omega \to \mathbf{R}^2$ が可測空間 $(\Omega, \mathcal{F})$ 上の 2 次元実可測関数であることと，その成分 $f = (f_1, f_2)$, $f_k :\ \Omega \to \mathbf{R}$, $k = 1,2$, が全て実可測関数であることは同値か？

（略解．同値．$\pi_i(x_1, x_2) = x_i$ で定義される射影関数 $\pi_i :\ \mathbf{R}^2 \to \mathbf{R}$ は連続関数だから，$f_i = \pi_i \circ f$ と 問 2, 問 3 より可測．逆に f_i が可測ならば,

$$\mathcal{G} = \{A \in \mathcal{B}_2 \mid f^{-1}(A) \in \mathcal{F}\} \supset \mathcal{B}_2$$

を示せばよい．そのためには $\mathcal{G}$ が σ 加法族であって，$\mathbf{R}^2$ の長方形を全て含めば十分．前者は 問 1 と同様．後者は

$$f^{-1}([a_1, b_1] \times [a_2, b_2]) = \bigcap_{i=1}^{2} f_i^{-1}([a_i, b_i]) \in \mathcal{F}$$

より成立.）

問 5.　$f_k :\ \Omega \to \mathbf{R}$, $k = 1,2$, が可測空間 $(\Omega, \mathcal{F})$ 上の実可測関数で，$g :\ \mathbf{R}^2 \to \mathbf{R}$ が $(\mathbf{R}^d, \mathcal{B}_d)$ 上の 2 次元実可測関数のとき，$h(\omega) = g(f_1(\omega), f_2(\omega))$ で定義される $h :\ \Omega \to \mathbf{R}$ も実可測関数となることを証明せよ．

（略証．問 4 より 問 3 に帰着する.）

問 5 を $X_i = f_i$, $i = 1,2$, と $g(x,y) = x - y$ に用いれば,

$X_i\ \Omega \to \mathbf{R}$ が実確率変数ならば $X_1 - X_2$ も実確率変数となることが分かる．同様に四則演算は全て関数の可測性を保つ．特に，$\{X_1 > X_2\} = \{\omega \in \Omega \mid X_1(\omega) > X_2(\omega)\} \in \mathcal{F}$.

問 6.　$\{X_1 > X_2\} \in \mathcal{F}$ を直接証明するのに，

$$\{X_1 > X_2\} = \bigcup_{c \in \mathbf{R}} \{X_1 > c\} \cap \{c > X_2\} \in \mathcal{F} \quad (?)$$

としてはいけない理由を指摘し，この式を修正してこの方針の証明を完成せよ．

問 1 の直接適用により，次を得る．

命題 3 $\{X_k\}$ が実確率変数列ならば $\inf_{k \in \mathbf{N}} X_k$, $\sup_{k \in \mathbf{N}} X_k$, も実確率変数，よってさらに，

$$\varliminf_{k \to \infty} X_k = \sup_{n \in \mathbf{N}} \inf_{k \geq n} X_k\,, \quad \varlimsup_{k \to \infty} X_k = \inf_{n \in \mathbf{N}} \sup_{k \geq n} X_k$$

も実確率変数．従って

$$\{X_k\text{が収束する}\ \} = \{-\infty < \varlimsup_{k \to \infty} X_k = \varliminf_{k \to \infty} X_k < \infty\} \in \mathcal{F},$$

即ち，極限が存在する $\omega \in \Omega$ の全体は可測集合であって，その上で $\lim_{k \to \infty} X_k = \varlimsup_{k \to \infty} X_k$ は実確率変数.

1.2.3　期待値（積分）．

定義の周辺．　一般の非負確率変数の期待値（非負可測関数の積分）を定義するときに，増加階段関数近似列の積分で定義した．ここで近似列のとり方によらないことに注意する．

命題 4 非負確率変数 X の期待値 $\mathrm{E}[\ X\ ]$ は定義の増大単関数近似列のとり方によらない．即ち，$\tilde{X}_1 \leq \tilde{X}_2 \leq \cdots$, が単関数列で $\lim_{n \to \infty} \tilde{X}_n = X$ *(*各点収束*)* ならば，$\lim_{n \to \infty} \mathrm{E}[\ \tilde{X}_n\ ] = \mathrm{E}[\ X\ ]$

証明. $X_n \leq X$ および $\tilde{X}_n \to X$ (いずれも各点で) だから，$\epsilon > 0$ ならば $A_{mn} = \{\tilde{X_m} > X_n - \epsilon\} \subset \Omega$ は $m \to \infty$ で増大しながら Ω に近づく：

$$A_{1n} \subset A_{2n} \subset \cdots, \quad \bigcup_{m=1}^{\infty} A_{mn} = \Omega.$$

よって確率の連続性から

$$\lim_{m \to \infty} \mathrm{P}[\ A_{mn}\ ] = 1\,.$$

故に

$$\exists m_0;\ (\forall m \geq m_0)\ \mathrm{P}[\ A_{mn}^c\ ](\max_{\omega \in \Omega} X_n(\omega) + \epsilon) < \epsilon\,.$$

これより

$$\tilde{X}_m(\omega) \geq \chi_{A_{mn}}(\omega)\,(X_n(\omega) - \epsilon) \geq X_n(\omega) - \epsilon - (\max_{\omega \in \Omega} X_n(\omega) + \epsilon)\chi_{A_{mn}^c}\,.$$

$\tilde{X}_n$ が増大列であることと合わせて

$$\lim_{m \to \infty} \mathrm{E}[\ \tilde{X}_m\ ] \geq \mathrm{E}[\ X_n\ ] - 2\epsilon.$$

左辺は変数 n, ϵ を含まないから

$$\lim_{m\to\infty} \mathrm{E}[\,\tilde{X}_m\,] \geq \lim_{n\to\infty} \mathrm{E}[\,X_n\,].$$

逆向きの不等号も証明できるので等号が成り立つ. □

命題 5 X, Y が可積分関数のとき以下が成り立つ.

非負性： $X = 0$, *a.e.,* ならば $\mathrm{E}[\,X\,] = 0$. $X = Y$, *a.e.,* ならば $\mathrm{E}[\,X\,] = \mathrm{E}[\,Y\,]$.

単調性： $X \geq 0$, *a.e.,* ならば $\mathrm{E}[\,X\,] \geq 0$. $X \geq Y$, *a.e.,* ならば $\mathrm{E}[\,X\,] \geq \mathrm{E}[\,Y\,]$. $|\mathrm{E}[\,X\,]| \leq \mathrm{E}[\,|X|\,]$.

線形性： $a, b \in \mathbf{R}$ に対して $\mathrm{E}[\,aX + bY\,] = a\mathrm{E}[\,X\,] + b\mathrm{E}[\,Y\,]$.

証明. **非負性：** 測度 0 の集合からの単関数の積分への寄与は 0 だから[9] .

単調性： 非負可測関数に対する標準的な近似列（2^{-n} 刻み，n 上限）を考えれば，$0 \leq X \leq Y$ ならば $0 \leq X_n \leq Y_n$ だから，$0 \leq \mathrm{E}[\,X_n\,] \leq \mathrm{E}[\,Y_n\,]$. 非負とは限らない場合，符号に応じて積分範囲を分割して考えればよい. 最後の性質は $\pm X \leq |X|$ から.

線形性： 可測関数の線形結合が可測関数であることは既出だから積分を考えてよい. 定数倍は定数倍で単関数近似できるので， $a = b = 1$ としてよい.

非負可測関数の和はそれぞれの単関数近似の和が単関数近似となる. 単関数における積分の線形性は明らかだから，非負可測関数における積分の線形性を得る.

非負とは限らない場合は，X, Y, $X + Y$ の符号で場合分けする. 例えば $X > 0$, $Y < 0$, $X + Y > 0$ の場合，$\mathrm{E}[\,X + Y\,] + \mathrm{E}[\,-Y\,] = \mathrm{E}[\,X\,]$ を示せば， $\mathrm{E}[\,-Y\,] = -\mathrm{E}[\,Y\,]$ は定義だから $\mathrm{E}[\,X + Y\,] = \mathrm{E}[\,X\,] + \mathrm{E}[\,Y\,]$ を得る. つまり，適当に移行して 3 項とも正になるようにして証明すれば非負可測関数の場合に帰着する. 今の例では，$X + Y$ と $-Y$ の単関数近似 $(X + Y)_n$, $(-Y)_n$ を考えれば $(X + Y)_n + (-Y)_n$ は X の単関数近似となる.

□

注. $\pm\infty$ が混在せず一方だけならば確率変数の値としてこれらを許す. ◇

収束定理．

関数列の積分の根本は**単調収束定理**である. 非負関数を**加えていく**正項級数型の極限と積分は無条件に交換できる. 級数が各点で絶対収束していれば正負の部分に分けることで交換できるが，問題は級数の形でない場合（級数の形で絶対収束が見えない場合）. このために多くの定理があるが**優収束定理**が決定版.
その準備として非常に多用される **Fatou の補題**があるが，この補題自体は仮定を満たす例の中に trivial な評価しか得られないものがある，という意味で「不完全な」補題である.

[9] $X = Y$, a.e., のとき，完備測度ならば X の可積分性のみ仮定すれば Y の可測性が自動的. ここでは Borel 測度を主に扱うので，Y の可測性も仮定に必要.

定理 6 X_n を増大実数値可測関数列で $X = \lim_{n\to\infty} X_n$, *a.e.,* も実数値関数とする．このとき $\lim_{n\to\infty} \mathrm{E}[\,X_n - X_1\,] = \mathrm{E}[\,X - X_1\,]$．特に，$\mathrm{E}[\,X_1\,] > -\infty$ ならば $\lim_{n\to\infty} \mathrm{E}[\,X_n\,] = \mathrm{E}[\,X\,]$.

注. $\mathrm{E}[\,X_1\,] > -\infty$ から $X_1 > -\infty$, a.e., なので $(X_n - X_1) + X_1$ が $\infty - \infty$ 型になることはない．◇

証明. $Y_n = X_{n+1} - X_n$, $n = 1, 2, 3, \cdots$, は非負可測関数なので，各 Y_n の非負増大単関数近似列を Y_{nm}, $m = 1, 2, 3, \cdots$, とし，$Z_{n1} = Y_{n1}$ および $Z_{nm} = Y_{n,m} - Y_{n,m-1}$, $m = 2, 3, \cdots$, とおくと，Z_{nm} は非負単関数.

積分が増大単関数近似列によらないことと積分の線形性から

$$\mathrm{E}[\,Y_n\,] = \lim_{m\to\infty} \mathrm{E}[\,Y_{nm}\,] = \lim_{m\to\infty} \sum_{k=1}^{m} \mathrm{E}[\,Z_{nk}\,] = \sum_{k=1}^{\infty} \mathrm{E}[\,Z_{nk}\,].$$

よって

$$\mathrm{E}[\,X_n - X_1\,] = \sum_{m=1}^{n-1} \sum_{k=1}^{\infty} \mathrm{E}[\,Z_{mk}\,]. \tag{3}$$

他方，非負単関数 $\{Z_{nm} \mid n, m \in \mathbf{N}\}$ を一列に並べ直して $\{\tilde{Z}_N \mid N \in \mathbf{N}\}$ と書くと，正項級数は和の順序の交換が任意に行えるから，Ω の各点で

$$X - X_1 = \lim_{n\to\infty} X_n - X_1 = \sum_{m=1}^{\infty} \sum_{k=1}^{\infty} Z_{mk} = \sum_{N=1}^{\infty} \tilde{Z}_N\,.$$

右辺第 2 項の級数の部分和の数列は $X - X_1$ の非負増大単関数近似列だから,

$$\mathrm{E}[\,X - X_1\,] = \sum_{N=1}^{\infty} \mathrm{E}[\,\tilde{Z}_N\,] = \sum_{m=1}^{\infty} \sum_{k=1}^{\infty} \mathrm{E}[\,Z_{mk}\,].$$

これを (3) と見比べれば定理の前半を得る.

両辺非負なので $\mathrm{E}[\,X_1\,] > -\infty$ ならばこれを両辺に加えて定理の後半を得る[10]．□

増加列でない場合（級数ならば正負混在の場合）の極限と積分の順序交換定理の重要な十分条件に定理 23 があるが，その十分条件になっている**優収束定理**が直感的で扱いやすいため，最初に習う.（[3, §13 定理 13.6]，[9, §4.4 定理 4] では Lebesgue の収束定理という名前.）その準備として，証明が直感的に分かりやすい Fatou の補題が教科書では必ず紹介される.

補題 7 実数値可測関数列 Y_n, $n \in \mathbf{N}$, について $\mathrm{E}[\,\inf_{n\in\mathbf{N}} Y_n\,]$ が $\mathbf{R} \cup \{+\infty\}$ に存在するならば（たとえば非負値関数列ならば）$\mathrm{E}[\,\varliminf_{n\to\infty} Y_n\,] \le \varliminf_{n\to\infty} \mathrm{E}[\,Y_n\,]$.

注. $\{Y_m^+ = 0\}$ 上で $-Y_m^- = Y_m \ge \inf Y_n$ だから，$\mathrm{E}[\,\inf_{n\in\mathbf{N}} Y_n\,] > -\infty$ から

$$\mathrm{E}[\,-Y_m^-\,] = \mathrm{E}[\,-Y_m^-;\ Y_m^+ = 0\,] \ge \mathrm{E}[\,\inf Y_n;\ Y_n^+ = 0\,] \ge \mathrm{E}[\,-(\inf Y_n)^-\,] > -\infty$$

なので，$\mathrm{E}[\,Y_n\,] \in \mathbf{R} \cup \{+\infty\}$, $n \in \mathbf{N}$.　◇

[10] [9, §4.1 (E9)] の証明は，Z_{nm} の並べ替えをせずに Y_{nm} から各点収束部分列を取り出すために，Borel–Cantelli の定理を使う．ここでは，非負増大単関数近似列にとことん戻って Z_{nm} の並べ替えから近似列変更を行う [3, §13 定理 13.2] の証明を採用した．深く考えると並べ替えのところで coupling argument 的な高級さがあるが,「非負増大関数列の積分と極限の交換は無条件」というルベーグ積分の極意が見える，初等的な証明なのでこのほうが良い証明だと思う.

なお，$X_1 \ge 0$, a.e., が通常の仮定だが，哲弥が仮定を緩めた．ルベーグ積分と極限が交換できないのは積分値が $\infty - \infty$ の問題を起こす場合だけである，という点が明確な仮定のほうがいいと思う．Fatou の lemma や優収束定理において仮定が必要な理由も同様である.

証明. $X_n = \inf_{k\geq n} Y_k,\ n \in \mathbf{N}$, とおくと，増大可測関数列だから $\varliminf_{n\to\infty} Y_n = \lim_{n\to\infty} X_n$ と定理 6 と $X_n \leq Y_n$ より，

$$\mathrm{E}[\ \varliminf_{n\to\infty} Y_n\ ] = \mathrm{E}[\ \lim_{n\to\infty} X_n\ ] = \lim_{n\to\infty} \mathrm{E}[\ X_n\ ] \leq \varliminf_{n\to\infty} \mathrm{E}[\ Y_n\ ].$$

□

注. $\mathrm{E}[\ \inf_{n\in\mathbf{N}} Y_n\ ] > -\infty$ という仮定は哲弥による．関数列が各点収束するとき，仮定成立時の結論の不等号が生じるには $\mathrm{E}[\ \sup_{n\in\mathbf{N}} Y_n\ ] = \infty$ が必要なことが優収束定理から分かる．積分値の $\infty - \infty$ だけが極限と積分の順序交換を妨害する，という経験則の現れ． ◇

定理 8 (優収束定理) X_n を実数値可測関数列で $\exists \lim_{n\to\infty} X_n,\ a.e.$, であってそれも実数値関数とする．$|X_n| \leq Y,\ a.e.$, なる可積分関数 Y があれば，$\lim_{n\to\infty} \mathrm{E}[\ X_n\ ] = \mathrm{E}[\ \lim_{n\to\infty} X_n\ ]$.

証明. $Y_n = Y \pm X_n \geq 0$ に 補題 7 を適用して可積分性 $\mathrm{E}[\ Y\ ] < \infty$ と X_n の各点収束を用いると

$$\pm\mathrm{E}[\ \lim_{n\to\infty} X_n\ ] \leq \varliminf_{n\to\infty} \mathrm{E}[\ \pm X_n\ ].$$

よって

$$\varlimsup_{n\to\infty} \mathrm{E}[\ X_n\ ] \leq \mathrm{E}[\ \lim_{n\to\infty} X_n\ ] \leq \varliminf_{n\to\infty} \mathrm{E}[\ X_n\ ].$$

□

注. 各点（概収束）収束するのに期待値が収束しない例： $X_n = n\chi_{(0,1/n)}$. ◇

1.2.4 分布．

分布． d 次元 Borel 空間 $(\mathbf{R}^d, \mathcal{B}_d)$ 上の確率測度を**分布**という．$[0,1]$ など $\mathbf{R}^d$ の部分空間の場合も分布という．

基本は $\Omega = [0,1]$ 上のルベーグ測度 $([0,1], \mathcal{B}_1([0,1]), \mu_1)$．[11]．

ここで $A \in \mathcal{B}_1$ に対して $\mathcal{B}_1(A) = \{F \cap A \mid F \in \mathcal{B}_1\}$ は A 上の Borel 集合族（A に相対的な開集合を全て要素に持つ最小の σ 加法族）．

$[0,1]$ 上のルベーグ測度が確率測度であることは（測度になっていることを既知とすれば）明らか．d 次元に一般化できることも明らか．$\Omega = [0,1]^d$ 上の d 次元ルベーグ測度も確率測度である．

また，一般に $\Omega = [a,b]$ $(a < b)$ の上のルベーグ測度 μ_1 に対して $\mathrm{P}[\ A\ ] = \dfrac{1}{b-a}\mu_1(A),\ A \in \mathcal{B}_1$, で P を定義すると $([a,b], \mathcal{B}_1([a,b]))$ 上の確率測度になることも明らか．

以上を総称して（例えば最後の例では $[a,b]$ 上の）**一様分布** という．

一様でない分布も容易に作れる．$\mathbf{R}$ 上の非負実（ルベーグ）可測関数 f が積分可能で $Z = \int_{\mathbf{R}} f(x)\,d\mu_1(x) > 0$ のとき，

$$\mathrm{P}[\ A\ ] = \frac{1}{Z}\int_A f(x)\,d\mu_1(x) = \frac{1}{Z}\int f(x)\,\chi_A(x)\,dx \tag{4}$$

で定義される集合関数 P は $(\mathbf{R}, \mathcal{B}_1)$ 上の確率測度になる．

[11] Ω は $(0,1)$ や $[0,1)$, $(0,1]$ などでも測度空間としてはルベーグ測度としては何も違いがない（測度 0 の集合は事実上無視される）ので，問題に応じて他の便宜から使い分ける．

実用上の分布の多くがこのようにして定義されている．詳しく言うと (4) によって定義される分布（確率測度）の全体は**ルベーグ測度に対して絶対連続な分布**の全体に一致する（後者は $\mu_1(A)=0$ ならば必ず $\mathrm{P}[\,A\,]=0$ となる分布のこと）．この事実は測度論の Radon–Nikodym の定理の直接の結果である．絶対連続な分布 P に対応する (4) の f はルベーグ測度 0 の集合を除いて一意的に決まる．これを**密度関数**という．一つの点の測度が全て 0 である分布を**連続な分布**という．連続だがルベーグ測度に対して絶対連続でない分布を**特異分布**という．ルベーグ測度に対して特異な分布は，存在する．

例：

正規分布： $N(m,v)$，但し m は実数定数，v は正定数．$(\mathbf{R},\mathcal{B}_1,\mathrm{P})$;

$$\mathrm{P}[\,A\,]=\frac{1}{\sqrt{2\pi v}}\int_A e^{-\frac{1}{2v}(x-m)^2}dx.$$

コーシー分布： $(\mathbf{R},\mathcal{B}_1,\mathrm{P})$; m は実数定数，a は正定数.

$$\mathrm{P}[\,A\,]=\frac{a}{\pi}\int_A \frac{1}{a^2+(x-m)^2}dx.$$

d 次元正規分布： $N(m,v)$，但し m は d 次元実定縦ベクトル，v は $d\times d$ 対称行列で，固有値が全て正のもの．$(\mathbf{R}^d,\mathcal{B}_d,\mathrm{P})$;

$$\mathrm{P}[\,A\,]=\frac{1}{\sqrt{(2\pi)^d\det v}}\int_A e^{-\frac{1}{2}{}^t(x-m)v^{-1}(x-m)}dx.$$

（d 次元積分であることに注意.）

一般に有限測度 μ（全空間の測度 $\mu(\Omega)$ が実数値＝無限でない）に対して $\mathrm{P}[\,A\,]=\frac{1}{\mu(\Omega)}\mu(A)$ とおくことで確率測度が μ が定義されているのと同じ可測空間の上に定義できる．有限測度を定数倍して確率測度にするときの定数を**規格化定数**という.

平均，分散，モーメント． $(\mathbf{R},\mathcal{B}_1,\mathrm{P})$ を分布とする．$m=\int_{\mathbf{R}} x\mathrm{P}[\,dx\,]$, $v=\int_{\mathbf{R}}(x-m)^2\mathrm{P}[\,dx\,]$, をそれぞれ分布の**平均**および**分散**という．また，$n\in\mathbf{Z}_+$ に対して $M_n=\int_{\mathbf{R}} x^n\mathrm{P}[\,dx\,]$ を分布の n **次モーメント**という.

確率変数の分布． 実確率変数 X は $A\in\mathcal{B}_1$ に対して

$$\mathrm{P}[\,X\in A\,]=\mathrm{P}[\,X^{-1}(A)\,]=\mathrm{P}[\,\{\omega\in\Omega\mid X(\omega)\in A\}\,]$$

が決まる.

> $\mathrm{P}[\,X\in A\,]$ を X が A に入る**確率**という.

命題 9 μ: $\mathcal{B}_1\to\mathbf{R}$ を

$$\mu(A)=\mathrm{P}[\,X\in A\,]=\mathrm{P}[\,X^{-1}(A)\,],\quad A\in\mathcal{B}_1\,, \tag{5}$$

で定義すれば，μ は確率測度になる.

証明. 非負，規格化，σ 加法性，全て定義から得る. □

> $(\mathbf{R},\mathcal{B}_1)$ 上の確率測度 $\mathrm{P}\circ X^{-1}$ を X **の分布 (法則)** という.

$X:\ \Omega \to \mathbf{R}$ が実確率変数であれば，元の確率空間 $(\Omega, \mathcal{F}, \mathrm{P})$ が何であっても X によって $(\mathbf{R}, \mathcal{B}_1)$ 上の確率測度に移される．値の分布が問題になる限り，元の空間は見る必要がないことに注意．d 次元実確率変数でも以上は直ちに拡張される．

$X = (X_1, \cdots, X_k)$ のとき， $\mathrm{P} \circ X^{-1}$ を $X_1, \cdots, X_k$ の**結合分布**ともいう．

注. この講義では $\mathbf{R}$ 上の分布しかほとんど扱わない．これは確率変数列の特性関数 (§2.4)，特に，独立確率変数列の特性関数 (§3.2.4) を扱うときに話を「同様に」ですませてしまうことになる．実際，気をつけるべきことはこの講義の範囲では少ないと思うが，将来各自が多次元分布について詳細な検討が必要になったときは，教科書等で勉強しなおして頂きたい． ◇

命題 10 $\mu = \mathrm{P} \circ X^{-1}$ を実確率変数 X の分布，$f:\ \mathbf{R} \to \mathbf{R}$ を実可測関数とするとき $\mathrm{E}[\,f(X)\,] = \int_\Omega f(X) d\mathrm{P}$ が存在するならば

$$\mathrm{E}[\,f(X)\,] = \int_{\mathbf{R}} f d\mu.$$

証明. $A \in \mathcal{B}_1$ に対して $f = \chi_A$ とすると，

$$\int_\Omega \chi_A(X) d\mathrm{P} = \int_\Omega \chi_A(X(\omega)) d\mathrm{P}[\,\omega\,] = \mathrm{P}[\,X \in A\,] = \mu(A) = \int_{\mathbf{R}} \chi_A(x) \mu(dx).$$

この線形結合によって，積分の線形性から，f が階段関数の場合も成り立つ．極限をとれば f が非負可測関数の場合に成り立つ．f の正の部分 $f_+ = f \vee 0$ と負の部分 $f_- = f \wedge 0$ に分ければ，一般の可測関数の場合も成り立つ． □

命題 10 は n 次元実確率変数と $f:\ \mathbf{R}^n \to \mathbf{R}$ に拡張できる．命題 10 から，特に，X の分布 $\mu = \mathrm{P} \circ X^{-1}$ の平均 m と分散 v は

$$m = \int_{\mathbf{R}} x \mu(dx) = \mathrm{E}[\,X\,], \quad v = \int_{\mathbf{R}} x^2 \mu(dx) - m^2 = \mathrm{E}[\,(X - \mathrm{E}[\,X\,])^2\,] = \mathrm{V}[\,X\,],$$

即ち，

X の期待値と分散はそれぞれ X の分布 $P \circ X^{-1}$ の期待値と分散に等しい．

分布関数． μ を $(\mathbf{R}, \mathcal{B}_1)$ 上の測度，X を μ を分布とする実数値確率変数とする．記述の簡単のため $\mathbf{R}$ に $\pm\infty$ は許さないことにする．

$F(x) = \mu((-\infty, x]) = \mathrm{P}[\,X \leq x\,]$, $x \in \mathbf{R}$, で定義された関数 $F:\ \mathbf{R} \to [0,1]$ を X（または μ）の**分布関数**という．

分布関数は分布（$(\mathbf{R}, \mathcal{B}_1)$ 上の確率測度）の性質を調べるのには簡単で強力だが，一般の空間上の確率測度には使えないのが難点．この講義でも Prokhorov の定理 (定理 32 の十分性のほう) で証明に使いたいところだが，この講義では言及する余裕がないので，Prokhorov の定理の証明とともに，紹介だけしておく．

定理 11 $F:\ \mathbf{R} \to [0,1]$ が分布関数ならば以下が成り立つ．

単調性． $x \le y$ ならば $F(x) \le F(y)$.

右連続性． $\lim_{x \downarrow a} F(x) = F(a)$.

値域． $\lim_{x\to\infty} F(x) = 1, \quad \lim_{x\to-\infty} F(x) = 0$.

逆に $F:\ \mathbf{R} \to [0,1]$ が以上を満たすならば，ある確率空間とその上の実数値確率変数をとって，F をその分布関数にできる．しかもそのような確率変数の法則は一意的である．

証明は略す.

2 確率変数列の収束．

2.1 期待値に関する種々の不等式．

以下断らなければ全て実数値確率変数（可測関数）とする．

Tchebichev の不等式： X が期待値を持てば（可積分ならば）$|X|$ が大きくなる確率は小さい.

Jensen の不等式： $f:\ \mathbf{R} \to \mathbf{R}$ が下に凸 (convex) ならば X を f で写すことで X の値の大きいところの影響が強調される.（平均を先に取った方が f で先に移した平均よりも小さい.）

Minkowski の不等式： $p \ge 1$ のとき $\|X\|_p = \mathrm{E}[\,|X|^p\,]^{1/p}$ は三角不等式を満たす（L^p ノルム）.

Hölder の不等式： Schwarz の不等式の一般化．関数の積の期待値を L^p ノルムで上から押さえる.

命題 12 (Tchebichev) $f:\ \mathbf{R} \to \mathbf{R}_+$ が非負可測関数で $x \ge 0$ で増加とする．$x \ge 0$ ならば

$$\mathrm{E}[\,f(X)\,] \ge f(x)\mathrm{P}[\,X \ge x\,].$$

注. 例えば $p > 0$ のとき $f(x) = x^p \vee 0$ とし X を $|X|$ にして用いると $\mathrm{E}[\,|X|^p\,] \ge x^p \mathrm{P}[\,|X| \ge x\,]$, $x > 0$. ◇

証明.

$$\mathrm{E}[\,f(X)\,] = \mathrm{E}[\,f(X);\ X \ge x\,] + \mathrm{E}[\,f(X);\ X < x\,] \ge \mathrm{E}[\,f(X);\ X \ge x\,] \ge \mathrm{E}[\,f(x);\ X \ge x\,]$$
$$\ge f(x)\mathrm{P}[\,X \ge x\,].$$

□

命題 13 (Jensen) X が確率変数で，$f:\ \mathbf{R} \to \mathbf{R}$ が *convex*（下に凸），即ち，

$$0 \le t \le 1 \implies f(tx + (1-t)y) \le tf(x) + (1-t)f(y),\ x, y \in \mathbf{R}, \tag{6}$$

を満たすならば $f(\mathrm{E}[\,X\,]) \le \mathrm{E}[\,f(X)\,]$.

注. 母関数による漸近解析に重要になることがある (§A.4). ◇

証明. 略証：$c = \mathrm{E}[\,X\,]$ とおくとき，$(c, f(c))$ を通る直線 $y = a(x-c)+f(c)$ で $f(x) \ge a(x-c)+f(c)$, $x \in \mathbf{R}$, を満たすものがある．積分の単調性と線形性から $\mathrm{E}[\,f(X)\,] \ge f(c) = f(\mathrm{E}[\,X\,])$. □

命題 14 (Hölder の不等式．) $p^{-1}+q^{-1}=1$, $1<p<\infty$, のとき $\|X\|_p<\infty$, $\|Y\|_q<\infty$ ならば

$$\mathrm{E}[\,|XY|\,] \le \|X\|_p \|Y\|_q\,.$$

ここで $\|X\|_p = \mathrm{E}[\,|X|^p\,]^{1/p}$ は L^p **ノルム**.

$p=q=2$ のとき **Schwarz の不等式**という.

証明. **補題 15** $a, b \ge 0$ のとき $\dfrac{1}{p}a^p + \dfrac{1}{q}b^q \ge ab$.

証明. $ab=0$ なら当然だから $ab \ne 0$ とする．$f(x) = \dfrac{1}{p}x^p + \dfrac{1}{q} - x$, $x \ge 0$, は $x=1$ で最小値 0 をとるので $f(x) \ge 0$, $x \ge 0$．$x = ab^{-q/p}$ とおけば主張を得る. □

$\|X\|_p = 0$ ならば $X=0$, a.e., となるので $XY=0$, a.e., で主張は成立．よって，$0<\|X\|_p<\infty$, $0<\|Y\|_q<\infty$, としてよいが，$a = \dfrac{|X|}{\|X\|_p}$, $b = \dfrac{|Y|}{\|Y\|_q}$, として 補題 15 を使い，両辺の期待値をとると，

$$1 = \frac{1}{p} + \frac{1}{q} \ge \frac{\mathrm{E}[\,|XY|\,]}{\|X\|_p\|Y\|_q}\,.$$

□

命題 16 (Minkowski の不等式) $p \ge 1$ ならば $\|X+Y\|_p \le \|X\|_p + \|Y\|_p\,.$

証明. $p=1$ のときは各点の三角不等式 $|X+Y| \le |X|+|Y|$ の期待値をとればよい．$p>1$ とする．$\|X+Y\|_p = 0$ ならば当然成り立つので $\|X+Y\|>0$ とする.

$$|X+Y|^p \le |X+Y|^{p-1}|X| + |X+Y|^{p-1}|Y|$$

の期待値をとり，右辺に Hölder の不等式を用いると

$$\mathrm{E}[\,|X+Y|^{p-1}|X|\,] \le \mathrm{E}[\,|X+Y|^{(p-1)q}\,]^{1/q}\|X\|_p = \|X+Y\|_p^{p-1}\|X\|_p$$

などから

$$\|X+Y\|_p^p \le \|X+Y\|_p^{p-1}(\|X\|_p + \|Y\|_p).$$

$\|X+Y\|>0$ を仮定したので，主張を得る. □

例．Hölder の不等式から次を得る.

$$\|X\|_p \le \|X\|_{p'}, \quad 0<p<p'. \tag{7}$$

2.2 種々の収束の定義．

確率変数とは（可測）関数のことだから，確率変数列の収束とは関数列の収束のことである．初等的な微積分学では各点収束かせいぜい一様収束しか登場しないが，定義域の空間に確率測度という構造があるので，種々の視点からの収束が考えられる．複数の収束を考える利点は：

(i) 複数の（単純な強弱関係にない）収束が言えれば非常に強力な結果を得るかもしれない，

(ii) 強い結果を得るのが難しいとき，先にそれより弱い収束の結果を得る．それだけで言えることもあるだろうし，いったん収束が言えればより強い収束を言うために必要な追加条件はさほど強くない場合がある．

確率変数列の標準的な収束の概念には以下のものがある．

$$\begin{array}{l} \text{概収束} \searrow \\ L^p\ \text{収束} \nearrow \end{array} \text{確率収束} \;\rightarrow\; \text{法則収束} \tag{8}$$

矢印は無条件に成り立つ強弱（十分条件／必要条件）関係を示す．

各矢印がどの程度の違いであるか（逆向き不成立の反例と成立するケース）はそれぞれ準備を要するので順次後述する．

2.2.1 概収束．

気持ち：　概収束　＝　各点収束 (a.e.)
定義：収束するという事象の確率が 1．

確率空間 $(\Omega, \mathcal{F}, \mathrm{P})$ 上の実確率変数列 X_n, $n = 1, 2 \cdots$, が実数値確率変数（$\pm\infty$ はとらない） Y に**概収束**するとは

$$\mathrm{P}[\lim_{n\to\infty} X_n = Y\,] = 1$$

となることを言う．§1.1.3 の記号法によって

$$X_n \longrightarrow Y,\ a.e.,$$

と書くのが普通である．

注. (i) 命題 3 より，

$$\{\lim_{n\to\infty} X_n = Y\} = \{\omega \in \Omega \mid \lim_{n\to\infty} X_n(\omega) = Y(\omega)\}$$

は事象（可測集合）なので確率が定義できるし，極限が確率変数（可測関数）になることも分かっている．

(ii) 確率変数は Ω 上の実数値（概収束では極限が無限大の確率 0 なので，事実上 $\pm\infty$ の値は無視できる）関数だから，概収束でいう Y への収束は関数の各点収束である．ただし，収束しない ω があってもその合計が確率 0 ならば許す点で各点収束より「わずかに」弱い[12]．

◇

[12] 確率論では確率 0 の事象は無視する（起きない現象である）という「気持ち」だから,「弱い」というより，概収束が自然な定義であって,「本当の」各点収束は確率論では扱うべからざる概念というほうが正しいだろう．

2.2.2 L^p 収束.

定義： L^p 収束 = 差の p 乗の積分が消えること.
意義：関数の集合が Banach 空間 (完備線形空間) になるオーソドックスなノルム収束.

確率変数 X に対して, $\|X\|_p = (\mathrm{E}[\,|X|^p\,])^{1/p}$ と書く (L^p–**ノルム**). $X_n \in L^p, n \in \mathbf{N}$, が $\displaystyle\lim_{n\to\infty} \|X_n - X\|_p = 0$ を満たすとき，X_n は X に p **次平均収束**するという.

注意: $|\mathrm{E}[\,X\,]| \le \mathrm{E}[\,|X|\,]$ だから, L^1 収束 $\displaystyle\lim_{n\to\infty} \mathrm{E}[\,|X_n - X|\,] = 0$ は, 期待値の収束 $\displaystyle\lim_{n\to\infty} \mathrm{E}[\,X_n\,] = \mathrm{E}[\,X\,]$ よりも強い.（即ち， L^1 収束すれば期待値は収束する.）

Banach 空間 （復習）. ルベーグ積分で既知のように，可測関数に対して $f = g$, a.e., のとき $f \sim g$ とおくと $\sim$ は関数の集合上の同値条件になるので， $\sim$ で同値類に類別できる．$\|X\|_p < \infty$ を満たす可測関数の集合を，この同値類の集合と読み直して，さらに，自然な意味で線形空間とみなしたとき，$\|\cdot\|_p$ はノルム（非負 ($\|f\| \ge 0$)，一意 ($\|f\| = 0$ ならば $f = 0$)，一次 ($\|af\| = |a|\|f\|$)，三角不等式 ($\|f+g\| \le \|f\| + \|g\|$)，が成り立つ実数値関数）になる．関数の集合がノルムが定義する距離 $\rho(f,g) = \|f-g\|_p$ (非負，一意，対称，三角不等式，が成り立つ 2 変数関数) に関して完備なときこれを **Banach 空間**という．$\|X\|_p < \infty$ なる関数の集合は $\|\cdot\|_p$ をノルムとして Banach 空間になる.（以上についてはルベーグ積分論を復習のこと.）

2.2.3 確率収束.

任意の $\epsilon > 0$ に対して
$$\lim_{n\to\infty} \mathrm{P}[\,|X_n - Y| \ge \epsilon\,] = 0$$
が成り立つとき， $X_n \to Y$, in prob., 等と書いて，$\{X_n\}$ が Y に**確率収束**するという.

命題 17 *(i)* 概収束すれば確率収束する.
(ii) L^p 収束すれば確率収束する.

証明. (i) 優収束定理と概収束の仮定から
$$\lim_{n\to\infty} \mathrm{P}[\,|X_n - Y| \ge \epsilon\,] = \lim_{n\to\infty} \mathrm{E}[\,\chi_{|X_n - Y} \ge \epsilon\,] = \mathrm{E}[\,\lim_{n\to\infty} \chi_{|X_n - Y} \ge \epsilon\,] = 0$$
が任意の $\epsilon > 0$ に対して成り立つので主張が成り立つ.
(ii) チェビシェフの不等式から
$$\epsilon^p \lim_{n\to\infty} \mathrm{P}[\,|X_n - Y| \ge \epsilon\,] \le \lim_{n\to\infty} \mathrm{E}[\,|X_n - Y|^p\,] = 0, \quad \epsilon > 0\,.$$
□

2.2.4 法則収束（弱収束）.

μ と μ_n, $n = 1, 2, 3, \cdots$, を分布（実数上の Borel 確率測度）とする.

定理 18 以下は同値．これが成り立つとき，μ_n は μ に弱収束する，という．

(i) 有界連続関数 $f:\ \mathbf{R}\to\mathbf{R}$ に対して

$$\lim_{n\to\infty}\int f\,d\mu_n = \int f\,d\mu. \tag{9}$$

(ii) 任意の開集合 G に対して $\varliminf_{n\to\infty}\mu_n(G)\geq\mu(G)$.

(iii) 任意の閉集合 F に対して $\varlimsup_{n\to\infty}\mu_n(F)\leq\mu(F)$.

(iv) $\mu(\partial A)=0$ を満たす[13] 任意のボレル集合 A に対して

$$\lim_{n\to\infty}\mu_n(A)=\mu(A). \tag{10}$$

(v) 台有界の連続関数 $f:\ \mathbf{R}\to\mathbf{R}$ に対して *(9)* が成り立つ．

定理 *18* **の証明**. **(i) から (ii)．** $\mathbf{R}^d$ では任意の開集合は増加閉集合列の極限だから，任意の $\epsilon>0$ に対して，$F\subset G$ を満たす閉集合 F であって，

$$\mu(G)<\mu(F)+\epsilon$$

を満たすものがある．有界連続関数 f を

$$f(x)\begin{cases} =0, & x\in G^c,\\ \in[0,1], & x\in G\cap F^c,\\ =1, & x\in F,\end{cases}$$

となるようにとる[14] と，(9) から

$$\varliminf_{n\to\infty}\mu_n(G)\geq\lim_{n\to\infty}\int f(x)\mu_n(dx)=\int f(x)\mu(dx)\geq\mu(F)\geq\mu(G)-\epsilon.$$

故に $\varliminf_{n\to\infty}\mu_n(G)\geq\mu(G)$.

(ii) と (iii) の同値性． (ii) と (iii) が同値なことは，補集合を考えれば明らか.

(ii) と (iii) から (iv)． $\mu(A^o)\leq\varliminf_{n\to\infty}\mu_n(A^o)\leq\varliminf_{n\to\infty}\mu_n(A)\leq\varlimsup_{n\to\infty}\mu_n(A)\leq\varlimsup_{n\to\infty}\mu_n(\bar{A})\leq\mu(\bar{A})$.

(iv) から (v)． $\epsilon>0$ を任意にとる．

f が台有界な有界連続関数なので，初等解析学の議論によって，有限個の区間 $J_k,\ k=1,\cdots,N$, を選んで階段関数列 $\tilde{f}=\sum_{k=1}^N c_k\chi_{J_k}$ で一様近似できる．即ち，$0\leq f-\tilde{f}\leq\epsilon/2$ (各点) とできる．ここで $\max_k c_k\leq\sup_x|f(x)|+\epsilon/2$. （有限個でいいというのは，compactness argument.）しかも，μ は有限測度なので不連続点は高々可算個だから，連続点が稠密にあるので,（必要なら近似を少しゆるめるだけで）J_k の端点は全て μ の連続点にすることができる（最初に $\epsilon/4$ 近似してから $\epsilon/2$ 近似に戻ればよい）．従って特に，$\mu(\partial J_k)=0,\ k=1,\cdots,N$.

$\mu_n,\ \mu$ が分布（確率）なので，

$$\sum_{k=1}^N c_k\mu_n(E_k)=\int\tilde{f}\,d\mu_n\leq\int f\,d\mu_n\leq\int\tilde{f}\,d\mu_n+\epsilon=\sum_{k=1}^N c_k\mu_n(E_k)+\epsilon.$$

[13] $\partial A=\bar{A}\cap\bar{A^c}$ (A の境界点からなる集合). 極限測度 μ に対してのみ条件が付いていることに注意.「途中」は無関係.

[14] [9, 5章 §4 定理 1] ではここまでを 1 行で済ませているが，$\Omega=\mathbf{R}^n$ の位相的性質に依存する同値条件なので，ていねいに書いてあるほうがいいのではないか？ここの記述は [2] に従った.

および

$$\sum_{k=1}^{N} c_k \mu(E_k) \le \int f\,d\mu \le \sum_{k=1}^{N} c_k \mu(E_k) + \epsilon.$$

(10) から $\lim_{n\to\infty} \mu_n(E_k) = \mu(E_k)$ なので,

$$\begin{aligned}
\int f\,d\mu - \epsilon &\le \sum_{k=1}^{N} c_k \mu(E_k) \\
&\le \varliminf_{n\to\infty} \int f\,d\mu_n \le \varlimsup_{n\to\infty} \int f\,d\mu_n \\
&\le \sum_{k=1}^{N} c_k \mu(E_k) + \epsilon \le \int f\,d\mu + \epsilon.
\end{aligned}$$

これが任意の $\epsilon > 0$, $\delta > 0$ に対して成り立つから (9) を得る.

(v) から (i). （台有界で連続なら，有界な関数だから，(i) から (v) は当然．自明でない逆のほうを示す.）台有界連続関数で近似し，誤差が確率の連続性で消えることを示す.

任意の $m > 0$ に対して,

$$g(x) = f(x),\ |x| \le m, \quad g(x) = 0,\ |x| > m+1,$$

を満たす compact support 連続関数 g が存在する．f が有界なので $c = \sup_x |f(x)-g(x)| < \infty$.

$$h(x) \quad \begin{cases} = c, & |x| \le m-1, \\ \in [0,c], & m-1 < |x| < m, \\ = 0, & |x| \ge m, \end{cases}$$

なる compact support 連続関数 h をとれば $|f(x) - g(x)| \le c - h(x)$, $x \in \mathbf{R}$. このとき,

$$\int |f(x) - g(x)| \mu(dx) \le c\mu([-m,m]^c).$$

また (9) を h に適用すれば,

$$\int |f(x) - g(x)| \mu_n(dx) \le c - \int h(x)\mu_n(dx) \ \longrightarrow\ c - \int h(x)\mu(dx), \quad n \to \infty.$$

よって

$$\varlimsup_{n\to\infty} \int |f(x) - g(x)| \mu_n(dx) \le c\mu([-m+1, m-1]^c).$$

g に (9) を適用すれば

$$\lim_{n\to\infty} \int g(x)\mu_n(dx) = \int g(x)\mu(dx).$$

以上を合わせると

$$\varlimsup_{n\to\infty} \left| \int f(x)\mu_n(dx) - \int f(x)\mu(dx) \right| \le 2c\mu([-m+1, m-1]^c).$$

m は任意で左辺は m によらないから， $m \to \infty$ として確率の連続性を使えば (9) が成り立つことが分かる.

□

注. (i) 弱収束の定義は実数上の確率測度に限らず，もっとはるかに一般の空間上の確率測度でもできるが，以上の同値条件の成立には，下にある空間にいくつかの位相的性質が必要．弱収束の概念が有効なのはだいたい Poland 空間 (可分な完備距離空間) くらいというコンセンサスがあるらしい.

(ii) 分布 ($\mathbf{R}^n$ 上の分布でも可) の場合は以上の他に，分布関数 $\mu((-\infty, x])$ による同値条件もある [9, §5.4 定理 1].
(iii) 特性関数による同値条件は特に重要なので別に扱う (§2.4).

$\diamond$

定義 1 確率変数 $X_n,\ n \in \mathbf{N},\ X,$ の分布を $\mu_n,\ n \in \mathbf{N},\ \mu,$ とするとき， X_n が X に法則収束（弱収束）するとは μ_n が μ に弱収束することを言う.

鍵となる条件式 (9) は

$$\lim_{n\to\infty} \mathrm{E}[\, f(X_n)\,] = \mathrm{E}[\, f(X)\,]$$

と書け，例えば， (10) は

$$\lim_{n\to\infty} \mathrm{P}[\, X_n \in A\,] = \mathrm{P}[\, X \in A\,]$$

と書ける.

命題 19 確率収束すれば法則収束する

証明. $\epsilon > 0$ を任意にとる． $f : \mathbf{R} \to \mathbf{R}$ を台有界連続関数とすると，有界かつ一様連続だから，

$$M = \sup_{x\in\mathbf{R}} |f(x)| < \infty$$

かつ

$$\exists \delta > 0;\ (\forall x, y \in \mathbf{R})\ |x-y| < \delta \Longrightarrow\ |f(x) - f(y)| < \frac{1}{2}\epsilon.$$

X_n が X に確率収束しているから，この δ に対して，

$$\exists n_0;\ n \ge n_0\ \Longrightarrow\ \mathrm{P}[\, |X_n - X| \ge \delta\,] < \frac{1}{4M}\epsilon.$$

よって

$$\begin{aligned} &|\mathrm{E}[\, f(X_n)\,] - \mathrm{E}[\, f(X)\,]| \le \mathrm{E}[\, |f(X_n) - f(X)|\,] \\ &\le \mathrm{E}[\, |f(X_n) - f(X)|;\ |X_n - X| < \delta\,] + 2M\mathrm{P}[\, |X_n - X| \ge \delta\,] \le \epsilon,\quad n \ge n_0. \end{aligned}$$

$\epsilon > 0$ は任意だったから，定理 18 より X_n は X に弱収束する. $\square$

以上によって，種々の収束概念の強弱関係は (8) のようになる．
なお，次のことが知られている．

定理 20 (Skorohod) X_n が X に法則収束するならば，$(0, 1)$ 上のルベーグ確率空間の上の確率変数 Y_n, Y を，X_n と Y_n の法則， X と Y の法則，がそれぞれ等しく，Y_n が Y に概収束する，ようにとることができる．

分布関数を使った証明が [9, 5 章 §5] にあるが，ここでは省略する．

法則収束しても確率収束しない例． 法則収束は確率変数の分布（値の散らばり具合）の収束なので，確率変数たちが同じ Ω 上の関数でなくても考えることができる．逆に言えば，そういう状況では法則収束していても他の収束は考えようがないことになる．

X_n が同じ Ω 上の確率変数の場合でも簡単に例が作れる．例えば $\mathrm{P}[\, X = 0\,] = \mathrm{P}[\, X = 1\,] = 1/2$ とし， $X_n = 1 - X,\ n \in \mathbf{N},$ とおくと，分布は全て等しいから特に法則収束しているが，定義から $|X_n - X| = 1$ が恒等的に成り立つので， $\epsilon < 1$ だと $\mathrm{P}[\, |X_n - X| > \epsilon\,] = 1$ となって確率収束していない．

2.3 概収束と他の収束の関係．

「関係」はいろいろ考えられるが，代表的なものをいくつか．

2.3.1 L^p 収束と概収束の関係．

一方で収束してももう一方で収束しない例． 定理 8 の直後の注のとおり，概収束しても期待値収束しない例 $X_n = n\chi_{(0,1/n)}$ がある．$|\mathrm{E}[\,X\,]| \le \mathrm{E}[\,|X|\,]$ なので，当然 L^1 収束しない例にもなっている．L^p 収束 $(p \ge 1)$ もしていないし，部分列も L^p 収束しない．

L^p 収束しても概収束しない例もある．$n \in \mathbf{Z}_+$, $r = 0, 1, \cdots, 2^n - 1$, に対して $A_{n,r} = (r2^{-n}, (r+1)2^{-n}] \subset (0, 1]$ とおき，$N = 2^n + r$ のとき $X_N = \chi_{A_{n,r}}$ とおくと，X_N は $N \to \infty$ で任意の $1 \le p < \infty$ に対して $X = \chi_\emptyset$ に L^p 収束するが，概収束はしない（$(0, 1]$ 上どの点でも収束しない）．

L^p 収束から概収束について言えること．

命題 21 *$f \in L^p$, $f_n \in L^p$, $n \in \mathbf{N}$, かつ $\lim_{n\to\infty} \|f_n - f\|_p = 0$ ならば，適当な部分列をとって $\lim_{k\to\infty} f_{n(k)}(x) = f(x)$, a.e.-$x$, とできる．*

即ち L^p 収束していれば，概収束する部分列がとれる．

証明. $\{f_n\}$ は L^p の中で収束するから Cauchy 列，即ち，$\lim_{n,m\to\infty} \|f_n - f_m\|_p = 0$．$n(k)$, $k = 1, 2, 3, \cdots$, を単調増加で $\|f_n - f_{n(k)}\|_p < 2^{-k}$, $n > n(k)$, となるようにとれるので，特に，$\|f_{n(k+1)} - f_{n(k)}\|_p < 2^{-k}$, $k = 1, 2, 3, \cdots$. $\tilde{f}_k = f_{n(k)}$, $k = 1, 2, 3, \cdots$, とおく．$g_n = |\tilde{f}_1| + \sum_{j=1}^{n-1} |\tilde{f}_{j+1} - \tilde{f}_j| \in L^p$ とおくと，非負値単調増加で，各点での複素数の三角不等式と $\tilde{f}_n$ のとり方から $\|g_n\|_p \le \|\tilde{f}_1\|_p + 1$．単調収束定理から

$$\|\lim_{n\to\infty} g_n\|_p = \lim_{n\to\infty} \|g_n\|_p \le \|\tilde{f}_1\|_p + 1 < \infty$$

が存在．即ち，$g(x) = \lim_{n\to\infty} g_n(x)$ が a.e.-x で存在して $g \in L^p$．$|g(x)| < \infty$ から $\tilde{f}_1 + \sum_{j=1}^{\infty} (\tilde{f}_{j+1} - \tilde{f}_j) = \lim_{n\to\infty} \tilde{f}_n = \tilde{f}$ は a.e.-x で絶対収束して $|\tilde{f}| \le g$，即ち $\tilde{f} \in L^p$.

$|\tilde{f}(x) - \tilde{f}_n(x)| \le g(x)$ なので，優収束定理 (定理 8) が使えて，$\lim_{n\to\infty} \|\tilde{f} - \tilde{f}_n\|_p = 0$ を得る．仮定と合わせて $\|f - \tilde{f}\|_p = 0$ となるので，$f = \tilde{f}$, a.e.. 従って

$$\lim_{k\to\infty} f_{n(k)} = \lim_{n\to\infty} \tilde{f}_n = \tilde{f} = f, \ a.e..$$

□ [15]

2.3.2 一様可積分性（概収束から L^1 収束へ）．

> X_n, $n \in \mathbf{N}$, が $\lim_{K\to\infty} \sup_{n\in\mathbf{N}} \mathrm{E}[\,|X_n|;\ |X_n| \ge K\,] = 0$ を満たすとき**一様可積分**であるという．
> （既に例示したように）概収束だけでは L^p 収束について何も言えないが，一様可積分性があれば L^1 収束する．

一様可積分の同値条件.

命題 22 *$\{X_n\}$ が一様可積分であることと次の 2 条件の同時成立は同値.*

[15] [3, §22 定理 22.2] の証明.

(i) $\sup_{n\in\mathbf{N}} \mathrm{E}[\,|X_n|\,] < \infty$.
(ii) 任意の $\epsilon > 0$ に対して $\delta > 0$ がとれて,

$$\mathrm{P}[\,F_n\,] < \delta,\ F_n \in \mathcal{F},\ n = 1, 2, \cdots \implies \sup_{n\in\mathbf{N}} \mathrm{E}[\,|X_n|;\ F_n\,] < \epsilon.$$

証明. （必要性.）一様可積分ならばある K で $\sup_{n\in\mathbf{N}} \mathrm{E}[\,|X_n|;\ |X_n| \geq K\,] \leq 1$. よって,

$$\mathrm{E}[\,|X_n|\,] = \mathrm{E}[\,|X_n;\ |X_n| \geq K\,] + \mathrm{E}[\,|X_n;\ |X_n| < K\,] \leq 1 + K\mathrm{P}[\,|X_n| < K\,] \leq 1 + K < \infty.$$

また, 任意の $\epsilon > 0$ に対して $K > 0$ が存在して $\sup_{n\in\mathbf{N}} \mathrm{E}[\,|X_n|;\ |X_n| \geq K\,] < \epsilon/2$. $\delta = \epsilon/(2K)$ とおくと, $\mathrm{P}[\,F_n\,] < \delta$ ならば,

$$\mathrm{E}[\,|X_n|;\ F_n\,] \leq \mathrm{E}[\,|X_n|;\ \{|X_n| < K\} \cap F_n\,] + \mathrm{E}[\,|X_n|;\ |X_n| \geq K\,] \leq K\mathrm{P}[\,F_n\,] + \epsilon/2 < \epsilon.$$

（十分性.）$\epsilon > 0$ を任意にとり, 第２仮定で対応する $\delta > 0$ をとる. $M = \sup_{n\in\mathbf{N}} \mathrm{E}[\,|X_n|\,]$ とおき, $M/K < \delta$ を満たすように K を選ぶと, チェビシェフの不等式から $\mathrm{P}[\,|X_n| \geq K\,] \leq K^{-1}\mathrm{E}[\,|X_n|\,] < \delta$ なので, 第２仮定から $\sup_{n\in\mathbf{N}} \mathrm{E}[\,|X_n|;\ |X_n| \geq K\,] < \epsilon$. □

概収束確率変数列における一様可積分の特徴付け.

定理 23 可積分な確率変数列 X_n, $n = 1, 2, \cdots$, が確率変数 X に概収束しているとき次の３条件は同値である.

(i) X_n, $n \in \mathbf{N}$, は一様可積分.
(ii) $\lim_{n\to\infty} \mathrm{E}[\,|X_n - X|\,] = 0$.
(iii) $\lim_{n\to\infty} \mathrm{E}[\,|X_n|\,] = \mathrm{E}[\,|X|\,] < \infty$.

注. (i) もちろん, 第２条件が成り立つとき, $\mathrm{E}[\,X_n\,]$ は $\mathrm{E}[\,X\,]$ に収束. 定理 8 も参照のこと. しかし, 概収束してかつ期待値が収束しても一様可積分とは言えない.
(ii) 第３条件の $\mathrm{E}[\,|X|\,] < \infty$ は, もちろん, X も可積分, という意味.

◇

証明. (i) から (ii). Fatou の補題から $\mathrm{E}[\,|X|\,] \leq \varliminf_{n\to\infty} \mathrm{E}[\,|X_n|\,] < \infty$.

先ず $\{X_n - X\}$ も一様可積分であることを示す.

$$|X_n - X| \geq K \ \Rightarrow\ |X| \geq \frac{K}{2} \text{ or } |X_n| \geq \frac{K}{2} > |X|,$$

を用いれば,

$$\begin{aligned}
&\sup_{n\in\mathbf{N}} \mathrm{E}[\,|X_n - X|;\ |X_n - X| \geq K\,] \\
&\leq \sup_{n\in\mathbf{N}} \mathrm{E}[\,|X_n|;\ |X_n| \geq K/2\,] + \sup_{n\in\mathbf{N}} \mathrm{E}[\,|X_n|;\ |X| \geq K/2\,] \\
&\ + \sup_{n\in\mathbf{N}} \mathrm{E}[\,|X|;\ |X| \geq K/2\,] + \sup_{n\in\mathbf{N}} \mathrm{E}[\,|X|;\ |X_n| \geq K/2 > |X|\,]
\end{aligned}$$

となり, $K \to \infty$ で, 一様可積分の仮定から第１項, 命題 22 と $\mathrm{E}[\,|X|\,] < \infty$ から第２項, $\mathrm{E}[\,|X|\,] < \infty$ から第３項, が消える. 第４項は被積分関数を X_n に取り替えられるので一様可積分の仮定から消える. よって $\{X_n - X\}$ は一様可積分.

よって 命題 22 から，任意の $\epsilon > 0$ に対して $\delta > 0$ がとれて，$\mathrm{P}[\ F_n\ ] < \delta$ ならば $\mathrm{E}[\ |X_n - X|;\ F_n\ ] < \epsilon/2$, $n \in \mathbf{Z}$. 概収束することからこの δ に対して N が存在して $n \geq N$ ならば $\mathrm{P}[\ |X - X_n| > \epsilon/2\ ] < \delta$ とできる．以上から

$$\mathrm{E}[\ |X_n - X|\ ] = \mathrm{E}[\ |X_n - X|;\ |X_n - X| > \epsilon/2\ ] + \mathrm{E}[\ |X_n - X|;\ |X_n - X| \leq \epsilon/2\ ] \leq \epsilon/2 + \epsilon/2 = \epsilon.$$

(ii) から (iii)．$|\mathrm{E}[\ |X_n|\ ] - \mathrm{E}[\ |X|\ ]| \leq \mathrm{E}[\ |X_n - X|\ ]$ だから可積分性と絶対値の期待値の収束の両方を得る．

(iii) から (i)．$Y^K = Y\chi_{|Y| \leq K}$ などと書くことにすると，概収束の仮定と有界収束定理から $\mathrm{P}[\ |X| = K\ ] = 0$ を満たす K に対して $\displaystyle\lim_{n\to\infty} \mathrm{E}[\ |X_n^K|\ ] = \mathrm{E}[\ |X^K|\ ]$. これと仮定から，

$$\lim_{n\to\infty} \mathrm{E}[\ |X_n|;\ |X_n| \geq K\ ] = \mathrm{E}[\ |X|;\ |X| \geq K\ ].$$

他方 X が可積分だから任意の $\epsilon > 0$ に対して $K_0 > 0$ を選んで $K \geq K_0$ ならば $\mathrm{E}[\ |X|;\ |X| \geq K\ ] < \epsilon$ とできる．$K \geq K_0$ を $\mathrm{P}[\ |X| = K\ ] = 0$ を満たすように選んでおくと，N が存在して $n \geq N$ ならば

$$\mathrm{E}[\ |X_n|;\ |X_n| \geq K\ ] < 2\epsilon$$

とできる．各 n 毎に $\mathrm{E}[\ |X_n|;\ |X_n| \geq K_n\ ] < 2\epsilon$ を満たす K_n を選んでおくと，

$$\sup_{n\in\mathbf{Z}} \mathrm{E}[\ |X_n|;\ |X_n| \geq \max\{K_1, \cdots, K_{N-1}, K\}\ ] \leq 2\epsilon.$$

よって $\{X_n\}$ は一様可積分.

□

2.3.3　確率収束と概収束の関係．

確率収束しても概収束しない例．　L^p 収束しても概収束しない例に同じ.

$(0,1]$ 上の一様分布（ルベーグ測度空間）$(\Omega, \mathcal{F}, \mathrm{P}) = ((0,1], \mathcal{B}_1, \mu_1)$ を考える．$n = 1, 2, 3, \cdots$, と $k = 1, 2, \cdots, 2^n$ に対して $\xi_{n,k} = \chi_{((k-1)2^{-n}, k2^{-n}]}$ で確率変数 $\xi_{n,k}$ を定義する.（短い区間だけ 1 で残りは 0 となる関数.）これを辞書式順序 $\xi_{1,1}$, $\xi_{1,2}$, $\xi_{2,1}$, $\xi_{2,2}$, $\cdots$, に並べて $X_1, X_2, \cdots$ とおく．Y を恒等的に 0 となる関数 $Y = \chi_{\emptyset}$ とすると，$1 > \epsilon > 0$ に対して $\mathrm{P}[\ |\xi_{n,k} - Y| \geq \epsilon\ ] = \mathrm{P}[\ \xi_{n,k} = 1\ ] = 2^{-n}$ だから，$\displaystyle\lim_{n\to\infty} \mathrm{P}[\ |X_n - Y| \geq \epsilon\ ] = 0$．即ち $X_n \to Y$, in prob.. しかし，任意の $\omega \in (0,1] = \Omega$ に対して，$X_n(\omega)$ は無限回 0 と 1 になるから収束しない．特に X_n は Y に概収束しない.

確率収束から概収束について言えること．　集合列 $\{A_k\}$ に対して

$$\underline{\lim}_{k\to\infty} A_k = \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k\,, \quad \overline{\lim}_{k\to\infty} A_k = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k\,,$$

と書くことがある.

次の準備的定理は，測度論の復習である.

命題 24 (Borel–Cantelli の第１定理．) 事象列 $\{A_k\}$ に対して $\displaystyle\sum_{k=1}^{\infty} \mathrm{P}[\ A_k\ ] < \infty$ ならば

$$\mathrm{P}[\ \overline{\lim}_{k\to\infty} A_k\ ] = 0 \text{ 即ち，} \quad \mathrm{P}[\ \underline{\lim}_{k\to\infty} A_k^c\ ] = 1\,.$$

証明. $\displaystyle\sum_{k=1}^{\infty} \mathrm{P}[\ A_k\ ] < \infty$ から $\displaystyle\lim_{n\to\infty} \sum_{k=n}^{\infty} \mathrm{P}[\ A_k\ ] = 0$ が従う．これに確率の劣加法性と連続性 命題 1 を使い，$\displaystyle\overline{\lim}_{k\to\infty} A_k = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k$ に注意すれば主張を得る．□

命題 25 $X_n \to Y$ *in prob.*, ならば $\{X_n\}$ の部分列が存在して Y に概収束する．

証明. 各 $k \in \mathbf{N}$ に対して $\epsilon = 2^{-k}$ として確率収束の定義を用いると，

$$\mathrm{P}[\ |X_{n_k} - Y| > 2^{-k}\ ] < 2^{-k}$$

となる増加列 $\{n_k\ |, k = 1, 2, \cdots\}$ がとれることが分かる．よって $A_k = \{|X_{n_k} - Y| > 2^{-k}, k \in \mathbf{N},$ とおくと，$\displaystyle\sum_{k=1}^{\infty} \mathrm{P}[\ A_k\ ] < \infty$ である．Borel–Cantelli 第 1 定理から

$$\mathrm{P}[\ \varliminf_{k\to\infty} A_k^c\ ] = 1\,.$$

他方 $\displaystyle\varliminf_{k\to\infty} A_k^c = \bigcup_{n=1}^{\infty}\bigcap_{k=n}^{\infty} A_k^c$ だから，$\omega \in \displaystyle\varliminf_{k\to\infty} A_k^c$ ならばある整数 $n_0 = n_0(\omega)$ が存在して $k \ge n_0$ ならば $\omega \in A_k^c$，即ち $|X_{n_k}(\omega) - Y(\omega)| \le 2^{-k}$. これは $\displaystyle\lim_{k\to\infty} X_{n_k}(\omega) = Y(\omega)$ を意味する．確率 1 でこれが起こるから， $X_{n_k} \to Y$, a.e. $(k \to \infty)$ である． □

2.4 特性関数．

2.4.1 定義．

複素数値関数 $X:\ \Omega \to \mathbf{C}$ が確率変数であるとは，実部と虚部が共に実数値確率変数であることをいい，期待値（積分）は $\mathrm{E}[\ X\ ] = \mathrm{E}[\ \mathbf{Re}(X)\ ] + \sqrt{-1}\mathrm{E}[\ \mathbf{Im}(X)\ ]$ で（これらの積分が存在する X に対して）定義する．

ほとんど至るところ実数値をとる確率変数 X に対して $\varphi_X(t) = \mathrm{E}[\ e^{\sqrt{-1}Xt}\ ]$ で定義される $\varphi_X:\ \mathbf{R} \to \mathbf{C}$ を X の特性関数という．

X の分布 (5) を μ とすれば，$\varphi_X(t) = \displaystyle\int e^{\sqrt{-1}xt}\mu(dx)$．これを μ **の特性関数**ともいう．

$|e^{\sqrt{-1}Xt}| \le 1$ だから,（X の値域が（確率 0 を除いて）実数に入っていれば期待値は存在する．

確率変数の特性関数とは，確率変数の分布密度（が存在するときはそ）のフーリエ変換のことである．

多成分 ($\mathbf{R}^n$ 値) 確率変数はこの講義ではほとんど触れないが，中心極限定理のときに触れるのでコメントだけしておく．

$X = (X_1, X_2)$ の特性関数は $\varphi_X:\ \mathbf{R}^2 \to \mathbf{C}$ と 2 変数になるだけである．即ち， $\varphi_X(t_1, t_2) = \mathrm{E}[\ e^{\sqrt{-1}(X_1t_1 + X_2t_2)}\ ]$．これから，特に， $Y = X_1 + X_2$ の特性関数は $\varphi_Y(t) = \varphi_X(t, t)$ と書けることが分かる．

2.4.2 簡単な性質．

φ_X を確率変数 X の特性関数とする．

$\varphi_X(0) = 1$，$|\varphi_X(t)| \le 1$，$\varphi_X(-t) = \overline{\varphi_X(t)}$ (複素共役)，はいずれも簡単．

命題 26 *(i)* φ_X は一様連続関数である．

(ii) 任意の $N \in \mathbf{N}$, $z_k \in \mathbf{C}$, $t_k \in \mathbf{R}$, $k = 1, \cdots, N$, に対して

$$\sum_{j,k=1}^{N} z_k \bar{z}_j \varphi_X(t_k - t_j) \ge 0\,.$$

（正定符号性）

(iii) （モーメントとの関係.）ある $n \in \mathbf{N}$ に対して $\mathrm{E}[\,|X|^n\,] < \infty$ ならば

$$\varphi_X^{(n)}(t) = \sqrt{-1}^n \mathrm{E}[\,X^n e^{\sqrt{-1}Xt}\,]$$

が存在して t の連続関数.
偶数次のモーメントについては逆が成り立つ．即ち，$\varphi_X^{(2n)}(0)$ が存在すれば $\mathrm{E}[\,|X|^{2n}\,] < \infty$.
（証明に対称差を用いるので奇数次では不成立．反例が知られている *[9, §5.2 例 1].*）

証明. (i) $|\varphi_X(t+h) - \varphi_X(t)| \le \mathrm{E}[\,|e^{\sqrt{-1}(t+h)X} - e^{\sqrt{-1}\,t\,X}|\,] = \mathrm{E}[\,|(e^{\sqrt{-1}hX} - 1)e^{\sqrt{-1}tX}|\,] = \mathrm{E}[\,|e^{\sqrt{-1}hX} - 1|\,]$.
$|e^{\sqrt{-1}hX} - 1| \le 2$ だから，優収束定理から期待値（積分）と極限 $h \to 0$ を交換できる．すなわち，$\lim_{h\to 0} \mathrm{E}[\,|e^{\sqrt{-1}hX} - 1|\,] = \mathrm{E}[\,\lim_{h\to 0} |e^{\sqrt{-1}hX} - 1|\,] = 0$．この収束は t によらないから，$|\varphi_X(t+h) - \varphi_X(t)|$ は $h \to 0$ のとき t について一様に 0 に収束する.
(ii)

$$\sum_{j,k=1}^{N} z_k \bar{z}_j \varphi_X(t_k - t_j) = \mathrm{E}[\,|\sum_{k=1}^{N} z_k e^{\sqrt{-1}t_k X}|^2\,] \ge 0\,.$$

(iii) $n = 1$ のとき.

$$\left|\frac{1}{h} e^{\sqrt{-1}Xt}(e^{\sqrt{-1}Xh} - 1)\right| \le |X|$$

だから優収束定理より

$$\frac{d}{dt}\varphi_X(t) = \sqrt{-1}\mathrm{E}[\,Xe^{\sqrt{-1}Xt}\,]$$

が存在．t についての連続性も $|\mathrm{E}[\,Xe^{\sqrt{-1}Xt}\,]| \le \mathrm{E}[\,|X|\,]$ と優収束定理から同様.
(7) から $0 < n' \le n$ なら $\mathrm{E}[\,|X|^{n'}\,] < \infty$ に注意．n で定理が成り立つとし，$\mathrm{E}[\,|X|^{n+1}\,] < \infty$ を仮定する．$\varphi_X^{(n)}(t) = \sqrt{-1}^n \mathrm{E}[\,X^n e^{\sqrt{-1}Xt}\,]$ において，$n = 1$ のときと同様に

$$\left|\frac{1}{h}(\varphi_X^{(n)}(t+h) - \varphi_X^{(n)}(t))\right| \le \mathrm{E}[\,|X|^{n+1}\,]$$

だから優収束定理より $h \to 0$ が存在．連続性も同様．n についての帰納法より前半の主張が成り立つ.
後半．n 次の対称差 $\varphi_{X,n}$ を帰納的に,

$$\varphi_{X,1}(t) = h^{-2}(\varphi_X(t+h) - 2\varphi_X(t) + \varphi_X(t-h))$$

および,

$$\varphi_{X,k+1}(t) = h^{-2}(\varphi_{X,k}(t+h) - 2\varphi_{X,k}(t) + \varphi_{X,k}(t-h))$$

で定義すると,

$$\varphi_{X,n}(t) = \mathrm{E}[\,e^{\sqrt{-1}Xt}(\sqrt{-1}\frac{2}{h}\sin\frac{hX}{2})^{2n}\,]$$

が帰納的に分かる．$h \to 0$ とすると対称差は $2n$ 階微分に収束．Fatou の補題より右辺を変形すると,

$$\infty > |\varphi_X^{(2n)}(0)| = \lim_{h\to 0} \mathrm{E}[\,\left(\frac{2}{h}\sin\frac{hX}{2}\right)^{2n}\,] \ge \mathrm{E}[\,X^{2n}\,].$$

□

2.4.3 直積測度と Fubini の定理（復習）.

$(\Omega_i,\mathcal{F}_i)$, $i=1,2$, を可測空間，その直積空間を $(\Omega,\mathcal{F})=(\Omega_1\times\Omega_2,\mathcal{F}_1\times\mathcal{F}_2)$ とする．

集合の直積 $\Omega=\Omega_1\times\Omega_2$ は周知と思う．σ 加法族の直積とは矩形集合 $E\times F$, $E\in\mathcal{F}_1$, $F\in\mathcal{F}_2$, を全て含む最小の σ 加法族を表す.

$A\subset\Omega$ と $\omega_1\in\Omega_1$ に対して $A_{\omega_1}=\{\omega_2\in\Omega_2\mid(\omega_1,\omega_2)\in A\}$ を A の ω_1 における**切り口**という. また， $X:\ \Omega\to\mathbf{R}\cup\{\pm\infty\}$ に対して，ω_1 を固定して Ω_2 上の実関数と見たものを $X(\omega_1,\cdot)$ などと書く.

測度空間 $(\Omega_i,\mathcal{F}_i,\mathrm{P}_i)$, $i=1,2$, の直積空間を

$$(\Omega,\mathcal{F},\mathrm{P})=(\Omega_1\times\Omega_2,\mathcal{F}_1\times\mathcal{F}_2,\mathrm{P}_1\times\mathrm{P}_2)$$

とする．ここで，直積測度 $\mathrm{P}=\mathrm{P}_1\times\mathrm{P}_2$ とは，矩形集合 $E\times F$ に対して

$$\mathrm{P}[\,E\times F\,]=\mathrm{P}_1[\,E\,]\mathrm{P}_2[\,F\,],\quad E\in\mathcal{F}_1,\ F\in\mathcal{F}_2,$$

を満たす測度のことである[16] .

そのような確率測度が存在する（しかもただ一つに決まる）ことは，ルベーグ測度の存在などと同様に Hopf の拡張定理によって証明される.（即ち，σ 加法性を持つ有限加法的測度になることを示せば，外測度が定義できて，それに関して可測集合の族をとって，そこに制限すれば測度になる.）以上は測度論で既知とする.

定理 27 (Fubini の定理) *(i)* $(\Omega_i,\mathcal{F}_i)$, $i=1,2$, を可測空間，その直積空間を $(\Omega,\mathcal{F})=(\Omega_1\times\Omega_2,\mathcal{F}_1\times\mathcal{F}_2)$ とする．$A\in\mathcal{F}$ ならば，全ての $\omega_1\in\Omega_1$ に対して切り口 A_{ω_1} は $\mathcal{F}_2$–可測である.
また， $X:\ \Omega\to\mathbf{R}\cup\{\pm\infty\}$ が $\mathcal{F}$–可測ならば，全ての $\omega_1\in\Omega_1$ に対して $X(\omega_1,\cdot)$ は $\mathcal{F}_2$–可測.
(ii) $(\Omega_i,\mathcal{F}_i,\mathrm{P}_i)$, $i=1,2$, を可測空間，その直積空間を

$$(\Omega,\mathcal{F},\mathrm{P})=(\Omega_1\times\Omega_2,\mathcal{F}_1\times\mathcal{F}_2,\mathrm{P}_1\times\mathrm{P}_2)$$

とする．$X:\ \Omega\to\mathbf{R}$ が非負値 $\mathcal{F}$–可測関数ならば， $\displaystyle\int_{\Omega_i}X(\omega_1,\omega_2)\,\mathrm{P}_i[\,d\omega_i\,]$, $i=1,2$, は $\mathcal{F}_i$–可測関数（但し値として $+\infty$ を許す）であり，

$$\begin{aligned}\int_{\Omega}X(\omega_1,\omega_2)\,\mathrm{P}[\,d\omega_1d\omega_2\,]&=\int_{\Omega_1}\left(\int_{\Omega_2}X(\omega_1,\omega_2)\,\mathrm{P}_2[\,d\omega_2\,]\right)\mathrm{P}_1[\,d\omega_1\,]\\&=\int_{\Omega_2}\left(\int_{\Omega_1}X(\omega_1,\omega_2)\,\mathrm{P}_1[\,d\omega_1\,]\right)\mathrm{P}_2[\,d\omega_2\,].\end{aligned}\tag{11}$$

また，X が非負とは限らない可積分な関数ならば， $\displaystyle\int_{\Omega_j}X(\omega_1,\omega_2)\,\mathrm{P}_j[\,d\omega_j\,]$ は Ω_i 上の可積分関数で，*(11)* が成り立つ.

証明. (i) A が矩形集合 $A=A_1\times A_2$, $A_i\in\mathcal{F}_i$, $i=1,2$, ならば明らか.

$$\mathcal{A}=\{A\subset\Omega\mid A_{\omega_1}\in\mathcal{F}_2,\ \omega_1\in\Omega_1\}$$

とおくと $\mathcal{A}$ は矩形集合を全て含む σ 加法族だから $\mathcal{A}\supset\mathcal{F}$. よって，$\mathcal{F}$ の要素は $\mathcal{A}$ の性質を持つ.

後半． $Y(\cdot)=X(\omega_1,\cdot):\ \Omega_2\to\mathbf{R}$ とおくとき，$A\in\mathcal{B}_1$ に対して $Y^{-1}(A)\in\mathcal{F}_2$ を言えばよい.

$$\omega_2\in Y^{-1}(A)\ \Leftrightarrow\ X(\omega_1,\omega_2)\in A\ \Leftrightarrow\ \omega_2\in(X^{-1}(A))_{\omega_1}\,.$$

即ち $Y^{-1}(A)=(X^{-1}(A))_{\omega_1}$. $X^{-1}(A)\in\mathcal{F}$ なので，前半より，切り口である $Y^{-1}(A)\in\mathcal{F}_2$.

[16] ルベーグ積分のように，$\mathrm{P}_1[\,\Omega_1\,]=\infty$ か $\mathrm{P}_2[\,\Omega_2\,]=\infty$ だと $0\times\infty$ が生じるおそれがあるので，もう少し丁寧に言わないといけないが，確率論では測度は有限なので，この点は問題がない.

(ii) 次の順序で示す [9, 4 章 §2 定理 3].

(a) $X = \chi_\Lambda$, $\Lambda \in \mathcal{F}$ のとき.

i. Λ が矩形集合 $\Lambda_1 \times \Lambda_2$ のとき．直接示せる.

ii. Λ が矩形集合の有限和のとき積分の線形性から.
そういう集合全体を $\mathcal{A} \subset \mathcal{F}$ とすると，直積 σ 加法族の定義から $\sigma[\mathcal{A}] = \mathcal{F}$ であることに注意.

iii. Λ が一般の場合．定理の成り立つ $\mathcal{F}$ の要素全体を $\mathcal{M} \subset \mathcal{F}$ とすると $\mathcal{M} \supset \mathcal{A}$ を上で言った．$\mathcal{M}$ が単調族であることをいうと，単調族定理から $\mathcal{M} \supset \sigma[\mathcal{A}] = \mathcal{F}$．従って $\mathcal{M} = \mathcal{F}$ となって，任意の $\Lambda \in \mathcal{F}$ が定理の性質を持つ.
$\mathcal{M}$ が単調族であることをいうには，$\mathcal{M}$ の増大列 Λ_n, $n \in \mathbf{N}$, に対して，χ_{Λ_n} は増大しながら $\chi_{\bigcup_n \Lambda_n}$ に近づくし，その切り口も同様であるから，単調収束定理と，$\mathcal{M}$ の定義から $\displaystyle\bigcup_{n\in\mathbf{N}} \Lambda_n \in \mathcal{M}$ を得る．また，補集合に関しても閉じているので減少列の極限に関しても閉じていることになるので，単調族である.

(b) X が単関数のとき．積分の線形性から.

(c) X が非負 $\mathcal{F}$–可測関数のとき．単関数増大極限で近似して単調収束定理から.

(d) X が可積分のとき．正の部分と負の部分に分ける.

□

2.4.4 P. Lévy の反転公式.

$\displaystyle\varphi_X(t) = \int e^{\sqrt{-1}xt}\mu(dx)$ であるが，特性関数の話では確率変数 X の分布であることは用いないので，しばらく分布 μ の特性関数ということにして，

$$\varphi(t) = \int_{\mathbf{R}} e^{\sqrt{-1}xt}\mu(dx)$$

とおく.

反転公式は，要するにフーリエ逆変換のことである．結果として，特性関数が分布を定める（特性関数が一致すれば分布が一致する）ことが分かるので，重要である.

定理 28 (P. Lévy の反転公式) $\displaystyle\varphi(t) = \int_{\mathbf{R}} e^{\sqrt{-1}xt}\mu(dx)$ *を分布 (1 次元ボレル確率測度)* μ の特性関数とすると，$a < b$ に対して，

$$\lim_{T\to\infty} \frac{1}{2\pi}\int_{-T}^{T} \frac{e^{-\sqrt{-1}bt} - e^{-\sqrt{-1}at}}{-\sqrt{-1}t}\varphi(t)\,dt = \mu((a,b)) + \frac{1}{2}(\mu(\{a\}) + \mu(\{b\})).$$

系 29 分布 μ, $\tilde{\mu}$ の特性関数 φ, $\tilde{\varphi}$ が $\varphi = \tilde{\varphi}$ を満たせば $\mu = \tilde{\mu}$.

定理 *28* を仮定したときの系 *29* の証明. 任意の a に対して，定理 28 から $\mu(\{a\}) = \tilde{\mu}(\{a\})$ を得て，従って任意の a に対して，定理 28 から $\mu((-\infty,a]) = \tilde{\mu}((-\infty,a])$ を得る．$\mathcal{B}_1 = \sigma[\{(-\infty,a] \mid a \in \mathbf{R}\}]$ なので $\mu = \tilde{\mu}$.

□

補題 30 *(i)*

$$\int_0^T e^{-ut}\sin t\,dt = \frac{1}{u^2+1} - \frac{e^{-uT}}{u^2+1}(\cos T + u\sin T). \tag{12}$$

(ii)

$$\lim_{T\to\infty}\int_0^T \sin xt\frac{dt}{t} = \begin{cases} \dfrac{\pi}{2}, & x>0,\\ 0, & x=0,\\ -\dfrac{\pi}{2}, & x<0\,. \end{cases} \tag{13}$$

証明. (i) $J=\displaystyle\int_0^T e^{-ut}\sin t\,dt$ とおく．2 回部分積分して，

$$J = 1-e^{-Tu}\cos T - ue^{-Tu}\sin T - u^2 J.$$

これより (12) を得る．

(ii) (12) より，

$$\int_0^T \sin t\frac{dt}{t} = \int_0^T \left(\int_0^\infty e^{-ut}du\right)\sin t\,dt$$
$$= \int_0^\infty \left(\int_0^T e^{-ut}\sin t\,dt\right)du = \int_0^\infty \frac{1}{u^2+1}\,du - \int_0^\infty \frac{e^{-uT}}{u^2+1}(\cos T+u\sin T)\,du.$$

第 2 項は $2\displaystyle\int_0^\infty e^{-uT}\,du = 2/T \longrightarrow 0\ (T\to\infty)$ で抑えられる．これより (13) を得る．

□

定理 *28* **の証明**. Fubini の定理と対称性より，

$$\frac{1}{2\pi}\int_{-T}^T \frac{e^{-\sqrt{-1}bt}-e^{-\sqrt{-1}at}}{-\sqrt{-1}t}\varphi(t)\,dt = \frac{1}{2\pi}\int_{-T}^T \frac{e^{-\sqrt{-1}bt}-e^{-\sqrt{-1}at}}{-\sqrt{-1}t}\int_{\mathbf{R}} e^{\sqrt{-1}xt}\mu(dx)\,dt$$
$$= \frac{1}{\pi}\int_{\mathbf{R}}\int_0^T \frac{\sin(x-b)t-\sin(x-a)t}{-t}dt\,\mu(dx).$$

t 積分は，変数変換して x, b, a を消去した後， $\pi/2$ までの積分で評価できることに注意すれば優収束定理が使えるので，(13) より，

$$\lim_{T\to\infty}\frac{1}{2\pi}\int_{-T}^T \frac{e^{-\sqrt{-1}bt}-e^{-\sqrt{-1}at}}{-\sqrt{-1}t}\varphi(t)\,dt = \frac{1}{\pi}\int_{\mathbf{R}}\int_0^\infty \frac{\sin(x-b)t-\sin(x-a)t}{-t}dt\,\mu(dx)$$
$$= \mu((a,b)) + \frac{1}{2}(\mu(\{a\})+\mu(\{b\})).$$

□

2.5 弱収束に関する道具的概念．

2.5.1 特性関数と弱収束．

特性関数による，弱収束の 定理 18 以外の同値条件．

弱収束（法則収束）は分布列（一般に測度の列）があれば考えることができるので，引き続き確率変数の言葉を suppress する．$\varphi(t)=\mathrm{E}[\,e^{\sqrt{-1}Xt}\,]$ や $\displaystyle\int f\,d\mu = \mathrm{E}[\,f(X)\,]$ などの対応で翻訳は容易である．

定理 31 (Glivenko の定理) φ, φ_n, $n\in\mathbf{N}$, をそれぞれ分布 μ, μ_n, $n\in\mathbf{N}$, の特性関数とする．このとき

$$\lim_{n\to\infty}\varphi_n(t)=\varphi(t),\quad t\in\mathbf{R},$$

*(*各点収束*)* と μ_n が μ に弱収束することは同値である．

注. (i) 弱収束から特性関数の各点収束は,（特性関数が有界連続関数なので）弱収束の定義そのもの. 逆が Glivenko の定理と呼ばれる．実は弱収束していると特性関数は広義一様収束するが，これは tightness という（弱収束する測度の列が持つ）重要な性質に由来するので後述する (§2.5.2).

(ii) [2] の証明は Fubini の定理と関数解析的手法のありかたを示す一つの初等的例題にもなっているので，詳しく取り上げたい．

関数解析的手法（基礎編）：（実）関数解析では,

(a) 極限と積分などの順序交換のできる形に主張を書き換えておく,

(b) 扱いやすい近似列の極限として一般の関数を扱う,

のが基礎的な方針である．

◇

定理 *31* **の証明**. 注に書いたように弱収束から特性関数の各点収束は当然なので，特性関数の収束から弱収束を証明する．

> 次の順序で証明する．
>
> (i) 可積分関数 g を用いて $f(x) = \int e^{\sqrt{-1}yx} g(y)dy$ と書ける場合，つまり可積分関数のフーリエ変換.（Fubini の定理による.）
>
> (ii) Compact support 連続関数 f の場合．上記の形の関数で一様近似できることを示す．

$g:\ \mathbf{R} \to \mathbf{R}$ を可積分関数とし，$f(x) = \int e^{\sqrt{-1}yx} g(y)dy$ とおく．先ず，この形の f に対して (9) を証明する．Fubini の定理から

$$
\begin{aligned}
&\left| \int f(x)\mu_n(dx) - \int f(x)\mu(dx) \right| = \left| \int (\varphi_n(y) - \varphi(y)) g(y) dy \right| \\
&\le \int |\varphi_n(y) - \varphi(y)|\, |g(y)| dy.
\end{aligned}
$$

特性関数は有界（絶対値 1 以下），g が可積分，だから有収束定理から積分と極限を交換できて，φ_n の各点収束の仮定から (9) を得る．

次に，compact support 連続関数 f が上記のような関数 f で一様近似できることを言う．即ちこの場合も (9) が成り立つことを証明する．具体的には,

$$
g_n(z) = \frac{1}{2\pi} e^{-z^2/(2n)} \int e^{-\sqrt{-1}zy} f(y)dy, \quad h_n(x) = \int e^{\sqrt{-1}zx} g_n(z)dz,
$$

とおくとき，g_n が可積分，$h_n \to f$ が一様収束，を示す．

Fubini の定理から

$$
\begin{aligned}
h_n(x) &= \frac{1}{2\pi} \int\int e^{\sqrt{-1}z(x-y) - z^2/(2n)} dz\, f(y)dy = \sqrt{\frac{n}{2\pi}} \int e^{-n(x-y)^2/2} f(y)dy \\
&= \sqrt{\frac{1}{2\pi}} \int e^{-t^2/2} f(x + \frac{t}{\sqrt{n}}) dt.
\end{aligned}
$$

よって，積分範囲を $|t| \le a$ と $|t| > a$ に分けて,

$$
\begin{aligned}
|h_n(x) - f(x)| &\le \sqrt{\frac{1}{2\pi}} \int e^{-t^2/2} |f(x + \frac{t}{\sqrt{n}}) - f(x)| dt \\
&\le \sup_{|t| \le a} |f(x + \frac{t}{\sqrt{n}}) - f(x)| + \frac{2 \sup_x |f(x)|}{\sqrt{2\pi}} \int_{|t|>a} e^{-t^2/2} dt.
\end{aligned}
$$

f は compact support 連続関数なので, $\sup_x |f(x)| < \infty$, かつ，一様連続になるので，右辺第１項は $n \to \infty$ で x について一様に消える．最後に $a > 0$ は任意で，左辺は a によらないから，$a \to \infty$ を考えれば第２項も消える．よって，

$$\overline{\lim_{n\to\infty}} \sup_x |h_n(x) - f(x)| = 0\,.$$

□

注. (i) [3, §24 定理 24.1 の直前 $L \subset L_\infty$, §29 定理 29.4 系 2] によれば，可積分関数のフーリエ変換は compact support 連続関数を一様近似することは分かる（上記証明の該当部分はその復習になっている）が，包含しているかどうかについては言及していないようにみえる.（逆向きの包含関係がないことはすぐ分かるが.）実は，包含関係はない．即ち，compact support 連続関数であって可積分関数のフーリエ変換で書けないものがある.

なお，可積分関数のフーリエ変換が compact support 連続関数を一様近似するという証明に用いられた積分核は，急減少関数が L^p 関数空間で稠密であることに広く用いられる [3, §25,§29].

(ii) Tightness argument は極限測度の構成（存在証明）のためにあるので，極限が（確率測度として）存在することが分かっている場合は，tightness argument はいらないはずである．しかし，西尾真喜子，佐藤担，Shiryaev の Probability(Springer)[17] などの教科書ではこの種の命題の証明は tightness argument (Prokhorov Theorem) に基づいて行われる.（もちろん，極限が確率測度であることが前もって分かっていない場合の定理と同時に証明してしまおうというねらいからだが.）

Tightness argument を使わない証明は，伊藤清の確率論（岩波基礎数学選書）§2.6, §2.5 にある[18]．なお，t についての広義一様性は仮定に必要ない.

参考までに [6, §13.1 定理 13.2] の証明を引用する．§2.5.2 の議論 (Prokhorov Theorem) を前提とする．特性関数の各点収束を仮定する．$\epsilon > 0$ とする．特性関数は連続関数なので，$|1-\varphi(t)| < \frac{\epsilon}{20}$, $|t| \le A$, となる $A > 0$ がある．他方，有界収束定理より，ある n_0 があって，

$$\left|\int_0^A \mathbf{Re}(\varphi_n(t) - \varphi(t))dt\right| \le \frac{A\epsilon}{20}\,,\quad n \ge n_0\,,$$

となる．以上と切断不等式命題 34 より，$n \ge n_0$ ならば

$$\mu_n([-A^{-1}, A^{-1}]^c) \le \frac{10}{A}\int_0^A (1 - \mathbf{Re}(\varphi_n(t)))dt \le \epsilon$$

となる．よって $\{\mu_n\}$ は弱収束する部分列を持つ．その極限分布の特性関数は φ_n の極限 φ であり，部分列によらない．従って，極限分布も部分列によらないから，分布の収束が言えた.

◇

2.5.2 Tightness と弱収束.

定義 2 分布からなるある集合 $\mathcal{M}$ が **tight** であるとは，$\mathcal{M}$ の中の任意の無限列が分布に弱収束する部分列を持つことを言う.

注. (i) 極限分布は部分列によって違ってもよいし，分布でさえあれば $\mathcal{M}$ の中になくてもよい.
(ii) $\mathbf{R}^d$ の分布でよいが，簡単のため以下では $\mathbf{R}$ 上の分布とする.

◇

[17] 千代延先生の情報.
[18] 千代延先生に教わった.

定理 32 (Prokhorov の定理) 分布からなるある集合 $\mathcal{M}$ が *tight* であるための必要十分条件は

$$(\forall\epsilon > 0)\ \exists K \subset \mathbf{R}:\ compact;\ (\forall\mu \in \mathcal{M})\mu(K) > 1 - \epsilon. \tag{14}$$

(14) が必要であることだけ証明する.

証明. **必要であること.** $\mathcal{M}$ が tight なのに (14) が成り立たないとする. 即ち, $\epsilon > 0$ が存在して, $n = 1, 2, 3, \cdots$ に対して $\mu_n \in \mathcal{M}$ がとれて $\mu_n([-n, n]) \le 1 - \epsilon$. 仮定から μ_n , $n = 1, 2, 3, \cdots$, は弱収束部分列 $\mu_{n(k)}$, $k = 1, 2, 3, \cdots$, を持つ. その極限を μ とする.

開集合 $G_m = (-m, m)$ に対して, $n(k) \ge m$ なる k に対しては $G_m \subset [-n(k), n(k)]$ だから $\mu_{n(k)}(G_m) \le 1 - \epsilon$, $n(k) \ge m$. 定理 18 から

$$\mu(G_m) \le \varliminf_{k\to\infty} \mu_{n(k)}(G_m) \le 1 - \epsilon, \quad m \in \mathbf{N}.$$

よって $\mu(\mathbf{R}) = \lim_{m\to\infty} \mu(G_m) \le 1 - \epsilon$ となるが, これは μ が確率測度であることに矛盾する. □

十分であること.

(14) が成り立つとき $\mu_n \in \mathcal{M}$, $n \in \mathbf{N}$, が収束部分列を持つことを証明する. こちらのほうが極限測度を構成しないといけないので高度である.

[15, §VIII.6], [9, §5.4], [13, §2.2 (2.5), §2.9 (9.2)] は Helly の方法によって証明する：

(i) 対角線論法で, 全ての有理点で収束するような分布関数の部分列があることを言う.
(ii) 単調性を利用して無理点に拡張する（$F(x) = \inf_{q>x} F_0(q)$）と, 分布関数になるための条件のうち $F(\infty) - F(-\infty) = 1$ 以外の全ての性質 ($F(x+h) - F(x) \ge 0$, $F(x+0)) = F(x)$) が満たされる.
(iii) (14) から最後の性質（全測度が 1 ということ）が言える.（即ち, $x < -K$ ならば $F(x) < \epsilon$, および $x > K$ ならば $F(x) > 1 - \epsilon$.）これで F がある確率測度の分布関数になっていることが分かる.
(iv) F の連続点 x に対して $F_{n(k)}(x) \to F(x)$ を言う. これが弱収束の同値条件になることが知られている. [19]

詳しい証明は分布関数の知識を要するのでここでは省略する (§A.2.2 参照).

定理 31 で測度の弱収束から特性関数の各点収束が従うことを注意した（実際は定義から明らか）が, 実は特性関数の広義一様収束が従う. 極限測度が与えられていても, 一様収束のためには tightness が本質となる.

---

[19] [26, 定理 2.3.4, 定理 2.3.5] は一般的な Poland 空間上の確率測度の場合の証明までカバーしている.

(i) Poland 空間では可算稠密集合がとれるので一列に並べて, k 番目との距離を小数第 k 位とすることでコンパクト距離空間の分布に写す.
(ii) コンパクト距離空間の場合に Prokhorov の定理を証明する.
  (a) コンパクト距離空間上の連続関数の集合は max ノルムで可分距離空間になる.
  (b) 可算個の連続関数を与えたとき, それらの積分の列が全て収束するような元の測度列 (をコンパクト距離空間に写したもの) の部分列が存在する (対角線論法). それが求める収束部分列.
  (c) 上記の積分の列の極限は全ての連続関数に拡張できる. 極限を与える操作を連続関数の集合上の写像とみると, それは正値加法的線形汎関数なので Rietz の定理から測度の積分で書ける. それが求める極限測度.
  (d) 確率測度になることは (14) から.
(iii) コンパクト距離空間から戻ってくる. (14) で確率が 1 になるコンパクト集合可算個の和集合 A を選べる. コンパクト空間に写すための写像が 1:1 onto conti であることに注意すると, A では逆像が可測になる. それで戻ってくる. あとはコンパクト集合で近似して閉集合による弱収束の定義を証明する.

定理 33 $\varphi,\ \varphi_n,\ n \in \mathbf{N}$, をそれぞれ分布 $\mu,\ \mu_n,\ n \in \mathbf{N}$, の特性関数とする．$\mu_n$ が μ に弱収束するならば φ_n は φ に広義一様収束する．

証明. [9, 5 章 §5 定理 3] もあるが，[26, 定理 3.1.7] に従う．定理 31 で特性関数の各点収束は分かっている．広義一様収束を言うには任意の $L > 0$ に対して $\varphi_n(t),\ n \in \mathbf{N}$, が $t \in [-L, L]$ で一様有界，同程度連続であることを言えばよい (Ascoli–Arzela の定理)．一様有界性は $|\varphi_n(t)| \leq 1$ だから明らか．同程度連続性をいう．

$\epsilon > 0$ を任意にとる．$\{\mu_n,\ n \in \mathbf{N}\}$ は tight だから，定理 32（の「必要性」）から $K > 0$ がとれて，

$$\mu_n([-K, K]^c) < \epsilon, \quad n \in \mathbf{N}.$$

$|h| < \epsilon/K$ ならば任意の $|x| \leq K$ に対して

$$|e^{\sqrt{-1}hx} - 1| = \left| x \int_0^h e^{\sqrt{-1}sx} ds \right| \leq K|h| < \epsilon$$

なので，

$$\begin{aligned} &|\varphi_n(t+h) - \varphi_n(t)| \leq \int_{[-K,K]} |e^{\sqrt{-1}hx} - 1|\, \mu_n(dx) + \int_{[-K,K]^c} |e^{\sqrt{-1}hx} - 1|\, \mu_n(dx) \\ &< \epsilon + 2\mu_n([-K, K]^c) < 3\epsilon. \end{aligned}$$

よって同程度連続．

□

Tightness argument は，極限測度の候補が前もって分かっていないとき，それを構成する十分条件を与える定理の証明で使われる．条件が特性関数で書かれるときは次の切断不等式が一つの補助手段となる．

命題 34 (切断不等式．) φ を分布 μ の特性関数とするとき， $A > 0$ ならば

$$\mu([-A^{-1}, A^{-1}]^c) \leq \frac{10}{A} \int_0^A (1 - \mathbf{Re}(\varphi(t)))dt.$$

注. 左辺の区間上端と右辺の積分範囲上限が逆数であることに注意．理論物理学（量子力学）で「不確定性原理」と呼ばれるいくつかの数学的事実の総称に含まれる現象の一つ． ◇

証明. $\dfrac{\sin u}{u}$ は対称関数で， $0 \leq u \leq \pi$ では単調減少，$u \geq \pi$ では $\dfrac{1}{\pi} < \dfrac{1}{2} = \sin \dfrac{\pi}{6} < \sin 1$ 以下，よって $|u| \geq 1$ で $\sin 1$ 以下．これより，

$$\begin{aligned} &\frac{1}{A} \int_0^A (1 - \mathbf{Re}(\varphi(t)))dt = \int_{\mathbf{R}} (1 - \frac{1}{A} \int_0^A \cos(tx)dt)\mu(dx) \\ &= \int_{\mathbf{R}} (1 - \frac{\sin Ax}{Ax})\mu(dx) \geq \int_{[-A^{-1},A^{-1}]^c} (1 - \frac{\sin Ax}{Ax})\mu(dx) \\ &\geq (1 - \sin 1)\mu([-A^{-1}, A^{-1}]^c) \geq \frac{1}{10}\mu([-A^{-1}, A^{-1}]^c)\,. \end{aligned}$$

最後の変形で $\sin 1 < \sin \frac{\pi}{3} = \frac{\sqrt{3}}{2} < \frac{9}{10}$ を使った [6, §12.3 定理 12.6]. □

定理 35 分布の列 $\mu_n,\ n = 1, 2, 3, \cdots$, *(*$\mu_n$ の特性関数を φ_n とおく*)* と関数 $f:\ \mathbf{R} \to \mathbf{R}$ に対して，もし，$t \in \mathbf{R}$ の各点で $\displaystyle\lim_{n\to\infty} \varphi_n(t) = f(t)$ であってしかも，

(i) $f(t)$ が $t = 0$ で連続 *(Lévy の定理)* か，または，

(ii) $t=0$ のある近傍で一様収束,

ならば, μ_n はある分布 μ に弱収束し，しかも，その特性関数は f である．

証明. (i) [6, §13.1 定理 13.3]，[9, 5 章 §5 定理 4][20] ．

$\epsilon>0$ を任意にとる．f が原点で連続なので $A=A(\epsilon)>0$ があって，$|t|<A$ ならば $|1-f(t)|<\epsilon$ とできるので,

$$\left|\int_0^A (1-\mathbf{Re}(f(t)))dt\right| < A\epsilon.$$

他方 φ_n が f に各点収束するので有界収束定理から自然数 $N=N(A)$ が存在して $n\ge N$ ならば

$$\left|\int_0^A \mathbf{Re}(\varphi_n(t)-f(t))dt\right| < A\epsilon.$$

命題 34 から

$$\begin{aligned}\mu_n([-A^{-1},A^{-1}]^c) \le\ & \frac{10}{A}\left(\int_0^A (1-\mathbf{Re}(f(t)))dt+\int_0^A \mathbf{Re}(f(t)-\varphi_n(t))dt\right)\\ <\ & 20\epsilon,\quad n\in\mathbf{N}.\end{aligned}$$

よって Prokhorov の定理 (定理 32 の「十分性」) から $\{\mu_n \mid n\in\mathbf{N}\}$ は tight になって，任意の部分列に弱収束部分列がある．

その極限分布の特性関数は定理 31 から f なので，部分列によらない．系 29 から極限分布自体が部分列によらない．よって μ_n は収束して，その弱収束極限分布の特性関数は f である．

(ii) 特性関数が原点近傍で一様収束していれば，極限もその近傍で連続関数になるから上の結果より弱収束が言える．直接証明が [6, §13.1 定理 13.3] にある．

□

注. 定理 35 にあるように何らかの付加条件がなければ（特性関数の各点収束だけでは)，極限関数が何かの確率測度の特性関数にはならない.（Tightness によって極限測度を構成するのだから，そのための条件が必要なのは証明からは明らかだが.）例：$\varphi_n(t)=\dfrac{1}{2n}\displaystyle\int_{-n}^{n} e^{\sqrt{-1}tx}dx$ は $n\to\infty$ で $\varphi=\chi_{\{0\}}$ に各点収束するが，φ は確率測度の特性関数にならない． ◇

次の定理は与えられた関数が何らかの分布の特性関数になる必要十分条件を（分布の言葉を全く使わずに）与えるという意味で必ず言及があるが，この講義では恐らく用いないので証明は [9, §5.5 定理 6] などに譲り省略.

定理 36 (Bochner の定理) $\varphi:\ \mathbf{R}\to\mathbf{C}$ が以下を満たせば，ある分布の特性関数になっている．

(i) $\varphi(0)=1$, $|\varphi(t)|\le 1$, $\varphi(-t)=\overline{\varphi(t)}$, $t\in\mathbf{R}$.

(ii) φ は一様連続関数.

(iii) Positive definiteness, 即ち，任意の自然数 n，複素数 z_k , $k=1,2,\cdots,n$, と $t_k\in\mathbf{R}$, $k=1,\cdots,n$, に対して $\displaystyle\sum_{j,k=1}^{n} z_j\bar{z}_k\varphi(t_j-t_k)\ge 0$.

[20] [9] に引用されている切断不等式は見た目が違うが，少し移項すると同内容であることが分かる．

なお [26, 定理 3.2.2] では切断不等式の代わりに tightness の十分条件として $\displaystyle\lim_{\epsilon\downarrow 0}\sup_n \int (1-e^{-\epsilon|x|^2})\mu_n(dx)=0$ [26, 補題 3.2.1] を用いる．

証明. 証明の方針．g が有界連続可積分関数ならば $\displaystyle\int\int \varphi(t-s)g(t)\bar{g}(s)dt\,ds$ はリーマン和で近似できるので第 3 条件から非負．特に $g(t) = e^{-2|t|^2/v-\sqrt{-1}tx}$ とすれば，$\displaystyle f_n(x) = \frac{1}{2\pi}\int \varphi(t)e^{-|t|^2/n-\sqrt{-1}tx}$ が非負で積分 1，即ち分布の密度になっていることが分かる．($\displaystyle\int_{\mathbf{R}}(1-\cos t)\frac{dt}{d^2} = 1$ を使う.）その特性関数は $n \to \infty$ で φ に収束し，仮定から連続だから定理 35 が使える． □

3 独立確率変数列．

3.1 準備．

3.1.1 独立という概念の「気持ち」．

独立という概念は，気持ちの上では，一方の情報から他方についての情報が得られない，ことをいう．

例えば，喫煙と寿命は独立でないとすると，喫煙者であるという情報があれば保険料を上げるべきだし，喫煙者でない人の保険料は下げるのが適当である．また，去年の宝くじと今年の宝くじは独立だとすると,「当たりくじが出ました」の看板を見ても，そこで宝くじを買うのが得策にもならないし損にもならないことになる．この「気持ち」を定式化[21] したものが確率論における独立の概念になる．

上記は 2 つの事象 A, B が独立であることを，一方の事象 A が生じる確率がもう一方 B も起きるかどうかによらないことで定義することになる．既に知っている（または容易に分かる）ように，自然な定義は

$$\mathrm{P}[\,A \mid B\,] = \frac{\mathrm{P}[\,A\cap B\,]}{\mathrm{P}[\,B\,]} = \mathrm{P}[\,A\,]$$

であるが，これを書き換えると $\mathrm{P}[\,A\cap B\,] = \mathrm{P}[\,A\,]\mathrm{P}[\,B\,]$ となる．これを確率変数の分布

$$\mathrm{P}[\ \cdot\ ] \ \mapsto\ \mathrm{P}\circ(X_1,X_2)^{-1}(\cdot) = \mathrm{P}[\,(X_1,X_2)\in\cdot\,],$$

に適用すれば確率変数列の独立性の定義に至る．

他方，この定義から A, B が独立ならば A, B^c が独立になる．σ 加法族は補集合について閉じていることを思い出せば，独立性の本質は事象や確率変数ではなく σ 加法族にあることが分かる．そこで事象の独立性が定義される．

事象の独立性の定義から**情報としての σ 加法族**という視点が明確になる．この視点は確率過程の一般論の定式化の際に重要になる,

無限個の独立確率変数を考えるということは無限次元空間上の関数を考えていることになる．無限次元空間の上の解析は 20 世紀以降の重要な研究課題なので，無限個の確率変数の解析は重要である．その中で独立確率変数列は確率論にとって分かりやすい（解析しやすい）無限次元という，研究の出発点や計算できる具体例としての重要性がある．

---

[21] well-defined に書くこと．

3.2 独立性．

3.2.1 事象の独立性．

確率空間 $(\Omega,\mathcal{F},\mathrm{P})$ の事象 $A,B\in\mathcal{F}$ が**独立**とは

$$\mathrm{P}[\,A\cap B\,]=\mathrm{P}[\,A\,]\mathrm{P}[\,B\,] \tag{15}$$

となることをいう．

事象が独立であること (15) は

$$\mathrm{P}[\,A\cap B^c\,]=\mathrm{P}[\,A\,]\mathrm{P}[\,B^c\,] \tag{16}$$

と同値である．

3.2.2 確率変数の独立性．

確率空間 $(\Omega,\mathcal{F},\mathrm{P})$ 上の実確率変数 $X_i,\ i=1,2,\cdots,n,$ が**独立**とは

$$\mathrm{P}[\,X_1\in A_1,\cdots,X_n\in A_n\,]=\prod_{i=1}^n \mathrm{P}[\,X_i\in A_i\,],\quad A_i\in\mathcal{B}_1,\ i=1,2,\cdots,n. \tag{17}$$

が成り立つことをいう．

いままで分布は $(\mathbf{R},\mathcal{B}_1)$ 上の確率測度のこととしてきたが，ここで一般化して $(\mathbf{R}^n,\mathcal{B}_n)$ 上の確率測度のことも分布 (n 次元分布) ということにする．分布に関する既出の定理は殆どが自明の拡張があるが，詳細は略す．X_i の分布 $\mu_i=\mathrm{P}\circ X_i^{-1},\ i=1,\cdots,n,$ (5) と結合分布（$X=(X_1,\cdots,X_n)$ の分布）$\mu=\mathrm{P}\circ X^{-1}:\ \mathcal{B}_n\to[0,1]$ を用いれば (17) は

$$\mu(A_1\times A_2\times\cdots A_n)=\prod_{i=1}^n \mu_i(A_i),\quad A_i\in\mathcal{B}_1,\ i=1,2,\cdots,n. \tag{18}$$

とも書ける．即ち，

$X_1,\cdots,X_n$ が独立であることと，その結合分布 μ が X_i の分布 μ_i たちの直積測度 (§2.4.3) であることは同値である．

X_i たちが独立で $f_i:\ \mathbf{R}\to\mathbf{R}$ が可測関数ならば，$f_i(X_i)$ たちも独立になる．なぜなら，

$$\begin{aligned}&\mathrm{P}[\,f_1(X_1)\in A_1,f(X_2)\in A_2\,]=\mathrm{P}[\,X_1\in f_1^{-1}(A_1),X_2\in f_2^{-1}(A_2)\,]\\&=\prod_{i=1}^2\mathrm{P}[\,X_i\in f_i^{-1}(A_i)\,]=\prod_{i=1}^2\mathrm{P}[\,f_i(X_i)\in A_i\,]\end{aligned}$$

だから．

X_i たちが d 次元確率変数であっても，さらに，i によって次元が違っても全く同様に定義する．

確率変数の無限列の独立性は有限部分列の独立性で定義する．

即ち，$X_1,\cdots$ が独立であるとは，任意の $n\in\mathbf{N}$ に対して $X_1,\cdots,X_n$ が独立であることをいう．

3.2.3 σ 加法族の独立性．

$(\Omega,\mathcal{F},\mathrm{P})$ を確率空間とする．

$\mathcal{F}$ の部分 σ 加法族 $\mathcal{A}_i,\ i=1,2,\cdots,n$ が**独立**であるとは，

$$\mathrm{P}[\ \bigcap_{i=1}^{n} A_i\ ] = \prod_{i=1}^{n} \mathrm{P}[\ A_i\ ],\quad A_i \in \mathcal{A}_i,\ i=1,2,\cdots,n,$$

が成り立つことをいう．部分 σ 加法族の無限列が独立とはその任意の部分有限列が独立なことをいう．

事象の独立性，確率変数の独立性との関係． $\sigma[X] = \{X^{-1}(A) \mid A \in \mathcal{B}_1\}$ は σ 加法族になる．これを X が生成する（作る）σ 加法族という．

例：定数関数が作る σ 加法族は最小の σ 加法族 $\{\emptyset,\Omega\}$ であり，χ_A が作る σ 加法族は $\{\emptyset, A, A^c, \Omega\}$ である．

事象，確率変数，有限加法族の独立性は全て，それらが生成する σ 加法族の独立性の言葉で書ける．

命題 37 *(i)* 事象 $A,\ B$ が独立であることと $\sigma[A]$ と $\sigma[B]$ が独立であることは同値である．

(ii) 確率変数 $X_i,\ i=1,2,\cdots,$ （有限列または無限列））が独立であることと $\sigma[X_i],\ i=1,2,\cdots,$ が独立であることは同値である．

(iii) 有限加法族の列 $\mathcal{A}_i,\ i=1,2,\cdots,$ （有限列または無限列）が独立であることと $\sigma[\mathcal{A}_i],\ i=1,2,\cdots,$ が独立であることは同値である．

証明. (i) と (ii) は定義を直接確かめればよいので各自の演習とする．

(iii) は σ 加法族の有限加法族による近似定理定理 2 を要する．

$F_i \in \sigma[\mathcal{A}_i],\ i=1,2,\cdots,$ とする．定理 2 から，任意の $\epsilon > 0$ に対して $\mathrm{P}[\ F_i \triangle A_i\ ] < \epsilon 2^{-n}$ なる $A_i \in \mathcal{A}_i$ がある．このとき

$$\left|\mathrm{P}[\ \bigcap_{i=1}^{n} F_i\ ] - \mathrm{P}[\ \bigcap_{i=1}^{n} A_i\ ]\right| \le \mathrm{P}[\ \left(\bigcap_{i=1}^{n} F_i\right) \triangle \left(\bigcap_{i=1}^{n} A_i\right)\ ] \le \sum_{i=1}^{n} \mathrm{P}[\ F_i \triangle A_i\ ] < 2^{-1}\epsilon,$$
$$\left|\prod_{i=1}^{n} \mathrm{P}[\ F_i\ ] - \prod_{i=1}^{n} \mathrm{P}[\ A_i\ ]\right| \le \sum_{i=1}^{n} |\mathrm{P}[\ F_i\ ] - \mathrm{P}[\ A_i\ ]| \le \sum_{i=1}^{n} \mathrm{P}[\ F_i \triangle A_i\ ] < 2^{-1}\epsilon,$$

$A_i,\ i=1,\cdots,n,$ の独立性を使うと

$$\left|\mathrm{P}[\ \bigcap_{i=1}^{n} F_i\ ] - \prod_{i=1}^{n} \mathrm{P}[\ F_i\ ]\right| < \epsilon$$

を得るが，$\epsilon > 0$ は任意なので $F_i,\ i=1,\cdots,n,$ の独立性を得る． □

注. (iii) で有限加法族であることは本質である．一般には集合族 $\mathcal{A}_1$ と $\mathcal{A}_2$ が独立でも $\sigma[\mathcal{A}_1]$ と $\sigma[\mathcal{A}_1]$ が独立とは限らない．例えば， $\mathcal{A}_1 = \{A\},\ \mathcal{A}_2 = \{B, C\},\ A = \{(1,1,2),(1,2,1),(2,1,1),(2,2,2)\}$, $B = \{(1,1,2),(1,2,1),(2,1,2),(2,2,1)\},\ C = \{(1,1,1),(1,1,2)\},\ \mathrm{P}[\ (i,j,k)\ ] = 1/8,\ i,j,k \in \{1,2\}$.
◇

命題 37 の意味で次の標語を得る．

σ 加法族の独立性の定義は事象の（初等的な意味の）独立性の定義の拡張であり，確率変数列の独立性は確率変数が作る σ 加法族の独立性である．この意味で独立性という情報を持っているのは σ 加法族である．

発展：独立性に関するその他の注．

(i) 独立な確率変数列は存在する．（そのような対象を想定しても論理上無内容ではない，ということ．）例えば，直積測度空間 $(\prod_{i=1}^{\infty}\mathbf{R}, \prod_{i=1}^{\infty}\mathcal{B}_1, \prod_{i=1}^{\infty}\mu_i)$ 上で第 i 成分への射影を $\pi_i :\ \mathbf{R}^{\infty} \to \mathbf{R}$ とするとき，$X_i = \pi_i,\ i = 1, 2, \cdots,$ とおけばよい．

(ii) 無限列の独立性はその任意の有限部分列の独立性で定義するが，有限列の独立性はその任意の部分列の独立性より強い概念である．例えば [9, §1.3 例 1]，$\Omega = \{1, 2, 3, 4\}$, $\mathcal{F} = 2^{\Omega}$, $\mathrm{P}[\ i\ ] = \frac{1}{4}$, $i = 1, 2, 3, 4,$ で定義された確率空間 $(\Omega, \mathcal{F}, \mathrm{P})$ において，$A = \{1, 2\}$, $B = \{2, 3\}$, $C = \{3, 1\}$, とおくとき，$\chi_A, \chi_B, \chi_C,$ はどの 2 つも互いに独立な実確率変数だが $\chi_A, \chi_B, \chi_C,$ は独立ではない．

3.2.4 独立性に関するよく知られた性質．

定理 38 (乗法定理) $n \in \mathbf{N}$ に対して確率変数 $X_k,\ k = 1, \cdots, n,$ を独立とする．各 X_k が可積分 ($\mathrm{E}[\ |X_k|\ ] < \infty$) ならば $\prod_{k=1}^{n} X_k$ も可積分で $\mathrm{E}[\ \prod_{k=1}^{n} X_k\ ] = \prod_{k=1}^{n} \mathrm{E}[\ X_k\ ]$.

証明. $n = 2$ のときと，ある n のとき成り立っていれば，$\prod_{k=1}^{n} X_k$ と X_{n+1} は独立でそれぞれ可積分だから $n = 2$ の場合を適用して $n + 1$ でも成り立つ．よって $n = 2$ のときを証明すればよい．

$n = 2$ のとき，確率変数が独立とはその結合分布が各確率変数の分布の直積測度に等しいことだから，Fubini の定理より

$$\mathrm{E}[\ |X_1X_2|\ ] = \int |x_1x_2|\mathrm{P}_{X_1}[\ dx_1\ ]\mathrm{P}_{X_2}[\ dx_2\ ] = \mathrm{E}[\ |X_1|\ ]\mathrm{E}[\ |X_2|\ ] < \infty$$

でまず積の可積分性が得られ，

$$\mathrm{E}[\ X_1X_2\ ] = \int x_1x_2\mathrm{P}_{X_1}[\ dx_1\ ]\mathrm{P}_{X_2}[\ dx_2\ ] = \mathrm{E}[\ X_1\ ]\mathrm{E}[\ X_2\ ]$$

から (可積分性も含めて) 成り立つ [9, §4.2 定理 4 p.88]. □

注. 独立でなければそれぞれ可積分でも積が可積分とは限らない [9, §4.2 定理 4 p.88] ので，可積分性の主張も無内容ではない．例： $\mathrm{P}[\ X = n\ ] = cn^{-3},\ n \in \mathbf{N},\ Y = X,$ のとき，$\mathrm{E}[\ X\ ] = \mathrm{E}[\ Y\ ] = c\sum_{n\in\mathbf{N}} n^{-2} < \infty$ だが，$\mathrm{E}[\ XY\ ] = c^2 \sum_{n\in\mathbf{N}} n^{-1} = \infty$. ◇

系 39 確率変数 $X_k,\ k = 1, \cdots, n,$ が独立で，それぞれ 2 次までのモーメントが有限（期待値が定義できるので，*implicit* に可積分 $\mathrm{E}[\ |X_k|\ ] < \infty$ も仮定していることに注意）ならば，$\mathrm{V}[\ \sum_{k=1}^{n} X_k\ ] = \sum_{k=1}^{n} \mathrm{V}[\ X_k\ ]$.

証明. 分散の定義と乗法定理より明らか [9, §4.2 定理 5 p.88]. □

定理 40 (Kac の定理) 確率変数 $X_k,\ k=1,\cdots,n,$ が独立であることと特性関数の乗法性

$$\mathrm{E}[\ e^{\sqrt{-1}\sum_{j=1}^n X_j t_j}\ ] = \prod_{j=1}^n \mathrm{E}[\ e^{\sqrt{-1}X_j t_j}\ ],\quad t_j\in\mathbf{R},\ j=1,2,\cdots,n,$$

が成り立つことは同値.

注. 独立性から特性関数の乗法性は自明 [9, §5.1 ϕ 9]. ◇

証明. 定義で注意したように，X_i たちが独立で $f_i:\ \mathbf{R}\to\mathbf{R}$ たちが可測関数ならば $f_i(X_i)$ たちも独立なので，定理 38 より特に特性関数の乗法性を得る．

逆は [26, 定理 4.1.9] に従う．$\mathbf{R}^n$ の分布を議論するので概略のみとする．

$X=(X_1,\cdots,X_n)$ の分布を μ，X_k の分布を μ_k とする．仮定から

$$\int e^{\sqrt{-1}\sum_{j=1}^n x_j t_j}\mu(dx_1\times\cdots\times dx_n) = \mathrm{E}[\ e^{\sqrt{-1}\sum_{j=1}^n X_j t_j}\ ] = \prod_{j=1}^n \mathrm{E}[\ e^{\sqrt{-1}X_j t_j}\ ]$$
$$= \prod_{j=1}^n \int e^{\sqrt{-1}x_j t_j}\mu_j(dx_j) = \int e^{\sqrt{-1}\sum_{j=1}^n x_j t_j}\mu_1\times\cdots\times\mu_n(dx_1\cdots dx_n).$$

最後の変形は Fubini の定理，最後の測度は直積測度．特性関数が一致するから測度が一致する (系 29 の $\mathbf{R}^n$ 版)．結合分布が分布の直積だから，(18) 以下によって $\{X_k\}$ は独立. □

独立確率変数の和の分布は（各確率変数の）分布のたたみこみ (convolution) で表される．

命題 41 $X_i:\ \Omega\to\mathbf{R},\ i=1,2,$ を独立な確率変数とし，その分布と特性関数をそれぞれ，$\mu_i=\mathrm{P}\circ X_i,\ \varphi_i(t)=\mathrm{E}[\ e^{\sqrt{-1}X_i t}\ ]=\int_{\mathbf{R}} e^{\sqrt{-1}xt}\mu_i(dx),\ t\in\mathbf{R},\ i=1,2,$ とする．このとき，確率変数 $Y=X_1+X_2$ の特性関数 φ と分布 μ は，$\varphi(t)=\varphi_1(t)\varphi_2(t),\ t\in\mathbf{R},$ と

$$\mu(A)=\int_{\mathbf{R}}\mu_1(A-x)d\mu_2(x)=\int_{\mathbf{R}}\mu_2(A-x)d\mu_1(x),\ \ A\in\mathcal{B}_1, \tag{19}$$

である．

注. (19) で定義される μ を μ_1 と μ_2 の畳み込みという. ◇

証明. 特性関数については定理 40(のやさしい向き) で注意したとおり，乗法定理から当然である．

測度については，(18) に戻って Fubini の定理を用いると，

$$\begin{aligned}\mu(A) &= \int_{x_1+x_2\in A} d\mu((x_1,x_2)) = \int_{\mathbf{R}^2}\chi_A(x_1+x_2)d\mu((x_1,x_2))\\ &= \int_{\mathbf{R}}\left(\int_{\mathbf{R}}\chi_{A-x_2}(x_1)d\mu_1(x_1)\right)d\mu_2(x_2)\\ &= \int_{\mathbf{R}}\mu_1(A-x)d\mu_2(x).\end{aligned}$$

（1 と 2 を入れ替えた式も Fubini から当然得られる.） □

Borel–Cantelli の第 1 定理命題 24 は事象列 $\{A_k\}$ に対して

$$\sum_{k=1}^{\infty}\mathrm{P}[\ A_k\ ]<\infty \implies \mathrm{P}[\ \overline{\lim_{k\to\infty}}\ A_k\ ]=0$$

を主張するが，これは無条件に成り立った．第 2 定理は $\displaystyle\sum_{k=1}^{\infty} \mathrm{P}[\, A_k \,] = \infty$ の場合を扱うという意味でこれと相補的だが，無条件には成り立たない．

定理 42 (Borel–Cantelli の第 2 定理) **独立な**事象列 $\{A_k\}$ に対して

$$\sum_{k=1}^{\infty} \mathrm{P}[\, A_k \,] = \infty \implies \mathrm{P}[\, \varlimsup_{k\to\infty} A_k \,] = 1\,.$$

証明. 確率の連続性と独立性，および条件式を順に使えば

$$\begin{aligned} \mathrm{P}[\, \bigcap_{k=n}^{\infty} A_k^c \,] &= \lim_{N\to\infty} \mathrm{P}[\, \bigcap_{k=n}^{N} A_k^c \,] = \lim_{N\to\infty} \prod_{k=n}^{N} (1 - \mathrm{P}[\, A_k \,]) \\ &\le \lim_{N\to\infty} \exp(-\sum_{k=n}^{N} \mathrm{P}[\, A_k \,]) = 0\,. \end{aligned}$$

よって

$$\mathrm{P}[\, \left(\varlimsup_{k\to\infty} A_k\right)^c \,] = \lim_{n\to\infty} \mathrm{P}[\, \bigcap_{k=n}^{\infty} A_k^c \,] = 0\,.$$

□

3.3 独立確率変数列の和の極限定理．

この節では断らなければ確率空間 $(\Omega, \mathcal{F}, \mathrm{P})$ 上の独立確率変数列 $X_n,\ n \in \mathbf{N}$, の部分和の平均値の列 $\displaystyle\frac{1}{n}\sum_{k=1}^{n} X_k,\ n \in \mathbf{N}$, を考える．

独立確率変数列の和の極限定理：

(i) 大数の法則
(ii) 中心極限定理
(iii) …

3.3.1 大数の法則 ー 簡単な場合．

大数の法則：期待値の等しい独立確率変数列 $X_n,\ n \in \mathbf{N}$, の n 個の平均は $n \to \infty$ で期待値に収束する．但し，適当な条件の下で．
概収束を示したとき大数の強法則，確率収束しか示さなければ大数の弱法則，という．

十分条件を与えるのが大数の法則という定理．弱収束だと，一つのサンプルに対しては何も言えない．現実の世界（テストを何度も受ける，何度も宝くじを買う）はたった一つのサンプル．これについて言える強法則のほうが極限定理の理念としては適切（強い結果）．

あきらかに $\displaystyle\frac{1}{n}\sum_{k=1}^{n} X_k \to a,\ n \to \infty$, と $\displaystyle\frac{1}{n}\sum_{k=1}^{n} (X_k - a) \to 0,\ n \to \infty$, は（概収束，確率収束，法則収束いずれでも）同値なので，$\mathrm{E}[\, X_k \,] = 0$ として十分なので，主にそういう形で定理を書く．応用上は注意.

「真実」としては確率変数の分散が列全体で有界ならば十分である．これまでの講義の範囲で簡単に証明できる [10, §10 定理 13] のは

(i) 大数の弱法則は分散有界の条件があれば，チェビシェフの不等式だけで証明できる．
(ii) 十分強い条件をおけば優収束定理だけで大数の強法則を証明できる．
(iii) 分散有界の条件があれば，マルチンゲールの L^2 maximal inequality の簡易版で大数の強法則を証明できる．

定理 43 (大数の弱法則，大数の強法則 (簡単な場合)) 独立確率変数列 $X_n,\ n \in \mathbf{N}$, の各 X_n の分布が全て X_1 と等しく[22]， $\mathrm{E}[\ X_1\ ]=0,\ \mathrm{E}[\ X_1^2\ ]=\sigma^2<\infty$ とする．このとき，

(i) $\dfrac{1}{n}\displaystyle\sum_{k=1}^{n} X_k$ は 0 に確率収束する．
(ii) 実は概収束もする．
(iii) さらに $\mathrm{E}[\ X_1^4\ ]=\delta^2<\infty$ ならば， $\dfrac{1}{n}\displaystyle\sum_{k=1}^{n} X_k$ は 0 に概収束する．

証明. (i) $W_n=\displaystyle\sum_{k=1}^{n} X_k$ とおき，$\epsilon>0$ を任意にとる．乗法定理と $\mathrm{E}[\ X_k\ ]=0,\ k\in\mathbf{N}$, より，

$$\mathrm{E}[\ \left(\frac{1}{n}W_n\right)^2\ ]=\frac{1}{n^2}\sum_{j,k=1}^{n}\mathrm{E}[\ X_jX_k\ ]=\frac{1}{n^2}\sum_{j\neq k}\mathrm{E}[\ X_j\ ]\mathrm{E}[\ X_k\ ]+\frac{1}{n^2}\sum_{j=1}^{n}\mathrm{E}[\ X_j^2\ ]=\frac{1}{n}\sigma^2. \tag{20}$$

チェビシェフの不等式から

$$\epsilon^2\mathrm{P}[\ \left|\frac{1}{n}W_n\right|\geq\epsilon\ ]\leq\mathrm{E}[\ \left(\frac{1}{n}W_n\right)^2\ ]$$

なので

$$\lim_{n\to\infty}\mathrm{P}[\ \left|\frac{1}{n}W_n\right|\geq\epsilon\ ]=0\,.$$

よって $\dfrac{1}{n}W_n$ は 0 に確率収束する．

(ii) $W_n=\displaystyle\sum_{k=1}^{n} X_k$ とおく．(20) から，

$$\sum_{n=1}^{\infty}\mathrm{E}[\ \left(\frac{1}{n^2}W_{n^2}\right)^2\ ]=\sum_{n=1}^{\infty}\frac{1}{n^2}\sigma^2<\infty.$$

なので，$\displaystyle\sum_{n=1}^{\infty}\left(\frac{1}{n^2}W_{n^2}\right)^2<\infty$, a.s., したがって $\displaystyle\lim_{n\to\infty}\frac{1}{n^2}W_{n^2}=0$, a.s..
$M_n=\displaystyle\max_{n^2<k\leq(n+1)^2}|W_k-W_{n^2}|$ とおくと，$n^2<k\leq(n+1)^2$ に対して

$$\left|\frac{1}{k}W_k\right|\leq\frac{1}{k}(M_n+|W_{n^2}|)\leq\frac{1}{n^2}M_n+\frac{1}{n^2}|W_{n^2}|$$

となるから $\displaystyle\lim_{n\to\infty}\frac{1}{n^2}M_n=0$, a.s., を証明すればよい．
(20) と同様に，$k>n^2$ のとき，

$$\mathrm{E}[\ (W_k-W_{n^2})^2\ ]=\sum_{j,k=n^2+1}^{k}\mathrm{E}[\ X_jX_k\ ]=\sum_{j\neq k}\mathrm{E}[\ X_j\ ]\mathrm{E}[\ X_k\ ]+\sum_{j=n^2+1}^{k}\mathrm{E}[\ X_j^2\ ]=\sigma^2\,(k-n^2)$$

となることに注意．加えて，$M_n^2\leq\displaystyle\sum_{n^2<k\leq(n+1)^2}(W_k-W_{n^2})^2$ に注意すると

$$\mathrm{E}[\ M_n^2\ ]\leq\sum_{n^2<k\leq(n+1)^2}\mathrm{E}[\ (W_k-W_{n^2})^2\ ]=\sigma^2\sum_{n^2<k\leq(n+1)^2}(k-n^2)\leq\sigma^2\,(2n+1)^2.$$

[22] この状況を「i.i.d. (independent, identical distribution) 確率変数列」と略称することが多い．

よって，

$$\sum_{n=1}^{\infty} \mathrm{E}[\ \left(\frac{M_n}{n^2}\right)^2\ ] \le \sum_{n=1}^{\infty} \frac{(2n+1)^2}{n^4}\sigma^2 < \infty$$

だから，先ほどと同様の議論で， $\displaystyle\lim_{n\to\infty} \frac{1}{n^2} M_n = 0$, a.s., を得る．よって概収束もする．

(iii) (7) から $|\mathrm{E}[\ X_1^3\ ]| \le \mathrm{E}[\ X_1^4\ ]^{3/4} < \infty$ に注意しておく．（このことから，W_n^4 を展開したときの $\mathrm{E}[\ X_i^3 X_j\ ]$ の形の項が独立性で期待値の積に分けられて結果として 0 とできる.）

乗法定理と $\mathrm{E}[\ X_k\ ] = 0$ から

$$\mathrm{E}[\ \left(\frac{1}{n} W_n\right)^4\ ] = \frac{1}{n^4}\left(\sum_{k=1}^{n} \mathrm{E}[\ X_k^4\ ] + 3\sum_{k=1}^{n} \sum_{j\ne k, 1\le j\le n} \mathrm{E}[\ X_k^2\ ]\mathrm{E}[\ X_j^2\ ]\right) = \frac{\delta^2}{n^3} + 3(n-1)\frac{\sigma^4}{n^3}$$

となるので，単調収束定理から

$$\mathrm{E}[\ \sum_{n=1}^{\infty} \left(\frac{1}{n} W_n\right)^4\ ] = \sum_{n=1}^{\infty} \mathrm{E}[\ \left(\frac{1}{n} W_n\right)^4\ ] < \infty.$$

期待値の定義から

$$\sum_{n=1}^{\infty} \left(\frac{1}{n} W_n\right)^4 < \infty,\ a.e.,$$

となるがこれは,

$$\lim_{n\to\infty} \left(\frac{1}{n} W_n\right)^4 = 0,\ a.e.,$$

即ち,

$$\lim_{n\to\infty} \frac{1}{n} W_n = 0,\ a.e.,$$

を意味する．

□

3.3.2 コルモゴロフ不等式による大数の強法則の証明．

分散有界の条件だけで大数の強法則を証明するには，マルチンゲール不等式 (§C.1.3) の特別な場合にあたるコルモゴロフの不等式を用いる．

この節では X_k たちは独立だが，同分布でなくてもよい．

定理 44 (Kolmogorov の不等式) 自然数 n に対して，$X_k,\ k = 1, \cdots, n,$ が独立で，$\mathrm{E}[\ X_k\ ] = 0$, $\mathrm{E}[\ X_k^2\ ] < \infty,\ k = 1, \cdots, n,$ とする．このとき，

$$\mathrm{P}[\ \max_{k=1,\cdots,n} \left|\sum_{i=1}^{k} X_i\right| \ge a\ ] \le \frac{1}{a^2}\mathrm{E}[\ \left(\sum_{k=1}^{n} X_k\right)^2\ ],\quad a > 0\,.$$

注. $n = 1$ のときは単なるチェビシェフの不等式だが，証明を見れば分かるように， $n \ge 2$ ではもっと深い内容が入っている．

他方，独立性は $\mathrm{E}[\ W_k\, W_n;\ \Lambda_k\ ] = \mathrm{E}[\ W_k^2;\ \Lambda_k\ ]$ の形でしか使わない．これは W_n のマルチンゲール性があれば十分である． ◇

証明. $W_k = \sum_{i=1}^{k} X_i,\ k = 1, \cdots, n$，および，$\Lambda = \{\omega \in \Omega \mid \max_{k=1,\cdots,n} |W_k(\omega)| \geq a\}$，とおいて，$|W_k(\omega)|$ が初めて a 以上になる k によって Λ を分割する：

$$\Lambda = \bigcup_{k=1}^{n} \Lambda_k; \quad \Lambda_k = \{\omega \in \Omega \mid |W_1(\omega)| < a, \cdots, |W_{k-1}(\omega)| < a,\ |W_k(\omega)| \geq a\}.$$

($j \neq k$ ならば $\Lambda_j \cap \Lambda_k = \emptyset$ に注意．)

$$\begin{aligned}
&\mathrm{E}[\, W_n^2 \,] \geq \mathrm{E}[\, W_n^2;\ \Lambda \,] = \sum_{k=1}^{n} \mathrm{E}[\, W_n^2;\ \Lambda_k \,] \\
&= \sum_{k=1}^{n} \mathrm{E}[\, W_k^2;\ \Lambda_k \,] + 2\sum_{k=1}^{n} \mathrm{E}[\, W_k\, \chi_{\Lambda_k}\, (X_{k+1} + \cdots + X_n)\,] + \sum_{k=1}^{n} \mathrm{E}[\, (X_{k+1} + \cdots + X_n)^2;\ \Lambda_k \,] \\
&\geq a^2 \sum_{k=1}^{n} \mathrm{P}[\, \Lambda_k \,] + 2\sum_{k=1}^{n} \mathrm{E}[\, W_k\, \chi_{\Lambda_k}\, (X_{k+1} + \cdots + X_n)\,].
\end{aligned}$$

において，$W_k\, \chi_{\Lambda_k}$ は $\mathrm{E}[\, X_k^2 \,] < \infty$ よりその積分は有限であって，$X_1, \cdots, X_k$ の可測関数なので $X_{k+1} + \cdots + X_n$ と独立だから，

$$\mathrm{E}[\, W_k\, \chi_{\Lambda_k}\, (X_{k+1} + \cdots + X_n)\,] = \sum_{j=k+1}^{n} \mathrm{E}[\, W_k\, \chi_{\Lambda_k} \,]\mathrm{E}[\, X_j \,] = 0\,.$$

よって

$$\mathrm{E}[\, W_n^2 \,] \geq a^2 \mathrm{P}[\, \Lambda \,].$$

□

定理 45 $X_n,\ n \in \mathbf{N}$, が独立で，$\mathrm{E}[\, X_n \,] = 0,\ n \in \mathbf{N}$, かつ $\sum_{n=1}^{\infty} \mathrm{V}[\, X_n \,] < \infty$ を満たせば，$W_n = \sum_{k=1}^{n} X_k$ は概収束（かつ $p \geq 1$ に対して L^p 収束）する．

証明. $v_n = \mathrm{V}[\, X_n \,],\ n \in \mathbf{N}$, とおくと $\mathrm{E}[\, X_n \,] = 0$ だから $n \geq m$ に対して $\mathrm{E}[\, (W_n - W_m)^2 \,] = \sum_{k=m+1}^{n} v_k$. 増加列 $p_n, n \in \mathbf{N}$, を $\sum_{k=p_n}^{\infty} v_k < 2^{-3n}$ を満たすように選ぶ. $\Lambda_n = \{\sup_{k \geq p_n} |W_k - W_{p_n}| > 2^{-n}\}$ とおくと，コルモゴロフの不等式 定理 44 により，

$$\mathrm{P}[\, \Lambda_n \,] = \lim_{\ell \to \infty} \mathrm{P}[\, \max_{p_n \leq k \leq \ell} |W_k - W_{p_n}| > 2^{-n} \,] \leq 2^{2n} \sum_{k=p_n}^{\infty} v_k < 2^{-n}.$$

ゆえに $\sum_{n=1}^{\infty} \mathrm{P}[\, \Lambda_n \,] < \infty$.

よって Borel–Cantelli の第 1 定理 命題 24 からほとんど全ての $\omega \in \Omega$ に対して $N = N(\omega)$ を選んで $n \geq N$ ならば $\omega \notin \Lambda_n$ とできる．このとき

$$j, k \geq p_n \geq N(\omega) \implies |W_k(\omega) - W_j(\omega)| \leq |W_k(\omega) - W_{p_n}(\omega)| + |W_j(\omega) - W_{p_n}(\omega)| \leq 2^{1-n}.$$

よって，概収束も L^p 収束も成り立つ [9, 6 章 §2 定理 2]. □

定理 45 は級数の各点での収束を言っているだけで，各点での絶対収束を意味しない．例：X_n, $n \in \mathbf{N}$, を i.i.d. で $\mathrm{P}[\, X_1 = 1 \,] = \mathrm{P}[\, X_1 = -1 \,] = \dfrac{1}{2}$ とすると $\sum_{n=1}^{\infty} \dfrac{X_n}{n}$ は概収束する．なぜなら，$X'_n = X_n/n,\ n \in \mathbf{N}$, は独立で，定理 45 の条件を満たすのでその和が概収束する．しかし，もちろん，各項の絶対値をとった級数は各点で発散する．

次の補題は級数に関する初等的な性質だが，大数の強法則の証明に使う．

補題 46 (Kronecker の補題) $\{a_n\}$ を増加正項数列で $\lim_{n\to\infty} a_n = \infty$ なるものとする．$\lim_{n\to\infty} \sum_{k=1}^n \frac{x_k}{a_k}$ が存在すれば

$$\lim_{n\to\infty} \frac{1}{a_n} \sum_{k=1}^n x_k = 0\,.$$

証明. 途中の計算の便宜のため $a_0 = 0$ とおく．$y_k = \sum_{i=1}^k \frac{x_i}{a_i}$ とおくと，

$$\begin{aligned}\sum_{k=1}^n x_k &= \sum_{k=1}^n a_k(y_k - y_{k-1}) = \sum_{k=1}^n \sum_{j=1}^k (a_j - a_{j-1})(y_k - y_{k-1})\\ &= \sum_{j=1}^n \sum_{k=j}^n (a_j - a_{j-1})(y_k - y_{k-1}) = \sum_{j=1}^n (a_j - a_{j-1})(y_n - y_{j-1}).\end{aligned}$$

$\{y_k\}$ は収束するから有界．即ち，$K = \sup_{k\in\mathbf{N}} |y_k| < \infty$．$\{a_j\}$ が非負で増加することにも注意して，$1 \le p < n$ に対して，

$$\left|\sum_{k=1}^n x_k\right| \le 2Ka_p + \sum_{j=p+1}^n (a_j - a_{j-1})|y_n - y_{j-1}| \le 2Ka_p + a_n \max_{p<j\le n} |y_n - y_{j-1}|.$$

$\epsilon > 0$ を任意にとる．$\{y_k\}$ が収束するから，$\sup_{m,n>p} |y_n - y_m| < \epsilon$ を満たす $p = p(\epsilon)$ が存在する．さらに $\lim_{n\to\infty} a_n = \infty$ から $a_n > 2Ka_p/\epsilon,\ n \ge n_0$, を満たす $n_0 = n_0(\epsilon)$ が存在する．$n \ge n_0$ のとき，

$$\left|\frac{1}{a_n}\sum_{k=1}^n x_k\right| \le 2\epsilon$$

となるから，主張を得る [9, 6 章 §2 定理 3 内補題]． □

定理 47 (コルモゴロフの大数の強法則) $X_n,\ n \in \mathbf{N}$, を独立確率変数列（同分布でなくてもよい）とする．$\mathrm{E}[\,X_n\,] = 0,\ n \in \mathbf{N}$, かつ $\sum_{n=1}^\infty \frac{\mathrm{E}[\,X_n^2\,]}{n^2} < \infty$ ならば $\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^n X_k = 0,\ a.e.$.

特に，独立確率変数列（同分布でなくてもよい） $X_n,\ n \in \mathbf{N}$, が $\mathrm{E}[\,X_n\,] = a,\ n \in \mathbf{N}$, および $\sup_{n\in\mathbf{N}} \mathrm{V}[\,X_n\,] < \infty$ を満たせば，$\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^n X_k = a,\ a.e.$.

証明. $X'_n = \frac{1}{n} X_n,\ n \in \mathbf{N}$, とおけば $\{X'_n\}$ は定理 45 の条件を満たし，その結論に補題 46 を使えば主張に至る（後半は前半から明らか）． □

同じ準備の下で，i.i.d. 確率変数列の場合には，大数の強法則成立の条件を精密化して，必要十分条件にまで持っていくことができる．ただし，証明が少し長いので，[9, 6 章 §2 定理 4] に任せて証明は省略する．

定理 48 (Khintchin) $X_n,\ n \in \mathbf{N}$, を $i.i.d.$ とする．$1 \le p < 2$ に対して $\lim_{n\to\infty} \frac{1}{n^{1/p}} \sum_{k=1}^n (X_k - m) = 0,\ a.e.$, となる必要十分条件は $\mathrm{E}[\,X_1\,] = m$ かつ $\mathrm{E}[\,|X_1|^p\,] < \infty$.

3.3.3 中心極限定理．

大数の法則：独立確率変数の平均値がある定数に殆ど必ず近づく．　→　近づく速さについては何も言っていない．
中心極限定理：分布で見たときの近づく「速さ」．

(i) 平均値の分布の「広がり」が $O(n^{-1/2})$.
(ii) 分布は正規分布．(即ち，正規分布からのずれは $o(n^{-1/2})$.)

[9, 6 章 §3 定理 1] は同分布とは限らない場合に適用可能な証明である．内容的には証明できる準備が整っているが，技術的に多少入り組んでいる印象なので教科書に譲ってここでは省略する[23]．

定理 49 (中心極限定理　ー　i.i.d. の場合) $X_n,\ n \in \mathbf{N}$, を *i.i.d.* 確率変数列で $\mathrm{E}[\,X_1\,] = 0,\ 0 < \mathrm{V}[\,X_1\,] = v < \infty$ とする．このとき，$T_n = \dfrac{1}{\sqrt{n}} \displaystyle\sum_{k=1}^{n} X_k,\ n \in \mathbf{N}$, の分布は平均 0 分散 v の正規分布 $N(0,v)$ に近づく．(正規分布の定義は §*1.2.4* 参照.)

証明. [7, §5.6 定理 5.26] に従う．$g(x) = \dfrac{1}{x^2}(e^{\sqrt{-1}x} - 1 - \sqrt{-1}x + \dfrac{1}{2}x^2),\ x \in \mathbf{R}$, とおく．乗法定理と同分布性および $\mathrm{E}[\,X_1\,] = 0,\ \mathrm{V}[\,X_1\,] = v$ から

$$
\begin{aligned}
\varphi_n(t) &= \mathrm{E}[\,e^{\sqrt{-1}T_n t}\,] = \left(\mathrm{E}[\,e^{\sqrt{-1}t\frac{1}{\sqrt{n}}X_1}\,]\right)^n \\
&= \left(1 + \sqrt{-1}t\frac{1}{\sqrt{n}}\mathrm{E}[\,X_1\,] - \frac{t^2}{2n}\mathrm{E}[\,X_1^2\,] + \frac{t^2}{n}\mathrm{E}[\,X_1^2\,g(t\frac{1}{\sqrt{n}}X_1)\,]\right)^n \\
&= \left(1 - \frac{vt^2}{2n} + \frac{t^2}{n}\mathrm{E}[\,X_1^2\,g(t\frac{1}{\sqrt{n}}X_1)\,]\right)^n .
\end{aligned}
$$

g について，虚部と実部に分けてテーラーの定理を使えば，

$$
|\cos x - (1 - \frac{1}{2}x^2)| \le \frac{x^4}{24} \wedge (2 + \frac{1}{2}x^2), \quad |\sin x - x| \le \left|\frac{x^3}{6}\right| \wedge |1 + x|.
$$

これより，$\lim_{x\to 0} g(x) = 0$ と $|g(x)| \le 1,\ x \in \mathbf{R}$, を得る．$\mathrm{E}[\,X_1^2\,] < \infty$ だから優収束定理から

$$
\lim_{n\to\infty} \mathrm{E}[\,X_1^2\,g(t\frac{1}{\sqrt{n}}X_1)\,] = 0\,.
$$

即ち，

$$
\lim_{n\to\infty} \varphi_n(t) = \lim_{n\to\infty} \left(1 - \frac{vt^2}{2n} + o(n^{-1})\right)^n = e^{-vt^2/2} = \int e^{-x^2/(2v) + \sqrt{-1}xt} dx.
$$

即ち，φ_n は $N(0,v)$ の特性関数に各点収束する．よって定理 31 より，T_n の分布は $N(0,v)$ に近づく． □

3.4 言わなかったこと．

かってな確率空間を与えたとき，仮定された性質を持つ確率変数が存在するかどうかは当然チェックすべきことである．例えば，独立な確率変数が複数個（特に無限個）存在するかどうかにこれまで

[23] 実験や観測で誤差を正規分布であると仮定することが多いが，中心極限定理はその根拠となる．即ち，独立な小さな確率的な影響が集積して目に見える誤差を生む場合は，その誤差の分布は中心極限定理から正規分布に近いと判断できる．そのときは，攪乱要因に独立性は仮定しても同分布性は仮定できないので，このような一般化に意味があると思われる．
　確率論基礎教科書の題材としてはこの一般化はいらない気がするが，確率論自体の興味として教科書の一般化が興味深いのかどうか私は知らない．

答えなかった．収束定理自体は「しかじかの確率変数列があればしかじかの結果が成り立つ」という形なので何ら矛盾はないが,「しかじかの確率変数列」がなければ，定理は全く意味がなくなる．この問題は直積測度空間の構成によって肯定的に解決するが，その詳細は演習とする．空間の構造に立ち入ることなく議論が展開できることが確率論の表現上の「強み」だと個人的に思うからである．（もちろん「立ち入ることなく」といっても，実体がなければ意味がないのだから，実体の存在に対する嗅覚は持っていることが前提で，単に，論文を書くとき，講義をするとき，長い導入を避けることができる，と言っているだけである.）

補遺．

実際の講義では，時間数を考えると，以下の中からいくつかを選んで講義することになる．どの程度講義するかは，どれくらい詳しく（証明をきちんと）やるかに当然よるが，せいぜい一つか二つであろう．むしろ受講生諸君が各自の興味に応じて選んでとりかかるのがよいと思う.

A 母関数（ラプラス変換）と Tauber 型定理．

[15, §VIII.8, §VIII.9, §XIII.5], [18, 19, 20], [23, §2, 3], [7, 4 章], [27, §3.4] などから母関数を用いて確率測度の漸近的振る舞いを調べる一つの考え方[24] を，あまり統一的でなく，とり上げる．本当はランダムウォーク (§B) を含む確率過程の解析 (path の「形状」の確率解析) に強力だが，ここでは入門的な話にとどめてランダムウォークには立ちいらない.

A.1 Tauber 型定理．

$c>0$ を定数とするとき，密度 $\rho(x)=e^{-cx}\chi_{[0,\infty)}$ の指数分布（$\mathbf{R}$ 上の確率測度） $d\mu=\rho\,dx$ を分布に持つ確率変数 X の母関数 $\mathrm{E}[\,e^{tX}\,]=\displaystyle\int e^{tx}d\mu(x)=\frac{1}{c-t}$, $t<c$, は $t=c$ で発散する．この逆の問題を考える．即ち，

母関数 $\mathrm{E}[\,e^{tX}\,]$ が $t<c$ で存在 (可積分) し，$t=c$ で発散するならば $\mathrm{P}[\,X>x\,]$ は $x\to\infty$ で e^{-cx} に近い振る舞いをするだろうか？

母関数 (Laplace 変換)

$$F(\lambda)=\int_0^\infty e^{-\lambda x}\mu(dx)$$

の振る舞いから確率測度 μ の振る舞いを求める定理を Tauber 型定理と言う.

元々は，母関数が原点付近 $(\lambda\downarrow 0)$ でべきの order で振る舞えば，確率測度が遠方 $(x\uparrow\infty)$ でべきの order で振る舞う，という形の定理を指す．[15, §XIII.5], [18, §0] に Hardy–Littlewood の Tauber 型定理および Karamata (1930) による regularly varying function の導入による証明の整理として言及があるのが，この古典的な Tauber 型定理である.

これに対して， [18], [23, §2, 3] などの 20 世紀後半の研究は large deviation principle などとの関連で，母関数の指数 order の振る舞いから確率測度の振る舞い，即ち $\log \mathrm{P}[\,X>x\,]$ の漸近形を与える指数型の Tauber 型定理に比重が移っている．これは熱統計力学の canonical ensemble と micro canonical ensemble の自由エネルギーの対応（Legendre 変換）に相当する order である.

[21] によれば，λ と x について注目する位置に関して次の 3 つに分類されるとのことである.

[24] 確率論で large deviation principle と呼ばれる研究手法の中で頻出．統計物理学で canonical ensemble と micro canonical ensemble の対応などを与える，ルジャンドル変換として知られる方法.

λ	x	
$+0$	$+\infty$	Kohlbecker 1958
$+\infty$	$+0$	de Brujin 1959
$-\infty$	$+\infty$	Kasahara 1978

最後のケースについて §A.3.4 でとりあげる．

A.2　正則変動関数．

A.2.1　定義．

Tauber 型定理の成り立つ自然な関数のクラスとして regularly varying funnction (正則変動関数) という概念が Karamata によって導入された [15, §VIII.8], [22, §1.1].

関数 F が $(+\infty$ で$)$ 正則変動するとは，ある $a>0$ に対して $[a,\infty)$ 上で定義された正値可測関数 $(F\colon\ [a,\infty)\to(0,\infty))$ であって，

$$(\forall\lambda>0)\ \lim_{x\to\infty}\frac{F(\lambda x)}{F(x)}=\lambda^{\rho} \tag{21}$$

が成り立つような実数値 ρ が存在することを言う．このとき ρ を正則変動の指数 (index) と言う．
特に $\rho=0$ のとき，即ち，

$$(\forall\lambda>0)\ \lim_{x\to\infty}\frac{F(\lambda x)}{F(x)}=1, \tag{22}$$

のとき， F は slowly varying function (緩慢変動関数) と言う．

$x\to\infty$ で収束する関数， $\log x, \log\log x$， などは緩慢変動関数である．指数 ρ の正則変動関数に対して $F(x)=x^{\rho}L(x)$ で L を定義することにより，正則変動関数に関する議論は緩慢変動関数の問題に帰着できる．

指数 ρ の正則変動関数の典型例は x^{ρ} である．正則変動しない関数の例としては $e^{\pm x}$, $2+\sin x$ などがある．

A.2.2　ラプラス変換．

命題 50 右連続非減少な関数 $f\colon\ \mathbf{R}\to\mathbf{R}$ は不連続点が高々可算個である．

証明. n,N を自然数として $(-N,N)$ にある高さ $1/n$ 以上の不連続点の個数を $m(n,N)$ とおくと，f は非減少なので $m(n,N)\le(f(N)-f(-N))n<\infty$. よって不連続点の個数 $\displaystyle\sum_{N,n=1}^{\infty}m(n,N)$ は高々可算個. □

実数上の右連続非減少な実数値関数は $\mathbf{R}$ 上のボレル測度を定める．例えば有界な右連続非減少関数が（有限測度，特に確率測度を与えるので）典型例である．

定理 51 $F\colon\ \mathbf{R}\to\mathbf{R}$ を右連続非減少な関数とする．このとき， $\mu((x,y])=F(y)-F(x),\ x<y,$ を満たす $\mathbf{R}$ 上のボレル測度が一意的に存在する．

注. (i) μ を F に対応する Lebesgue–Stieltjes 測度といい，μ に関する積分 $\displaystyle\int\cdot d\mu(x)$ を $\displaystyle\int\cdot dF(x)$ と書く．

(ii)

$$\lim_{x\to-\infty}F(x)=0,\quad \lim_{x\to\infty}F(x)=1 \tag{23}$$

のとき，明らかに対応する Lebesgue–Stiletjes 測度は分布 (１次元ボレル確率測度) になる．このとき F を μ の分布関数という．

(iii) F が $\mathbf{R}_+ = [0,\infty)$ 上で定義されているとき．対応する $\mathbf{R}_+$ 上の測度が構成できるのも同様である．このような場合も定理 51 に含まれているものとして特に断らないことにする．また必要に応じて，断りなく，$F(x) = 0,\ x < 0,$ で $\mathbf{R}$ 上の分布とみなす．

(iv) 逆に，有限なボレル測度 μ に対して $F(x) = \mu((-\infty, x])$ とおけば F は有界右連続非減少関数になる．

◇

証明. 詳しくは教科書に譲るが，本質的に Hopf の拡張定理である．例えば確率測度の場合の [9, §2.4 定理 1]．a, b が F の不連続点でないような $(-\infty, b]$, $(a, b]$, $(a, \infty]$, の形の区間の素な有限和で表される集合全てからなる集合族を $\mathcal{A}$ とすれば，$\mathcal{A}$ は有限加法族で，不連続点は高々可算個だから $\sigma(\mathcal{A}) = \mathcal{B}_1$.

$\mu((a, b]) = F(b) - F(a)$ を自然に拡張して得られる $\mathcal{A}$ 上の集合関数 μ は well-defined であり，測度論における拡張定理の成り立つ条件（非負値性と $\mathcal{A}$ 上の σ 加法性）を満たすことが示せる．実際，前者は明らかで，後者は確率の連続性（$\{A_n\} \subset \mathcal{A}$ が減少列のとき，$\displaystyle\bigcap_{n=1}^{\infty} A_n = \emptyset$ ならば $\displaystyle\lim_{n\to\infty} \mu(A_n) = 0$）で置き換えられる（拡張定理の証明は例えば [9, §2.3 定理 1]）．確率の連続性の対偶を示す証明が [9, §2.4 定理 1] にある．

□

右連続非減少な $\mathbf{R}_+$ 上の関数 F に対応する測度のラプラス変換とは

$$G(t) = \int_0^{\infty} e^{-tx} dF(x) \tag{24}$$

のことをいう ([15, §VIII.6]).

ここで， F が有界ならば $t \geq 0$ で (24) が存在するが，一般にはそうとは限らない．しかし，もしある $t = a$ で収束すれば任意の $t \geq a$ で収束するのは明らか. (24) が $t \geq a$ に対して収束するとき $[a, \infty)$ 上の関数 G を F に対応する測度のラプラス変換と言う．このとき

$$G(t + a) = \int_0^{\infty} e^{-tx} e^{-ax} dF(x), \quad t \geq 0, \tag{25}$$

は有限測度 $e^{-ax}dF(x)$ のラプラス変換であり，従って確率測度 $e^{-ax}dF(x)/G(a)$ のラプラス変換である．このようにして，有限測度または確率測度のラプラス変換に関する結果はラプラス変換の存在する測度に拡張できる（[15, XIII.1] では移動原理と呼んでいる）.

漸近的性質の前に，ラプラス変換が元の測度を特徴づける（一意的に定める）ことを [15, XIII.1] に従って示しておく.

定理 52 ラプラス変換が一致すれば測度は一致する.

一般に有限測度でない場合でもよく，ラプラス変換がある $[a, \infty)$ で収束して一致すれば測度は一致する.

証明. まず， F が (23) を満たすとする．F が定める分布を P とし，$Y(x) = e^{-x}$ で $\mathbf{R}_+$ 上の確率変数を定めると，$Q = \mathrm{P} \circ Y^{-1}$ は $(0, 1]$ 上の分布．P のラプラス変換 $G(t) = \int_0^{\infty} e^{-tx} \mathrm{P}[\,dx\,] = \mathrm{E}[\,Y^t\,]$ だから有界確率変数 Y のモーメントである．ラプラス変換が一致すれば Y のモーメントが一致し

て,（有界だから）特性関数 $\mathrm{E}[\, e^{\sqrt{-1}tY}\,] = \sum_{n=1}^{\infty} \frac{1}{n!} t^n \mathrm{E}[\, Y^n\,]$ が全ての $t \in \mathbf{R}$ で存在して一致し，系 29 から Q も一致するので，$\mathrm{P} = Q \circ Y^{-1^{-1}}$ も一致する．

分布でない場合も (25) の前後の注から分布の場合に帰着する． □

漸近的性質を記述するために弱収束の概念の一般の測度列への拡張を行う．未知の測度への収束を結論したいので， tightness argument を，§2.5.2 の 定理 32 で紹介した Helly の方法で，用意する．以下， [15, §VIII.1, §VIII.6] に従う．(ただし，定理 54 の証明の細部は [9, §5.4 定理 2] から補った.）

右連続非減少な $\mathbf{R}$ 上の関数の列 $\{F_n\}$ が右連続非減少な関数 F に弱収束するとは，$a < b$ を F の連続点とするとき $\lim_{n\to\infty}(F_n(b) - F_n(a)) = F(b) - F(a)$ が成り立つこととする[25]．((23) が F_n, F に対して成り立っていれば $a = -\infty$ だけで十分だが，確率測度の分布関数に収束しない場合はそうはいかない.）

補題 53 *(23)* を満たす関数の列 $F_n :\ \mathbf{R} \to [0,1]$, $n \in \mathbf{N}$, は任意の可算個の実数 $\{a_n\}$ （例えば有理数全体）の上で同時に収束 *(*$\pm\infty$ を許す*)* する部分列を含む．

証明. 対角線論法. □

定理 54 F_n たちは分布関数，即ち，*(23)* を満たす $\mathbf{R}$ 上の右連続非減少関数とする．このとき，ある右連続非減少関数 F に弱収束する部分列がある．

証明. $\{a_n\}$ を有理数全体として 補題 53 を適用すると，全ての有理数で収束する部分列 $\{F_{n(k)}\}$ が存在する．$F_0(a_j) = \lim_{k\to\infty} F_{n(k)}(a_j)$, $j \in \mathbf{N}$, とおき，$F(x) = \inf_{a_j > x} F_0(a_j)$, $x \in \mathbf{R}$, で F を定義すると，自動的に右連続非減少関数になる．

次に，$\{a_{j(k)}\}$ が減少して x に収束するならば

$$F(x) = \lim_{k\to\infty} F_0(a_{j(k)}) \tag{26}$$

となることに注意．なぜなら，任意の $a_p > x$ に対して k を十分大きく取れば $a_p > a_{j(k)}$ となるので， F_0 の単調性から

$$F_0(a_p) \ge F_0(a_{j(k)}) \ge \lim_{k\to\infty} F_0(a_{j(k)}) \ge F(x).$$

ここで $a_p > x$ について下限を取れば (26) を得る．

残っているのは F の連続点 x に対して $\lim_{k\to\infty} F_{n(k)}(x) = F(x)$ を言うこと．連続点なので，任意の $\epsilon > 0$ に対して $F(x) > F(a_p) - \epsilon$ となる $a_p > x$ があるので， $a_j > a_p$ を a_p に十分近く取れば (26) と $F_n(x)$ の x についての単調性より

$$F(x) > F_0(a_j) - 2\epsilon = \lim_{k\to\infty} F_{n(k)}(a_j) - 2\epsilon \ge \varlimsup_{k\to\infty} F_{n(k)}(x) - 2\epsilon.$$

同様に $a_m < x$ を $F(x) < F(a_m) + \epsilon$ を満たすように取れば $a_m < a_j < x$ なるとき

$$F(x) < F_0(a_j) + \epsilon = \lim_{k\to\infty} F_{n(k)}(a_j) + \epsilon \le \varliminf_{k\to\infty} F_{n(k)}(x) + \epsilon.$$

よって $\lim_{k\to\infty} F_{n(k)}(x) = F(x)$. □

注. §2.5.2 の 定理 32 では，このあと，(14) から $x < -K$ ならば $F(x) < \epsilon$，および $x > K$ ならば $F(x) > 1 - \epsilon$, を得て $F(\infty) - F(-\infty) = 1$ （全測度が 1 ということ）を言う． ◇

[25] 本当は次の (27) が成り立つことを弱収束と言うのだが，この講義録内の引用の都合でこちらを弱収束と言うことにする（同値だから問題は起きない）．

定理 55 F_n たちは分布関数，即ち，*(23)* を満たす $\mathbf{R}$ 上の右連続非減少な関数とする．F_n がある右連続非減少関数 F に弱収束することと，任意の $\displaystyle\lim_{x\to\pm\infty} f(x)=0$ を満たす連続関数 f に対して

$$\lim_{n\to\infty}\int f(x)dF_n(x)=\int f(x)dF(x) \tag{27}$$

が成り立つことは同値である．

証明. [15, §VIII.1] による．

F_n が F に弱収束しているとする．$0\le F(x)\le 1$, $x\in\mathbf{R}$, となることに注意．

任意の $\epsilon>0$ に対して，端点が F の連続点であるような十分大きな区間 A をとって，$|f(x)|<\epsilon$, $x\in A^c$, とできる．A を端点が F の連続点であるような小さな区間 $I_1,\cdots,I_r$, に分けて，各小区間での f の変動は ϵ 未満とすると，A^c で 0, 各 I_k で定数なる関数 σ であって，$|f(x)-\sigma(x)|<\epsilon$, $x\in\mathbf{R}$, とできる．よって

$$\left|\int(f(x)-\sigma(x))dF_n(x)\right|<\epsilon,\quad\left|\int(f(x)-\sigma(x))dF(x)\right|<\epsilon.$$

$\displaystyle\int\sigma dF_n$ は $F_n(I_k)$ の線形結合だから，$n\to\infty$ で弱収束の仮定から $F(I_k)$ の線形結合，即ち，$\displaystyle\int\sigma dF$ に収束する．$\epsilon>0$ は任意だから (27) が成り立つ．

次に (27) が成り立つとする．定理 54 より，$\{F_n\}$ には弱収束部分列がある．このとき，前半と (27) より，その弱収束極限 $\tilde F$ に対して $\displaystyle\int f(x)d\tilde F(x)=\int f(x)dF(x)$ となり，$d\tilde F=dF$ は部分列によらない（ f を調節することにより，分布関数の等式 $\tilde F_n(x)=F(x)$ にいくらでも近づけられる）．よって，$\{F_n\}$ は F に弱収束する．

□

定理 56 (連続定理 [15, §XIII.1]) *(i)* $n\in\mathbf{N}$ に対して，F_n は $\mathbf{R}_+$ 上の分布関数（*(23)* を満たす右連続非減少関数）として，ラプラス変換を G_n とおく．

このとき，F_n が右連続非減少関数 F に弱収束すれば $n\to\infty$ で $G_n(t)$ は F のラプラス変換 $G(t)$ に $t>0$ の各点で収束する．

逆に，$G_n(t)$ がある関数 $G(t)$ に $t>0$ の各点で収束するならば F_n は $\mathbf{R}_+$ 上の有界右連続非減少関数 F に弱収束し，そのラプラス変換は G である．

さらに，F が確率分布であることは $\lim_{t\to 0}G(t)=0$ と同値．

(ii) $n\in\mathbf{N}$ に対して F_n は右連続非減少な $\mathbf{R}_+$ 上の関数であって，そのラプラス変換を G_n とする．ある $a\ge 0$ に対して $t>a$ の各点で $\displaystyle\lim_{n\to\infty}G_n(t)=G(t)$ ならば G は右連続非減少なある $\mathbf{R}_+$ 上の関数 F のラプラス変換であって，F_n は F に弱収束する．

逆に，F_n が F に弱収束して，$\{G_n(a)\}$ が有界ならば F のラプラス変換を G とするとき $t>a$ の各点で $\displaystyle\lim_{n\to\infty}G_n(t)=G(t)$.

証明. (i) 前半は 定理 55 より明らか．後半は，先ず 定理 54 から収束部分列 $F_{n(k)}$ と弱収束極限 F をとることができる．仮定より $\displaystyle\lim_{k\to\infty}G_{n(k)}(t)=G(t)$ なので，F のラプラス変換は前半の結果より G となる．よって，定理 52 より F は収束部分列によらないから，F_n は F に弱収束する．F が $\mathbf{R}_+$ 上の有界右連続非減少関数であることは F_n の極限だから明らか．

最後の主張はラプラス変換の定義から明らか．

(ii) $t_0>a$ を一つとると (25) の前後の注より，$G_n(t+t_0)/G_n(t_0)$ は分布

$$d\mu_n(x)=\frac{1}{G_n(t_0)}e^{-t_0x}dF_n(x)$$

のラプラス変換である．前項を $d\mu_n$ に適用すればよい．

逆に，F_n が F に弱収束して，$\{G_n(a)\}$ が有界とする．弱収束の定義から（必要なら 定理 55 のように積分範囲を細かく分割する議論を適用して）任意の $t > a$ に対して，$\lim_{n\to\infty} G_n(t) = G(t)$.

□

A.2.3 Tauber 型定理．

定理 57 (Karamata の Tauber 型定理) $U:\ \mathbf{R}_+ \to \mathbf{R}$ を非減少関数とし，測度 dU のラプラス変換

$$w(x) = \int_{\mathbf{R}_+} e^{-xu} dU(u)$$

が全ての $x > 0$ で有限であるとする．このとき，任意の $\rho \geq 0$ と緩慢変動関数 L に対して以下が成り立つ．

(i) (x が $+0$ の付近で) $w(x) = x^{-\rho} L(x^{-1})$ ならば

$$U(x) \asymp \frac{1}{\Gamma(\rho+1)} x^{\rho} L(x), \quad x \to \infty,$$

(ii) (x が $+\infty$ の付近で) $w(x) = x^{-\rho} L(x)$ ならば

$$U(x) \asymp \frac{1}{\Gamma(\rho+1)} x^{\rho} L(x^{-1}), \quad x \to +0\,.$$

ここで $f(x) \asymp g(x),\ x \to a,$ とは $\lim_{x\to a} \dfrac{f(x)}{g(x)} = 1$ であることを言い，また，$\Gamma(y) = \displaystyle\int_0^\infty e^{-x} x^{y-1} dx$ はガンマ関数である．

証明. [15, §XIII.5] と [22, §2.2] に証明がある．両者は内容は同じだが後者のほうが記述が丁寧なので，以下，そちらを引用する．

後半も同様なので，前半のみ証明する．$x > 0$ に対して $U_x^*(u) = w(x)^{-1} U(u/x)$ で $U_x^*:\ \mathbf{R}_+ \to \mathbf{R}$ を定義すると，これも非減少で，変数変換 $u = xy$ によって

$$\int_{\mathbf{R}_+} e^{-\lambda u} dU_x^*(u) = \frac{w(\lambda x)}{w(x)}, \quad x > 0,$$

となるが，w に対する仮定から右辺は $x \to +0$ で $\displaystyle\int_{\mathbf{R}_+} e^{-\lambda u} dG(u)$ に収束する．ここで $G(u) = \dfrac{u^\rho}{\Gamma(\rho+1)}$. 従って，定理 56 より

$$\lim_{x\to+0} U_x^*(u) = G(u), \quad u \geq 0,$$

を得る．$v = ux^{-1}$ によって x を v に変数変換すると，

$$U(v) \asymp \frac{1}{\Gamma(\rho+1)} v^{p} L(vu^{-1}) \asymp \frac{1}{\Gamma(\rho+1)} v^{p} L(v), \quad v \to \infty.$$

□

定理 57 はラプラス変換 w から分布関数 U の性質を導く．もう少し条件を付ければ，分布関数の（ルベーグ測度に関する）密度関数の性質を得ることができる．

定理 58 ある $A>0$ があって U は $[A,\infty)$ 上で定義された正値関数であって，十分大きな y において単調な関数 $u(y)$ があって $U(x)=\int^x u(y)\,dy,\ x\geq A,$ が成り立っているとする．このとき $\rho\geq 0$ と緩慢増加関数 L に対して，*(*x が $+\infty$ の付近で*)* $U(x)=x^\rho L(x)$ ならば

$$\lim_{x\to\infty}\frac{xu(x)}{U(x)}=\rho\,.$$

証明. u が遠方で単調非減少な場合のみ証明する（単調非増加な場合も同様）．$\beta\geq\alpha$ が十分大きければ

$$\frac{U(t\beta)-U(t\alpha)}{U(t)}=\int_{t\alpha}^{t\beta}\frac{u(y)}{U(t)}\,dy.$$

よって

$$\frac{t(\beta-\alpha)u(t\beta)}{U(t)}\geq\frac{U(t\beta)-U(t\alpha)}{U(t)}\geq\frac{t(\beta-\alpha)u(t\alpha)}{U(t)}$$

となる．ここで $t\to\infty$ とすると，右辺から

$$\frac{\beta^\rho-\alpha^\rho}{\beta-\alpha}\geq\overline{\lim_{t\to\infty}}\frac{tu(t\alpha)}{U(t)}$$

となるが，右辺は β によらないので $\beta\downarrow\alpha$ として

$$\rho\alpha^{\rho-1}\geq\overline{\lim_{t\to\infty}}\frac{tu(t\alpha)}{U(t)}\,.$$

同様に先ほどの式の左辺からは

$$\underline{\lim_{t\to\infty}}\frac{tu(t\beta)}{U(t)}\geq\rho\beta^{\rho-1}\,.$$

α,β は任意なので，任意の $c>0$ に対して

$$\lim_{t\to\infty}\frac{tu(tc)}{U(t)}\geq\rho c^{\rho-1}\,.$$

$x=tc$ とおいて $U(t)=t^\rho L(t)$ を用いれば主張を得る． □

Tauber 型定理はべき級数の収束半径付近での性質からべき級数の係数の漸近的性質についても言う [15, §XIII.5].

定理 59 $q_n\geq 0,\ n=0,1,2,3,\cdots,$ とし，$Q(s)=\sum_{n=0}^{\infty}q_ns^n$ の収束半径を 1 とする．また， L を *(*$x\to\infty$ で*)* 緩慢変動関数とし， $\rho\geq 0$ とする．このとき

$$Q(s)\asymp\frac{1}{(1-s)^\rho}L((1-s)^{-1}),\quad s\to 1-,$$

ならば

$$\sum_{k=0}^{n}q_k\asymp\frac{1}{\Gamma(\rho+1)}n^\rho L(n),\quad n\to\infty,$$

である．特に， $\{q_n\}$ が単調で $\rho>0$ ならばさらに

$$q_n\asymp\frac{1}{\Gamma(\rho)}n^{\rho-1}L(n),\quad n\to\infty,$$

となる．

証明. $dU(x)=\sum_{n=0}^{\infty}q_n\delta(x-n)dx$ とし，w をそのラプラス変換とすると， $w(\lambda)=Q(e^{-\lambda})$ となるので，定理 57 より最初の主張を得る．

次に， $u(x)=q_{[x]}$ として 定理 58 を用いれば，対応する U のラプラス変換は上記 w を用いて $w(\lambda)\dfrac{1-e^{-\lambda}}{\lambda}$ と書けるが，これは $\lambda\to +0$ のとき $\asymp w(\lambda)$ となるので，あとの主張も得る． □

A.2.4 正則変動関数の表現定理．

§A.3 で使うので，正則変動関数の著しい性質に言及しておく．[22, §1.2] と [15, §VIII.9] に記述があるが，前者のほうが記述が丁寧[26] なので，以下この節は前者から引用する[27]．$R(x) = x^\rho L(x)$ を考えることによって，緩慢変動関数のみやれば十分である．

定理 60 (一様収束性) L が緩慢変動関数ならば，任意の $0 < a < b < \infty$ に対して *(22)* が $\lambda \in [a, b]$ に対して一様に成り立つ．

定理 61 (表現定理) ある $A > 0$ に対して L が $[A, \infty)$ 上で緩慢変動関数ならば，$B \geq A$ があって，

$$L(x) = \exp\left(\eta(x) + \int_B^x \frac{\epsilon(t)}{t}\,dt\right), \quad x \geq B, \tag{28}$$

という形に書ける．ここで，$\eta\colon [B, \infty) \to \mathbf{R}$ は $x \to \infty$ で収束する有界可測関数，また $\epsilon\colon [B, \infty) \to \mathbf{R}$ は連続関数であって $\displaystyle\lim_{x\to\infty} \epsilon(x) = 0$ が成り立つ．

次の性質は系 76 で用いるので [22, §1.5] から引用する．

定理 62 (漸近的逆関数 (asymptotic inverse)) 指数 $\gamma > 0$ の正則変動関数 $R_1(x) = x^\gamma L_1(x)$ に対して，

$$\lim_{x\to\infty} \frac{1}{x} R_1(R_2(x)) = \lim_{x\to\infty} \frac{1}{x} R_2(R_1(x)) = 1 \tag{29}$$

を満たす指数 $1/\gamma$ の正則変動関数 $R_2(x) = x^{1/\gamma} L_2(x)$ が存在する．しかも， R_2 は以下の意味で漸近的に一意的に決まる：R_3 も *(29)* のどちらかに対応する極限が 1 になり，かつ， $\displaystyle\lim_{x\to\infty} R_3(x) = \infty$ とすると， $\displaystyle\lim_{x\to\infty} R_2(x)/R_3(x) = 1$．

これらを証明するために，$f(x) = \log L(e^x)$ とおく．

定義から，f はある γ に対して $[\gamma, \infty)$ 上で定義された実数値関数で，

$$(\forall \mu \in \mathbf{R}) \quad \lim_{x\to\infty} (f(x+\mu) - f(x)) = 0 \tag{30}$$

が成り立つ．f に対して以下の補題を用意する．

補題 63 収束 *(30)* は μ について任意の有界閉区間で一様である．

証明. 先ず $\mu \in [0, 1]$ で一様であることを背理法で証明する．

収束が一様でないとすると， $\epsilon > 0$ と $\displaystyle\lim_{n\to\infty} x_n = \infty$ なる数列 $\{x_n\}$ と $\mu_n \in [0, 1], n \in \mathbf{Z}_+$, なる数列 $\{\mu_n\}$ がとれて

$$|f(x_n + \mu_n) - f(x_n)| \geq \epsilon, \quad n \in \mathbf{Z}_+, \tag{31}$$

が成り立つ．集合 $U_n \subset \mathbf{R}, V_n \subset \mathbf{R}, n \in \mathbf{Z}_+$, を

$$\begin{aligned} U_n &= \{\mu \in [0, 2] \mid |f(x_m + \mu) - f(x_m)| < \tfrac{1}{2}\epsilon,\ m \geq n\} \\ V_n &= \{\lambda \in [0, 2] \mid |f(x_m + \mu_m + \lambda) - f(x_m + \mu_m)| < \tfrac{1}{2}\epsilon,\ m \geq n\} \end{aligned} \tag{32}$$

で定義する．明らかに U_n, V_n たちは可測集合で，$\{U_n\}$, $\{V_n\}$ はそれぞれ増加集合列であり，(30) から

$$\bigcup_{n\in\mathbf{Z}_+} U_n = \bigcup_{n\in\mathbf{Z}_+} V_n = [0, 2]$$

なので，ルベーグ測度を m で表すことにすると，十分大きな N に対して

$$m(U_N) > \frac{3}{2}, \quad m(V_N) > \frac{3}{2}.$$

---

[26] 「丁寧」すぎる？

[27] 前者はほとんど初等解析的な証明で，後者は測度論的色彩の濃い証明だから，それぞれに若干の特徴があるが，後者は書き方が少し不親切なようだ．

$V'_N = V_N + \mu_N$ とおくと $m(V'_N) = m(V_N) > 3/2$ かつ $U_N \subset [0,3], V'_N \subset [0,3]$. よって $U_N \cap V'_N \neq \emptyset$ でなければならない．即ち，

$$\exists \mu \in U_N;\ \mu - \mu_N \in V_N\,.$$

この μ に対して，(32) から，

$$|f(x_N + \mu) - f(x_N)| - f(x_N)| < \frac{1}{2}\epsilon,$$
$$|f(x_N + \mu) - f(x_N)| - f(x_N + \mu_N)| < \frac{1}{2}\epsilon,$$

となるので

$$|f(x_N + \mu_N) - f(x_N)| < \frac{1}{2}\epsilon.$$

これは (31) に矛盾する．

一般の区間 $\mu \in [a,b]$ に関する一様収束性は，$\tilde{f}(x) = f((b-a)x + a)$ とおくと，$\tilde{f}$ も (30) が成り立つので以上の結果を f を $\tilde{f}$ として適用できて，変数変換 $y = x/(b-a), \nu = (\mu - a)/(b-a)$ で

$$f(x+\mu) - f(x) = \tilde{f}(y+\nu) - \tilde{f}(y) + f(x+a) - f(x)$$

となるので，$\nu \in [0,1]$ の場合と μ によらない 1 点 a の場合に帰着する． □

補題 64 ある $X \geq \gamma$ があって，f は任意の $X' \geq X$ に対して $[X, X']$ で有界であり，従って，特に（可測有界だから）f はそこで可積分である．

証明. 補題 63 からある X があって

$$|f(x+\mu) - f(x)| < 1, \quad x \geq X,\ \mu \in [0,1].$$

よって任意の $y \in [X, X+1]$ に対して $x = X, \mu = y - X$ とおくことで $|f(y)| \leq |f(X)| + 1$ を得る．さらに同様に $y \in [X+1, X+2]$ に対して $|f(y)| \leq |f(X+1)| + 1 \leq |f(X)| + 2$ を得る．一般に，任意の正整数 k に対して $|f(x)| \leq |f(X)| + k$ が $x \in [X+k-1, X+k]$ に対して成り立つので，$x \in [X, X+k]$ で成り立つ． □

補題 65 X を 補題 *64* と等しく取ると $x \geq X$ に対して

$$f(x) = c(x) + \int_X^x \epsilon(t)dt$$

の形に書ける．ここで $c,\ \epsilon$ は可測関数で，任意の $X' \geq X$ に対して $[X, X']$ で有界であり，また，$\lim_{x\to\infty} c(x) = c$ なる c が存在し，$\lim_{x\to\infty} \epsilon(x) = 0$ が成り立つ．

証明. 補題 64 より，

$$f(x) = \int_x^{x+1} (f(x) - f(t))dt + \int_X^x (f(t+1) - f(t))dt + \int_X^{X+1} f(t)dt, \quad x \geq X.$$

右辺各項をそれぞれ

$$= \delta + \int_X^x \epsilon(t)dt + c$$

とおくと，(30) と 補題 63 から，それぞれ

$$\lim_{t\to\infty} \epsilon(t) = 0,$$

および

$$\lim_{x\to\infty} \delta(x) = \lim_{x\to\infty} \int_0^1 (f(x) - f(x+\mu))d\mu = 0\,.$$

よって，$c(x) = \delta(x) + c$ とおけばよい． □

補題 66 $X \geq X$ が存在して，$x \geq X$ に対して

$$f(x) = c^*(x) + \int_X^x \epsilon^*(t)dt$$

の形に書ける．ここで c^*, ϵ^* は 補題 *65* と同じ性質を持つ．即ち，可測関数で，任意の $X' \geq X$ に対して $[X, X']$ で有界であり，また，$\lim_{x\to\infty} c^*(x) = c^*$ なる c^* が存在し，$\lim_{x\to\infty} \epsilon^*(x) = 0$ が成り立つ．

しかも，ϵ^* は連続関数である．

証明. $\epsilon^*(x) = \epsilon(x) - (c(x+1) - c(x)) = f(x+1) - f(x) - (c(x+1) - c(x))$, および,

$$c^*(x) = c(x) + \int_X^x (c(t+1) - c(t))dt = c(x) + \int_x^{x+1} c(t)dt - \int_X^{X+1} c(t)dt$$

とおく．[28] 補題 65 より，c^*, ϵ^* は主張の性質を全て持つことも明らかである ($c^* = 2c - \int_X^{X+1} c(t)dt$). □

定理 *60*, **定理** *61* **の証明**. $f(x) = \log L(e^x)$ 即ち，$L(x) = \exp f(\log x)$ とおけば，前者は補題 63 から，後者は補題 66 から，それぞれ明らか．

後者については $\eta(x) = c^*(\log x)$, $\epsilon(x) = \epsilon^*(\log x)$, $B = \exp X$, とおけばよい． □

定理 *62* **の証明**. 定理 61 からある C_1 に対して $x \geq C_1$ で $R_1(x) = K_1(x)r_1(x)$ とできる．ここで K_1 は $\exists K_1 = \lim_{x\to\infty} K_1(x) > 0$ を満たす正値関数で，r_1 は 0 に収束する連続関数 ϵ_1 を用いて

$$r_1(x) = \exp(\int_{C_1}^x (\gamma + \epsilon_1(t))\frac{dt}{t})$$

と書ける．ここで C_1 を十分大きく選んで $\gamma + \epsilon_1(t) > 0, t \geq C_1$, とする．従って，$r_1$ は連続で狭義増加して ∞ に発散するので，同様の性質を持つ逆関数 r_2 が $x \geq C_2 = r_1(C_1)$ で存在する．即ち，

$$r_1(r_2(x)) = r_2(r_1(x)) = x.$$

r_1, r_2 は微分可能で

$$r_1'(r_2(x))r_2'(x) = 1 = r_2'(r_1(x))r_1'(x)$$

だが，r_1 の表示を代入すれば

$$\frac{r_2'(r_1(x))r_1(x)}{r_2(r_1(x))} = \frac{1}{\gamma + \epsilon_1(x)}$$

となるので，$x = r_2(t)$ として

$$t(\log r_2)'(t) = \gamma^{-1} + \epsilon_2(t).$$

ここで ϵ_2 は連続で $\lim_{t\to\infty} \epsilon_2(t) = 0$ を満たす．したがって $x \geq C_2$ で

$$r_2(x) = x^{1/\gamma} \exp(\int_{C_2}^x \epsilon_2(t)\frac{dt}{t}).$$

$R_2(x) = K_1^{-1/\gamma} r_2(x)$ が (29) を満たす．ここで第 2 の等号を示すのには 定理 60 を用いる．

漸近的一意性を示すために，$R_1(R_3(x)) = x(1+\epsilon(x))$ とおく．ここで $\lim_{x\to\infty} \epsilon(x) = 0$．このとき，$R_2(R_1(R_3(x))) = R_2(x(1+\epsilon(x)))$ に $\lim_{x\to\infty} \frac{1}{x} R_2(R_1(x)) = x$ と 定理 60 を用いれば，主張を得る．もう一方の場合も同様である． □

---

[28] ここの証明はかってに書き換えた．[22] の証明は回りくどい上に X も X^* に変更している．本文のようにおけばいいと思うのだが…

A.3 指数型の Tauber 型定理．

指数型の振る舞いの場合でも，べき振る舞いまで精密に求めることで，$\log \mathrm{P}[\, X > x \,]$ ではなく $\mathrm{P}[\, X > x \,]$ の漸近形（比が 1 に収束するという意味）を求める研究も多いが，ここでは $\log \mathrm{P}[\, X > x \,]$ の漸近形だけを概観する．（例えば，ガンマ関数の漸近形は [7, 定理 4.2] で与えられるが，ここでは [7, 定理 4.1] までを目標にする，ということ．）

A.3.1 ガンマ関数の対数の漸近形．

母関数 $\mathrm{E}[\, e^{tX} \,]$ で $t \to \infty$ 等を考えるので，正実数上の確率測度の極限定理には特に有用である．[7, §4] に従ってガンマ関数を例に取る．

ガンマ関数 $\Gamma(t) = \displaystyle\int_0^\infty e^{-x} x^{t-1} dx \ (t > 0)$ の $t \to \infty$ での漸近形（比が 1 に収束するという意味で）が

$$\Gamma(t) \asymp \sqrt{2\pi t} t^t e^{-t}, t \to \infty,$$

であることは良よく知られている．もっとも増大度の大きい因子 $t^t = e^{t \log t}$ は変数変換 $x = yt$ で得る：

$$t^{-t}\Gamma(t) = \int_0^\infty e^{t(-y+\log y)} \frac{dy}{y}\,. \tag{33}$$

ここではその次に増大度の大きい因子 e^{-t} を目標にする．即ち，

目標：$\log(t^{-t}\Gamma(t)) \asymp -t,\ t \to \infty.$

定理 67 $[a, b]$ に値をとる確率変数 X の分布 μ の密度 $\rho = \dfrac{d\mu}{dx}$ が $[a, b]$ で連続関数とする．さらに $f\colon\ [a, b] \to \mathbf{R}$ が連続関数で

$$I(t) = \mathrm{E}[\, e^{tf(X)} \,] \ \ \left(= \int_a^b e^{tf(x)} \mu(dx) \ \right)$$

が $t \geq 0$ に対して存在するとする．

このとき

$$\lim_{t\to\infty} \frac{1}{t} \log I(t) = \sup\{f(x) \mid x \in [a, b],\ \rho(x) > 0\}. \tag{34}$$

証明. 任意の A と $\epsilon > 0$ に対してチェビシェフの不等式から

$$I(t) = \mathrm{E}[\, e^{tf(X)} \,] \geq e^{t(A-\epsilon)} \mathrm{P}[\, f(X) > A - \epsilon \,]$$

A を (34) の右辺にとればこの式の右辺は正だから，

$$\frac{1}{t} \log I(t) \geq A - \epsilon + \frac{1}{t} \log \mathrm{P}[\, f(X) > A - \epsilon \,]$$

なので

$$\underline{\lim}_{t\to\infty} \frac{1}{t} \log I(t) \geq A - \epsilon.$$

$\epsilon > 0$ は任意だから

$$\underline{\lim}_{t\to\infty} \frac{1}{t} \log I(t) \geq A.$$

他方 $x \in [a, b]$ かつ $\rho(x) > 0$ ならば $A \geq f(x)$ だから

$$I(t) = \mathrm{E}[\, e^{tf(X)} \,] \leq e^{tA} \mathrm{P}[\, [a, b] \,] = e^{tA}$$

となって，$\displaystyle \overline{\lim}_{t\to\infty} \frac{1}{t} \log I(t) \leq A$. よって (34) が成り立つ [7, 定理 4.1]. □

注. 証明から $[a,b]$ で有限測度なら十分で確率測度である必要はない. ◇

ガンマ関数 (33) に戻る. $0<a<1<b<\infty$ を固定し, μ を密度 $\rho(x)=\dfrac{1}{x\log(b/a)}$ を持つ $[a,b]$ 上の確率測度とする.(定理 67 の注によれば, 全測度を 1 にしなくてもいいので, $\log(b/a)$ でわざわざ割って, 確率測度にするのは無駄.) さらに $f(x)=-x+\log x$ として定理 67 を適用すると

$$\lim_{t\to\infty}\frac{1}{t}\log\int_a^b e^{t(-y+\log y)}\frac{dy}{y}=\sup\{-x+\log x\mid x\in[a,b]\}.$$

右辺は増減表を書けば $a<1<b$ なので -1 に等しい.

他方 $0<a<1$ だから $t>1$ で

$$\int_0^a e^{t(-y+\log y)}\frac{dy}{y}=\int_0^a e^{-ty}y^{t-1}dy\le\int_0^a e^{-ty}dy=\frac{1}{t}(1-e^{-at}).$$

よって

$$\lim_{t\to\infty}\frac{1}{t}\log\int_0^a e^{t(-y+\log y)}\frac{dy}{y}=0\,.$$

また, 増減表から $y>3$ ならば $-y/2+\log y<0$ だから, $b>3$ ならば,

$$\int_b^\infty e^{t(-y+\log y)}\frac{dy}{y}\le\int_b^\infty e^{-ty/2}dy=\frac{2}{t}e^{-bt/2}.$$

よって

$$\lim_{t\to\infty}\frac{1}{t}\log\int_b^\infty e^{t(-y+\log y)}\frac{dy}{y}\le-\frac{b}{2}<-\frac{3}{2}\,.$$

以上を (33) に代入して, log をとってから極限をとるので, 3 つの積分からの寄与のうち最大のものしか残らないことに注意すれば

$$\lim_{t\to\infty}\frac{1}{t}\log(t^{-t}\Gamma(t))=-1,$$

即ち目標の評価を得た.

A.3.2 チェビシェフの不等式.

本題に入る前に, やさしいが出発点となる注意をしておく. 一般に, 母関数から確率については, チェビシェフの不等式によって, 容易に一方向の評価（不等式）を得る. 次の命題はその簡単な例題である.

命題 68 X の分布が連続（即ち, 分布関数 $\mathrm{P}[\,X\le x\,]$ が $x\in\mathbf{R}$ の連続関数とする. $\mathrm{E}[\,e^{tX}\,]$ が $t<c$ で存在し, $t=c$ で発散するならば,

$$\overline{\lim_{x\to\infty}}\frac{1}{x}\log\mathrm{P}[\,X\ge x\,]=-c. \tag{35}$$

証明. チェビシェフの不等式から

$$e^{tx}\mathrm{P}[\,X\ge x\,]\le\mathrm{E}[\,e^{tX}\,],\quad t<c.$$

両辺の log をとって x で割り, $\overline{\lim\limits_{x\to\infty}}$ を考えれば

$$\overline{\lim_{x\to\infty}}\frac{1}{x}\log\mathrm{P}[\,X\ge x\,]\le-t,\quad t<c.$$

$t<c$ で常に成り立つから

$$\overline{\lim_{x\to\infty}}\frac{1}{x}\log\mathrm{P}[\,X\ge x\,]\le-c.$$

逆に,
$$\varlimsup_{x\to\infty}\frac{1}{x}\log \mathrm{P}[\,X\ge x\,] < -c$$
ならば $\epsilon>0$, $x_0\ge 0$ が存在して,
$$\mathrm{P}[\,X\ge x\,]\le e^{-(c+\epsilon)x},\quad x\ge x_0\,.$$
部分積分（[3, 定理 20.2] で $F(x)=\mathrm{P}[\,X\ge x\,]$ として，連続性の仮定を使う）によって
$$\mathrm{E}[\,e^{cX}\,]=\int_0^\infty e^{cx}d(-\mathrm{P}[\,X\ge x\,]) = -e^{cx}\mathrm{P}[\,X\ge x\,]|_0^\infty + c\int_0^\infty e^{cx}\mathrm{P}[\,X\ge x\,]dx<\infty$$
となって， $\mathrm{E}[\,e^{xX}\,]=\infty$ に矛盾する． □

指数型の Tauber 型定理では，このようなチェビシェフの不等式による一方向からの評価が，正しい漸近形を与えるのが普通である．従って問題の難しさは，逆向きの同じ order の評価をいかに得るかである．

A.3.3 凸関数とルジャンドル変換．

指数型 Tauber 型定理の証明のために凸関数 ((6)) とルジャンドル変換という初等的な概念を導入する．これは §A.4 でも使う．

I を $\mathbf{R}$ またはその連結部分集合（例えば §A.3.4 では $(0,\infty)$）とする．恒等的に ∞ でない関数 $f:\ I\to\mathbf{R}\cup\{\infty\}$ が下に凸な関数であるとは
$$0\le\lambda\le 1 \quad\Longrightarrow\quad f(\lambda x+(1-\lambda)y)\le \lambda f(x)+(1-\lambda)f(y),\ x,y\in I, \tag{36}$$
を満たすことをいう．ここで (36) は $\infty\le\infty$ も許す．

命題 69 $f:\ I\to\mathbf{R}\cup\{\infty\}$ を（恒等的に ∞ でない）下に凸な関数とするとき以下が成り立つ．

(i) f は I の区間 $J\subset I$ で実数値をとり，$I\setminus J$ では ∞ である．
(ii) f は上記 J の内部 J^o で連続関数である．
(iii) $x\in J^o$ に対して
$$f(y)\ge f(x)+m(y-x),\quad y\in I, \tag{37}$$
を満たす $m\in\mathbf{R}$ がある．

注. (i) I の外では $f=\infty$ とおくことで， I は $\mathbf{R}$ にとっておいていいように思うが， §A.3.4 で $I=(0,\infty)$ に取っているのでここではこのままにしておく．
(ii) ここでの定義だけでは J の端点が J に入るかどうか（J が I の閉区間かどうか）は何も言えないと思う．また， J の端点で（J の中から近づくとき）連続関数かどうかも言えないと思う．従って， §A.6 における rate function の定義の中の lower semicontinuous という要請は下に凸な関数というよりも多くの要請が入っている，と思う．命題 70 に示したように，ルジャンドル変換になっていれば lower semicontinuous なので，これが rate function の典型であろう．
◇

証明. (i) (36) から
$$x<z<y \quad\Longrightarrow\quad f(z)<f(x)\vee f(y) \tag{38}$$
だから，$f(x)\in\mathbf{R}$ なる x は連結集合 $J\subset I$ をなす．

(ii) J が 1 点ならば主張に意味がない (trivial) から内点のある区間とし，f が $c \in J^o$ で不連続とする．

J の定義より $f(c) < \infty$ だから (38) より，c を含む J の閉区間で有界である．$\inf_{n\in\mathbf{N}} |f(c)-f(x_n)| > 0$ が成り立つような c に収束する列 $\{x_n\} \subset J$ が存在する．部分列を取ることで $\{x_n\}$ は単調であるとしてよい．$[x_1, c]$ (または $[c, x_1]$) で f が有界だから部分列をとること (compactness argument) で $\{f(x_n)\}$ の収束する部分列があるので，$\{f(x_n)\}$ も収束しているとしてよい．

もし，さらに，

$$\lim_{n\to\infty} f(x_n) - f(c) > 0 \tag{39}$$

を満たすならば，十分大きい n を考えることで最初からある $\delta > 0$ に対して $2\delta > f(x_n) - f(c) > \delta, n \in \mathbf{N}$, としてよい．$\lim_{n\to\infty} x_n = c$ だから，ある n_0 と $0 < \lambda < 1/2$ がとれて $x_{n_0} - c = \lambda(x_1 - c)$ とできる.（ $\{x_n\}$ は単調と仮定したので x_{n_0} は x_1 と c の間にある.）即ち $x_{n_0} = \lambda x_1 + (1-\lambda)c$. このとき

$$f(\lambda x_1 + (1-\lambda)c) = f(x_{n_0}) > \delta + f(c) > 2\delta\lambda + f(c) > \lambda f(x_1) + (1-\lambda) f(c)$$

これは (36) に矛盾するから，下に凸であるという仮定に反する．よって (39) を満たす $\{x_n\}$ はとれないことが分かる[29]．

従って，f が c で不連続とすると，c に収束する単調な $\{x_n\}$ で $f(c) - \lim_{n\to\infty} f(x_n) > 0$ を満たすものがあり，上と同じ議論で $2\delta > f(c) - f(x_n) > \delta > 0, n \in \mathbf{N}$, としてよい．しかも，$c$ に関して $\{x_n\}$ と反対側にある x に対しては，$\overline{\lim}_{x\to c;\ \{x_n\}\text{と反対側}} f(x) \le f(c)$．$\{x_n\}$ と反対側の点 $(a, f(a))$ をとる．$n \in \mathbf{N}$ に対して，$\lambda = \dfrac{a-c}{a-x_n}$ とおくと，

$$\lambda f(x_n) + (1-\lambda) f(a) \le f(c) - \delta + (1-\lambda)(f(a) + \delta - f(c)).$$

x_n を c に十分近く取って，$(1-\lambda)(f(a) + \delta - f(c)) < \delta/2$ を満たすように選べば，

$$\lambda f(x_n) + (1-\lambda) f(a) < f(c) = f(\lambda x_n + (1-\lambda) a)$$

これは (36) に矛盾する．

(iii) (36) から $y < x < z$ ならば $f(x) \le f(z)\dfrac{x-y}{z-y} + f(y)\dfrac{z-x}{z-y}$．従って $x \in \mathbf{R}$ に対して

$$\sup_{y<x} \frac{f(x)-f(y)}{x-y} \le \inf_{z>x} \frac{f(z)-f(x)}{z-x}.$$

両辺の間に a があれば (等号を許す)，$y < x$ ならば左辺から $f(y) \ge f(x) + a(y-x)$，$y > x$ ならば右辺から $f(y) \ge f(x) + a(y-x)$，よって常にこれが成り立つ．

□

もう一つ準備を行う．$I \subset \mathbf{R}$ を区間（有限または無限，端点不問）とし，$f\colon I \to \mathbf{R}$ に対して

$$g(x) = \sup_{t\in I} (tx - f(t)), \quad x \in \mathbf{R}, \tag{40}$$

なる関数 $g\colon \mathbf{R} \to \mathbf{R} \cup \{\infty\}$ を考える．

I^c で $f = \infty$ とすることで，$f\colon \mathbf{R} \to \mathbf{R} \cup \{\infty\}$ とみなして，sup を $t \in \mathbf{R}$ でとっても同値だから，以下自由にそのように読み替える．

対応 $f \mapsto g$ を Legendre 変換といい，物理の熱力学で重要な概念の一つである．

命題 70 *(i)* g は下に凸である．従って特に 命題 *69* より，値が有限の区間の内部では連続関数である．

[29] (以上の議論は c が J の端点であってもかまわない．しかし以下の議論はそうはいかない．

(ii) g は全ての $x\in\mathbf{R}$ で *lower semicontinuous* である．即ち，$\varliminf_{y\to x} g(y)\ge g(x),\ x\in\mathbf{R}$．（ $g(x)=\infty$ のときは $\lim_{y\to x} g(y)=\infty$ の意味で連続.）従って特に値が有限の区間の境界でも有限区間側から連続である．

(iii) f が下に凸ならば，ルジャンドル変換を 2 度行うと I^o で f に戻る．

注. 命題 70 は [8, §H] から引用した．ただし，端点に関する主張は追加した．端点でも lower semicontinuous であるだけでなく，有限値をとる側から近づけば連続になると思うが，あとで使わないので考察を省略.

Lower semicontinuous というのは large deviation principle の rate function であるための要請の一つ §A.6．ルジャンドル変換が lower semicontinuous であることが rate function の定義の動機の背景にあることが伺えるのでここで言及した． ◇

証明. (i) まず,

$$f(t)+g(x)=\sup_{s\in\mathbf{R}}(f(t)+sx-f(s))\ge tx,\quad t,x\in\mathbf{R}. \tag{41}$$

に注意する．

$0\le\lambda\le 1$ および $x,y\in\mathbf{R}$ とする．(41) より

$$\lambda g(x)+(1-\lambda)g(y)\ge -\lambda f(t)+\lambda tx-(1-\lambda)f(t)+(1-\lambda)ty=t(\lambda x+(1-\lambda)y)-f(t)$$

が任意の t に対して成り立つから， t について上限をとると

$$\lambda g(x)+(1-\lambda)g(y)\ge g(\lambda x+(1-\lambda)y)$$

を得る．即ち，(36) に照らすと g は下に凸．よって 命題 69 より K^o で連続である．

(ii) まず $g(x)<\infty$ の場合を証明する．定義より

$$g(y)\ge yt-f(t),\quad t\in\mathbf{R},\ y\in\mathbf{R},$$

および

$$g(x)\le xt_m-f(t_m)+\frac{1}{m},\quad m\in\mathbf{N},$$

を満たす数列 $\{t_m\}$ が存在する．

$$g(x)-g(y)\le xt_m-f(t_m)+\frac{1}{m}-(yt_m-f(t_m))\le(x-y)t_m+\frac{1}{m}\,,\quad m\in\mathbf{N},\quad y\in\mathbf{R},$$

だから

$$g(x)-\varliminf_{y\to x} g(y)\le\frac{1}{m}\,,\quad m\in\mathbf{N}.$$

左辺は m によらないから $\varliminf_{y\to x} g(y)\ge g(x)$.

次に $g(x)=\infty$ の場合を証明する．

$$xt_m-f(t_m)\ge m,\quad m\in\mathbf{N},$$

を満たす数列 $\{t_m\}$ が存在する．上と同様に

$$g(y)\ge yt_m-f(t_m)\ge m+(y-x)t_m$$

を得るので

$$\varliminf_{y\to x} g(y)\ge m\,,\quad m\in\mathbf{N}.$$

左辺は m によらないから $\lim_{y\to x\to\infty} g(y)=\infty$.

(iii) $f : \mathbf{R} \to \mathbf{R} \cup \{\infty\}$ が下に凸とし,

$$g(x) = \sup_{t \in \mathbf{R}} (tx - f(t)), \quad x \in \mathbf{R},$$

および

$$h(t) = \sup_{x \in \mathbf{R}} (xt - g(x)), \quad t \in \mathbf{R},$$

とおく.

まず, (41) より $f(t) \geq \sup_{x \in \mathbf{R}} (tx - g(x)) = h(t), t \in \mathbf{R}$. 次に, (41) と (37) より ((37) を適用するために f が下に凸という仮定を使う) ある a が存在して,

$$h(t) \geq xt - g(x) = \inf_{s \in \mathbf{R}} (xt - sx + f(s)) \geq \inf_{s \in \mathbf{R}} (xt - sx + f(t) + a(s-t)), \quad x \in \mathbf{R},\ t \in I^o.$$

$x = a$ と選ぶと

$$h(t) \geq f(t),\ t \in I^o.$$

以上より I^o で $h = f$ を得る.

□

§A.3.4 で用いる性質は若干仮定が強化されている [18, §1]. (以下, 下に凸ではなく上に凸な関数で書く.)

命題 71 $\mathbf{R}_+ = [0, \infty)$ 上で実数値をとる関数 f: $\mathbf{R}_+ \to \mathbf{R}$ が非減少上に凸とする. このとき以下が成り立つ.

(i) $0 < \gamma \leq 1$ ならば $f(\xi^\gamma)$ も上に凸である.

(ii) $-f$ のルジャンドル変換 *(40)* *(*$g(x) = \sup_{\xi > 0} (x\xi + f(\xi))$*)* は $x > 0$ で $g(x) = \infty$ となり, $g(0) = \lim_{\xi \to \infty} f(\xi)$, および, $x \leq 0$ で非減少.

証明. (i) x^γ は上に凸なので $0 \leq \alpha \leq 1$ に対して

$$(\alpha\xi_1 + (1-\alpha)\xi_2)^\gamma \geq \alpha{\xi_1}^\gamma + (1-\alpha){\xi_2}^\gamma.$$

$0 < \alpha < 1$, $\xi_1 < \xi_2$, に対して g は非減少上に凸なので

$$g((\alpha\xi_1 + (1-\alpha)\xi_2)^\gamma) \geq g(\alpha{\xi_1}^\gamma + (1-\alpha){\xi_2}^\gamma) \geq \alpha g({\xi_1}^\gamma) + (1-\alpha) g({\xi_2}^\gamma).$$

(ii) f が非減少だから $x > 0$ のとき

$$g(x) = \sup_{\xi > 0} (x\xi + f(\xi)) = \lim_{\xi \to \infty} (x\xi + f(\xi)) = \infty.$$

$g(0)$ も同様.

$\xi > 0$ だから $0 \geq x' > x$ のとき

$$x'\xi + f(\xi) > x'\xi + f(\xi)$$

となって g は非減少.

□

A.3.4 指数型 Tauber 型定理．

以下の記号を固定する：

$\alpha > 0$ とし，正の実数に対して定義された，定数でない非減少実数値関数 $f\colon (0,\infty) \to \mathbf{R}$ はある $\beta > \alpha$ に対して $f(\xi^\beta)$ が上に凸であるとする．
命題 71 より， $f(\xi^\alpha)$ も上に凸．$-f(\xi^\alpha)$ のルジャンドル変換を

$$g(x) = \sup_{\xi>0}(f(\xi^\alpha) + x\xi) \tag{42}$$

とする．

命題 71 より $g(x)$ は $x \le 0$ で考えれば十分で，そこで非減少，かつ，命題 70 より下に凸．さらに次が成り立つ．

命題 72 $g(x) \in \mathbf{R}$ *(有限), $x<0$．従って特に命題 69 より* $x < 0$ で連続関数．$x = 0$ でも連続．さらに $x < 0$ で狭義増加で $g(x) > f(+0) = \lim_{\xi\downarrow 0} f(\xi)$．$g(-\infty) = f(+0)$ とおけば，g は $[-\infty, 0]$ 上の $\mathbf{R} \cup \{\pm\infty\}$ に値を取る連続関数．

証明. $-f(\xi^\beta)$ が下に凸で $(0,\infty)$ 上実数値だから命題 69 より（(37) で例えば $x = 1$ として）$f(\eta^\beta) \le f(1) + m(\eta - 1), \eta > 0,$ を満たす $m \in \mathbf{R}$ がある． f が非減少だから $m \ge 0$．よって $-\infty < x < 0$ で

$$g(x) = \sup_{\eta>0}(f(\eta^\beta) + x\eta^{\beta/\alpha}) \le f(1) - m + \sup_{\eta>0}(m\eta + x\eta^{\beta/\alpha}) = f(1) - m + m\frac{\beta-\alpha}{\beta}\eta_0 < \infty.$$

ここで

$$\eta_0 = \left(\frac{m\alpha}{-x\beta}\right)^{\alpha/(\beta-\alpha)}.$$

即ち $g(x) \in \mathbf{R}$, $-\infty < x < 0$．よって特に命題 69 より $x < 0$ で連続関数．

$g(0) = f(\infty) \ge \sup_{\xi>0}(x\xi + f(\xi^\alpha)) = g(x)$, $x < 0$．他方， $f(+\infty) = \infty$ ならば任意の $N > 0$ に対して $\xi_0 > 0$ がとれて $f(\xi_0^\alpha) \ge N + 1$． $0 > x > -\xi_0^{-1}$ のとき

$$g(x) \ge x\xi_0 + f(\xi_0^\alpha) \ge N$$

となるから g は $x = 0$ で連続．$f(+\infty) < \infty$ ならば任意の $\epsilon > 0$ に対して ξ_0 がとれて $f(\xi_0^\alpha) \ge f(+\infty) - \dfrac{\epsilon}{2}$． $0 > x > -\dfrac{\epsilon}{2\xi_0}$ のとき

$$g(x) \ge x\xi_0 + f(\xi_0^\alpha) \ge f(\infty) - \epsilon$$

となるから，やはり g は $x = 0$ で連続．

f は非減少で定数でないから $f(\eta_1^\beta) > f(+0)$ なる $\eta_1 > 0$ がある． $0 < \eta_2 < \eta_1$ と $0 < a < 1$ に対して

$$f((a\eta_2 + (1-a)\eta_1)^\beta) \ge af(\eta_2^\beta) + (1-a)f(\eta_1^\beta)$$

だが， $\eta_2 \downarrow 0$ とし， $\eta_1 > \eta > 0$ に対して $a = 1 - \eta/\eta_1$ とおくと

$$f(\eta^\beta) - f(+0) \ge \frac{\eta}{\eta_1}(f(\eta_1^\beta) - f(+0)),\quad 0 < \eta < \eta_1,$$

を得る．$\beta > \alpha$ なので $\eta > 0$ を十分小さく取ることで右辺を $-x\eta^{\beta/\alpha}$ より真に大きくできる．よって

$$g(x) \ge f(\eta^\beta) + x\eta^{\beta/\alpha} > f(+0).$$

$x<0$ とする．$f(\eta_1^\beta)>f(+0)$ なる $\eta_1>0$ をとる．$f(\eta^\beta)$ が上に凸だから $0<\eta<\eta_1$ ならば $f(\eta^\beta)$ は $(0,f(+0))$ と $(\eta_1,f(\eta_1^\beta))$ を結ぶ直線より上にあるので，特に $f(\eta^\beta)>f(+0)$．他方で $f(+0)$ の定義から $0<\eta_2<\eta_1$ なる η_2 で $f(\eta_2^\beta)<f(\eta_1^\beta)$ なるものがあるから，

$$f(+0)<f(\eta_2^\beta)<f(\eta_1^\beta),\ 0<\eta_2<\eta_1,$$

なる η_2 がある．$m_1=\dfrac{f(\eta_1^\beta)-f(\eta_2^\beta)}{\eta_1-\eta_2}>0$ および $0<\eta_3<\eta_2$ に対して

$$\gamma(\eta_3)=\frac{f(\eta_2^\beta)-f(\eta_3^\beta)}{\eta_2-\eta_3}$$

とおくと，$f(\eta^\beta)$ が上に凸だから，$\gamma(\eta_3)\ge m_1,\,0<\eta_3<\eta_2,$ であって，η_3 に関して減少．よって $0<\eta_3<\eta_3'<\eta_2$ のとき

$$\begin{aligned}&(f({\eta_3'}^\beta)-m_1\eta_3')-(f({\eta_3}^\beta)-m_1\eta_3)=(\eta_2-\eta_3)(\gamma(\eta_3)-m_1)-(\eta_2-\eta_3')(\gamma(\eta_3')-m_1)\\&\ge(\eta_3'-\eta_3)(\gamma(\eta_3)-m_1)\ge 0,\end{aligned}$$

即ち $f({\eta_3}^\beta)-m_1\eta_3$ は $0<\eta_3<\eta_2$ で増加．他方 $m_1\eta+x\eta^{\beta/\alpha}$ は十分小さい η に対して増加するから結局 $f(\eta^\beta)+x\eta^{\beta/\alpha}$ は十分小さいが正の η に対して増加． $0<\eta<\eta_x$ で増加とすると，

$$g(x)=\sup_{\eta\ge\eta_x}(f(\eta^\beta)+x\eta^{\beta/\alpha})=\sup_{\xi\ge\xi_x}(f(\xi^\alpha)+x\xi).$$

ここで $\xi_x=\eta_x^{\beta/\alpha}$．$x<x'<0$ のとき

$$f(\xi^\alpha)+x'\xi\ge f(\xi^\alpha)+x\xi+(x'-x)(\xi_x\wedge\xi_{x'}),\quad \xi\ge\xi_x\wedge\xi_{x'}.$$

両辺の $\displaystyle\sup_{\xi\ge\xi_x\wedge\xi_{x'}}$ をとれば

$$g(x')\ge g(x)+(x'-x)(\xi_x\wedge\xi_{x'})>g(x),$$

即ち， g は $x<0$ で狭義単調増加.

□

命題 73 各 $-\infty<A<0$ に対して，

$$f(\xi^\alpha)+A\xi=g(A) \tag{43}$$

の $\infty>\xi>0$ における解がただ一つ存在し，$f(+0)<B<g(A)$ に対して

$$f(\xi^\alpha)+A\xi=B \tag{44}$$

の $\infty>\xi>0$ における解がちょうど二つ存在する．

証明. 命題 69 より $f(\xi^\alpha)$ は $\xi>0$ で連続だから，最大値の原理より $g(A)$ の定義 (42) の sup はある $0\le\xi_0\le\infty$ で attain する：

$$g(A)=f(\xi_0^\alpha)+A\xi_0\,.$$

$f(\eta^\beta)$ が上に凸なので $A<0$ のとき

$$\lim_{\xi\to\infty}(f(\xi^\alpha)+A\xi)=\lim_{\eta\to\infty}f(\eta^\beta)+A\eta^{\beta/\alpha}=-\infty.$$

また命題 72 より

$$\lim_{\xi\downarrow 0}(f(\xi^\alpha)+A\xi)=f(+0)<g(A).$$

よって $\xi_0 \neq 0$, かつ $\xi_0 \neq \infty$. 即ち $\infty > \xi > 0$ に解が存在する．

Uniqueness を証明するために $\xi = \lambda_i$, $i = 1, 2$, の二つの解があるとする．$g(A)$ の定義より $f(\xi^\alpha) + A\xi \le g(A)$, $\xi > 0$．他方左辺は ξ の関数として上に凸なので $\lambda_1 \le \xi \le \lambda_2$ で左辺は $g(A)$ 以上．よって $\lambda_1 \le \xi \le \lambda_2$ で $f(\xi^\alpha) + A\xi = g(A)$. 従って $f(\eta^\beta) = g(A) - A\eta^{\beta/\alpha}$, $\lambda_1^{\alpha/\beta} \le \eta \le \lambda_2^{\alpha/\beta}$. $A < 0$ なので右辺は下に凸だから $f(\eta^\beta)$ が上に凸という仮定に矛盾．よって解は一つ．

後半は，(43) の解を λ_0 とおくと，$f(\xi^\alpha) + A\xi$ が上に凸であることから，$0 < \xi < \lambda_0$ で増加，$\xi > \lambda_0$ で減少だから．

ξ	0		λ_1		λ_0		λ_2		∞
$f(\xi^\alpha) + A_2\xi$	$f(+0)$	$\nearrow$	$g(A_1)$	$\nearrow$	$g(A_2)$	$\searrow$	$g(A_1)$	$\searrow$	$-\infty$

$\square$

以上の記号の下で，[18] の主結果が成り立つ．

定理 74 ([18, Theorem 1]) $\mu(dx)$ を $(0, \infty)$ 上の有限ボレル測度，L を緩慢増加関数，$\phi(\lambda) = \lambda^\alpha L(\lambda)$ とし，

$$F(\lambda) = \int_0^\infty e^{\lambda f(x/\phi(\lambda))} \mu(dx)$$

とおく．このとき，

(i) ある $-\infty \le A_1 \le A_2 \le 0$ に対して

$$A_1 \le \varliminf_{x\to\infty} \frac{1}{x} \log \mu((\phi(x), \infty)) \le \varlimsup_{x\to\infty} \frac{1}{x} \log \mu((\phi(x), \infty)) \le A_2$$

が成り立つならば

$$g(A_1) \le \varliminf_{\lambda\to\infty} \frac{1}{\lambda} \log F(\lambda) \le \varlimsup_{\lambda\to\infty} \frac{1}{\lambda} \log F(\lambda) \le g(A_2). \tag{45}$$

(ii) 逆に，ある $-\infty \le A_1 \le A_2 < 0$ に対して *(45)* が成り立つ（g が非減少であることに注意）ならば

$$\frac{\lambda_2}{\lambda_1} A_2 \le \varliminf_{x\to\infty} \frac{1}{x} \log \mu((\phi(x), \infty)) \le \varlimsup_{x\to\infty} \frac{1}{x} \log \mu((\phi(x), \infty)) \le A_2\,.$$

ここで，$\lambda_1 \le \lambda_2$ は $f(\xi^\alpha) + A_2\xi = g(A_1)$ の二つの解（命題 *73* 参照）．$A_1 = -\infty$ のときは $\dfrac{\lambda_2}{\lambda_1} A_2 = -\infty$ と約束する．

(iii) $f(+\infty) < \infty$ かつ $\displaystyle\varliminf_{\lambda\to\infty} \frac{1}{\lambda} \log F(\lambda) \ge B > f(+0)$ ならば

$$\varliminf_{x\to\infty} \frac{1}{x} \log \mu((\phi(x), \infty)) \ge \frac{1}{\lambda_3}(B - f(+\infty)).$$

ここで $\lambda_3 = \displaystyle\sup_{f(\lambda^\alpha) < B} \lambda$.

注. 命題 73 の証明の最後の増減表から A_1 を止めて $A_2 \uparrow 0$ とすると

$$\lim_{A_2 \uparrow 0} \frac{\lambda_2}{\lambda_1} A_2 = \frac{1}{\lambda_3}(g(A_1) - f(+\infty))$$

を得るから，最後のケースは２番目のケースの極限形とみなせる． $\diamond$

系 75 *(i)* $-\infty \le A \le 0$ に対して，$\displaystyle\varlimsup_{x\to\infty} \frac{1}{x} \log \mu((\phi(x), \infty)) = A$ と $\displaystyle\varlimsup_{\lambda\to\infty} \frac{1}{\lambda} \log F(\lambda) = g(A)$ は同値である．

(ii) $-\infty \le A < 0$ に対して，$\displaystyle\lim_{x\to\infty} \frac{1}{x} \log \mu((\phi(x), \infty)) = A$ と $\displaystyle\lim_{\lambda\to\infty} \frac{1}{\lambda} \log F(\lambda) = g(A)$ は同値である．さらに $f(+\infty) < \infty$ ならば $A < 0$ の仮定はいらない．

証明. 前半は 命題 72 より g が狭義単調増加だから．後半は $A_1 = A_2$ ならば $\lambda_1 = \lambda_2$ (命題 73 参照) だから． □

系 76 ([18, Theorem 2]) $0 < \alpha < 1$ のとき，ψ を $x/\phi(x)$ の漸近的逆関数とすると，

(i) ある $-\infty \le -A_1 \le -A_2 \le 0$ に対して

$$-A_1 \le \varliminf_{x\to\infty} \frac{1}{x} \log \mu((\phi(x), \infty)) \le \varlimsup_{x\to\infty} \frac{1}{x} \log \mu((\phi(x), \infty)) \le -A_2$$

が成り立つならば

$$\begin{aligned}(1-\alpha)\left(\frac{\alpha}{A_1}\right)^{\alpha/(1-\alpha)} &\le \varliminf_{\lambda\to\infty} \frac{1}{\psi(\lambda)} \log \int_0^\infty e^{\lambda x}\mu(dx) \le \varlimsup_{\lambda\to\infty} \frac{1}{\psi(\lambda)} \log \int_0^\infty e^{\lambda x}\mu(dx) \\ \le (1-\alpha)\left(\frac{\alpha}{A_2}\right)^{\alpha/(1-\alpha)} &.\end{aligned} \tag{46}$$

(ii) 逆に，ある $-\infty < -A_1 \le -A_2 < 0$ に対して *(46)* が成り立つならば

$$-\frac{\lambda_2}{\lambda_1} A_2 \le \varliminf_{x\to\infty} \frac{1}{x} \log \mu((\phi(x), \infty)) \le \varlimsup_{x\to\infty} \frac{1}{x} \log \mu((\phi(x), \infty)) \le -A_2\,.$$

ここで， $\lambda_1 \le \lambda_2$ は $\xi^\alpha - \xi = (1-\alpha)\left(\frac{\alpha A_2}{A_1}\right)^{\alpha/(1-\alpha)}$ の二つの解.

証明. $f(x) = x$ のとき $g(x) = (1-\alpha)\left(\frac{\alpha}{-x}\right)^{\alpha/(1-\alpha)}$． ψ は指数 $1/(1-\alpha)$ の正則変動関数 (定理 62) だから主張を得る． □

系 77 ([18, Theorem 3]) $a(x)$, $x \ge 0$, は $a(0) = 0$ を満たす非減少右連続関数である $\lambda < \infty$ に対して $\int_0^\infty e^{-\lambda x} da(x) < \infty$ とする． x_0 を a の連続点とする．$\alpha > 0$ のとき以下が成り立つ.

(i) ある $-\infty \le -A_1 \le -A_2 \le 0$ に対して

$$-A_1 \le \varliminf_{x\downarrow 0} x^{1/\alpha} \log a(x) \le \varlimsup_{x\downarrow 0} x^{1/\alpha} \log a(x) \le -A_2$$

が成り立つならば

$$\begin{aligned}-(1+\alpha)\left(\frac{A_1}{\alpha}\right)^{\alpha/(1+\alpha)} &\le \varliminf_{\lambda\to\infty} \lambda^{-1/(1+\alpha)} \log \int_0^\infty e^{-\lambda x} da(x) \\ \le \varlimsup_{\lambda\to\infty} \lambda^{-1/(1+\alpha)} \log \int_0^\infty e^{-\lambda x} da(x) &\le -(1+\alpha)\left(\frac{A_2}{\alpha}\right)^{\alpha/(1+\alpha)}.\end{aligned} \tag{47}$$

(ii) 逆に，ある $-\infty \le -A_1 \le -A_2 < 0$ に対して *(47)* が成り立つならば

$$-\frac{\lambda_2}{\lambda_1} A_2 \le \varliminf_{x\downarrow 0} x^{1/\alpha} \log a(x) \le \varlimsup_{x\downarrow 0} x^{1/\alpha} \log a(x) \le -A_2\,.$$

ここで， $\lambda_1 \le \lambda_2$ は $x^{-\alpha} + x = (1+\alpha)\left(\frac{\alpha A_2}{A_1}\right)^{-\alpha/(1+\alpha)}$ の二つの解.

(iii) $\varliminf_{\lambda\to\infty} \lambda^{-1/(1+\alpha)} \log \int_0^\infty e^{-\lambda x} da(x) \ge -B \in [0,\infty])$ ならば $\varliminf_{x\downarrow 0} x^{1/\alpha} \log a(x) \ge -B^{(1+\alpha)/\alpha}$.

証明. $b(x) = a(1/x_0) - a(1/(x+0))$, $x > x_0$, および $b(x) = 0$, $0 \le x \le x_0$, とおけば db は Stieltjes 測度である．(42) で $f(x) = -1/x$ とすると，$g(x) = -(1+\alpha)\left(\dfrac{-x}{\alpha}\right)^{\alpha/(1-\alpha)}$．よって，定理 74 から，例えば,

$$\overline{\lim_{x\to\infty}}\, x^{-1/\alpha}\log(b(\infty) - b(x)) = A \in [-\infty, 0]$$

と

$$\overline{\lim_{\lambda\to\infty}}\, \lambda^{-1/(1+\alpha)}\log\int_0^\infty e^{-\lambda/x}db(x) = -(1+\alpha)\left(\frac{-A}{\alpha}\right)^{\alpha/(1+\alpha)}$$

は同値になる．これを書き換えると,

$$\overline{\lim_{x\downarrow 0}}\, x^{1/\alpha}\log a(x) = A$$

と

$$\overline{\lim_{\lambda\to\infty}}\, \lambda^{-1/(1+\alpha)}\log\int_0^\infty e^{-\lambda x}da(x) = -(1+\alpha)\left(\frac{-A}{\alpha}\right)^{\alpha/(1+\alpha)}$$

は同値になる．

以下同様. □

この節の残りで 定理 74 の証明を行う．

最初の補題はチェビシェフの不等式である．

補題 78 *(i)* $\overline{\lim_{\lambda\to\infty}}\dfrac{1}{\lambda}\log F(\lambda) \ge g(\overline{\lim_{x\to\infty}}\dfrac{1}{x}\log\mu((\phi(x),\infty)))$.

(ii) $\underline{\lim_{\lambda\to\infty}}\dfrac{1}{\lambda}\log F(\lambda) \ge g(\underline{\lim_{x\to\infty}}\dfrac{1}{x}\log\mu((\phi(x),\infty)))$.

証明. 後半も同様なので前半のみ証明する．

$A = \overline{\lim_{x\to\infty}}\dfrac{1}{x}\log\mu((\phi(x),\infty))$ とおく．$A = -\infty$ のときは $g(-\infty) = f(+0)$ から明らかなので $A > -\infty$ とする．$\xi > 0$ に対して

$$F(\lambda) \ge \int_{\phi(\xi\lambda)}^\infty e^{\lambda f(x/\phi(\lambda))}\mu(dx) \ge e^{\lambda f(\phi(\xi\lambda)/\phi(\lambda))}\mu((\phi(\xi\lambda),\infty)).$$

よって

$$\overline{\lim_{\lambda\to\infty}}\frac{1}{\lambda}\log F(\lambda) \ge f(\xi^\alpha) + A\xi.$$

□

次の補題が本質である．

補題 79 $-\infty < A < 0$ に対して $\overline{\lim_{x\to\infty}}\dfrac{1}{x}\log\mu((\phi(x),\infty)) \le A$ ならば，λ_0 を *(43)* の唯一の解とするとき，以下が成り立つ．

(i) $\overline{\lim_{\lambda\to\infty}}\dfrac{1}{\lambda}\log\displaystyle\int_{\phi(\eta\lambda)}^\infty e^{\lambda f(x/\phi(\lambda))}\mu(dx) \le f(\eta^\alpha) + A\eta$, $\eta > \lambda_0$.

(ii) $\overline{\lim_{\lambda\to\infty}}\dfrac{1}{\lambda}\log\displaystyle\int_0^{\phi(\eta\lambda)} e^{\lambda f(x/\phi(\lambda))}\mu(dx) \le f(\eta^\alpha) + A\eta$, $0 < \eta < \lambda_0$.

証明. $0 < \eta < \lambda_0$ の場合を証明する.（$\eta > \lambda_0$ も同様である.）

$$h_i(x,\delta) = f((1+\delta)x^{\alpha(1+(-1)^i\delta)}) + A(1-\delta)x, \quad i = 1, 2,$$

とおくと，これらは $\delta \in [0,1)$ について連続で $0 \le \delta \le (\frac{\beta}{\alpha} - 1) \wedge 1$ のとき x について上に凸．$\eta \ne \lambda_0$ ならば $h_i(\eta, 0) < h_i(\lambda_0, 0)$ なので，$0 < \eta < \lambda_0$ ならば，$c > 0$ と $0 < \delta_0 < (\frac{\beta}{\alpha} - 1) \wedge 1$ が存在して

$$\frac{h_i(\eta,\delta) - h_i(\lambda_0,\delta)}{\eta - \lambda_0} \ge c > 0, \quad \delta \in (0, \delta_0),\ i = 1, 2.$$

他方，h_i は上に凸なので

$$h_i(x,\delta) \le h_i(\eta,\delta) + \frac{h_i(\eta,\delta)-h_i(\lambda_0,\delta)}{\eta-\lambda_0}(x-\eta), \quad 0<x<\eta.$$

よって，$h(x,\delta)=h_1(x,\delta)\vee h_2(x,\delta)$ とおくと，

$$h(x,\delta) \le h(\eta,\delta)+c(x-\eta), \quad 0<x<\eta,\ 0<\delta<\delta_0\,.$$

次に，表現定理 (定理 61) と仮定より，任意の $\delta>0$ に対して $N_\delta>0$ が存在して

$$\frac{\phi(y)}{\phi(x)} \le \begin{cases} (1+\delta)\left(\dfrac{y}{x}\right)^{\alpha(1+\delta)}, & y\ge x\ge N_\delta\,, \\ (1+\delta)\left(\dfrac{y}{x}\right)^{\alpha(1-\delta)}, & x\ge y\ge N_\delta\,, \end{cases}$$

および

$$\mu((\phi(x),\infty)) \le e^{A(1-\delta)x}, \quad x\ge N_\delta,$$

が成り立つ．

$\epsilon>0$ を任意にとり，

$$\eta_k=\eta-\epsilon k, \quad k=1,2,\cdots,$$

とおく．$\eta_{k+1}\xi\ge N_\delta$ かつ $\xi\ge N_\delta$ のとき，

$$\begin{aligned}&\int_{\phi(\eta_{k+1}\xi)}^{\phi(\eta_k\xi)} e^{\xi f(x/\phi(\xi))}\mu(dx) \le \exp(\xi f(\phi(\eta_k\xi)/\phi(\xi)))\,\mu((\phi(\eta_{k+1}\xi),\infty)) \\ &\le \exp(\xi(f((1+\delta)\eta_k^{\alpha(1\pm\delta)})+A(1-\delta)\eta_{k+1})) \le \exp(\xi(h(\eta_k,\delta)-A(1-\delta)\epsilon)) \\ &\le \exp(\xi(h(\eta,\delta)-k\epsilon c-A(1-\delta)\epsilon)).\end{aligned}$$

よって $\xi\ge N_\delta$ のとき

$$\int_{\phi(N_\delta+\epsilon)}^{\phi(\eta\xi)} e^{\xi f(x/\phi(\xi))}\mu(dx) \le \sum_{k;\ \eta_{k+1}\xi\ge N_\delta}\int_{\phi(\eta_{k+1}\xi)}^{\phi(\eta_k\xi)} e^{\xi f(x/\phi(\xi))}\mu(dx) \le \frac{1}{1-e^{-c\epsilon\xi}}\,e^{\xi(h(\eta,\delta)-A(1-\delta)\epsilon)}.$$

これより

$$\varlimsup_{\xi\to\infty}\frac{1}{\xi}\log\int_{\phi(N_\delta+\epsilon)}^{\phi(\eta\xi)} e^{\xi f(x/\phi(\xi))}\mu(dx) \le h(\eta,\delta)-A(1-\delta)\epsilon$$

となるから，

$$\varlimsup_{\xi\to\infty}\frac{1}{\xi}\log\int_0^{\phi(\eta\xi)} e^{\xi f(x/\phi(\xi))}\mu(dx) \le f(+0)\vee(h(\eta,\delta)-A(1-\delta)\epsilon).$$

ここで $f(+0)$ は $\int_0^{\mathrm{const.}}$ からの寄与を $\log(A+B)\le\log 2+(\log A\vee\log B)$ と f の単調性を用いて計算すれば得られる．

$\epsilon\downarrow 0,\ \delta\downarrow 0$ とすれば主張を得る． □

残りの 3 つの補題は補題 79 をラプラス変換の評価にまとめたものである．

補題 80 $-\infty<A<0$ に対して $\varlimsup_{x\to\infty}\dfrac{1}{x}\log\mu((\phi(x),\infty))\le A$ ならば，$\varlimsup_{\lambda\to\infty}\dfrac{1}{\lambda}\log F(\lambda)\le g(A)$.

証明. λ_0 を補題 79 と同様 (43) の解とし，λ_1 と λ_2 を $0<\lambda_1<\lambda_0<\lambda_2<\infty$ を満たすように選ぶ．このとき，

$$\begin{aligned}&\varlimsup_{\xi\to\infty}\frac{1}{\xi}\log\int_{\phi(\lambda_1\xi)}^{\phi(\lambda_2\xi)} e^{\xi f(x/\phi(\xi))}\mu(dx) \le \varlimsup_{\xi\to\infty}\frac{1}{\xi}\log(\exp(\xi f(\phi(\lambda_2\xi)/\phi(\xi)))\,\mu((\phi(\lambda_1\xi),\infty))) \\ &\le f(\lambda_2^\alpha)+A\lambda_1\,.\end{aligned}$$

よって補題 79 から

$$\varlimsup_{\lambda\to\infty}\frac{1}{\lambda}\log F(\lambda) \le \max\{f(\lambda_1^\alpha)+A\lambda_1,\ f(\lambda_2^\alpha)+A\lambda_1,\ f(\lambda_2^\alpha)+A\lambda_2\}.$$

$\lambda_1\uparrow\lambda_0$ および $\lambda_2\downarrow\lambda_0$ として λ_0 の定義を思い出せば主張を得る． □

補題 81 $-\infty < A < 0$ に対して $\overline{\lim}_{x\to\infty} \frac{1}{x}\log\mu((\phi(x),\infty)) \le A$ であって，かつ，$\underline{\lim}_{\lambda\to\infty} \frac{1}{\lambda}\log F(\lambda) \ge B > f(+0)$ ならば，$\underline{\lim}_{x\to\infty} \frac{1}{x}\log\mu((\phi(x),\infty)) \ge \frac{\lambda_2}{\lambda_1}A$. ここで $0 < \lambda_1 \le \lambda_2 < \infty$ は *(44)* の2つの解.

証明. 仮定と補題 80 から $B \le g(A)$ を得るので，(44) は2つの解を持ち，$B = g(A)$ のとき，そのときに限り重解になる．η_1 と η_2 を $0 < \eta_1 < \lambda_1 \le \lambda_2 < \eta_2 < \infty$ を満たすように選ぶと，補題 79 から，

$$\overline{\lim_{\lambda\to\infty}} \frac{1}{\lambda}\log\int_0^{\phi(\eta_1\lambda)} e^{\lambda f(x/\phi(\lambda))}\mu(dx) \le f(\eta_1^\alpha) + A\eta_1 < f(\lambda_1^\alpha) + A\lambda_1 = B,$$
$$\overline{\lim_{\lambda\to\infty}} \frac{1}{\lambda}\log\int_{\phi(\eta_2\lambda)}^{\infty} e^{\lambda f(x/\phi(\lambda))}\mu(dx) \le f(\eta_2^\alpha) + A\eta_2 < f(\lambda_2^\alpha) + A\lambda_2 = B,$$

を得る．第2の仮定と合わせると，

$$\underline{\lim}_{\lambda\to\infty} \frac{1}{\lambda}\log\int_{\phi(\eta_1\lambda)}^{\phi(\eta_2\lambda)} e^{\lambda f(x/\phi(\lambda))}\mu(dx) \ge B.$$

他方，補題 80 の証明と同様に

$$\underline{\lim}_{\lambda\to\infty} \frac{1}{\lambda}\log\int_{\phi(\eta_1\lambda)}^{\phi(\eta_2\lambda)} e^{\lambda f(x/\phi(\lambda))}\mu(dx) \le f(\eta_2^\alpha) + \eta_1 \underline{\lim}_{\lambda\to\infty} \frac{1}{\lambda}\log\mu((\phi(\lambda),\infty))$$

を得るので，

$$\underline{\lim}_{\lambda\to\infty} \frac{1}{\lambda}\log\mu((\phi(\lambda),\infty)) \ge \frac{1}{\eta_1}(B - f(\eta_2^\alpha)).$$

λ_2 が (44) の解であることに注意すれば，$\eta_1 \uparrow \lambda_1$, $\eta_2 \downarrow \lambda_2$ によって主張を得る． □

補題 82 $-\infty \le A \le 0$ に対して $\overline{\lim}_{x\to\infty} \frac{1}{x}\log\mu((\phi(x),\infty)) \le A$ と $\overline{\lim}_{\lambda\to\infty} \frac{1}{\lambda}\log F(\lambda) \le g(A)$ は同値.

証明. g が単調なので補題 78 から後者から前者を得る．そこで前者から後者を導く．$g(0) = f(+\infty)$ だから $A = 0$ のときは自明．$-\infty < A < 0$ のときは補題 80 で証明されている．$A = -\infty$ のときは A を $A' > -\infty$ に取り替えた式が成り立つから今示したことから

$$\overline{\lim_{\lambda\to\infty}} \frac{1}{\lambda}\log F(\lambda) \le \inf_{0>A'>-\infty} g(A') = g(-\infty) = g(A).$$

□

定理 *74* の証明. 最初の主張は補題 78 と補題 82 から，第2の主張は補題 81 と補題 82 から，最後の主張は，

$$F(\xi) \le e^{\xi f(\phi(\lambda\xi)/\phi(\xi))}\mu((0,\phi(\lambda\xi))) + e^{\xi f(\infty)}\mu((\phi(\lambda\xi),\infty)), \quad \xi > 0,\ \lambda > 0,$$

から $f(\lambda^\alpha) < B$ のとき

$$\lambda \underline{\lim}_{x\to\infty} \frac{1}{x}\log\mu((\phi(\lambda\xi),\infty)) \ge -f(\infty) + B.$$

□

2000/12 追加．

Kasahara の結果は，次の対応で統計力学と熱力学の関係の定式化になっている．但し結果が漸近的逆関数を使って書かれていることと，ラプラス変換で x の符号が逆なので，少し対応がみにくい．[19, 20, 21] のほうがルジャンドル変換とラプラス変換の対応が統計力学と熱力学の関係としてみやすいと思うので，結果のみ引用しておく．

$\{U_\lambda\}_{\lambda\geq 1}$ を $U_\lambda(0)=0$ なる非減少右連続関数の族で

$$\log\int_0^\infty e^{-\lambda sx}dU_\lambda(x)<\infty,\quad s>0,\ \lambda\geq 1,$$

を満たすとする．

定理 83 ([21, 定理 1,2])

$$\varlimsup_{\lambda\to\infty}\frac{1}{\lambda}\log\int_0^\infty e^{-\lambda sx}dU_\lambda(x)<\infty,\quad s>0,$$

であって，さらに，非減少な $\phi^*:\ (0,\infty)\to[-\infty,\infty)$ があって

$$\lim_{\lambda\to\infty}\frac{1}{\lambda}\log U_\lambda(x)=\phi^*(x)$$

が ϕ^* の全ての連続点で成り立つならば，

$$\lim_{\lambda\to\infty}\frac{1}{\lambda}\log\int_0^\infty e^{-\lambda sx}dU_\lambda(x)=\sup_{x>0}(\phi^*(x)-sx),\quad s>0.$$

逆に，

$$\phi(s)=\lim_{\lambda\to\infty}\frac{1}{\lambda}\log\int_0^\infty e^{-\lambda sx}dU_\lambda(x),\quad s>0,$$

が存在し，連続微分可能で $\lim_{s\downarrow 0}\phi'(s)=-\infty,\ \lim_{\infty\to\infty}\phi'(s)=0,$ を満たすならば，

$$\lim_{\lambda\to\infty}\frac{1}{\lambda}\log U_\lambda(x)=\inf_{s>0}(\phi(s)+sx),\quad x>0\,.$$

前半がアーベル型，後半がタウバー型定理である．ルジャンドル変換から元の関数が再現できるための条件として後半での ϕ の連続微分可能性が本質的になる．

ϕ が連続微分可能でない場合のタウバー型定理は次のようになる．

定理 84 ([21, 定理 3])

$$\phi(s)=\lim_{\lambda\to\infty}\frac{1}{\lambda}\log\int_0^\infty e^{-\lambda sx}dU_\lambda(x),\quad s>0,$$

が存在するとする．

$$\phi^*(x)=\inf_{s>0}(\phi(s)+sx),\quad x>0$$

とおくと，上に凸な非減少関数であって，ϕ が狭義に上に凸になる任意の点 $x_0>0$ と任意の $x\geq x_0$ に対して

$$\phi^*(x_0)\leq\varliminf_{\lambda\to\infty}\frac{1}{\lambda}\log U_\lambda(x)\leq\varlimsup_{\lambda\to\infty}\frac{1}{\lambda}\log U_\lambda(x)\leq\phi^*(x).$$

x が狭義に上に凸な点では先の定理に帰着する． x が ϕ^* の直線部分にあるときに，x_0 を直線部分の下端点にとらざるを得ないというのが趣旨である．[30]

統計力学ー熱力学との対応は明らかである：

λ	測度 U	変数 x	漸近形 ϕ^*	ラプラス変換	変数 s	漸近形 ϕ
自由度	MCE	energy	entropy	CE	T^{-1}	$-T^{-1}$ free energy

(MCE: micro canonical ensemble, CE: canonical ensemble, T: temperature.)

より一般的な形，および，[18] との対応は [20] にある．

[30] x が直線部分にあるということは，熱力学では温度が一定のまま内部エネルギーが増加することなので， x_0 で相転移が起きて，$x-x_0$ の潜熱を蓄えている状況である．と思うが，詳しくは検討していない．

A.4 独立確率変数の和の平均 ー Cramér の公式.

この節は [23, §3] に従う.

X_n, $n \in \mathbf{N}$, を i.i.d. で，その分布 μ は遠方で十分早く decay するが，正のところで 0 にはならないとする．正確には，以下を仮定する．

(i) X_1 の母関数

$$M(t) = \mathrm{E}[\ e^{tX_1}\ ] = \int e^{tx} d\mu(x) \tag{48}$$

は t のある区間 (全区間ならもちろんよい) で存在.

(ii) (定理 86 の下からの評価の簡単のため)

$$\mathrm{P}[\ X_1 > x\ ] > 0,\ \ \mathrm{P}[\ X_1 < x\ ] > 0,\ \ x \in \mathbf{R}. \tag{49}$$

> $\frac{1}{n} W_n = \frac{1}{n} \sum_{k=1}^{n} X_k$, $n \in \mathbf{N}$, の分布を問題にする．$\frac{1}{n} W_n$ の $\mathrm{E}[\ X_1\ ]$ 付近 $O(n^{-1/2})$ 近傍の分布は中心極限定理で分かるが，$\mathrm{P}[\ \frac{1}{n} W_n \geq y\ ]$ の振る舞いを問題にする．

$\log M$ のルジャンドル変換 (40) を I とする：

$$I(x) = \sup_{t \in \mathbf{R}} (tx - \log M(t)),\ \ x \in \mathbf{R}. \tag{50}$$

命題 85 *(i)* $\log M$ も I も *(*値が有限の区間で*)* 連続関数で下に凸.

(ii) $I(x) \geq 0$, $x \in \mathbf{R}$.

(iii) $m = \mathrm{E}[\ X_1\ ]$ とおく（母関数があるから m はもちろん存在）と，$I(m) = 0$，即ち，m で最小値をとる．さらに，$x < m$ で減少 *(*非増加*)* し，$x > m$ で増加 *(*非減少*)* する．

(iv) $I(x) = \begin{cases} \sup_{t>0} (tx - \log M(t)), & x > m, \\ \sup_{t<0} (tx - \log M(t)), & x < m. \end{cases}$

証明. (i) 仮定の下で優収束定理から M は $M < \infty$ を満たす区間で連続関数．$M > 0$ だから $\log M$ も連続．同じく優収束定理から M は ($\log M$ も) 何回でも微分可能．特に，

$$\frac{d^2}{dt^2} \log M(t) = \frac{1}{2} \frac{\mathrm{E}[\ (X - \tilde{X})^2 e^{t(X+\tilde{X})}\ ]}{\mathrm{E}[\ e^{t(X+\tilde{X})}\ ]} \geq 0, \quad t \in \mathbf{R}.$$

ここで $\tilde{X}$ は X と同じ分布で独立な確率変数．この式から $\log M$ は下に凸.

I は $\log M$ のルジャンドル変換だから 命題 70 より下に凸.

(ii) $M(0) = 1$ より $I(x) \geq 0$, $x \in \mathbf{R}$.

(iii) Jensen の不等式 命題 13 から $M(t) \geq e^{mt}$, $t \in \mathbf{R}$, だから $mt - \log M(t) \leq 0$, $t \in \mathbf{R}$．よって $I(m) \leq 0$ だが，既に $I(x) \geq 0$ を証明しているから $I(m) = 0$，即ち，m で最小値をとる．

I は下に凸なので，$x < m$ で減少し，$x > m$ で増加する．

(iv) $x > m$ で $t \leq 0$ ならば上の証明の途中の評価から

$$tx - \log M(t) \leq tm - \log M(t) \leq 0,\ \ t \in \mathbf{R},$$

となって，$I(t) \geq 0$ は既に分かっているから，$t > 0$ で sup をとれば十分．同様に $x < m$ で $t \geq 0$ も sup に参加させる必要はない．

□

定理 86 (Cramér) $X_n,\ n \in \mathbf{N}$, を *i.i.d.* で *(49)* を満たし，母関数 $M(t) = \mathrm{E}[\ e^{tX_1}\ ]$ が t のある区間で存在するとする．

$\log M$ のルジャンドル変換 I を *(50)* で定義し，$\dfrac{1}{n}W_n = \dfrac{1}{n}\displaystyle\sum_{k=1}^{n} X_k$, $n \in \mathbf{N}$, とおく．このとき以下が成り立つ．

$F \subset \mathbf{R}$ が閉集合ならば

$$\varlimsup_{n\to\infty} \frac{1}{n} \log \mathrm{P}[\ \frac{1}{n}W_n \in F\ ] \le -\inf_{x\in F} I(x). \tag{51}$$

$G \subset \mathbf{R}$ が開集合ならば

$$\varliminf_{n\to\infty} \frac{1}{n} \log \mathrm{P}[\ \frac{1}{n}W_n \in G\ ] \ge -\inf_{x\in G} I(x). \tag{52}$$

注. (i) 従って特に，$\displaystyle\inf_{x\in A^o} I(x) = \inf_{x\in \bar{A}} I(x)$ なるボレル集合 $A \in \mathcal{B}_1$ に対しては

$$\lim_{n\to\infty} \frac{1}{n} \log \mathrm{P}[\ \frac{1}{n}W_n \in A\ ] = -\inf_{x\in A} I(x).$$

このように独立確率変数の平均に対して，確率測度の漸近形を母関数から導くことができる．

特に 命題 85 と合わせると，例えば

$$\lim_{n\to\infty} (\mathrm{P}[\ \frac{1}{n}W_n \le a\ ])^{1/n} = \begin{cases} \displaystyle\inf_{t<0} \mathrm{E}[\ e^{t(X_1-a)}\ ], & a < m, \\ 1, & a \ge m. \end{cases}$$

と，母関数の漸近的振る舞いから確率測度の漸近的振る舞いが得られる様子が分かる．

(ii) (49) が成り立っていなくても結論は成り立つ．次節指数分布の例はそれを前提にして説明する．

◇

証明.

$$\begin{aligned} \mathrm{P}[\ \frac{1}{n}W_n \ge y\ ] &= \int_y^{\infty} d(\mathrm{P}\circ(W_n/n)^{-1}) \\ &\le e^{-ty}\mathrm{E}[\ e^{tW_n/n}\ ] = e^{-ty}\mathrm{E}[\ e^{X_1 t/n}\ ]^n = e^{-ty} M\left(\frac{t}{n}\right)^n, \quad t \in \mathbf{R}. \end{aligned}$$

t に nt を代入して両辺の $\displaystyle\inf_{t>0}$ をとると，命題 85 より $y > m$ のとき，

$$\frac{1}{n} \log \mathrm{P}[\ \frac{1}{n}W_n \ge y\ ] \le -I(y), \quad y > m. \tag{53}$$

$y < m$ のときは $\inf_{t<0}$ をとることで同様の不等式が成り立つ．

F が閉集合とする．

$m \in F$ ならば，確率が 1 以下だから $\log \mathrm{P}[\ \dfrac{1}{n}W_n \in F\ ] \le 0$．他方で， 命題 85 から $\displaystyle\inf_{x\in F} I(x) = I(m) = 0$．よってこのとき (51) が成り立つ．

$m \notin F$ ならば，F^c の中で m を含む連結成分 (開区間) を (y_1, y_2) とおくと，$F \subset (-\infty, y_1] \cup [y_2, \infty)$ だから，(53) より

$$\varlimsup_{n\to\infty} \frac{1}{n} \log \mathrm{P}[\ \frac{1}{n}W_n \in F\ ] \le -(I(y_1) \wedge I(y_2)).$$

（$y_1 = -\infty$ または $y_2 = \infty$ の場合は，考慮対象になる項は一方のみ.）命題 85 より，$I(x)$ は $x < m$ で減少し，$x > m$ で増加するから，

$$I(y_1) \wedge I(y_2) = \inf_{y \in F} I(y)$$

なので，このときも (51) が成り立つ.

補題 87 任意の $\delta > 0$ と $y \in \mathbf{R}$ に対して

$$\varliminf_{n\to\infty} \frac{1}{n} \log \mathrm{P}[\, \frac{1}{n} W_n \in (y-\delta, y+\delta) \,] \geq -I(y) \tag{54}$$

ならば *(52)* が成り立つ.

証明. $\epsilon > 0$ を任意に選んで，$I(y) < \inf_{x\in G} I(x) + \epsilon$ を満たす $y \in G$ をとり，$(y-\delta, y+\delta) \subset G$ となるように $\delta > 0$ を選ぶ．すると，仮定から

$$\varliminf_{n\to\infty} \frac{1}{n} \log \mathrm{P}[\, \frac{1}{n} W_n \in G \,] \geq \varliminf_{n\to\infty} \frac{1}{n} \log \mathrm{P}[\, \frac{1}{n} W_n \in (y-\delta, y+\delta) \,] > -\inf_{x\in G} I(x) - \epsilon.$$

$\epsilon > 0$ は任意だから (52) が成り立つ. □

補題 87 より (54) を証明すれば十分である．$\delta > 0$, $y \in \mathbf{R}$, とする．仮定 (49) から,

$$\lim_{|t|\to\infty} \frac{1}{|t|} \log M(t) = \infty$$

なので，$I(y) = \sup_{t\in\mathbf{R}} (ty - \log M(t))$ の sup は途中のどこか ($\pm\infty$ でなく実数値) $t_0 = t_0(y)$ で達成する．即ち,

$$I(y) = t_0 y - \log M(t_0), \quad y = \frac{M'(t_0)}{M(t_0)}.$$

$$\mu_{t_0}(A) = \frac{1}{M(t_0)} \mathrm{E}[\, e^{t_0 X_1}; A \,], \quad A \in \mathcal{B}_1,$$

で定義される分布[31] μ_{t_0} を考える.

$$\int x \mu_{t_0}(dx) = \frac{1}{M(t_0)} \mathrm{E}[\, X_1 e^{t_0 X_1} \,] = \frac{M'(t_0)}{M(t_0)} = y$$

なので，大数の法則から任意の $\delta_1 > 0$ に対して

$$\lim_{n\to\infty} \int_{|\frac{1}{n}(x_1+\cdots+x_n)-y|<\delta_1} \mu_{t_0}(dx_1)\cdots\mu_{t_0}(dx_n) = 1\,. \tag{55}$$

他方，$\delta_1 < \delta$ ならば

$$\begin{aligned}
&\mathrm{P}[\, \frac{1}{n} W_n \in (y-\delta, y+\delta) \,] \geq \mathrm{P}[\, \left| \frac{1}{n} \sum_{k=1}^{n} X_k - y \right| < \delta_1 \,] \\
&\geq e^{-n(t_0 y + |t_0|\delta_1)} \mathrm{E}[\, e^{t_0 W_n};\ |\frac{1}{n} W_n - y| < \delta_1 \,] \\
&= e^{-n(t_0 y + |t_0|\delta_1)} M(t_0)^n \int_{|\frac{1}{n}(x_1+\cdots+x_n)-y|<\delta_1} \mu_{t_0}(dx_1)\cdots\mu_{t_0}(dx_n).
\end{aligned}$$

ここに (55) と t_0 の定義を使えば,

$$\varliminf_{n\to\infty} \frac{1}{n} \log \mathrm{P}[\, \frac{1}{n} W_n \in (y-\delta, y+\delta) \,] \geq -I(y) - \delta_1 |t_0|.$$

$0 < \delta_1 < \delta$ ならば成り立つので (54) が成り立つ. □

[31] 統計力学で出てくる canonical ensemble!?

A.5 指数分布の和の分布の極限．

Cramér の定理の例として指数分布の和の分布を取り上げる．

定理 88 次の公式が成り立つ[32]．

$$\lim_{n\to\infty}\left(\sum_{k=n}^{\infty}\frac{n^k x^k}{k!}\right)^{1/n} = \begin{cases} e\cdot x, & 0\le x<1,\\ e^x, & x\ge 1,\end{cases}$$

注. どこが難しそうか？

$$G_n(x) = \begin{cases} e^{-x}\displaystyle\sum_{k=n}^{\infty}\frac{x^k}{k!} = 1-e^{-x}\sum_{k=0}^{n-1}\frac{x^k}{k!}, & x\ge 0,\\ 0, & x<0,\end{cases} \tag{56}$$

とおくと，求める公式は

$$\lim_{n\to\infty} G_n(nx)^{1/n} = \begin{cases} xe^{1-x}, & 0\le x\le 1,\\ 1, & x>1\,.\end{cases}$$

まず x が 1 より大きいか小さいかで大きく関数形が変わる．さらに，

- $\displaystyle\lim_{n\to\infty} G_n(x) = 0$, $x\ge 0$,
- $\displaystyle\lim_{x\to\infty} G_n(x) = 1$, $n\ge 0$,

は定義から容易に分かる．従って， x を xn に置き換えて $n\to\infty$ とすることにより x と n を同時に大きくしたらどうなるかが微妙である． ◇

定理 *88* **の証明**. $X_n,\ n\in\mathbf{N}$, を i.i.d. で $\mathrm{P}[\,X_1\in A\,] = \displaystyle\int_A e^{-x}dx,\ A\in\mathcal{B}_1$, を満たすものとし，$W_n = X_1+\cdots+X_n$ とする．

$M(t) = \mathrm{E}[\,e^{tX_1}\,] = \dfrac{1}{1-t}$, $t<1$, だから，

$$I(x) = \sup_{t\in\mathbf{R}}(tx-\log M(t)) = \begin{cases} x-1-\log x, & x>0,\\ \infty, & x\le 0\,.\end{cases}$$

定理 86 より (注を前提にして)

$$\lim_{n\to\infty}\frac{1}{n}\log\mathrm{P}[\,\frac{1}{n}W_n<a\,] = -(\inf_{0<x<a}(x-1-\log x)\wedge\infty) = \begin{cases} -\infty, & a\le 0,\\ 1-a+\log a, & 0<a<1,\\ 0, & a\ge 1\,.\end{cases} \tag{57}$$

他方，直接計算することにより，

$$d\mathrm{P}\circ(W_n/n)^{-1}(x) = \chi_{[0,\infty)}\frac{n^n}{(n-1)!}e^{-nx}x^{n-1}dx \tag{58}$$

だから，

$$\lim_{n\to\infty}\frac{1}{n}\log\int_0^a\frac{n^n}{(n-1)!}e^{-nx}x^{n-1}dx = \begin{cases} 1-a+\log a, & 0<a<1,\\ 0, & a\ge 1\,.\end{cases}$$

[32] この形に書いたのは [27, §3.4] が最初．そこにこの公式と binary tree の関係も触れている．

少し書き換える．

$$\int_0^a \frac{n^n}{(n-1)!}e^{-nx}x^{n-1}dx = \frac{1}{(n-1)!}\int_0^{an} e^{-y}y^{n-1}dy = g_n(an);$$
$$g_n(x) = \frac{1}{(n-1)!}\int_0^x e^{-y}y^{n-1}dy,\ x \geq 0\,.$$

他方，(56) の G_n は $x \geq 0$ で $G_n(0) = 0$ と $G'_n(x) = \dfrac{1}{(n-1)!}e^{-x}x^{n-1}$ を満たすので $g_n = G_n$．以上から，$n \to \infty$ とすると 定理 88 を得る．

(58) は，例えば特性関数を用いて次のように証明できる．

$$\mathrm{E}[\, e^{\sqrt{-1}W_n t/n}\,] = \left(\int_0^\infty e^{-x+\sqrt{-1}xt/n}dx\right)^n = \left(1-\sqrt{-1}\frac{t}{n}\right)^{-n},\quad t \in \mathbf{R}.$$

他方，

$$\int_0^\infty e^{\sqrt{-1}xt}\frac{n^n}{(n-1)!}e^{-nx}x^{n-1}dx = \frac{n^n}{(n-1)!}\frac{1}{(n-\sqrt{-1}t)^n}\int_0^\infty e^{-y}y^{n-1}dy$$
$$= \left(1-\sqrt{-1}\frac{t}{n}\right)^{-n},\quad t \in \mathbf{R}.$$

□

§A.6 でより正確に言うが，定理 86 のような評価が成り立つとき $\{P_n\}$ は large deviation principle を満たす，という．(57) に即して説明すると，大数の法則から $\dfrac{1}{n}W_n$ は期待値 $m = \mathrm{E}[\, X_1\,]$ に概収束する．しかも中心極限定理から m のまわりの $O(\sqrt{n})$ くらいの広がりに確率は集中している．これが $a \geq 1 = m$ のときの (57) の意味である．

逆に，m から $O(1)$ 離れたところの分布の様子を与えるのが，(57) の $0 < a < 1$ の non-trivial な振る舞いである．即ち，そこでは

$$\mathrm{P}[\, \frac{1}{n}W_n < a\,])^{1/n} \simeq e^{-n(a-1-\log a)}$$

のように，指数関数的に小さくなる，ということを主張している．確率の大部分がある平均値の周辺から離れたところの確率の一般的な振る舞い，という意味で LDP と名付けられていると思われる．

A.6 Large deviation principle.

独立確率変数の和の LDP については §A.4 でみたように解決している．現在 LDP というときは，研究対象はもっぱら測度空間または関数空間上の確率測度，即ち，測度または関数に値を取る確率変数，言い換えると確率（超）過程である．そのような複雑な集合に対しても一般的に成り立ちうる形の「結論」(このような結論を目指して定理を作ろう，という提唱？）を定式化したのが LDP である．

Brown 運動，特にその先にある確率微分方程式，あるいは統計力学，などと関連させると，深い話題があるが，ここではとても言及できない．

[23, §2] から一般的な LDP を参考までに引用するにとどめる．

Ω を Poland 空間 (完備可分距離空間)，$\mathcal{F}$ をボレル集合族，P_ϵ, $\epsilon > 0$, を $(\Omega, \mathcal{F})$ 上の確率測度の族とする．主に想定しているのは $\epsilon \downarrow 0$ で P_ϵ が単位分布 δ_{x_0} に弱収束する状況である．

例えば実数上の密度 ρ を持つ確率測度 $\mathrm{P}[\, dx\,] = \rho(x)dx$ に対して $\mathrm{P}_n[\,(a,\infty)\,] = \mathrm{P}[\,(na,\infty)\,]$, $n \in \mathbf{Z}_+$, とすると $n \to \infty$ で P_n は δ_0 に弱収束する．§A.3.1 で取り上げた問題はこの状況である．

$\{\mathrm{P}_\epsilon\}$ が rate function I の large deviation principle に従うとは, $I:\ \Omega\to[0,\infty]$ が以下を満たすことを言う[33].

(i) I は lower semicontinuous, 即ち任意の $\epsilon>0$ と $\omega\in\Omega$ に対して $\delta>0$ がとれて, $d(\omega,\omega')<\delta$ ならば $f(\omega')>f(\omega)-\epsilon$.

(ii) 各 $\ell<\infty$ に対して $\{\omega\in\Omega\mid I(\omega)\le\ell\}$ が Ω のコンパクト集合.

(iii) $F\subset\Omega$ が閉集合ならば
$$\varlimsup_{n\to\infty}\epsilon\log\mathrm{P}_\epsilon[\,F\,]\le-\inf_{x\in F}I(x).$$

(iv) $G\subset\Omega$ が開集合ならば
$$\varliminf_{n\to\infty}\epsilon\log\mathrm{P}_\epsilon[\,G\,]\ge-\inf_{x\in G}I(x).$$

注. 命題 70 よりルジャンドル変換は lower semicontinuous なので rate function の資格がある.(命題 69 の注より, 単に下に凸というだけでは不十分.) 例えば (50) の I は $\log M$ のルジャンドル変換. ◇

δ_{x_0} に収束する状況では $I(x_0)=0$ と $I(x)>0,\ x\ne x_0$, が典型的な期待される状況である.

定義と母関数の関係を象徴する簡単な定理がある.

定理 89 $\{\mathrm{P}_\epsilon\}$ が *rate function* I の *large deviation principle* に従うとする. このとき, Ω 上の任意の有界連続関数 F に対して
$$\lim_{\epsilon\downarrow0}\epsilon\log\left(\int_\Omega\exp\left(\frac{1}{\epsilon}F\right)\,d\mathrm{P}_\epsilon\right)=\sup_{\omega\in\Omega}(F(x)-I(x)).$$
証明は長くないが, [23, §2 定理 2.2] にゆだねる.

この表式を見ると, $\epsilon\downarrow0$ は一般的にはスピン系の統計力学でいう熱力学的極限, $-I$ は熱力学でいうエントロピーに対応しているらしい.

また, Cramér の定理は $\epsilon\downarrow0$ $(n\to\infty)$ が i.i.d. r.v. の熱力学的極限になっているという意味で平均場近似に, I はそのときの自由エネルギーを温度で割った F/T の $\beta=1/T$ に関するルジャンドル変換(即ちエントロピー)に対応しているらしい.

もちろん, 抽象化されているので, スケール変換の取り方によって, 流体力学極限もこの範疇で研究されている. 物理との対応は一対一ではない.

B 1次元ランダムウォーク.

B.1 定義と説明.

$d\in\mathbf{N}$ と $x_0\in\mathbf{Z}^d$ に対して, 原点の隣の格子点の集合 $S=\{x\in\mathbf{Z}^d\mid|x|=1\}$ とおくとき, S に値をとる i.i.d. 確率変数列 $X_n:\Omega\to H,\ n\in\mathbf{Z}_+$, が S の $2d$ 個の要素全てを等しい確率 $\mathrm{P}[\,X_1=h\,]=(2d)^{-1}$ でとるとき, その和が作る確率変数列 $W_n=x_0+\sum_{k=1}^{n}X_k,\ n\in\mathbf{Z}_+$, を d 次元 simple random walk という.

[33] 以下の定義は母関数との関係が不鮮明な点が私は不満である. 技術的には Tauber 型定理の無限次元版という性格のものだと思われるので(あるいは, 統計力学からルジャンドル変換もしくは micro-canonical ensemble と canonical ensemble の対応の抽象化として得られたものだと思われるので), I が母関数から得られる点が explicit なほうがいいと感じる. 恐らく, しかし, それができたならば principle (指導原理) ではなく定理と書かれたことであろう. 無限次元測度空間では簡単に行かないことが多いらしい.

さいころをふって格子点上を動くすごろく（あるいはでたらめに歩く酔っぱらい）の時刻 n における位置を W_n と書く，という描像から random walk という名前が由来する．

$d=1$ の場合は §3.3 の独立確率変数列の和の議論（で，$\{\pm 1\}$ に値をとるもっとも単純な場合）に他ならないが，視点が，個々の確率変数の分布の一般化などよりも，分布は単純なものにしておいて，連続極限を含む極限のより精密な性質（特に，単に W_n という値の漸近的性質を見るのでなく，path の性質，即ち $\{W_1,\cdots,W_n\}$ 全体に依存する「図形の確率論」）や，独立でない場合などに重点を移すときに random walk という言葉が使われる印象がある．背景には拡散現象の物理学や，より遠い背景には臨界現象の統計力学からの刺激がある．

Random walk というとき何を指すかは人によって多少の違いがあるように感じるが，もっとも広く言えば，Markov chain のことを指す.（Markov chain とは W_{n+1} が $W_n=a$ で条件付ければ W_k, $k=1,\cdots,n-1$, と独立になることをいう.）例えば $\{W_n\}$ が独立確率変数列の部分和の列の場合が random walk の典型である．W_n が $\mathbf{Z}^d$ に値をとるときは，後者を random walk と呼ぶのが普通のようである．

以下，ここでは $\{X_n\}$ を i.i.d. で ± 1 を確率 1/2 でとる確率変数，$W=(W_1,W_2,\cdots)$; $W_n = x_0+\sum_{k=1}^{n} X_k$, $n\in\mathbf{Z}_+$, を $x_0\in\mathbf{Z}$ から出発する simple random walk，とする．x_0 は特に断らなければ 0 にとることが多い．

$W(\omega)$, $(W_1,\cdots,W_n)(\omega)$ を平面内の折れ線に翻訳しよう．

平面内の連続した折れ線グラフで $(0,x_0)$ から出発して傾き $\pm 45^\circ$ で長さ $\sqrt{2}$ の線分を x 軸正方向につないでつくったものを walk (path) と言う．Random walk のイメージから各線分要素を１歩という．Path の全体を $\mathcal{W}$ とおく．そのうち最初の n 歩の部分を n 歩の path と呼んで，その全体を $\mathcal{W}_n$ とおく．$\sharp\mathcal{W}_n=2^n$ は明らか．

$\omega\in\Omega$ を１つとる．$k=1,2,\cdots$ に対して $(k-1,W_{k-1}(\omega))$ と $(k,\cdots,W_k(\omega))$ を結んで折れ線グラフをつくることで $W:\ \Omega\to\mathcal{W}$ なる対応を得る．また，各 $n\in\mathbf{Z}_+$ 毎に $(W_1,\cdots,W_n):\ \Omega\to\mathcal{W}_n$ なる対応がある．定義から全ての path がそれぞれ確率 2^{-n} で現れる．

Simple random walk の研究に有益でそれぞれ重要な一般化がある概念や性質を紹介しながら，簡単（だが，自明ではない）性質をいくつか紹介する．

B.2 Stopping time と区間からの脱出．

§C.1.2 でも再述するが，独立して読めるように，同じことをここでも記述しておく．

確率空間 $(\Omega,\mathcal{F},\mathrm{P})$ 上の確率変数列 X_n, $n\in\mathbf{N}$, に対して確率変数 $\tau:\ \Omega\to\mathbf{N}\cup\{\infty\}$ が stopping time であるとは，各 $n\in\mathbf{N}$ に対して $\{\sigma\le n\}$ が $X_1,\cdots,X_n$ が生成する σ 加法族 $\mathcal{F}_n=\sigma[X_1,\cdots,X_n]$ の要素になっていることをいう．

ここで $\sigma[X_1,\cdots,X_n]$ は $X_1,\cdots,X_n$ が生成する σ 加法族，即ち $X_i^{-1}(\mathcal{B}_1)$, $i=1,\cdots,n$, たちを含む最小の σ 加法族．§C.1.1 も参照．

$\mathcal{F}_1\subset\mathcal{F}_2\subset\mathcal{F}_3\subset\cdots$ だから，τ が stopping time になることは

$$\{\tau=n\}\in\mathcal{F}_n,\quad n\in\mathbf{N},$$

と同値である．即ち，τ が stopping time であるとは，時刻（無限大を許す）に値をとる確率変数であって，τ がある時刻になることがその時刻までの path によって決まることをいう．

典型例：Hitting time. １次元 simple random walk W_n と $A\subset\mathbf{Z}$ に対して $\tau=\inf\{n\in\mathbf{Z}_+\mid W_n\in A\}$ なる τ を A の hitting time という（条件不成立なら $\tau=\infty$ の約束）．Hitting time は stopping time である．

情報は σ 加法族で決まるので，一般に，以上の定義を次のように拡張する．可測空間 $(\Omega,\mathcal{F})$ と $\mathcal{F}$ の部分 σ 加法族の列で増大するもの $\mathcal{F}_1\subset\mathcal{F}_2\subset\mathcal{F}_3\subset\cdots$ が与えられたとき，$\tau:\ \Omega\to\mathbf{N}\cup\{\infty\}$ が stopping time であるとは，

$$\{\tau\le n\}\in\mathcal{F}_n,\quad n\in\mathbf{N},$$

が成り立つことをいう．

> Stopping time は σ 加法族の増大列によって定まるので，σ 加法族列が複数出てくるときは $\{\mathcal{F}_n\}$–Stopping time とも書く．

Stopping time の一つの興味深い使い方は W_τ という確率変数の期待値を考えるときである．

定理 90 (Wald) $\{X_n\}$ が *i.i.d.* で $\mathrm{E}[\,|X_1|\,]<\infty$ とし，$\mathcal{F}_n=\sigma[X_1,\cdots,X_n]$, $W_n=X_1+\cdots+X_n$, $n\in\mathbf{N}$ とおく．τ は $\mathrm{E}[\,\tau\,]<\infty$ を満たす $\{\mathcal{F}_n\}$*–stopping time* とする．このとき

$$\mathrm{E}[\,W_\tau\,]=\mathrm{E}[\,X_1\,]\mathrm{E}[\,\tau\,].$$

注. i.i.d. の和でなくてもマルチンゲールならば，関連する結果を（上の notation 設定では右辺が 0 という形で）得る（Doob の任意停止定理）． ◇

証明. $\mathrm{E}[\,\tau\,]<\infty$ だから $\tau\neq\infty$, a.s. に注意．即ち，最初から $\tau\in\mathbf{N}$ としてよいので，$\chi_\Omega=\sum_{n=1}^{\infty}\chi_{\tau=n}$, a.s.

[13, §3.1 Theorem (1.6)] に従う．まず $X_k\geq 0$, $k=1,\cdots,n$, の場合を扱う．和の順序交換が自由になるので

$$\begin{aligned}\mathrm{E}[\,W_\tau\,]&=\sum_{n=1}^{\infty}\mathrm{E}[\,W_n\chi_{\tau=n}\,]=\sum_{n=1}^{\infty}\sum_{k=1}^{n}\mathrm{E}[\,X_k\chi_{\tau=n}\,]\\&=\sum_{k=1}^{\infty}\sum_{n=k}^{\infty}\mathrm{E}[\,X_k\chi_{\tau=n}\,]=\sum_{k=1}^{\infty}\mathrm{E}[\,X_k\chi_{\tau\geq k}\,].\end{aligned}$$

ここで，$\{\tau\geq k\}=\{\tau<k-1\}^c\in\mathcal{F}_{k-1}$ だから，特に X_k と独立であることに注意し，また $\mathrm{E}[\,X_k\,]=\mathrm{E}[\,X_1\,]$ だから，

$$\mathrm{E}[\,W_\tau\,]=\sum_{k=1}^{\infty}\mathrm{E}[\,X_k\,]\mathrm{P}[\,\tau\geq k\,]=\mathrm{E}[\,X_1\,]\mathrm{E}[\,\tau\,].$$

一般の場合，この結果，特に $|X_k|$ に対して以上の証明が成り立つので，

$$\infty>\mathrm{E}[\,|X_1|\,]\mathrm{E}[\,\tau\,]=\sum_{k=1}^{\infty}\sum_{n=k}^{\infty}\mathrm{E}[\,|X_k|\chi_{\tau=n}\,].$$

従って，元の級数は絶対収束するので，和の順序交換がやはり自由になるので $\mathrm{E}[\,W_\tau\,]=\mathrm{E}[\,X_1\,]\mathrm{E}[\,\tau\,]$ を得る． □

0 から出発する simple random walk の場合に応用する．即ち，$\{X_n\}$ を i.i.d. で $\mathrm{P}[\,X_1=-1\,]=\mathrm{P}[\,X_1=1\,]=1/2$ なるものとし，$W_n=\sum_{k=1}^{n}X_k$ とする．

0 をはさんで 2 つの整数 $-a<0<b$ をとる．0 から出発する simple random walk が（行ったり来たりしながらやがて）$-a$ または b にたどり着いたらそこで止まる，という状況を考える．

硬貨を投げて表なら右裏なら左に行くすごろくで，$-a$ に着いたら A の勝ち，b に着いたら B の勝ち，というゲームを想像すればよい．問題は，A, B それぞれの勝つ確率である．

Simple random walk を stopping time $\tau=\inf\{n\in\mathbf{N}\mid W_n\in\{-a,b\}\}$ までの時間で見る．定理 90 が使えるには $\mathrm{E}[\,\tau\,]<\infty$ が必要．その証明を後回しにして，定理 90 を使うと，$\mathrm{E}[\,X_1\,]=0$ なので，

$$b\mathrm{P}[\,W_\tau=b\,]-a\mathrm{P}[\,W_\tau=-a\,]=\mathrm{E}[\,W_\tau\,]=\mathrm{E}[\,X_1\,]\mathrm{E}[\,\tau\,]=0\,.$$

($\mathrm{E}[\,\tau\,] < \infty$ なので $\tau \in \mathbf{N}$, a.e., に注意．）$\mathrm{P}[\,W_\tau = b\,] + \mathrm{P}[\,W_\tau = -a\,] = 1$ と合わせると，

$$\mathrm{P}[\,W_\tau = -a\,] = \frac{b}{a+b}, \quad \mathrm{P}[\,W_\tau = b\,] = \frac{a}{a+b}\,. \tag{59}$$

こうやってこのすごろくの勝敗確率が計算できる！ちょうど内分点をベクトル表示するときの係数と等しくなっていて，覚えやすい.（簡単な理解ができる結果なのかどうか分からなかった.）

残っていた $\mathrm{E}[\,\tau\,] < \infty$ の証明.

$-a < x < b$ ならば $b + a$ 歩まっすぐ進めば必ず開区間 $(-a, b)$ から出るので，$\mathrm{P}[\,(x + W_{b+a}) \in (-a, b)\,] \le 1 - 2^{-(b+a)}$．このまっすぐ進行評価を n 回繰り返すと $\mathrm{P}[\,\tau > n(a+b)\,] \le (1 - 2^{-(b+a)})^n$. 特に，

$$\mathrm{P}[\,\tau = \infty\,] = \mathrm{P}[\,\bigcap_{n=1}^{\infty} \tau > n(a+b)\,] = \lim_{n\to\infty} \mathrm{P}[\,\tau > n(a+b)\,] = 0\,.$$

よって

$$\mathrm{E}[\,\tau\,] = \sum_{m=1}^{\infty} m\mathrm{P}[\,\tau = m\,] \le \sum_{n=1}^{\infty} (a+b)^2 n (1 - 2^{-(b+a)})^{n-1} < \infty.$$

特に，このことから，1 次元 simple random walk が各点にいつかは到達するか？という問に答えられる.

$T_a = \inf\{n \in \mathbf{N} \mid W_n = -a\}$, $T_b = \inf\{n \in \mathbf{N} \mid W_n = b\}$ とおくと (59) は $\mathrm{P}[\,T_a < T_b\,] = \dfrac{b}{a+b}$ を意味する．ここで $b \to \infty$ とすると，

$$\mathrm{P}[\,T_a < \infty\,] \ge \sup_{b>0} \mathrm{P}[\,T_a < T_b\,] = 1, \quad a < 0\,.$$

対称性と $T_0 = 0$ から全ての $x \in \mathbf{Z}$ に対して，$T_x < \infty$, a.e., 即ち，

1 次元 simple random walk はどの点も殆ど必ず hit する.

一方，hit するまでの平均時間は（最初から hit している原点以外は）∞ であることも分かる．なぜなら，もし $\mathrm{E}[\,T_x\,] < \infty$ ならば定理 90 から $x = \mathrm{E}[\,W_{T_x}\,] = \mathrm{E}[\,X_1\,]\mathrm{E}[\,T_x\,] = 0$ となって，$x \ne 0$ では矛盾するから.

B.3 Reflection principle と arcsine law.

この節は [13, §3.3] に従う[34]

§B.1 に説明したように， simple random walk $\{W_n\}$ の sample を平面内の折れ線 (path) に対応させせることができて，n 歩目までの各 path は 2^{-n} の確率で現れる.

$n, x \in \mathbf{N}$ に対して $W_n = x$ となる sample の集合（明らかに可測集合）について考える．平面内の $(0, 0)$ から (n, x) への折れ線（であって，一歩進む毎に x 方向に 1，y 方向に ± 1 進むもの）の集合である.

$a = (n + x)/2$, $b = (n - x)/2$ とおけば $n = a + b$, $x = a - b$ である．$n - x$ が偶数ならば， $(0, 0)$ から (n, x) への path の総数 $N_{n,x}$ は $N_{n,x} = \begin{pmatrix} n \\ a \end{pmatrix}$ であり，奇数ならば $N_{n,x} = 0$.

Simple random walk ではどの n 歩の path も出現する確率は 2^{-n} なので，以上のことから特に，$2n$ 歩で出発点 $W_0 = 0$ に戻る確率を $u_{2n} = \mathrm{P}[\,W_{2n} = 0\,]$ と書くと，

$$u_{2n} = \mathrm{P}[\,W_{2n} = 0\,] = 2^{-2n} N_{2n,0} = \begin{pmatrix} 2n \\ n \end{pmatrix} 2^{-2n} \asymp \frac{1}{\sqrt{\pi n}}, \quad n = 0, 1, 2, \cdots. \tag{60}$$

[34] [13, §3.3] は「Feller 3 章に従う」と書いてあるので,「[15, §III.1] に従う」というのが正しい.

ここで漸近形（$n \to \infty$ で比の極限が 1 に収束するという意味で）は Stirling の公式

$$n! \asymp \sqrt{2\pi n}e^{-n}n^n$$

を用いた．また，一般に

$$\mathrm{P}[\, W_{2n} = 2k \,] = 2^{-2n}N_{2n,2k} = \binom{2n}{n+k}2^{-2n}, \quad k = 0, 1, \cdots, n. \tag{61}$$

次の命題は，単純ながら，path の到達点と本数の関係を調べるときに有効な普遍的方法である．

命題 91 (Reflection principle) *$n, x, y \in \mathbf{N}$ ならば，$(0, -x)$ から (n, y) への path の本数は，$(0, x)$ から (n, y) への path のうち x 軸と原点以外で交点を持つ（触る）ものの本数に等しい．*

証明. $(0, -x)$ から (n, y) への path を (i, s_i), $i = 0, \cdots, n$, とする（$s_0 = -x$, $s_n = y$）．$-x < 0$, $y > 0$ なので必ず x 軸を横切る．はじめて x 軸を横切る時を K（$s_K = 0$）とする：$K = \inf\{k \in \mathbf{N} \mid s_k = 0\}$. 時刻 K までの path を x 軸に関して対称移動すれば $(0, x)$ から (n, y) への path を得る．この変換は明らかに単射かつ全射であるから，主張を得る． □

Reflection principle の応用の有名な例題として，選挙のとき常にリードを保ったまま勝つ確率がある．

定理 92 (The Ballot Theorem[35]) *二人の候補者 A, B について一票ずつ開票して最後にそれぞれ a, b 票ずつ得票して A が勝つ $(a > b)$ とする．A が常にリードしながら勝つ票の出方 $(A, B$ の並び方$)$ の総数と経過は問わずにとにかく $a > b$ で勝つ票の出方の総数の比は $\dfrac{a-b}{a+b}$ である[36]．*

証明. $x = a - b$, $n = a + b$ とおく．票の出方の総数は $N_{n,x}$ である．

常にリードしながら勝つ票の出方は $(1, 1)$ から (n, x) への path のうち x 軸に触らないものの総数と等しいことは明らか．Reflection principle によって，これは次に等しい．

$$N_{n-1,x-1} - N_{n-1,x+1} = \binom{n-1}{a-1} - \binom{n-1}{a} = \frac{a-b}{a+b}N_{n,x}.$$

これより，主張を得る． □

‘Arcsine law’ とは simple random walk が期間内で出発点に**最後に戻ってくる時刻**の分布の漸近形に関する主張である．これを証明するために次の 2 つの補題を用意する[37]．引き続き $W_0 = 0$ とする．

補題 93 *(i)* $\mathrm{P}[\, W_1 \neq 0, \cdots, W_{2n} \neq 0 \,] = \mathrm{P}[\, W_{2n} = 0 \,]$ $(= u_{2n})$.

(ii) $L_{2n} = \sup\{m \in \{0, 1, \cdots, 2n\} \mid W_m = 0\}$ とおく（最後の帰還時刻）とき，$\mathrm{P}[\, L_{2n} = 2k \,] = u_{2k}u_{2n-2k}$.

証明. (i) 主張の左辺は

$$2\mathrm{P}[\, W_1 > 0, \cdots, W_{2n} > 0 \,] = 2\sum_{k=1}^{\infty} \mathrm{P}[\, W_1 > 0, \cdots, W_{2n-1} > 0, W_{2n} = 2k \,]$$

36 [15, §3.3] によれば 1878 年に初めて証明されたとのこと．
36 Ballot とは (無記名) 投票 (用紙) のこと．
37 これ以後，hitting time で条件付けて，独立性を使って確率を分解した後，hitting time を時刻の原点に取り直す，という式変形を証明に用いるが，この変形ができることを (強)Markov 性という．

に等しい．$p_{n,x} = \mathrm{P}[\,W_n = x\,] = N_{n,x}2^{-n}$ とおくと，$(0,0)$ から $(2n,2k)$ まで行く path で正時刻に 0 に戻らないものは 定理 92 より $N_{2n-1,2k-1} - N_{2n-1,2k+1}$ に等しいから，

$$2\mathrm{P}[\,W_1 > 0,\cdots,W_{2n-1} > 0, W_{2n} = 2k\,] = p_{2n-1,2k-1} - p_{2n-1,2k+1}\,.$$

よって主張の左辺は，対称性と独立性から

$$p_{2n-1,1} = \frac{1}{2}(p_{2n-1,1} + p_{2n-1,-1}) = \mathrm{P}[\,W_{2n} = 0\,].$$

(ii) $\mathrm{P}[\,L_{2n} = 2k\,] = \mathrm{P}[\,W_{2k} = 0, W_{2k+1} \neq 0,\cdots,W_{2n} \neq 0\,]$ だから，独立性と上の主張から明らか. □

出発点への最初の帰還時刻の分布は容易に分かる．

系 94 (First return) $f_{2n} = \mathrm{P}[\,W_1 \neq 0,\cdots,W_{2n-1} \neq 0, W_{2n} = 0\,]$, $n \in \mathbf{N}$, とおく（出発点への最初の帰還時刻）.）

このとき， $f_{2n} = u_{2n-2} - u_{2n} = \dfrac{1}{2n-1}u_{2n} = \dfrac{1}{2n}u_{2n-2}$.

証明. 奇数時刻に 0 に戻ることがないことに注意すると，補題 93 より，

$$\begin{aligned}&f_{2n} = \mathrm{P}[\,W_2 \neq 0, W_4 \neq 0,\cdots,W_{2n-2} \neq 0\,] - \mathrm{P}[\,W_2 \neq 0, W_4 \neq 0,\cdots,W_{2n-2} \neq 0, W_{2n} \neq 0\,]\\ &= \mathrm{P}[\,W_{2n-2}\,] - \mathrm{P}[\,W_{2n}\,] = u_{2n-2} - u_{2n}\,.\end{aligned}$$

これに具体形 (60) を適用すれば明らか. □

Hitting time $\tau_x = \min\{n > 0 \mid W_n = x\}$ を $x \in \mathbf{Z}$ に対して定義すると，$\{\tau_0 = n\} = \{W_1 \neq 0,\cdots,W_{n-1} \neq 0, W_n = 0\}$ なので， 補題 93 の最初の主張は $\mathrm{P}[\,\tau_0 > 2n\,] = u_{2n}$ と書ける．

また系 94 の f_{2n} は

$$f_{2n} = \mathrm{P}[\,\tau_0 = 2n\,] \tag{62}$$

となる．

定理 95 (Arcsine law for the last visit to 0**)** $0 < a < b < 1$ ならば

$$\lim_{n\to\infty} \mathrm{P}[\,a \le \frac{L_{2n}}{2n} \le b\,] = \frac{1}{\pi}\int_a^b \frac{dx}{\sqrt{x(1-x)}} \quad \left(= \frac{2}{\pi}\int_{\sqrt{a}}^{\sqrt{b}} \frac{dy}{\sqrt{1-y^2}} = \frac{2}{\pi}\arcsin y \Big|_{\sqrt{a}}^{\sqrt{b}} \right).$$

即ち， $L_{2n}/(2n)$ は分布関数 $\mu([0,x]) = \dfrac{2}{\pi}\arcsin\sqrt{x}$ なる $[0,1]$ 上の分布に法則収束する．

Ballot 定理では各整数時刻で常に値が正になる確率を計算するが，折れ線が x 軸の正の側にある時間幅の総長も arcsine law が成り立つ．既に定義した $\tau_x = \min\{n > 0 \mid W_n = x\}$ に関して次が成り立つ．

補題 96 *(i)* $\mathrm{P}[\,\tau_1 = 2n-1\,] = \mathrm{P}[\,\tau_{-1} = 2n-1\,] = f_{2n}\,.$

(ii) $\mathrm{P}[\,\tau_1 > 2n\,] = \mathrm{P}[\,\tau_{-1} > 2n\,] = u_{2n}\,.$

(iii) $u_{2n} = \displaystyle\sum_{\ell=1}^{n} f_{2\ell}u_{2n-2\ell}\,.$

証明. (i) 独立性と $y = 1$ で折り返す reflection principle と (61) から

$$\begin{aligned}&\mathrm{P}[\,\tau_1 = 2n-1\,] = \mathrm{P}[\,W_1 \le 0,\cdots,W_{2n-3} \le 0, W_{2n-2} = 0\,]\mathrm{P}[\,W_1 = 1\,]\\ &= \frac{1}{2}(\mathrm{P}[\,W_{2n-2} = 0\,] - \mathrm{P}[\,W_{2n-2} = 0,\ \exists k \in \{1,2,\cdots,2n-3\};\ W_k = 1\,])\\ &= \frac{1}{2}(\mathrm{P}[\,W_{2n-2} = 0\,] - \mathrm{P}[\,W_{2n-2} = 2\,]) = 2^{-2n+1}(\binom{2n-2}{n-1} - \binom{2n-2}{n}) = \frac{1}{2n}u_{2n-2}\\ &= f_{2n}\,.\end{aligned}$$

最後の変形で系 94 を用いた．

(ii) 系 94 より $f_{2n} = u_{2n-2} - u_{2n}$. よって

$$u_{2n} = u_0 - \sum_{k=1}^{n} f_{2k} = 1 - \sum_{k=1}^{n} \mathrm{P}[\,\tau_1 = 2k-1\,] = \mathrm{P}[\,\tau_1 > 2n\,].$$

(iii) 独立性と (62) から,

$$u_{2n} = \mathrm{P}[\,W_{2n} = 0\,] = \sum_{\ell=1}^{n} \mathrm{P}[\,\tau_0 = 2\ell, W_{2n} = 0\,] = \sum_{\ell=1}^{n} \mathrm{P}[\,\tau_0 = 2\ell\,]\mathrm{P}[\,W_{2n-2\ell} = 0\,]$$
$$= \sum_{\ell=1}^{n} f_{2\ell} u_{2n-2\ell}\,.$$

□

定理 97 (滞在時間の arcsine law) η_{2n} を, $[0,2n]$ の時間帯に正の側にいる単位線分の本数（$W_t > 0$ となる時間幅）とする．即ち,

$$\eta_{2n} = \sharp\{k \in [0,2n) \mid W_k \vee W_{k+1} > 0\}.$$

このとき $\mathrm{P}[\,\eta_{2n} = 2k\,] = u_{2k}u_{2n-2k}$, $k = 0,1,\cdots,n$.

（従って, $\eta_{2n}/(2n)$ も $n \to \infty$ で $L_{2n}/(2n)$ と同じ分布に法則収束する．つまり,「正または負一方に偏る」のが普通.）

証明. 補題 96 から

$$\mathrm{P}[\,\eta_{2n} = 0\,] = \mathrm{P}[\,\tau_1 > 2n\,] = u_{2n} = u_0 u_{2n}\,.$$

よって $k=0$ のとき主張は成り立つ．対称性から $k=n$ についても主張は成り立つ．特に $n=1$ のときは全ての k について主張が成り立つ.

$n-1$ まで主張が成り立つとする．独立性と帰納法の仮定と (62) と補題 96 から

$$\mathrm{P}[\,\eta_{2n} = 2k\,] = \sum_{\ell=1}^{n} \sum_{e\in\{-1,1\}} \mathrm{P}[\,W_1 = e, \tau_0 = 2\ell, \eta_{2n} = 2k\,]$$
$$= \sum_{\ell=1}^{n} (\mathrm{P}[\,W_1 = 1, \tau_0 = 2\ell\,]\mathrm{P}[\,\eta_{2n-2\ell} = 2k-2\ell\,] + \mathrm{P}[\,W_1 = -1, \tau_0 = 2\ell\,]\mathrm{P}[\,\eta_{2n-2\ell} = 2k\,])$$
$$= \sum_{\ell=1}^{n} (\frac{1}{2} f_{2\ell} u_{2k-2\ell} u_{2n-2k} + \frac{1}{2} f_{2\ell} u_{2k} u_{2n-2k-2\ell})$$
$$= \frac{1}{2} u_{2n-2k} \sum_{\ell=1}^{k} f_{2\ell} u_{2k-2\ell} + \frac{1}{2} u_{2k} \sum_{\ell=1}^{n-k} f_{2\ell} u_{2n-2k-2\ell}$$
$$= u_{2k} u_{2n-2k}$$

となって, n のときも成り立つ.

□

C　離散時刻マルチンゲール.

この節は [24, 10 章] と [27, §4] に基づく.

C.1 マルチンゲール．

C.1.1 確率変数が生成する σ 加法族．

§B.2 でも触れたが，ここにまとめておく．

> 可測空間 $(\Omega, \mathcal{F})$ 上の実可測関数（実確率変数） $X_i:\ \Omega \to \mathbf{R}$, $i = 1, 2, \cdots, n$, に対して，
> $$\sigma[X_1, \cdots, X_n] = \{(X_1, \cdots, X_n)^{-1}(A) \mid A \in \mathcal{B}_n\} \subset \mathcal{F}$$
> は X_i, $i = 1, 2, \cdots, n$, を可測にする最小の σ 加法族である．これを $X_1, \cdots, X_n$ が**生成する σ 加法族**という．

命題 98 $\sigma[X_1, \cdots, X_n]$ が σ 加法族であり，$X_i^{-1}(\mathcal{B}_1) \subset \sigma[X_1, \cdots, X_n]$ 及び最小性が成り立つ．

証明．略証．

$$\sigma[X_1, \cdots, X_n] = \{\{\omega \in \Omega \mid (X_1(\omega), \cdots, X_n) \in A\} \mid A \in \mathcal{B}_n\}$$

が σ 加法族であることは容易．

$$\sigma[X_1, \cdots, X_n] \supset \sigma[\bigcup_{i=1}^{n} \{X_i^{-1}(A) \mid A \in \mathcal{B}_1\}]$$

は $X_i^{-1}(A) = (X_1, \cdots, X_n)^{-1}(\mathbf{R} \times \cdots A \cdots \times \mathbf{R})$ から．

最小性 $(\sigma[X_1, \cdots, X_n] \subset \mathcal{C})$ は，$\tilde{\mathcal{B}} = \{A \in \mathcal{B}_n \mid (X_1, \cdots, X_n)^{-1}(A) \in \mathcal{C}\}$ とおくと $\tilde{\mathcal{B}}$ は矩形集合を全て含む σ 加法族になるので $\mathcal{B}_n$ を含むから． □

このイメージに合うのは情報が増大していく場合，即ち，可測空間 $(\Omega, \mathcal{F})$ の σ 加法族の列 $\mathcal{F}_n$, $n \in \mathbf{Z}_+$, が filtration であるとは，$\mathcal{F}_0 \subset \mathcal{F}_1 \subset \cdots (\subset \mathcal{F})$ が成り立つことをいう．$\mathcal{F}_n = \sigma[X_1, \cdots, X_n]$, $n \in \mathbf{N}$, および $\mathcal{F}_0 = \{\emptyset, \Omega\}$ とすれば $\mathcal{F}_n$, $n \in \mathbf{Z}_+$, は filtration である．以下では filtration としてこの場合のみ考える．

Filtration $\{\mathcal{F}_n\}$ が与えられたとき，確率変数列 $\{X_n\}$ が $\{\mathcal{F}_n\}$ に adapted とは，各 n に対して X_n が $\mathcal{F}_n$–可測であることをいう．もちろん X_n は $\sigma[X_1, \cdots, X_n]$–可測である．

Filtration と adaptedness は連続確率過程にも拡張できることは想像できるだろうが，この講義ではそれは触れない．

C.1.2 Stopping time．

§B.2 でも定義しているが，独立に読めるようにここでも重複して定義しておく．

確率空間 $(\Omega, \mathcal{F}, \mathrm{P})$ 上の確率変数列 X_n, $n \in \mathbf{N}$, に対して確率変数 $\tau:\ \Omega \to \mathbf{N} \cup \{\infty\}$ が stopping time であるとは，各 $n \in \mathbf{N}$ に対して $\{\sigma \le n\}$ が $X_1, \cdots, X_n$ が生成する σ 加法族 $\mathcal{F}_n = \sigma[X_1, \cdots, X_n]$ の要素になっていることをいう．

ここで $\sigma[X_1, \cdots, X_n]$ は $X_1, \cdots, X_n$ が生成する σ 加法族，即ち $X_i^{-1}(\mathcal{B}_1)$, $i = 1, \cdots, n$, たちを含む最小の σ 加法族．

$\mathcal{F}_1 \subset \mathcal{F}_2 \subset \mathcal{F}_3 \subset \cdots$ だから，τ が stopping time になることは

$$\{\tau = n\} \in \mathcal{F}_n, \quad n \in \mathbf{N},$$

と同値である．即ち，

> 気持ち： τ が stopping time であるとは，時刻（無限大を許す）に値をとる確率変数であって，τ がある時刻になることがその時刻までの path によって決まることをいう．

直感的には τ が stopping time とは，ゲームからの離脱時刻ということ．時刻 n 終了時に離脱するかどうかは n までのゲームの情報 $\mathcal{F}_n$ で決まるという意味で $\{\tau = n\} \in \mathcal{F}_n$ であるべきである．

典型例は，adapted process $\{X_n\}$ とボレル集合 $B \in \mathcal{B} = \sigma(\{[-\infty, a] \mid a \in \mathbf{R}\})$ に対する hitting time $\tau = \inf\{n \geq 0 \mid X_n \in B\}$．但し， $\inf \emptyset = \infty$．具体的な stopping time としては殆ど hitting time しかない．他方，全ての stopping time についてあることが成り立つと，$\{X_n\}$ に関してある性質が成り立つ，といった非構成的な使い方も多い．

C.1.3 マルチンゲール．

マルチンゲールという概念は本来は条件付き期待値という概念を定義して，それに基づいて定義するのが自然であり,「公平な賭」という名の由来が分かる．

このためには Radon-Nikodym の定理などの，測度論からの準備が必要になる．確率過程論，特に確率積分や確率微分方程式の研究においてはたいへん重要な概念だが，この講義ではそこまで踏み込めない．他の教科書を読む場合は，そのことに注意する必要がある．

また，確率微分方程式では連続時刻マルチンゲールを扱うが，ここでは離散時刻 (stochastic chain) の場合のみを扱う．

以下では，離散時間確率過程，特に random walk について，準備の少なくてすむ定義ですませる．

確率空間 $(\Omega, \mathcal{F}, \mathrm{P})$ に filtration，即ち，増大する σ 加法族の列 $\mathcal{F}_n,\ n \in \mathbf{Z}_+$, が与えられているとする．各 $n \in \mathbf{Z}_+$ に対して $\mathcal{F}_n$–可測確率変数 W_n であるような ($\{\mathcal{F}_n\}$–adapted な) 確率変数列 $\{W_n\}$ がマルチンゲールであるとは，$\mathrm{E}[\,|W_n|\,] < \infty,\ n \in \mathbf{Z}_+$, (可積分) であって，

$$\mathrm{E}[\,W_{n+1}; B\,] = \mathrm{E}[\,W_n; B\,],\ B \in \mathcal{F}_n\,,\quad n \in \mathbf{Z}_+, \tag{63}$$

が成り立つことをいう．

このとき（積分の定義の一般論を思い出すと）， $W_n F$ が可積分になるような任意の $\mathcal{F}_n$–可測可積分関数 $F:\ \Omega \to \mathbf{R}$ に対して,（$F\,W_{n+1}$ も可積分で）

$$\mathrm{E}[\,F\,W_{n+1}\,] = \mathrm{E}[\,F\,W_n\,] \tag{64}$$

が成り立つ．

(63) から直ちに次を得る．

命題 99 $X_n,\ n \in \mathbf{Z}_+$, が独立確率変数列，$\mathcal{F}_n = \sigma[X_0, \cdots, X_n]$, $W_n = \sum_{k=0}^{n} X_k,\ n \in \mathbf{Z}_+$, のとき，$\{W_n\}$ がマルチンゲールであることは $\mathrm{E}[\,X_n\,] = 0,\ n \in \mathbf{N}$, と同値である *(*$X_0$ は別格，積分定数みたいなもの*)*.

証明. $\mathrm{E}[\,X_n; B\,] = 0,\ B \in \mathcal{F}_{n-1},\ n \in \mathbf{N}$, と同値であることは明らか．$\{X_n\}$ が独立なので，$B \in \mathcal{F}_{n-1}$ ならば $\mathrm{E}[\,X_n; B\,] = \mathrm{E}[\,X_n\,]\mathrm{P}[\,B\,]$ だから，これは $\mathrm{E}[\,X_n\,] = 0$ と同値. □

同様に，確率変数列 $\{W_n \mid n \in \mathbf{Z}_+\}$ が優マルチンゲール (supermartingale) であるとは，adapted, 可積分，かつ,

$$\mathrm{E}[\,W_{n+1}; B\,] \leq \mathrm{E}[\,W_n; B\,],\ B \in \mathcal{F}_n\,,\quad n \in \mathbf{Z}_+,$$

を満たすことを言い，確率変数列 $\{W_n \mid n \in \mathbf{Z}_+\}$ が劣マルチンゲール (submartingale) であるとは，adapted，可積分，かつ,

$$\mathrm{E}[\,W_{n+1}; B\,] \geq \mathrm{E}[\,W_n; B\,],\ B \in \mathcal{F}_n\,,\quad n \in \mathbf{Z}_+,$$

特に劣マルチンゲールならば， $\mathrm{E}[\, W_{n+1}\,] \ge \mathrm{E}[\, W_n\,]$，即ち，期待値は増える．ほかについても同様．
例：

(i) 定数を加えても [優/劣/] マルチンゲール性に変化はない．
(ii) 命題 99 の記号の下で， $\{W_n\}$ がマルチンゲールならば $\{{W_n}^2\}$ は劣マルチンゲール．
(iii) W_n が優マルチンゲールならば $W_n \wedge a$ も優マルチンゲール．

C.2 Doob の optional stopping．

この節は（「勝ち逃げ」の小節を除いて）[24, §10.9–§10.11] からの引用である．

C.2.1 マルチンゲール変換．

$\{W_n\}$ が賭やゲームにおける累積配当率を表すとしよう．つまり， $X_n = W_n - W_{n-1}$ が n 回目のゲーム終了時の単位掛け金あたりの配当である．

確率変数列 C_n, $n = 1, 2, 3, \cdots$, が predictable (previsible) とは C_n が $\mathcal{F}_{n-1}$–可測であること，即ち，$\{C_n \le x\} \in \mathcal{F}_{n-1}$, $x \in \mathbf{R}$． C_0 は定義されない．

例えば $\{W_n\}$ が adapted のとき $Y_n = W_{n-1}$ とおけば prdictable になるから，表面的には定義だけの問題とも言える．しかし，気持ちには意味がある．$X_n = W_n - W_{n-1}$ を n 回目のゲームの賭の配当率，C_n を n 回目のゲームへの掛け金とすると，掛け金は $n-1$ 回目終了時までのゲームに関する情報 $\mathcal{F}_{n-1}$ で決めるしかないから，C_n は $\mathcal{F}_{n-1}$–可測，即ち predictable である．

このとき，時刻 n （ゲーム終了時点）の全財産を Y_n と書けば，$Y_n = Y_{n-1} + C_n(W_n - W_{n-1})$ となり，adapted な確率変数列になる．しかも， $\{W_n\}$ が martingale で，Y_n が可積分（になるような C_n ）ならば $\{Y_n\}$ も martingale になるし，$\{W_n\}$ が submartingale で $C_n \ge 0$, a.e., $n \in \mathbf{N}$, で可積分性が保証されるならば $\{Y_n\}$ も submartingale である．$Y_0 = 0$ のとき，この Y_n を

$Y_n = \displaystyle\sum_{k=1}^{n} C_k(W_k - W_{k-1}) = (C \bullet W)_n$ とおいて，W の C による martingale 変換と呼ぶ．
Martingale 変換は確率積分 $\displaystyle\int C\, dW$ の離散類推である．

Martingale の martingale 変換が martingale になる十分条件を書いておく．

定理 100 *(i)* C を非負有界（$0 \le C_n \le K$, *a.e.*, $n \in \mathbf{N}$）な *predictable process* とする．W が優 *[*劣*]* マルチンゲールならば $C \bullet W$ も時刻 0 で 0 の優 *[*劣*]* マルチンゲール．
(ii) C を有界（$|C_n| \le K$, *a.e.*, $n \in \mathbf{N}$）な *predictable process* とする．W がマルチンゲールならば $C \bullet W$ も時刻 0 で 0 のマルチンゲール．
(iii) 上で有界の仮定を $C_n \in \mathcal{L}^2$, $n \in \mathbf{N}$, と $W_n \in \mathcal{L}^2$, $n \in \mathbf{Z}_+$, に置き換えてもよい．

証明. 最初の 2 つは

$$\mathrm{E}[\, |(C \bullet W)_n|\,] = \mathrm{E}[\, |C_n W_n|\,] + \sum_{k=1}^{n-1} \mathrm{E}[\, (|C_k| + |C_{k+1}|)|W_k|\,]$$

から，

$$\mathrm{E}[\, |(C \bullet W)_n|\,] \le 2K \sum_{k=1}^{n} \mathrm{E}[\, |W_k|\,]$$

マルチンゲールの仮定に可積分性 $\mathrm{E}[\, |W_k|\,] < \infty$ があるので，$(C \bullet W)_n$ も可積分．既に adaptedness とマルチンゲール性は言及したとおりだから，主張を得る．

最後の主張は Schwarz の不等式から

$$\mathrm{E}[\,|C_k W_k|\,] \le \sqrt{\mathrm{E}[\,|C_k|^2\,]\mathrm{E}[\,|W_k|^2\,]} < \infty$$

となって，可積分性を得る．

□

可積分性を保証する条件に神経質になるのは，勝ち逃げという反例を除く必要があるからである．

C.2.2 気にしていること ― 公平な賭と勝ち逃げ．

多数回繰り返す（各回は独立）公平な賭で必ず「勝つ」方法がある．勝ち逃げである．つまり負けが勝ちより多い間はゲームを続け，ゲーム前に比べて資産が増えたらゲームから立ち去るのである．（初回に勝ったらその賭博場からはすぐに足を洗う．）

資産を独立確率変数 X_n の和 $W_n = \sum_{k=0}^{n} X_n$ で表せば， $\tau = \inf\{n \ge 0 \mid W_n > 0\}$ とおけば，明らかに τ は $\sigma[\{X_n\}]$–stopping time である．掛け金に最小単位 (例えば 1 円) があれば，つまり例えば $X_n \in \mathbf{Z}$ ならば，$\mathrm{E}[\,W_\tau;\ \tau < \infty\,] \ge 1 > 0$ である．もし $\tau < \infty$, a.e., ならば勝ち逃げは成り立つ !?

例えば 1 次元 simple random walk は $\tau < \infty$ が成り立つ．即ち，どの点も確率 1 で有限の時刻で hit する．この性質を 1 次元 simple random walk は recurrent である，という．ちなみに

この講義でしなかった重要な話：d 次元 simple random walk は $d \le 2$ のとき recurrent，$d > 2$ のとき transient である．

この話は統計力学との関連などで重要だが，割愛する．

話を元に戻す．勝ち逃げが成り立つということは現実的にも賭をつまらなくするが，理論的にも扱いにくい[38]．例えば 1 次元 simple random walk(一般に公平な賭) では，$\mathrm{E}[\,W_n\,] = 0$, $n \in \mathbf{Z}_+$, だから，$\mathrm{E}[\,W_n\,] \ne \mathrm{E}[\,W_\tau\,]$ を意味する．$\mathrm{E}[\,W_\tau\,] = \mathrm{E}[\,W_0\,]$ は，たとえ $\mathrm{P}[\,\tau < \infty\,] = 1$ であっても無条件には保証されないというわけだ．

「数学公式」としては，もし

$$\mathrm{E}[\,W_0\,] = \mathrm{E}[\,W_\tau\,] \tag{65}$$

が保証されていれば，τ が計算しづらい stopping time であってもそれを計算できてしまうので嬉しい．そこで，(65) が成り立つための十分条件を考える．(無条件に成り立たないことは「勝ち逃げ」の例で分かっている．言い換えれば,「勝ち逃げ」を許さない「いじめではない」ルールを見つける，というのが §C.2.3 の定理の意図である．）そのような十分条件を与える一つの重要な定理が任意停止定理である．

C.2.3 Doob の optional stopping．

マルチンゲールの非常に重要な使い方として次のような，stopping time との組み合わせによる新しいマルチンゲールを導く方法がある．

補題 101 *W を (super)martingale, τ を stopping time とすると $W^\tau = \{W_{\tau\wedge n} \mid n \in \mathbf{N}\}$ も (super)martingale．特に， $\mathrm{E}[\,W_{\tau\wedge n}\,] \le (=)\,\mathrm{E}[\,W_0\,]$, $n \in \mathbf{N}$.*

証明. $C_n = \chi_{n\le\tau} = \begin{cases} 1, & n \le \tau, \\ 0, & \text{otherwise}, \end{cases}$ ，とおく．即ち，時刻 τ まで 1 単位ずつ賭けてそのあとゲームから離脱することを考える．$(C \bullet W)_n = W_{\tau\wedge n} - W_0 = {W^\tau}_n - W_0$ となるので，定理 100 から補題を得る．

□

[38] もっとも，τ が有界でなければ，勝つまでにとてつもない時間がかかる可能性があって，現実的ではない．

この補題の重要な点は martingale ということに implicit な条件を除いて可積分性関係の特別な条件を必要としないという点である．

しかし，おもしろいのは添字が純粋に stopping time τ だけで書ける場合，即ち，$\mathrm{E}[\,W_\tau\,] = \mathrm{E}[\,W_0\,]$ の十分条件である．

次の定理はやさしい場合として，τ か W_n か増分 X_n が有界の場合を扱う．

定理 102 (任意停止定理 (有界な場合)) W を *(super)martingale,* τ を *stopping time* とすると次のいずれかが成り立てば W_τ は可積分でかつ $\mathrm{E}[\,W_\tau\,] \le (=) \mathrm{E}[\,W_0\,]$：

(i) τ が有界 $(\tau \le N,\ a.s.)$.
(ii) $\tau < \infty,\ a.s.,$ かつ W が有界 $(|W_n| \le K,\ a.s.,\ n \in \mathbf{N})$.
(iii) $\mathrm{E}[\,\tau\,] < \infty$，かつ $X_n = W_n - W_{n-1}$ に対して $|X_n| \le K,\ a.s.,\ n \in \mathbf{Z}_+$.
(iv) Supermartingale の場合で $W_n \ge 0,\ a.s.,\ n \in \mathbf{Z}_+,$ かつ $\tau < \infty,\ a.s..$

証明. 優マルチンゲールの場合を証明すれば，W を $-W$ に置き換えることで，マルチンゲールの場合も成り立つ．

補題 101 より，$W_{\tau\wedge n}$ も可積分で

$$\mathrm{E}[\,W_{\tau\wedge n}\,] \le \mathrm{E}[\,W_0\,]. \tag{66}$$

よって，

(i) (66) で $n = N$ とおけばよい．
(ii) (66) で $n \to \infty$ として，$\tau < \infty$ と $|W_n| \le K$ から有界収束定理を使えばよい．
(iii) 過程から τ は可積分関数であって，

$$|W_{\tau\wedge n} - W_0| \le \sum_{k=1}^{\tau\wedge n} |X_k| \le K\tau$$

だから優収束定理から (66) で $n \to \infty$ とすればよい．
(iv) $\tau < \infty$, a.s., と Fatou の補題と (66) から

$$\mathrm{E}[\,W_\tau\,] = \mathrm{E}[\,\varliminf_{n\to\infty} W_{\tau\wedge n}\,] \le \mathrm{E}[\,W_0\,].$$

□

系 103 W を $|W_n - W_{n-1}| \le K,\ a.s.,\ n \in \mathbf{N},$ を満たすマルチンゲール，C を *predictable* かつ有界 $(|C_n| \le K,\ n \in \mathbf{N})$，$\tau$ を可積分 $(\mathrm{E}[\,\tau\,] < \infty)$ な *stopping time* とする．

このとき $\mathrm{E}[\,(C \bullet W)_\tau\,] = 0.$

証明. 定理 100 より $C \bullet W$ は時刻 0 で 0 なるマルチンゲールで，仮定から，定理 102 の第 3 条件を満たす．よって $\mathrm{E}[\,(C \bullet W)_\tau\,] = 0.$

□

系 103 等の optional stopping theorem に出てくる $\mathrm{E}[\,\tau\,] < \infty$ が成り立つ十分条件を挙げる．

定理 104 τ が

$$\exists N \in \mathbf{N},\ \exists \epsilon > 0;\ (\forall n \in \mathbf{N})(\forall B \in \mathcal{F}_n)\ \frac{\mathrm{P}[\,(\tau \le n + N) \cap B\,]}{\mathrm{P}[\,B\,]} > \epsilon,\ a.s.,$$

を満たす *stopping time* ならば $\mathrm{E}[\,\tau\,] < \infty.$

証明. $\{\tau > n\} = \{\tau \le n\}^c \in \mathcal{F}_n$ だから，

$$\begin{aligned} 1 - \mathrm{P}[\,\tau = 0\,] &= \sum_{k=0}^{\infty} \mathrm{P}[\,kN < \tau \le (k+1)N\,] + \mathrm{P}[\,\tau = \infty\,] \\ &> \epsilon \sum_{k=0}^{\infty} \mathrm{P}[\,\tau > kN\,] + \mathrm{P}[\,\tau = \infty\,] \\ &\ge \epsilon \sum_{n=0}^{\infty} \mathrm{P}[\,\tau = n\,] \frac{n-1}{N} + \infty \mathrm{P}[\,\tau = \infty\,]. \end{aligned}$$

これよりまず $\tau < \infty$, a.e., となり，かつ，

$$1 > \frac{\epsilon}{N}(\mathrm{E}[\,\tau\,] - 1)$$

だから， $\mathrm{E}[\,\tau\,] < \infty$.

□

定理 102 では $\mathrm{E}[\,W_\tau\,] \le (=) \mathrm{E}[\,W_0\,]$ を示しただけで，新しいマルチンゲールを作ってはいない．実際は，適当な条件の下で $\tau \le \sigma$ ならば $\mathrm{E}[\,W_\tau; B\,] \le (=) \mathrm{E}[\,W_\sigma; B\,]$，$B \in \mathcal{F}_\tau$, 即ち，stopping time の列に対してマルチンゲールになること，も言える[39]．この種の定理は確率過程における確率積分（マルチンゲール変換の連続極限）で重要な役割を果たす.

C.2.4 文豪ザル．

あるサルがアルファベット大文字をタイプし続ける．毎回 26 字全てを等確率で打つとするとき，平均何回タイプを打てば ABRACADABRA という 11 文字の文字列が出てくるか？

最初の 11 文字がこの文字列になる確率は 26^{-11} だが，だからといって初めてこの文字列が出てくるまで平均 26^{11} 回タイプする，と答えるのは，悪い予想ではないが，正解ではない！

原理的には数え上げで計算できるはずだが，実行すれば大変なことになる．ここでは martingale を用いた解法を紹介しよう．

タイプをでたらめに永久に打ち続ける確率空間を考える．

$$\begin{aligned} &r = 26,\ \Sigma = \{A, B, \cdots, Z\}, \\ &\Omega = \{\omega = (\omega_1, \omega_2, \cdots) \mid \omega_i \in \Sigma\ (i \in \mathbf{N})\}, \\ &\mathcal{F}_n = \sigma[\{\{\omega \in \Omega \mid \omega_i = v_i,\ i = 1, 2, \cdots, n\} \mid (v_1, \cdots, v_n) \in \Sigma^n\}],\ \mathcal{F} = \sigma[\bigcup_{n \in \mathbf{N}} \mathcal{F}_n], \\ &n \in \mathbf{N},\ A \in \mathcal{F}_n \text{ に対して } \mathrm{P}[\,A\,] = r^{-n} \sharp A. \end{aligned}$$

文字列のどれでもいいから n 個分の文字を指定すると，それらの値がその通り生じる確率は 2^{-n} になる，という意味である．

ω_n は n 文字目に猿がタイプする文字とする．$\mathcal{F}_n$ は最初の n 文字 $(\omega_1, \cdots, \omega_n)$ で表せる集合（$(\omega_1, \cdots, \omega_n)$ を決めると決まる集合）の全体であり，そういう意味で時刻 n（n 文字目）までの情報を表す．$\{\mathcal{F}_n\}$ は filtration である．

k 文字の単語 $w_1 w_2 \cdots w_k$ が初めて現れる時刻を T とする：$T = \inf\{n \ge k \mid \omega_{n-k+1} = w_1, \omega_{n-k+2} = w_2, \cdots, \omega_n = w_k\}$. T は確率変数になる（$\{T \le n\}$ は $(\omega_1, \cdots, \omega_n)$ で表せるので $\mathcal{F}_n$ の要素である）．

n 文字目までどう決まっていても，そのあと $N = k$ 文字が $w_1, \cdots, w_k$ になる確率は $\epsilon = r^{-k} > 0$ なので，定理 104 より，$\mathrm{E}[\,T\,] < \infty$，従って特に $T < \infty$, a.e..

$\mathrm{E}[\,T\,]$ を求めたい．次のような方法を採る．

[39] [13] はこれもこめて optional stopping theorem(任意停止定理) と呼び，[24] や [9] はこれを optional sampling theorem(任意抽出定理) と呼んでいるようだ．

各時刻 $n \geq 1$ に対して，時刻 n （の直前）に，猿の打つ文字の予想を立てる賭けに参加するギャンブラーの時刻 m 財産を ${W^{(n)}}_m$ とする．このギャンブラーは最初 1 単位の財産を持っており，時刻 n に文字 w_1，以下時刻 $n+k-1$ に文字 w_k までその都度全財産を賭ける．賭けは敗れれば没収され，従って，以後 0 のまま，勝てば財産は r 倍になる．式で書けば $W_0 = 1$ であって，

$$
{W^{(n)}}_m = \begin{cases} 1, & m < n, \\ r\,{W^{(n)}}_{m-1}\,\chi_{\omega_m = w_{m-n+1}}, & n \leq m < n+k, \\ {W^{(n)}}_{n+k-1}, & m \geq n+k\,. \end{cases}
$$

定義から明らかに， $0 \leq {W^{(n)}}_m(\omega) \leq r^k$ が任意の n, m, ω に対して成り立つ（つまり，有界）．また，定理 104 が $N = k, \epsilon = r^{-k}$ として成り立つ．そして $n \leq m < n+k$ のとき $B \in \mathcal{F}_{m-1}$ に対して各賭の独立性から，

$$
\mathrm{E}[\,{W^{(n)}}_m; B\,] = r\,\mathrm{E}[\,W^{(n)}_{m-1}\,\chi_{\omega_m = w_{m-n+1}}\,\chi_B\,] r\,\mathrm{P}[\,\omega_m = w_{m-n+1}\,]\,\mathrm{E}[\,W^{(n)}_{m-1}; B\,]\mathrm{E}[\,W^{(n)}_{m-1}; B\,]
$$

となるので，$W^{(n)}_m$, $m \in \mathbf{Z}_+$, は martingale である（$m < n$, $m \geq n+k$ では明らか）．よって，定理 102 より，$\mathrm{E}[\,{W^{(n)}}_T\,] = \mathrm{E}[\,{W^{(n)}}_0\,] = 1$ となる．

$$
a_i = \begin{cases} 1, & w \text{ の最初と最後の } i \text{ 文字が等しいとき}, \\ 0, & \text{otherwise}, \end{cases}
$$

とおく． T の定義から

$$
{W^{(n)}}_{T\wedge(n+k-1)} = \begin{cases} 0, & T \geq n+k, \\ r^k, & T = n+k-1, \\ r^{T-n+1} a_{T-n+1}\,, & n \leq T < n+k-1, \\ 1, & T < n. \end{cases}
$$

従って[40]，

$$
1 = P(T = n+k-1) r^k + \sum_{i=1}^{k-1} P(T = n+i-1) a_i r^i + P(T < n)
$$

だから

$$
P(T \geq n) = P(T = n+k-1) r^k + \sum_{i=1}^{k-1} P(T = n+i-1) a_i r^i.
$$

これを用いると

$$
\mathrm{E}[\,T\,] = \sum_{n \geq 1} n P(T = n) = \sum_{n \geq 1} P(T \geq n) = P(T \geq k) r^k + \sum_{i=1}^{k-1} P(T \geq i) a_i r^i.
$$

最低 k 文字タイプしないと単語は出てこないので $n \leq k-1$ ならば $P(T = n) = 0$，従って $n \leq k$ ならば $P(T \geq n) = 1$ となる．よって

$$
\mathrm{E}[\,T\,] = r^k + \sum_{i=1}^{k-1} a_i r^i. \tag{67}
$$

ABRACADABRA については $r = 26$, $k = 11$, $a_i = 1$ となるのは $i = 1, 4$，だから

$$
\mathrm{E}[\,T\,] = 26^{11} + 26^4 + 26\,.
$$

最初の山勘 26^{11} との差はわずかだが， 0 ではない！

[40] [24] はこのように ${W^{(n)}}_T$ を用いる解法を期待しているように私は感じたが，本当は ${W^{(n)}}_{T\wedge(n+k-1)}$ を考えることで同じ式を得ることができて，以下の議論が成り立つ．このやり方だと前節の結果のうち 補題 101 さえあれば十分である．つまり，文豪ザルの問題はそれほど簡単な問題だ，ということ．

でも，それだと optional sampling theorem の紹介が全然できないので，[24] の構成に従って ${W^{(n)}}_T$ でやることにした．

マルチンゲールの紹介としては以上で十分だが，最後に，結果について書いておきたい．(67) を見ると，文字列に（最初と最後で）繰り返しがあるほど実現するまで時間がかかることになる．例えば AB が実現するには $\mathrm{E}[\,T\,] = 26^2$ でいいのに，AA が実現するには $\mathrm{E}[\,T\,] = 26^2 + 26$ と，余分の時間がかかる．この現象の「気持ち」を考えてみる．

ある時刻 t に A が出たとき，次の時刻 $t+1$ に AB が実現しなかったとすると，それは時刻 $t+1$ に B 以外が出たことを意味するが，その中には A が出るケースもある．そのとき時刻 $t+2$ に AB が完成する確率が正になる．他方，ある時刻 t に A が出たとき，次の時刻 $t+1$ に AA が実現しなかったとすると，それは時刻 $t+1$ に A 以外が出たことを意味するので，時刻 $t+2$ に AA が完成する確率は 0 である．よって，文字列に繰り返しがあると，それだけ失敗したときに引き続いて文字列が完成するチャンスが低くなるのである．公式 (67) はそのことを教えてくれる．

C.3　横断数とマルチンゲール収束定理．

定理 102 では，適当な有界性の下で $\mathrm{E}[\,W_\tau\,] = \mathrm{E}[\,W_0\,]$ を示した．確率積分などを扱う連続時間確率過程の場合，有界性より弱い可積分性で話を進められるほうが確率解析が強力になる．特に，L^2 有界マルチンゲール ([24, 12 章], [13, §4.4]) や一様可積分マルチンゲール ([24, 14 章], [13, §4.5]) が重要である．

この講義は以上には立ち入らないが，マルチンゲールがこれらの解析で役に立つのは，確率変数列の極限の存在を示す手段があるからである．それが横断数とマルチンゲール収束定理である ([13, §4.2], [24, 11 章])．

この話のためには， Radon–Nikodym の定理によって条件付き期待値と呼ばれる確率変数の存在を示しておいたほうが次の命題の証明がすっきりするので望ましい．（従って，確率積分などに進む場合は Radon-Nikodym の定理は不可避である.）

命題 105 $\{W_n\}$ がマルチンゲールで f が下に凸か，$\{W_n\}$ が劣マルチンゲールで f が下に凸な増加関数ならば，$\{f(W_n)\}$ は劣マルチンゲール．特に（$f(x) = (x-a)\vee 0$ として），$\{W_n\}$ が劣マルチンゲールならば，$(W_n - a)_+ = (W_n - a)\vee 0,\ n \in \mathbf{Z}_+$, も劣マルチンゲール．

注. 横断数不等式で使う．◇

証明. 略証：命題 13 と同様の証明を行う．

$f(x) \geq a(x-c) + f(c)$ において $x = W_{n+1},\ c = W_n$, として，両辺の B における積分をとる．$a = a(c)$ は c だけで決まる．しかも f の c における左接線にとれるので可測関数．

$$\mathrm{E}[\,f(W_{n+1}); B\,] \geq \mathrm{E}[\,W_{n+1}a(W_n); B\,] - \mathrm{E}[\,W_n a(W_n); B\,] + \mathrm{E}[\,f(W_n); B\,].$$

(64) によって， $\{W_n\}$ がマルチンゲールならば $\mathrm{E}[\,W_{n+1}a(W_n); B\,] = \mathrm{E}[\,W_n a(W_n); B\,]$，$\{W_n\}$ が劣マルチンゲールで f が増加ならば ($a \geq 0$ となって)，(64) に至るのと同様の単関数近似によって，$\mathrm{E}[\,W_{n+1}a(W_n); B\,] \geq \mathrm{E}[\,W_n a(W_n); B\,]$．いずれにしても，

$$\mathrm{E}[\,f(W_{n+1}); B\,] \geq \mathrm{E}[\,f(W_n); B\,]$$

だから， $\{f(W_n)\}$ も劣マルチンゲールである．

□

確率連鎖 $W_n,\ n \in \mathbf{Z}_+$, と $a < b$ に対して，$\{W_n\}$ が $[a,b]$ を上向きに横断する回数を定義する．まず，stopping time の列 $T_1, \cdots$, を次のように定義する．

T_1 は初めて $W_m \le a$ となる m，T_2 は T_1 以後で初めて $W_m \ge b$ となる m，T_3 は T_2 以後で初めて $W_m \le a$ となる m，一般に $k = 1, 2, 3, \cdots$, に対して ($T_0 = -1$ として)，$T_{2k-1} = \inf\{m > T_{2k-2} \mid W_m \le a\}$，$T_{2k} = \inf\{m > T_{2k-1} \mid W_m \ge b\}$.

T_j が stopping time であることは明らかだが，さらに，

$$\{T_{2k-1} < m \le T_{2k}\} = \{T_{2k-1} \le m-1\} \cap \{T_{2k} \le m-1\}^c \in \mathcal{F}_{m-1}$$

なので，$C_m = \begin{cases} 1, & \exists k;\ T_{2k-1} < m \le T_{2k}, \\ 0, & \text{otherwise}, \end{cases}$ とおくと，$C_m \in \mathcal{F}_{m-1}$ 即ち，predictable．$\{C_m\}$ は $[a, b]$ を上向きに横断中のとき 1 そうでないとき 0 である．

$U_N = U_N[a, b] = \sup\{k \in \mathbf{N} \mid T_{2k} \le N\}$ は時刻 N までに完成した上向き横断回数．

命題 106 (Upcrossing inequality) $\{W_m\}$ が劣マルチンゲールならば

$$(b-a)\mathrm{E}[\, U_n \,] \le \mathrm{E}[\, (W_n - a)_+ \,] - \mathrm{E}[\, (W_0 - a)_+ \,].$$

証明. $Y_m = a + (W_m - a)_+$ とおく ($c_+ = c \vee 0$) と，命題 105 より，$\{Y_m\}$ は劣マルチンゲールで，$[a, b]$ を W_m と同じ回数上向きに横断する．しかも

$$(b-a)U_n \le (C \bullet Y)_n \tag{68}$$

である．なぜなら，$[a, b]$ を上向きに横断する毎に右辺は $b - a$ 以上増える，最後の完成しなかった横断は Y_m の定義のしかたから右辺に与える寄与は非負，だから．

$Y_n - Y_0 = (C \bullet Y)_n + ((1-C) \bullet Y)_n$ において，定理 100 より $(1-C) \bullet Y$ も時刻 0 で 0 の劣マルチンゲールだから

$$\mathrm{E}[\, ((1-C) \bullet Y)_n \,] \ge \mathrm{E}[\, ((1-C) \bullet Y)_0 \,] = 0\,. \tag{69}$$

よって

$$\mathrm{E}[\, (C \bullet Y)_n \,] \le \mathrm{E}[\, Y_n - Y_0 \,]. \tag{70}$$

(68) と (70) を合わせると主張を得る．

□

注. (69) は，劣マルチンゲールで賭をやると「どうやっても」勝ってしまう，ことを表す． ◇

定理 107 (マルチンゲール収束定理) W が $\sup_{n \in \mathbf{Z}_+} \mathrm{E}[\, W_{n+} \,] < \infty$ を満たす劣マルチンゲールならば，W_n は概収束してその極限は可積分である．

注. この収束定理の特徴は，次の証明で，仮定の下で W の上方横断数が有限になることが本質である点である．他の収束定理にはこのような視点がなかった． ◇

証明. $(W_n - a)_+ \le W_{n+} + |a|$ なので 命題 106 から

$$\mathrm{E}[\, U_n \,] \le \frac{|a| + \mathrm{E}[\, W_{n+} \,]}{b-a}\,.$$

U_n は単調なので $n \to \infty$ である確率変数 U ($[a, b]$ を上方に横断する全回数) に $+\infty$ を許せば収束する．仮定 $\sup \mathrm{E}[\, W_{n+} \,] < \infty$ と単調収束定理から $\mathrm{E}[\, U \,] < \infty$ だから，$U < \infty$, a.s.. a, b は任意だから

$$\mathrm{P}[\, \bigcup_{a,b \in \mathbf{Q}} \{\, \underline{\lim}_{n\to\infty} W_n < a < b < \overline{\lim}_{n\to\infty} W_n \} \,] = 0\,.$$

即ち，$\varliminf_{n\to\infty} W_n = \varlimsup_{n\to\infty} W_n$, a.s., となるから，$W := \lim_{n\to\infty} W_n$ が殆ど至る所存在する．Fatou の補題から $\mathrm{E}[\, W_+ \,] \le \varliminf_{n\to\infty} \mathrm{E}[\, W_{n+} \,] < \infty$.

$\{W_n\}$ が劣マルチンゲールなので,

$$\mathrm{E}[\, W_{n-} \,] = \mathrm{E}[\, W_{n+} \,] - \mathrm{E}[\, W_n \,] \le \mathrm{E}[\, W_{n+} \,] - \mathrm{E}[\, W_0 \,]$$

だから Fatou の補題から,

$$\mathrm{E}[\, W_- \,] \le \varliminf_{n\to\infty} \mathrm{E}[\, W_{n-} \,] \le \sup_{n\in\mathbf{Z}_+} \mathrm{E}[\, W_{n+} \,] - \mathrm{E}[\, W_0 \,] < \infty.$$

□

系 108 $\{W_n\}$ が非負優マルチンゲールならば $n\to\infty$ で概収束し，極限 W は $\mathrm{E}[\, W \,] \le \mathrm{E}[\, W_0 \,]$ を満たす.

証明. $Y_n = -W_n \le 0$ は劣マルチンゲールで $\mathrm{E}[\, Y_{n+} \,] = 0$ を満たすので 定理 107 より可積分関数 $-W$ に収束する．$\mathrm{E}[\, W_0 \,] \ge \mathrm{E}[\, W_n \,]$ だから，Fatou の補題から不等式も得る．

□

C.4 マルチンゲールでできることの例．

[24, §15] から，例を紹介する．但し，ここでは導出はいっさい行わない．前節の意義を想像することや，教科書に取りかかる動機として使ってもらえることを期待する．

C.4.1 羊たちの沈黙．

[24, §15.3–5]．The Mabinogion （Wales の伝説 Peredur ap Efrawg 物語）の中に白と黒の羊の話があるそうだ．各時刻 $1, 2, 3, \cdots,$ に羊が一頭ランダムに鳴く．すると，鳴いた羊の色と反対の色の羊が一頭鳴いた羊と同色に変わる．人は各時刻 $0, 1, 2, \cdots,$ （の直後）に白い羊を何匹でも消すことができる．目的は，全部黒い羊になったとき（それで最終状態）黒い羊の数の期待値を最大にすることである．

このようにランダムな現象に途中で人が手を加えて何かの期待値を最善にしようとする問題を stochastic control と呼ぶ．あらゆる現実の制御は，必ず制御不能な撹乱があるという意味で stochastic control である．マルチンゲール理論を用いることで羊の話について最善策が決定できる．この問題は証明の計算がかなり長い．結果だけ引用しておく．

結論（最善策）: 各時刻毎に黒い羊の方が多い（か黒い羊が残っていない）ときは何もしない．さもなければ，白い羊を黒い羊の数から一頭引いた数に減らす.

この策の下で黒白 k 頭ずつから始めたときの，最後の黒い羊の数の期待値 E_k は，$\lim_{k\to\infty} E_k - (2k + \frac{1}{4}\pi - \sqrt{\pi k} = 0$ を満たすことが知られている．例えば一万頭ずつから始めたときおおよそ 19824 頭残る．

C.4.2 「離散版」ブラックショールズの公式．

[24, §15.1–2]

Black–Scholes の公式はある基本的なオプション (株の「買い方」についての契約) の価格 (価値) を決める公式である．ブラックショールズの公式は連続時刻の公式であり，その証明にはマルチンゲール表現定理と Cameron–Martin–Girsanov 変換定理が必要な，確率過程論の本格的な応用を要する．

他方，この離散時刻アナロジーには確率論は不要で，組み合わせ論で答を得る．

利子 r が一定の債券 (国債)(bond) と価格が変動する株 (stock) が１種類ずつの経済を考える．N を自然数として，時刻 $0,\cdots,N$ に債券と株の値段が変動するとする．１単位の価格は，時間帯 $(n,n{+}1)$ で，債券が $B_n=(1{+}r)^n B_0$, 株が S_n とする．時刻 0 に全資産を x，資産配分 (portfolio) を (株, 債券) $=(A_0,V_0)$ 単位とする．時刻 1 の直前までに portfolio の組み替え (投資) を行って (A_1,V_1) にする．従って

$$A_0S_0+V_0B_0=A_1S_0+V_1B_0=x.$$

一般に時間帯 $(n,n+1)$ における全資産を X_n, portfolio を (A_n,V_n), 価格を (B_n,V_n) とすると，

$$X_n=A_nS_n+V_nB_n=A_{n+1}S_n+V_{n+1}B_n \tag{71}$$

となる．株価の「利子」を $R_n=\dfrac{S_n-S_{n-1}}{S_{n-1}}$ で定義する[41]．

以上の定義で，割引いた資産 $Y_n=(1+r)^{-n}X_n$ は

$$Y_n-Y_{n-1}=(1+r)^{-n}A_nS_{n-1}(R_n-r)$$

となる．

さらに，R_n の取りうる値は $-1<a<r<b<\infty$ を満たす a,b だけとする．

European option に対する Black–Scholes 公式を導く．European option とは時刻 0 の直後に契約する．契約内容は時刻 N の直後に価格 K (striking price) で株を買う権利である．(権利，というのは実際の株価が安ければ，この契約からは買わないという意味.) １株当たりの時刻 N 直後のこの契約の利益は $(S_N-K)_+$ となる．問題は，時刻 0 でこの契約の１株当たりの値段を決める[42] ことである．

Black–Scholes は以下の hedging strategy を満たす解の有無という問題として定式化し，それに解を与えた．それが Black–Scholes の公式である．与えられたオプションに対する，初期資産 x の下での hedging strategy とは predictable な $\{A_n\}$, $\{V_n\}$ による portfolio management scheme $\{(A_n,V_n)\}$ であって，$X_0=x$，(71)，$X_n\ge 0,\ n=0,\cdots,N$，$X_N=(S_N-K)_+$，を (あらゆる株価変動の組み合わせに対して) 満たすものをいう．つまり，時刻 N において与えられたオプションを portfolio によって (破産せずに) 再現することをいう．

Black–Scholes の公式の根拠は次の定理である．これを述べるために，以下の量を定義する．$R_n=\dfrac{a+b}{2}+\dfrac{b-a}{2}\epsilon_n$, $p=\dfrac{r-a}{b-a}$, $Z_0=0$, $Z_n-Z_{n-1}=\epsilon_n-2p+1$, で ϵ_n, p, $\{Z_n\}$ を定義する．

定理 109 *Hedging strategy* は

$$x=\mathrm{E}[\,(1+r)^{-N}(S_N-K)_+\,]=:x_0$$

が成り立つとき，そのときに限り，存在する．ここで $\mathrm{E}[\ \cdot\ ]$ は，$\{\epsilon_n\}$ を ±1 に値をとる *i.i.d.* で $\mathrm{P}[\,\epsilon_1=1\,]=p$ なるものとする確率空間における期待値．

このとき *hedging strategy* は一つに決まり，$A_n<0$ は起きない (空売り *(short selling)* がない)．

この定理によって，時刻 0 におけるヨーロッパオプションの価格は x_0 と決めるべきである，というのが BS 公式である．

離散版 BS 公式 定理 109 の証明は難しくないようだが，教科書に譲る．

D Brown 運動門前編．

20050112 注．本論からいきなりこれを読んでも不親切．初等講義としては，以下のような手順でやるべきであろう．

---

[41] 債券と違って分母が時刻 $n-1$ の値になっているがこれは [24] だけなのか，普通なのか？

[42] もちろん，理論をやるときは，公平になるように決める，という意味．

(i) 準備

(a) 直積空間上の測度（直積測度とは限らない）の構成と整合条件の話．測度の拡張定理との関係を強調.

(b) 整合条件を満たす有限直積空間上の測度列が与えられたときの無限直積空間上の測度の構成 (Kolmogorov)

(ii) 定義

(a) ブラウン運動の直感的背景

(b) 数学的にとらえたものとしてのブラウン運動の定義（性質全部で定義と呼ぶことも，唯一性を満たす最小限で定義とすることも，間にもいろんな選択があるが，いちばん手頃なもの＝存在定理が最小限のもの＝ [14] に従う.

(c) 構成という言葉の意味（存在定理ということ）．確率空間と確率変数を与える

(d) ４つの主な方法（連続変形定理 [14]，直交関数展開 [4]，不変原理 [12]，くりこみ群 [17]）．ここでは連続変形定理で.

(iii) 連続変形定理経由の構成

(a) 定義から W の分布だけでいいこと．確率 1 で連続関数になることが知られている．そこで最初から $\Omega = C^0$ で作ってみせる.

(b) ガウス積分の復習，特に整合性の確認.

(c) $\Omega = \mathbf{R}^n = \{(w(1), \cdots, w(n))\}$ 上のガウス測度，$\Omega\{w:\ \mathbf{Z}_+ \to \mathbf{R}\}$ 上のガウス測度，$\Omega\{w:\ \mathbf{Q}_+ \to \mathbf{R}\}$ 上のガウス測度.（有限個＝ cylinder sets に対して定義しておいて整合性を確認，可算和の加法族を含む最小の σ 加法族へ拡張）.

(d) 連続変形定理（ガウス測度であることから $2n$ 次のモーメント条件を得て，それが連続性を保証するのに十分なこと）．$\Omega\{w:\ \mathbf{Q}_+ \to \mathbf{R}\}$ の中で確率 1 で一様連続なので，そこに制限すれば $C^0(\mathbf{R}_+ \to \mathbf{R})$ への埋め込み ι が存在して可測．$\mathrm{P} \circ \iota^{-1}$ で C^0 上の測度を得る.

D.1　確率過程の「気持ち」.

この節は [28, §5]，[27, §5.1] から引用した.

人生は偶然の連続，という考え方がある.「人生万事塞翁が馬」という言い回しは，人生いつまでたってもその先が予測不能，という意味である．必然性のないようにみえる突然の出来事の連続で，不運（幸運）にも人生がころころ変わる.

RPG (role playing game) は各場面ごとに選択肢があって，一人の主人公についていろいろな冒険が可能である．各選択場面でさいころを振って選択肢を決めるとすると，ある時刻 t における RPG の主人公の状態（例えば存在位置）を W_t と書いたとき， W_t が各 t で確率変数になっている，といえる.

すごろくというゲームは洋の東西を問わない[43]．人生の中から普遍的な要素としての偶然性を抽出した象徴的なゲームだから人類に普遍的に受け入れられた，と見るのは考え過ぎか.

> 実数 $t \in \mathbf{R}$ で番号づけられた確率変数の族（集まり）$\{W_t\}_{t\in\mathbf{R}}$ を確率過程 (stochastic process) という.

族のパラメータ (添字) を時刻変数ということが多いが，時刻を表さなくてもかまわない．例えば，実験データ W_n, $n = 1, 2, \cdots,$ は大勢の人が同じ実験をやったとき，制御不能な撹乱要因 (測定誤差) のために値が異なる場合，これを確率変数とみてデータ処理を行うことが考えられる．n は誰が行った実験かを区別する．また，画像データ W_x, $x \in [a, b)$, はデータに望まない信号 (ノイズ) が加わって本来の画像がランダムに乱される場合，各点 x 毎の値（例えば白黒画像ならば輝度）W_x が確率変数と考えて分析せきるかもしれない．　このときは x は空間的な位置を表す変数である.（平面画像

[43] 例えばイギリスには snakes and ladders と呼ばれるすごろくがある．蛇の頭の絵まで来ると尾まで逆戻りし，はしごの下端に来ると上まで進める.

を扱う場合は $W_{(x,y)}$ のように座標2変数を parameter とする必要が生じうる．この講義ではそこまでは扱えない.）

時刻変数の定義域は実数の部分集合でもよい． $[0,1]$, $[0,\infty)$ などもよく用いられる．時刻変数が $\mathbf{Z}_+$ などの離散集合の時は，確率連鎖 (stochastic chain) と言うことも多い．この場合は添字を n（には限らないが）にして，W_n, $n=0,1,2,\cdots$, と記すことが多い．独立確率変数列は確率連鎖の例である．

もちろん確率変数というからには，確率空間 $(\Omega,\mathcal{F},\mathrm{P})$ が与えられていて，各 $t\in\mathbf{R}$ 毎に $W_t:\Omega\to\mathbf{R}$ と考えている．なお，今まで確率変数は実数値関数としてきたが，実数値でない場合にも拡張できる．W_t がある可測空間 Σ を値域とするとき，

$$W_t:\ \Omega\to\Sigma,\ \ t\in\mathbf{R},$$

Σ をこの確率過程の **状態空間** と呼ぶ.

§B の random walk は確率過程 (確率連鎖) の重要な例である．広い意味では random walk とは独立確率変数列 $\{X_n\}$ の部分和の列 $W_n=x_0+\sum_{k=1}^{n}X_k$, $n\in\mathbf{Z}_+$ を指す．§B でも触れたように，random walk $\{W_n\}$ が与えられると $(k-1,W_{k-1})$ と (k,W_k) を $k=1,2,\cdots$ に対してつなぐことで $\omega\in\Omega$ に平面上の path (この場合は折れ線，一般的には $\mathbf{R}_+$ 上の連続関数) を対応させることができる．

例えば，$\mathbf{R}_+$ 上の実数値連続関数全体を $\mathcal{W}$ とおいて path 空間ということにする.「折れ線」の対応によって $W=(W_0,W_1,\cdots):\ \Omega\to\mathcal{W}$ と考えることができる．

より積極的に，このものの見方が「気持ち」以上に確率論的内容を持つために，W を path 空間に値をとる確率変数 (可測関数) と考えることができるか？ブラウン運動などの確率過程の研究のためには当然避けて通れない視点である．しかし， W を確率変数 (可測関数) とみるためには若干の準備を要する．W を可測関数とみるためには，先ず path 空間に可測空間の構造が必要である．独立確率変数列の極限定理では有限の n について $(X_1,\cdots,X_n)$ を考え，そこから和 W_n を作って実数値確率変数列の極限を考えていた．この場合は値域は直積可測空間 $(\mathbf{R}^n,\mathcal{B}_n)$ にとればいい．Random walk について有限の時刻までで考えるならばこれで十分.（以下の注意はいらない.）しかし，連続時間の確率過程や random walk で無限の未来まで一度で考えるときなどの一般的な path 空間では無限次元になるので，工夫が必要である．確率変数の値域は $(\mathbf{R}^n,\mathcal{B}_n)$ の範疇を越える．例えばブラウン運動は値域として連続関数の集合 C^0 を想定している．

では，値域の拡張が必要であるとして，どの程度まで拡張するのが「適切」か？[44] 一般的な「常識」として，特に弱収束（法則収束）を扱うときの観点などから，確率変数の値域としては，Poland 空間 (完備可分距離空間) で，σ 加法族として (Poland 空間の距離が定める開集合族を含む最小の σ 加法族としての)Borelσ 加法族をとれれば，この講義の結果を含む多くの確率論の一般論が使える．

Poland 空間はその可算直積空間に直積位相が入り，再び Poland 空間となる，という長所がある．実際 (X_i,d_i), $i=1,2,3,\cdots$, を集合とその距離の列とすると，$X=\prod_{i=1}^{\infty}X_i$ に距離 $d(x,y)=\sum_{i=1}^{\infty}2^{-i}(d_i(x,y)\wedge 1)$ が入り，(X,d) は Poland 空間になる．この距離は，有限個の成分に対して $d_i(x_i,y_i)<\epsilon$, $i=1,\cdots,k$, を満たす y の集合，という形の集合を開基（任意の開集合が，これらの形の集合の和集合で与えられる）とする位相を与える．この位相は成分毎の収束が (X,d) の収束になっている．$\mathbf{R}$, $\mathbf{R}^n$, 数列空間 $\mathbf{R}^\infty$, 連続関数の集合 $C^0([0,1])$ は全て Poland 空間である.（$C^0([0,1])$ については §D.3 に譲る.）

具体的には，random walk の場合は path 空間 $\mathcal{W}$ の距離として，各時刻毎に W_n が値をとる空間の距離 ρ_n に対して $\rho(w_1,w_2)=\sum_{k=1}^{\infty}2^{-n}(1\wedge\rho(w_{1,n},w_{2,n}))$ を距離とすればよい．

[44] 「適切」は ill-defined な概念だが，広げすぎると強い結果が得られないし，狭すぎると汎用性のある結果を得られない，という意味である．

この定義は simple random walk のように，各時刻でとりうる値が $\mathbf{Z}$ など離散的な場合は，有限個の時刻で値が一致するもの全体を開集合と定義する位相を与えることが分かる．

１次元 simple random walk の値域空間とその中での W の分布 $\mathrm{P}\circ W^{-1}$ を具体的に構成する試みの中で，上に書いたことの内容を説明する．

$$\Sigma=\{H,T\},\ r=\sharp\Sigma=2,\ \Omega=\{\omega=(\omega_1,\omega_2,\cdots)\mid \omega_i\in\Sigma\ (i\in\mathbf{N})\},$$
$$\mathcal{F}_n=\sigma[\{\{\omega\in\Omega\mid \omega_i=v_i,\ i=1,2,\cdots,n\}\mid (v_1,\cdots,v_n)\in\Sigma^n\}]\ \mathcal{F}=\sigma[\bigcup_{n\in\mathbf{N}}\mathcal{F}_n],$$
$$n\in\mathbf{N},\ A\in\mathcal{F}_n\ \text{に対して}\ \mathrm{P}[\,A\,]=r^{-n}\sharp A.$$

どの回でもいいから n 回分の硬貨の裏表を指定すると，結果がそうなる確率は 2^{-n} になる，という意味である．例えば $n=1$ の例で言うと，

$$\mathrm{P}[\,\{\omega\mid\omega_n=H\}\,]=\mathrm{P}[\,\{\omega\mid\omega_k=T\}\,]=1/2$$

が全ての k について成り立つ．これは,（いかさまでない，まともな）硬貨を（無限回）投げたとき，第 n 回目に表が出れば $w_n=T$, 裏が出れば $w_n=H$ とおいたことに相当する．

無限次元で新たに注意すべきことは，このように定義したときに確率空間が定義できることは証明を要する点である．$\bigcup_n\mathcal{F}_n$ の上だけで P が定義されているが，これを $\mathcal{F}$ に拡張できるかどうかの確認が必要である．

原理的にはこれがすめば $(\Omega,\mathcal{F},\mathrm{P})$ 上の任意の確率変数の期待値が計算できることになる．しかし，P が密度を持たないので，素朴な高校時代の計算は注意を要することがある．

Sample を一つ，例えば $\omega=(+-++\cdots)$ と固定しよう（簡単のため符号だけ書いた）．各 n 毎に $W_n(\omega)$ が定まる．これを $n\in\mathbf{Z}_+$ の関数と見ることができる．n を横軸に，$W_n(\omega)$ を縦軸にとればグラフに書ける．関数 $W_\cdot(\omega)$ は各時刻毎にゲームのコマがどこにいるかをたどっていく．つまり，コマの軌跡（道のり）を表す.（‘$\cdot$’ は，そこが関数の変数を代入する場所であることを表す便法）．$W_\cdot(\omega)$ を ω に対する**道 (sample path, sample function)** と呼んだ．即ち，確率過程には二つの見方があることになる：

(i) $n\in\mathbf{Z}_+$ を固定したとき，$W_n(\cdot):\ \Omega\to\mathbf{Z}$ （確率変数）,
(ii) $\omega\in\Omega$ を固定したとき，$W_\cdot(\omega):\ \mathbf{Z}_+\to\mathbf{Z}$ （時間の関数，即ち道）.

第１の見方が確率過程の定義であり，世の中は各時刻 n 毎にランダムで制御しきれない確率変数である，という「人の立場」と言えるかもしれない．すると，第２の見方は賽の目 ω を全て知っている「神の立場」ということになろうか．第２の見方で眺めると，Ω から関数の集まり（関数空間）への関数 $W_\cdot$ があることになる．Ω には確率が定義されているので，関数の集合が事象であるような確率空間を考えることになる，とも言える．

例では，異なる ω には異なる path $W_\cdot(\omega)$ が対応し，無数の path が実現しうる．これが「確率過程らしい」確率過程である．この例は **１次元酔歩 (one-dimensional simple random walk)** と呼ばれる．一本道を一歩毎にでたらめに行ったり来たりする，という描像である．例は単に酔歩（酔っぱらいの行動）やすごろくを表すだけではない．ω_n は各時刻 n 毎にランダムに値をとり，しかも，時刻間で独立なので（P の定義と独立の定義から $\omega_1,\,\omega_2,\,\cdots,$ が独立であることが示せる），本来 0 になるべき実験データ（あるいはゼロ信号）に加わったランダムな誤差（ノイズ）とみることができる.（比較的）簡単だが大変重要な，典型的な例である．

ω_n を実験データとみなすと，平均 $n^{-1}W_n$ に興味がある．大数の法則によれば $n^{-1}W_n(\omega)$ を計算すると，$n\to\infty$ の極限では確率 1 で 0 になる：$n^{-1}\lim_{n\to\infty}W_n=0$, a.s..

例では，時間がたつほどいろいろな道が可能になり，$n\to\infty$ を考えることが難しく見える．実際やさしくはない．しかし，$\{W_n\}$ は独立確率変数の和であることを生かすと，いろいろなことが分かる．一般に n が大きくなっていくときの確率過程の傾向を与える結果は極限定理と呼ばれる．大数の法則も極限定理の一つ．中心極限定理によれば $n^{-1/2}W_n$ の分布は $n\to\infty$ のとき正規分布 $N(0,1)$

に収束する．定理 86 のような大偏差値原理を用いれば，W_n が 0 からはずれているところの振る舞いが分かる．

さらに，path 空間を考えたので，与えられた path の集合に W が入る確率を考えることができるが，構成論的場の量子論と呼ばれる物理学の分野で連続極限と呼ばれる極限の取り方（または臨界現象の統計力学でスケーリング極限と呼ばれる極限の取り方）を考えることで，より精密な極限定理が得られる．

D.2 Brown 運動とは何か．

この節は [28, §7]，[27, §5.2] から引用した．

自然現象は時間が連続であり，いつでも「ノイズ」が生じうる．連続時間の場合，有限の時間間隔でも無限個の確率変数があり，sample path $W_t(\omega)$ の t に関する連続性のような，短い時間での変化を問う問題が生じる．

時間変数も状態空間も連続な確率過程の代表例，Brown 運動，を取り上げる．Brown 運動は，確率過程論の出発点になった代表例中の代表例であり，多くの興味深い，特徴的な性質を持つ．応用上の重要性も格別である．自然現象は時間的にも空間的にも連続なものが多いから，当然かも知れない．

1827 年，花粉の中の微小な粉が水中で不規則な運動を行うことが顕微鏡観察で発見された．Brown 運動の研究の出発点となったこの運動は，発見した植物学者にちなみ，Brown（ブラウン）運動と呼ばれる．Einstein（アインシュタイン）等の研究により，不規則な運動の原因は水分子の熱運動による撹乱であることが確立した．19 世紀末に「観測できないものは存在しない」という哲学から，見ることのできない分子や原子を否定した物理学者・科学哲学者 Mach（マッハ）が，Brown 運動を顕微鏡で見せられて分子の存在を認めた，という科学史上のエピソードもあるらしい．Brown 運動は，Brown 運動の数学モデルであると同時に，通信回線等の（アナログ）電気回路の熱雑音のモデルでもある．名前は，数学としての Brown 運動研究の端緒を与えた Wiener に由来する．

実際の世界は 3 次元空間だが，ここでは簡単のため 1 次元 Brown 運動に話を限る．

時間が連続な確率過程 $W_t,\ t \ge 0$, が 1 次元標準 Brown 運動であるとは，次の条件を満たすことである [1].

(i) 状態空間（W_t の値域）が実数上のボレル可測空間 $(\mathbf{R}, \mathcal{B}_1)$ で，sample path （時間 t の関数としての $W_t(\omega)$）は連続関数である．

(ii) 独立増分性（$t_1 < t_2 < \cdots < t_n$ とするとき，$W_{t_n} - W_{t_{n-1}}, W_{t_{n-1}} - W_{t_{n-2}}, \cdots, W_{t_1} - W_{t_0}$, が独立なこと）．

(iii) 定常性（$W_t - W_s$ の分布が $t - s$ だけで決まること）．

(iv) 分散の存在 $\mathrm{E}[\,W_1 - W_0\,] = 0$，かつ，$\mathrm{V}[\,W_1 - W_0\,] = 1$.

(v) 始点 $W_0 = 0$, a.s..

以上の性質を持つ確率過程が存在することは証明されている．次節以降は，存在証明に関する話題である．

Brown 運動 (確率過程 $W_t,\ t \in [0,1]$) の像測度 (連続関数空間上分布) $\mathrm{P} \circ W^{-1}$ を Wiener 測度という．

Brown 運動のように種々の性質が成り立つ確率過程が存在すること自体が，深い内容である．(安易にいろんな性質を仮定すると，そんな話のうまい確率過程は存在しない，ということになりかねない.)

なお，第 4 条件（分散と期待値の有限性）は，定数倍を除いて他の条件から得られるとのことである（分散が無限大ならば第 1 条件の path の連続性が成り立たない，ということ）．実際，分布の形まで決まってしまう．

定理 110 実数値をとる確率過程 W_t が上記 *Brown* 運動の第４条件以外を満たせば，ある実数 m 及び $v \geq 0$ がとれて，任意の $t > s$ に対して $W_t - W_s$ は，平均 $\mathrm{E}[\, W_t - W_s \,] = m\,(t-s)$，分散 $\mathrm{V}[\, W_t - W_s \,] = v\,(t-s)$ の正規分布に従う．

逆に，1 次元標準 Brown 運動の定義の条件のうち第１条件（sample path の連続性）の代わりに定理 110 の結論を Brown 運動の特徴付けの中心とすることもできるそうだ [12, §9] （[11, §3.2] も参照）．

定理 111 以下の条件を満たす確率過程 $W_t,\, t \geq 0,$ は標準 *Brown* 運動である．
状態空間（W_t の値域）が実数上のボレル可測空間 $(\mathbf{R}, \mathcal{B}_1)$ で，W_t の分布は期待値 0，分散 t の正規分布 $\mathrm{P}[\, W_t \leq a \,] = \dfrac{1}{\sqrt{2\pi t}} \displaystyle\int_{-\infty}^{a} e^{-x^2/(2t)} dx$ であり，定常性（$s < t$ のとき $W_t - W_s$ と W_{t-s} の分布が等しい）独立増分性を満たす（$t_1 < t_2 < \cdots < t_n$ とするとき，$W_{t_n} - W_{t_{n-1}}$, $W_{t_{n-1}} - W_{t_{n-2}}, \cdots, W_{t_1} - W_{t_0}$, が独立）．

つまり，定常性と独立増分性[45] の下では sample path の連続性と X_1 の分布が正規分布であることが同値になる！(一件不思議に見えるが，中心極限定理は，和の分布が正規分布に近づくことを言っている．つまり正規分布は小刻みな連続的な変化の中に自然に現れる分布である．逆に，正規分布であることから導かれる $\mathrm{E}[\, (W_t - W_s)^{2n} \,] = |t-s|^n$ が確率 1 で sample path の連続性を保証する．これが連続変形定理.）

結論として，Brown 運動の種々の確率や期待値などは正規分布の密度関数を用いて計算できる．一見抽象的に見えるかも知れない定義から具体的な公式を得る．

特に $s = 0,\, y = 0$ とおけば，定義より $W_0 = 0$ なので，$\mathrm{P}[\, W_0 = 0 \,] = 1$．これから，$\mathrm{P}[\, [W_t \in A \,] = \mathrm{P}[\, W_t \in A \mid W_0 = 0 \,]$．よって，

$$\mathrm{P}[\, W_t \in A \,] = \int_{x \in A} \rho_{0,t}(x)\, dx\,, \quad \rho_{0,t}(x) = \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{x^2}{2\,t}\right). \tag{72}$$

この式は，時刻 0 に原点 0 から出発 $(W_0 = 0)$ した Brown 運動の t 時間後の位置 W_t が平均 0 分散 t の正規分布に従うことを意味する．少し嘘を言えば，時刻 t に点 x にいる相対的な確率（確率密度）が，正規分布で表される．

水の中に一滴のインキを落とすと，次第に薄まりながら広がっていく．金属棒や板の一箇所を一瞬暖めると，次第に熱は広がって全体が一様な温度になる．これらの現象は，分子の熱運動という不規則な運動の，巨視的な効果である．微視的には（インキ粒子一つは） Brown 運動で記述される不規則な運動だが，多数の粒子が集まってインキ一滴となると，微視的な運動の生じる確率に応じた割合で，密度分布が実現する（これは粒子間の運動が独立であるという仮定の下で正しい）．例えば，長細い容器に入った水にインキを落とし，落とした点を原点 0，時刻を 0 とする．t 時間後点 x でのインキの密度（薄さ） $u(t,x)$ は (72) の確率密度に比例する；$u(x,t) \propto \rho_{0,t}(x)$．時間がたつにつれて山が低くなる正規分布となる．言い換えると密度関数は拡散方程式に従う．微視的 (sample path) には不規則な運動でも，巨視的（密度）にはなめらかな関数になる．

著者の力量不足により，この講義では Brown 運動の解析的性質の詳細についてはとても述べられない．

Brown 運動の sample path は連続関数になる，とあるが，いたるところ微分不可能であること（もっと強く 1/2–Hölder 連続であること）が知られている．従って，普通の意味での「微分」 $\frac{dW_t}{dt}$ は存在しない．ということは,「積分」 $\displaystyle\int f(W_s) dW_s$ も普通の Lebesgue–Stieltjes 積分の意味では存在し

[45] [11, §3.2] や [12, §9] では独立増分性と呼んでいる．他方 [1] では対応する性質を加法性と呼んでいる．前者は n 個の区間での増分が独立，後者は最後の区間の増分がそれ以前の区間の増分たちと独立，という違いらしいが，正確には知らない．Brown 運動に限れば全て成り立つ．

ない．しかし，確率積分という別の定義の意味で内容のある「積分」が定義できることが分かっている．確率積分はマルチンゲールと呼ばれる確率過程に拡張されている．

他方，インキ分子の（なめらかでない）Brown 運動と，その多数の集まり（期待値）としての，拡散方程式に従う（なめらかな）密度関数の関係に触れた．この間をつなぐ数学が推移半群と生成作用素の議論である．ここには Brown 運動の持つマルコフ性が反映する．

このように Brown 運動の先には非常に大きな解析学の領域が広がっている．以上の内容については，例えば [11, §3.2, §8.4] に極めて簡単なあらすじが書かれている．

D.3 連続関数の集合の上の確率空間．

時間を $t \geq 0$ とする代わりに，ここでは，有限区間に限る．$[0,1]$ の時間内の Brown 運動 $W = \{W_t \mid t \in [0,1]\}$, をつくりたい（存在を証明したい）．[12, §9 定理 9.2 の前の説明] によれば，現代の基礎教科書で採用される存在証明はだいたい２種類あるらしい[46]．

[46] [12, §9 最後の注] にもあるように [16] にはさらに少し違う方法が複数紹介されている．

[16, §1.4] では C^0 の有限加法族上で Wiener 測度の条件を満たす集合関数を構成し（これは容易），σ 加法性を証明することで（ルベーグ測度の構成と本質的に同じ方法）拡張定理によって C^0 上の測度を得る．σ 加法性の証明の際に連続性を経由するので，実質は連続変形定理経由と同じ方法である．

[16, §1.5] では，$L^2([0,1])$ 空間が Hilbert 空間であることを用いて，正規直交規定の線形結合で係数をランダムにすることを考える方法を紹介している． N. Wiener の構成方法は正規直交基底として三角関数を用いていた．

Brown 運動（Wiener 測度）の標準的な構成方法.（ポイントは，極限測度が連続関数上の測度として作れることをどのように保証するかの一点である.）

(i) 連続変形定理を経由する方法 [13, §7.1].
$[0,1]$ の中に dense な可算集合 T_0 をとる（例えば有理数.）$\mathbf{R}^{T_0}$ は Poland 空間の可算直積だから Poland 空間になる（各点収束位相）. Kolmogorov の拡張定理で Wiener 測度の条件を満たす測度が $\mathbf{R}^{T_0}$ 上に作れる. この sample が確率 1 で T_0 上一様連続であることを言う（ここで $W_t - W_s$ が正規分布に従うこと，特に $\mathrm{E}[\ (W_t - W_s)^4\ ] = C|t-s|^2$ を使う.）. このとき，path が連続になるように $[0,1]$ 上に (一意的に) 拡張する map で，測度を $C^0([0,1])$ 上の測度にすることができる.
おそらく今日確率論の講義ではもっとも標準的な定義であろう.

(ii) ランダムウォークの連続極限による方法 [12, §9].
$C^0([0,1])$ に一様収束位相 $d(w_1,w_2) = \max_{t\in[0,1]} |w_1(t) - w_2(t)|$ を入れれば Poland 空間になる. 1 次元 random walk(を線型内挿して $C^0([0,1])$ に値をとる確率過程とみたもの) の弱収束極限として Brown 運動を得る[47]. 確率変数または測度の列が（未知の測度に）弱収束することを証明するのに tightness argument を用いる.
$C^0([0,1])$ 上の Borel 測度について，測度列の tightness の条件と，極限分布が部分列によらないこと (2 つの測度が一致すること) を示す使いやすい条件，が必要なので，確率論では最初に出てくる方法ではないが，ランダムウォークの連続極限という意味合いがはっきりするので，数理物理学のイメージ（を確率論に翻訳したもの）からは恐らくいちばん直訳に近い.

(iii) 直交関数展開を利用する方法 [9, 問題六.8]，[16, §1.5 Problem 3].
X_k, $k = 0,1,2,\cdots$, を i.i.d. で標準正規分布 $N(0,1)$ を分布に持つとするとき，$B(t) = \dfrac{t}{\sqrt{\pi}} X_0 + \sqrt{\dfrac{2}{\pi}} \sum_{k=0}^{\infty} \sum_{j=3}^{2^{k+1}} X_j \dfrac{1}{j} \sin jt$ は $t \in [0,\pi]$ に関して (確率 1 で) 一様収束する.
概収束を経由することで，もっとも初等的にできるので，初期の教科書はここから入ったようだが，今日では極限の連続性がポイントであることを明示するために，および Hölder 連続性などの path property が構成可能性に本質であることを明示するために[48]，連続変形定理から入ることが多いようだ.

[12, §9 定理 9.2] の周辺には，一見全く違う両者のアプローチが，本質的に同じ内容量を持つ作業であることを説明している. 第 2 の方法で，極限分布の一意性は有限次元分布 (有限個の時刻の W_t の結合分布) をみれば十分. $C^0([0,1])$ 上の測度の列の tightness の条件の中に極限測度の sample の連続性を保証する条件が入っているので， $C^0([0,1])$ の中で考えても，第 1 の方法と同様に sample の連続性の作業をさぼることはできない.

以下，最後の 2 小節で，最初の二つの構成方法をそれぞれ概観する.（詳しいことは省くが，あらすじと文献の引用を行う.）

D.4 連続変形定理.

この節は [13, §7.1] の引用である.

どうせ §D.2 の定義から 定理 110 の性質が出てくるとのことだから，[13, §7.1] に従って，最初から，path の連続性以外の全てを $\mathbf{R}^{\mathbf{Q}_+}$ 上で満たすように作ろう. すると有限個の時刻 $T = \{t_1,\cdots,t_n\}$

[48] Random walk の一歩の分布は期待値と分散が有限ならば何でもよいことが知られている（Donsker's invariance theorem. 中心極限定理で，元の分布が何でも極限分布が正規分布になったことと対応する）.
[48] だと思う.

($n \in \mathbf{N}$) における path の位置の結合分布 $\mu_T = \mu_{t_1,\cdots,t_n}$ は正規分布を用いて書けねばならない．即ち，$A_i \in \mathcal{B}_1$, $i = 1, \cdots, n$, に対して

$$\mu_T(A_1 \times \cdots \times A_n) = \mu_{t_1,\cdots,t_n}(A_1 \times \cdots \times A_n) = \int_{A_1} dx_1 \cdots \int_{A_n} dx_n \prod_{i=1}^n p_{t_i - t_{i-1}}(x_{i-1}, x_i). \quad (73)$$

ここで，$x_0 = 0$ (x_0 を x とおけば，原点ではなく x を始点とする Brown 運動を得る) および $t_0 = 0$ であり，さらに，

$$p_t(a,b) = \frac{1}{\sqrt{2\pi t}} e^{-(b-a)^2/(2t)}. \quad (74)$$

(74) からあらわに計算することで

$$\int_{\mathbf{R}} p_{t_2-t_1}(x,y) p_{t_3-t_2}(y,z)\, dy = p_{t_3-t_1}(x,z). \quad (75)$$

この式は，測度の族 $\{\mu_T \mid T \subset [0,1],\ \sharp T < \infty\}$ が整合していること，即ち，$T = \{t_1, \cdots, t_n\}$ のとき

$$\mu_T(A_1 \times \cdots \times A_{j-1} \times \mathbf{R} \times A_{j+1} \times \cdots \times A_n) = \mu_{T\setminus\{t_j\}}(A_1 \times \cdots \times A_{j-1} \times A_{j+1} \times \cdots \times A_n) \quad (76)$$

を表す．

この測度を $[0,1]$ 上の実数値連続関数上の測度に拡張できるだろうか？即ち，$[0,1]$ 上の関数全体の集合 Ω_0 上の確率測度 ν があって，

$$A = \{w \in \Omega_0 \mid w(t_i) \in A_i,\ 1 \le i \le n\} \quad (77)$$

に対して

$$\nu(A) = \mu_T(A_1 \times \cdots \times A_n) \quad (78)$$

が成り立ち，しかも，$\nu(C^0([0,1])) = 1$ が成り立つようにできるだろうか？もし，拡張できれば，(78) と有限次元分布の性質から，加法性，定常性，分散の存在，始点，は求めるものになっているから，ν が求めるものになる．

実はこの時点で若干の注意が必要である．詳しくは [13, §7.1 定理 1.3, 練習 1.1]，[12, §9 Separable stochastic processes] 等を参照．そこに紹介されている，やってはいけない議論 (間違ってはいないが連続関数上の測度につながらない議論)．$\mathbf{R}$ 上の関数の全体 Ω_0 を，「有限次元集合」即ち，(77) の形の集合 ($A_i \in \mathcal{B}_1$, $i = 1, \cdots, n$, $0 \le t_1 < \cdots < t_n$, $n \in \mathbf{N}$) 全体が生成する σ 加法族 $\mathcal{F}_0$ で可測空間とする．整合性条件 (76) があると，Kolmogorov の拡張定理 ([12, Appendix II]，[13, §A.7] 等を参照) によって，$(\Omega_0, \mathcal{F}_0)$ 上に確率測度 ν が構成できて，(77) の A に対して (78) が成り立つ．

これが求める拡張である，と思いたいところだが，しかし，これではダメである．なぜなら，連続関数の上の測度にしたいのであるが，連続関数という性質は可測にならない（即ち連続関数の全体という集合は $\mathcal{F}_0$ に入らない）から当然 $\nu(C^0([0,1])) = 1$ も言えないからである．証明は [13, §A.7]）．

$C^0([0,1])$ を可測にするに，σ 加法族の構造を取り替える．$[0,1]$ の中に dense な可算集合 T_0 をとる．例えば以下では定理 112 での簡単のため [13, §7.1] に従って 2 進有理数 $T_0 = \{m2^{-n} \in [0,1] \mid m, n \in \mathbf{Z}_+\}$ にとる．

$\mathbf{R}^{T_0}$ は Poland 空間の可算直積だから Poland 空間になる（各点収束位相）．Kolmogorov の拡張定理で (78) を満たす測度 ν が $\mathbf{R}^{T_0}$ 上に作れる．

この sample が確率 1 で T_0 上一様連続であることを言う（この部分がもっとも系の詳細を使うので後述）．Path を連続になるように $[0,1]$ 上に (一意的に) 拡張する map $\iota:\ \mathbf{R}^{T_0} \to C^0([0,1])$ で，測度を $C^0([0,1])$ 上の測度にすることができる．ι の可測性は

$$\begin{aligned}&\iota^{-1}(\{w \in C^0([0,1]) \mid \max_{t\in[0,1]} |w(t)| \le \delta\}) = \{w \in \mathbf{R}^{T_0} \mid |w(t)| \le \delta,\ t \in T_0\} \\ &= \bigcap_{t \in T_0} \{w \in \mathbf{R}^{T_0} \mid |w(t)| \le \delta\} \in \mathcal{F}_0\end{aligned}$$

から.（このことから $C^0([0,1])$ のボレル集合（ max ノルムで距離が入っている）B に対して $\iota^{-1}(B) \in \mathcal{F}_0$ となるので,）$\nu \circ \iota$ は $C^0([0,1])$ 上のボレル確率測度である[49]．作り方から，Wiener 測度になっている．即ち $w \in C^0([0,1])$ に対して $W_t(w) = w(t)$, $t \in [0,1]$, で定義される確率過程は Brown 運動になる．

$\mathbf{R}^{T_0}$ 上の確率測度 ν の sample が確率 1 で T_0 上一様連続であることを言うことが，連続な path を持つ確率過程としての Brown 運動の存在定理の証明におけるもっとも中心的な定量的評価である．次の定理は，$W_t - W_s$ の高次モーメント (べきの期待値) が時間のべきで抑えられることから個別の path 毎に (sample 毎に) $W_t - W_s$ も時間のべきで抑えられることを言う．

定理 112 $T_0 = \{m2^{-n} \in [0,1] \mid m, n \in \mathbf{Z}_+\}$ 上の確率過程 W_t, $t \in T_0$, がある $0 < \alpha < \beta$ と $K > 0$ に対して $\mathrm{E}[\,|W_s - W_t|^\beta\,] \le K|t-s|^{1+\alpha}$, $t, s \in T_0$, を満たすとする．このとき $0 < \gamma < \alpha/\beta$ に対して $C = C(w)$ が存在して

$$|W_q - W_r| \le C|q-r|^\gamma, \quad q, r \in T_0\,,\ a.e..$$

証明. [13, §7.1 定理 1.5] の証明を引用する．

添字をみやすくするために $W(x) = W_x$ のように書く．$\eta > 0$, $I_n = \{(i,j) \mid 0 \le i \le j \le 2^n,\ 0 < j - i \le 2^{n\eta}\}$, $G_n = \{|W(j2^{-n}) - W(i2^{-n})| \le ((j-i)2^{-n})^\gamma,\ (i,j) \in I_n\}$, とおく．

チェビシェフの不等式から

$$\nu(G_n^c) \le \sum_{(i,j)\in I_n} ((j-i)2^{-n})^{-\beta\gamma}\mathrm{E}[\,|W(j2^{-n}) - W(i2^{-n})|^\beta\,].$$

$\alpha - \beta\gamma > 0$ と $\sharp I_n \le 2^n 2^{n\eta}$ と仮定から

$$\nu(G_n^c) \le K\sharp I_n(2^{n\eta}2^{-n})^{-\beta\gamma+1+\alpha} = K2^{-n\lambda}.$$

ここで $\lambda = (1-\eta)(1+\alpha-\beta\gamma) - (1+\eta)$ とおいた．$\gamma < \alpha/\beta$ なので $\eta > 0$ を十分小さく取れば $\lambda > 0$ とできる．

補題 113 $A = 3\,2^{(1-\eta)\gamma}/(1-2^{-\gamma})$ とおく．$H_N = \bigcap_{n \ge N} G_n$ 上では

$$|W(q) - W(r)| \le A|q-r|^\gamma, \quad q, r \in T_0\,; |q-r| < 2^{-(1-\eta)N}.$$

補題 113 を仮定して定理の証明を続ける．

$$\nu(H_N^c) \le \sum_{n=N}^\infty \nu(G_n^c) \le K\sum_{n=N}^\infty 2^{-n\lambda} = \frac{K}{1-2^{-\lambda}}2^{-N\lambda}.$$

なので，確率 1 の $w \in \Omega_0$ に対して $\delta = \delta(w) > 0$ がとれて[50]，

$$|W(q) - W(r)| \le A|q-r|^\gamma, \quad q, r \in T_0\,; |q-r| < \delta(w).$$

一般の $q, r \in T_0$ に対しては，$|s_i - s_{i-1}| < \delta(w)$, $i = 1, \cdots n$, を満たすように $q = s_0 < s_1 < \cdots < s_n = r$ と分割し，$C(w) = 1 + \delta(w)^{-1}$ とおくとき $|W(q) - W(r)| \le AC(w)|q-r|^\gamma$ を得る．ここで $0 < \gamma < 1$ のとき三角不等式の「逆」$p^\gamma + q^\gamma \le (p+q)^\gamma$, $p, q \ge 0$, が成り立つことを用いた．

[49] $C^0([0,1])$ ではここでの可測集合族は，max ノルムで距離を入れたときのボレル集合族と一致することは，有限次元直積集合の $C^0([0,1])$ への逆像が両者を生成することで示される．

[50] ここは途中の計算が省略されている．$\nu(\bigcup_{N=1}^\infty \bigcap_{n=N}^\infty G_n) = 1$ となるから，確率 1 の $w \in \Omega_0$ に対して $N = N(w)$ が存在して，$n \ge N(w)$ ならば $w \in G_n$．区間 $[q, r]$ を 2^{-n}, $n \ge N$, の点で最小限の分割を施して，各分割毎に G_n の性質を使う，とうまく行ったと思う．

あとは補題 113 を証明すればよい．

$q, r \in T_0$, $0 < r - q < 2^{-(1-\eta)N}$, とする．$m \ge N$ を $2^{-(m+1)(1-\eta)} \le r - q < 2^{-m(1-\eta)}$ が成り立つように選び，$m < r(1) < \cdots < r(\ell)$, $m < q(1) < \cdots < q(k)$ を $r = j2^{-m} + 2^{-r(1)} + \cdots + 2^{-r(\ell)}$, $q = i2^{-m} - 2^{-q(1)} - \cdots - 2^{-q(k)}$, が成り立つように選ぶ．

$0 < r - q < 2^{-m(1-\eta)}$ なので $j - i < 2^{m\eta}$ となって，H_N 上では

$$|W(i2^{-m}) - W(j2^{-m})| \le (2^{m\eta}2^{-m})^{\gamma}.$$

H_N 上では三角不等式から

$$|W(q) - W(i2^{-m})| \le \sum_{h=1}^{k}(2^{-q(h)})^{\gamma} \le \sum_{h=m}^{\infty}(2^{-\gamma})^{h} \le C_{\gamma}2^{-\gamma m}.$$

ここで $C_{\gamma} = \dfrac{1}{1-2^{-\gamma}} > 1$．この計算を繰り返すと，

$$|W(r) - W(j2^{-m})| \le C_{\gamma}2^{-\gamma m}.$$

以上の 3 式を合わせると， $2^{-m(1-\eta)} \le 2^{1-\eta}|r-q|$ なので

$$|W(q) - W(r)| \le 3C_{\gamma}2^{-\gamma m(1-\eta)} \le 3C_{\gamma}2^{(1-\eta)\gamma}|r-q|^{\gamma}.$$

□

D.5 Random walk の連続極限．

この節は [12, §9]，および [28, §7]，[27, §5.1, §5.2] からの引用である．

Random walk の連続極限とは，長いすごろくを遠くから眺めることである．

Brown 運動の作り方としては理論的にもっとも単純とは言えないかもしれないが，ある観点からは極めて興味深い，作り方を説明する．

§B で 1 次元 random walk を紹介した．単に random walk というと，人や状況によって多少指す対象の範囲が変わると思うが，平たく言えば，整数時刻の実数値確率過程 Z_n, $n \in \mathbf{Z}_+$, で増分が独立同分布，即ち，i.i.d. 確率変数列 ξ_k, $k \in \mathbf{N}$, に対して $Z_n = x_0 + \sum_{k=1}^{n}\xi_k$, と書けるものを指す．(記号 W は Brown 運動にとっておきたいので Z を用いた．) 以下， $x_0 = 0$ とするが，出発点の変更だけなので，対応する結果は一般の x_0 について当然成り立つ．

特に，1 次元 simple random walk といえば，ξ_1 が硬貨投げ ($\mathrm{P}[\,\xi_1 = 1\,] = \mathrm{P}[\,\xi_1 = -1\,] = 1/2$) を指す．

時刻が離散的だが，これを連続にし，線型内挿によって，sample を連続関数にする．(単にこれからの話を，簡単のために連続関数の集合だけで考える目的．連続関数にしなくても同じ結果を得るが，連続関数より大きな集合を定義しないといけないのがわずらわしいだけのこと.)

定理 114

$$W_{\lambda}(t,\omega) = \frac{1}{\sqrt{\lambda}}Z_{[\lambda t]}(\omega) + (\lambda t - [\lambda t])\frac{1}{\sqrt{\lambda}}\xi_{[\lambda t]+1}(\omega),\ \omega \in \Omega,\ \lambda > 0,\ t \in [0,1],$$

は $\lambda \to \infty$ のとき，*Brown* 運動 W_t に法則収束する．

（ここで $[\lambda t]$ は，実数 λt の整数部分をとる *Gauss* の記号.）

法則収束は §D.4 の Wiener 測度 ν を用いて，$\nu(\partial A) = 0$ を満たす任意のボレル可測集合 $A \subset C^0([0,1])$ に対して，$\lim_{\lambda\to\infty} \mathrm{P}[\,W_{\lambda} \in A\,] = \nu(A)$ が成り立つ，という意味である (§D.4 参照).

定理 114 は，simple random walk という，計算機でもシミュレーションしやすく，数学的にも直感的にも比較的簡単な確率過程の極限として，Brown 運動を「作る」ことができることを表している．時空ともに離散的な確率過程が時空ともに連続的な Brown 運動に近づく，というのも興味深いが，上記の定理は，次の定理の意味で，Brown 運動の「自己相似性」を表す[51]．

定理 115 *Brown* 運動 W_t の分布は，任意の正の実数 a に対して *Brown* 運動を時空についてスケール変換（拡大・縮小）した確率過程 $a^{-1/2}\,W_{[a\,t]}$ の分布に等しい．

証明. 定理 114 において，λ を $\lambda\,a$ と置き換えて $\lambda\to\infty$ としても（極限をとる変数の変数変換だから）構わない．即ち，$(\lambda\,a)^{-1/2}\,W_{a\lambda}(t,\omega)$ は $\lambda\to\infty$ のとき Brown 運動 W_t に法則収束する．これは，定理 114 において $t\to a\,t,\,W\to a^{-1/2}\,W$ と置き換えたものになっている．よって，$\lambda\to\infty$ で $a^{-1/2}\,W_{a\,t}$ にも法則収束する．故に，$a^{-1/2}\,W_{a\,t}$ と W_t の分布は等しい． □

定理 114 は W_t の時間と空間をうまく呼応させてスケール変換（拡大・縮小）した後，極限をとることになっている．このような極限の取り方を scaling limit と呼ぶ．他方， 定理 115 は Brown 運動が，時間と空間をうまく呼応させてスケール変換すると，再び Brown 運動に戻る（分布が不変である）こと，を表す．これを分布のスケール不変性 (scale invariance) という．物理用語では統計的自己相似性 (statistical self-similarity) ともいう.「統計的」という修飾語は，一つ一つの見本関数 sample path の関数としての自己相似性ではなく，分布（確率）の自己相似性に言及していることからつく．

大雑把に言うと，Brown 運動を顕微鏡で 3 倍拡大してビデオに録画し，$1/9=1/3^2$ 倍のスロー再生で再生画像を見ると,「元の Brown 運動らしく」見える，ということである．拡大と早回し（時空のスケール変換）の関係は古典力学でもおなじみである．怪獣映画やアクション映画で，ビルが爆破されて破片が飛び散る（放物運動）シーンを，ミニチュア撮影ですませるには，スローモーションで再生する必要がある．違いは，放物運動の場合は拡大が 3 倍ならばスピードは $1/\sqrt{3}$ 倍のスロー再生，ということである．2 乗か 1/2 乗かの違いがある．

「統計的」という修飾語にもう一つの特徴がある．放物運動は時空の適切なスケール変換で正確に元の運動に一致する（自己相似である）．Brown 運動の sample path は自己相似ではない．不規則な運動が拡大・スロー再生でぴたっと一致したら，かえっておかしい．あくまで「典型的な振る舞いが元の path と同じ」に過ぎない．しかし，そもそも，Brown 運動の典型的な sample path の振る舞いとはなにか？これは直感的にすら容易ではない．諸君が各自計算機で sample path を発生させてそれをいくつか図示したときに，共通するものを直感したならば，それが Brown 運動に対する諸君の直感ということになろうか.

Brown 運動は自然現象（例えば，分子の大きさの熱運動に由来する現象を扱う統計力学の分野）を表すには単純過ぎて，十分当てはまらない現象が多いことが分かっている．しかし，Brown 運動とは一致しなくても，統計的自己相似性には関連すると見られる現象が普遍的に知られている．この意味で，定理 115 は Brown 運動を越えて興味ある結果である．

定理 114 は，次の定理の特別な場合である．

---

[51] 連続極限の方法がスケール不変性を見やすいかのような記述になっているが，誤解を招く．$\{a^{-1/2}W_{at}\}$ が標準 Brown 運動の定義を満たすことはガウス分布の性質から容易に分かるので，分布の唯一性に基づいて，スケール不変性は自動的に得られる．つまりどの方法で構成してもスケール不変性は分かる．

定理 116 (Donsker's invariance principle) $\xi_k,\ k\in\mathbf{N}$, が *i.i.d.* でその分布が平均 0, 分散 $0<\sigma^2<\infty$ とする．このとき，

$$W_n(t,\omega)=\frac{1}{\sigma\sqrt{n}}\sum_{k=1}^{[nt]}\xi_k(\omega)+(nt-[nt])\frac{1}{\sigma\sqrt{n}}\xi_{[nt]+1}(\omega),\ \omega\in\Omega,\ n>0,\ t\in[0,1],$$

は $n\to\infty$ のとき，*Brown* 運動 W_t に法則収束する．

即ち ξ_1 は硬貨投げであることは本質ではなく，分散が有限なら分布は何でもかまわない．それゆえ invariance principle と呼ばれる.（物理用語で普遍性 universality と呼ばれる広い性質の一つである.）

Invariance principle については [12, §10, Theorem 10.1] に譲る．以下では，random walk の極限として Brown 運動を得る方法を，[12, §9, Theorem 9.1] から引用する（Brown 運動の存在の別証明）.

定理 117 $\xi_k,\ k\in\mathbf{N}$, が *i.i.d.* でその分布が平均 0, 分散 1 の正規分布とする．このとき

$$W_n(t,\omega)=\frac{1}{\sigma\sqrt{n}}\sum_{k=1}^{[nt]}\xi_k(\omega)+(nt-[nt])\frac{1}{\sigma\sqrt{n}}\xi_{[nt]+1}(\omega),\ \omega\in\Omega,\ n>0,\ t\in[0,1],$$

は $n\to\infty$ のとき法則収束する．
ここで，$C^0([0,1])$ に max ノルムで距離を入れた位相空間のボレル加法族を考えて可測空間とする．この空間に値を取る確率変数列 W_n の法則収束（この空間上のボレル確率測度の列の弱収束）の意味である．
極限分布は *Brown* 運動の性質を満たす．即ち，*Brown* 運動の存在証明になっている．

証明の全貌は [12] に譲り，あらすじのみ記す．

これから Brown 運動の存在を証明しようというのだから，極限測度はまだ分かっていない．そこで tightness argument を用いて，とにかく $C^0([0,1])$ 上の確率測度に収束する，ということを言う．それが Brown 運動の性質を満たすことは，むしろやさしい．

(i) $C^0([0,1])$ 上のボレル確率測度列の弱収束をいうために，確率測度の族の tightness を $\mathbf{R}$ 上のボレル確率測度の場合にならって定義する．
一般に距離空間 Ω 上の確率測度の族 Π が tight であるとは任意の $\epsilon>0$ に対してコンパクト集合 $K\subset\Omega$ がとれて，任意の $P\in\Pi$ に対して $P(K)>1-\epsilon$ が成り立つことをいう [12, §6].

(ii) $\{W_n\}$ が tight であることを証明するのに便利な定理を用意する．

定理 118 ([12, §8, 定理 8.4]) $\xi_i,\ i\in\mathbf{N}$, を確率変数の列とし，

$$W_n(t,\omega)=\frac{1}{\sigma\sqrt{n}}\sum_{k=1}^{[nt]}\xi_k(\omega)+(nt-[nt])\frac{1}{\sigma\sqrt{n}}\xi_{[nt]+1}(\omega)$$

とおく *($\sigma>0$ は正定数)*.

$$(\forall\epsilon>0)\ \exists\lambda>1,\ \exists n_0\in\mathbf{N};\ (\forall n\geq n_0)\ \mathrm{P}[\ \max_{i\leq n}|\sum_{j=k+1}^{k+i}\xi_j|\geq\lambda\sigma\sqrt{n}\ ]\leq\epsilon\lambda^{-2},$$

が成り立つならば，$\{W_n\}$ は *tight* である．

詳細は教科書に任せるが，定理 118 は次の定理の簡単な変形である．

$w \in C^0([0,1])$ の modulus of continuity m_w を次で定義する．

$$m_w(\delta) = \sup_{|s-t|<\delta} |w(s) - w(t)|, \quad 0 < \delta \le 1\,.$$

$m_w(\delta)$ は w の関数として連続ゆえ可測なことに注意.

定理 119 ([12, §8, 定理 8.2]) $C^0([0,1])$ 上の確率測度の列 $\{\mathrm{P}_n \mid n \in \mathbf{N}\}$ が *tight* であることは次の 2 条件の同時成立と同値である．

(a) $(\forall\eta > 0)\ \exists a;\ \mathrm{P}_n[\ \{w \in C^0([0,1]) \mid |w(0)| > a\}\ ] \le \eta,\ n \in \mathbf{N}\,.$
(b) $(\forall\epsilon, \eta > 0)\ \exists\delta;\ 0 < \delta < 1,\ \exists n_0;\ \mathrm{P}_n[\ \{w \in C^0([0,1]) \mid m_w(\delta) \ge \epsilon\}\ ] \le \eta,\ n \ge n_0\,.$

証明. $\{\mathrm{P}_n\}$ が tight とする．
$\epsilon, \eta > 0$ をとる．Tight だから $\mathrm{P}_n[\ K\ ] > 1 - \eta, n \in \mathbf{N}$, なるコンパクト集合 $K \subset C^0([0,1])$ がとれる．Ascoli–Arzelà 定理から，十分大きな a と十分小さな $\delta > 0$ に対して，

$$K \subset \{w \in C^0([0,1]) \mid |w(0)| \le a,\ m_x(\delta) < \epsilon\}$$

となるので，2 条件を得る．
次に定理の 2 条件を仮定する．完備可分距離空間では一つの確率測度は tight である.（証明は確率論というよりは位相空間論の基礎事項なので [12, §1, 定理 1.4，Appendix I] に任せて省略する.）だから，定理の第 2 条件の n_0 は最初から 1 としてよい．
$\eta > 0$ をとる．仮定から， $a > 0$ と $\delta_k > 0,\ k \in \mathbf{N}$, を $A = \{w \mid |w(0)| \le a\}$， $A_k = \{w \mid m_w(\delta_k) < k^{-1}\},\ k \in \mathbf{N}$, とおくとき，

$$\mathrm{P}_n[\ A\ ] \ge 1 - \frac{1}{2}\eta,\ n \in \mathbf{N},\ \ \mathrm{P}_n[\ A_k\ ] \ge 1 - \eta 2^{-k-1},\ k, n \in \mathbf{N},$$

を満たすように取ることができる．$K = \overline{A \cap \bigcap_{k=1}^{\infty} A_k}$ とおくと，$\mathrm{P}_n[\ K\ ] \ge 1 - \eta,\ n \in \mathbf{N}$, であり，かつ，Ascoli–Arzelà の定理から K はコンパクトである．よって $\{\mathrm{P}_n\}$ は tight である．
□

なお，Ascoli–Arzelà 定理は次のように書かれる．証明は [12, Appendix I].

定理 120 $A \subset C^0([0,1])$ に対して $\bar{A}$ がコンパクトになることと次の 2 条件の同時成立は同値である．

(a) $\sup_{w \in A} |w(0)| < \infty.$
(b) $\lim_{\delta\to 0} \sup_{x \in A} m_x(\delta) = 0\,.$

(iii) いまの場合に $\{W_n\}$ が 定理 118 の仮定を満たすことを証明する．この箇所で 定理 112 に相当する連続性の評価が入る.（連続な確率過程 W_n の極限であっても，極限では動きが速くなりすぎて不連続になったり遅すぎて止まってしまう可能性があるので，連続性の評価は必要である.）Tightness に関する長い準備の下でこの部分の証明は定理 112 のときより短い．$\mathrm{P}[\ \max_{i \le n} |\sum_{k=1}^{i} \xi_k| \ge 2\lambda\sqrt{n}\ ]$ に対して定理 44 の証明と類似の分解による上からの評価を導き，$\sum_{k=1}^{n} \xi_k$ についての絶対値 3 次

モーメントの有限性を用いてチェビシェフの不等式を適用する．これより

$$(\forall\epsilon > 0)\ \exists\lambda > 1;\ (\forall n \in \mathbf{N})\ \mathrm{P}[\ \max_{i\le n} |\sum_{k=1}^{i} \xi_k| \ge 2\lambda\sqrt{n}\ ] \le \epsilon\lambda^{-3}$$

を得るので， 定理 118 を適用すれば $\{W_n\}$ は tight となる [12, §9].

(iv) Tight ならば任意の部分列に対して弱収束する部分列がとれる．これは次の定理による．

定理 121 ([12, 定理 6.1, 6.2]) *Poland*空間上の確率測度の族については *tight* と *relatively compact* は同値である．

さらに，弱収束極限測度が常に一致するならば，元の列が弱収束することになり，極限測度の存在が言える．

ここで次の事実に注意する．

定理 122 $C^0([0,1])$ 上に 2 つのボレル確率測度があったとき，有限次元分布が全て一致すれば，それらは一致する．

証明. $C^0([0,1])$ 上の確率測度 ν の有限次元分布とは，$w \in C^0([0,1])$ と $t_1,\cdots,t_n \in [0,1]$ $(n \in \mathbf{N})$ に対して，$\pi_{t_1,\cdots,t_n}:\ w \mapsto (w(t_1),\cdots,w(t_n)) \in \mathbf{R}^n$ とおくとき，像測度 ($(\mathbf{R}^n,\mathcal{B}_n)$ 上の測度) $\nu\circ\pi^{-1}_{t_1,\cdots,t_n}$ を指す．π は連続写像なので n 次元ボレル集合 A に対して $\nu\circ\pi^{-1}_{t_1,\cdots,t_n}(A)$ は $C^0([0,1])$ のボレル集合である．逆に， $C^0([0,1])$ の閉球は対応する有限次元集合の極限である． $C^0([0,1])$ が可分なので，開集合は可算個の開球（従って可算個の閉球でも可）の和である．よって，有限次元集合たちの全体 (有限加法族をなす) は $C^0([0,1])$ のボレル σ 加法族を生成する．よって，有限次元分布が一致する 2 つの確率測度は一致する [12, §3]. □

いまの場合，部分列によらず有限次元分布が一致することは，各 W_n の有限次元分布が（多次元の）正規分布であることが分かるので，分散の極限の一致だけ確認すれば良く，やさしい [12, §9].

(v) Brown 運動の性質を満たすことを証明する（各 W_n の有限次元分布が（多次元の）正規分布であることが分かるので，分散の収束だけ確認すれば良く，やさしい [12, §9]）.

なお，上で確率測度の弱収束と 2 つの確率測度の一致で必要な条件が大きく異なる (tightness) ことに注意 [12, §7 終わり].

参考文献

[1] 数学辞典，第 3 版，岩波書店.

[2] 伊藤清，確率論，岩波基礎数学選書，1991.

３分冊の岩波講座として長く親しまれた．[9] が出るまでは,「測度論に基づく確率論」の初学者向け本格的教科書は日本では殆どこれしかなかったのではないだろうか．いまでも勉強になるものが証明の随所にある気がするが，いつ注文しても絶版なのはたいへん残念.

[3] 伊藤清三，ルベーグ積分入門，数学選書 4，裳華房，1996.

確率論に密着した測度論の教科書．確率論への接続を考えるならば測度論の講義の入門教科書として適当.

[4] 熊谷 隆，共立出版（新しい解析学の流れ），2003.

標準的な教科書の一冊．マルチンゲールの応用や抵抗回路によるディリクレ形式の表現などが特徴的．新しい教科書の代表として挙げておく.

[5] 小針あき[52] 宏，確率・統計入門，岩波，1973.

古典的な範囲の確率論．古典統計学も含む．測度論に基づく確率論には触れていないので，現代の数学科の専門科目の確率論の教科書とは言えないが，いろんな基礎公式を初等的に証明しているので，記述や計算の初等化という点で参考になる．

[6] 佐藤坦，測度から確率へ，共立出版，1994.

これも「西尾以後」の確率論の基礎教科書の一つ．見かけはやさしいが内容は標準的教科書一冊分.

[7] 志賀徳造，ルベーグ積分論から確率論，共立講座 21 世紀の数学 10，共立出版，2000.

これも「西尾以後」の確率論の基礎教科書の一つ．ガンマ関数の漸近公式を例にして，母関数と漸近解析について一章をあてているのが特色．著者の研究の関係で voter model に言及がある．

[8] 田崎晴明，熱力学，新物理シリーズ，培風館， 2000.

物理の一分野熱力学の基礎教科書．ここではルジャンドル変換の説明を引用しただけだが，熱力学の基礎教科書として推薦できる良い本.

[9] 西尾真喜子，確率論，実教出版，1978.

20 世紀後半の比較的長い間，日本人の書いた数学科の確率論の基礎教科書というと [2] 以降は殆どこれしかなかったようだ．内容や証明は [15] に準拠しているように思う.

[10] 樋口保成，パーコレーション — ちょっと変わった確率論入門，遊星社，1992.

表題の通りの教科書だが，日本の数学にはなかなかこういう本が出ない．パーコレーションという一点で話が進むのに，厳密数学としての確率論の基礎教科書にもなっている．パーコレーション特有の部分に関しては読みこなすのに腕力が必要だが，確率論の教科書としても self-contained である．薄い本なのになぜかやり遂げている．おもしろい.

[11] 福島正俊，確率論，裳華房，1998.

本格的な定義を最小限にして，ブラウン運動まで言及する，というスタイル．一般的定義を最小限にするため，言及できない重要事項は証明しない，という点で数学科の教科書としては self-contained でない，ということになるが，確率論の最初から始めて最小限の講義時間でブラウン運動について何が説明できるか，という課題の参考になりそうに思う.

[12] P. Billingsley, *Convergence of probability measures,* John Wiley & Sons, 1968.

本講義の水準を越える advanced な内容．確率過程 (関数空間上の確率測度) の弱収束の詳しいていねいな教科書．Brown 運動の連続極限による構成を引用した.

[13] R. Durrett, *Probability: Theory and examples,* 2nd ed., Wadsworth, 1996.

英語の確率論の本格的現代的教科書の一つ．基礎から始めながらブラウン運動までカバーする．後半は本講義の水準を越える advanced な内容．Random walk や Brown 運動のコンパクトな記述を引用した.

[14] R. Durrett, *Stochastic calculus: A practical introduction,* CRC Press, 1996.

[13] の続き．Brown 運動の構成から始まって，その基礎性質，確率微分方程式，拡散，と基礎事項を解説している.

[52] Unicode 774D 目見

[15] W. Feller, *An introduction to probability theory and its applications,* vol. 1,2, 3rd ed., John Wiley & Sons, 1968.

W. フェラー (卜部舜一訳) , 確率論とその応用１上・下, ２上・下, 紀伊国屋書店, 1988.

20 世紀中頃の確率論の定番の教科書の一つだったようにみえる．盛りだくさんの内容．用語が少し古くなっているのと，証明の記述が時々丁寧でない点で読みづらいが，内容や証明はその後の基礎教科書の元ネタになっている気がする．影響力の大きかった教科書のようだ．

Random walk の取り扱いに詳しい点は今でも使える．

[16] K. Ito, H. P. McKean, Jr., *Diffusion processes and their sample paths,* Springer, 1996.

古典的名著とのことである．本講義の水準を超えるだけでなく，最初から読者に多くを要求する書き方で，初学者にはつらい．

1 次元 Brown 運動に関する基礎事項は非常にたくさんのことが書かれている気がする (決定版だと聞いた気もする)．

[17] 服部哲弥，ランダムウォークとくりこみ群 — 確率論から数理物理学へ —，共立出版, 2004.8.

ブラウン運動を構成する第４の方法（F. Knight の方法）の背後にあるくりこみ群の思想を確率連鎖に関して解説した教科書．ブラウン運動そのものは扱わないが，それ以外に必要な全てを詳しく説明してあるので，連続極限の構成は容易なはずである．連続極限については同書の参考文献の中の F. B. Knight の論文や B. Hambly, K. Hattori, T. Hattori の論文を参照.

[18] Y. Kasahara, *Tauberian theorems of exponential type*, J. Math. Kyoto Univ. **18-2** (1978) 209–219.

指数型の Tauber 型定理の論文．たいへん整理されて書かれた決定版.

[19] N. Kosugi, *Tauberian theorems of exponential type and its application to mutliple convolution,* J. Math. Kyoto Univ., **39** (1999) 331–346, (Errata) J. Math. Kyoto Univ., **40** (2000) 203–203,

[20] N. Kosugi, *Tauberian theorems of exponential type on limits of oscillation,* J. Math. Kyoto Univ., **39** (1999) 783–792.

[21] Y. Kasahara, N. Kosugi, *Remarks on Tauberian theorem of exponential type and Frenchel-Legendre transform,* Osaka Journal of Mathematics (2002).

[22] E. Seneta, *Regularly varying functions*, Lecture notes in mathematics **508**, Springer, 1976.

正則変動関数と Tauber 型定理の決定版の教科書．[15] の記述より丁寧なのでお勧めだが，絶版なのが残念.

[23] S. R. S. Varadhan, *Large deviations and applications,* Society for industrial and applied mathematics, 1994.

本講義の水準を越える advanced な内容．母関数の適用例として，この本の中から唯一初等的な節を引用した.

[24] D. Williams, *Probability with martingales*, Cambridge University Press, 1991.

日本の標準的教科書と同様に離散確率過程 (stochastic chain) までだが，マルチンゲールをできるだけ早く導入することに特徴を出している．テーマになっている例は，パズル的ながら Doob の任意停止定理の適用例としてとても教育的.

[25] 長田博文，名古屋大学講義ノート，1999.

[26] 楠岡成雄（服部久美子記），東京大学講義ノート，1986-10.

[27] 服部哲弥，信州大学集中講義ノート，http://www.math.nagoya-u.ac.jp/~hattori , 1998.

[28] 服部哲弥, 宇都宮大学講義ノート, http://www.math.nagoya-u.ac.jp/~hattori , 1995.