測度論的確率論　講義ノート[1]

2016 年版

加藤　賢悟[2]

[1]First version: 2015 年 1 月 8 日. This version: 平成 29 年 4 月 25 日. ちゃんと校正していないので，誤植・間違いがあると思います．何かコメントがあればメール下さい．

[2]東京大学大学院経済学研究科．〒 113-0033 東京都文京区本郷 7-3-1. E-mail: kkato@e.u-tokyo.ac.jp.

前書き

本講義ノートは，東京大学大学院経済学研究科「測度論的確率論」(通年 4 単位) のために用意されたものである．だいたい Durrett (2010) の Chapters 1-3, 5, 6, 8 (測度論，大数の法則，中心極限定理，マルチンゲール，マルコフ・チェイン，Brown 運動) をカバーしているが，詳細はかなり異なる．Durrett (2010) がカバーしていないトピックもいくつか加えた (例えば，距離空間上の確率測度の弱収束理論)．測度論の知識は仮定しないが，統計学コース向けの講義なので，中級レベルの数理統計学 (例えば，竹村 (1991) レベル) には慣れていることを仮定している．従って，具体的な分布とその計算はある程度知っているものと仮定して，講義ノートではあまり紹介していない．

Durrett (2010) とともに，“first year graduate probability” の標準的な教科書として，Resnick (1998), Williams (1991), Billinsgley (1995), Chung (2001), Breiman (1968) を挙げておく．その他に，オンラインの講義ノートではあるが，Dembo (2015) も参考になる．この中では，Resnick (1998) が一番平易だと思う．数学的な背景知識が少なくても読めるように配慮されているが，それでも確率論のスタンダードな内容はちゃんとカバーしている良書である．Williams (1991) はマルチンゲールに重点を置いた確率論の入門書である．測度論の基本的な定理 (π-λ 定理，Carathéodory の拡張定理，無限直積確率空間の構成) に関してもちゃんと証明を載せている．また，ところどころ深いことも書いてある良書である．何か副読本を，というのであれば，Resnick (1998) か Williams (1991) を勧める．その他に日本語の教科書だと舟木 (2004) がわかりやすい．

もっと高級な確率論の教科書として，Dudley (2002) と Stroock (2011) を挙げておく．Dudley (2002) は確率論の教科書としてカバーしている範囲はそれほど広くないが，抽象度が高い．前半は位相空間論，測度論，関数解析をカバーしており，各定理に関してかなり一般的な仮定の下で簡潔な証明が与えられているので，手元にあると何かと便利である．また，各 Chapter の最後に，その Chapter で扱ったトピックの歴史的な経緯を解説していて，その部分だけでも読んでみる価値はある．Stroock (2011) はハイレベルな教科書である．読みやすくはないが，他の教科書にはない専門的なこと (例えば，抽象 Wiener 空間) も色々書いてあるので面白いといえば面白い[3]．その他，各トピックに関する参考文献を講義ノートのあちこちで紹介している．適宜参照されたい．

本講義ノートは最初に測度論の一般論を解説し，そのあと確率論に入る．重点は確率論の方なので，測度論の一般論には深く立ち入らない．測度論の標準的な教科書として (ごく一部ではあるが) 吉田 (2006), Folland (1999), Dudley (2002), Cohn (2013) を挙げておく．また，基本的な関数解析も勉強しておくと，確率論の理解が深まると思う．関数解析に関しては，前掲の測度論の教科書もある程度カバーしているが，その他に，黒田 (1980), 藤田・伊藤・黒田 (1991), Reed and Simon (1980) などが参考になる．

[3]余談だが，Stroock の本は Preface が面白い．Stroock (2011) の p.xix には，“except for those who find their way into poorly stocked library of some prison camp, few copies of this book will be read from cover to cover” と書かれている．

記号と慣例

- $\mathbb{R} = (-\infty, \infty)$, $\overline{\mathbb{R}} = [-\infty, +\infty]$, $\mathbb{R}_+ = [0, \infty)$, $\mathbb{C} = \{x + \sqrt{-1}y : x, y \in \mathbb{R}\}$, $\mathbb{N} = \{1, 2, \dots\}$, $\mathbb{Z} = \{0, \pm 1, \pm 2, \dots\}$, $\mathbb{Z}_+ = \{0, 1, 2, \dots\}$, $\mathbb{Q} = \{m/n : m, n \in \mathbb{Z}, n \neq 0\}$.
- $z \in \mathbb{C}$ に対して，$\overline{z}$ を z の複素共役とする．すなわち，$z = x + \sqrt{-1}y$ $(x, y \in \mathbb{R})$ なら，$\overline{z} = x - \sqrt{-1}y$ である．また，$|z|^2 = z\overline{z} = x^2 + y^2$ とする．さらに，$\mathrm{Re}\, z, \mathrm{Im}\, z$ をそれぞれ z の実部と虚部とする．
- 集合 S に対して，2^S は S の部分集合全体を表す．
- $a, b \in \mathbb{R}$ に対して，$a \wedge b = \min\{a, b\}, a \vee b = \max\{a, b\}$ と書く．
- 関数 $f : T \to \mathbb{R}$ に対して，$\|f\|_u = \sup_{t \in T} |f(t)|$ と書く．
- 特に断らない限り，添え字の n, m, k, ℓ etc. は整数の範囲を動き，添え字の s, t etc. は実数の範囲を動くものとする．
- $N(\mu, \sigma^2)$ は平均 μ, 分散 σ^2 をもつ正規分布，$Po(\lambda)$ はパラメータ λ をもつ Poisson 分布，$U(a, b)$ は (a, b) 上の一様分布を表す．
- 位相空間の部分集合 A に対して，$A^\circ, \overline{A}, \partial A$ をそれぞれ A の内部，閉包，境界とする．
- 写像 $f : X \to Y$ と集合 $A \subset Y$ に対して，$\{f \in A\} = f^{-1}(A) = \{x \in X : f(x) \in A\}$ と書く．$Y = \overline{\mathbb{R}}$ のときは，$\{f \leq y\} = \{x \in X : f(x) \leq y\}$ とも書く．
- $A := B$ と書いたら，A を B で定義する，という意味である．
- $\overline{\mathbb{R}}$ 上の計算ルールとして，

  $$
  \begin{aligned}
  &x \pm \infty = \infty \pm x = \pm\infty \quad (x \in \mathbb{R}),\\
  &x \cdot (\pm\infty) = (\pm\infty) \cdot x = \pm\infty \quad (x > 0),\ x \cdot (\pm\infty) = (\pm\infty) \cdot x = \mp\infty \quad (x < 0),\\
  &\infty + \infty = \infty,\ -\infty - \infty = -\infty,\ (\pm\infty) \cdot 0 = 0 \cdot (\pm\infty) = 0
  \end{aligned}
  $$

  と約束しておく．ただし，$\infty - \infty$ は定義されないとする．

ランダムなコメント

- Durrett 1.2.3 などとあるのは，Durrett (2010) の Exercise の番号を指している．
- あくまで講義ノートなので，しばしばインフォーマルな記述をしている．例えば，論理記号 ($\forall, \exists$ など) や略語 (s.t., r.v. など) はフォーマルな論文や本では使うべきではないとされるが，本講義ノートではしばしば使っている．

- 2015 年からの変更は以下の通りである．1. いくつか証明を省略していた部分を補った．これにより，講義ノートはある程度自己充足的になったが，それでも Durrett や Folland などのちゃんとした確率論・測度論の教科書をもっておくべきである．講義ノートはあくまでも講義の内容 $(+\alpha)$ 程度しかカバーしていないので，深みに欠けるためである．2. いくつかのトピックを新しく加えた．3. 演習問題を加えた．演習問題は Durrett と重ならないものを中心に集めたが，重なっているものもある．Durrett と講義ノートから宿題を出す予定である．

目 次

1 測度空間 **7**

2 $\mathbb{R}$ 上の **Lebesgue-Stieltjes** 測度 **16**

3 可測関数 **24**

4 積分 **30**

5 直積測度 **41**

6 確率空間 **48**

7 独立性 **61**

8 大数の弱法則 **70**

9 **Borel-Cantelli** の補題と **Kolmogorov** の **0-1** 法則 **76**

10 大数の強法則 **85**

11 弱収束 **92**

12 **Lindeberg-Feller** の **CLT** **100**

13 特性関数 **107**

14 モーメント法と **Edgeworth** 展開 **116**

15 無限分解可能分布 **124**

16 安定分布 **133**

17 極値分布 **138**

18 多次元 **CLT** **143**

19 条件付き期待値 **148**

20 正則条件付き分布と **Kolmogorov** の拡張定理 **161**

21 マルチンゲールと **a.s.** 収束 **171**

22 マルチンゲールの L^p 収束 **178**

23 バックワード・マルチンゲール **189**

24 任意停止定理 **196**

25 マルチンゲールに関するその他の話題 **201**

26 マルコフ・チェイン **208**

27 マルコフ・チェインの基本概念 **212**

28 極限定理 **220**

29 **Brown** 運動 **227**

30 連続時間のマルチンゲール **238**

31 **Brown** 運動のマルコフ性とマルチンゲール性 **247**

32 **Skorohod** の埋め込み定理とその応用 **253**

33 距離空間上の確率測度の弱収束 **263**

34 **Prohorov** の定理と **Prohorov** 距離 **277**

35 C 空間 **290**

36 D 空間と **Skorohod** 位相 **298**

37 D 空間における弱収束 **307**

38 解析集合 **314**

39 射影定理と可測選択定理 **319**

Part I. 測度論

確率論の数学的に厳密な基礎付けはKolmogorov (1933) による．Kolmogorovは確率論を測度論に枠組みの中で展開し，以降，Kolmogorov流の確率論が標準的なものとして受け入れられている．J.L. Doobはその著書 (Doob, 1953) の中で，“Probability theory is simply a branch of measure theory, with its own special emphasis and field of application” と言い切っている．このステートメントには異論もあるかもしれないが，そう考えた方が確率論の理解が単純であるのも確かである．

いずれにせよ，確率論を勉強するためには，測度論の勉強が必要になる．本パートは測度論の基本的な結果をカバーする．次のパート以降で必要となる測度論の結果をカバーするだけなので，その扱いは簡潔に留める．

1 測度空間

測度論では，素朴な意味での面積や体積の概念を抽象的な空間 X の部分集合に対して拡張することを考える．その拡張を測度と呼ぶ．のちのち明らかになっていくように，与えられた性質と両立する測度を X のすべての部分集合に対して定義できるかどうかという問題は一般にかなり難しい．そこで，制限された部分集合族，すなわち，σ-field 上に測度を定義する．

X を空でない集合とし，$\mathcal{A} \subset 2^X$ とする．$\mathcal{A}$ が field であるとは，次の 3 条件がみたされることを言う：

(i) $X \in \mathcal{A}$.

(ii) $A, B \in \mathcal{A} \Rightarrow A \cup B \in \mathcal{A}$.

(iii) $A \in \mathcal{A} \Rightarrow A^c := X \setminus A \in \mathcal{A}$.

$\mathcal{A}$ が field なら，任意の $A, B \in \mathcal{A}$ に対して，

$$A \cap B = (A^c \cup B^c)^c \in \mathcal{A},\ A \setminus B = A \cap B^c \in \mathcal{A}$$

となる．また，$\varnothing = X \setminus X$ より，

$$\varnothing \in \mathcal{A}$$

である．さらに，$A_i \in \mathcal{A}$ $(i = 1, \ldots, n)$ なら，帰納的に，

$$\bigcup_{i=1}^{n} A_i \in \mathcal{A},\ \bigcap_{i=1}^{n} A_i \in \mathcal{A}$$

である．

$\mathcal{A}$ が σ-field であるとは，次の 2 条件がみたされることを言う：

(i) $\mathcal{A}$ が field である．

(ii) $A_n \in \mathcal{A}\ (n = 1, 2, \dots) \Rightarrow \bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$.

このとき，可算個の $A_n \in \mathcal{A}\ (n = 1, 2, \dots)$ に対して，

$$\bigcap_{n=1}^{\infty} A_n = \left(\bigcup_{n=1}^{\infty} A_n^c \right)^c \in \mathcal{A}$$

である．$\mathcal{A}$ が σ-field のとき，$(X, \mathcal{A})$ を 可測空間 (measurable space) と呼ぶ．また，$\mathcal{A}$ に属する集合を 可測集合 (measurable set) (または $\mathcal{A}$ 可測集合) と呼ぶ．

Example 1.1. $\{\varnothing, X\}$, $\{\varnothing, A, A^c, X\}$, 2^X はすべて σ-field である．

Example 1.2 (field だが σ-field でない例)．$X = (0, 1]$ とし，$\mathcal{A} = \{\varnothing\} \bigcup \{\bigcup_{i=1}^{n} (a_i, b_i] : 0 \le a_1 < b_1 \le a_2 < \cdots \le a_n < b_n \le 1, n \in \mathbb{N}\}$ とおくと，$\mathcal{A}$ は field だが，σ-field でない．$\mathcal{A}$ が field であることを確認するのは容易である．一方，$\mathcal{A}$ が σ-field でないことは，

$$\left(0, \frac{1}{2}\right] \bigcup \left(\frac{1}{2} + \frac{1}{2^2}, \frac{1}{2} + \frac{1}{2^2} + \frac{1}{2^3}\right] \bigcup \cdots = \bigcup_{n=0}^{\infty} \left(1 - \frac{1}{2^{2n}}, 1 - \frac{1}{2^{2n+1}}\right]$$

が $\mathcal{A}$ に属していないことからわかる．

Lemma 1.1. (i) I を任意の添え字集合とする (I は非可算でもよい)．$\mathcal{A}_i, i \in I$ が σ-fields なら，$\bigcap_{i \in I} \mathcal{A}_i$ も σ-field である．(ii) 任意の集合族 $\mathcal{C} \subset 2^X$ に対して，$\mathcal{C}$ を含む最小の σ-field $\sigma(\mathcal{C})$ が存在する．さらに，そのような σ-field は一意である．

$\sigma(\mathcal{C})$ を $\mathcal{C}$ によって 生成される σ-field と呼ぶ．

Proof. (i) は明らか．(ii). $\sigma(\mathcal{C}) = \bigcap \{\mathcal{A} : \mathcal{A}$ は $\mathcal{C}$ を含む σ-field$\}$ とすればよい． □

σ-fields $\mathcal{A}_i, i \in I$ に対して，$\bigcup_{i \in I} \mathcal{A}_i$ は必ずしも σ-field とはならない．そこで，$\bigvee_{i \in I} \mathcal{A}_i$ を $\mathcal{A}_i, i \in I$ を含む最小の σ-field と定義する：

$$\bigvee_{i \in I} \mathcal{A}_i := \sigma \left(\bigcup_{i \in I} \mathcal{A}_i \right)$$

$\mathcal{A}$ を σ-field とする．写像 $\mu : \mathcal{A} \to [0, \infty]$ が 測度 (measure) であるとは，次の 2 条件がみたされることを言う：

(i) $\mu(\varnothing) = 0$.

(ii) $A_n \in \mathcal{A}\ (n = 1, 2, \dots)$ が排反 (すなわち，$A_n \cap A_m = \varnothing, n \ne m$) なら，

$$\mu \left(\bigcup_{n=1}^{\infty} A_n \right) = \sum_{n=1}^{\infty} \mu(A_n).$$

($\mu(A_n) \ge 0$ より，$\sum_{n=1}^{\infty} \mu(A_n)$ は $+\infty$ を許せば必ず存在する．)

測度は $+\infty$ に値をとってもよい．(ii) がみたされていて，少なくとも 1 つの $A \in \mathcal{A}$ に対して $\mu(A) < \infty$ なら，$\mu(\varnothing) = 0$ である ((ii) において，$A_1 = A, A_n = \varnothing, n \geq 2$ とすれば，$\mu(A) \geq \mu(A) + \mu(\varnothing)$ であるから，$\mu(\varnothing) = 0$)．よって，(i) の条件は $\mu \equiv +\infty$ なる場合を排除しているだけである．特に，μ が有限値なら，測度の定義において (i) の条件は不要である．(ii) の性質を測度の 可算加法性 (countable additivity) と呼ぶ．(ii) において，$A_n = \varnothing, n \geq m+1$ とすれば，$A_1, \ldots, A_m \in \mathcal{A}$ が排反なら，$\mu(A_1 \cup \cdots \cup A_m) = \mu(A_1) + \cdots + \mu(A_m)$ となる．また，$A, B \in \mathcal{A}, A \subset B$, & $\mu(A) < \infty$ なら，$\mu(A \cup (B \setminus A)) = \mu(A) + \mu(B \setminus A)$ より，$\mu(B \setminus A) = \mu(B) - \mu(A)$ となる．$(X, \mathcal{A})$ が可測空間，μ が $\mathcal{A}$ 上の測度のとき，$(X, \mathcal{A}, \mu)$ を 測度空間 (measure space) と呼ぶ．$\mu(X) < \infty$ のとき，μ は 有限 であると言い，$A_n \in \mathcal{A}, A_n \subset A_{n+1}$ $(\forall n), \bigcup_{n=1}^{\infty} A_n = X$, & $\mu(A_n) < \infty$ $(\forall n)$ をみたす集合列 $\{A_n\}$ が存在するとき，μ は σ-finite であると言う．明らかに，有限測度は σ-finite である．

$B \in \mathcal{A}$ に対して,

$$\mathcal{A}_B := \{A \cap B : \ A \in \mathcal{A}\} = \{A : A \in \mathcal{A}, A \subset B\}$$

とおくと，$\mathcal{A}_B$ は σ-field であって，$(B, \mathcal{A}_B, \mu|_{\mathcal{A}_B})$ は測度空間である．$(B, \mathcal{A}_B, \mu|_{\mathcal{A}_B})$ を $(X, \mathcal{A}, \mu)$ の B への 制限 (restriction) と呼ぶ．誤解の恐れのない場合は，$(B, \mathcal{A}_B, \mu|_{\mathcal{A}_B})$ の代わりに，$(B, \mathcal{A}_B, \mu)$ と書く．

Example 1.3. (1 点分布と計数測度).

- $x \in X, A \subset X$ に対して,

$$\delta_x(A) := 1_A(x) := \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases}$$

と定義する．$\mu = \delta_x$ は 2^X 上の測度である．δ_x を x の 1 点分布 (point mass)，または Dirac 測度 (Dirac measure) と呼ぶ.

- $x_n \in X, c_n \geq 0$ $(n = 1, 2, \ldots, N \leq \infty)$ に対して，$\mu = \sum_{n=1}^{N} c_n \delta_{x_n}$ は 2^X 上の測度である．μ が可算加法的であることは，初等解析の次の結果から直ちに従う (証明は演習問題とする)．$c_n = 1$ $(\forall n)$ のとき，$\mu = \sum_{n=1}^{N} \delta_{x_n}$ を $\{x_n\}_{n=1}^{N}$ 上の 計数測度 (counting measure) と呼ぶ.

Lemma 1.2. $a_{n,m} \geq 0, n, m \in \mathbb{N}$ を二重添え字をもつ非負数列とすると,

$$\sum_{m=1}^{\infty} \sum_{n=1}^{\infty} a_{n,m} = \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} a_{n,m} = \sup \left\{ \sum_{m=1}^{M} \sum_{n=1}^{N} a_{n,m} : N, M \in \mathbb{N} \right\}.$$

測度の基本的な性質．集合列 $A_n, n \in \mathbb{N}$ に対して，$A_n \uparrow A$ は，$A_n \subset A_{n+1}$ $(\forall n \in \mathbb{N})$ & $\bigcup_{n=1}^{\infty} A_n = A$ を意味する．$A_n \downarrow A$ は，$A_n \supset A_{n+1}$ $(\forall n \in \mathbb{N})$ & $\bigcap_{n=1}^{\infty} A_n = A$ を意味する.

Lemma 1.3. $(X, \mathcal{A}, \mu)$ を測度空間とする．(i) $A \subset B$ なら，$\mu(A) \leq \mu(B)$. (ii) $\mu(\bigcup_{n=1}^{\infty} A_n) \leq \sum_{n=1}^{\infty} \mu(A_n)$. (iii) $A_n \uparrow A$ なら，$\mu(A_n) \uparrow \mu(A)$. (iv) $A_n \downarrow A$ & $\mu(A_1) < \infty$ なら，$\mu(A_n) \downarrow \mu(A)$.

Proof. (i). $C = B \setminus A$ とおくと，A と C は排反であって，$B = A \cup C$ であるから，$\mu(B) = \mu(A) + \mu(C) \geq \mu(A)$.

(ii). $A = \bigcup_{n=1}^{\infty} A_n, B_1 = A_1, B_m = A_m \setminus \bigcup_{n=1}^{m-1} A_n, m \geq 2$ とおくと，$B_m, m \in \mathbb{N}$ は排反であって，$A = \bigcup_{m=1}^{\infty} B_m$．よって，$\mu(A) = \sum_{m=1}^{\infty} \mu(B_m) \leq \sum_{m=1}^{\infty} \mu(A_m)$.

(iii). $B_1 = A_1, B_n = A_n \setminus A_{n-1}, n \geq 2$ とおくと，$B_n, n \in \mathbb{N}$ は排反であって，$A = \bigcup_{n=1}^{\infty} B_n, A_m = \bigcup_{n=1}^{m} B_n, m \in \mathbb{N}$ であるから，$\mu(A) = \sum_{n=1}^{\infty} \mu(B_n) = \lim_m \sum_{n=1}^{m} \mu(B_n) = \lim_m \mu(A_m)$.

(iv). $A_n \downarrow A$ より，$A_1 \setminus A_n \uparrow A_1 \setminus A$ であるから，(iii) より $\mu(A_1 \setminus A_n) \uparrow \mu(A_1 \setminus A)$. よって，$A \subset A_n \subset A_1$ より，$\mu(A_n) = \mu(A_1) - \mu(A_1 \setminus A_n) \downarrow \mu(A_1) - \mu(A_1 \setminus A) = \mu(A)$. □

Remark 1.1. (iv) において，$\mu(A_1) < \infty$ という仮定は，$\infty - \infty$ が証明中に起こらないことを保証している．この仮定は本質的である．例えば，$X = \mathbb{N} = \{1, 2, \dots\}$ とし，$\mu = \sum_{n=1}^{\infty} \delta_n$ ($\mathbb{N}$ 上の計数測度) を考えると，$A_n := \mathbb{N} \setminus \{1, \dots, n\} \downarrow \varnothing =: A$ だが，$\mu(A_n) = \infty$ $(\forall n)$, $\mu(A) = 0$ となる．

1 点分布や計数測度より複雑な測度の構成を考える．次の定理は基本的である．

Theorem 1.1 (Carathéodory の拡張定理)**.** (i) $\mathcal{C} \subset 2^X$ を field とし，$\mu : \mathcal{C} \to [0, \infty]$ は次の性質 (a) と (b) をみたすとする：(a) $\mu(\varnothing) = 0$. (b) $A_n \in \mathcal{C}, n \in \mathbb{N}$ が排反であって，かつ $\bigcup_{n=1}^{\infty} A_n \in \mathcal{C}$ なら，$\mu(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mu(A_n)$ となる．このとき，μ は $\sigma(\mathcal{C})$ 上の測度に拡張できる．(ii) さらに，μ が $\mathcal{C}$ 上で σ-finite であれば，すなわち，$A_n \uparrow X, A_n \in \mathcal{C}$, & $\mu(A_n) < \infty$ $(\forall n \in \mathbb{N})$ をみたす集合列 $\{A_n\}$ が存在すれば，(i) の拡張は一意である．

Remark 1.2 (Carathéodory の拡張定理の重要性)**.** 一般に σ-field はかなり複雑であり，後に登場する $\mathbb{R}$ の Borel σ-field を例にとれば，そこに属していない $\mathbb{R}$ の集合 (あるにはある) を具体的に構成しようとすると結構大変である．従って，普通の関数とは異なり，各可測集合に対して値を指定することで測度を定義するというのは，1 点分布や計数測度のような単純な測度を除いて，一般には難しい．その代わり，比較的 “小さい” 集合族上で測度の候補を特徴づけて，拡張定理により測度を定義するということを行う．

(i) の証明はスケッチのみ与える．詳細は Durrett (2010, Appendix) を参照せよ．

Proof Sketch of Theorem 1.1 (i). 任意の $E \subset X$ に対して，μ の 外測度 (outer measure) を

$$\mu^*(E) = \inf \left\{ \sum_{n=1}^{\infty} \mu(A_n) : E \subset \bigcup_{n=1}^{\infty} A_n, A_n \in \mathcal{C} \ (n = 1, 2, \dots) \right\}$$

と定義する．$\mu^* = \mu$ on $\mathcal{C}$ を示すのは難しくない．従って，あとは μ^* が $\sigma(\mathcal{C})$ 上の測度であることを示せばよい．

μ^* が 2^X 上で単調性 ($E \subset F \Rightarrow \mu^*(E) \leq \mu^*(F)$) と可算劣加法性 ($\mu^*(\bigcup_n E_n) \leq \sum_n \mu^*(E_n)$) をみたすことを示すのは難しくない．$\mu^*$ が $\sigma(\mathcal{C})$ 上で可算加法的であることを示すために，まず μ^* が任意の $A \in \sigma(\mathcal{C})$ に対して，

$$\mu^*(E) = \mu^*(E \cap A) + \mu^*(E \setminus A), \quad \forall E \subset X \tag{$*$}$$

をみたすことを示す．これはまず $A \in \mathcal{C}$ に対して $(*)$ が成り立つことを示し，次に $(*)$ をみたす A の全体が σ-field であることを示せばよい．最後にこの性質と，すでに示した単調性と可算劣加法性より，μ^* が $\sigma(\mathcal{C})$ 上で可算加法性をみたすことが示される． □

(ii) の証明に π-λ 定理 (**重要!**) を用いる．まず定義をおさらいする．$\mathcal{P} \subset 2^X$ が π システムであるとは，$A, B \in \mathcal{P}$ なら，$A \cap B \in \mathcal{P}$ となることを言う．言葉で述べると，π システムとは有限個の積に関して閉じている部分集合族のことである．$\mathcal{L} \subset 2^X$ が λ システム であるとは，次の条件 (i)–(iii) をみたすことを言う：(i) $X \in \mathcal{L}$, (ii) $A, B \in \mathcal{L}, A \subset B$ なら，$B \setminus A \in \mathcal{L}$, (iii) $A_n \uparrow A, A_n \in \mathcal{L}$ なら，$A \in \mathcal{L}$.

Theorem 1.2 (π-λ 定理)**.** $\mathcal{P}$ を π システムとし，$\mathcal{L}$ を $\mathcal{P}$ を含む λ システムとする：$\mathcal{L} \supset \mathcal{P}$. このとき，$\sigma(\mathcal{P}) \subset \mathcal{L}$ である．

Proof. λ システムは σ-field と同様に積に関して閉じている．すなわち，$\mathcal{L}_i, i \in I$ が λ システムなら，$\bigcap_{i \in I} \mathcal{L}_i$ も λ システムである．そこで，$\ell(\mathcal{P})$ を $\mathcal{P}$ を含む最小の λ システムとすれば，$\mathcal{L} \supset \ell(\mathcal{P})$ である．ここで，$\ell(\mathcal{P})$ が σ-field であることを示す．これから定理の結論が従う．

$X \in \ell(\mathcal{P})$ と $A \in \ell(\mathcal{P}) \Rightarrow A^c \in \ell(\mathcal{P})$ は λ システムの定義からよい．さらに $\ell(\mathcal{P})$ が π システムであれば，$A, B \in \ell(\mathcal{P}) \Rightarrow A \cup B = (A^c \cap B^c)^c \in \ell(\mathcal{P})$ となるから，$A_n \in \ell(\mathcal{P})$ $(n = 1, 2, \dots)$ に対して，$\ell(\mathcal{P}) \ni \bigcup_{m=1}^n A_m \uparrow \bigcup_{m=1}^\infty A_m \in \ell(\mathcal{P})$ となる．よって，$\ell(\mathcal{P})$ が π システムであることを示せばよい．$A \in \mathcal{P}$ を任意に固定し，$\mathcal{D} = \{B \in \ell(\mathcal{P}) : A \cap B \in \ell(\mathcal{P})\}$ とおくと，$\mathcal{D}$ は $\mathcal{P}$ を含む λ システムであるから，$\mathcal{D} = \ell(\mathcal{P})$ を得る．$A \in \mathcal{P}$ は任意だったから，$A \in \mathcal{P}, B \in \ell(\mathcal{P}) \Rightarrow A \cap B \in \ell(\mathcal{P})$ が示された．そこで，今度は $B \in \ell(\mathcal{P})$ を任意に固定して，$\mathcal{E} = \{A \in \ell(\mathcal{P}) : A \cap B \in \ell(\mathcal{P})\}$ とおくと，上の結果より $\mathcal{E}$ は $\mathcal{P}$ を含む λ システムである．ゆえに $A, B \in \ell(\mathcal{P}) \Rightarrow A \cap B \in \ell(\mathcal{P})$ が示された． □

Proof of Theorem 1.1 (ii). ν をもう 1 つの拡張とする．$A \in \mathcal{C}$ を $\mu(A) < \infty$ とし ($\mu = \nu$ on $\mathcal{C}$ より，$\nu(A) < \infty$ でもある)，

$$\mathcal{L} = \{B \in \sigma(\mathcal{C}) : \mu(A \cap B) = \nu(A \cap B)\}$$

とおいて，$\mathcal{L}$ が $\mathcal{C}$ を含む λ システムであることを確認する．まず μ と ν は $\mathcal{C}$ 上で一致しているので，$\mathcal{C} \subset \mathcal{L}$ はよい．次に，$\mathcal{L}$ が λ システムであることは，$X \in \mathcal{L}$ は明らかであり，$B, C \in \mathcal{L}, B \subset C$ なら，$\mu(A \cap (C \setminus B)) = \mu(A \cap C) - \mu(A \cap B) = \nu(A \cap C) - \nu(A \cap B) = \nu(A \cap (C \setminus B))$

であるから，$C \setminus B \in \mathcal{L}$ となる (ここで，$\mu(A) = \nu(A) < \infty$ より，$\infty - \infty$ は起こらない). あとは，$B_n \uparrow B, B_n \in \mathcal{L}$ なら，$B \in \mathcal{L}$ となることを確認すればよいが，これは測度の性質から直ちに従う．以上より，$\mathcal{L}$ は $\mathcal{C}$ を含む λ システムであることが確認できた．一方，field は π システムであるから，π-λ 定理より，$\mathcal{L} = \sigma(\mathcal{C})$ を得る．すなわち，任意の $A \in \mathcal{C}$ s.t. $\mu(A) < \infty$ と $B \in \sigma(\mathcal{C})$ に対して，$\mu(A \cap B) = \nu(A \cap B)$ となる．ここで，各 $B \in \sigma(\mathcal{C})$ に対して，$A_n \cap B \uparrow B$ であるから，$\mu(B) = \lim_{n\to\infty} \mu(A_n \cap B) = \lim_{n\to\infty} \nu(A_n \cap B) = \nu(B)$． □

拡張定理の (ii) において，$\sigma(\mathcal{C})$ 上に拡張された測度 μ は，$\mathcal{C}$ 上での値から一意に決まるわけであるが，さらに次の定理が成り立つ．$A, B \subset X$ に対して，$A \Delta B := (A \setminus B) \cup (B \setminus A)$ を A と B の 対称差 (symmetric difference) と呼ぶ．

Theorem 1.3 (近似定理)**.** $\mathcal{C} \subset 2^X$ を field とし，μ を $(X, \sigma(\mathcal{C}))$ 上の測度であって，$\mathcal{C}$ 上で σ-finite とする．このとき，任意の $A \in \sigma(\mathcal{C})$ に対して，次をみたす集合列 $B_n \in \mathcal{C}$ $(n = 1, 2, \dots)$ が存在する：$\lim_n \mu(A \Delta B_n) = 0$.

Proof. 拡張定理とは独立に証明してみる．μ は有限測度と仮定してよい (なぜか)：$\mu(X) < \infty$. $\mathcal{A} = \sigma(\mathcal{C})$ とおく．このとき，$d(A, B) = \mu(A \Delta B), A, B \in \mathcal{A}$ とおくと，d は三角不等式をみたす：$d(A, C) \le d(A, B) + d(B, C)$ ($\because A \setminus C = A \cap C^c = (A \cap C^c \cap B) \cup (A \cap C^c \cap B^c) \subset (C^c \cap B) \cup (A \cap B^c) = (B \setminus C) \cup (A \setminus B)$ より，$A \Delta C \subset (A \Delta B) \cup (B \Delta C)$). ここで，

$$\mathcal{D} = \{A \in \mathcal{A} : \exists B_n \in \mathcal{C} \text{ s.t. } \lim_n \mu(A \Delta B_n) = 0\}$$

とおくと，$\mathcal{C} \subset \mathcal{D}$ である．$\mathcal{D}$ が σ-field であることを示せば，$\mathcal{D} \supset \sigma(\mathcal{C}) = \mathcal{A}$ となり，定理の結論が従う．まず，$d(A^c, B^c) = d(A, B)$ より，$A \in \mathcal{D} \Rightarrow A^c \in \mathcal{D}$ はよい．次に，$A_m \in \mathcal{D}$ $(m = 1, 2, \dots)$ とすると，各 A_m に対して，$\lim_n d(A_m, B_{m,n}) = 0$ となる集合列 $\{B_{m,n}\}_{n=1}^{\infty} \subset \mathcal{C}$ が存在する．$B_n^N = \bigcup_{m=1}^N B_{m,n}, A^N = \bigcup_{m=1}^N A_m$ とおくと，

$$\begin{aligned} d(B_n^N, A^N) &\le d(\cup_{m=1}^N B_{m,n}, A_1 \cup (\cup_{m=2}^N B_{m,n})) \\ &\quad + d(A_1 \cup (\cup_{m=2}^N B_{m,n}), A_1 \cup A_2 \cup (\cup_{m=3}^N B_{m,n})) \\ &\quad + \cdots + d(\cup_{m=1}^{N-1} A_m \cup B_{N,n}, \cup_{m=1}^N A_m) \\ &\le \sum_{m=1}^N d(B_{m,n}, A_m). \end{aligned}$$

ここで，$(A \cup B)\Delta(B \cup C) = \{(A \cup B) \cap (B^c \cap C^c)\} \cup \{(B \cup C) \cap (A^c \cap B^c)\} = (A \cap B^c \cap C^c) \cup (C \cap A^c \cap B^c) \subset A \Delta C$ を使った．ゆえに $\lim_n d(B_n^N, A^N) = 0$ であるから，$A^N \in \mathcal{D}$ であり，$\lim_N d(A^N, \cup_{m=1}^{\infty} A_m) = 0$ より，$\cup_{m=1}^{\infty} A_m \in \mathcal{D}$ を得る． □

測度空間の完備化

測度空間 $(X, \mathcal{A}, \mu)$ が 完備 (complete) であるとは，

$$S \subset N \in \mathcal{A}, \mu(N) = 0 \Rightarrow S \in \mathcal{A}$$

となることを言う．言葉で述べると，測度 0 の可測集合の部分集合が常に可測になるとき，測度空間を完備と言う．

じつはもともとの測度空間が完備でなくても，完備な拡張が常に存在する．$(X, \mathcal{A}, \mu)$ を (完備とは限らない) 測度空間とし，$\mathcal{N} = \{S \subset X : \exists N \in \mathcal{A} \text{ s.t. } S \subset N, \mu(N) = 0\}, \mathcal{A}^\mu := \sigma(\mathcal{A} \cup \mathcal{N})$ とおく．

Lemma 1.4. $B \in \mathcal{A}^\mu \Leftrightarrow \exists A \in \mathcal{A}, \exists S \in \mathcal{N} \text{ s.t. } B = A \cup S.$

Proof. $\mathcal{A}^* := \{A \cup S : A \in \mathcal{A}, S \in \mathcal{N}\}$ とおく．$\mathcal{A}^* \subset \mathcal{A}^\mu$ はよい．逆の包含関係を示す．そのために，$\mathcal{A}^*$ が σ-field であることを示せば，$\mathcal{A}, \mathcal{N} \subset \mathcal{A}^*$ より，$\mathcal{A}^\mu \subset \mathcal{A}^*$ が従う．$X \in \mathcal{A}^*$ は明らか．次に，$A \in \mathcal{A}, S \in \mathcal{N} \Rightarrow (A \cup S)^c \in \mathcal{A}^*$ を示す．$\mathcal{N}$ の定義より，$\exists N \in \mathcal{A}$ s.t. $S \subset N, \mu(N) = 0$ であって，$S^c = N^c \cup (N \cap S^c)$ であるから，

$$(A \cup S)^c = A^c \cap S^c = \underbrace{(A^c \cap N^c)}_{\in \mathcal{A}} \cup \underbrace{(N \setminus (A \cup S))}_{\subset N} \in \mathcal{A}^*.$$

あとは，

$$A_n \in \mathcal{A}, S_n \in \mathcal{N} \ (n = 1, 2, \dots) \Rightarrow \bigcup_{n=1}^{\infty} (A_n \cup S_n) \in \mathcal{A}^*$$

を示せばよいが，これは難しくない． □

写像 $\bar{\mu} : \mathcal{A}^\mu \to [0, \infty]$ を，

$$\bar{\mu}(B) := \mu(A), \ B = A \cup S, A \in \mathcal{A}, S \in \mathcal{N}$$

と定義する．

Theorem 1.4. $\bar{\mu}$ は $\mathcal{A}^\mu$ 上の well-defined な測度であって，$\mathcal{A}$ 上で μ に一致する．

Proof. $\bar{\mu}$ が well-defined なことだけ示す．$B = A_1 \cup S_1 = A_2 \cup S_2, A_1, A_2 \in \mathcal{A}, S_1, S_2 \in \mathcal{N}$ に対して，$A_1^c \cap S_1 \supset A_1^c \cap A_2 = A_2 \setminus A_1$ であるから，$A_2 \setminus A_1 \in \mathcal{N}$．同様に，$A_1 \setminus A_2 \in \mathcal{N}$．よって，$\mu(A_1) = \mu(A_2)$． □

$(X, \mathcal{A}^\mu, \bar{\mu})$ を $(X, \mathcal{A}, \mu)$ の 完備化 (completion) と呼ぶ．$\mathcal{A}^\mu$ を $\mathcal{A}$ の μ に関する完備化とも言う．$\mathcal{A}^\mu$ は μ がもともとの定義域を超えて問題なく拡張できる限界だと思えばよい [4]．

演習問題

Exercise 1.1. Lemma 1.2 を示せ．

[4] とはいえこのステートメントは正確ではない．例えば，Bogachev (2007, Section 1.12(v)) を参照せよ．また，次節の Lebesgue 非可測集合に関する議論も参照せよ．

Exercise 1.2. 任意の $\mathcal{A}_0 \subset 2^X$ に対して，$\mathcal{A}_0' = \{\bigcap_{j=1}^m B_j : B_j \text{ or } B_j^c \in \mathcal{A}_0 \ (j = 1, \dots, m), m \in \mathbb{N}\}$ とおくと，

$$\mathcal{A} = \left\{ \bigcup_{j=1}^m A_j : A_1, \dots, A_m \in \mathcal{A}_0', m \in \mathbb{N} \right\}$$

は $\mathcal{A}_0$ を含む最小の field であることを示せ．$\mathcal{A}_0$ が可算なら $\mathcal{A}$ も可算であることを示せ．

Exercise 1.3. $\mathcal{A}$ を σ-field であって，無限集合とする．$\mathcal{A}$ は可算になりうるか．

Exercise 1.4. $\mathcal{C}$ を X の部分集合族とする．このとき，各 $B \in \sigma(\mathcal{C})$ に対して，$B \in \sigma(\mathcal{C}_B)$ となる可算な集合族 $\mathcal{C}_B \subset \mathcal{C}$ が存在することを示せ．

Exercise 1.5. $\mathcal{A}$ を field とし，$\mu : \mathcal{A} \to \mathbb{R}_+$ を有限加法的 $(A, B \in \mathcal{A}, A \cap B = \varnothing \Rightarrow \mu(A \cup B) = \mu(A) + \mu(B))$ とする．このとき，μ が $\sigma(\mathcal{A})$ 上の測度に拡張できるためには，

$$A_n \downarrow \varnothing,\ A_n \in \mathcal{A} \Rightarrow \lim_{n \to \infty} \mu(A_n) = 0$$

が成り立つことが必要十分である．このことを示せ．

Exercise 1.6. $\mathcal{A}_1 \subset \mathcal{A}_2 \subset \cdots$ を σ-fields とする．

(a) $\bigcup_n \mathcal{A}_n$ が field であることを示せ．

(b) $\bigcup_n \mathcal{A}_n$ が σ-field とならないような例を構成せよ．

Exercise 1.7. $\mathcal{A}$ を σ-field とし，$D \notin \mathcal{A}$ とする．このとき，

$$\mathcal{A} \vee \sigma(\{D\}) = \{(A \cap D) \cup (B \cap D^c) : A, B \in \mathcal{A}\}$$

を示せ．

Exercise 1.8. $X = \mathbb{R},\ \mathcal{A} = \{A \subset \mathbb{R} : A \text{ or } A^c \text{ は可算 }\}$,

$$\mu(A) = \begin{cases} 0 & A \text{ が可算のとき} \\ 1 & A^c \text{ が可算のとき} \end{cases}$$

と定義する．このとき，$(X, \mathcal{A}, \mu)$ が測度空間になることを示せ．

Exercise 1.9. $\mathcal{A}$ を集合 $A \subset \{1, 2, \dots\}$ のうち，極限

$$d(A) := \lim_{n \to \infty} \frac{\mathrm{Card}(A \cap \{1, \dots, n\})}{n}$$

が存在するものを集めた集合族とする．

(a) d が $\mathcal{A}$ 上で可算加法的でないことを示せ.

(b) $\mathcal{A}$ は field でないことを示せ.

Exercise 1.10. Carathéodory の拡張定理において，μ が $\mathcal{C}$ 上で σ-finite でなければ，拡張は一意とは限らない．そのような例を構成せよ.

Exercise 1.11. Theorem 1.3 の証明を μ が $\mathcal{C}$ 上で σ-finite な場合にまで拡張せよ.

Exercise 1.12. μ を可測空間 $(X, \mathcal{A})$ 上の有限測度とし，μ^* を μ の外測度とする．すなわち，任意の $E \subset X$ に対して，$\mu^*(E) = \inf\{\mu(A) : E \subset A, A \in \mathcal{A}\}$ である．また，μ の内測度 (inner measure) μ_* を $\mu_*(E) = \mu(X) - \mu^*(X \setminus E), E \subset X$ と定義する．$\mathcal{A}^\mu$ を $\mathcal{A}$ の μ による完備化とすると，

$$E \in \mathcal{A}^\mu \Leftrightarrow \mu^*(E) = \mu_*(E)$$

を示せ.

2 $\mathbb{R}$ 上の Lebesgue-Stieltjes 測度

Carathéodory の拡張定理の重要な応用として，$\mathbb{R}$ 上の Lebesgue-Stieltjes 測度を構成する．$F : \mathbb{R} \to \mathbb{R}$ を右連続かつ非減少な関数とする．$\mathbb{R}$ の部分集合族 $\mathcal{D}$ を

$$\mathcal{D} = \{(a, b] : -\infty \le a < b < \infty\} \cup \{(a, \infty) : -\infty \le a < \infty\} \cup \{\varnothing\}$$

とおいて，$\mathcal{D}$ 上の非負関数 $\mu : \mathcal{D} \to [0, \infty]$ を

$$\begin{aligned}
&\mu((a, b]) := F(b) - F(a),\ -\infty \le a < b < \infty,\\
&\mu((a, \infty)) = F(\infty) - F(a),\ -\infty \le a < \infty,\\
&\mu(\varnothing) = 0
\end{aligned}$$

と定義する．ただし，$F(\infty) := \lim_{x\to\infty} F(x), F(-\infty) := \lim_{x\to-\infty} F(x)$ である．F は非減少なので，$F(\infty)$ と $F(-\infty)$ は $[-\infty, \infty]$ の範囲で存在する．

Theorem 2.1. μ は $\mathcal{B} := \sigma(\mathcal{D})$ 上の測度に一意に拡張できる．

Theorem 2.1 から決まる測度を，F に対応する Lebesgue-Stieltjes 測度 と呼ぶ．特に，$F(x) = x$ のとき，対応する Lebesgue-Stieltjes 測度を Lebesgue 測度 と呼び，それを λ と書くことにする．すなわち， Lebesgue 測度 λ とは，

$$\lambda((a, b]) = b - a,\ -\infty < a < b < \infty$$

をみたす $\mathcal{B}$ 上の測度である (そのような測度は一意に決まる)．

F が右連続かつ非減少なことは必要条件でもある．実際，μ が測度に拡張できるなら，μ は非負なので，F は非減少である．一方，$x_n \downarrow x$ に対して，

$$F(x_n) - F(a) = \mu((a, x_n]) \downarrow \mu((a, x]) = F(x) - F(a)$$

であるから，F は右連続でもある．

F を右連続かつ非減少とし，μ を対応する Lebesgue-Stieltjes 測度とする．F は非減少なので，各点 $x \in \mathbb{R}$ で左極限 $F(x-) = \lim_{y\uparrow x} F(y)$ が存在し，$\mu((a, x)) = F(x-) - F(a)$ である．また，F は右連続なので，$x \in \mathbb{R}$ が F の不連続点になるのは，$\mu(\{x\}) = F(x) - F(x-) > 0$ のとき，またそのときに限る．一般に $\mathbb{R}$ 上の非減少関数は不連続点を高々可算個しかもたないから (cf. Exercise 2.1)，$\mu(\{x\}) > 0$ となる x は高々可算個しかない．F が連続なら (e.g. Lebesgue 測度)，すべての $x \in \mathbb{R}$ に対して，$\mu(\{x\}) = 0$ である．

Proof of Theorem 2.1. $\mathcal{D}$ は field でないが，

$$\overline{\mathcal{D}} := \left\{ \bigcup_{m=1}^{n} A_m :\ A_1, \dots, A_n \in \mathcal{D} \text{ は排反}, n \in \mathbb{N} \right\}$$

は field であって，$\sigma(\overline{\mathcal{D}}) = \sigma(\mathcal{D}) = \mathcal{B}$ が成り立つ．そこで，μ を $\overline{\mathcal{D}}$ 上に

$$\mu\left(\bigcup_{m=1}^{n} A_m\right) := \sum_{m=1}^{n} \mu(A_m), \quad A_1, \ldots, A_n \in \mathcal{D} \text{ は排反}$$

と拡張して，Carathéodory の拡張定理を使う．残りの証明をいくつかのステップに分割する．以下の証明では，記号の簡単のため，$b = \infty$ のときは，$(a, b] = (a, \infty)$ と理解する．

Step 1. $\overline{\mathcal{D}}$ 上に拡張された μ が well-defined なことを示す．まず $-\infty \le a < b \le \infty$ に対して，$(a, b] = \bigcup_{m=1}^{n} (a_m, b_m]$ (排反な和) と表せるとき，一般性を失うことなく，$a = a_1 < b_1 = a_2 < \cdots = a_n < b_n = b$ と仮定してよい．このとき，$\mu((a, b]) = F(b) - F(a) = \sum_{m=1}^{n} \{F(b_m) - F(a_m)\} = \sum_{m=1}^{n} \mu((a_m, b_m])$. 次に，$A_m \in \mathcal{D}$ $(1 \le m \le n), B_k \in \mathcal{D}$ $(1 \le k \le \ell), \bigcup_{m=1}^{n} A_m = \bigcup_{k=1}^{\ell} B_k$ (両辺は排反な和) とすると，$A_m \cap B_k$ $(1 \le m \le n, 1 \le k \le \ell)$ は排反な区間であって (そのうちの何個かは $\varnothing$ でもよい)，$A_m = \bigcup_{k=1}^{\ell} (A_m \cap B_k), B_k = \bigcup_{m=1}^{n} (A_m \cap B_k)$ である．従って，

$$\begin{aligned}
\mu\left(\bigcup_{m=1}^{n} A_m\right) &= \sum_{m=1}^{n} \mu(A_m) = \sum_{m=1}^{n} \sum_{k=1}^{\ell} \mu(A_m \cap B_k) \\
&= \sum_{k=1}^{\ell} \sum_{m=1}^{n} \mu(A_n \cap B_k) = \sum_{k=1}^{\ell} \mu(B_k) = \mu\left(\bigcup_{k=1}^{\ell} B_k\right).
\end{aligned}$$

Step 2. $A, B \in \overline{\mathcal{D}}, A \cap B = \varnothing$ なら，$\mu(A \cup B) = \mu(A) + \mu(B)$ となることを示す．$A = \bigcup_{m=1}^{n} A_m, A_m \in \mathcal{D}$ $(1 \le m \le n), B = \bigcup_{k=1}^{\ell} B_k, B_k \in \mathcal{D}$ $(1 \le k \le \ell)$ と排反な和で表せるとき，$A_m \cap B_k = \varnothing$ $(1 \le m \le n, 1 \le k \le \ell)$ であるから，$\mu(A \cup B) = \sum_m \mu(A_m) + \sum_k \mu(B_k) = \mu(A) + \mu(B)$.

Step 3. $-\infty < a < b < \infty$ に対して，$(a, b] = \bigcup_{n=1}^{\infty} (a_n, b_n]$ (排反な和) と表せるとする．$\mu((a, b]) = \sum_{n=1}^{\infty} \mu((a_n, b_n])$ を示す．まず，$(a, b] \supset \bigcup_{m=1}^{n} (a_m, b_m]$ であるから，Step 2 の結果を 2 回使うと，$\mu((a, b]) \ge \mu(\bigcup_{m=1}^{n} (a_m, b_m]) = \sum_{m=1}^{n} \mu((a_m, b_m])$. $n \to \infty$ として，$\mu((a, b]) \ge \sum_{m=1}^{\infty} \mu((a_m, b_m])$. 逆の不等式を示す．$\varepsilon > 0$ を任意に固定する．F の右連続性より，$\exists \delta \in (0, b - a)$ s.t. $F(a + \delta) < F(a) + \varepsilon$. また，各 $m \in \mathbb{N}$ に対して，$\exists \eta_m > 0$ s.t. $F(b_m + \eta_m) < F(b_m) + \varepsilon 2^{-m}$. このとき，$[a + \delta, b] \subset \bigcup_{m=1}^{\infty} (a_m, b_m + \eta_m)$ であるから，Heine-Borel の定理より，$\exists n \in \mathbb{N}$ s.t. $(a + \delta, b] \subset [a + \delta, b] \subset \bigcup_{m=1}^{n} (a_m, b_m + \eta_m) \subset \bigcup_{m=1}^{n} (a_m, b_m + \eta_m]$. 従って，

$$\begin{aligned}
\mu((a, b]) &= F(b) - F(a) < F(b) - F(a + \delta) + \varepsilon = \mu((a + \delta, b]) + \varepsilon \\
&\le \mu\left(\bigcup_{m=1}^{n} (a_m, b_m + \eta_m]\right) + \varepsilon = \sum_{m=1}^{n} \{F(b_m + \eta_m) - F(a_m)\} + \varepsilon \\
&\le \sum_{m=1}^{n} \{F(b_m) - F(a_m)\} + 2\varepsilon \le \sum_{m=1}^{\infty} \mu((a_m, b_m]) + 2\varepsilon.
\end{aligned}$$

$\varepsilon > 0$ は任意だったから，$\mu((a,b]) \leq \sum_{m=1}^{\infty} \mu((a_m, b_m])$ を得る．

$a = -\infty$ or $b = \infty$ の場合，$\mu((a,b]) \geq \sum_{m=1}^{\infty} \mu((a_m, b_m])$ は同じ証明でよい．逆の不等式を示す．$a > -\infty$ かつ $b = \infty$ の場合，$\mu((a,k]) = \sum_{m=1}^{\infty} \mu((a_m, b_m] \cap (a,k]) \leq \sum_{m=1}^{\infty} \mu((a_m, b_m])$ であるから，$k \to \infty$ として，$\mu((a,\infty)) \leq \sum_{m=1}^{\infty} \mu((a_m, b_m])$ を得る．残りの場合も同様である．

Step 4. μ が $\overline{\mathcal{D}}$ 上で可算加法的なことを示す．$A_n \in \overline{\mathcal{D}}, n \in \mathbb{N}$ を排反かつ $A := \bigcup_{n=1}^{\infty} A_n \in \overline{\mathcal{D}}$ とする．$A, A_n \in \overline{\mathcal{D}}$ より，A, A_n はそれぞれ，$A = \bigcup_{k=1}^{\ell} B_k, B_k \in \mathcal{D}$ $(1 \leq k \leq \ell), A_n = \bigcup_{m=1}^{r_n} A_{n,m}, A_{n,m} \in \mathcal{D}$ $(1 \leq m \leq r_n)$ と排反な和で表せる．このとき，

$$B_k = \bigcup_{n=1}^{\infty} \bigcup_{m=1}^{r_n} \underbrace{(A_{n,m} \cap B_k)}_{\text{排反な区間}}, \quad A_n = \bigcup_{k=1}^{\ell} \bigcup_{m=1}^{r_n} \underbrace{(A_{n,m} \cap B_k)}_{\text{排反な区間}}$$

と表せるから，

$$\begin{aligned} \mu(A) &= \sum_{k=1}^{\ell} \mu(B_k) = \sum_{k=1}^{\ell} \sum_{n=1}^{\infty} \sum_{m=1}^{r_n} \mu(A_{n,m} \cap B_k) \quad (\because \text{ Step 3}) \\ &= \sum_{n=1}^{\infty} \underbrace{\sum_{k=1}^{\ell} \sum_{m=1}^{r_n} \mu(A_{n,m} \cap B_k)}_{=\mu(A_n)} . \quad (\because \text{ Lemma 1.2}) \end{aligned}$$

μ は $\overline{\mathcal{D}}$ 上で明らかに σ-finite なので，Carathéodory の拡張定理より定理の結論を得る． □

Theorem 1.3 より，Lebesgue-Stieltjes 測度 μ に対して，$A \in \mathcal{B}$ なら，任意の $\varepsilon > 0$ に対して，排反な区間 $D_1, \ldots, D_n$ が存在して，$\mu(A \Delta (\bigcup_{m=1}^{n} D_m)) < \varepsilon$ となる．μ が Lebesgue 測度のとき $(\mu = \lambda)$，$\lambda(A) < \infty$ なら，近似する区間 $D_1, \ldots, D_n$ は有界である．さらに 1 点集合の Lebesgue 測度は 0 なので，$D_1, \ldots, D_n$ は開区間にとれる．以上の結果を次の補題にまとめておく．

Lemma 2.1. λ を $(\mathbb{R}, \mathcal{B})$ 上の Lebesgue 測度とし，$A \in \mathcal{B}, \lambda(A) < \infty$ とする．このとき，任意の $\varepsilon > 0$ に対して，排反な有界開区間 $D_1, \ldots, D_n$ が存在して，$\lambda(A \Delta (\bigcup_{m=1}^{n} D_m)) < \varepsilon$ となる．

位相空間 X に対して，開集合全体を含む最小の σ-field を，X の Borel σ-field と呼ぶ．X の Borel σ-field に属する集合のことを，X の Borel 集合 と呼ぶ．また，Borel σ-field 上に定義された測度のことを Borel 測度 と呼ぶ．

Lemma 2.2. $\mathcal{B}$ は $\mathbb{R}$ の Borel σ-field に一致する．

Proof. $\mathcal{C}$ を $\mathbb{R}$ の Borel σ-field とする．$(a,b) = \bigcup_{n=1}^{\infty} (a, b - 1/n]$ より，$(a,b) \in \mathcal{B}$ である．$\mathbb{R}$ の任意の開集合は開区間の可算和で表せるので，$\mathcal{B}$ に属する．よって，$\mathcal{C} \subset \mathcal{B}$ である．逆に，$(a,b] = \bigcap_{n=1}^{\infty} (a, b + 1/n)$ であるから，$\mathcal{B}$ は開区間全体を含む最小の σ-field に一致する．従って，$\mathcal{B} \subset \mathcal{C}$ であるから，$\mathcal{B} = \mathcal{C}$ を得る． □

$\mathbb{R}^n$ の Borel σ-field を $\mathcal{B}(\mathbb{R}^n)$ とおく．$n=1$ のときと同様に，

$$\mathcal{B}(\mathbb{R}^n)=\sigma(\{(a_1,b_1]\times\cdots\times(a_n,b_n]:-\infty\le a_i<b_i<\infty,1\le i\le n\})$$

が成り立つ (演習問題)．

可測空間 $(X_i,\mathcal{A}_i),1\le i\le n$ に対して，

$$\{A_1\times\cdots\times A_n:A_i\in\mathcal{A}_i,1\le i\le n\}$$

から生成される σ-field を $\mathcal{A}_1\times\cdots\times\mathcal{A}_n$ と書いて，$\mathcal{A}_1,\ldots,\mathcal{A}_n$ の 直積 σ-field (product σ-field) と呼ぶ[5]．$\mathcal{A}_i=\mathcal{A},1\le i\le n$ のときは，$\mathcal{A}\times\cdots\times\mathcal{A}=\mathcal{A}^n$ と書く．

Lemma 2.3. $\mathcal{B}(\mathbb{R}^n)=\mathcal{B}^n:=\mathcal{B}\times\cdots\times\mathcal{B}$.

Proof. $n=2$ のとき補題を示す．$A\subset\mathbb{R}$ が開集合なら，$A\times\mathbb{R}$ は $\mathbb{R}^2$ の開集合であるから，$A\times\mathbb{R}\in\mathcal{B}(\mathbb{R}^2)$ である．そこで，$\mathcal{C}=\{A:A\times\mathbb{R}\in\mathcal{B}(\mathbb{R}^2)\}$ とおくと，$\mathcal{C}$ は $\mathbb{R}$ の開集合全体を含む σ-field であるから，$\mathcal{C}\supset\mathcal{B}$ である．よって，任意の $A\in\mathcal{B}$ に対して，$A\times\mathbb{R}\in\mathcal{B}(\mathbb{R}^2)$ である．同様に，任意の $B\in\mathcal{B}$ に対して，$\mathbb{R}\times B\in\mathcal{B}(\mathbb{R}^2)$ であるから，任意の $A\in\mathcal{B},B\in\mathcal{B}$ に対して，$A\times B=(A\times\mathbb{R})\cap(\mathbb{R}\times B)\in\mathcal{B}(\mathbb{R}^2)$ となる．従って，$\mathcal{B}\times\mathcal{B}\subset\mathcal{B}(\mathbb{R}^2)$ を得る．逆の包含関係は，$\mathcal{B}(\mathbb{R}^2)=\sigma(\{(a_1,b_1]\times(a_2,b_2]:-\infty\le a_i<b_i<\infty,i=1,2\})$ からわかる． □

以下，特に断らない限り，$\mathbb{R}^n$ の Borel σ-field を $\mathcal{B}(\mathbb{R}^n)$ の代わりに $\mathcal{B}^n$ と書き，$\mathbb{R}^n$ には $\mathcal{B}^n$ を入れる．

Lebesgue (非) 可測集合

$\mathbb{R}$ の Borel σ-field $\mathcal{B}$ の Lebesgue 測度 λ に関する完備化を $\mathcal{L}$ とおいたとき，$\mathcal{L}$ に属する集合を Lebesgue 可測集合 と呼ぶ．Lebesgue 可測集合の全体は十分に広い集合族であり，$\mathbb{R}$ のほぼすべての部分集合は Lebesgue 可測だと思ってもさしあたり問題ない．とは言え，(選択公理を仮定すると)Lebesgue 非可測集合は存在する (後述)．なお，Lebesgue 可測だが，Borel 非可測な集合 ($\mathcal{B}$ に属していない集合) も存在する (次節を参照)．つまり，$\mathcal{B}\subsetneq\mathcal{L}\subsetneq 2^{\mathbb{R}}$ である．λ の完備化も記号を変えずに λ と書くことにする．

以下，Lebesgue 非可測集合を構成しよう．$x,y\in[0,1)$ に対して，

$$x\sim y\Leftrightarrow x-y\in\mathbb{Q}$$

と定義すると，$\sim$ は同値関係である．この同値関係に関して各同値類から代表元を集めた集合を V とおく (選択公理を使っている)．集合 V は Vitali 集合 とも呼ばれる．Vitali 集合は Lebesgue 非可測である．このことを証明する．

[5] $\mathcal{A}_1,\ldots,\mathcal{A}_n$ の直積 σ-field を $\mathcal{A}_1\times\cdots\times\mathcal{A}_n$ と書く記法はある程度標準的であるが，これだと集合としての直積とまぎらわしいので，$\mathcal{A}_1\otimes\cdots\otimes\mathcal{A}_n$ と書く場合がある．この記法も標準的である．

Theorem 2.2. $V \notin \mathcal{L}$.

各 $r \in \mathbb{R}$ に対して，写像 $T_r : [0,1) \to [0,1)$ を

$$T_r(x) = x + r - \lfloor x + r \rfloor$$

と定義する ($\lfloor x \rfloor$ は x 以下の最大の整数である)．つまり，$T_r(x)$ は $x+r$ の小数部分である．

Lemma 2.4. $E \in \mathcal{L}, E \subset [0,1)$ なら，任意の $r \in [0,1)$ に対して，$\lambda(T_r(E)) = \lambda(E)$.

Proof. 次は認める (演習問題)：$A \in \mathcal{L}$ なら，任意の $x \in \mathbb{R}$ に対して，$A + x := \{y + x : y \in A\} \in \mathcal{L}$ であって，$\lambda(A + x) = \lambda(A)$. $A = E \cap [0, 1-r), B = E \cap [1-r, 1)$ とおいて，$A' = A + r, B' = B + (r-1)$ とおくと，$A, B \in \mathcal{L}$ より，$A', B' \in \mathcal{L}$ であって，$\lambda(A) = \lambda(A'), \lambda(B) = \lambda(B')$. 一方，$T_r(E) = A' \cup B'$ (排反な和) より，$T_r(E) \in \mathcal{L}$ であって，

$$\lambda(E) = \lambda(A) + \lambda(B) = \lambda(A') + \lambda(B') = \lambda(T_r(E))$$

を得る． □

Proof of Theorem 2.2. 仮に V が Lebesgue 可測なら，$T_r(V), r \in \mathbb{Q} \cap [0,1)$ は Lebesgue 可測集合であり，$\lambda(T_r(V)) = \lambda(V) = \alpha$ (say). V の定義より，$T_r(V), r \in \mathbb{Q} \cap [0,1)$ は排反であって，$\bigcup_{r \in \mathbb{Q} \cap [0,1)} T_r(V) = [0,1)$. ゆえに，$\alpha > 0$ なら，$\lambda([0,1)) = \sum_{r \in \mathbb{Q} \cap [0,1)} \lambda(T_r(V)) = \infty$ となり，$\alpha = 0$ なら，$\lambda([0,1)) = \sum_{r \in \mathbb{Q} \cap [0,1)} \lambda(T_r(V)) = 0$ となって，いずれにしても矛盾が生じる．従って，$V \notin \mathcal{L}$ である． □

Lebesgue 非可測集合の存在は，完備化によっては Lebesgue 測度を $2^{\mathbb{R}}$ にまで拡張できないことを意味している．もちろんこれは，その他の方法によって Lebesgue 測度を $2^{\mathbb{R}}$ に拡張できる可能性を排除するものではない．しかしながら，連続体仮説を仮定すると，Lebesgue 測度を $2^{\mathbb{R}}$ にまで拡張することは不可能であることが知られている．Dudley (2002, Appendix C) を参照せよ．

Example 2.1 (Cantor 集合)**.** $\mathbb{R}$ の 1 点集合の Lebesgue 測度は 0 なので，$\mathbb{R}$ の可算集合の Lebesgue 測度も 0 である．しかし，非可算集合であって，Lebesgue 測度が 0 のものは存在する．そのような集合の有名な例として，Cantor 集合 がある．まず閉区間 $[0,1]$ を 3 分割して，真ん中の開区間 $(\frac{1}{3}, \frac{2}{3})$ を取り除く．すると，残るのは排反な 2 つの閉区間の和 $[0, \frac{1}{3}] \cup [\frac{2}{3}, 1]$ である．各閉区間 $[0, \frac{1}{3}], [\frac{2}{3}, 1]$ をそれぞれ 3 分割して，再び真ん中の開区間 $(\frac{1}{9}, \frac{2}{9}), (\frac{7}{9}, \frac{8}{9})$ を取り除く．この操作を繰り返すと，n 回目の操作が終わった段階で，

$$1 + 2 + 2^2 + \cdots + 2^{n-1} = 2^n - 1$$

個の排反な開区間が取り除かれていて，残っているのは 2^n 個の長さ $1/3^n$ の排反な閉区間である．取り除いた開区間を左から順に並べたものを $J_{n,k}, 1 \le k \le 2^n - 1$ とし，$U_n = \bigcup_{k=1}^{2^n-1} J_{n,k}$ とおくと，

$$\lambda(U_n) = \frac{1}{3} + \frac{2}{3^2} + \cdots + \frac{2^{n-1}}{3^n} = 1 - \left(\frac{2}{3}\right)^n.$$

ここで，$U=\bigcup_{n=1}^{\infty}U_n$ とおくと，U は開集合であって，$U_n\uparrow U$ であるから，$\lambda(U)=1$ である．集合 $C=[0,1]\setminus U$ を Cantor 集合と呼ぶ．C はコンパクト集合であって，$\lambda(C)=1-\lambda(U)=0$ である．C が非可算であることを示す．まず各 $x\in[0,1]$ は 3 進数展開

$$x=\sum_{j=1}^{\infty}\frac{a_j}{3^j},\ a_j\in\{0,1,2\}$$

をもつ．$x=k3^{-n}$ $(1\le k<3^n, n\in\mathbb{N})$ のときは，3 進数展開は一意でない．例えば，

$$\frac{1}{3}=\frac{1}{3}+\frac{0}{3^2}+\frac{0}{3^3}+\cdots=\frac{0}{3}+\frac{2}{3^2}+\frac{2}{3^3}+\cdots,$$
$$\frac{2}{3}=\frac{2}{3}+\frac{0}{3^2}+\frac{0}{3^3}+\cdots=\frac{1}{3}+\frac{2}{3^2}+\frac{2}{3^3}+\cdots.$$

k を 3 の倍数でないとすれば，このような場合は，$a_n=0, a_j=2$ $(j>n)$，または $a_n=2, a_j=0$ $(j>n)$ と決めておくと，各 $x\in[0,1]$ に対して 3 進数展開が一意に決まる．このとき，

$$a_1=1\Leftrightarrow\frac{1}{3}<x<\frac{2}{3},$$
$$a_1\neq1\ \&\ a_2=1\Leftrightarrow\frac{1}{9}<x<\frac{2}{9}\ \text{or}\ \frac{7}{9}<x<\frac{8}{9},$$
$$\vdots$$

であるから，$C=\{\sum_{j=1}^{\infty}a_j3^{-j}:a_j\in\{0,2\}\}$ と表せる．いま，$x=\sum_{j=1}^{\infty}a_j3^{-j}$ $(a_j\in\{0,2\})$ に対して，$F(x)=\sum_{j=1}^{\infty}(a_j/2)2^{-j}$ と定義すれば，F は C から $[0,1]$ への全射であるから，C が非可算であることが示された．

前述の議論に現れた関数 F は，$x,y\in C, x<y$ に対して，x,y が除かれた区間の両端点でない限りは，$F(x)<F(y)$ となる (例えば，$x=1/3, y=2/3$ なら，$x=0/3+2/3^2+2/3^3+\cdots, y=2/3+0/3^2+0/3^3+\cdots$ であるから，$F(x)=0/2+1/2^2+1/2^3+\cdots=1/2=1/2+0/2^2+0/2^3+\cdots=F(y)$)．そこで，各 $n\in\mathbb{N}, 1\le k\le 2^n-1$ に対して，$x_{n,k}$ を $J_{n,k}$ の左端点とし，F を

$$F(x)=F(x_{n,k}),\ x\in J_{n,k}$$

と拡張すれば，F は $[0,1]$ から $[0,1]$ への非減少関数になる．また F は全射であるから不連続点をもちえないので，F は連続である．この関数 F を Cantor 関数 と呼ぶ．

演習問題

Exercise 2.1. $f:\mathbb{R}\to\mathbb{R}$ を非減少関数とすると，f の不連続点は高々可算であることを示せ．

ヒント：f が有界で，$|f(x)|\le M$ $(\forall x\in\mathbb{R})$ なら，任意の $\varepsilon>0$ に対して，$\mathrm{Card}\{x: f(x+)-f(x-)>\varepsilon\}\le 2M/\varepsilon$ となる．

Exercise 2.2. Lemma 2.1 の結論は一般の Lebesgue-Stieltjes 測度に対しても成り立つ．このことを示せ．

Exercise 2.3. $\mathbb{R}^n$ の Borel σ-field $\mathcal{B}(\mathbb{R}^n)$ に対して，定義にもとづいて，$\mathcal{B}(\mathbb{R}^n) = \sigma(\{(a_1, b_1] \times \cdots \times (a_n, b_n] : -\infty \leq a_i < b_i < \infty, 1 \leq i \leq n\})$ を示せ．

Exercise 2.4. $(X, \mathcal{A}), (Y, \mathcal{C})$ を可測空間とし，$E \in \mathcal{A} \times \mathcal{C}$ とする．このとき，可算個の $A_n \in \mathcal{A}, C_n \in \mathcal{C}, n = 1, 2, \ldots$ が存在して，$E \in \sigma(\{A_n \times C_n : n = 1, 2, \ldots\})$ となることを示せ．

Exercise 2.5. Lemma 2.3 を一般化して次を示せ．$(X, d_1), (Y, d_2)$ を可分な距離空間とし，$\mathcal{B}(X), \mathcal{B}(Y)$ をそれぞれ X, Y の Borel σ-field とする．また，$X \times Y$ の直積位相に関する Borel σ-field を $\mathcal{B}(X \times Y)$ とおく．このとき，$\mathcal{B}(X \times Y) = \mathcal{B}(X) \times \mathcal{B}(Y)$ を示せ．

Remark 2.1. Exercise 2.5 において，X, Y の可分性は本質的な仮定である．実際，X が連続体濃度より真に大きい濃度をもつ位相空間のとき[6]，$\mathcal{B}(X) \times \mathcal{B}(X) \subsetneq \mathcal{B}(X \times X)$ になる．$\mathcal{B}(X) \times \mathcal{B}(X) \subset \mathcal{B}(X \times X)$ の証明は Lemma 2.3 と同様である．$\mathcal{B}(X) \times \mathcal{B}(X) \neq \mathcal{B}(X \times X)$ を示そう．

$$\mathcal{E}_0 := \left\{ E = \bigcup_{i \in I} (A_i \times B_i) : A_i, B_i \subset X, i \in I, \mathrm{Card}(I) = \mathrm{Card}(\mathbb{R}) \right\},$$
$$\mathcal{E} := \{E \subset X \times X : E \in \mathcal{E}_0 \ \& \ E^c \in \mathcal{E}_0\}$$

とおくと，$\mathcal{E}$ は $\mathcal{B}(X) \times \mathcal{B}(X)$ を含む σ-field である．ここで，$X \times X$ の対角線集合 $D = \{(x, x) : x \in X\}$ は $X \times X$ の閉集合であるから，$D \in \mathcal{B}(X \times X)$ であるが，明らかに $D \notin \mathcal{E}$ である．

Exercise 2.6. σ-field $\mathcal{A}$ が 可算生成 (countably generated) であるとは，$\mathcal{A} = \sigma(\mathcal{C})$ となる可算集合 $\mathcal{C} \subset \mathcal{A}$ が存在することである．可分な距離空間の Borel σ-field は可算生成であることを示せ．特に，$\mathbb{R}^n$ の Borel σ-field $\mathcal{B}^n$ は可算生成である．

Exercise 2.7. 可算生成な σ-field の sub σ-field は可算生成とは限らない．例えば，$\mathcal{A} = \{A \subset \mathbb{R} : A \text{ or } A^c \text{ は可算 }\}$ は $\mathcal{B}$ の sub σ-field であるが，$\mathcal{A}$ は可算生成でない．このことを示せ ($\mathcal{A}$ が σ-field であることは Exercise 1.8 で示した)．

Exercise 2.8. X を位相空間とし，$X_0 \subset X$ とする．このとき，X_0 の相対位相に関する Borel σ-field $\mathcal{B}(X_0)$ は X の Borel σ-field $\mathcal{B}(X)$ の X_0 への制限 $\{A \cap X_0 : A \in \mathcal{B}(X)\}$ に一致することを示せ：$\mathcal{B}(X_0) = \{A \cap X_0 : A \in \mathcal{B}(X)\}$.

[6] 可分な距離空間 X は連続体濃度以下の濃度をもつ．$\{x_n\}$ を X の可算稠密集合とすれば，X の各点は $\{x_n\}$ の集積点であるから，その濃度は $\mathrm{Card}(2^{\mathbb{N}}) = \mathrm{Card}(\mathbb{R})$ 以下である．

Exercise 2.9. 内点をもたない $\mathbb{R}$ のコンパクト集合であって，Lebesgue 測度が正のものを構成せよ．

ヒント：$\{x_n\}$ を $[0,1]$ の可算稠密集合として，$[0,1]$ から $(x_n - \varepsilon/2^{n+1}, x_n + \varepsilon/2^{n+1}), n = 1, 2, \ldots$ をすべて取り除いた集合を考えよ．

上の例からわかるように，$\mathbb{R}$ の部分集合の境界は Lebesgue 測度 0 とは限らない．

Exercise 2.10. λ を $\mathbb{R}$ 上の Lebesgue 測度とする．$x, r \in \mathbb{R}$ と $B \subset \mathbb{R}$ に対して，

$$B + x := \{y + x : y \in B\},\ rB := \{ry : y \in B\}$$

と定義する．このとき，任意の $x, r \in \mathbb{R}$ と $B \in \mathcal{B}$ に対して，$B + x, rB \in \mathcal{B}$ であって，

$$\lambda(B + x) = \lambda(B),\ \lambda(rB) = |r|\lambda(B)$$

となることを示せ．B が Lebesgue 可測集合ならどうか．

Exercise 2.11. μ を $(\mathbb{R}, \mathcal{B})$ 上の測度とし，任意の $-\infty < a < b < \infty$ に対して，$\mu((a, b]) < \infty$ とする．いま，任意の $x \in \mathbb{R}$ と $B \in \mathcal{B}$ に対して，

$$\mu(B + x) = \mu(B)$$

ならば，適当な定数 $\alpha \geq 0$ が存在して，$\mu = \alpha\lambda$ となることを示せ．

Exercise 2.12. $V \subset [0,1)$ を Vitali 集合とすると，$E \subset V$ が Lebesgue 可測なら，$\lambda(E) = 0$ となることを示せ．

Exercise 2.13. $E \subset \mathbb{R}$ を Lebesgue 可測集合とし，$\lambda(E) > 0$ とする．このとき，E に含まれる Lebesgue 非可測集合が存在することを示せ．

ヒント：$E \subset [0,1)$ と仮定してよい．

3 可測関数

可測空間 $(X,\mathcal{A}),(Y,\mathcal{C})$ に対して，写像 $f:X\to Y$ が 可測 (measurable) (または $\mathcal{A}/\mathcal{C}$ 可測) であるとは，任意の $C\in\mathcal{C}$ に対して，$f^{-1}(C)\in\mathcal{A}$ となることを言う．写像の可測性を示すときに，いちいちすべての $C\in\mathcal{C}$ に対して $f^{-1}(C)\in\mathcal{A}$ をチェックする必要はない．

Lemma 3.1. $\sigma(\mathcal{E})=\mathcal{C}$ なら，$f:X\to Y$ が可測 $\Leftrightarrow f^{-1}(E)\in\mathcal{A}\ \forall E\in\mathcal{E}$.

Proof. $\Rightarrow$ は明らか．$\Leftarrow$ を示す．$\mathcal{C}'=\{C\in\mathcal{C}:f^{-1}(C)\in\mathcal{A}\}$ とおくと，$\mathcal{E}\subset\mathcal{C}'$ であって，$\mathcal{C}'$ は σ-field．よって，$\mathcal{C}'\supset\sigma(\mathcal{E})=\mathcal{C}$. □

Lemma 3.2. $(X,\mathcal{A}),(Y,\mathcal{C}),(Z,\mathcal{D})$ を可測空間とし，$f:X\to Y, g:Y\to Z$ を可測写像とすると，合成写像 $g\circ f:X\to Z$ は可測である．

Proof. $D\in\mathcal{D}$ に対して,

$$(g\circ f)^{-1}(D)=\{x:g(f(x))\in D\}=\{x:f(x)\in g^{-1}(D)\}=f^{-1}(g^{-1}(D)).$$

ここで，$g^{-1}(D)\in\mathcal{C}$ であるから，$f^{-1}(g^{-1}(D))\in\mathcal{A}$ である． □

実数値関数．$Y=\mathbb{R}$ の場合を考える．$\mathbb{R}$ には Borel σ-field $\mathcal{B}$ を入れる．すなわち，実数値関数 $f:X\to\mathbb{R}$ が可測 (または $\mathcal{A}$ 可測) であるとは，任意の $B\in\mathcal{B}$ に対して，$f^{-1}(B)\in\mathcal{A}$ となることを言う．

Lemma 3.3. 任意の実数値関数 $f:X\to\mathbb{R}$ に対して，次の (i)–(v) は同値である．

(i) f は可測.

(ii) $\forall a\in\mathbb{R},\{x\in X:f(x)\le a\}\in\mathcal{A}$.

(iii) $\forall a\in\mathbb{R},\{x\in X:f(x)<a\}\in\mathcal{A}$.

(iv) $\forall a\in\mathbb{R},\{x\in X:f(x)\ge a\}\in\mathcal{A}$.

(v) $\forall a\in\mathbb{R},\{x\in X:f(x)>a\}\in\mathcal{A}$.

Proof. (i) $\Leftrightarrow$ (ii) は $\mathcal{B}=\sigma(\{(-\infty,a]:a\in\mathbb{R}\})$ から従う．他も同様である． □

$\mathbb{R}^n$ にはその Borel σ-field $\mathcal{B}^n$ を入れると約束していることを思い出す.

Lemma 3.4. (i) 連続関数 $g:\mathbb{R}^n\to\mathbb{R}$ は可測．(ii) 任意の関数 $f_i:X\to\mathbb{R}$ $(1\le i\le n)$ に対して，$f=(f_1,\dots,f_n):X\to\mathbb{R}^n$ が可測 $\Leftrightarrow$ 各 $f_i:X\to\mathbb{R}$ $(1\le i\le n)$ が可測.

Proof. (i). $O \subset \mathbb{R}$ を開集合とすると，$g^{-1}(O)$ は $\mathbb{R}^n$ の開集合であるから，$g^{-1}(O) \in \mathcal{B}^n$.
(ii). $n = 2$ で示す．$\mathcal{B}^2 = \mathcal{B} \times \mathcal{B} = \sigma(\{A \times B : A, B \in \mathcal{B}\})$ より，

$$\begin{aligned} f \text{が可測} &\Leftrightarrow f^{-1}(A \times B) \in \mathcal{A}\ \forall A, B \in \mathcal{B} \\ &\Leftrightarrow f_1^{-1}(A) \cap f_2^{-1}(B) \in \mathcal{A}\ \forall A, B \in \mathcal{B} \\ &\Leftrightarrow\ f_i, i = 1, 2 \text{が可測} \end{aligned}$$

である． □

Example 3.1. Lemma 3.4 より，$g : \mathbb{R}^n \to \mathbb{R}$ を連続，$f_i : X \to \mathbb{R}$ $(1 \le i \le n)$ を可測とすると，$g(f_1, \ldots, f_n) : X \to \mathbb{R}$ は可測である．例えば，$f, g : X \to \mathbb{R}$ が可測なら，

$$fg,\ f + g,\ \max(f, g),\ \min(f, g)$$

はすべて可測である．

拡張実数値関数．応用上，関数の値域を拡張実数 $\overline{\mathbb{R}} = [-\infty, \infty]$ にまで許した方が都合がよい．例えば，各 f_n が X 上の実数値関数だったとしても，$\sup_n f_n$ は実数値になるとは限らないし，いちいち実数値になるように制限するのも面倒である．$\overline{\mathbb{R}}$ には適当な距離に関する Borel σ-field を入れる．$\varphi : \overline{\mathbb{R}} \to [-1, 1]$ を

$$\varphi(x) = \begin{cases} 1 & \text{if } x = \infty \\ \frac{x}{1+|x|} & \text{if } -\infty < x < \infty \\ -1 & \text{if } x = -\infty \end{cases}$$

と定義すると，φ は狭義単調増加かつ全射である．このとき，$\overline{\mathbb{R}}$ は

$$d(x, y) = |\varphi(x) - \varphi(y)|,\ x, y \in \overline{\mathbb{R}}$$

を距離にもつ距離空間となり，φ は $\overline{\mathbb{R}}$ から $[-1, 1]$ への等長同型 (距離を保つ全射) になる (従って，$\overline{\mathbb{R}}$ の位相は $[-1, 1]$ から “引き戻して” 考えればよい)．$\overline{\mathbb{R}}$ には，この距離に関する Borel σ-field $\mathcal{B}^*$ を入れる．拡張実数値関数 $f : X \to \overline{\mathbb{R}}$ が可測 (または $\mathcal{A}$ 可測) であるとは，任意の $B \in \mathcal{B}^*$ に対して，$f^{-1}(B) \in \mathcal{A}$ となることを言う．実数値関数のときと同様に，次の同値関係が成り立つ．

Lemma 3.5. 任意の拡張実数値関数 $f : X \to \overline{\mathbb{R}}$ に対して，次の (i)–(v) は同値である．

(i) f は可測．

(ii) $\forall a \in \mathbb{R}, \{x \in X : f(x) \le a\} \in \mathcal{A}$.

(iii) $\forall a \in \mathbb{R}, \{x \in X : f(x) < a\} \in \mathcal{A}$.

(iv) $\forall a \in \mathbb{R}, \{x \in X : f(x) \geq a\} \in \mathcal{A}$.

(v) $\forall a \in \mathbb{R}, \{x \in X : f(x) > a\} \in \mathcal{A}$.

Proof. (i) $\Leftrightarrow$ (ii) のみ示す．$\mathcal{C} = \sigma(\{[-\infty, a] : a \in \mathbb{R}\})$ とおく．$\mathcal{B}^* = \mathcal{C}$ を示せば十分である．$[-\infty, a), (a, b), (a, \infty]$ の形の区間はすべて $\mathcal{C}$ に属する．$\overline{\mathbb{R}}$ の任意の開集合はそのような区間の可算和で表せるので，$\mathcal{C}$ に属する．よって，$\mathcal{B}^* \subset \mathcal{C}$ を得る．$\mathcal{B}^* \supset \mathcal{C}$ は明らか． □

Remark 3.1. $f : X \to \mathbb{R}$ は $\overline{\mathbb{R}}$ への関数とも考えることができるが，f が $\mathbb{R}$ への関数として可測なことと，$\overline{\mathbb{R}}$ への関数として可測なことは同値である．

Lemma 3.6. $f_n : X \to \overline{\mathbb{R}}, n \in \mathbb{N}$ が可測なら，

$$\sup_n f_n,\ \inf_n f_n,\ \limsup_n f_n,\ \liminf_n f_n$$

はすべて可測である．$\lim_n f_n(x) = f(x)$ $(\forall x \in X)$ が存在すれば，f も可測である．

Proof. $\sup_n f_n$ と $\inf_n f_n$ が可測なことは，

$$\{\sup_n f_n \leq a\} = \bigcap_n \{f_n \leq a\} \in \mathcal{A},\ \{\inf_n f_n \geq a\} = \bigcap_n \{f_n \geq a\} \in \mathcal{A}$$

からわかる．$\limsup_n f_n$ と $\liminf_n f_n$ が可測なことは，$\limsup_n f_n = \inf_n \sup_{m \geq n} f_m$ と $\liminf_n f_n = \sup_n \inf_{m \geq n} f_m$ からわかる．$f = \lim_n f_n$ が存在すれば，$f = \limsup_n f_n$ であるから，f は可測である． □

次の補題の証明は演習問題とする．

Lemma 3.7. $f, g : X \to \overline{\mathbb{R}}$ を可測とする．(i) fg は可測．(ii) f, g が同時に，$f(x) = \infty, g(x) = -\infty$, または $f(x) = -\infty, g(x) = \infty$ とならなければ，$f + g$ は可測．

Remark 3.2. 本講義ノートでは，$f : \mathbb{R} \to \overline{\mathbb{R}}$ が可測と言ったら，f が $\mathcal{B}/\mathcal{B}^*$ 可測なことを意味する．$\mathbb{R}$ には Lebesgue 可測集合全体 $\mathcal{L}$ も入れることができて，測度論の教科書だと，$\mathcal{L}/\mathcal{B}^*$ 可測なことを Lebesgue 可測，$\mathcal{B}/\mathcal{B}^*$ 可測なことを Borel 可測と区別していることが多いが，ここでは Borel 可測なことを単に可測と言っている．以下で見るように，Lebesgue 可測であるが Borel 可測でない集合というのが存在するので，その指示関数を考えれば，Lebesgue 可測であるが Borel 可測でない関数というのも存在する．従って，Lebesgue 測度に関する積分 (次節を参照) を考える場合，Lebesgue 可測な関数を考えるのは自然であるが，確率論では特に指定していない確率測度 (全測度が 1 の測度) に関する積分を考えることになるので，Borel 可測関数を考える方が都合がよい (完備化は測度に依存した操作である！)．

Example 3.2 (Lebesgue 可測であるが Borel 可測でない集合)**.** Cantor 関数を用いて，Lebesgue 可測であるが Borel 可測でない集合を 具体的に 構成する．$F : [0,1] \to [0,1]$ を Cantor 関数とし，

$$H(x) = \frac{1}{2}(F(x) + x),\ x \in [0,1]$$

とおくと，H は $[0,1]$ からそれ自身への同相写像である．$C \subset [0,1]$ を Cantor 集合とすると，$H(C)$ はコンパクト集合であって，$\lambda(H(C)) = 1/2$ である．実際，前節の記号を用いると，$\lambda(H(J_{n,k})) = 1/(2 \cdot 3^n)$ であるから，$\lambda(H(U)) = (1/4)\sum_{n=1}^{\infty}(2/3)^n = 1/2$．よって，$\lambda(H(C)) = \lambda([0,1] \setminus H(U)) = 1/2$. そこで，$A$ を $H(C)$ に含まれる Lebesgue 非可測集合として (cf. Exercise 2.13)，$B = H^{-1}(A)$ とおくと，$B \subset C, \lambda(C) = 0$ であるから，B は Lebesgue 可測である．しかし，B は Borel 可測でない．実際，仮に B が Borel 可測なら，$H^{-1} : [0,1] \to [0,1]$ は連続なので，$(H^{-1})^{-1}(B) = H(B) = A$ は Borel 可測になるが，A は Lebesgue 非可測であったから矛盾が生じる．

じつは Lebesgue 可測だが Borel 非可測な集合の存在は，次の考察からもわかる．一般に可算生成な σ-field は連続体濃度をもつことが知られている (Folland, 1999, p. 41-42)．しかし，Cantor 集合 C は非可算な Borel 可測 (よって Lebesgue 可測) 集合であり，その Lebesgue 測度は 0 であるから，その部分集合は Lebesgue 可測である．よって，Lebesgue 可測集合の全体は連続体濃度より真に大きい濃度をもつから，その中には Borel 非可測な集合を含む．

この議論は非構成的であるが，Lebesgue 可測集合全体の中に Borel 非可測集合はかなりたくさんあるということがわかる．また，$\mathbb{R}$ のあらゆる非可算集合は Borel 非可測な部分集合をもつこともわかる．

演習問題

Exercise 3.1. X, Y を空でない集合とし，$f : X \to Y$ を写像とする．このとき，

$$f^{-1}(A^c) = (f^{-1}(A))^c,\ f^{-1}\left(\bigcup_{i\in I} A_i\right) = \bigcup_{i\in I} f^{-1}(A_i),\ f^{-1}\left(\bigcap_{i\in I} A_i\right) = \bigcap_{i\in I} f^{-1}(A_i)$$

を示せ．ただし，I は任意の添え字集合とする．

Exercise 3.2. 前問において，$A \subset X$ に対して，$f(A) = \{f(x) : x \in A\}$ と定義する．

(a) $f(\bigcup_{i\in I} A_i) = \bigcup_{i\in I} f(A_i),\ f\left(\bigcap_{i\in I} A_i\right) \subset \bigcap_{i\in I} f(A_i)$ を示せ．

(b) $f\left(\bigcap_{i\in I} A_i\right) \subsetneq \bigcap_{i\in I} f(A_i)$ となるような例を構成せよ．

(c) $f(A^c) \neq \varnothing, (f(A))^c \neq \varnothing, f(A^c) \cap (f(A))^c = \varnothing$ となるような例を構成せよ．

Exercise 3.3. $f, g : X \to \overline{\mathbb{R}}$ を可測とし，$A \in \mathcal{A}$ とする．関数 $h : X \to \overline{\mathbb{R}}$ を

$$h(x) = \begin{cases} f(x) & \text{if } x \in A \\ g(x) & \text{if } x \in A^c \end{cases}$$

と定義すると，h は可測であることを示せ．

Exercise 3.4. $A_1, \ldots, A_m$ を X の分割とする．すなわち，$A_1, \ldots, A_m$ は排反であって，$\bigcup_{i=1}^m A_i = X$ とする．このとき，関数 $f : X \to \mathbb{R}$ が $\sigma(\{A_1, \ldots, A_m\})$ 可測になるための必要十分条件は，ある $b_1, \ldots, b_m \in \mathbb{R}$ が存在して，$f = \sum_{i=1}^m b_i 1_{A_i}$ と表せることを示せ．

Exercise 3.5. I を非可算集合とし，$f_i : X \to \overline{\mathbb{R}}, i \in I$ を可測関数とする．このとき，$\sup_{i \in I} f_i$ は可測とは限らない．そのような例を構成せよ．

Exercise 3.6. Lemma 3.7 を示せ．

Exercise 3.7. $(X, \mathcal{A}, \mu)$ を測度空間とし，$f : X \to \overline{\mathbb{R}}$ を (可測かどうかわからない) 任意の関数，$g : X \to \overline{\mathbb{R}}$ を可測関数とする．いま，$N \in \mathcal{A}, \mu(N) = 0$ となる N が存在して，$f(x) = g(x)$ $(\forall x \in X \setminus N)$ とする．

(a) $(X, \mathcal{A}, \mu)$ が完備なとき，f が可測なことを示せ．

(b) $(X, \mathcal{A}, \mu)$ が完備でないときは，f は $\mathcal{A}/\mathcal{B}^*$ 可測とは限らない．そのような例を構成せよ．

Exercise 3.8. 可測関数 $f : X \to \mathbb{R}$ の像 $f(X)$ は Lebesgue 可測になるか．

Exercise 3.9. $f : X \to \mathbb{R}$ を可測関数とする．このとき，f のグラフ $\{(x, f(x)) \in X \times \mathbb{R} : x \in X\}$ は $\mathcal{A} \times \mathcal{B}$ 可測であることを示せ．

Exercise 3.10. 連続関数 $f : \mathbb{R} \to \mathbb{R}$ の像 $f(\mathbb{R})$ は $\mathcal{B}$ 可測であることを示せ．

Exercise 3.11. $f : \mathbb{R} \to \mathbb{R}$ を $|f|$ が可測になる関数とする．このとき，f は可測か．

Exercise 3.12. $\mathbb{R}$ から $\mathbb{R}$ への単調関数は可測であることを示せ．

Exercise 3.13. $\mathbb{R}$ から $\mathbb{R}$ への右連続関数と左連続関数は可測であることを示せ．

Exercise 3.14. (Jenrich, 1969). $\Theta \subset \mathbb{R}^d$ を空でないコンパクト集合とし，$F : X \times \Theta \to \mathbb{R}$ を次をみたす関数とする：$\theta \in \Theta$ を任意に固定すると，$F(x, \theta)$ は x の関数として $\mathcal{A}$ 可測であって，$x \in X$ を任意に固定すると，$F(x, \theta)$ は θ の関数として連続である．このとき，次をみたす $\mathcal{A}/\mathcal{B}^d$ 可測写像 $\hat{\theta} : X \to \Theta$ の存在を示せ：$F(x, \hat{\theta}(x)) = \inf_{\theta \in \Theta} F(x, \theta)$ $(\forall x \in X)$

Remark 3.3. 統計学において，推定量はランダムな関数の最大ないし最小化点で定義されることが多い (そのような推定量は M 推定量と呼ばれる) が，その場合，推定量を可測に選べるかどうかはそんなに明らかではない．Exercise 3.14 の結論は，推定量が可測に選べるための 1 つの十分条件を与える．もちろんこれは最良の条件ではなく，可測な推定量の存在はもっと弱い条件で成り立つ．ただし，その考察には測度論に関する深い議論を必要とする．第 39 節を参照せよ．

以下の問題では，(X, d) を距離空間とし，$\mathcal{B}(X)$ を X の Borel σ-field とする．

Exercise 3.15. 関数 $f : X \to \overline{\mathbb{R}}$ が 下半連続 (lower semicontinuous) であるとは，任意の $x \in X$ と x に収束する任意の点列 x_n に対して，$\liminf_n f(x_n) \geq f(x)$ となることを言う．また $-f$ が下半連続なとき，f を 上半連続 (upper semicontinuous) と言う．$f : X \to \overline{\mathbb{R}}$ が下半連続であるためには，任意の $\alpha \in \mathbb{R}$ に対して，$\{x : f(x) > \alpha\}$ が開集合となることが必要十分であることを示せ．特に，半連続関数は $\mathcal{B}(X)/\mathcal{B}^*$ 可測である．

Exercise 3.16. $f : X \to \mathbb{R}$ を任意の関数とし，D_f を f の不連続点全体とする．このとき，$D_f \in \mathcal{B}(X)$ を示せ．

ヒント:$\delta > 0$ に対して，$f^\delta(x) = \sup\{f(y) : d(x, y) < \delta\}, f_\delta(x) = \inf\{f(y) : d(x, y) < \delta\}$ とおくと，f^δ と f_δ はそれぞれ下半連続，上半連続である．そこで，$f^0 = \lim_{\delta \downarrow 0} f^\delta, f_0 = \lim_{\delta \downarrow 0} f_\delta$ とおくと，$D_f = \{x : f^0(x) \neq f_0(x)\}$ である．

Exercise 3.17. X の Borel σ-field $\mathcal{B}(X)$ はすべての連続関数 $f : X \to \mathbb{R}$ が可測になる X 上の最小の σ-field に一致することを示せ．

Remark 3.4. 位相空間 X に対して，すべての連続関数 $f : X \to \mathbb{R}$ が可測になる X 上の最小の σ-field のことを Baire σ-field と呼ぶ．X が距離空間のときは Baire σ-field は Borel σ-field に一致するが，一般には後者の方が大きい．Dudley (2002, Section 7.1) を参照せよ．

4 積分

$(X, \mathcal{A}, \mu)$ を測度空間とする．適当な可測関数 $f : X \to \overline{\mathbb{R}}$ に対して，積分

$$\int f d\mu$$

を定義したい．非負単関数，非負可測関数，一般の可測関数，の順に積分を定義していく．

Step 1. 非負単関数.

$$f = \sum_{i=1}^{n} a_i 1_{A_i}, \quad a_1, \ldots, a_n \in \mathbb{R}, A_1, \ldots, A_n \in \mathcal{A} \text{ は排反}, \tag{$*$}$$

なる形の関数 f を 単関数 (simple function) と呼ぶ．単関数の値域は有限集合である．逆に，可測関数 $f : X \to \mathbb{R}$ の値域が有限集合 $\{a_1, \ldots, a_n\}$ なら，$A_i = f^{-1}(\{a_i\}), i = 1, \ldots, n$ とおけば，$f = \sum_{i=1}^{n} a_i 1_{A_i}$ と表せる．ここで，$A_1, \ldots, A_n$ は排反であって，f の可測性より各 A_i は $\mathcal{A}$ 可測である．すなわち，単関数とは値域が有限集合である実数値可測関数に他ならない.

$a_i \geq 0$ $(1 \leq \forall i \leq n)$ のとき，$(*)$ の単関数に対して，

$$\int f d\mu := \sum_{i=1}^{n} a_i \mu(A_i)$$

と定義する．右辺は $+\infty$ にもなりうる．また，$0 \cdot \infty = 0$ と約束している．$\int f d\mu$ が well-defined であることを確認しよう．$f = \sum_i a_i 1_{A_i} = \sum_j b_j 1_{B_j}, a_i > 0$ $(\forall i), b_j > 0$ $(\forall j)$ とする．ただし，$\{A_i\}$ と $\{B_j\}$ はそれぞれ排反な $\mathcal{A}$ 可測集合列である．$\bigcup_i A_i = \bigcup_j B_j$ であるから，$A_i = \bigcup_j (A_i \cap B_j)$ (排反な和) であって，$A_i \cap B_j \neq \varnothing$ なら $a_i = b_j$ であるから，$a_i \mu(A_i) = \sum_j b_j \mu(A_i \cap B_j)$．従って，

$$\sum_i a_i \mu(A_i) = \sum_i \sum_j b_j \mu(A_i \cap B_j) = \sum_j b_j \sum_i \mu(A_i \cap B_j) = \sum_j b_j \mu(B_j).$$

最後の等号は，$B_j = \bigcup_i (A_i \cap B_j)$ (排反な和) による．これで非負単関数の積分が well-defined であることが示された.

Lemma 4.1. f, g を非負単関数とし，$c \geq 0$ を定数とすると，

$$\int f d\mu \geq 0, \ \int (cf) d\mu = c \int f d\mu, \ \int (f + g) d\mu = \int f d\mu + \int g d\mu.$$

Proof. 最後の等式のみ示す．$f = \sum_{i=1}^{n} a_i 1_{A_i}, g = \sum_{j=1}^{m} b_j 1_{B_j}$ とする．$A_0 = \bigcup_{j=1}^{m} B_j \setminus \bigcup_{i=1}^{n} A_i, B_0 = \bigcup_{i=1}^{n} A_i \setminus \bigcup_{j=1}^{m} B_j$ とおくと，$\bigcup_{i=0}^{n} A_i = \bigcup_{j=0}^{m} B_j$ (両辺は排反な和)．$a_0 =$

$b_0 = 0$ とおくと，$f + g = \sum_{i=0}^{n} \sum_{j=0}^{m} (a_i + b_j) 1_{A_i \cap B_j}$ と表せる．従って，

$$\begin{aligned}\int (f+g) d\mu &= \sum_{i=0}^{n} \sum_{j=0}^{m} (a_i + b_j) \mu(A_i \cap B_j) \\ &= \sum_{i=0}^{n} a_i \underbrace{\sum_{j=0}^{m} \mu(A_i \cap B_j)}_{=\mu(A_i)} + \sum_{j=0}^{m} b_j \underbrace{\sum_{i=0}^{n} \mu(A_i \cap B_j)}_{=\mu(B_j)} \\ &= \sum_{i=0}^{n} a_i \mu(A_i) + \sum_{j=0}^{m} b_j \mu(B_j) = \int f d\mu + \int g d\mu.\end{aligned}$$

□

また，$f \leq g$ なら，$g - f$ も非負単関数なので，

$$\int g d\mu = \int f d\mu + \int (g - f) d\mu \geq \int f d\mu$$

となる．

Step 2. 非負可測関数．$f : X \to [0, \infty]$ を可測関数とする．f の積分を

$$\int f d\mu := \sup \left\{ \int g d\mu : 0 \leq g \leq f,\ g \text{ は単関数} \right\}$$

と定義する．

Lemma 4.2. 任意の非負可測関数 $f : X \to [0, \infty]$ に対して，非負単関数の列 f_n であって，$f_n(x) \uparrow f(x)$ $(\forall x \in X)$ となるものが存在する．そのような f_n に対して，$\int f_n d\mu \uparrow \int f d\mu$.

Proof. $E_{n,j} = f^{-1}((j/2^n, (j+1)/2^n])$ $(j = 1, \ldots, 2^n n - 1), E_n = f^{-1}((n, \infty])$ とおき，

$$f_n := \sum_{j=1}^{2^n n - 1} (j/2^n) 1_{E_{n,j}} + n 1_{E_n}$$

と定義すると，$0 \leq f_n \uparrow f$.

g を $0 \leq g \leq f$ なる任意の単関数とする．h_n を $0 \leq h_n \uparrow f$ なる単関数とすると，$\int h_n d\mu \uparrow \exists c \in [0, \infty].\ c \geq \int g d\mu$ を示せばよい．$g = \sum_i a_i 1_{A_i}$ & $X = \bigcup_i A_i$ (排反な和) と表すと，各 i に対して，

$$\lim_n \int h_n 1_{A_i} d\mu \geq a_i \mu(A_i)$$

を示せば十分である (左辺は n に関して非減少なので極限が存在する)．$a_i = 0$ なら明らかなので，$a_i > 0$ と仮定する．$0 < \varepsilon < 1$ を任意に固定して，$B_{n,i} = \{x \in A_i : h_n(x) > (1 - \varepsilon) a_i\}$ とおくと，$B_{n,i} \uparrow A_i$. よって，$\lim_n \int h_n 1_{A_i} d\mu \geq (1 - \varepsilon) a_i \lim_n \mu(B_{n,i}) = (1 - \varepsilon) a_i \mu(A_i)$. $\varepsilon \downarrow 0$ として補題の結論を得る． □

次の事実は明らかであろう．f, g を非負可測とし，$a, b \geq 0$ を定数とすると，

$$\int f d\mu \geq 0, \ \int (af + bg) d\mu = a \int f d\mu + b \int g d\mu.$$

Step 3. 一般の可測関数．$f : X \to \overline{\mathbb{R}}$ を可測関数とする．

$$f^+ := \max\{f, 0\}, \ f^- := -\min\{f, 0\}$$

とおくと，f^+, f^- は非負可測であって，$f = f^+ - f^-$．$\int f^+ d\mu, \int f^- d\mu$ のどちらかが ∞ でないとき，積分 $\int f d\mu$ が定義できると言う．このとき，

$$\int f d\mu := \int f^+ d\mu - \int f^- d\mu$$

と定義する．また，f の積分が定義できるとき，$A \in \mathcal{A}$ に対して，

$$\int_A f d\mu := \int f 1_A d\mu$$

と定義する．ここで，$f^+ 1_A \leq f^+, f^- 1_A \leq f^-$ であるから，f の積分が定義できるなら，$f 1_A$ の積分は常に定義できる．

$\int |f| d\mu = \int f^+ d\mu + \int f^- d\mu < \infty$ のとき，f は 可積分 (integrable) であると言う．実数 $0 < p < \infty$ に対して，$|f|^p$ が可積分なとき，f は p 乗可積分であると言う．p 乗可積分な X 上の実数値関数の全体を $L^p(X, \mathcal{A}, \mu)$ とおく．

$$L^p(X, \mathcal{A}, \mu) := \left\{ f : f : X \to \mathbb{R}, \ f \text{ は } \mathcal{A} \text{ 可測}, \ \int |f|^p d\mu < \infty \right\}.$$

誤解の恐れがない場合は，$L^p(X, \mathcal{A}, \mu)$ を $L^p(\mu)$ または L^p と省略する．$a, b \geq 0$ に対して，$(a+b)^p \leq 2^p (a \vee b)^p \leq 2^p (a^p + b^p)$ より，$f, g \in L^p$ なら $f + g \in L^p$ である．よって，L^p は実ベクトル空間である．非負可測関数に対する積分の線形性より，次の定理が直ちに従う．

Theorem 4.1 (積分の線形性)**.** $L^1 \ni f \mapsto \int f d\mu$ は線形である．すなわち，任意の $f, g \in L^1$ と $a, b \in \mathbb{R}$ に対して，$\int (af + bg) d\mu = a \int f d\mu + b \int g d\mu$.

Remark 4.1 (複素数値関数の積分)**.** $\mathbb{C}$ の通常の位相に関する Borel σ-field を $\mathcal{B}(\mathbb{C})$ とおくと，複素数値関数 $f : X \to \mathbb{C}$ が $\mathcal{A}/\mathcal{B}(\mathbb{C})$ 可測なとき，それを単に可測と言う．f が可測なことは，その実部と虚部が $\mathbb{R}$ への関数として可測なことと同値であり ($\mathbb{R}^2$ への関数とみなせばよい)，さらに実部と虚部が $\mathbb{R}$ への関数として可積分なとき，f は可積分であると言う．f が可積分なことは，$|f|$ が $\mathbb{R}$ への関数として可積分なことと同値であり，そのとき，

$$\int f d\mu := \int (\mathrm{Re} f) d\mu + i \int (\mathrm{Im} f) d\mu, \ i = \sqrt{-1}$$

と定義する．従って，$\mathbb{C}$ への関数の積分は $\mathbb{R}$ への関数の積分に帰着するので，以降は $\mathbb{R}$ (or $\overline{\mathbb{R}}$) への関数の積分の性質を調べることにする．

Example 4.1. (積分の例).

- $\mu = \delta_x$ のとき，$\int f d\mu = f(x)$.
- $x_n \in X, c_n \geq 0$ $(n = 1, 2, \ldots, N \leq \infty), \mu = \sum_{n=1}^{N} c_n \delta_{x_n}$ のとき，$\int f d\mu = \sum_{n=1}^{N} c_n f(x_n)$.
- Lebesgue 測度．$\mathcal{L}$ を $\mathbb{R}$ の Lebesgue 可測集合の全体とし，$-\infty < a < b < \infty$ とする．このとき，Riemann 積分可能な有界関数 $f : [a, b] \to \mathbb{R}$ は自動的に $\mathcal{L}_{[a,b]}/\mathcal{B}$ 可測になる．f は有界であるから，λ に関して可積分であり，さらに

$$\underbrace{\int_a^b f(x)dx}_{\text{Riemann 積分}} = \int_{[a,b]} f d\lambda$$

が成り立つ．証明は後述する．逆に，$f(x) = 1_{\mathbb{Q}\cap[0,1]}(x), x \in [0, 1]$ は Riemann 積分可能でないが，$\mathbb{Q} \cap [0, 1] \in \mathcal{B}$ であるから，λ に関して積分が定義できて，さらに，$\lambda(\mathbb{Q} \cap [0, 1]) = 0$ より，$\int_{[0,1]} f d\lambda = 0$ となる．この意味で，Lebesgue 積分は Riemann 積分の一般化になっている．以下，$d\lambda(x) = dx$ と書く．

$(Y, \mathcal{C})$ をもう 1 つの可測空間として，$T : X \to Y$ を可測写像とする．このとき，

$$(\mu \circ T^{-1})(C) := \mu(T^{-1}(C)),\ C \in \mathcal{C}$$

と定義すると，写像 $\mu \circ T^{-1} : \mathcal{C} \to [0, \infty]$ は測度になる．これを T による μ の 像測度 (image measure) と呼ぶ.

Theorem 4.2 (変数変換公式)**.** 任意の可測関数 $f : Y \to \overline{\mathbb{R}}$ に対して，$\int f d(\mu \circ T^{-1})$ と $\int (f \circ T) d\mu$ のどちらかが定義されるならもう一方も定義されて，両者は等しい.

Proof. 積分の構成より，$f = 1_C, C \in \mathcal{C}$ の場合を考えればよい．この f に対して，$f \circ T = 1_{T^{-1}(C)}$ であるから，$\int f d(\mu \circ T^{-1}) = (\mu \circ T^{-1})(C) = \mu(T^{-1}(C)) = \int (f \circ T) d\mu$. □

測度 0 の集合の役割. $x \in X$ に関する命題が "almost everywhere" (a.e.) に成り立つとは，ある集合 $N \in \mathcal{A}, \mu(N) = 0$ が存在して，すべての $x \notin N$ に対してその命題が成り立つことを言う．例えば，$f \leq g$ a.e. と書いたら，ある集合 $N \in \mathcal{A}, \mu(N) = 0$ が存在して，$f(x) \leq g(x)$ $(\forall x \notin N)$ が成り立つことを意味する．積分に関する限り，測度 0 の集合上での違いは無視してよい.

Lemma 4.3. $f, g : X \to \overline{\mathbb{R}}$ は可測で，$f = g$ a.e. とする．このとき，$\int f d\mu, \int g d\mu$ のどちらかが定義されるならもう一方も定義されて，両者は等しい.

トリビアルな例として，$f = 0$ a.e. なら，$\int f d\mu = 0$ である.

Proof. $f(x) = g(x)$ $(\forall x \notin N)$ とする．任意の非負可測関数 h に対して，$\int h d\mu = \int_{X\setminus N} h d\mu$ を示せばよい．積分の構成より，$h = 1_A, A \in \mathcal{A}$ の場合を考えれば十分だが，このとき，$\int h d\mu = \mu(A) = \mu(A \setminus N) = \int_{X\setminus N} h d\mu$. □

Example 4.2. $f, g \geq 0$ or $f, g \in L^1$ に対して，$f \leq g$ a.e. なら，

$$\tilde{f}(x) = \begin{cases} f(x) & \text{if } x \in N^c \\ 0 & \text{if } x \in N \end{cases}, \quad \tilde{g}(x) = \begin{cases} g(x) & \text{if } x \in N^c \\ 0 & \text{if } x \in N \end{cases}$$

とおくと，$\tilde{f} = f$ a.e., $\tilde{g} = g$ a.e., $\tilde{f}(x) \leq \tilde{g}(x)$ $(\forall x \in X)$. 従って，

$$\int g d\mu = \int \tilde{g} d\mu = \int \tilde{f} d\mu + \int (\tilde{g} - \tilde{f}) d\mu \geq \int \tilde{f} d\mu = \int f d\mu.$$

可測関数の定義を a.e. に定義された関数にまで拡張し，拡張された意味での可測関数に対して積分を定義する．ある集合 $N \in \mathcal{A}, \mu(N) = 0$ に対して，N を除いた集合上で定義された関数 $f : X \setminus N \to \overline{\mathbb{R}}$ が，$f^{-1}(B) = \{x \in X \setminus N : f(x) \in B\} \in \mathcal{A}$, $\forall B \in \mathcal{B}^*$ をみたすとき，f を a.e. に定義された 可測関数と呼ぶ．a.e. に定義された可測関数 f に対して，

$$\tilde{f}(x) = \begin{cases} f(x) & \text{if } x \in X \setminus N \\ 0 & \text{if } x \in N \end{cases}$$

とおくと，$\tilde{f} : X \to \overline{\mathbb{R}}$ は可測であって，$\int \tilde{f} d\mu$ が定義できるとき，f の積分を

$$\int f d\mu := \int \tilde{f} d\mu$$

と定義する．a.e. に定義された可測関数は Fubini の定理に現れる．

積分の性質

実数 $1 \leq p < \infty$ に対して，$\|f\|_p = (\int |f|^p d\mu)^{1/p}$ と定義する．

Theorem 4.3 (Hölder の不等式)**.** $f, g : X \to \overline{\mathbb{R}}$ を可測関数とし, $p, q \in (1, \infty)$ を $p^{-1} + q^{-1} = 1$ なる実数とすると，$\int |fg| d\mu \leq \|f\|_p \|g\|_q$ が成り立つ.

$p = q = 2$ のとき，Hölder の不等式を Cauchy-Schwarz の不等式 と呼ぶ．Hölder の不等式は，左辺が $+\infty$ のときは右辺は $+\infty$ になるという意味で，f, g に対する可積分性の条件なしに成り立つ．以下，このような注意は繰り返さない．

Proof. $\|f\|_p = 0$ or $\|g\|_q = 0$ なら $fg = 0$ a.e. なので，$\|f\|_p > 0, \|g\|_q > 0$ の場合のみ考えればよい．また，$\|f\|_p = +\infty$ or $\|g\|_q = +\infty$ なら，$\|f\|_p \|g\|_q = +\infty$ となって，この場合は定理の結論は明らかなので，$\|f\|_p$ と $\|g\|_q$ は有限と仮定してよい．さら

に，f と g をそれぞれ $f/\|f\|_p$ と $g/\|g\|_q$ に取り替えて，$\|f\|_p = \|g\|_q = 1$ と仮定してよい．$ab \leq a^p/p + b^q/q$ $(a \geq 0, b \geq 0)$ より，$|fg| \leq |f|^p/p + |g|^q/q$．両辺を積分して，$\int |fg| d\mu \leq 1/p + 1/q = 1 = \|f\|_p \|g\|_q$ を得る． □

Theorem 4.4 (単調収束定理 (MCT))**.** $f_n : X \to [0, \infty], n \in \mathbb{N}$ を可測関数とし，$f_n(x) \uparrow f(x)$ $(\forall x \in X)$ とすると，f も可測であって，

$$\int f_n d\mu \uparrow \int f d\mu.$$

Proof. f が可測なのはよい．f_n に対して，単関数 $f_{n,m}$ で，$0 \leq f_{n,m} \uparrow f_n$ $(m \to \infty)$ となるものをとり，$g_n = \max\{f_{1,n}, \ldots, f_{n,n}\}$ とおくと，g_n は単関数で $0 \leq g_n \uparrow f$．従って，$\int g_n d\mu \uparrow \int f d\mu$．$g_n \leq f_n \leq f$ より，$\int g_n d\mu \leq \int f_n d\mu \leq \int f d\mu$．$n \to \infty$ として，$\int f d\mu \leq \liminf_n \int f_n d\mu \leq \limsup_n \int f_n d\mu \leq \int f d\mu$． □

MCT の重要な応用として，次の系を証明しよう．

Corollary 4.1. 可測関数 $f_n : X \to [0, \infty], n \in \mathbb{N}$ に対して，

$$\int \sum_{n=1}^{\infty} f_n d\mu = \sum_{n=1}^{\infty} \int f_n d\mu.$$

Proof. $g_n = \sum_{m=1}^{n} f_m$ とおくと，$0 \leq g_n \uparrow \sum_{m=1}^{\infty} f_m$ であるから，MCT より，

$$\sum_{m=1}^{n} \int f_m d\mu = \int g_n d\mu \uparrow \int \sum_{m=1}^{\infty} f_m d\mu.$$

□

つまり非負可測関数に対しては，可算無限和と積分は自由に交換できる．

Theorem 4.5 (Fatou の補題)**.** 可測関数 $f_n : X \to [0, \infty], n \in \mathbb{N}$ に対して，

$$\int \liminf_n f_n d\mu \leq \liminf_n \int f_n d\mu.$$

Proof. $g_n(x) := \inf\{f_m(x) : m \geq n\}$ とおくと，$g_n \uparrow \liminf_n f_n =: f$ であるから，MCT より，$\int f d\mu = \lim_n \int g_n d\mu$．$m \geq n$ に対して，$g_n \leq f_m$ であるから，$\int g_n d\mu \leq \int f_m d\mu$．右辺において $m \geq n$ に関する inf をとり $n \to \infty$ として，$\int f d\mu = \lim_n \int g_n d\mu \leq \liminf_n \int f_n d\mu$ を得る． □

Remark 4.2. MCT と Fatou の補題は 非負 可測関数列に対して成り立つ．

Theorem 4.6 (優収束定理 (DCT))**.** $g, f_n \in L^1, n \in \mathbb{N}$ とし，$|f_n(x)| \leq g(x)$ $(\forall x \in X, \forall n \in \mathbb{N})$, $f_n(x) \to f(x)$ $(\forall x \in X)$ とする．このとき，$f \in L^1$ であって，

$$\int f_n d\mu \to \int f d\mu.$$

Proof. $|f| \leq g$ より $f \in L^1$ である．一方，$g + f_n \geq 0$ であるから，Fatou の補題より，$\int (g+f)d\mu \leq \liminf_n \int (g+f_n)d\mu = \int g d\mu + \liminf_n \int f_n d\mu$．従って，$\int f d\mu \leq \liminf_n \int f_n d\mu$．同様にして，$\int f d\mu \geq \limsup_n \int f_n d\mu$． □

Example 4.3 (“悪い関数列” の例)．$(X, \mathcal{A}, \mu) = (\mathbb{R}, \mathcal{B}, \lambda)$ とし，次の関数列を考える．

(a) $f_n = n^{-1} 1_{(0,n)}$.

(b) $f_n = 1_{(n,n+1)}$.

(c) $f_n = n 1_{[0,1/n]}$.

(d) $f_n = 1_{[(n-2^k)/2^k, (n-2^k+1)/2^k]},\ k \geq 0, 2^k \leq n < 2^{k+1}$.

(a)–(c) では，それぞれ $f_n \to 0$ (一様，各点，a.e.) だが，$\int f_n dx = 1$ である．(d) の例では，$\int f_n dx = 2^{-k},\ 2^k \leq n < 2^{k+1}$ なので，$\int f_n dx \to 0$ だが，各 $x \in [0,1]$ に対して，$f_n(x) = 0$ となる n と $f_n(x) = 1$ となる n がそれぞれ無限個あるので，$f_n(x)$ は収束しない．

Theorem 4.7 (積分と微分の交換)**.** 関数 $f : X \times (a,b) \to \mathbb{R}$ を，$t \in (a,b)$ を任意に固定すると，$f(\cdot, t) \in L^1$ であって，$x \in X$ を任意に固定すると，$f(x,t)$ は t に関して微分可能とする．このとき，ある関数 $g \in L^1$ が存在して，$|\partial f(x,t)/\partial t| \leq g(x)$ $(\forall (x,t) \in X \times (a,b))$ なら，$t \mapsto \int f(x,t) d\mu(x)$ は微分可能であって，

$$\frac{d}{dt} \int f(x,t) d\mu(x) = \int \frac{\partial f(x,t)}{\partial t} d\mu(x),\ t \in (a,b).$$

Proof. $h_n \to 0, h_n \neq 0$ とする．平均値の定理より，

$$\exists \theta = \theta_{x,t,n} \in [0,1] \text{ s.t. } \frac{f(x,t+h_n) - f(x,t)}{h_n} = \frac{\partial f(x, t+\theta h_n)}{\partial t}$$

であるから，$|\{f(x,t+h_n) - f(x,t)\}/h_n| \leq g(x)$．よって，DCT より，

$$\begin{aligned} &\lim_n \frac{1}{h_n} \left\{ \int f(x,t+h_n) d\mu(x) - \int f(x,t) d\mu(x) \right\} \\ &\quad = \lim_n \int \frac{f(x,t+h_n) - f(x,t)}{h_n} d\mu(x) = \int \frac{\partial f(x,t)}{\partial t} d\mu(x) \end{aligned}$$

となる． □

Lebesgue 積分と Riemann 積分の関係．$f : [a,b] \to \mathbb{R}$ を有界な関数とする．まず Riemann 積分の定義を復習する．$a = t_0 < t_1 < \cdots < t_n = b$ に対して，$\Delta = \{t_j\}_{j=0}^n$ を $[a,b]$ の分割と呼び，各 j に対して，$M_j = \sup_{x \in [t_{j-1}, t_j]} f(x), m_j = \inf_{x \in [t_{j-1}, t_j]} f(x)$ とおく．

$$S_\Delta f = \sum_{j=1}^n M_j (t_j - t_{j-1}),\ s_\Delta f = \sum_{j=1}^n m_j (t_j - t_{j-1})$$

をそれぞれ過剰和，不足和と呼ぶ．ここで，

$$\inf_{\Delta} S_\Delta f = \sup_{\Delta} s_\Delta f$$

が成り立つとき，f は Riemann 積分可能であると言って，上式の共通の値を $\int_a^b f(x)dx$ と書く．$\mathcal{L}$ を $\mathbb{R}$ の Lebesgue 可測集合の全体とする．$\mathcal{L}_{[a,b]} = \{E \cap [a,b] : E \in \mathcal{L}\}$ である．

Theorem 4.8. $f : [a,b] \to \mathbb{R}$ が Riemann 積分可能な有界関数なら，f は $\mathcal{L}_{[a,b]}/\mathcal{B}$ 可測であって，$\int_a^b f(x)dx = \int_{[a,b]} f d\lambda$ (左辺は Riemann 積分).

Proof. $G_\Delta = \sum_{j=1}^n M_j 1_{(t_{j-1},t_j]}, g_\Delta = \sum_{j=1}^n m_j 1_{(t_{j-1},t_j]}$ とおくと，過剰和，不足和はそれぞれ $S_\Delta f = \int_{[a,b]} G_\Delta d\lambda, s_\Delta f = \int_{[a,b]} g_\Delta d\lambda$ と表せる．f は Riemann 積分可能であるから，$[a,b]$ の分割の適当な増加列 $\{\Delta_k\}$ (i.e., $\Delta_k \subset \Delta_{k+1}$) が存在して，$\lim_k S_{\Delta_k} f = \lim_k s_{\Delta_k} f = \int_a^b f(x)dx$ となる．このとき，$G_{\Delta_k}, g_{\Delta_k}$ はそれぞれ k に関して非増加，非減少であるから，$G = \lim_k G_{\Delta_k}, g = \lim_k g_{\Delta_k}$ が存在する．いま，$g \le f1_{(a,b]} \le G$ であって，$\{g_{\Delta_k}\}$ と $\{G_{\Delta_k}\}$ はそれぞれ一様有界であるから，DCT より，$\int_{[a,b]} G d\lambda = \int_{[a,b]} g d\lambda = \int_a^b f(x)dx$ を得る．従って，$\int_{[a,b]} (G-g)d\lambda = 0$ であるから，$G = g = f$ λ-a.e. を得る (cf. Exercise 4.2)．ここで，G は単関数の極限なので $\mathcal{L}_{[a,b]}/\mathcal{B}$ 可測であって，$\mathcal{L}_{[a,b]}$ の完備性より，f も $\mathcal{L}_{[a,b]}/\mathcal{B}$ 可測である (cf. Exercise 3.7)．またこのことから，$\int_{[a,b]} f d\lambda = \int_a^b f(x)dx$ も得る． □

演習問題

Exercise 4.1. $\mathcal{A}^\mu$ を $\mathcal{A}$ の μ に関する完備化とすると，任意の $\mathcal{A}^\mu$ 可測関数 $f : X \to \overline{\mathbb{R}}$ に対して，ある $\mathcal{A}$ 可測関数 $g : X \to \overline{\mathbb{R}}$ が存在して，$f = g$ a.e. となることを示せ．

Exercise 4.2. $f : X \to [0,\infty]$ を可測とし，$\int f d\mu = 0$ とする．このとき，$f = 0$ a.e. となることを示せ．

Exercise 4.3. 複素数値の可積分関数 $f : X \to \mathbb{C}$ に対して，$|\int f d\mu| \le \int |f| d\mu$ を示せ．

Exercise 4.4. $f, g \in L^1$ に対して，$\int_A f d\mu = \int_A g d\mu$ $(\forall A \in \mathcal{A})$ とする．このとき，$f = g$ a.e. となることを示せ．

Exercise 4.5. $1 \le r, p_1, \dots, p_k < \infty, 1/p_1 + \cdots + 1/p_k = 1/r$ とすると，$f_1 \in L^{p_1}, \dots, f_k \in L^{p_k}$ に対して，$\|f_1 \cdots f_k\|_r \le \|f_1\|_{p_1} \cdots \|f_k\|_{p_k}$ を示せ．

Exercise 4.6 (Minkowski の不等式)**.** $1 \le p < \infty, f, g \in L^p$ に対して，$\|f+g\|_p \le \|f\|_p + \|g\|_p$ を示せ．また，a.e. に等しい関数を同一視すれば，$(L^p, \|\cdot\|_p)$ は実ノルム空間になることを示せ．このことから，$\|\cdot\|_p$ は $\underline{L^p \text{ ノルム}}$ と呼ばれる．

Exercise 4.7. $0 < p < 1$ のときは，$f \mapsto (\int |f|^p d\mu)^{1/p}$ は三角不等式をみたさないので，ノルムにならない．しかし，$d_p(f,g) = \int |f-g|^p d\mu$ として，a.e. に等しい関数を同一視すれば，(L^p, d_p) は距離空間になることを示せ．

$p=\infty$ に対応する L^p 空間を定義する．可測関数 $f: X \to \mathbb{R}$ に対して，

$$\|f\|_\infty := \inf\{a \geq 0 : \mu(\{x : |f(x)| > a\}) = 0\} \quad (\inf \varnothing = \infty)$$

と定義する．すなわち，$\|f\|_\infty$ は，$|f| \leq a$ a.e. となる定数 a の下限である．$\|f\|_\infty < \infty$ なら，右辺の下限は達成される：$|f| \leq \|f\|_\infty$ a.e. なぜなら，$a = \|f\|_\infty$ とおくと，$\{x : |f(x)| > a\} = \bigcup_n \{x : |f(x)| > a + n^{-1}\}$ であって，定義より右辺の μ 測度は 0 であるからである．そこで，

$$L^\infty := L^\infty(X, \mathcal{A}, \mu) := \{f : f : X \to \mathbb{R},\ f \text{ は可測},\ \|f\|_\infty < \infty\}$$

と定義する．$f \in L^\infty$ なら，$g = f1_{\{|f| \leq \|f\|_\infty\}}$ とおくと，$|g(x)| \leq \|f\|_\infty\ \forall x \in X$ であって，$f = g$ a.e. をみたす．すなわち，L^∞ は有界可測関数に a.e. に等しい関数からなる集合である．

Exercise 4.8. a.e. に等しい関数を同一視すれば，$(L^\infty, \|\cdot\|_\infty)$ がノルム空間になることを示せ．さらに，$\|f_n - f\|_\infty \to 0$ となるためには，$\mu(E^c) = 0$ なる集合 $E \in \mathcal{A}$ が存在して，$\sup_{x \in E} |f_n(x) - f(x)| \to 0$ となることが必要十分であることを示せ．

Exercise 4.9. $f: X \to \mathbb{R}$ を可測関数とし，ある $0 < p < \infty$ に対して，$f \in L^p \cap L^\infty$ とする．このとき，$\|f\|_\infty = \lim_{q \to \infty} \|f\|_q$ を示せ．

Exercise 4.10. $0 < p < q \leq \infty$ とする．

(a) $\mu(X) < \infty$ なら，$L^q \subset L^p$ となることを示せ．

(b) $\mu(X) = \infty$ のときは，一般には $L^q \subset L^p$ でも $L^p \subset L^q$ でもない．$(X, \mathcal{A}, \mu) = (\mathbb{R}, \mathcal{B}, \lambda)$ のとき，次をみたす可測関数 $f, g : \mathbb{R} \to \mathbb{R}$ を構成せよ：$f \in L^p \setminus L^q, g \in L^q \setminus L^p$.

Exercise 4.11. μ を $\mathbb{N}$ 上の計数測度とし，$0 < p \leq \infty$ に対して，$\ell^p := L^p(\mathbb{N}, 2^{\mathbb{N}}, \mu)$ と定義する．すなわち，

$$\ell^p = \begin{cases} \{\{x_n\}_{n=1}^\infty : \sum_{n=1}^\infty |x_n|^p < \infty\} & \text{if } 0 < p < \infty \\ \{\{x_n\}_{n=1}^\infty : \sup_n |x_n| < \infty\} & \text{if } p = \infty \end{cases}$$

である．$0 < p < q \leq \infty$ に対して，$\ell^p \subsetneq \ell^q$ を示せ．

Exercise 4.12. 可測関数 $f: X \to [0, \infty]$ に対して，写像 $\nu : \mathcal{A} \to [0, \infty]$ を

$$\nu(A) = \int_A f d\mu,\ A \in \mathcal{A}$$

と定義する．

(a) ν は $(X, \mathcal{A})$ 上の測度であることを示せ．

(b) 可測関数 $g : X \to \overline{\mathbb{R}}$ に対して，$\int g d\nu$ と $\int g f d\mu$ のどちらかが定義されるならもう一方も定義されて，両者は等しいことを示せ．

Exercise 4.13. DCT を使って次の等式を示せ：

$$\lim_{n\to\infty} \int_0^n \left(1 - \frac{x}{n}\right)^n dx = 1.$$

Exercise 4.14. $f_n \to 0, \int f_n dx \to 0$ となるが，$|f_n| \le g\ (\forall n)$ をみたす $g \in L^1(\mathbb{R}, \lambda)$ が存在しない $\mathbb{R}$ 上の可測関数列 $\{f_n\}$ の例を構成せよ．

Exercise 4.15 (DCT の一般化)**.** $f_n, g_n \in L^1\ (n = 1, 2, \dots)$ とし，$|f_n| \le g_n\ (\forall n),\ f_n(x) \to f(x),\ g_n(x) \to g(x)\ (\forall x \in X)$ とする．このとき，$\int g_n d\mu \to \int g d\mu < \infty$ なら，$\int f_n d\mu \to \int f d\mu$ となることを示せ．

Exercise 4.16. $f \in L^1$ とする．このとき，任意の $\varepsilon > 0$ に対して，ある $\delta > 0$ が存在して，$A \in \mathcal{A}, \mu(A) < \delta$ なら，$\int_A |f| d\mu < \varepsilon$ となることを示せ．

Exercise 4.17. $f(x) = (\sin x)/x\ (x \neq 0)$ とおくと，$\lim_{x\to 0, x\neq 0} f(x) = 1$ より，f は $\mathbb{R}$ 上の連続関数に拡張できる．このとき，Lebesgue 積分の意味で $\int_{\mathbb{R}} f(x)dx$ を定義できるか．

Exercise 4.18. $(0,1) \cap \mathbb{Q} = \{q_1, q_2, \dots\}$ とし，$0 < \alpha < 1$ とする．このとき，$\sum_{n=1}^{\infty} 2^{-n} |x - q_n|^{-\alpha}$ は λ-a.e. $x \in (0,1)$ に対して収束することを示せ．

Exercise 4.19 (Egorov)**.** μ を有限測度，$f, f_n : X \to \mathbb{R}, n = 1, 2, \dots$ を可測関数とし，$f_n(x) \to f(x)\ \forall x \in X$ とする．このとき，任意の $\varepsilon > 0$ に対して，次をみたす集合 $E \in \mathcal{A}$ が存在することを示せ：$\mu(E) \le \varepsilon\ \&\ \sup_{x \in E^c} |f_n(x) - f(x)| \to 0$.

ヒント：$f \equiv 0$ としてよい．$E_{m,n} = \bigcup_{k=n}^{\infty} \{|f_k| > 1/m\}$ とおくと，$\forall m\ \exists n(m)$ s.t. $\mu(E_{m,n(m)}) \le \varepsilon/2^m$. そこで $E = \bigcup_{m=1}^{\infty} E_{m,n(m)}$ とせよ．

Exercise 4.20. Egorov の定理は μ が Lebesgue 測度のときはそのままでは成り立たない．反例を構成せよ．しかし，“$\sup_{x\in E^c} |f_n(x) - f(x)| \to 0$” の部分を “$\sup_{x \in [-A,A] \cap E^c} |f_n(x) - f(x)| \to 0\ (\forall A > 0)$” とすれば正しい．これを示せ．

ヒント：前問のヒントにおいて，$E_{m,n} = (\bigcup_{k=n}^{\infty} \{|f_k| > 1/m\}) \cap [-m, m]$ とせよ．

以下，$f : \mathbb{R} \to \mathbb{R}$ を Lebesgue 測度に関して可積分な関数とする．ある定数 $c_1, \dots, c_k \in \mathbb{R}$ と $a_0 < a_1 < \cdots < a_k$ が存在して，$q(x) = \sum_{j=1}^{k} c_j 1_{(a_{j-1}, a_j)}(x)$ の形に表される関数をステップ関数と呼ぶ．

Exercise 4.21. $\varepsilon > 0$ を given とする．

(a) 次をみたすステップ関数 q が存在することを示せ：$\int |f - q| dx < \varepsilon$.

ヒント：まず f を単関数で近似する．単関数表示に現れる各集合を Lemma 2.1 より開区間の和で近似する．

(b) 有界区間の外では 0 となる連続関数 $g : \mathbb{R} \to \mathbb{R}$ が存在して，$\int |f - g| dx < \varepsilon$ となることを示せ．

Remark 4.3 (Littlwood の 3 原理)**.** Littlewood の 3 原理 とは，おおまかに言って，次の (i)–(iii) のことである．

(i) $\mathbb{R}$ 上の Lebesgue 可測集合はだいたい区間の有限和である．

(ii) Lebesgue 測度に関する可積分関数はだいたい連続関数である．

(iii) 各点収束する関数列はだいたい一様収束する．

第 1 原理は Lemma 2.1 に対応する (各 Lebesgue 可測集合に対して，Lebesgue 測度 0 の差しかない Borel 集合が存在するので，Lemma 2.1 の A は Lebesgue 可測集合としても正しい)．第 2 原理は Exercise 4.21 (b) に対応する．第 3 原理は Egorov の定理 (cf. Exercise 4.19 & Exercise 4.20) に対応する．

Exercise 4.21 より，f の積分はステップ関数や連続関数の積分で近似できることがわかる．このような近似にもとづいて，f の積分の性質を調べることができる．

Exercise 4.22. (Riemann-Lebesgue の補題)

$$\lim_{n\to\infty} \int f(x) \cos(nx) dx = 0$$

を示せ．

Exercise 4.23.

$$\lim_{h\to 0} \int_{\mathbb{R}} |f(x+h) - f(x)| dx = 0$$

を示せ．

次の Lusin の定理も Littlewood の第 2 原理のバージョンと言える．

Exercise 4.24 (Lusin)**.** 任意の $\varepsilon > 0$ に対して，次をみたす集合 $E \in \mathcal{B}$ が存在することを示せ：$\lambda(E) \leq \varepsilon$ & f の E^c への制限 $f|_{E^c}$ が連続になる．

ヒント:各 n に対して，$\int |f - f_n| dx \leq \varepsilon/4^n$ なる連続関数 f_n が存在する．$E_n = \{|f - f_n| > 1/2^n\}$ とおくと，$\lambda(E_n) \leq \varepsilon/2^n$ である．あとは $E = \bigcup_{n=1}^{\infty} E_n$ とせよ．

Lusin の定理はもともとの f が E^c 上で連続であることを意味するわけではない．例えば，$1_{\mathbb{Q}}$ は $\mathbb{R}$ 上の関数としていたるところ不連続であるが，$\mathbb{R} \setminus \mathbb{Q}$ への制限は $(1_{\mathbb{Q}})|_{\mathbb{R}\setminus\mathbb{Q}} \equiv 0$ だから連続である．

5　直積測度

$(X,\mathcal{A},\mu),(Y,\mathcal{C},\nu)$ を σ-finite な測度空間とする．$\mathcal{D}=\{A\times C : A\in\mathcal{A}, C\in\mathcal{C}\}$ において，$\rho:\mathcal{D}\to[0,\infty]$ を

$$\rho(A\times C):=\mu(A)\nu(C),\ A\in\mathcal{A}, C\in\mathcal{C}$$

と定義する．ρ を $\mathcal{A}\times\mathcal{C}:=\sigma(\mathcal{D})$ 上の測度に拡張したい．

まず ρ が $\mathcal{D}$ 上で可算加法的になることを確認する．すなわち，排反な $A_n\times C_n\in\mathcal{D}, n\in\mathbb{N}$ に対して，$\bigcup_{n=1}^{\infty}(A_n\times C_n)\in\mathcal{D}$ なら，$\rho(\bigcup_{n=1}^{\infty}(A_n\times C_n))=\sum_{n=1}^{\infty}\rho(A_n\times C_n)$ となることを示す．ここで，$\bigcup_{n=1}^{\infty}(A_n\times C_n)\in\mathcal{A}\times\mathcal{C}$ より，ある $A\in\mathcal{A}, C\in\mathcal{C}$ が存在して，$\bigcup_{n=1}^{\infty}(A_n\times C_n)=A\times C$ と表せる．$A_n\times C_n, n\in\mathbb{N}$ は排反であるから，

$$1_A(x)1_C(y)=1_{A\times C}(x,y)=\sum_{n=1}^{\infty}1_{A_n\times C_n}(x,y)=\sum_{n=1}^{\infty}1_{A_n}(x)1_{C_n}(y).$$

両辺を y に関して積分すると，MCT より，

$$1_A(x)\nu(C)=\sum_{n=1}^{\infty}1_{A_n}(x)\nu(C_n).$$

次に，x について積分すると，再び MCT より，

$$\mu(A)\nu(C)=\sum_{n=1}^{\infty}\mu(A_n)\nu(C_n).$$

左辺 $=\rho(A\times C)$, 右辺 $=\sum_{n=1}^{\infty}\rho(A_n\times C_n)$ であるから，求める主張を得る．

あとの議論は Theorem 2.1 の証明と同様である．まず $\mathcal{D}$ は field でないが，

$$\overline{\mathcal{D}}=\left\{\bigcup_{m=1}^{n}D_m :\ D_1,\dots,D_n\in\mathcal{D}\text{ は排反}, n\in\mathbb{N}\right\}$$

は field であって，$\sigma(\overline{\mathcal{D}})=\sigma(\mathcal{D})=\mathcal{A}\times\mathcal{C}$ である．ρ を $\overline{\mathcal{D}}$ 上に

$$\rho\left(\bigcup_{m=1}^{n}D_m\right):=\sum_{m=1}^{n}\rho(D_m),\quad D_1,\dots,D_n\in\mathcal{D}\text{ は排反}$$

と拡張すれば，ρ は well-defined であって，$\overline{\mathcal{D}}$ 上で可算加法的であることが示せる．従って，Carathéodory の拡張定理より，次の定理を得る．

Theorem 5.1. ρ は $\mathcal{A}\times\mathcal{C}$ 上の測度に一意に拡張できる．

こうして構成された $\mathcal{A}\times\mathcal{C}$ 上の測度 ρ を μ と ν の 直積測度 (product measure) と呼び，$\rho=\mu\times\nu$ と書く．

この操作を繰り返せば，σ-finite な測度空間 $(X_i,\mathcal{A}_i,\mu_i), 1\le i\le n$ に対して，$\mathcal{A}_1\times\cdots\times\mathcal{A}_n$ 上の測度 ρ であって，

$$\rho(A_1\times\cdots\times A_n)=\mu_1(A_1)\cdots\mu(A_n),\ A_1\in\mathcal{A}_1,\ldots,A_n\in\mathcal{A}_n$$

をみたすものが一意に存在することが示せる．$\rho=\mu_1\times\cdots\times\mu_n=\prod_{i=1}^n\mu_i$ と書く．

$(X_i,\mathcal{A}_i)=(\mathbb{R},\mathcal{B}),\mu_i=\lambda$ のとき，$(\mathbb{R}^n,\mathcal{B}^n)$ 上の直積測度 $\lambda^n:=\lambda\times\cdots\times\lambda$ を n 次元 Lebesgue 測度と呼ぶ．

$$d\lambda^n(\boldsymbol{x})=d\boldsymbol{x}=dx_1\cdots dx_n,\ \boldsymbol{x}=(x_1,\ldots,x_n)$$

と書く．

Remark 5.1. $\mathcal{B}^n$ の λ^n に関する完備化を $\mathcal{L}(\mathbb{R}^n)$ とおくと，$\mathcal{L}(\mathbb{R}^n)$ に属する集合を n 次元 Lebesgue 集合と呼ぶ．注意として，

$$\mathcal{L}^n:=\mathcal{L}\times\cdots\times\mathcal{L}\subsetneq\mathcal{L}(\mathbb{R}^n)$$

である．$\mathcal{L}^n\subset\mathcal{L}(\mathbb{R}^n)$ はよい．また，例えば，$n=2$ のとき，$V\subset\mathbb{R}$ を Vitali 集合として，$A=V\times\varnothing$ とおくと，$A\notin\mathcal{L}^2$ であるが，$A\subset\mathbb{R}\times\varnothing,\lambda^2(\mathbb{R}\times\varnothing)=0$ より，$A\in\mathcal{L}(\mathbb{R}^2)$ である．ただし，$\mathcal{L}(\mathbb{R}^n)$ はこのあとほとんど登場しない (特に断わらない限り，$\mathbb{R}^n$ には $\mathcal{B}^n$ を入れる)．

Fubini の定理

$n=2$ の場合に戻り，$\mathcal{D}=\{A\times C:A\in\mathcal{A},C\in\mathcal{C}\},\ \rho=\mu\times\nu$ とする．μ,ν は σ-finite と仮定している．適当な条件の下で，直積測度に関する積分が逐次積分に等しいことを保証する定理 (群) を Fubini の定理 と呼ぶ．まず指示関数の積分を考える．

Theorem 5.2 (Fubini．指示関数の場合)**.** 任意の $E\in\mathcal{A}\times\mathcal{C}$ に対して，(i) $y\in Y$ を任意に固定すると，$1_E(x,y)$ は x の関数として $\mathcal{A}$ 可測，(ii) $y\mapsto\int 1_E(x,y)d\mu(x)$ は $\mathcal{C}$ 可測，(iii) $\rho(E)=\iint 1_E(x,y)d\mu(x)d\nu(y)$.

Remark 5.2. 議論は対称なので，(i)' $x\in X$ を任意に固定すると，$1_E(x,y)$ は y の関数として $\mathcal{C}$ 可測，(ii)' $x\mapsto\int 1_E(x,y)d\nu(y)$ は $\mathcal{A}$ 可測，(iii)' $\rho(E)=\iint 1_E(x,y)d\nu(y)d\mu(x)$，も成り立つ．特に，任意の $E\in\mathcal{A}\times\mathcal{C}$ に対して，

$$\iint 1_E(x,y)d\mu(x)d\nu(y)=\iint 1_E(x,y)d\nu(y)d\mu(x)$$

となる．Theorem 5.3 と Theorem 5.4 についても同様の注意が当てはまる．

Proof. μ, ν は σ-finite であるから，$\mu(X) < \infty, \nu(Y) < \infty$ と仮定して定理を証明すれば十分である．$\mathcal{E}$ を (i)–(iii) が成り立つ $E \in \mathcal{A} \times \mathcal{C}$ の全体とする．π-λ 定理を使って，$\mathcal{E} = \mathcal{A} \times \mathcal{C}$ を示す．まず $\mathcal{D}$ が π システムであることと，$\mathcal{D} \subset \mathcal{E}$ はよい．従って，$\mathcal{E}$ が λ システムであることを示せば，$\mathcal{E} \subset \sigma(\mathcal{D}) = \mathcal{A} \times \mathcal{C}$, すなわち，$\mathcal{E} = \mathcal{A} \times \mathcal{C}$ が従う．

Step 1. $X \times Y \in \mathcal{E}$ はよい．

Step 2. $E, F \in \mathcal{E}, E \subset F$ なら，$F \setminus E \in \mathcal{E}$ であることを示す．$1_{F \setminus E} = 1_F - 1_E$ であるから，(i) はよい．次に，$\int 1_{F \setminus E}(x, y) d\mu(x) = \int 1_F(x, y) d\mu(x) - \int 1_E(x, y) d\mu(x)$ であるから，(ii) もよい (ここで，$\mu(X) < \infty$ より，$\infty - \infty$ は起こらない)．最後に，$\iint 1_{F \setminus E}(x, y) d\mu(x) d\nu(y) = \iint 1_F(x, y) d\mu(x) d\nu(y) - \iint 1_E(x, y) d\mu(x) d\nu(y) = \rho(F) - \rho(E) = \rho(F \setminus E)$ であるから，(iii) もよい (再び，$\nu(Y) < \infty$ より，$\infty - \infty$ は起こらない)．

Step 3. $E_n \in \mathcal{E}, E_n \uparrow E$ なら，$E \in \mathcal{E}$ であることを示す．$1_{E_n} \uparrow 1_E$ より，(i) はよい．次に，MCT より，$\int 1_{E_n}(x, y) d\mu(x) \uparrow \int 1_E(x, y) d\mu(x)$ であるから，(ii) もよい．最後に，再び MCT より，$\int \{\int 1_{E_n}(x, y) d\mu(x)\} d\nu(y) \uparrow \int \{\int 1_E(x, y) d\mu(x)\} d\nu(y)$. $E_n \in \mathcal{E}$ より，左辺 $= \rho(E_n) \uparrow \rho(E)$ であるから，$\rho(E) = \iint 1_E(x, y) d\mu(x) d\nu(y)$ を得る． □

次に非負可測関数の積分を考える．

Theorem 5.3 (Fubini. 非負可測関数の場合)**.** 非負可測関数 $f : X \times Y \to [0, \infty]$ に対して，(i) $y \in Y$ を任意に固定すると，$f(x, y)$ は x の関数として $\mathcal{A}$ 可測，(ii) $y \mapsto \int f(x, y) d\mu(x)$ は $\mathcal{C}$ 可測，(iii) $\int f d\rho = \iint f(x, y) d\mu(x) d\nu(y)$.

Remark 5.3 (**重要!**)**.** 非負可測関数に対しては，直積測度に関する積分が常に逐次積分に一致する．また，対称性より，非負可測関数の逐次積分の順序は自由に交換できる．この事実はあとで頻繁に用いる．

Proof. Theorem 5.2 より，非負単関数に対して定理が成り立つことがわかる．一般の非負可測関数 f に対しては，非負単関数列で，$\uparrow f$ となるものとればよい． □

最後に一般の可測関数に対する Fubini の定理を述べる．

Theorem 5.4 (Fubini. 一般の可測関数の場合)**.** $f \in L^1(X \times Y, \mathcal{A} \times \mathcal{C}, \rho)$ に対して，(i) $y \in Y$ を任意に固定すると，$f(x, y)$ は x の関数として $\mathcal{A}$ 可測，(ii) $y \mapsto \int f(x, y) d\mu(x)$ は ν-a.e. に定義された $\mathcal{C}$ 可測関数であって，(iii) $\int f d\rho = \iint f(x, y) d\mu(x) d\nu(y)$.

Proof. $f = f^+ - f^-$ と分解する．(i) はよい．(ii) は Theorem 5.3 より，

$$y \mapsto \int f^+(x, y) d\mu(x), \ y \mapsto \int f^-(x, y) d\mu(x)$$

は $\mathcal{A}$ 可測であって，$f \in L^1(X \times Y, \mathcal{A} \times \mathcal{C}, \rho)$ より，

$$\iint f^+(x, y) d\mu(x) d\nu(y) = \int f^+ d\rho < \infty, \ \iint f^-(x, y) d\mu(x) d\nu(y) = \int f^- d\rho < \infty$$

である．従って，

$$y \mapsto \int f(x,y)d\mu(x) = \int f^+(x,y)d\mu(x) - \int f^-(x,y)d\mu(x)$$

は ν-a.e. に定義された $\mathcal{C}$ 可測関数である．実際，

$$F_1(y) = \int f^+(x,y)d\mu(x),\ F_2(y) = \int f^-(x,y)d\mu(x)$$

とおくと，$F_1 < \infty, F_2 < \infty$ ν-a.e. である．そこで，$N = \{F_1 = \infty\} \cap \{F_2 = \infty\}$ とおくと，$\nu(N) = 0$ であって，$\tilde{F}_i(y) = F_i(y)1_{Y\setminus N}(y), i = 1, 2$ と定義すれば，各 $\tilde{F}_i$ は $\mathcal{C}$ 可測であるから，$\tilde{F}_1 - \tilde{F}_2$ も $\mathcal{C}$ 可測である．$F_1 - F_2$ は $Y \setminus N$ 上で定義されていて，その上で $\tilde{F}_1 - \tilde{F}_2$ に一致するから，ν-a.e. に定義された $\mathcal{C}$ 可測関数であることが確かめられた．最後に，

$$\begin{aligned}\iint f(x,y)d\mu(x)d\nu(y) &= \iint f^+(x,y)d\mu(x)d\nu(y) - \iint f^-(x,y)d\mu(x)d\nu(y)\\ &= \int f^+d\rho - \int f^-d\rho = \int fd\rho\end{aligned}$$

より，(iii) もよい． □

Theorem 5.4 より，$f \in L^1(X \times Y, \mathcal{A} \times \mathcal{C}, \rho)$ に対して，

$$\iint f(x,y)d\mu(x)d\nu(y) = \iint f(x,y)d\nu(y)d\mu(x)$$

が成り立つ．$f \in L^1(X \times Y, \mathcal{A} \times \mathcal{C}, \rho)$ を確認するには，$|f|$ は非負なので，逐次積分

$$\iint |f(x,y)|d\mu(x)d\nu(y), \iint |f(x,y)|d\nu(y)d\mu(x)$$

のうちどちらか評価しやすい方を計算して，それが有限であることを示せばよい．

Fubini の定理において，μ, ν が σ-finite という仮定は本質的である．

Example 5.1. $X = (0,1), \mathcal{A} = \mathcal{B}_{(0,1)}, \mu = \lambda|_{\mathcal{B}_{(0,1)}}, Y = (0,1), \mathcal{C} = 2^Y, \nu(C) = \mathrm{Card}(C), C \in \mathcal{C}$ とする．ν は σ-finite でない．

$$f(x,y) = \begin{cases} 1 & \text{if } x = y \\ 0 & \text{if } x \neq y \end{cases}$$

とおくと，f は $\mathcal{A} \times \mathcal{C}$ 可測である．ところが，

$$\int_Y f(x,y)d\nu(y) = \mathrm{Card}(\{x\}) = 1 \text{ より，} \int_X \int_Y f(x,y)d\nu(y)d\mu(x) = 1.$$

一方，

$$\int_X f(x,y)d\mu(x) = \lambda(\{y\}) = 0 \text{ より，} \int_Y \int_X f(x,y)d\mu(x)d\nu(y) = 0.$$

$\mathbb{R}^n$ における変数変換

$U \subset \mathbb{R}^n$ を開集合とすると，写像 $\varphi = (\varphi_1, \ldots, \varphi_n) : U \to \varphi(U) \subset \mathbb{R}^n$ が 微分同相 (diffeomorphism) であるとは，φ が 1 対 1 な C^1 級写像であって，その Jacobi 行列

$$(D\varphi)(\boldsymbol{x}) = (\partial\varphi_i(\boldsymbol{x})/\partial x_j)_{1\le i,j\le n} = \begin{pmatrix} \partial\varphi_1(\boldsymbol{x})/\partial x_1 & \cdots & \partial\varphi_1(\boldsymbol{x})/\partial x_n \\ \vdots & \ddots & \vdots \\ \partial\varphi_n(\boldsymbol{x})/\partial x_1 & \cdots & \partial\varphi_n(\boldsymbol{x})/\partial x_n \end{pmatrix}$$

が各 $\boldsymbol{x} \in U$ において正則であることを言う．このとき，逆関数定理より，$\varphi(U)$ は開集合であって，$\varphi^{-1} : \varphi(U) \to U$ も C^1 級である．$(D\varphi)(\boldsymbol{x})$ の行列式

$$J_\varphi(\boldsymbol{x}) := \det(D\varphi)(\boldsymbol{x})$$

を φ の Jacobian と呼ぶ.

Theorem 5.5. $U \subset \mathbb{R}^n$ を開集合とし，$\varphi : U \to \varphi(U) \subset \mathbb{R}^n$ を微分同相とする．このとき，$\mathcal{B}^n_{\varphi(U)}/\mathcal{B}$ 可測関数 $f : \varphi(U) \to \mathbb{R}$ に対して，$f \ge 0$ or $f \in L^1(\varphi(U), \lambda^n)$ なら，

$$\int_{\varphi(U)} f(\boldsymbol{y})d\boldsymbol{y} = \int_U f(\varphi(\boldsymbol{x}))|J_\varphi(\boldsymbol{x})|d\boldsymbol{x}$$

が成り立つ.

この定理の証明は，Folland (1999, Theorem 2.47) を参照せよ．もちろん，Theorem 5.5 は特殊な場合には，Riemann 積分に対する変数変換公式に帰着する.

演習問題

Exercise 5.1. $E \in \mathcal{B}^n$ を $\lambda^n(\mathbb{R}^n \setminus E) = 0$ なる集合とする．このとき，E は $\mathbb{R}^n$ で稠密であることを示せ.

Exercise 5.2. $f : \mathbb{R}^n \to \mathbb{R}_+$ を連続関数とする．このとき，$\int_{\mathbb{R}^n} f(\boldsymbol{x})d\boldsymbol{x} = 0$ なら $f(\boldsymbol{x}) = 0$ $(\forall \boldsymbol{x} \in \mathbb{R}^n)$ であることを示せ.

Exercise 5.3. $(X, \mathcal{A}, \mu), (Y, \mathcal{C}, \nu)$ を σ-finite な測度空間とし，$(X \times Y, \mathcal{A} \times \mathcal{C}, \rho)$ をその直積測度空間とする $(\rho = \mu \times \nu)$. 可測関数 $f : X \times Y \to \overline{\mathbb{R}}$ に対して，次の 3 条件が同値なことを示せ.

(i) $f = 0$ ρ-a.e.

(ii) μ-a.e. x に対して，$f(x, y) = 0$ ν-a.e. y.

(iii) ν-a.e. y に対して，$f(x, y) = 0$ μ-a.e. x.

Exercise 5.4. F, G をそれぞれ $\mathbb{R}$ 上の右連続かつ非減少な関数とし，μ, ν をそれぞれ F, G に対応する Lebesgue-Stieltjes 測度とする．積分 $\int \cdot \, d\mu$ を $\int \cdot \, dF$ とも書くことにする．

(a) 次の等式を示せ．

$$\int_{(a,b]} \{F(y) - F(a)\} dG(y) = (\mu \times \nu)(\{(x, y) : a < x \leq y \leq b\}).$$

(b) 次の等式を示せ．

$$\int_{(a,b]} F(y) dG(y) + \int_{(a,b]} G(x) dF(x) = F(b)G(b) - F(a)G(a) + \sum_{a<x\leq b} \mu(\{x\})\nu(\{x\}).$$

Exercise 5.5. μ を $(\mathbb{R}, \mathcal{B})$ 上の有限測度とし，$F(x) = \mu((-\infty, x])$ とおく．このとき，

$$\int \{F(x + c) - F(x)\} dx = c\mu(\mathbb{R})$$

を示せ．

Exercise 5.6. $a, b > 0$ に対して，

$$\left| \int_0^b \frac{\sin x}{x} e^{-ax} dx - \frac{\pi}{2} \right| \leq \arctan(a) + \delta(b)$$

を示せ．ただし，$\lim_{b\to\infty} \delta(b) = 0$ である．特に，

$$\lim_{b\to\infty} \int_0^b \frac{\sin x}{x} dx = \frac{\pi}{2}, \quad \lim_{a\downarrow 0} \int_0^\infty \frac{\sin x}{x} e^{-ax} dx = \frac{\pi}{2}$$

となる．

ヒント：$e^{-xy} \sin x$ を x と y に関して積分する．

Exercise 5.7. $\boldsymbol{x} \in \mathbb{R}^n, r \in \mathbb{R}$ と $B \in \mathcal{B}^n$ に対して，$B + \boldsymbol{x} := \{\boldsymbol{y} + \boldsymbol{x} : \boldsymbol{y} \in B\}, rB := \{r\boldsymbol{y} : \boldsymbol{y} \in B\} \in \mathcal{B}^n$ を示せ．また，

$$\lambda^n(B + \boldsymbol{x}) = \lambda^n(B), \ \lambda^n(rB) = |r|^n \lambda^n(B)$$

を示せ．$B \in \mathcal{L}(\mathbb{R}^n)$ ならどうか．

Exercise 5.8 (表面測度)**.** $\mathbb{R}^n$ の標準ノルムを $|\boldsymbol{x}| = \sqrt{\sum_{m=1}^n x_m^2}$ と書く．$n \geq 2$ とし，$\mathbb{S}^{n-1} = \{\boldsymbol{x} \in \mathbb{R}^n : |\boldsymbol{x}| = 1\}$ を $\mathbb{R}^n$ の単位球面とする．$\mathbb{R}^n$ から決まる $\mathbb{S}^{n-1}$ の相対位相に関する Borel σ-field を $\mathcal{B}(\mathbb{S}^{n-1})$ とおき，$0 \leq a < b$ と $A \in \mathcal{B}(\mathbb{S}^{n-1})$ に対して，$C_{a,b}(A) = \{r\boldsymbol{x}' : r \in (a, b], \boldsymbol{x}' \in A\}$ とおく．

(a) $C_{a,b}(A) \in \mathcal{B}^n$ を示せ．さらに，$\sigma(A) := n\lambda^n(C_{0,1}(A)), A \in \mathcal{B}(\mathbb{S}^{n-1})$ と定義すると，σ が $(\mathbb{S}^{n-1}, \mathcal{B}(\mathbb{S}^{n-1}))$ 上の有限測度になることを示せ．σ を $\mathbb{S}^{n-1}$ 上の 表面測度 (surface measure) と呼ぶ．

(b) 可測関数 $f : \mathbb{R}^n \to \mathbb{R}$ に対して，$f \geq 0$ or $f \in L^1(\lambda^n)$ なら，

$$\int_{\mathbb{R}^n} f(\boldsymbol{x}) d\boldsymbol{x} = \int_0^\infty \left\{ \int_{\mathbb{S}^{n-1}} f(r\boldsymbol{x}') d\sigma(\boldsymbol{x}') \right\} r^{n-1} dr$$

を示せ．

(c) $\sigma(\mathbb{S}^{n-1}) = 2\pi^{n/2}/\Gamma(n/2)$ を示せ．ただし，$\Gamma(\alpha)$ はガンマ関数である：

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx.$$

Remark 5.4. $\tilde{\sigma}(A) := \sigma(A)/\sigma(\mathbb{S}^{n-1}), A \in \mathcal{B}(\mathbb{S}^{n-1})$ とおけば $\tilde{\sigma}$ は $(\mathbb{S}, \mathcal{B}(\mathbb{S}^{n-1}))$ 上の確率測度 (全測度が 1 の測度) である．$\tilde{\sigma}$ のことを $\mathbb{S}^{n-1}$ 上の一様分布とも呼ぶ．

Exercise 5.9 (Okamoto (1973))**.** $f(x_1, \ldots, x_n)$ を $x_1, \ldots, x_n$ の多項式とし，恒等的に 0 でないとする．このとき，集合

$$N_n = \{(x_1, \ldots, x_n) : f(x_1, \ldots, x_n) = 0\}$$

の n 次元 Lebesgue 測度は 0 であることを示せ：$\lambda^n(N_n) = 0$.

Part II. 確率空間と大数の法則

本パートから確率論に入る．まず，確率論の基本的な概念を一通り紹介したあと，大数の弱法則と強法則を証明する．大数の法則は結果そのものも重要であるが，その証明に使われるテクニック，特に Borel-Cantelli の補題，に習熟することも重要である．

6 確率空間

確率空間 (probability space) $(\Omega, \mathcal{F}, P)$ とは全測度が 1 の測度空間である．すなわち，$(\Omega, \mathcal{F}, P)$ は測度空間であって，$P(\Omega) = 1$ をみたすものである．P を確率測度 (probability measure, p.m.) と呼ぶ．$\mathcal{F}$ に属する集合を事象 (event) とも言う．慣例として，確率空間では，a.e. の代わりに，“a.s.” (almost surely) と書く．次の補題は，Carathéodory の拡張定理の一意性の証明から直ちに従う (π-λ 定理の簡単な，しかし重要な応用である)．

Lemma 6.1. ある π システム上で一致する 2 つの p.m.'s は，その π システムが生成する σ-field 上でも一致する．

可測空間 $(S, \mathcal{S})$ が与えられたとき，可測写像 $X : \Omega \to S$ を S 値確率変数 (S-valued random variable (r.v.)) と呼ぶ．単に r.v. と書いたら $\mathbb{R}$-valued r.v. を意味するものとする．S-valued r.v. X の像測度 $P \circ X^{-1}$ を $\mathcal{L}(X)$ と書いて，X の分布 (distribution) と呼ぶ．$\mathcal{L}(X) = \mathcal{L}(Y)$ のとき，$X \stackrel{d}{=} Y$ と書く．

直積 σ-field の定義より，可測空間 $(S_i, \mathcal{S}_i), 1 \le i \le n$ が与えられたとき，写像

$$X = (X_1, \dots, X_n) : \Omega \to \left(\prod_{i=1}^n S_i, \prod_{i=1}^n \mathcal{S}_i \right)$$

が $(\prod_{i=1}^n S_i)$-valued r.v. であることは，各 X_i が S_i-valued r.v. であることと同値である．証明は，Lemma 3.4 と全く同様である． 念のために，$n = 2$ の場合に確認してみよう．$\mathcal{S}_1 \times \mathcal{S}_2 = \sigma(\{B_1 \times B_2 : B_i \in \mathcal{S}_i, i = 1, 2\}$ であるから，X が $\mathcal{S}_1 \times \mathcal{S}_2$ 可測なことは，$X^{-1}(B_1 \times B_2) \in \mathcal{F}$ $(\forall B_i \in \mathcal{S}_i, i = 1, 2)$ と同値である．ここで，$X^{-1}(B_1 \times B_2) = X_1^{-1}(B_1) \cap X_2^{-1}(B_2)$ であって，これが任意の $B_i \in \mathcal{S}_i$ $(i = 1, 2)$ に対して $\mathcal{F}$ に属していることは，各 X_i が $\mathcal{F}/\mathcal{S}_i$ 可測なことと同値である．

$(X_1, \dots, X_n)$ の分布 $\mathcal{L}(X_1, \dots, X_n)$ を $X_1, \dots, X_n$ の同時分布 (joint distribution) と呼ぶ．特に，各 X_i が $\mathbb{R}$-valued r.v. のとき，$(X_1, \dots, X_n)$ を確率ベクトル (random vector) と呼ぶ．

分布関数. r.v. $X : \Omega \to \mathbb{R}$ に対して，

$$F(x) := P(X \le x),\ x \in \mathbb{R}$$

を X の分布関数 (distribution function, d.f.) と呼ぶ．d.f. F は次の性質をみたす．

(i) F は右連続かつ非減少.

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

さらに，$F(x-) = \lim_{y\uparrow x} F(y) = P(X < x), F(x) - F(x-) = P(X = x)$ である．

(i), (ii) と Theorem 2.1 より，$\mathcal{L}(X)$ は d.f. F から一意に決まる (π-λ 定理を直接適用してもよい)．逆に，関数 $F : \mathbb{R} \to [0,1]$ が (i) と (ii) をみたすなら，Theorem 2.1 より，

$$\mu((a,b]) = F(b) - F(a),\ -\infty \le a < b < \infty$$

をみたす $(\mathbb{R}, \mathcal{B})$ 上の p.m. μ (Lebesgue-Stieltjes 測度) が一意に決まる．そこで，確率空間を $(\Omega, \mathcal{F}, P) = (\mathbb{R}, \mathcal{B}, \mu)$ とし，X を $X(\omega) = \omega, \omega \in \Omega$ とすれば，

$$\mu(X \le x) = \mu((-\infty, x]) = F(x).$$

すなわち，X は F を d.f. にもつ r.v. である．以上の議論より，次の定理を得る．

Theorem 6.1. 関数 $F : \mathbb{R} \to [0,1]$ がある r.v. の d.f. になるための必要十分条件は，それが (i) と (ii) をみたすことである．このとき，F を d.f. にもつ分布が一意に決まる．

以下，多くの場合，分布と d.f. を同一視する．

Theorem 6.1 の証明は，与えられた分布をもつ r.v. を適当な確率空間に具体的に構成している．この議論は μ が一般の可測空間 $(S, \mathcal{S})$ 上の測度であっても成り立つ．すなわち，

- 確率空間を $(\Omega, \mathcal{F}, P) = (S, \mathcal{S}, \mu)$ とし，
- X を $X(\omega) = \omega,\ \omega \in \Omega$ とすれば，

X は S-valued r.v. であって，$\mathcal{L}(X) = \mu$ となる．

与えられた d.f. をもつ r.v. の別の構成法. 与えられた d.f. をもつ r.v. は $(0,1)$ 上の一様乱数を用いて発生させることができる．d.f. F に対して，F の 分位点関数 (quantile function) F^{-1} を

$$F^{-1}(p) := \inf\{x \in \mathbb{R} : F(x) \ge p\}$$

と定義する．F^{-1} は左連続かつ非減少である．分位点関数に関して不案内な場合は，竹村 (1991, p.16-19) を参照せよ．$U(0,1)$ は $(0,1)$ 上の Lebesgue 測度 (一様分布) を表す.

Theorem 6.2. $U \sim U(0,1)$ に対して，$F^{-1}(U) \sim F$.

Proof. $F^{-1}(p) \le x \Leftrightarrow p \le F(x)$ であるから，$P(F^{-1}(U) \le x) = P(U \le F(x)) = F(x)$. □

密度関数. d.f. F がある非負可測関数 $f : \mathbb{R} \to \mathbb{R}_+$ を用いて，

$$F(x) = \int_{-\infty}^{x} f(y)dy \qquad (*)$$

と表せるとき，F を 絶対連続 (absolutely continuous) であると言って，f を F の 密度 (density) と呼ぶ．定義より，

$$\int f(x)dx = 1 \qquad (**)$$

である．このとき，$\mu = \mathcal{L}(X)$ は，

$$\mu(A) = \int_A f(x)dx,\ A \in \mathcal{B}$$

と表せる．F の密度 f は λ-a.e. に一意である．なお，後述する Lebesgue の微分定理 (Theorem 25.4) より，F は λ-a.e. に微分可能であって，$F' = f$ λ-a.e. となる．

逆に，$(**)$ をみたす非負可測関数 $f : \mathbb{R} \to \mathbb{R}_+$ が与えられたとき，F を $(*)$ と定義すれば，F は d.f. になる．$(**)$ をみたす非負可測関数を 確率密度関数 (probability density function) と呼ぶ．

Example 6.1 (連続だが絶対連続でない d.f. の例)．絶対連続な d.f. は連続であるが，その逆は成り立たない．例えば，$F : [0,1] \to [0,1]$ を Cantor 関数とし，F を $F(x) = 0$ $(x < 0)$, $F(x) = 1$ $(x > 1)$ として $\mathbb{R}$ 上の関数に拡張すると，拡張された F は連続な d.f. である．仮に F が絶対連続として，f をその密度とすると，f は $\mathbb{R} \setminus [0,1]$ 上では (a.e.) に 0 である．さらに，Cantor 関数の構成より，$\lambda(U) = 1$ なる開集合 $U \subset [0,1]$ が存在して，F は各 $x \in U$ の十分小さい近傍上で一定となるから，$f = 0$ a.e. $x \in U$ である．ゆえに $f = 0$ a.e. $x \in \mathbb{R}$ となるが，これは $\int_{\mathbb{R}} f(x)dx = 1$ に反する．この例に現れた d.f. F に対応する分布を Cantor 分布 と呼ぶ．

期待値. $\overline{\mathbb{R}}$-valued r.v. $X : \Omega \to \overline{\mathbb{R}}$ に対して，積分 $\int X dP$ が存在するとき，その 期待値 (expectation) を

$$E[X] := \int X dP$$

と定義する．積分の線形性より，X, Y が r.v.'s で，$E[|X|] < \infty, E[|Y|] < \infty$ なら，

$$E[aX + bY] = aE[X] + bE[Y],\ a, b \in \mathbb{R}$$

が成り立つ．

変数変換公式より，期待値は分布に関する積分に帰着する．X を S-valued r.v. とし，$g : S \to \mathbb{R}$ を可測とすると，$g(X)$ は r.v. であって，$E[g(X)]$ が定義できるなら，

$$E[g(X)] = \int g d\mu,\ \mu = \mathcal{L}(X)$$

となる．さらに，r.v. $X : \Omega \to \mathbb{R}$ の d.f. F が密度 f をもてば，

$$E[g(X)] = \int g dF = \int g(x)f(x)dx$$

となる (cf. Exercise 4.12).

ある $p > 0$ に対して，$E[|X|^p] < \infty$ なら，$0 < \forall q \leq p$ に対して，$|X|^q \leq 1 + |X|^p$ より，$E[|X|^q] \leq 1 + E[|X|^p] < \infty$ である．$E[X^2] < \infty$ のとき，X の 分散 $\mathrm{Var}(X)$ を

$$\mathrm{Var}(X) := E[(X - E[X])^2] = E[X^2] - (E[X])^2$$

と定義する．また，$E[X^2] < \infty$, $E[Y^2] < \infty$ のとき，Cauchy-Schwarz の不等式より，$E[|XY|] \leq (E[X^2])^{1/2}(E[Y^2])^{1/2} < \infty$ である．このとき，X, Y の 共分散 $\mathrm{Cov}(X, Y)$ を

$$\mathrm{Cov}(X, Y) := E[(X - E[X])(Y - E[Y])] = E[XY] - E[X]E[Y]$$

と定義する．$\mathrm{Cov}(X, X) = \mathrm{Var}(X)$ である．

I を $\mathbb{R}$ の区間とすると，$\varphi : I \to \mathbb{R}$ が 凸関数 (convex function) であるとは，任意の $x, y \in I$ と $\theta \in [0, 1]$ に対して，

$$\varphi(\theta x + (1 - \theta)y) \leq \theta\varphi(x) + (1 - \theta)\varphi(y)$$

となることを言う．

Theorem 6.3 (Jensen の不等式)．I を $\mathbb{R}$ の区間とし，$\varphi : I \to \mathbb{R}$ を凸関数とする．また，X を $P(X \in I) = 1$ なる r.v. とする．このとき，$\varphi(X)$ は a.s. に定義された r.v. である．さらに，$E[|X|] < \infty$ なら，$E[\varphi(X)]$ は定義でき，$-\infty < E[\varphi(X)] \leq \infty$ であって，$\varphi(E[X]) \leq E[\varphi(X)]$ が成り立つ．

Jensen の不等式の前に，凸関数に関する基本的な結果を述べておく．まず，$x, y, z \in I, x < y < z$ に対して，$\theta = (z - y)/(z - x)$ とおくと，$y = \theta x + (1 - \theta)z$ であるから，$\varphi(y) \leq \theta\varphi(x) + (1 - \theta)\varphi(z)$ である．これを書き直して，

$$\frac{\varphi(y) - \varphi(x)}{y - x} \leq \frac{\varphi(z) - \varphi(x)}{z - x}$$

を得る．同様にして，$x, y, z \in I, x < y < z$ に対して，

$$\frac{\varphi(y) - \varphi(x)}{y - x} \leq \frac{\varphi(z) - \varphi(y)}{z - y}$$

となる．x, y が I の内点なら，$\inf I < a < b < x < y < c < d < \sup I$ に対して，

$$\frac{\varphi(b) - \varphi(a)}{b - a} \leq \frac{\varphi(y) - \varphi(x)}{y - x} \leq \frac{\varphi(d) - \varphi(c)}{d - c}$$

であるから，

$$|\varphi(y)-\varphi(x)| \le |y-x| \cdot \max\left\{ \left|(\varphi(b)-\varphi(a))/(b-a)\right|, \left|((\varphi(d)-\varphi(c))/(d-c)\right| \right\}$$

である．よって，I の内部の各有界区間上で φ は Lipschitz 連続である．特に，φ は I の内部で連続である (端点では連続とは限らない)．

次に，I の内点 x に対して，$y \mapsto (\varphi(y)-\varphi(x))/(y-x)$ は $(x,\infty)\cap I$ 上で非減少であるから，右微分

$$D_+\varphi(x) := \lim_{y \downarrow x} \frac{\varphi(y)-\varphi(x)}{y-x}$$

が存在する．同様に $y \mapsto (\varphi(y)-\varphi(x))/(y-x)$ は $(-\infty,x)\cap I$ 上で非減少であるから，左微分

$$D_-\varphi(x) := \lim_{y \uparrow x} \frac{\varphi(y)-\varphi(x)}{y-x}$$

も存在して，$-\infty < D_-\varphi(x) \le D_+\varphi(x) < \infty$ となる．このとき，任意の $a \in [D_-\varphi(x), D_+\varphi(x)]$ と $y \in I$ に対して，

$$\varphi(y) \ge \varphi(x) + a(y-x)$$

となる．すなわち，次の補題を得る．

Lemma 6.2. I を $\mathbb{R}$ の区間とし，$\varphi : I \to \mathbb{R}$ を凸関数とする．また，c を I の内点とする．このとき，$\exists a \in \mathbb{R}$ s.t. $\varphi(x) \ge \varphi(c) + a(x-c)$ $(\forall x \in I)$ となる．

Remark 6.1. この補題は c が I の端点のときは成り立たない．例えば，$I = \mathbb{R}_+, \varphi(x) = 1_{\{0\}}(x), x \in \mathbb{R}_+$ という関数を検討してみればよい．

次の系は条件付き期待値に対する Jensen の不等式の証明に用いる．この系の証明は演習問題とする．

Corollary 6.1. I を開区間とし，$\varphi : I \to \mathbb{R}$ を凸関数とする．このとき，高々可算個の $a_n, b_n \in \mathbb{R}$ が存在して，$\varphi(x) = \sup_n (a_n x + b_n)$ $(\forall x \in I)$ と表せる．

Proof of Theorem 6.3. $\varphi(X)$ が a.s. に定義された r.v. であることを確認する．φ が可測であることを示せばよいが，これは φ の I の内部での連続性から明らかである．

次に，$c = E[X]$ とおく．c が I の端点なら，$X = c$ a.s. であるから，定理の結論は明らかである．よって，c が I の内点の場合を考えればよい．このとき，Lemma 6.2 より，$\exists a \in \mathbb{R}$ s.t. $\varphi(X) \ge \varphi(c) + a(X-c)$ であるから，$E[\varphi(X)^-] < \infty$ である．さらに，両辺の期待値をとって，$E[\varphi(X)] \ge \varphi(c) + a(E[X]-c) = \varphi(c)$ を得る． □

Example 6.2. $\varphi(x) = x \log x, x > 0$ に対して，$\varphi(0) = \lim_{x \downarrow 0} \varphi(x) = 0$ として，φ を $\mathbb{R}_+$ に拡張すれば，φ は $\mathbb{R}_+$ 上の凸関数であって，さらに，$\varphi(x) \ge -e^{-1}$ $\forall x \in \mathbb{R}_+$ である．よって，可積分な r.v. $X \ge 0$ に対して，$E[\varphi(X)]$ は定義でき，$E[\varphi(X)] \ge \varphi(E[X])$ となる．右辺は有限なので，$H(X) := E[\varphi(X)] - \varphi(E[X])$ は well-defined であって，$H(X) \ge 0$ である．$H(X)$ を X のエントロピーと呼ぶ．

Jensen の不等式の多変数への拡張を述べよう．$C \subset \mathbb{R}^k$ が 凸集合 であるとは，

$$\boldsymbol{x}, \boldsymbol{y} \in C, \theta \in [0,1] \Rightarrow \theta\boldsymbol{x} + (1-\theta)\boldsymbol{y} \in C$$

がみたされることを言う．$\varphi : C \to \mathbb{R}$ が凸関数であるとは，任意の $\boldsymbol{x}, \boldsymbol{y} \in C$ と $\theta \in [0,1]$ に対して，

$$\varphi(\theta\boldsymbol{x} + (1-\theta)\boldsymbol{y}) \leq \theta\varphi(\boldsymbol{x}) + (1-\theta)\varphi(\boldsymbol{y})$$

となることを言う．

Theorem 6.4 (多変数の Jensen の不等式)**.** $C \subset \mathbb{R}^k$ を Borel 可測な凸集合とし，$\varphi : C \to \mathbb{R}$ を凸関数とする．また，$\boldsymbol{X} = (X_1, \ldots, X_k)$ を C に値をとる $\mathbb{R}^k$-valued r.v. とし，各 $1 \leq j \leq k$ に対して，$E[|X_j|] < \infty$ とする．このとき，$\varphi(\boldsymbol{X})$ が r.v. なら，$E[\varphi(\boldsymbol{X})]$ が定義でき，$-\infty < E[\varphi(\boldsymbol{X})] \leq \infty$ であって，$\varphi(E[\boldsymbol{X}]) \leq E[\varphi(\boldsymbol{X})]$ が成り立つ． ただし，$E[\boldsymbol{X}] = (E[X_1], \ldots, E[X_k])$ である．

いくつか注意を述べておく．$k \geq 2$ では，凸集合は必ずしも Borel 可測でない．例えば，$D = \{(x,y) \in \mathbb{R}^2 : x^2 + y^2 < 1\}$ とし，$S \subset \partial D = \{(x,y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$ を $\mathbb{R}^2$ の Borel 非可測な集合とすると，$C = D \cup S$ は Borel 非可測な凸集合である．そのような S が存在することは，濃度の比較からわかる．すなわち，$\mathcal{B}^2$ は可算生成であるから，連続体濃度をもつが，$2^{\partial D}$ は連続体濃度よりも真に大きい濃度をもつから，∂D の部分集合のうち Borel 非可測なものが存在する．

また，φ は C の内部では連続であるが，C の境界では連続とは限らない．さらに，C が Borel 可測であっても，$\varphi(\boldsymbol{X})$ が r.v. になるとは限らない．例えば，先ほどの例で，$C = D \cup \partial D = \{(x,y) : x^2 + y^2 \leq 1\}$ とし，$S \subset \partial D$ を Borel 非可測な集合とする．ここで，$\varphi(x,y) = 1_S(x,y), (x,y) \in C$ とおくと，φ は凸関数であるが可測ではない．

Theorem 6.4 の証明は Dudley (2002, Theorem 10.2.6) を参照せよ．

Theorem 6.5 (Markov の不等式)**.** r.v. $X \geq 0$ に対して，$P(X \geq t) \leq E[X]/t$ $(\forall t > 0)$.

Proof. $E[X] \geq E[X1_{\{X \geq t\}}] \geq tP(X \geq t)$. □

$\varphi : \mathbb{R}_+ \to \mathbb{R}_+$ を非減少であって，$\varphi(x) > 0$ $(\forall x > 0)$ とする．X を r.v. とし，$t > 0$ とすると，$|X| \geq t$ なら $\varphi(|X|) \geq \varphi(t)$ であるから，Markov の不等式より，

$$P(|X| \geq t) \leq P(\varphi(|X|) \geq \varphi(t)) \leq \frac{E[\varphi(|X|)]}{\varphi(t)}.$$

特に，$\varphi(x) = x^2$ とすれば，

$$P(|X| \geq t) \leq t^{-2}E[X^2]$$

となる．これを Chebyshev の不等式 とも呼ぶ．

確率変数の収束. $X, X_n, n \in \mathbb{N}$ を r.v.'s とする．$\lim_n X_n = X$ と書いたら，$\lim_n X_n$ が存在して X に等しい，という意味だと約束すると，

$$\{\lim_n X_n = X\} = \{\limsup_n X_n = \liminf_n X_n = X\} \in \mathcal{F}.$$

そこで，$P(\lim_n X_n = X) = 1$ のとき，X_n は X に a.s. 収束 する (converge almost surely) と言う．このとき，$X_n \to X$ a.s. と書く．

X_n が X に 確率収束 する (converge in probability) とは，任意の $\varepsilon > 0$ に対して，$\lim_n P(|X_n - X| > \varepsilon) = 0$ となることを言う．このとき，$X_n \xrightarrow{P} X$ と書く．

$0 < p < \infty$ に対して，X_n が X に L^p 収束する (converge in L^p) とは，$X_n, X \in L^p$ & $\lim_n E[|X_n - X|^p] = 0$ となることを言う．このとき，$X_n \to X$ in L^p と書く．

Lemma 6.3. (i) $X_n \to X$ a.s. or $X_n \to X$ in $L^p \Rightarrow X_n \xrightarrow{P} X$. (ii) $X_n \xrightarrow{P} X$ & $X_n \xrightarrow{P} Y \Rightarrow X = Y$ a.s.

Proof. (i). $X_n \to X$ a.s. なら，$1_{\{|X_n - X| \} > \varepsilon} \to 0$ a.s. よって，DCT より，$P(|X_n - X| > \varepsilon) = E[1_{\{|X_n - X| > \varepsilon\}}] \to 0$. $X_n \to X$ in L^p なら，Markov の不等式より，$P(|X_n - X| > \varepsilon) \le \varepsilon^{-p} E[|X_n - X|^p] \to 0$.

(ii). $P(|X - Y| > \varepsilon) \le P(|X - X_n| > \varepsilon/2) + P(|X_n - Y| > \varepsilon/2) \to 0$ より，$P(|X - Y| \le \varepsilon) = 1$. $\varepsilon_m \downarrow 0$ のとき，$\{|X - Y| \le \varepsilon_m\} \downarrow \{X = Y\}$ であるから，$P(X = Y) = \lim_m P(|X - Y| \le \varepsilon_m) = 1$. □

Remark 6.2. 一般に，$X_n \xrightarrow{P} X \not\Rightarrow X_n \to X$ in L^p である．実際，X を $E[|X|] = \infty$ なる r.v. とし，$X_n = n^{-1}X$ とおくと，$X_n \xrightarrow{P} 0$ だが，$E[|X_n|] = \infty$. 確率収束からモーメントの収束を保証するには，以下で述べる一様可積分性なる条件が必要 (かつ十分) である．また，一般に $X_n \xrightarrow{P} X \not\Rightarrow X_n \to X$ a.s. である (cf. Exercise 9.5).

Lemma 6.4. $f : \mathbb{R} \to \mathbb{R}$ が連続，$X_n \xrightarrow{P} X$ なら，$f(X_n) \xrightarrow{P} f(X)$. f が有界連続なら，$f(X_n) \to f(X)$ in L^1.

Proof. $\varepsilon, \eta > 0$ を任意に固定する．$M > 0$ を十分大きく選んで，$P(|X| > M) \le \eta$ とすると，$\{|X_n| > M + 1\} \subset \{|X| > M\} \cup \{|X_n - X| > 1\}$ より，十分大きな n に対して，$P(|X_n| > M + 1) \le P(|X| > M) + P(|X_n - X| > 1) \le 2\eta$. f は $[-M - 1, M + 1]$ 上で一様連続であるから，$\exists \delta > 0$ s.t. $|x - y| < \delta,\ x, y \in [-M - 1, M + 1] \Rightarrow |f(x) - f(y)| < \varepsilon$. 従って，十分大きな n に対して，

$$\begin{aligned} P(|f(X_n) - f(X)| < \varepsilon) &\ge P(\{|X_n - X| < \delta\} \cap \{|X_n| \le M + 1\} \cap \{|X| \le M + 1\}) \\ &\ge P(|X_n - X| < \delta) - P(|X_n| > M + 1) - P(|X| > M + 1) \\ &\ge 1 - 4\eta. \end{aligned}$$

f が有界連続なら，$f(X_n) \xrightarrow{P} f(X)$ であるから，

$$\begin{aligned} E[|f(X_n) - f(X)|] &= E[|f(X_n) - f(X)|1_{\{|f(X_n)-f(X)|\geq\varepsilon\}}] \\ &\quad + E[|f(X_n) - f(X)|1_{\{|f(X_n)-f(X)|<\varepsilon\}}] \\ &\leq 2\|f\|_u P(|f(X_n) - f(X)| \geq \varepsilon) + \varepsilon = o(1) + \varepsilon. \end{aligned}$$

従って，$E[|f(X_n) - f(X)|] \to 0$. □

一様可積分性. r.v.'s $\{X_i : i \in I\}$ が一様可積分 (uniformly integrable) であるとは

$$\lim_{M\to\infty} \underbrace{\sup_{i\in I} E[|X_i|1_{\{|X_i|>M\}}]}_{(*)} = 0$$

となることを言う．このとき，M を十分大きくとると，$(*) \leq 1$ となるから，

$$\sup_{i\in I} E[|X_i|] \leq M + 1 < \infty$$

となる．一様可積分性はモーメントの収束を保証する．

Theorem 6.6 (Vitali)**.** $\{X_n : n \in \mathbb{N}\} \subset L^1, X_n \xrightarrow{P} X$ なら，次の (i)–(iii) は同値である．

(i) $\{X_n : n \in \mathbb{N}\}$ は一様可積分．

(ii) $X_n \to X$ in L^1.

(iii) $E[|X_n|] \to E[|X|] < \infty$.

Proof. (i) ⇒ (ii).

$$\varphi_M(x) = \begin{cases} M & \text{if } x > M \\ x & \text{if } |x| \leq M \\ -M & \text{if } x \leq -M \end{cases}$$

とおく．φ_M は有界連続である．

$$|X_n - X| \leq |X_n - \varphi_M(X_n)| + |\varphi_M(X_n) - \varphi_M(X)| + |\varphi_M(X) - X|$$

と分解すると，$|x - \varphi_M(x)| = (|x| - M)^+ \leq |x|1_{\{|x|>M\}}$ であるから，

$$\begin{aligned} E[|X_n - X|] &\leq \underbrace{E[|\varphi_M(X_n) - \varphi_M(X)|]}_{\to 0 \ (n\to\infty)} + E[|X_n|1_{\{|X_n|>M\}}] + E[|X|1_{\{|X|>M\}}], \\ \limsup_n E[|X_n - X|] &\leq \sup_n E[|X_n|1_{\{|X_n|>M\}}] + E[|X|1_{\{|X|>M\}}]. \end{aligned}$$

$\{X_n\}$ は一様可積分であるから，$M \to \infty$ とすれば，右辺第 1 項 $\to 0$ である．あとは，$X \in L^1$ が言えれば，DCT より，右辺第 2 項 $\to 0$ $(M \to \infty)$ が言えて，$X_n \to X$ in L^1 を得る．

$X \in L^1$ を示そう．まず任意の $L > 0$ に対して，$x \mapsto |x| \wedge L$ は有界連続であるから，$E[|X| \wedge L] = \lim_n E[|X_n| \wedge L]$. $\{X_n\}$ の一様可積分性より，$\sup_n E[|X_n|] < \infty$ であるから，$E[|X_n| \wedge L] \leq \sup_m E[|X_m|]$. あとは，Fatou の補題より，$E[|X|] \leq \liminf_{L\to\infty} E[|X| \wedge L] \leq \sup_m E[|X_m|] < \infty$ を得る．

(ii) $\Rightarrow$ (iii). 明らか．

(iii) $\Rightarrow$ (i).

$$\psi_M(x) = \begin{cases} |x| & \text{if } 0 \leq |x| \leq M-1 \\ 0 & \text{if } |x| > M \\ \text{linear} & \text{if } M-1 < |x| \leq M \end{cases}$$

とおくと，DCT より，十分大きな M に対して，$E[|X|] \leq E[\psi_M(X)] + \varepsilon/2$. $\lim_n E[|X_n|] = E[|X|]$, $\lim_n E[\psi_M(X_n)] = E[\psi_M(X)]$ であるから，十分大きな n に対して，

$$E[|X_n| 1_{\{|X_n|>M\}}] \leq E[|X_n|] - E[\psi_M(X_n)] \leq E[|X|] - E[\psi_M(X)] + \varepsilon/2 \leq \varepsilon$$

となる．従って，定理が示された． □

再配分不等式

ちょっと寄り道になるが，分布関数に関連する話題として，関数の再配分 (rearrangement) を考察する．$f : (0,1) \to \mathbb{R}$ を $(\mathcal{B}_{(0,1)}/\mathcal{B})$ 可測関数とし，f を確率空間 $((0,1), \mathcal{B}_{(0,1)}, \lambda)$ 上の r.v. とみなして (λ は $(0,1)$ 上の Lebesgue 測度である)，その d.f. を F_f とおく．すなわち，$F_f(t) = \lambda(\{x \in (0,1) : f(x) \leq t\})$ である．このとき，

$$f^*(x) := F_f^{-1}(x) := \inf\{t : F_f(t) \geq x\}, \quad x \in (0,1)$$

を f の 非減少再配分 (nondecreasing rearrangement) と呼ぶ．非減少再配分は次の性質をみたす．

(a) f^* は非減少．

(b) $f \overset{d}{=} f^*$.

(c) f が非減少なら，$f = f^*$ a.e.

(d) $1_{\{f^*>t\}} = (1_{\{f>t\}})^*$.

(e) 左連続な非減少関数 $G : \mathbb{R} \to \mathbb{R}$ に対して，$(G \circ f)^* = G \circ f^*$.

(a) と (b) は明らか．f が非減少なら，f の連続点 x に対して，$f(x) = f^*(x)$ となり，f の不連続点は高々可算個しかないので，(c) が従う．(d) に関しては，$f^*(x) \le t \Leftrightarrow x \le F_f(t)$ より，$\{f^* > t\} = \{f^* \le t\}^c = (F_f(t), 1)$ であるから，$1_{\{f^*>t\}} = 1_{(F_f(t),1)} = (1_{\{f>t\}})^*$．(e) に関しては，$G^+(t) = \sup\{x : G(x) \le t\}$ とおくと，$G(x) \le t \Leftrightarrow x \le G^+(t)$ であるから，$F_{G\circ f}(t) = F_f \circ G^+$．さらに，$F_f \circ G^+(t) \ge x \Leftrightarrow G^+(t) \ge f^*(x) \Leftrightarrow t \ge G \circ f^*(x)$ であるから，$(G \circ f)^* = G \circ f^*$．

Theorem 6.7 (Hardy-Littlewood)**.** 有界可測関数 $f, g : (0,1) \to \mathbb{R}$ に対して，

$$\int_0^1 f(x)g(x)dx \le \int_0^1 f^*(x)g^*(x)dx.$$

Proof. まず，f と g は非負と仮定する．Fubini の定理より，

$$\int_0^1 f(x)g(x)dx = \int_0^\infty \int_0^\infty \left\{ \int_0^1 1_{\{f>s\}}(x)1_{\{g>t\}}(x)dx \right\} dsdt$$

であり，さらに，(d) より，

$$\int_0^1 f^*(x)g^*(x)dx = \int_0^\infty \int_0^\infty \left\{ \int_0^1 1^*_{\{f>s\}}(x)1^*_{\{g>t\}}(x)dx \right\} dsdt$$

であるから，f と g が指示関数の場合に不等式を示せば十分である．いま，$f = 1_A, g = 1_B$, $A, B \in \mathcal{B}_{(0,1)}$ のとき，$\int_0^1 f(x)g(x)dx = \lambda(A \cap B)$ であり，一方，$f^* = 1_{(1-\lambda(A),1)}, g^* = 1_{(1-\lambda(B),1)}$ であるから，$\int_0^1 f^*(x)g^*(x)dx = \min\{\lambda(A), \lambda(B)\} \ge \lambda(A \cap B)$ を得る．

次に，f と g が非負とは限らないときは，$h_f = f - \inf f, h_g = g - \inf g$ とおくと，h_f, h_g は非負であって，$h_f^* = f^* - \inf f, h_g^* = g^* - \inf g$ である．あとは，$f \overset{d}{=} f^*, g \overset{d}{=} g^*$ より，$\int_0^1 f(x)dx = \int_0^1 f^*(x)dx, \int_0^1 g(x)dx = \int_0^1 g^*(x)dx$ であるから，

$$\begin{aligned}
&\int_0^1 f(x)g(x)dx \\
&= \int_0^1 h_f(x)h_g(x)dx + (\inf f)\int_0^1 g(x)dx + (\inf g)\int_0^1 f(x)dx - (\inf f)(\inf g) \\
&\le \int_0^1 h_f^*(x)h_g^*(x)dx + (\inf f)\int_0^1 g^*(x)dx + (\inf g)\int_0^1 f^*(x)dx - (\inf f)(\inf g) \\
&= \int_0^1 f^*(x)g^*(x)dx
\end{aligned}$$

を得る． □

特に，$f, g : (0,1) \to \mathbb{R}$ が有界可測なら，$\int_0^1 \{f^*(x) - g^*(x)\}^2 dx \le \int_0^1 \{f(x) - g(x)\}^2 dx$ となる．もっと一般に次の定理が成り立つ．$1 \le p \le \infty$ に対して，$\|\cdot\|_p$ を $(0,1)$ 上の Lebesgue 測度に関する L^p ノルムとする．

Theorem 6.8. 有界可測関数 $f, g : (0,1) \to \mathbb{R}$ に対して，

$$\|f^* - g^*\|_p \leq \|f - g\|_p, \quad 1 \leq \forall p \leq \infty$$

が成り立つ．

Proof. $1 \leq p < \infty$ に対して，

$$\begin{aligned}|f(x) - g(x)|^p &= p\int_{g(x)}^{\infty} \{(f(x) - t)^+\}^{p-1} dt + p\int_{f(x)}^{\infty} \{(g(x) - t)^+\}^{p-1} dt \\ &= p\int \left[\{(f(x) - t)^+\}^{p-1}\{1 - 1_{\{g>t\}}(x)\} + \{(g(x) - t)^+\}^{p-1}\{1 - 1_{\{f>t\}}(x)\}\right] dt.\end{aligned}$$

ここで，t を固定して，$h(x) = \{(f(x) - t)^+\}^{p-1}$ とおくと，(e) より，$h^*(x) = \{(f^*(x) - t)^+\}^{p-1}$ であるから，前定理と (d) より，

$$\int_0^1 \{(f(x) - t)^+\}^{p-1} 1_{\{g>t\}}(x) dx \leq \int_0^1 \{(f^*(x) - t)^+\}^{p-1} \underbrace{1^*_{\{g>t\}}(x)}_{=1_{\{g^*>t\}}(x)} dx$$

を得る．また，$\int_0^1 \{(f(x) - t)^+\}^{p-1} dx = \int_0^1 \{(f^*(x) - t)^+\}^{p-1} dx$ である．もう一方の項に対しても同様の操作を適用して，$\|f - g\|_p \geq \|f^* - g^*\|_p$ を得る．$p = \infty$ に対しては，$p \to \infty$ とすればよい． □

Theorem 6.8 の統計学における応用を述べよう．いま，$(0,1)$ 上の (可測) 関数 f の推定に興味があり，適当な推定値 $\hat{f}$ があるとする．f が非減少なら，推定値 $\hat{f}$ も非減少になるように選ぶのが自然であるが，標準的なノンパラメトリック推定法 (例えば，カーネル法) を適用しただけでは，$\hat{f}$ の単調性までは保証されない．しかし，$\hat{f}$ の非減少再配分 $\hat{f}^*$ は，その構成から非減少であり，L^p ノルムの誤差の意味でもともと推定値より悪くはならない：$\|\hat{f}^* - f\|_p \leq \|\hat{f} - f\|_p$．詳細に関しては，例えば，Chernozhukov et al. (2009) を参照せよ．

演習問題

Exercise 6.1. $X : \Omega \to \mathbb{R}$ が単関数のとき，その分位点関数をなるべく明示的に表現せよ．

Exercise 6.2. 連続な d.f. は一様連続であることを示せ．

Exercise 6.3. F を d.f. とする．点 $x \in \mathbb{R}$ が F の サポート (support) に属するとは，任意の $\varepsilon > 0$ に対して，$F(x + \varepsilon) - F(x - \varepsilon) > 0$ となることを言う．F のサポートに属する点全体を F のサポートと呼ぶ．

(a) F のサポートは閉集合であることを示せ．

(b) F が連続なら，F のサポートは孤立点を含まないことを示せ.

$(\mathbb{R}, \mathcal{B})$ 上の p.m. μ に対して，そのサポートを μ の d.f. のサポートで定義する.

Exercise 6.4. F を絶対連続な d.f. とし，連続な密度関数 f をもつとする．このとき，F のサポートは $\{f > 0\}$ の閉包に一致することを示せ.

適当な可算集合 $\{a_j\} \subset \mathbb{R}$ と $b_j > 0, \sum_j b_j = 1$ なる $\{b_j\}$ が存在して，$\mu = \sum_j b_j \delta_{a_j}$ と表される p.m. μ を 離散分布 (discrete distribution) と呼ぶ.

Exercise 6.5. $\mathbb{R}$ 全体をサポートにもつ離散分布を構成せよ.

Exercise 6.6. 与えられた (空でない) 閉集合 $C \subset \mathbb{R}$ をサポートにもつ分布を構成せよ.

Exercise 6.7. Cantor 分布のサポートは Cantor 集合であることを示せ.

Exercise 6.8. $X \sim F$ とし，F は連続とする．このとき，$Y = F(X)$ は $[0, 1]$ 上の一様分布に従うことを示せ.

$A \in \mathcal{F}$ が P の アトム (atom) であるとは，$P(A) > 0$ であって，

$$B \subset A, B \in \mathcal{F} \Rightarrow P(B) = P(A) \text{ or } P(B) = 0$$

となることを言う.

Exercise 6.9. $(\mathbb{R}, \mathcal{B})$ 上の p.m. μ に対して，μ がアトムをもたないための必要十分条件は，$\mu(\{x\}) = 0\ \forall x \in \mathbb{R}$ であることを示せ.

Exercise 6.10. $(\Omega, \mathcal{F}, P)$ はアトムをもたないとし，$A \in \mathcal{F}, P(A) > 0$ とする.

(a) 任意の $\varepsilon > 0$ に対して，次の条件をみたす $B \in \mathcal{F}$ が存在することを示せ : $B \subset A, 0 < P(B) < \varepsilon$.

(b) 任意の $0 < a < P(A)$ に対して，次の条件をみたす $B \in \mathcal{F}$ が存在することを示せ : $B \subset A, P(B) = a$.

Exercise 6.11. $(\Omega, \mathcal{F}, P)$ を確率空間とし，P^* を P の外測度とする．すなわち，

$$P^*(A) = \inf\{P(B) : B \in \mathcal{F}, B \supset A\},\ A \subset \Omega$$

である．$\Omega_0 \subset \Omega$ を $P^*(\Omega_0) = 1$ なる集合とし ($\Omega_0 \notin \mathcal{F}$ でもよい)，$\mathcal{G} = \{B \cap \Omega_0 : B \in \mathcal{F}\}$ とおく．このとき，写像 $Q : \mathcal{G} \to [0, 1]$ を，$A = B \cap \Omega_0, B \in \mathcal{F}$ に対して，$Q(A) = P(B)$ と定義すれば，Q は $(\Omega_0, \mathcal{G})$ 上の well-defined な p.m. であることを示せ.

Exercise 6.12. Corollary 6.1 を示せ.

Exercise 6.13. $x_1, \ldots, x_n \in \mathbb{R}_+$ と $1 \leq p < \infty$ に対して，

$$\left(\frac{1}{n} \sum_{i=1}^{n} x_i \right)^p \leq \frac{1}{n} \sum_{i=1}^{n} x_i^p$$

を示せ．また，$0 < p < 1$ のときは逆の不等式が成り立つことを示せ．

Exercise 6.14 (Paley-Zygmund の不等式)**.** $X \geq 0$ を r.v. とし，$E[X] < \infty$ とする．このとき，任意の $\theta \in [0, 1]$ に対して，

$$P(X > \theta E[X]) \geq (1 - \theta)^2 \frac{(E[X])^2}{E[X^2]}$$

を示せ．

ヒント：$X = X1_{\{X \leq \theta E[X]\}} + X1_{\{X > \theta E[X]\}}$ と分割し，第 2 項に Cauchy-Schwarz の不等式を適用する．

Exercise 6.15. X を r.v. とする．

(a) $(E[|X|^p])^{1/p}$ は $p > 0$ に関して非減少であることを示せ．

(b) (Littlewood の不等式). $0 < p < q < r$ に対して，

$$E[|X|^q] \leq (E[|X|^p])^{(r-q)/(r-p)} (E[|X|^r])^{(q-p)/(r-p)}$$

を示せ．例えば，$E[X^2] \leq (E[|X|])^{2/3} (E[X^4])^{1/3}$ である．

Exercise 6.16. $\{X_i : i \in I\}$ を r.v.'s とする．ある $1 < p < \infty$ が存在して，$\sup_{i \in I} E[|X_i|^p] < \infty$ なら，$\{X_i : i \in I\}$ は一様可積分であることを示せ．

Exercise 6.17. r.v.'s $\{X_i : i \in \mathbb{N}\}$ が一様可積分なら，$\{n^{-1} \sum_{i=1}^{n} X_i : n \in \mathbb{N}\}$ も一様可積分であることを示せ．

7 独立性

$(\Omega, \mathcal{F}, P)$ を確率空間とする．σ-fields $\mathcal{F}_1, \ldots, \mathcal{F}_n \subset \mathcal{F}$ が 独立 (independent) であるとは，任意の $A_1 \in \mathcal{F}_1, \ldots, A_n \in \mathcal{F}_n$ に対して，

$$P(A_1 \cap \cdots \cap A_n) = P(A_1) \cdots P(A_n)$$

となることを言う．一般の添え字集合 I に対して，σ-fields $\mathcal{F}_i \subset \mathcal{F}, i \in I$ が独立であるとは，任意の相異なる $i_1, \ldots, i_n \in I$ に対して，$\mathcal{F}_{i_1}, \ldots, \mathcal{F}_{i_n}$ が独立になることを言う．次の補題は σ-fields の独立性を証明するときによく用いる結果である．

Lemma 7.1. $\mathcal{A}_1, \ldots, \mathcal{A}_n$ は Ω を含む π システムで，任意の $A_1 \in \mathcal{A}_1, \ldots, A_n \in \mathcal{A}_n$ に対して，

$$P(A_1 \cap \cdots \cap A_n) = P(A_1) \cdots P(A_n)$$

が成り立つとする．このとき，$\sigma(\mathcal{A}_1), \ldots, \sigma(\mathcal{A}_n)$ は独立である．

Proof. $\Omega \in \mathcal{A}_i$ であるから，$A_{i_1} \in \mathcal{A}_{i_1}, \ldots, A_{i_k} \in \mathcal{A}_{i_k}, 1 \leq i_1 < \cdots < i_k \leq n$ に対して，

$$P(A_{i_1} \cap \cdots \cap A_{i_k}) = P(A_{i_1}) \cdots P(A_{i_k})$$

となる．

$A_2 \in \mathcal{A}_2, \ldots, A_n \in \mathcal{A}_n$ を任意に固定し，$F = A_2 \cap \cdots \cap A_n, \mathcal{L} = \{A : P(A \cap F) = P(A)P(F)\}$ とおくと，$\mathcal{L}$ は $\mathcal{A}_1$ を含む λ システムであるから，π-λ 定理より，$\mathcal{L} \supset \sigma(\mathcal{A}_1)$ を得る．従って，任意の $A_1 \in \sigma(\mathcal{A}_1), A_i \in \mathcal{A}_i$ $(2 \leq i \leq n)$ に対して，$P(A_1 \cap A_2 \cap \cdots \cap A_2) = P(A_1)P(A_2 \cap \cdots \cap A_n) = P(A_1)P(A_2) \cdots P(A_n)$ となる．同じ操作を繰り返して結論を得る． □

独立な σ-fields は小分けにしても独立である．

Lemma 7.2. $\mathcal{F}_{i,j}$, $1 \leq i \leq n, j \in J_i$ を独立な σ-fields とし，$\mathcal{G}_i = \bigvee_{j \in J_i} \mathcal{F}_{i,j} := \sigma(\bigcup_{j \in J_i} \mathcal{F}_{i,j})$ とおく．このとき，$\mathcal{G}_i, 1 \leq i \leq n$ は独立である．

Proof. 各 $1 \leq i \leq n$ に対して，集合族 $\mathcal{A}_i \subset 2^\Omega$ を

$$\mathcal{A}_i = \left\{ \bigcap_{j \in F} A_{i,j} : F \subset J_i \text{ は有限}, A_{i,j} \in \mathcal{F}_{i,j} \ (j \in F) \right\}$$

と定義すると，$\mathcal{A}_i$ は π システムであって，$\mathcal{A}_i \supset \bigcup_j \mathcal{F}_{i,j}$ である．あとは Lemma 7.1 を適用すればよい． □

集合 $A_i \in \mathcal{F}, i \in I$ が独立であるとは，σ-fields $\{\varnothing, A_i, A_i^c, \Omega\}, i \in I$ が独立になることを言う．この定義は，初等的な定義と同値である．

Lemma 7.3. $A_1, \ldots, A_n \in \mathcal{F}$ が独立 $\Leftrightarrow$ 任意の相異なる $i_1, \ldots, i_m \in \{1, \ldots, n\}$ に対して，$P(A_{i_1} \cap \cdots \cap A_{i_m}) = P(A_{i_1}) \cdots P(A_{i_m})$.

Proof. $\{A_i, \Omega\}$ は π システムであって，$\sigma(\{A_i, \Omega\}) = \{\varnothing, A_i, A_i^c, \Omega\}$ なので，あとは Lemma 7.1 を適用すればよい． □

次に，S-valued r.v. と σ-field の独立性を定義する．$(S, \mathcal{S})$ を可測空間とし，X を S-valued r.v. とする．このとき，

$$\sigma(X) := \{X^{-1}(B) : B \in \mathcal{S}\} \subset \mathcal{F}$$

は σ-field である．$\sigma(X)$ を X により生成される σ-field と呼ぶ．S-valued r.v. X と $\mathcal{F}$ の sub σ-field $\mathcal{G}$ が独立であるとは，$\sigma(X)$ と $\mathcal{G}$ が独立なことを言う．すなわち，任意の $B \in \mathcal{S}$ と $A \in \mathcal{G}$ に対して，

$$P(\{X \in B\} \cap A) = P(X \in B)P(A)$$

となるとき，X と $\mathcal{G}$ は独立であると言う．$(T, \mathcal{T})$ をもう 1 つの可測空間とし，$g : S \to T$ を可測写像とすると，合成写像は可測性を保存するから，$Y = g(X)$ は T-valued r.v. になる．このとき，$Y^{-1}(B) = X^{-1}(g^{-1}(B))$ であるから，$\sigma(Y) \subset \sigma(X)$ である．従って，

$$X \text{ と } \mathcal{G} \text{ が独立なら，} Y = g(X) \text{ と } \mathcal{G} \text{ も独立}$$

である．同様の注意は以降の議論にもあてはまる．

Remark 7.1. $\sigma(X)$ は X のもっている “情報” だと考えると，最後の注意は自然に理解できる：X を Y に移すと，その分だけ情報が失われる $(\sigma(Y) \subset \sigma(X))$.

Example 7.1 (トリビアルな例)**.** $\mathcal{G} = \{\varnothing, \Omega\}$ なら，あらゆる S-valued r.v. X は $\mathcal{G}$ と独立である．

最後に，r.v.'s の独立性を定義する．$(S_i, \mathcal{S}_i), i \in I$ を可測空間とし，各 X_i を S_i-valued r.v. とする．このとき，$X_i, i \in I$ が独立であるとは，σ-fields $\sigma(X_i), i \in I$ が独立になることを言う．すなわち，$X_i, i \in I$ が独立であるとは，任意の相異なる $i_1, \ldots, i_n \in I$ に対して，

$$P(X_{i_1} \in B_1, \ldots, X_{i_n} \in B_n) = P(X_{i_1} \in B_1) \cdots P(X_{i_n} \in B_n),\ B_1 \in \mathcal{S}_1, \ldots, B_n \in \mathcal{S}_n,$$

となることである．

Remark 7.2. $\sigma(X_i, i \in I) := \bigvee_{i \in I} \sigma(X_i)$ と定義する．

$I = \{1, \ldots, n\}$ とする．直積空間 $\prod_{i=1}^n S_i$ には直積 σ-field $\prod_{i=1}^n \mathcal{S}_i$ を入れると約束すると，変数変換公式と直積測度の構成より，

$$X_1, \ldots, X_n \text{ が独立 } \Leftrightarrow\ \mathcal{L}(X_1, \ldots, X_n) = \mathcal{L}(X_1) \times \cdots \times \mathcal{L}(X_n)$$

が成り立つ (直接 π-λ 定理を使っても同値性を証明できる).

以下，$(S_i, \mathcal{S}_i) = (\mathbb{R}, \mathcal{B})$ $(1 \le i \le n)$ の場合を考える．($\mathbb{R}$-valued) r.v.'s $X_1, \dots, X_n$ が独立ならば，変数変換公式と Fubini の定理より，(i) $X_i \ge 0$ $(1 \le \forall i \le n)$, or (ii) $E[|X_i|] < \infty$ $(1 \le \forall i \le n)$ のいずれかがみたされるとき，

$$E[X_1 \cdots X_n] = E[X_1] \cdots E[X_n]$$

が成り立つ．

Example 7.2. r.v.'s X, Y は独立で，$X \sim F$, $Y \sim G$ とする．このとき，

$$P(X = Y) = E[1_{\{X=Y\}}] = \iint 1_{\{x=y\}} dF(x) dG(y) = E[F(Y) - F(Y-)].$$

F が連続なら，右辺 $= 0$ である．

次の補題は，($\mathbb{R}$-valued) r.v.'s の独立性が初等的な定義と一致することを示している．

Lemma 7.4.

r.v.'s $X_1, \dots, X_n$ が独立

$\Leftrightarrow P(X_1 \le x_1, \dots, X_n \le x_n) = P(X_1 \le x_1) \cdots P(X_n \le x_n)$, $-\infty < x_1, \dots, x_n \le \infty$.

Proof. 各 i に対して，$\{\{X_i \le x\} : -\infty < x \le \infty\}$ は Ω を含む π システムであって，$\sigma(\{\{X_i \le x\} : -\infty < x \le \infty\}) = \sigma(X_i)$ なので，あとは Lemma 7.1 を適用すればよい． □

Theorem 7.1 (畳み込み)**.** 独立な r.v.'s $X \sim F$, $Y \sim G$ に対して，(i) $X + Y$ の d.f. は

$$P(X + Y \le z) = F * G(z) := \int F(z - y) dG(y).$$

(ii) F が密度 f をもてば，$X + Y$ の d.f. は密度

$$h(z) = \int f(z - y) dG(y)$$

をもつ．(iii) さらに G が密度 g をもてば，h は

$$h(z) = \int f(z - y) g(y) dy. \qquad (*)$$

$F * G$ を F と G の 畳み込み (convolution) と呼ぶ．また，$(*)$ の右辺を f と g の畳み込みと呼び，$f * g$ と書く．

$$f * g(z) = \int f(z - y) g(y) dy = \int f(y) g(z - y) dy.$$

Proof. (i). X, Y の独立性と Fubini の定理より，

$$P(X+Y \le z) = \iint 1_{\{x+y\le z\}} dF(x)dG(y) = \int F(z-y)dG(z).$$

(ii). $F(z-y) = \int_{-\infty}^{z} f(x-y)dx$ であるから，Fubini の定理より，

$$\int F(z-y)dG(y) = \int \int_{-\infty}^{z} f(x-y)dG(y) = \int_{-\infty}^{z} \int f(x-y)dG(y)dx.$$

(iii). cf. Exercise 4.12. □

独立な r.v.'s の構成

$(\mathbb{R}, \mathcal{B})$ 上の p.m.'s $\mu_i, i \in I$ が与えらえたとき，独立な r.v.'s $X_i, i \in I$ であって，各 X_i が μ_i を分布にもつようなものを 同じ確率空間上に 構成できるであろうか．I は有限か可算であるとしよう．I が有限 $I = \{1, \dots, n\}$ のときは，話は簡単である．確率空間を $(\Omega, \mathcal{F}, P) = (\mathbb{R}^n, \mathcal{B}^n, \prod_{i=1}^n \mu_i)$ とし，各 X_i を 座標写像 (coordinate map)

$$X_i(\omega) = \omega_i,\ \omega = (\omega_1, \dots, \omega_n) \in \Omega$$

とすれば，直積測度の定義より，

$$P(X_1 \in B_1, \dots, X_n \in B_n) = (\mu_1 \times \cdots \times \mu_n)(B_1 \times \cdots \times B_n) = \mu_1(B_1) \cdots \mu_n(B_n)$$

となるから，$X_1, \dots, X_n$ が求める r.v.'s である．

では，I が可算無限集合 $I = \{1, 2, \dots\}$ の場合はどうであろうか．この場合，話はそう簡単ではない．というのも，有限個の測度に対して直積測度が構成できることは示したが，それが無限個にまで拡張できるかは (すぐには) わからないからである．(可算) 無限個の独立な r.v.'s を同じ確率空間上に定義する必要がなぜあるかと言うと，大数の強法則 (後述) のステートメントを思い起こしてみればよい．大数の強法則は，$X_1, X_2, \dots$ を期待値が有限な i.i.d. r.v.'s とすると，

$$\left\{ \lim_n \frac{1}{n} \sum_{i=1}^{n} X_i = E[X_1] \right\}$$

なる事象の確率が 1 であることを述べている．このステートメントが意味をもつには，無限個の X_i がすべて同じ確率空間上に定義されている必要がある．

可算無限個の独立な r.v.'s を同時に構成するには，確率空間 $(\mathbb{R}, \mathcal{B}, \mu_i), i \in \mathbb{N}$ の無限直積空間を構成すればよいが，ややテクニカルなので，本節では，直接に無限直積確率空間を構成するのではなく，一様確率変数の 2 進数展開を利用した直感的に分かりやすい (と思われる) 構成法を紹介する．後々のために，ある程度広いクラスの可測空間に値をとる独立な r.v.'s の構成を考える．

可測空間 $(S,\mathcal{S})$ と $(T,\mathcal{T})$ が同型 (isomorphic) であるとは，次をみたす写像 $\varphi : S \to T$ が存在することを言う:φ は全単射であって，φ と φ^{-1} は可測である．このとき，φ を同型写像 (isomorphism) と呼ぶ．可測空間 $(S,\mathcal{S})$ が標準 Borel 空間 (standard Borel space) であるとは，$\mathbb{R}$ の Borel 集合 E が存在して，$(S,\mathcal{S})$ が $(E,\mathcal{B}_E)$ と同型になることを言う．

Theorem 7.2. 各 $i \in \mathbb{N}$ に対して，$(S_i,\mathcal{S}_i)$ を標準 Borel 空間とし，μ_i を $(S_i,\mathcal{S}_i)$ 上の p.m. とする．このとき，適当な確率空間 $(\Omega,\mathcal{F},P)$ 上に，独立な S_i-valued r.v.'s $X_i, i \in \mathbb{N}$ であって，$\mathcal{L}(X_i) = \mu_i$ となるものを構成できる．

Proof. 確率空間 $(\Omega,\mathcal{F},P)$ を

$$\Omega = (0,1),\ \mathcal{F} = \mathcal{B}_{(0,1)},\ P = U(0,1) = (0,1) \text{ 上の一様分布 (Lebesgue 測度)}$$

とする．証明を 2 つのステップに分ける．

Step 1. $(S_i,\mathcal{S}_i) = (\mathbb{R},\mathcal{B})$ の場合に定理の結論を示す．$\omega \in \Omega$ を 2 進数展開する．

$$\omega = 0.\varepsilon_1(\omega)\varepsilon_2(\omega)\varepsilon_3(\omega)\cdots = \frac{\varepsilon_1(\omega)}{2} + \frac{\varepsilon_2(\omega)}{2^2} + \frac{\varepsilon_3(\omega)}{2^3} + \cdots .$$

$\omega \in \{k2^{-n} : 1 \le k < 2^n, n \in \mathbb{N}\} =: N$ のとき，$\varepsilon_n(\omega), n \in \mathbb{N}$ は一意に決まらない．例えば，

$$0.100\cdots = \frac{1}{2} + 0 + \cdots = 0 + \frac{1}{2} + \frac{1}{2^2} + \cdots = 0.011\cdots .$$

しかし，$P(N) = 0$ なので，$\omega \in N$ に対して，$\varepsilon_n(\omega), n \in \mathbb{N}$ はどう決めてもよい．ここでは，$\omega = k2^{-n}$ (k: odd) のとき，$\varepsilon_n(\omega) = 0, \varepsilon_m(\omega) = 1$ $(\forall m > n)$ と決めておくと，$\varepsilon_1, \varepsilon_2, \ldots$ は独立な r.v.'s であって，$P(\varepsilon_i = 0) = P(\varepsilon_i = 1) = 1/2$ となる．

次に，

$$\begin{pmatrix} \delta_{1,1} & \delta_{1,2} & \delta_{1,3} & \cdots \\ \delta_{2,1} & \delta_{2,2} & \delta_{2,3} & \cdots \\ \delta_{3,1} & \delta_{3,2} & \delta_{3,3} & \cdots \\ \delta_{4,1} & \delta_{4,2} & \delta_{4,3} & \cdots \\ \vdots & \vdots & \vdots & \end{pmatrix} = \begin{pmatrix} \varepsilon_1 & \varepsilon_3 & \varepsilon_6 & \cdots \\ \varepsilon_2 & \varepsilon_5 & \cdots & \\ \varepsilon_4 & \varepsilon_8 & \cdots & \\ \varepsilon_7 & \cdots & \cdots & \\ \vdots & \vdots & & \end{pmatrix} \qquad (**)$$

として $\delta_{i,j},\ i,j \in \mathbb{N}$ を定義すると，$\delta_{i,j},\ i,j \in \mathbb{N}$ は独立であって，$P(\delta_{i,j} = 0) = P(\delta_{i,j} = 1) = 1/2$ となる．$U_i = \sum_{j=1}^{\infty} \delta_{i,j}/2^j$ とおくと，$U_i, i \in \mathbb{N}$ は独立であって，$U_i \sim U(0,1)$ となる．μ_i の d.f. を F_i として，$X_i := F_i^{-1}(U_i)$ とおくと，$X_i, i \in \mathbb{N}$ は独立であって，各 X_i は $\mathcal{L}(X_i) = \mu_i$ をみたす．これで Step 1 の証明が終わった．

Step 2. 一般の場合に定理を示す．$\varphi_i : S_i \to E_i \in \mathcal{B}$ を同型写像とする．$\nu_i := \mu_i \circ \varphi_i^{-1}$ とおくと，ν_i は $(E_i,\mathcal{B}_{E_i})$ 上の p.m. である．$\nu_i(B) := \nu_i(B \cap E_i), B \in \mathcal{B}$ として，ν_i を $(\mathbb{R},\mathcal{B})$ に拡張すると，Step 1 より，独立な ($\mathbb{R}$-valued) r.v.'s $Y_i, i \in \mathbb{N}$ であって，$\mathcal{L}(Y_i) = \nu_i$ となるものを構成できる．そこで，各 i に対して $x_i \in S_i$ に任意にとって，

$$X_i := \begin{cases} \varphi_i^{-1}(Y_i) & \text{if } Y_i \in E_i \\ x_i & \text{if } Y_i \in E_i^c \end{cases}$$

とおくと，$X_i, i \in \mathbb{N}$ は独立である．さらに，$P(Y_i \in E_i) = 1$ より，

$$P(X_i \in B) = P(Y_i \in \varphi_i(B)) = \mu_i \circ \varphi_i^{-1}(\varphi_i(B)) = \mu_i(B).$$

以上より，定理の結論を得る． □

定義からすぐには明らかではないが，標準 Borel 空間はかなり一般的なクラスの可測空間を含む．例えば，$(\mathbb{R}^k, \mathcal{B}^k)$ は標準 Borel 空間である．完備かつ可分な距離空間のことを Poland 空間 (Polish space) と呼ぶ (より一般には，完備な距離によって距離化可能な可分位相空間のことを Poland 空間と呼ぶ)．

Theorem 7.3. S が Poland 空間であって，$\mathcal{S}$ がその Borel σ-field なら，$(S, \mathcal{S})$ は標準 Borel 空間である．

Proof. S が有限集合なら定理の結論は明らかなので，S は無限集合とする．S の距離を d とし，$\{x_n\}$ を S の稠密な集合とする．また，$T = [0,1]^{\mathbb{N}} = [0,1] \times [0,1] \times \cdots$ を $[0,1]$ の可算直積とし，T に距離

$$e((a_n), (b_n)) = \sum_{n=1}^{\infty} \frac{1}{2^n} |a_n - b_n|, \ (a_n), (b_n) \in T$$

を入れる [7]．証明を 2 つのステップに分ける．

Step 1. S は T の Borel 集合と同相であることを示す．$g : S \to [0,1]^{\mathbb{N}}$ を

$$g(x) := \left(\frac{d(x, x_n)}{1 + d(x, x_n)} \right)_{n \in \mathbb{N}}, \ x \in S$$

と定義すれば，g は S から $g(S) =: B$ への同相写像である．B が T の Borel 集合であることを確認する．これは次の補題から従う．

Lemma 7.5. S, T を距離空間とし，S を完備とする．このとき，$g : S \to B \subset T$ が同相写像なら，B は T の Borel 集合である．

Proof of Lemma 7.5. $f = g^{-1} : B \to S$ として，

$$w_f(y) := \inf\{\mathrm{diam}_S(f(U \cap B)) : U \ni y, \ U\text{: open}\}, \ y \in T$$

と定義する．ただし，$A \subset S$ に対して，$\mathrm{diam}_S(A) = \sup\{d(x,y) : x, y \in A\}$ は A の直径である ($\mathrm{diam}(\varnothing) = 0$ とする)．ここで，$y \in \partial B, w_f(y) = 0$ とし，$y_n \to y, y_n \in B$ なる y_n を任意に選び，$x_n = f(y_n)$ とおくと，$\{x_n\}$ は S の Cauchy 列である ($\because w_f(y) = 0$ より，$\forall \varepsilon > 0, \exists U \ni y$: open s.t. $\mathrm{diam}_S(f(U \cap B)) < \varepsilon$．$y_n \to y, y_n \in B$ より，十分大きな n に対して，$y_n \in U \cap B$ となるから，十分大きな n, m に対して，$d(x_n, x_m) < \varepsilon$ となる)．よって，S の完備性より，$\exists x \in S$ s.t. $d(x_n, x) \to 0$ となる．このとき，$y_n = g(x_n) \to g(x)$ と

[7] $[0,1]^{\mathbb{N}}$ は Hilbert 立方体とも呼ばれる．

なるから，$y = g(x)$，すなわち，$y \in B$ を得る．以上の議論より，$B = \overline{B} \cap \{y : w_f(y) = 0\}$ が示された．$\overline{B}$ は閉集合より T の Borel 集合であって，$y \mapsto w_f(y)$ は上半連続であるから，$\{y : w_f(y) = 0\}$ も T の Borel 集合である．よって補題の結論を得る． □

Step 2. Step 1 より，T の Borel σ-field を $\mathcal{B}(T)$ とおいたとき，$(T, \mathcal{B}(T))$ と $([0,1], \mathcal{B}_{[0,1]})$ が同型であることを示せば十分である．Theorem 7.2 の証明で用いたルールに従うと，$[0,1]$ の各点に対して 2 進数展開が一意に決まる．そこで，$(a_n) \in T$ に対して，各 a_n を 2 進数展開する：$a_n = \sum_{m=1}^{\infty} \delta_{n,m}/2^m$. このようにして構成した $\delta_{n,m},\ n, m \in \mathbb{N}$ に対して，$(**)$ から $\varepsilon_n, n \in \mathbb{N}$ を定義すれば，写像 $(a_n) \mapsto \sum_{n=1}^{\infty} \varepsilon_n/2^n$ は T から $[0,1]$ への同型写像を与える． □

Theorem 7.2 の言い換え．$\prod_{i=1}^{\infty} S_i = S_1 \times S_2 \times \cdots$ の シリンダー集合 とは

$$B \times \prod_{i=n+1}^{\infty} S_i = \{(x_1, x_2, \dots) : (x_1, \dots, x_n) \in B\},\ B \in \prod_{i=1}^{n} \mathcal{S}_i, n \in \mathbb{N}$$

なる形の集合である．シリンダー集合全体の生成する σ-field を $\prod_{i=1}^{\infty} S_i$ の シリンダー σ-field と呼び，$\prod_{i=1}^{\infty} \mathcal{S}_i$ と書く：

$$\prod_{i=1}^{\infty} \mathcal{S}_i := \sigma\left(\left\{B \times \textstyle\prod_{i=n+1}^{\infty} S_i : B \in \prod_{i=1}^{n} \mathcal{S}_i, n \in \mathbb{N}\right\}\right).$$

すなわち，シリンダー σ-field とは，各 $n \in \mathbb{N}$ に対して，$\prod_{i=1}^{\infty} S_i$ から $\prod_{i=1}^{n} S_i$ への射影 $(x_1, x_2, \dots) \mapsto (x_1, \dots, x_n)$ が可測になる最小の σ-field である．この特徴づけから，

$$\begin{aligned}\prod_{i=1}^{\infty} \mathcal{S}_i &= \sigma\left(\left\{B_1 \times \cdots \times B_n \times \textstyle\prod_{i=n+1}^{\infty} S_i : B_1 \in \mathcal{S}_1, \dots, B_n \in \mathcal{S}_n, n \in \mathbb{N}\right\}\right) \\ &= \sigma\left(\left\{\textstyle\prod_{i=1}^{n-1} \mathcal{S}_i \times B_n \times \prod_{i=n+1}^{\infty} S_i : B_n \in \mathcal{S}_n, n \in \mathbb{N}\right\}\right)\end{aligned}$$

とも表せることがわかる (詳細は各自確かめよ).

Theorem 7.4. 各 $(S_i, \mathcal{S}_i)$ が標準 Borel 空間なら，$(\prod_{i=1}^{\infty} S_i, \prod_{i=1}^{\infty} \mathcal{S}_i)$ 上の p.m. μ であって，

$$\mu(B_1 \times \cdots \times B_n \times \textstyle\prod_{i=n+1}^{\infty} S_i) = \mu_1(B_1) \cdots \mu_n(B_n),\ B_1 \in \mathcal{S}_1, \dots, B_n \in \mathcal{S}_n, n \in \mathbb{N} \quad (*)$$

をみたすものが存在する．そのような μ は一意である．$\mu = \prod_{i=1}^{\infty} \mu_i$ と書く．

Proof. 一意性の証明は π-λ 定理を使えばよい．存在を示す．$(\Omega, \mathcal{F}, P),\ X_i, i \in \mathbb{N}$ を Theorem 7.2 で構成したものとすると，$X = (X_1, X_2, \dots) : \Omega \to \prod_{i=1}^{\infty} S_i$ は可測写像であって，像測度 $\mu = P \circ X^{-1}$ は $(*)$ をみたす． □

Remark 7.3. ここではTheorem 7.2からTheorem 7.4を証明したが，逆に，Theorem 7.4からTheorem 7.2を証明することもできる．$(\Omega, \mathcal{F}, P) = (\prod_{i=1}^{\infty} S_i, \prod_{i=1}^{\infty} \mathcal{S}_i, \mu)$ をTheorem 7.4で構成したものとし，X_i を座標写像 $X_i(\omega_1, \omega_2, \dots) = \omega_i, i \in \mathbb{N}$ にとれば，$X_i, i \in \mathbb{N}$ は独立であって，$\mathcal{L}(X_i) = \mu_i$ となる．

Remark 7.4. Theorem 7.2 (or 7.4) は各 $(S_i, \mathcal{S}_i)$ が一般の可測空間でも成り立つ (Dudley, 2002, Theorem 8.2.2)．しかし，標準 Borel 空間はこのあとのいくつかの議論において本質的な役割を果たす．

演習問題

Exercise 7.1. r.v.'s $\{X_n : n \in \mathbb{N}\}$ が独立であるためには，各 n に対して，$(X_1, \dots, X_n)$ と X_{n+1} が独立になることが必要十分である．このことを示せ．

Exercise 7.2. (a) r.v. X に対して，$\sigma(X)$ は可算生成であることを示せ．(b) $\mathcal{G}$ を $\mathcal{F}$ の可算生成な sub σ-field とすると，ある r.v. X が存在して，$\mathcal{G} = \sigma(X)$ となることを示せ．

Exercise 7.3. X とr.v.とし，$f, g : \mathbb{R} \to \mathbb{R}$ を非減少関数であって，$E[|f(X)|] < \infty, E[|g(X)|] < \infty, E[|f(X)g(X)|] < \infty$ とする．このとき，

$$E[f(X)g(X)] \geq E[f(X)]E[g(X)]$$

を示せ．また，f が非減少で g が非増加なら逆の不等式が成り立つこと示せ．

Exercise 7.4. $(\mathbb{R}, \mathcal{B})$ 上の p.m. μ に対して，そのサポートを $\mathrm{supp}(\mu)$ とおく．$(\mathbb{R}, \mathcal{B})$ 上の p.m.'s μ, ν に対して，

$$\mathrm{supp}(\mu * \nu) = \overline{\mathrm{supp}(\mu) + \mathrm{supp}(\nu)}$$

を示せ．ただし，$A, B \subset \mathbb{R}$ に対して，$A + B := \{x + y : x \in A, y \in B\}$ である．

Exercise 7.5. $X_n, n = 1, 2, \dots$ を i.i.d. r.v.'s とし，$P(X_n = 0) = P(X_n = 2) = 1/2$ とする．このとき，

$$\sum_{n=1}^{\infty} \frac{X_n}{3^n}$$

は a.s. に有限な値に収束し，Cantor 分布に従うことを示せ．さらに，Cantor 分布の平均と分散を計算せよ．

Exercise 7.6. $\mathbb{R}$ の可算直積 $\mathbb{R}^{\mathbb{N}}$ の直積位相は，距離

$$d(x, y) = \sum_{j=1}^{\infty} \frac{1}{2^j}(1 \wedge |x_j - y_j|), \quad x = (x_j), y = (y_j) \in \mathbb{R}^{\mathbb{N}}$$

によって距離化可能であり，この距離に関して $\mathbb{R}^{\mathbb{N}}$ は完備かつ可分であることを示せ．また，$\mathbb{R}^{\mathbb{N}}$ の Borel σ-field を $\mathcal{B}(\mathbb{R}^{\mathbb{N}})$ とおくと，$\mathcal{B}(\mathbb{R}^{\mathbb{N}})$ はシリンダー σ-field $\mathcal{B}^{\mathbb{N}}$ に一致することを示せ．

Exercise 7.7. $(S, \mathcal{S})$ が標準 Borel 空間なら，その可算直積 $(S^{\mathbb{N}}, \mathcal{S}^{\mathbb{N}})$ も標準 Borel 空間であることを示せ．

Exercise 7.8. $f, g : \mathbb{R} \to \mathbb{R}_+$ を確率密度関数とし，g を有界とする．このとき，$f * g$ は有界連続であることを示せ．

Exercise 7.9 (Steinhaus)**.** $E \subset \mathbb{R}$ を Lebesgue 可測な集合とし，$\lambda(E) > 0$ とする．このとき，$E - E := \{x - y : x, y \in E\}$ は原点のある開近傍を含むことを示せ．

ヒント：まず $E \in \mathcal{B}, \lambda(E) < \infty$ としてよいことを確認せよ．次に，$f = 1_E/\lambda(E), g = 1_{-E}/\lambda(E)$ とおくと，f, g はそれぞれ確率密度関数である．この f, g に対して前問の結果を適用せよ．

Remark 7.5 (Lebesgue 非可測集合に関する補足)**.** Steinhaus の定理を用いて，$I = [0, 1]$ の部分集合であって，Lebesgue 外測度が 1 で内測度が 0 となるものを構成してみよう．λ を I 上の Lebesgue 測度とし，λ^* と λ_* をそれぞれ外測度，内測度とする：$\lambda^*(E) := \inf\{\lambda(B) : B \in \mathcal{B}_I, B \supset E\}$, $\lambda_*(E) := 1 - \lambda^*(I \setminus E)$. α を無理数とし，$A = \{m + n\alpha : m, n \in \mathbb{Z}\}$, $B = \{2m + n\alpha : m, n \in \mathbb{Z}\}$ とおく．$x, y \in \mathbb{R}$ に対して，

$$x \sim y \Leftrightarrow x - y \in A$$

と定義すると，$\sim$ は同値関係になる．この同値関係に関して各同値類から代表元を集めた集合を C とおく (選択公理を使っている)．ここで，$D = C + B$ とおく．

Theorem 7.5. $\lambda^*(D \cap I) = 1$ & $\lambda_*(D \cap I) = 0$.

Proof. まず，A が $\mathbb{R}$ で稠密であることを示す．各正整数 n に対して，一意な整数 m_n が存在して，$0 \le m_n + n\alpha < 1$ となる．$x_n = m_n + n\alpha$ とおく．いま，任意の開区間 $U \subset \mathbb{R}$ に対して，$\lambda(U) > 1/N$ となるように正整数 N を選ぶ．このとき，$x_1, \ldots, x_{N+1}$ のうちあるペア x_i, x_j が存在して，$|x_i - x_j| < 1/N$ となる．$x_i - x_j$ のある整数倍は U に属していなくてはならないが，$A \supset \{m(x_i - x_j) : m \in \mathbb{Z}\}$ であるから，$A \cap U \neq \varnothing$ となる．これは A が $\mathbb{R}$ で稠密であることを意味する．同様に，$B, B + 1 := \{x + 1 : x \in B\} = \{(2m + 1) + n\alpha : m, n \in \mathbb{Z}\}$ も $\mathbb{R}$ で稠密である．

次に，$x_i \in C, y_i \in B$ $(i = 1, 2)$ に対して，$x_i - x_j$ は 0 か A に属していないかのいずれかであり，一方，$y_1 - y_2 \in B$ であるから，$x_1 - x_2 + (y_1 - y_2) \notin B + 1$ となる．従って，$(D - D) \cap (B + 1) = \varnothing$ となり，$B + 1$ が $\mathbb{R}$ で稠密なことより，$D - D$ は開区間を含みえない．よって，Steinhaus の定理より，$\lambda_*(D \cap I) = 0$ を得る．同様に，$(\mathbb{R} \setminus D) - (\mathbb{R} \setminus D) = (C + B + 1) - (C + B + 1) = (C + B) - (C + B)$ は $B + 1$ と排反なので，$\lambda_*(I \setminus (D \cap I)) = 0$. □

8　大数の弱法則

r.v.'s $\{X_i : i \in \mathbb{N}\}$ に対して，$S_n = \sum_{i=1}^n X_i$ とおく．ある正規化定数の列 $a_n > 0, b_n \in \mathbb{R}$ に対して，確率収束

$$\frac{S_n - b_n}{a_n} \xrightarrow{P} 0,\ n \to \infty$$

が成り立つことを 大数の弱法則 (weak law of large number, WLLN) と呼ぶ．確率収束の代わりに，a.s. 収束

$$\frac{S_n - b_n}{a_n} \to 0 \quad a.s.,\ n \to \infty$$

が成り立つことを 大数の強法則 (strong law of large number, SLLN) と呼ぶ．通常，大数の法則は (期待値 $E[S_n]$ の存在は仮定するとして)，$a_n = n, b_n = E[S_n]$ の場合を指す．正規化定数が同じ場合，SLLN は WLLN を含意する．ただし，多くの場合，確率収束を示すより a.s. 収束を示す方がずっと難しい (WLLN は成り立つが，SLLN は成り立たない例が存在する．次節を参照せよ)．SLLN については次節以降に扱うとして，本節は WLLN のいくつかのバージョンとその応用例を紹介する．

次の定理は $\{X_i\}$ が独立でなくても成り立つ．

Theorem 8.1 (WLLN)**.** $E[X_i] = \theta, \mathrm{Var}(X_i) \le C < \infty\ (\forall i \ge 1), \mathrm{Cov}(X_i, X_j) = 0\ (\forall i \ne j)$ とする．ただし，θ と C は定数である．このとき，$S_n/n \to \theta$ in L^2，従って，$S_n/n \xrightarrow{P} \theta$.

Proof. $\mathrm{Var}(S_n) = \sum_i \mathrm{Var}(X_i) + \sum_{i \ne j} \mathrm{Cov}(X_i, X_j) \le Cn$．よって，$E[|S_n/n - \theta|^2] = n^{-2}\mathrm{Var}(S_n) \le Cn^{-1} \to 0.$ □

Example 8.1. $\{X_i\}$ が i.i.d. r.v.'s であって，$E[X_1^2] < \infty$ なら，定理の条件がみたされるので，$S_n/n \to E[X_1]$ in L^2 となる．

Example 8.2 (Weierstrass の近似定理)**.** $f : [0,1] \to \mathbb{R}$ を連続とし，

$$f_n(x) = \sum_{m=0}^n \binom{n}{m} x^m (1-x)^{n-m} f(m/n)$$

とおく．f_n を f の n 次 Bernstein 多項式 と呼ぶ．このとき，

$$\sup_{x \in [0,1]} |f_n(x) - f(x)| \to 0,\ n \to \infty$$

が成り立つ．

Proof. $\{X_i\}$ を成功確率 p の独立な Bernoulli 試行とし，$S_n = \sum_{i=1}^n X_i$ とおくと，

$$P(S_n = m) = \binom{n}{m} p^m (1-p)^{n-m}.$$

よって，$E[f(S_n/n)] = f_n(p)$. $M = \sup_{x\in[0,1]} |f(x)|$ とおく．$\varepsilon > 0$ を任意に固定する．f は $[0,1]$ 上で一様連続であるから，$\exists\delta > 0$ s.t. $|x-y| < \delta \Rightarrow |f(x) - f(y)| < \varepsilon$. 一方，

$$P(|S_n/n - p| \geq \delta) \leq n^{-2}\delta^{-2} \operatorname{Var}(S_n) = n^{-1}\delta^{-2}p(1-p) \leq \frac{1}{4n\delta^2}.$$

$A = \{|S_n/n - p| < \delta\}, Y = f(S_n/n) - f(p)$ とおくと，

$$|\underbrace{E[f(S_n/n)]}_{=f_n(p)} - f(p)| = |E[Y]| \leq E[|Y|] \leq E[|Y|1_A] + E[|Y|1_{A^c}]$$
$$\leq \varepsilon + 2MP(A^c) \leq \varepsilon + \frac{M}{2n\delta^2}.$$

最右辺は p に依存しないので，$\limsup_n \sup_{p\in[0,1]} |f_n(p) - f(p)| \leq \varepsilon$. □

Remark 8.1 (C 空間の可分性と Stone-Weierstrass の定理)**.** $C([0,1])$ を $[0,1]$ 上の実数値連続関数の全体として，$C([0,1])$ に一様距離 $d_{\sup}(f,g) = \sup_{x\in[0,1]} |f(x) - g(x)|$ を入れると，Weierstrass の近似定理は $[0,1]$ 上の多項式の全体が $C([0,1])$ で稠密になることを示している．また，任意の多項式は有理係数をもつ多項式によって $[0,1]$ 上で一様に近似できるが，有理係数をもつ多項式は可算個しかないので，$C([0,1])$ が可分であることがわかる．なお Weierstrass の近似定理の相当な拡張として，Stone-Weierstrass の定理というのがある．本節の最後を参照せよ．

次の定理は (i.i.d. を仮定しているが) 期待値が存在しなくても成り立つ．

Theorem 8.2. $\{X_i\}$ を i.i.d. r.v.'s とし，$xP(|X_1| > x) \to 0$ $(x \to \infty)$ とする．このとき，$\theta_n = E[X_1 1_{\{|X_1|\leq n\}}]$ とおくと，$S_n/n - \theta_n \xrightarrow{P} 0$.

まず次の補題を示す．この補題は Fubini の定理の単純な応用であるが，今後随所に現れる重要な事実である．

Lemma 8.1. r.v. $Y \geq 0$ と実数 $0 < p < \infty$ に対して，$E[Y^p] = \int_0^\infty py^{p-1}P(Y > y)dy$.

Proof. Fubini の定理より，

$$E[Y^p] = \int Y^p dP = \int_\Omega \int_0^Y py^{p-1}dydP = \int_\Omega \int_0^\infty 1_{\{y<Y\}}py^{p-1}dydP$$
$$= \int_0^\infty py^{p-1} \int_\Omega 1_{\{Y>y\}}dPdy = \int_0^\infty py^{p-1}P(Y > y)dy.$$

□

Proof of Theorem 8.2. $\bar{X}_{n,i} = X_i 1_{\{|X_i|\leq n\}}, \bar{S}_n = \sum_{i=1}^n \bar{X}_{n,i}$ とおく．

$$P(|S_n/n - \theta_n| > \varepsilon) \leq P(S_n \neq \bar{S}_n) + P(|\bar{S}_n/n - \theta_n| > \varepsilon) = I + II.$$

$I \le P(X_i \neq \bar{X}_{n,i}, 1 \le \exists i \le n) \le \sum_{i=1}^n P(|X_i| > n) = nP(|X_1| > n) \to 0$. 一方，Markov の不等式より，$II \le \varepsilon^{-2} n^{-2} \operatorname{Var}(\bar{S}_n) \le \varepsilon^{-2} n^{-1} E[\bar{X}_{n,1}^2]$. ここで，

$$E[\bar{X}_{n,1}^2] = 2\int_0^\infty yP(|\bar{X}_{n,1}| > y)dy \le 2\int_0^n yP(|X_1| > y)dy.$$

$g(y) = yP(|X_1| > y)$ とおくと，$0 \le g(y) \le y, g(y) \to 0$ $(y \to \infty)$ であるから，$\sup_y g(y) < \infty$. よって，DCT より，$n^{-1}\int_0^n g(y)dy = \int_0^1 g(ny)dy \to 0$. □

Corollary 8.1. $\{X_i\}$ を i.i.d. r.v.'s とし，$E[|X_1|] < \infty$ とする．このとき，$S_n/n \xrightarrow{P} E[X_1]$ $(n \to \infty)$.

Proof. DCT より，$xP(|X_1| > x) = E[x1_{\{|X_1|>x\}}] \le E[|X_1|1_{\{|X_1|>x\}}] \to 0$ $(x \to \infty)$. 一方，$\theta_n = E[X_1 1_{\{|X_1| \le n\}}] \to E[X_1]$ $(n \to \infty)$ であるから，$S_n/n \xrightarrow{P} E[X_1]$ を得る． □

Stone-Weierstrass の定理. 寄り道となるが，Weierstrass の近似定理の拡張である，Stone-Weierstrass の定理を証明してみよう．後者の証明は前者にもとづいている．K をコンパクトな Hausdorff 空間とし，$C(K)$ を K 上の実数値連続関数の全体とする．$C(K)$ には一様距離 $d_{\sup}(f,g) = \sup_{x \in K} |f(x) - g(x)|$ を入れる．$\mathcal{A} \subset C(K)$ が部分代数であるとは，$f, g \in \mathcal{A}, a, b \in \mathbb{R} \Rightarrow af + bg, fg \in \mathcal{A}$ となることであり，$\mathcal{A}$ が K のあらゆる 2 点を分離するとは，任意の $x, y \in K, x \neq y$ に対して，$f(x) \neq f(y)$ となる $f \in \mathcal{A}$ が存在することを言う．

Theorem 8.3 (Stone-Weierstrass の定理)**.** $\mathcal{A} \subset C(K)$ を部分代数とし，K のあらゆる 2 点を分離し，かつ定数関数 $f \equiv 1$ を含むとする．このとき，$\mathcal{A}$ は $(C(K), d_{\sup})$ において稠密である．

Proof. $\mathcal{A}$ の $d_{\sup}$ に関する閉包を $\overline{\mathcal{A}}$ とおくと，$\overline{\mathcal{A}}$ も部分代数である．$\overline{\mathcal{A}} = C(K)$ を示せばよい．証明を 3 つのステップに分ける．

Step 1. $f, g \in \overline{\mathcal{A}} \Rightarrow f \vee g, f \wedge g \in \overline{\mathcal{A}}$ を示す．$f \vee g = (1/2)(f+g) + (1/2)|f-g|, f \wedge g = (1/2)(f+g) - (1/2)|f-g|$ より，$f \in \overline{\mathcal{A}} \Rightarrow |f| \in \overline{\mathcal{A}}$ を示せばよい．f は有界なので，$\exists M > 0$ s.t. $|f| \le M$. $|f|/M \in \overline{\mathcal{A}}$ なら $|f| = M(|f|/M) \in \overline{\mathcal{A}}$ となるので，$|f| \le 1$ と仮定してよい．Weierstrass の近似定理より，$[-1, 1]$ 上の連続関数 $t \mapsto |t|$ は多項式で一様に近似できる．また，$\overline{\mathcal{A}}$ が 1 を含む部分代数なことから，f の多項式は $\overline{\mathcal{A}}$ に属する．よって，$|f| \in \overline{\mathcal{A}}$ が示された．

Step 2. 次に，与えられた $x, y \in K, x \neq y$ と $c, d \in \mathbb{R}$ に対して，$h(x) = c, h(y) = d$ となる $h \in \mathcal{A}$ が存在することを示す．実際，$f \in \mathcal{A}$ を $f(x) \neq f(y)$ となるように選び，$a = (c-d)/(f(x) - f(y)), b = c - af(x)$ ととれば，$af(x) + b = c, af(y) + b = d$ となる．

Step 3. 定理の結論を示そう．$h \in C(K)$ とし，$x \in K$ を任意に固定する．任意の $y \in K$ に対して，Step 2 より，$h_y(x) = h(x), h_y(y) = h(y)$ となる $h_y \in \mathcal{A}$ が存在する．このとき，任意の $\varepsilon > 0$ に対して，y の適当な開近傍 U_y が存在して，$h_y(w) > h(w) - \varepsilon$ $(\forall w \in U_y)$

となる．$U_y, y \in K$ は K を被覆するが，K はコンパクトであったから，適当な有限部分被覆 $U_{y(1)}, \ldots, U_{y(n)}$ が存在する．そこで，$g_x = \max_{1 \le j \le n} h_{y(j)}$ とおくと，Step 1 より，$g_x \in \overline{\mathcal{A}}$ である．さらにその構成より，$g_x(x) = h(x), g_x(w) > h(w) - \varepsilon$ $(\forall w \in K)$ となる．

次に各 $x \in K$ に対して，x の適当な開近傍 V_x が存在して，$g_x(w) < h(w) + \varepsilon$ $(\forall w \in V_x)$ となる．再び K のコンパクト性より，適当な有限開被覆 $V_{x(1)}, \ldots, V_{x(m)}$ がとれる．そこで，$g = \min_{1 \le j \le m} g_{x(j)}$ とおくと，$g \in \overline{\mathcal{A}}$ & $d_{\sup}(g, h) < \varepsilon$ となる．ε は任意だったから，$h \in \overline{\mathcal{A}}$ を得る．以上より，定理が示された． □

ところで，証明からは K が Hausdorff であるという仮定がエッセンシャルか明らかでないが，K が Hausdorff でないと，K のあらゆる 2 点を分離する部分代数が存在しないので，この部分で Hausdorff 性を使っている．

演習問題

Exercise 8.1. $x \in \mathbb{R}_+$ と $t \in (0, \infty)$ に対して，

$$\lim_n \sum_{0 \le k \le nx} \frac{(nt)^k}{k!} = \begin{cases} 1 & \text{if } x > t \\ 0 & \text{if } x < t \end{cases}$$

を示せ．

ヒント：独立な $X_1, X_2, \cdots \sim Po(t)$ に対して，左辺は $P(S_n \le x)$ に等しい．

Exercise 8.2. F を $\mathbb{R}_+$ に集中した d.f. とする (i.e., $F(0-) = 0$)．このとき，

$$\psi(t) = \int_{\mathbb{R}_+} e^{-ty} dF(y), \ t \in \mathbb{R}_+$$

を F の Laplace 変換 と呼ぶ．前問の結果を用いて，F の連続点 $x \in \mathbb{R}_+$ に対して，

$$\lim_n \sum_{0 \le k \le nx} \frac{(-n)^k}{k!} \psi^{(k)}(n) = F(x)$$

を示せ．この結論から Laplace 変換から F が一意に決まることを説明せよ．

$P(\varepsilon = \pm 1) = 1/2$ なる r.v. ε のことを Rademacher r.v. と呼ぶ．

Exercise 8.3. (対称化不等式)．$1 \le p < \infty$ を given とする．また，$X_1, \ldots, X_n$ を独立な r.v.'s とし，$E[X_i] = 0$ $(1 \le \forall i \le n)$ とする．$\varepsilon_1, \ldots, \varepsilon_n$ を独立な Rademacher r.v.'s とし，$X_1, \ldots, X_n$ と独立とすると，

$$\frac{1}{2^p} E\left[\left|\sum_{i=1}^n \varepsilon_i X_i\right|^p\right] \le E\left[\left|\sum_{i=1}^n X_i\right|^p\right] \le 2^p E\left[\left|\sum_{i=1}^n \varepsilon_i X_i\right|^p\right]$$

を示せ．

Exercise 8.4. $\varepsilon_1, \dots, \varepsilon_n$ を独立な Rademacher r.v.'s とし，$a_1, \dots, a_n \in \mathbb{R}$ を定数とする．ただし，$a_1, \dots, a_n$ のうち少なくとも 1 つは 0 でないとする．

(a) (Hoeffding の不等式)．任意の $x > 0$ に対して，

$$P\left(\left|\sum_{i=1}^n a_i\varepsilon_i\right| > x\right) \le 2\exp\left(-\frac{x^2}{2\sum_{i=1}^n a_i^2}\right)$$

を示せ．

(b) (Khinchin の不等式)．任意の $1 \le p < \infty$ に対して，p のみに依存する定数 A_p, B_p が存在して，

$$A_p\left(\sum_{i=1}^n a_i^2\right)^{1/2} \le \left(E\left[\left|\sum_{i=1}^n a_i\varepsilon_i\right|^p\right]\right)^{1/p} \le B_p\left(\sum_{i=1}^n a_i^2\right)^{1/2}$$

が成り立つことを示せ．

ヒント：Littlewood の不等式と Hoeffding の不等式．じつは A_p は p と独立になるようにとれる．

Exercise 8.5. (Marcinkiewicz-Zygmund の不等式)．$1 \le p < \infty$ を given とする．また，$X_1, \dots, X_n$ を独立な r.v.'s とし，$E[X_i] = 0$ $(1 \le \forall i \le n)$ とする．このとき，p のみに依存する定数 A_p, B_p が存在して，

$$A_pE\left[\left(\sum_{i=1}^n X_i^2\right)^{p/2}\right] \le E\left[\left|\sum_{i=1}^n X_i\right|^p\right] \le B_pE\left[\left(\sum_{i=1}^n X_i^2\right)^{p/2}\right]$$

が成り立つことを示せ．さらに，

$$E\left[\left(\sum_{i=1}^n X_i^2\right)^{p/2}\right] \le \begin{cases} \sum_{i=1}^n E[|X_i|^p] & \text{if } 1 \le p < 2 \\ n^{p/2-1}\sum_{i=1}^n E[|X_i|^p] & \text{if } p \ge 2 \end{cases}$$

を示せ．

ヒント：対称化不等式と Khinchin の不等式．

特に，$\{X_i\}$ が平均 0 の i.i.d. r.v.'s であって，ある $2 \le p < \infty$ に対して，$E[|X_i|^p] < \infty$ なら，$E[|n^{-1}\sum_{i=1}^n X_i|^p] = O(n^{-p/2})$ である．よって，Markov の不等式より，$\varepsilon > 0$ に対して，

$$P\left(\left|\frac{1}{n}\sum_{i=1}^n X_i\right| > \varepsilon\right) \le \varepsilon^{-p}O(n^{-p/2}).$$

すなわち，X_i がより多くの有限なモーメントをもつほど，$n^{-1}\sum_{i=1}^n X_i$ はその平均のまわりに集中する．さらに，有限な指数モーメントをもてば，収束のスピードは n に関して指数的に速い．

Exercise 8.6. $\{X_i\}$ を有限な平均をもつ i.i.d. r.v.'s とする.

(a) $\Lambda := \{\lambda \in \mathbb{R} : E[e^{\lambda(X_1 - E[X_1])}] < \infty\}$ が空でない区間であることを示せ. また, $\psi(\lambda) := \log E[e^{\lambda(X_1 - E[X_1])}]$ が Λ 上の凸関数であることを示せ.

(b) $x > 0$ に対して,

$$P\left(\left|\frac{1}{n}\sum_{i=1}^{n}(X_i - E[X_1])\right| > x\right) \leq 2\exp\left\{-n\sup_{\lambda\in\Lambda}\{\lambda x - \psi(\lambda)\}\right\}$$

を示せ. X_1 が自由度 k の χ^2 分布に従うとき, 右辺を評価せよ.

Exercise 8.7. $C_{per}([0, 2\pi]) = \{f : f : [0, 2\pi] \to \mathbb{R},\ f$ は連続, $f(0) = f(2\pi)\}$ とおくと, $\{1, \cos(nx), \sin(nx), n = 1, 2, \dots\}$ の有限 1 次結合の全体は $(C_{per}([0, 2\pi]), \|\cdot\|_u)$ において稠密であることを示せ.

ヒント : Stone-Weierstrass の定理.

9 Borel-Cantelli の補題と Kolmogorov の 0-1 法則

本節では，SLLN を証明する上で重要な役割を果たす Borel-Cantelli の補題を証明する．Borel-Cantelli の補題はそれ自身いくつもの有用な結果を導く．その応用例をいくつか紹介する．最後に，やや寄り道となるが，独立性に関する重要な結果として Kolmogorov の 0-1 法則を証明する．

Borel-Cantelli の補題. $A_n \subset \Omega$ $(n = 1, 2, \dots)$ に対して，

$$\begin{aligned}
\limsup A_n &:= \{A_n \text{ i.o.}\} := \bigcap_{m\geq 1} \bigcup_{n\geq m} A_n \\
&= \{\omega : \text{無限に多くの } n \text{ に対して } \omega \in A_n\}, \\
\liminf A_n &:= \{A_n \text{ ev.}\} := \bigcup_{m\geq 1} \bigcap_{n\geq m} A_n \\
&= \{\omega : \text{有限個の } n \text{ を除いて } \omega \in A_n\}
\end{aligned}$$

と定義する．“i.o.” は infinitely often，“ev.” は eventually の意味である．

$$\{A_n \text{ i.o.}\}^c = (\limsup_n A_n)^c = \liminf A_n^c = \{A_n^c \text{ ev.}\}$$

である．

Theorem 9.1 (Borel-Cantelli の補題)**.** (i) $\sum_n P(A_n) < \infty \Rightarrow P(A_n \text{ i.o.}) = 0$, i.e., $P(A_n^c \text{ ev.}) = 1$. (ii) 独立な事象 $A_n, n \in \mathbb{N}$ に対して，$\sum_n P(A_n) = \infty \Rightarrow P(A_n \text{ i.o.}) = 1$.

Proof. (i) $N = \sum_n 1_{A_n}$ とおくと，Fubini の定理より，$E[N] = \sum_n P(A_n) < \infty$. 従って，$N < \infty$ a.s., i.e., $P(A_n \text{ i.o.}) = 0$.

(ii) $m < M < \infty$ とする．独立性と $1 - x \leq e^{-x}$ より，

$$\begin{aligned}
P\left(\bigcap_{n=m}^{M} A_n^c\right) &= \prod_{n=m}^{N} (1 - P(A_n)) \leq \prod_{n=m}^{M} \exp(-P(A_n)) \\
&= \exp\left(-\sum_{n=m}^{M} P(A_n)\right) \to 0, \ M \to \infty.
\end{aligned}$$

$\bigcap_{n=m}^{M} A_n^c \downarrow \bigcap_{n=m}^{\infty} A_n^c$ より，$P(\bigcap_{n=m}^{\infty} A_n^c) = 0$, i.e., $P(\bigcup_{n=m}^{\infty} A_n) = 1$. $\bigcup_{n=m}^{\infty} A_n \downarrow \limsup_n A_n$ より，$P(\limsup_n A_n) = 1$. □

Example 9.1. (i) の逆は正しくない．$P = U(0,1), A_n = (0, 1/n)$ とすると，$\limsup_n A_n = \{0\}$ であるから，$P(A_n \text{ i.o.}) = 0$．しかし，$\sum_n P(A_n) = \sum_n (1/n) = \infty$.

Remark 9.1. 証明から明らかなように，(i) は p.m. に限らず一般の測度でも成り立つ．すなわち，μ を $(\Omega, \mathcal{F})$ 上の測度とすると，$\sum_n \mu(A_n) < \infty$ なら $\mu(A_n \text{ i.o.}) = 0$ となる．

Borel-Cantelli の補題の応用例.

Theorem 9.2. $X_n, n \in \mathbb{N}$ を r.v.'s とする．次の (i)–(iii) は同値である.

(i) X_n が確率収束する.

(ii) $\forall \varepsilon > 0, \lim_{n,m\to\infty} P(|X_n - X_m| > \varepsilon) = 0.$

(iii) ある r.v. X が存在して，X_n の任意の部分列 $X_{n(m)}$ は X に a.s. 収束するさらなる部分列 $X_{n(m_k)}$ をもつ.

次の事実を思い出す (証明は演習問題とする).

Lemma 9.1 (Urysohn の部分列原理)**.** 数列 a_n に対して，ある a が存在して，a_n の任意の部分列が a に収束するさらなる部分列をもてば，$a_n \to a$.

Proof of Theorem 9.2. (i) $\Leftrightarrow$ (ii), (i) $\Leftrightarrow$ (iii) を示す.

(i) $\Rightarrow$ (ii). $X_n \xrightarrow{P} X$ のとき,

$$P(|X_n - X_m| > \varepsilon) \le P(|X_n - X| > \varepsilon/2) + P(|X_m - X| > \varepsilon/2)$$

より，$m, n \to \infty$ のとき，左辺 $\to 0$.

(ii) $\Rightarrow$ (i). 部分列 $n(m)$ を次のように選ぶ．まず $n(1) = 1$ とし，$n(m-1)$ まで決まったら，$n(m)$ を

$$n(m) = \inf\left\{M > n(m-1) : P(|X_k - X_\ell| > 2^{-m}) \le 2^{-m},\ \forall k, \ell \ge M\right\}$$

と決める．$A_m = \{|X_{n(m+1)} - X_{n(m)}| > 2^{-m}\}$ とおくと，$\sum_m P(A_m) \le \sum_m (1/2^m) < \infty$ であるから，Borel-Cantelli の補題より，$P(A_m^c \text{ ev.}) = 1$．$\omega \in \{A_m^c \text{ ev.}\}$ に対しては，十分大きな m に対して，$|X_{n(m+1)}(\omega) - X_{n(m)}(\omega)| \le 2^{-m}$ となるから,

$$|X_{n(m+\ell)}(\omega) - X_{n(m)}(\omega)| \le 2^{-m} + \cdots + 2^{-m-\ell+1} \le 2^{-m+1},\ \forall \ell \ge 1.$$

$m, \ell \to \infty$ のとき右辺 $\to 0$ であるから，$\{X_{n(m)}(\omega)\}_{m=1}^\infty$ は $\mathbb{R}$ の Cauchy 列である．よって，$\mathbb{R}$ の完備性より，有限な $\lim_m X_{n(m)}(\omega)$ が存在する．ここで,

$$X(\omega) = \begin{cases} \lim_m X_{n(m)}(\omega) & \text{if } \omega \in \{A_m^c \text{ ev.}\} \\ 0 & \text{otherwise} \end{cases}$$

とおけば，X は r.v. であって [8]，$X_{n(m)} \to X$ a.s.

$X_n \xrightarrow{P} X$ を示す.

$$P(|X_n - X| > \varepsilon) \le P(|X_n - X_{n(m)}| > \varepsilon/2) + P(|X_{n(m)} - X| > \varepsilon/2). \qquad (*)$$

[8](その確認) $Y := \limsup_m X_{n(m)}$(これは常に存在する) とおくと，Y は ($\overline{\mathbb{R}}$-valued) r.v. であるから，$X = Y 1_{\{A_m^c \text{ ev.}\}}$ は r.v. である.

$\delta > 0$ を given とすると，(ii) の条件より，$\exists n_1, m_1 \in \mathbb{N}$ s.t. $n \geq n_1, m \geq m_1 \Rightarrow P(|X_n - X_{n(m)}| > \varepsilon/2) \leq \delta/2$. また，a.s. 収束するなら確率収束するので，$\exists m_2 \in \mathbb{N}$ s.t. $m \geq m_2 \Rightarrow P(|X_{n(m)} - X| > \varepsilon/2) \leq \delta/2$. ゆえに，$n \geq n_1, m \geq m_1 \vee m_2$ に対して，$(*)$ の右辺 $\leq \delta$, すなわち，$\limsup_n P(|X_n - X| > \varepsilon) \leq \delta$.

(i) $\Rightarrow$ (iii). $X_n \xrightarrow{P} X$ ならば，X_n の任意の部分列 $X_{n(m)}$ は X に確率収束する．(i) と (ii) の同値性と，(ii) $\Rightarrow$ (i) の証明に現れた議論より，$X_{n(m)}$ は X に a.s. 収束するさらなる部分列 $X_{n(m_k)}$ をもつ．

(iii) $\Rightarrow$ (i). $a_n = P(|X_n - X| > \varepsilon)$ とおくと，a.s. 収束するなら確率収束するので，a_n の任意の部分列は 0 に収束するさらなる部分列をもつ．よって，$a_n \to 0$. □

Corollary 9.1. $X_n \xrightarrow{P} X$ なら，X に a.s. 収束する部分列 $X_{n(m)}$ が存在する．

Remark 9.2. L^0 を r.v.'s の全体とし，

$$d(X,Y) = E\left[\frac{|X-Y|}{1+|X-Y|}\right],\ X, Y \in L^0$$

と定義すると，a.s. に等しい r.v.'s を同一視すれば，この d は L^0 上の距離を与える (Durrett 2.3.8). さらに，Durrett 2.3.9 より，

$$d(X_n, X) \to 0 \Leftrightarrow X_n \xrightarrow{P} X$$

であるから，d は確率収束を距離化している (Theorem 9.2 より，(L^0, d) が完備な距離空間になることもわかる)．一方，a.s. 収束は距離化可能でないこともわかる．つまり，

$$\rho(X_n, X) \to 0 \Leftrightarrow X_n \to X\ a.s.$$

をみたす L^0 上の距離 ρ は存在しない．実際，そのような距離 ρ が仮に存在したとして，X_n を $X_n \xrightarrow{P} X$ であるが a.s. 収束はしない r.v.'s とする (cf. Exercise 9.5)．$a_n = \rho(X_n, X)$ とおくと，$X_n \xrightarrow{P} X$ と Theorem 9.2 より，a_n の任意の部分列 $a_{n(m)}$ に対して，さらなる部分列 $a_{n(m_k)}$ が存在して，$a_{n(m_k)} \to 0$ $(k \to \infty)$ となる．しかし，これは $a_n \to 0$ を意味するので，ρ の定義より，$X_n \to X$ a.s. となって矛盾が生じる．

Example 9.2 (確率収束に対する DCT と Fatou の補題)**.** 確率収束に対しても DCT が成り立つ．すなわち，$X_n \xrightarrow{P} X, \exists Y \in L^1$ s.t. $|X_n| \leq Y$ なら，$X \in L^1$ & $E[X_n] \to E[X]$.

Proof. X_n の任意の部分列 $X_{n(m)}$ に対して，さらなる部分列 $X_{n(m_k)}$ が存在して，$X_{n(m_k)} \to X$ a.s. となるから，$X \in L^1$ であって，DCT より，$E[X_{n(m_k)}] \to E[X]$. 極限が部分列のとり方に依存しないから，$E[X_n] \to E[X]$ を得る. □

同様に，確率収束に対しても Fatou の補題が成り立つ．すなわち，$X_n \geq 0, X_n \xrightarrow{P} X$ なら，$E[X] \leq \liminf_n E[X_n]$ となる (Durrett 2.3.6)．

次に，$1 \leq p < \infty$ に対して，L^p 空間が Banach 空間であることを示そう．完備なノルム空間のことを Banach 空間 と呼び，完備な内積空間のことを Hilbert 空間 と呼ぶ．

Theorem 9.3. $1 \leq p < \infty$ とし，$L^p = L^p(\Omega, \mathcal{F}, P)$ とする．また，r.v. X に対して，$\|X\|_p = (E[|X|^p])^{1/p}$ と定義する．このとき，a.s. に等しい r.v.'s を同一視すれば，$(L^p, \|\cdot\|_p)$ は実 Banach 空間になる．さらに，L^2 は内積 $\langle X, Y \rangle = E[XY]$ に関して実 Hilbert 空間になる．

Proof. $\|\cdot\|_p$ がノルムであることはよい．完備性を示す．$\|X_n - X_m\|_p \to 0$ $(n, m \to \infty)$ とする．このとき，$X_n - X_m \xrightarrow{P} 0$ $(n, m \to \infty)$ であるから，ある r.v. X が存在して，$X_n \xrightarrow{P} X$ となる．ここで，n を固定して $m \to \infty$ とすると，$X_n - X_m \xrightarrow{P} X_n - X$ となるから，確率収束に対する Fatou の補題より，$E[|X_n - X|^p] \leq \liminf_m E[|X_n - X_m|^p]$ を得る．$n \to \infty$ のとき，右辺 $\to 0$ であるから，$X \in L^p$ と $X_n \to X$ in L^p を得る． □

Remark 9.3. L^p 空間はほとんどの場合，無限次元である．従って，L^p の有界閉集合はコンパクトとは限らない (逆は正しい)．これは，(実/複素) ノルム空間の閉単位球がコンパクトであるためにはそのノルム空間が有限次元であることが必要十分であるという，関数解析の基本的な定理から従う．確率空間が十分大きい場合，L^p 空間の閉単位球がコンパクトでないことは，次の例からわかる：いま，$(\Omega, \mathcal{F}, P)$ は独立かつ L^p ノルムが 1 の r.v.'s $X_1, X_2, \ldots$ をもつとする．このとき，$n \neq m$ に対して，$\|X_n - X_m\|_p = \|X_1 - X_2\|_p > 0$ であるから，$\{X_n\}$ は L^p において収束部分列をもたない．

次の定理は SLLN の逆である (cf. Corollary 10.1)．

Theorem 9.4. $\{X_i\}$ を i.i.d. r.v.'s とすると，$E[|X_1|] = \infty$ なら，

$$P(S_n/n \text{ が有限な極限をもつ}) = 0.$$

Proof. $\{X_i\}$ は同一分布をもつから，

$$\infty = E[|X_1|] = \int_0^\infty P(|X_1| > x)dx = \sum_{n=0}^\infty \int_n^{n+1} P(|X_1| > x)dx \leq 1 + \sum_{n=1}^\infty P(|X_n| > n).$$

いま，$\{X_n > n\}, n \in \mathbb{N}$ は独立であるから，Borel-Cantelli の補題より，$P(|X_n| > n \text{ i.o.}) = 1$. 次に，$C = \{S_n/n$ が有限の極限をもつ $\}$ とおくと，$S_n/n - S_{n+1}/(n+1) = S_n/n(n+1) - X_{n+1}/(n+1)$ と，$\forall \omega \in C, S_n(\omega)/n(n+1) \to 0$ より，

$$\omega \in C \cap \{|X_n| > n \text{ i.o.}\} \Rightarrow \left| \frac{S_n(\omega)}{n} - \frac{S_{n+1}(\omega)}{n+1} \right| > 2/3 \quad \text{i.o.}$$

しかしこれは C の定義に矛盾するから，$C \cap \{|X_n| > n \text{ i.o.}\} = \varnothing$. $P(|X_n| > n \text{ i.o.}) = 1$ より，$P(C) = P(C \cap \{|X_n| > n \text{ i.o.}\}) + P(C \cap \{|X_n| > n \text{ i.o.}\}^c) = 0$. □

Example 9.3 (WLLN は成り立つが，SLLN は成り立たない例)**.** Theorem 8.2 と Theorem 9.4 より，WLLN は成り立つが，SLLN は成り立たない例を簡単に構成できる．例えば，$\{X_i\}$ を i.i.d. r.v.'s とし，X_1 の分布は対称 ($X_1 \stackrel{d}{=} -X_1$) であり，$P(|X_1| > x) = e/(x \log x)$ $(x \geq e)$ とする．このとき，$xP(|X_1| > x) = e/(\log x) \to 0$ $(x \to \infty)$ であり，分布の対称性より，$\theta_n = E[X_1 1_{\{|X_1| \leq n\}}] = 0$ であるから，Theorem 8.2 より，$S_n/n \stackrel{P}{\to} 0$. 一方，$E[|X_1|] = \infty$ であるから，Theorem 9.4 より，S_n/n が有限な極限をもつ確率は 0 である．さらに，Kolmogorov の 0-1 法則 (後述) を使うと，

$$\limsup_n \frac{S_n}{n} = \infty \quad a.s., \quad \liminf_n \frac{S_n}{n} = -\infty \quad a.s.$$

となることも示すことができる．対称性より，最初の主張のみ示せばよい．任意の $A > 0$ に対して，$E[|X_1|/A] = \infty$ であるから，$P(|X_n| > An \text{ i.o.}) = 1$ であり，$|X_n| > An$ のとき，$|S_n - S_{n-1}| = |X_n| > An$ となるから，$|S_n| > An/2$ or $|S_{n-1}| > An/2$ のいずれかが起こる．以上の議論より，$P(|S_n| > An/2 \text{ i.o.}) = 1$．特に，$\{S_n > An \text{ i.o.}\}$ と $\{-S_n > An \text{ i.o.}\}$ のいずれかは正の確率をもつが，$S_n \stackrel{d}{=} -S_n$ より，結局，$P(S_n > An \text{ i.o.}) > 0$ である．さらに，$\{S_n > An \text{ i.o.}\}$ は tail event (後述) であるから，Kolmogorov の 0-1 法則より，

$$P(S_n > An/2 \text{ i.o.}) = 1.$$

$A > 0$ は任意であったから，$\Omega_0 = \bigcap_{m=1}^{\infty} \{S_n/n > m \text{ i.o.}\}$ とおくと，$P(\Omega_0) = 1$ であり，$\omega \in \Omega_0$ に対して，$\limsup_n S_n(\omega)/n = \infty$.

Example 9.4 (可積分性の特徴づけ)**.** Borel-Cantelli の補題は，r.v. の可積分性の興味深い特徴づけを導く．その一例として，i.i.d. r.v.'s $\{X_i\}$ に対して，

$$E[|X_1|] < \infty \Leftrightarrow P(|X_n| > n \text{ i.o.}) = 0 \Leftrightarrow X_n/n \to 0 \ a.s.$$

を示してみる．まず，

$$E[|X|] = \int_0^{\infty} P(|X| > x)dx = \sum_{n=0}^{\infty} \int_n^{n+1} P(|X| > x)dx$$

より，$\sum_{n=0}^{\infty} P(|X| > n+1) \leq E[|X|] \leq \sum_{n=0}^{\infty} P(|X| > n)$ となるから，

$$E[|X|] < \infty \Leftrightarrow \sum_{n=1}^{\infty} P(|X| > n) < \infty.$$

次に，$\{X_i\}$ は同一分布に従うから，$\sum_n P(|X_1| > n) = \sum_n P(|X_n| > n)$. $\{X_n > n\}, n \in \mathbb{N}$ は独立であるから，Borel-Cantelli の補題より，

$$E[|X_1|] < \infty \Leftrightarrow P(|X_n| > n \text{ i.o.}) = 0.$$

あとは，$X_n(\omega)/n \to 0 \Rightarrow \omega \in \{|X_n| \leq n \text{ ev.}\}$ は明らかなので，$E[|X_1|] < \infty \Rightarrow X_n/n \to 0$ a.s. を示せばよい． $Y_i = X_i/\varepsilon$ とおくと，$E[|Y_1|] < \infty$ より，$1 = P(|Y_n| \leq n \text{ ev.}) = P(|X_n|/n \leq \varepsilon \text{ ev.})$. すなわち，$\limsup_n |X_n|/n \leq \varepsilon$ a.s. ε は任意だったから，$\varepsilon = \varepsilon_m = 1/m \downarrow 0$ $(m \to \infty)$ として，$X_n/n \to 0$ a.s. を得る． □

Example 9.5 (表が続く回数)．$Z_n, n \in \mathbb{Z}$ を i.i.d. r.v.'s とし, $P(Z_n = 0) = P(Z_n = 1) = 1/2$ とする．$\{Z_n = 1\}$ を n 回目にコインの表が出る事象，$\{Z_n = 0\}$ を n 回目に裏が出る事象とする．n 回目の時点で表が続いている回数を ℓ_n とおく：

$$\ell_n = \max\{m : Z_{n-m+1} = \cdots = Z_n = 1\}.$$

また，1 から n 回目までの時点で表が続いている回数の最大値を L_n とおく：

$$L_n = \max_{1 \le m \le n} \ell_m.$$

$P(\ell_n = k) = P(Z_{n-k} = -1, Z_{n-k+1} = \cdots = Z_n = 1) = (1/2)^{k+1}$ である．

$$L_n / \log_2 n \to 1 \ a.s.$$

を示す．まず，$P(\ell_n \ge (1+\varepsilon)\log_2 n) \le n^{-(1+\varepsilon)}$ であるから，Borel-Cantelli の補題より $P(\ell_n < (1+\varepsilon)\log_2 n \text{ ev.}) = 1$. すなわち，$\exists N_\varepsilon \in \mathcal{A}, P(N_\varepsilon) = 0$ s.t. $\forall \omega \in \Omega \setminus N_\varepsilon, \exists n_0 = n_0(\omega) \in \mathbb{N}$ with $n > n_0 \Rightarrow \ell_n(\omega)/\log_2 n < 1 + \varepsilon$. 従って，$\limsup_n L_n(\omega)/\log_2 n \le 1 + \varepsilon$ $(\forall \omega \in \Omega \setminus N_\varepsilon)$. $N = \bigcup_{m=1}^{\infty} N_{1/m}$ とおくと，$P(N) = 0$ であり，$\limsup_n L_n(\omega)/\log_2 n \le 1 + 1/m$ $(\forall \omega \in \Omega \setminus N, \forall m \in \mathbb{N})$. 従って，$\limsup_n L_n/\log_2 n \le 1$ a.s.

あとは，$\liminf_n L_n/\log_2 n \ge 1$ a.s. を示せばよい．$k_n = \lfloor (1-\varepsilon)\log_2 n \rfloor$ とおき，n を k_n で割った商を q_n，余りを r_n とする．$n = q_n k_n + r_n$．$\{1, \ldots, n\}$ を $q_n + 1$ 個のブロックに分割する．

$$\underbrace{\{1, \ldots, k_n\}}_{I_{n,1}}, \underbrace{\{k_n + 1, \ldots, 2k_n\}}_{I_{n,2}}, \ldots, \underbrace{\{(q_n - 1)k_n + 1, \ldots, q_n k_n\}}_{I_{n,q_n}}, \{q_n k_n + 1, \ldots, n\}.$$

ここで，$P(\ell_{jk_n} \ge k_n) \ge P(Z_i = 1 \ \forall i \in I_{n,j}) = 2^{-k_n} \ge n^{-(1-\varepsilon)}$. 十分大きな n に対して，$q_n \ge n/\log_2 n$ であり，そのとき，

$$\begin{aligned} P(L_n \le (1-\varepsilon)\log_2 n) &= P(\ell_m \le (1-\varepsilon)\log_2 n, 1 \le \forall m \le n) \\ &\le P(\ell_{jk_n} \le (1-\varepsilon)\log_2 n, 1 \le \forall j \le k_n) \\ &\le (1 - n^{-(1-\varepsilon)})^{q_n} \le e^{-n^{\varepsilon}/\log_2 n}. \quad (\because (1 - x/n)^n \le e^{-x}) \end{aligned}$$

右辺の n に関する和は有限なので，Borel-Cantelli の補題より，

$$P(L_n > (1-\varepsilon)\log_2 n \text{ ev.}) = 1.$$

$\varepsilon > 0$ は任意だったから，$\liminf_n L_n/\log_2 n \ge 1$ a.s. を得る． □

Kolmogorov の 0-1 法則

$X_1, X_2, \dots$ を独立な r.v.'s とする．

$$\mathcal{F}'_n := \sigma(X_n, X_{n+1}, \dots) := \bigvee_{m \geq n} \sigma(X_m)$$

とおくと，$\mathcal{F}'_n$ は $X_n, X_{n+1}, \dots$ が可測になる Ω 上の最小の σ-field である．このとき，

$$\mathcal{T} := \bigcap_n \mathcal{F}'_n$$

を tail σ-field と呼び，$\mathcal{T}$ に属する事象を tail event と呼ぶ．直感的には，tail event は有限個の X_m を取り除いても不変な事象と理解できる．

Example 9.6. (i) $S_n = \sum_{m=1}^n X_m$ とおくと，$\{\lim_n S_n \text{ が存在する }\} \in \mathcal{T}$. (ii) $A_n \in \sigma(X_n)$ $(n = 1, 2, \dots)$ に対して，$\limsup A_n \in \mathcal{T}$. (iii) $\{\limsup_n X_n > a\} \in \mathcal{T}$.

Proof. いずれも直感的には明らかであるが，形式的に証明してみる．

(i). $A = \{\lim_n S_n \text{ が存在する }\}$ とおくと，任意の $k \in \mathbb{N}$ に対して，

$$A = \bigcap_{r>0, r\in\mathbb{Q}} \bigcup_{N \geq k} \bigcap_{n,m \geq N} \{|S_n - S_m| < r\}.$$

$n, m \geq k$ のとき，$\{|S_n - S_m| < r\} \in \sigma(X_{n\wedge m+1}, \dots, X_{n\vee m}) \subset \mathcal{F}'_{k+1}$ より，$A \in \mathcal{F}'_{k+1}$ であり，k は任意だったから，$A \in \bigcap_k \mathcal{F}'_k = \mathcal{T}$ を得る．

(ii). $\limsup A_n = \bigcap_{m\geq 1} \bigcup_{n \geq m} A_n = \bigcap_{m \geq k} \bigcup_{n \geq m} A_n$ より，$\limsup A_n \in \mathcal{F}'_k$ であり，k は任意であるから，$\limsup_n A_n \in \mathcal{T}$ を得る．

(iii). (ii) において，$A_n = \{X_n > a\} \in \sigma(X_n)$ とおくと，$\limsup A_n = \{A_n \text{ i.o.}\} = \{\limsup_n X_n > a\}$. □

Theorem 9.5 (Kolmogorov の 0-1 法則)**.** $P(A) \in \{0, 1\}$ $\forall A \in \mathcal{T}$.

Proof. まず，$\mathcal{F}'_1$ と $\mathcal{T}$ が独立なことを示す．$\bigcup_k \sigma(X_1, \dots, X_k)$ と $\mathcal{T}$ は Ω を含む π システムであるから，$\mathcal{F}'_1$ と $\mathcal{T}$ が独立であることを示すには，任意の $A \in \bigcup_k \sigma(X_1, \dots, X_k)$ と $B \in \mathcal{T}$ に対して，$P(A \cap B) = P(A)P(B)$ を示せばよいが，ある $k \in \mathbb{N}$ に対して，$A \in \sigma(X_1, \dots, X_k)$ であり，$\mathcal{T} \subset \mathcal{F}'_{k+1} = \sigma(X_{k+1}, X_{k+2}, \dots)$ であるから，Lemma 7.2 より，$P(A\cap B) = P(A)P(B)$ を得る．次に，$A \in \mathcal{T}$ に対して，$\mathcal{T} \subset \mathcal{F}'_1$ より $A \in \mathcal{F}'_1$ でもあるから，A は A 自身と独立である．従って，$P(A) = P(A\cap A) = P(A)^2$ より，$P(A) \in \{0, 1\}$ を得る． □

Remark 9.4. $\mathcal{F}$ の sub σ-field $\mathcal{G}$ が 自明 (trivial) であるとは，$P(A) \in \{0, 1\}$ $(\forall A \in \mathcal{G})$ となることを言う．Kolmogorov の 0-1 法則は tail σ-field が自明になることを述べている．次の事実を確認しておく．

Lemma 9.2. $\mathcal{G}$ を自明な σ-field とすると，$\mathcal{G}$ 可測な r.v. X, i.e., $X^{-1}(B) \in \mathcal{G}$ $(\forall B \in \mathcal{B})$ なる r.v. は a.s. に定数に等しい．

Proof. $\{X \leq x\} \in \mathcal{G}$ であるから，$P(X \leq x) \in \{0,1\}$. すなわち，X の d.f. は $\{0,1\}$ にしか値をとらない．これは X が a.s. に定数に等しいことを意味する． □

演習問題

Exercise 9.1. Lemma 9.1 を示せ．

Exercise 9.2. $\liminf_n A_n \subset \limsup_n A_n$ を示せ．$\liminf_n A_n \subsetneq \limsup_n A_n$ となるような例を構成せよ．

Exercise 9.3. $\limsup_n 1_{A_n} = 1_{\limsup_n A_n}$, $\liminf_n 1_{A_n} = 1_{\liminf A_n}$ を示せ。

Exercise 9.4. $\limsup_n P(A_n) \leq P(\limsup_n A_n), P(\liminf_n A_n) \leq \liminf_n P(A_n)$ を示せ．

Exercise 9.5. X_n を $P(X_n = 1) = p_n, P(X_n = 0) = 1 - p_n$ なる r.v. とする．このとき，$X_n \xrightarrow{P} 0$ となるための p_n に関する必要十分条件と，$X_n \to 0$ a.s. となるための p_n に関する必要十分条件を導出せよ．また，$X_n \xrightarrow{P} 0$ であるが，$X_n \to 0$ a.s. とはならない p_n を構成せよ．

Exercise 9.6. 任意の r.v.'s X_n に対して，ある定数列 $c_n > 0$ が存在して，$X_n/c_n \to 0$ a.s. となることを示せ．

Exercise 9.7. $X_1, X_2, \ldots$ を独立な r.v.'s とする．このとき，$\sup_n X_n < \infty$ a.s. であるためには，ある $A > 0$ が存在して，$\sum_n P(X_n > A) < \infty$ となることが必要十分であることを示せ．

Exercise 9.8. $\{X_i\}$ を i.i.d. r.v.'s とすると，次の (1)–(3) が同値なことを示せ．(1) $E[|X_1|] < \infty$; (2) $\max_{1 \leq i \leq n} |X_i|/n \to 0$ a.s.; (3) $E[\max_{1 \leq i \leq n} |X_i|]/n \to 0$.

Exercise 9.9. $X_1, X_2, \cdots \sim N(0,1)$ i.i.d. とする．$\phi(x) = (2\pi)^{-1/2} e^{-x^2/2}$ を $N(0,1)$ の密度関数とする．

(a)
$$\lim_{x \to +\infty} \frac{P(X_1 > x)}{x^{-1}\phi(x)} = 1$$
を示せ．

(b)
$$\limsup_n (X_n/\sqrt{2 \log n}) = 1 \text{ a.s.}$$
を示せ．

Exercise 9.10 (Kochen-Stone). $\{A_n\} \subset \mathcal{F}$ とし，$\sum_n P(A_n) = \infty$ とする．このとき，

$$P(A_n \text{ i.o.}) \geq \limsup_n \frac{\{\sum_{m=1}^n P(A_m)\}^2}{\sum_{\ell=1}^n \sum_{m=1}^n P(A_\ell \cap A_m)}$$

を示せ．$\{A_n\}$ が独立なら，右辺は $= 1$ なので，上の不等式は Borel-Cantelli の補題 (ii) の一般化である．

ヒント：$\sum_{n=m}^M 1_{A_n}$ に対して，Paley-Zygmund の不等式を適用する．

Exercise 9.11 (Riesz-Fischer). Theorem 9.3 は一般の測度空間に対しても成り立つ．$(S, \mathcal{S}, \mu)$ を測度空間とし，$1 \leq p < \infty$ に対して，$L^p = L^p(S, \mathcal{S}, \mu), \|f\|_p = (\int |f|^p d\mu)^{1/p}$ とすると，a.e. に等しい関数を同一視すれば，$(L^p, \|\cdot\|_p)$ は実 Banach 空間になる．また，$p = 2$ のときは，L^2 は $\langle f, g \rangle = \int f g d\mu$ を内積として実 Hilbert 空間になる．これらを示せ．

ヒント：L^p の Cauchy 列 $\{f_n\}$ に対して a.e. 収束する適当な部分列 $\{f_{n_k}\}$ が存在する：$f_{n_k} \to f$ a.e. このとき，$\|f_n - f\|_p \to 0$ を示せ．

Exercise 9.12. 前問において，μ が σ-finite であって，$\mathcal{S}$ が可算生成なら，$(L^p, \|\cdot\|_p)$ は可分であることを示せ．

Exercise 9.13. さらに，a.e. に等しい関数を同一視すれば，$(L^\infty, \|\cdot\|_\infty)$ は実 Banach 空間になることを示せ．しかし，L^∞ は μ が σ-finite で $\mathcal{S}$ が可算生成であったとしても，可分とは限らない．例えば，μ を $\mathbb{N}$ 上の計数測度とすると，$\ell^\infty := L^\infty(\mathbb{N}, 2^{\mathbb{N}}, \mu)$ は可分でないことを確認せよ．

Exercise 9.14. $A, B \in \mathcal{F}$ に対して，$d(A, B) = P(A \Delta B)$ と定義すると，$P(A \Delta B) = 0$ なる A, B を同一視すれば，$(\mathcal{F}, d)$ は完備な距離空間であることを示せ．

ヒント：d が距離であることは Theorem 1.3 で示した．完備性を示すために，

$$d(A \Delta B) = E[|1_A - 1_B|]$$

なる表現を使え．

10　大数の強法則

本節は SLLN とそれに関連する結果をいくつか紹介する．

Theorem 10.1 (SLLN)**.** $\{X_i : i \in \mathbb{N}\}$ を i.i.d. r.v.'s とし, $E[|X_1|] < \infty$ とする．このとき，$S_n/n \to E[X_1]$ a.s. $(n \to \infty)$.

証明は次の定理と Kronecker の補題による．

Theorem 10.2 (1 級数定理)**.** $\{X_i\}$ を独立な r.v.'s とし，$E[X_i] = 0, E[X_i^2] < \infty$ $(\forall i)$ とする．このとき，$\sum_{n=1}^{\infty} E[X_n^2] < \infty$ なら，a.s. に S_n は有限な値に収束する．

cf. Kolmogorov の 0-1 法則．Theorem 10.2 の証明のために，次の Kolmogorov の最大不等式を示す．

Theorem 10.3 (Kolmogorov の最大不等式)**.** $X_1, \ldots, X_n$ を独立な r.v.'s とし，$E[X_i] = 0, E[X_i^2] < \infty$ $(1 \le \forall i \le n)$ とする．このとき，あらゆる $x > 0$ に対して，

$$P\left(\max_{1\le k\le n} |S_k| \ge x\right) \le \frac{E[S_n^2]}{x^2}$$

が成り立つ．

Proof. $A_k = \{|S_k| \ge x, |S_j| < x, 1 \le \forall j < k\}$ とおく．$A_1, \ldots, A_n$ は排反であるから，

$$\begin{aligned} E[S_n^2] &\ge \sum_{k=1}^{n} E[S_n^2 1_{A_k}] = \sum_{k=1}^{n} E[(S_k + (S_n - S_k))^2 1_{A_k}] \\ &\ge \sum_{k=1}^{n} E[S_k^2 1_{A_k}] + 2\sum_{k=1}^{n} E[(S_n - S_k)S_k 1_{A_k}]. \end{aligned}$$

$S_k 1_{A_k}$ は $\sigma(X_1, \ldots, X_k)$ 可測であり，一方，$S_n - S_k$ は $\sigma(X_{k+1}, \ldots, X_n)$ 可測であるから，両者は独立である．従って，

$$E[(S_n - S_k)S_k 1_{A_k}] = \underbrace{E[S_n - S_k]}_{=0} E[S_k 1_{A_k}] = 0.$$

A_k 上で $|S_k| \ge x$ であることと，$A_1, \ldots, A_n$ が排反であることより，

$$E[S_n^2] \ge \sum_{k=1}^{n} E[S_k^2 1_{A_k}] \ge x^2 \sum_{k=1}^{n} P(A_k) = x^2 P\left(\bigcup_{k=1}^{n} A_k\right) = x^2 P\left(\max_{1\le k\le n} |S_k| \ge x\right).$$

最後の等号は A_k の定義から従う． □

Proof of Theorem 10.2. Kolmogorov の最大不等式より，

$$P\left(\max_{M\le m\le N}|S_m-S_M|>\varepsilon\right)\le\varepsilon^{-2}E[(S_N-S_M)^2]=\varepsilon^{-2}\sum_{n=M+1}^{N}E[X_n^2].$$

$N\to\infty$ のとき，$\{\max_{M\le m\le N}|S_m-S_M|>\varepsilon\}\uparrow\{\sup_{m\ge M}|S_m-S_M|>\varepsilon\}$ より，

$$P\left(\sup_{m\ge M}|S_m-S_M|>\varepsilon\right)\le\varepsilon^{-2}\sum_{n=M+1}^{\infty}E[X_n^2]\to 0,\ M\to\infty.$$

$w_M=\sup_{m\ge M}|S_m-S_M|$ とおくと，$w_M\xrightarrow{P}0$. ここで次の補題を用いる．

Lemma 10.1. $Y_n\ge 0$ が非増加で $Y_n\xrightarrow{P}0$ なら，$Y_n\to 0$ a.s.

Proof of Lemma 10.1. $Y_n\xrightarrow{P}0$ より，$\exists Y_{n(k)}\to 0$ a.s. $\omega\in\{Y_{n(k)}\to 0\}$ に対して，$\forall\varepsilon>0,\exists k\in\mathbb{N}$ s.t. $Y_{n(k)}(\omega)\le\varepsilon$. Y_n は非増加であるから，$n\ge n(k)$ に対して，$Y_n(\omega)\le Y_{n(k)}(\omega)\le\varepsilon$. □

定理の証明に戻ると，$w_M\ge 0$ は非増加で $w_M\xrightarrow{P}0$ であるから，$w_M\to 0$ a.s., i.e., S_n は a.s. に Cauchy 列である．よって，求める結論を得る． □

Lemma 10.2 (Kronecker の補題)**.** $a_n\uparrow\infty$ であって $\sum_{n=1}^{\infty}x_n/a_n$ が収束するなら，

$$\frac{1}{a_n}\sum_{m=1}^{n}x_n\to 0.$$

Proof. $a_0=b_0,b_m=\sum_{k=1}^{m}x_k/a_k$ とおくと，$x_m=a_m(b_m-b_{m-1})$ であるから，簡単な計算により，$a_n^{-1}\sum_{m=1}^{n}x_m=b_n-\sum_{m=1}^{n}(a_m-a_{m-1})b_{m-1}/a_n$. 仮定より，$\lim_n b_n=:b$ が存在する．一方，$\sum_{m=1}^{n}(a_m-a_{m-1})b_{m-1}/a_n\to b$ を示すのは難しくない． □

Proof of Theorem 10.1. $Y_k=X_k1_{\{|X_k|\le k\}},T_n=\sum_{k=1}^{n}Y_k$ とおくと，$\sum_{k=1}^{\infty}P(|X_k|>k)\le E[|X_1|]<\infty$ となるから，Borel-Cantelli の補題より，$P(X_k\ne Y_k\text{ i.o.})=0$. よって，$|S_n/n-T_n/n|\to 0$ a.s. となる．$Z_k=Y_k-E[Y_k]$ とおくと，

$$E[Z_k^2]\le E[Y_k^2]=2\int_0^{\infty}yP(|Y_k|>y)dy\le 2\int_0^{k}yP(|X_1|>y)dy$$

であるから，

$$\begin{aligned}\sum_{k=1}^{\infty}E[Z_k^2]/k^2&\le 2\sum_k k^{-2}\int_0^{\infty}1_{\{y<k\}}yP(|X_1|>y)dy\\&=2\int_0^{\infty}\left\{\sum_k k^{-2}1_{\{y<k\}}\right\}yP(|X_1|>y)dy\\&=2\int_0^{\infty}(\sum_{k>y}k^{-2})yP(|X_1|>y)dy.\end{aligned}$$

ここで，$\sum_{k>y} k^{-2} = O(y^{-1})$ $(y \to \infty)$, $\int_0^\infty P(|X_1| > y)dy = E[|X_1|] < \infty$ であるから，$\sum_{k=1}^\infty Z_k/k$ は a.s. に収束する．従って，Kronecker の補題より，

$$\frac{T_n}{n} - \underbrace{\frac{1}{n}\sum_{k=1}^n E[X_1 1_{\{|X_1|\le k\}}]}_{=\theta_n} = \frac{1}{n}\sum_{k=1}^n Z_k \to 0 \ a.s.$$

DCT より，$E[X_1 1_{\{|X_1|\le k\}}] \to E[X_1]$ $(k \to \infty)$ となるから，$\theta_n \to E[X_1]$ $(n \to \infty)$ である．従って，$T_n/n \to E[X_1]$ a.s. を得る． □

次の結果は，期待値 $E[X_1]$ が存在していさえすれば，それが発散していても SLLN が成り立つことを示している．

Corollary 10.1. $\{X_i\}$ を i.i.d. r.v.'s とする．このとき，$E[X_1^+] = \infty, E[X_1^-] < \infty$ なら，$S_n/n \to \infty$ a.s. となる．

Proof. $M > 0, X_i^M = X_i \wedge M$ とおくと，$E[|X_i^M|] < \infty$ より，$S_n^M/n := \sum_{i=1}^n X_i^M/n \to E[X_1^M]$ a.s. さらに，$X_i \ge X_i^M$ より，

$$\liminf_n S_n/n \ge \liminf_n S_n^M/n = E[X_1^M] \ a.s.$$

ここで，Fatou の補題より，$\infty = E[X_1^+] \le \liminf_M E[(X_1^M)^+]$ であるから，$E[X_1^M] = E[(X_1^M)^+] - E[(X_1^M)^-] \to \infty$. 従って，$\liminf_n S_n/n = \infty$ a.s. □

Remark 10.1. 対称性より，$E[X_1^+] < \infty, E[X_1^-] = \infty$ なら，$S_n/n \to -\infty$ a.s. となる．

Corollary 10.2 (収束レート)**.** $\{X_i\}$ を i.i.d. r.v.'s とし, $E[X_1] = 0, E[X_1^2] < \infty$ とする．このとき，各 $\varepsilon > 0$ に対して，$S_n/n^{1/2}(\log n)^{1/2+\varepsilon} \to 0$ a.s. となる．

Proof. $a_n = n^{1/2}(\log n)^{1/2+\varepsilon}$, $n \ge 2$, $a_1 = 1$ とおくと，

$$\sum_n E[(X_n/a_n)^2] = E[X_1^2]\left\{1 + \sum_{n=2}^\infty \frac{1}{n(\log n)^{1+2\varepsilon}}\right\} < \infty.$$

よって，$\sum_n X_n/a_n$ は a.s. に収束する．Kronecker の補題より，$S_n/n \to 0$ a.s. を得る． □

収束レートに関しては，さらに次の定理が成り立つことが知られている．

Theorem 10.4 (Hartman-Winter の重複対数の法則)**.** $\{X_i\}$ を i.i.d. r.v.'s とし, $E[X_i] = 0, E[X_i^2] = 1$ とする．このとき，

$$\limsup_n \frac{S_n}{\sqrt{2n\log\log n}} = 1, \ \liminf_n \frac{S_n}{\sqrt{2n\log\log n}} = -1.$$

が成り立つ．

Remark 10.2. $\log\log n$ が現れるので，重複 “対数” と呼ばれる．重複 “大数” ではない．重複対数の法則はあとで Skorohod の埋め込み定理を用いて証明する．それとは異なる方針の証明は，Stroock (2011, Theorem 1.5.9) を参照せよ．

$E[|X_1|^p] < \infty, 1 < p < 2$ のときは次の定理が成り立つ．

Theorem 10.5 (Marcinkiewicz-Zygmund)**.** $\{X_i\}$ を i.i.d. r.v.'s とし，$E[X_1] = 0$ とする．また，$1 < p < 2$ とする．このとき，$S_n/n^{1/p} \to 0$ a.s. となるための必要十分条件は $E[|X_1|^p] < \infty$ となることである．

Proof. 必要性は演習問題とする．十分性を示す．$Y_k = X_k 1_{\{|X_k| \le k^{1/p}\}}, T_n = \sum_{k=1}^n Y_k$ とおくと，

$$\sum_{k=1}^n P(Y_k \neq X_k) = \sum_{k=1}^n P(|X_k|^p > k) = \sum_{k=1}^\infty P(|X_1|^p > k) \le E[|X_1|^p] < \infty$$

であるから，Borel-Cantelli の補題より，$P(Y_k \neq X_k \text{ i.o.}) = 0$. 従って，$T_n/n^{1/p} \to 0$ a.s. を示せばよい．そのために，Kronecker の補題を使って，$(T_n - E[T_n])/n^{1/p} \to 0$ a.s. を示し，次に，$E[T_n]/n^{1/p} \to 0$ を示す．まず，

$$\sum_{n=1}^\infty \mathrm{Var}(Y_n/n^{1/p}) \le \sum_{n=1}^\infty E[Y_n^2/n^{2/p}] = \sum_{n=1}^\infty n^{-2/p} \int_0^\infty 2yP(|Y_n| > y)dy.$$

定義より，$|Y_n| \le n^{1/p}$ であり，$P(|Y_n| > y) \le P(|X_n| > y) = P(|X_1| > y)$ であるから，

$$\begin{aligned}
&\le \sum_{n=1}^\infty n^{-2/p} \int_0^{n^{1/p}} 2yP(|X_1| > y)dy \\
&= \sum_{n=1}^\infty n^{-2/p} \sum_{k=1}^n \int_{(k-1)^{1/p}}^{k^{1/p}} 2yP(|X_1| > y)dy \\
&= \sum_{k=1}^\infty \int_{(k-1)^{1/p}}^{k^{1/p}} \left(\sum_{n=k}^\infty n^{-2/p} \right) 2yP(|X_1| > y)dy. \quad (\because \text{Fubini の定理})
\end{aligned}$$

ここで，$\sum_{n=k}^\infty n^{-2/p} = O((k-1)^{(p-2)/p})$ $(k \to \infty)$. 一方，

$$\begin{aligned}
\sum_{k=2}^\infty \int_{(k-1)^{1/p}}^{k^{1/p}} (k-1)^{(p-2)/p} yP(|X_1| > y)dy &\le \sum_{k=2}^\infty \int_{(k-1)^{1/p}}^{k^{1/p}} y^{p-1} P(|X_1| > y)dy \\
&= \int_1^\infty y^{p-1} P(|X_1| > y)dy < \infty.
\end{aligned}$$

従って，$\sum_n \mathrm{Var}(Y_n/n^{1/p}) < \infty$ であるから，$\sum_n (Y_n - E[Y_n])/n^{1/p}$ は a.s. に収束する．よって，Kronecker の補題より，$(T_n - E[T_n])/n^{1/p} \to 0$ a.s. を得る．次に，$E[Y_k]$ を評価

する．$E[X_k]=0$ より，$E[Y_k]=-E[X_1 1_{\{|X_1|>k^{1/p}\}}]$ なので，

$$\begin{aligned}|E[Y_k]| &\le E[|X_1|1_{\{|X_1|>k^{1/p}\}}] = k^{1/p}E[(|X_1|/k^{1/p})1_{\{|X_1|>k^{1/p}\}}] \\ &\le k^{1/p}E[(|X_1|/k^{1/p})^p 1_{\{|X_1|>k^{1/p}\}}] \quad (\because |X_1|/k^{1/p}>1 \text{ if } 1_{\{|X_1|>k^{1/p}\}}=1) \\ &= k^{1/p-1}E[|X_1|^p 1_{\{|X_1|^p>k\}}].\end{aligned}$$

$E[|X_1|^p 1_{\{|X_1|^p>k\}}]\downarrow 0$ $(k\to\infty)$ より，$\forall\varepsilon>0, \exists k_0\in\mathbb{N}$ s.t. $k\ge k_0 \Rightarrow E[|X_1|^p 1_{\{|X_k|>k\}}]\le \varepsilon$. 以上より，$n\to\infty$ のとき，

$$\begin{aligned}n^{-1/p}\sum_{k=1}^{n}|E[Y_k]| &\le n^{-1/p}\Big(\sum_{k=1}^{k_0-1}+\sum_{k=k_0}^{n}\Big) \\ &\le o(1)+\varepsilon n^{-1/p}\underbrace{\sum_{k=k_0}^{n}k^{-1/p+1}}_{=O(n^{1/p})} = o(1)+\varepsilon O(1).\end{aligned}$$

$\varepsilon>0$ は任意だったから，$n^{-1/p}\sum_{k=1}^{n}E[Y_k]\to 0$. □

SLLN の応用例

Example 10.1 (再生理論)**.** $\{X_i\}$ を i.i.d. 正値 r.v.'s とし，$T_n=\sum_{i=1}^{n}X_i$ とおく．T_n は n 回目の事象が起こる時刻と考えることができる．例えば，X_i を電球 i の持続時間とすると，T_n は n 番目の電球が切れる時間であって，$N_t=\sup\{n\ge 1: T_n\le t\}$ は時刻 t までに切れる電球の個数である．

Theorem 10.6. $E[X_1]=\theta\le\infty$ とおくと，$t\to\infty$ のとき，

$$N_t/t\to 1/\theta \quad a.s.$$

ただし，$1/\infty=0$ とする．

Proof. SLLN より，$T_n/n\to\theta$ a.s. N_t の定義より，$T_{N_t}\le t<T_{N_t+1}$ であるから，

$$\frac{T_{N_t}}{N_t}\le\frac{t}{N_t}\le\frac{T_{N_t+1}}{N_t+1}\frac{N_t+1}{N_t}.$$

$T_n<\infty$ $(\forall n)$ より，$N_t\uparrow\infty$ $(t\to\infty)$. $\Omega_0=\{T_n/n\to\theta,\ N_t\uparrow\infty\}$ とおくと，$P(\Omega_0)=1$ であって，$\forall\omega\in\Omega_0, T_{N_t(\omega)}(\omega)/N_t(\omega)\to\theta, (N_t(\omega)+1)/N_t(\omega)\to 1$. 従って，$t/N_t\to\theta$ a.s. を得る． □

Example 10.2 (Glivenko-Cantelli の定理)**.** $X_1, X_2, \cdots\sim F$ i.i.d. とする．$X_1,\ldots,X_n$ の経験分布関数 (empirical distribution function) を

$$\hat{F}_n(x):=\hat{F}_n(x,\omega):=\frac{1}{n}\sum_{i=1}^{n}1_{\{X_i(\omega)\le x\}}$$

と定義する．ここで，$E[\hat{F}_n(x)] = F(x)$ であるから，各 $x \in \mathbb{R}$ に対して，$\hat{F}_n(x) \to F(x)$ a.s. が成り立つ． さらに，a.s. に一様収束が成り立つ．

Theorem 10.7 (Glivenko-Cantelli)**.** $\sup_{x\in\mathbb{R}} |\hat{F}_n(x) - F(x)| \to 0$ a.s.

証明の前に，次の補題を示す．

Lemma 10.3. $p \in (0,1)$ に対して，$F(F^{-1}(p)) \geq p, F(F^{-1}(p)-) \leq p$.

Proof. $F^{-1}(p)$ の定義より，$x_m \downarrow F^{-1}(p), F(x_m) \geq p$ なる数列 x_m が選べる．ここで，F の右連続性より，$F(F^{-1}(p)) = \lim_m F(x_m) \geq p$. 次に，$F^{-1}(p)$ の定義より，$x < F^{-1}(p)$ に対して，$F(x) < p$ となる．$x \uparrow F^{-1}(p)$ として，$F(F^{-1}(p)-) \leq p$ を得る． □

Proof of Theorem 10.7. d.f. の右連続性より，$\sup_{x\in\mathbb{R}} |\hat{F}_n(x) - F(x)| = \sup_{x\in\mathbb{Q}} |\hat{F}_n(x) - F(x)|$ となるので，左辺は可測である．SLLN より，各 $x \in \mathbb{R}$ に対して，

$$\hat{F}_n(x-) = \frac{1}{n}\sum_{i=1}^{n} 1_{\{X_i<x\}} \to P(X_1 < x) = F(x-) \quad a.s.$$

となる (除外集合は x に依存する！)．$k \geq 2$ に対して，$x_{0,k} = -\infty$, $x_{j,k} = F^{-1}(j/k), 1 \leq j \leq k-1$, $x_{k,k} = \infty$ として，

$$\Delta_{n,k} = \max_{1\leq j\leq k-1} |\hat{F}_n(x_{j,k}) - F(x_{j,k})| \vee \max_{1\leq j\leq k-1} |\hat{F}_n(x_{j,k}-) - F(x_{j,k}-)|$$

とおくと，$\lim_n \Delta_{n,k} = 0$ a.s. である．$\Omega_k = \{\lim_n \Delta_{n,k} = 0\}$ とおく．

いま，$x_{j,k}$ の定義より，$F(x_{j,k}-) - F(x_{j-1,k}) \leq j/k - (j-1)/k = k^{-1}$ となるから，各 $x \in (x_{j-1,k}, x_{j,k})$ に対して，

$$\begin{aligned}
\hat{F}_n(x) &\leq \hat{F}_n(x_{j,k}-) \leq F(x_{j,k}-) + \Delta_{n,k} \\
&\leq F(x_{j-1,k}) + k^{-1} + \Delta_{n,k} \leq F(x) + k^{-1} + \Delta_{n,k}, \\
\hat{F}_n(x) &\geq \hat{F}_n(x_{j-1,k}) \geq F(x_{j-1,k}) - \Delta_{n,k} \\
&\geq F(x_{j,k}-) - k^{-1} - \Delta_{n,k} \geq F(x) - k^{-1} - \Delta_{n,k}.
\end{aligned}$$

従って，$\sup_{x\in\mathbb{R}} |\hat{F}_n(x) - F(x)| \leq k^{-1} + \Delta_{n,k}$. $\tilde{\Omega} = \bigcap_{k=2}^{\infty} \Omega_k$ とおくと，

$$\limsup_n \sup_{x\in\mathbb{R}} |\hat{F}_n(x,\omega) - F(x)| \leq k^{-1}, \ \forall \omega \in \tilde{\Omega}, \forall k \geq 2.$$

$P(\tilde{\Omega}) = 1$ より，$k \to \infty$ として結論を得る． □

Remark 10.3. 大数の強法則はペアワイズに独立な r.v.'s に対しても成り立つ．すなわち，$\{X_i\}$ をペアワイズに独立な r.v.'s (i.e., 各 $i \neq j$ に対して，X_i と X_j は独立) であって，同一分布に従うとする．このとき，$E[|X_1|] < \infty$ なら，$S_n/n \to E[X_1]$ a.s. が成り立つ (Etemedi, 1981). 詳細は，Durrett (2010, Section 2.4) を参照せよ．

演習問題

Exercise 10.1. Theorem 10.5 の必要性の部分を示せ.

Exercise 10.2. 連続関数 $f : [0,1] \to \mathbb{R}$ に対して,

$$\lim_{n\to\infty} \int_0^1 \cdots \int_0^1 f((x_1 \cdots x_n)^{1/n}) dx_1 \cdots dx_n$$

を計算せよ.

Exercise 10.3. $\{X_i\}$ を i.i.d. r.v.'s とし, $E[|X_1|] < \infty, E[X_1] = 0$ とする. $\{c_i\} \subset \mathbb{R}$ が有界なら, $n^{-1}\sum_{i=1}^n c_i X_i \to 0$ a.s. となることを示せ.

r.v. X が対称であるとは, $X \stackrel{d}{=} -X$ となることを言う.

Exercise 10.4 (Lévy の不等式)**.** $X_1, \ldots, X_n$ を独立な対称 r.v.'s とし, τ_n を $\tau_n(\omega) = \min\{1 \le m \le n : |X_m(\omega)| = \max_{1\le\ell\le n} |X_\ell(\omega)|\}$ と定義する. $Y_n(\omega) = X_{\tau_n(\omega)}(\omega), \hat{S}_n = S_n - Y_n$ とおく. このとき,

$$(\hat{S}_n, Y_n) \stackrel{d}{=} (-\hat{S}_n, Y_n)$$

を示せ. これから,

$$\begin{aligned} P(Y_n \ge x) &\le P(Y_n \ge x, \hat{S}_n \ge 0) + P(Y_n \ge x, \hat{S}_n \le 0) \\ &= 2P(Y_n \ge x, \hat{S}_n \ge 0) \le 2P(S_n \ge x) = P(|S_n| \ge x) \end{aligned}$$

を得る.

Exercise 10.5. $\{X_i\}$ を対称な i.i.d. r.v.'s とする. このとき,

$$\exists C > 0 \text{ s.t. } \lim_n P(|S_n/n| \le C) = 1 \Rightarrow \lim_n nP(|X_1| \ge n) = 0$$

を示せ.

Theorem 8.2 とあわせて, 対称な i.i.d. r.v.'s $\{X_i\}$ に対しては,

$$S_n/n \xrightarrow{P} 0 \Leftrightarrow \lim_n nP(|X_1| \ge n) = 0$$

が成り立つ.

Part III. 中心極限定理

本パートのトピックは中心極限定理 (CLT) である．CLT は，適当な仮定のもとで，正規化した標本平均が標本サイズが大きいとき “近似的に” 正規分布に従うことを述べるものである．もちろん，“近似的に正規分布に従う”，というステートメントはその意味を厳密に定義する必要がある．以下，分布の収束と CLT に関連したトピックを考察していく．

11 弱収束

区間 I 上の関数 H に対して，$\mathcal{C}(H) = \{x \in I : \ H \text{ は } x \text{ で連続}\}$ と定義する．

$F, F_n, n \in \mathbb{N}$ を d.f.'s とすると，F_n が F に 弱収束する (converge weakly) とは，あらゆる $x \in \mathcal{C}(F)$ に対して，$\lim_n F_n(x) = F(x)$ となることを言う．このとき，$F_n \overset{w}{\to} F$ と書く．$(\mathbb{R}, \mathcal{B})$ 上の p.m.'s $\mu, \mu_n, n \in \mathbb{N}$ に対して，その d.f.'s をそれぞれ $F, F_n, n \in \mathbb{N}$ とおくと，$F_n \overset{w}{\to} F$ のとき，μ_n は μ に弱収束すると言って，$\mu_n \overset{w}{\to} \mu$ と書く．r.v.'s $X_n \sim F_n, n \in \mathbb{N}$, $X \sim F$ に対して，$F_n \overset{w}{\to} F$ のとき，X_n は X に 分布収束する (converge in distribution) と言って，$X_n \overset{d}{\to} X$ と書く．やや記号の濫用であるが，$X_n \overset{d}{\to} \mu$ と書いたりもする．もちろん，この意味は，$X \sim \mu$ に対して，$X_n \overset{d}{\to} X$ となることである．

F_n の弱収束極限は存在すれば一意である．なぜなら，$F_n \overset{w}{\to} F$, $F_n \overset{w}{\to} G$ とすると，$F(x) = G(x)\ \forall x \in \mathcal{C}(F) \cap \mathcal{C}(G) =: C$ だが，F, G の不連続点は高々可算個しかないので，C は $\mathbb{R}$ で稠密である．よって，$\forall x \in \mathbb{R},\ \exists x_m \in C$ s.t. $x_m \downarrow x$ であり，d.f. の右連続性より，

$$F(x) = \lim_m F(x_m) = \lim_m G(x_m) = G(x).$$

従って，$F = G$ である．

弱収束の定義において，F の不連続点に対しては，F_n は F に収束しなくてもよい．

Example 11.1. $X \sim F, X_n = X + 1/n$ とすると，F の不連続点 x に対して，

$$F_n(x) := P(X_n \le x) = P(X \le x - 1/n) = F(x - 1/n) \to F(-x) \neq F(x).$$

ただし，F の連続点 x に対しては，$F_n(x) \to F(x)$ であるから，$F_n \overset{w}{\to} F$.

一方，F が連続なとき，弱収束は各点収束を意味する．じつはもっと強いことが言える．

Theorem 11.1 (Pólya. Durrett 3.2.9)**.** $F_n \overset{w}{\to} F$ とし，F を連続とする．このとき，

$$\sup_{x \in \mathbb{R}} |F_n(x) - F(x)| \to 0,\ n \to \infty$$

となる．

分布収束は d.f.'s の弱収束によって定義されるので，X_n と X が同じ確率空間で定義されている必要はない．従って，分布収束は，a.s. 収束，L^p 収束，確率収束のいずれも含意しない．しかし，$F_n \xrightarrow{w} F$ のとき，F_n, F をそれぞれ d.f. にもつ r.v.'s Y_n, Y (Y_n, Y は同じ確率空間に定義されている必要がある) であって，$Y_n \to Y$ a.s. となるものを構成できる．

Theorem 11.2 (Skorohod の表現定理)**.** $F_n \xrightarrow{w} F$ のとき，$\exists Y_n \sim F_n, \exists Y \sim F$ s.t. $Y_n \to Y$ a.s.

Remark 11.1. 言い換えると，$X_n \xrightarrow{d} X$ のとき，適当な確率空間 (X_n, X が定義された確率空間とは違ってもよい) 上に $Y_n \overset{d}{=} X_n, Y \overset{d}{=} X$ なる r.v.'s であって，$Y_n \to Y$ a.s. なるものを構成できる．$X_n \to X$ a.s. は 言えない ので注意する．

Skorohod の表現定理の証明に次の定理を使う．

Theorem 11.3. $F_n \xrightarrow{w} F$ なら，$F_n^{-1}(p) \to F^{-1}(p)$ $(\forall p \in \mathcal{C}(F^{-1}))$.

Remark 11.2. この定理自体，結構重要な結果である．例えば，$X_1, X_2, \cdots \sim F$ i.i.d. とし，$X_1, \ldots, X_n$ の経験分布関数を $\hat{F}_n$ とすると，Glivenko-Cantelli の定理より，$\sup_{x \in \mathbb{R}} |\hat{F}_n(x) - F(x)| \to 0$ a.s. 一様収束が成り立つ ω に対して，$\hat{F}_n(\cdot, \omega) \xrightarrow{w} F$ $(n \to \infty)$ となるので，各 $p \in \mathcal{C}(F^{-1})$ に対して，

$$\hat{F}_n^{-1}(p) \to F^{-1}(p) \quad a.s.$$

すなわち，$\hat{F}_n^{-1}(p)$ は $F^{-1}(p)$ の一致推定量となる．

Proof of Theorem 11.3. $\varepsilon > 0, p \in \mathcal{C}(F^{-1})$ を任意に固定する．F の不連続点は高々可算個しかないので，$(F^{-1}(p) - \varepsilon, F^{-1}(p)) \cap \mathcal{C}(F) \neq \varnothing$. $x \in (F^{-1}(p) - \varepsilon, F^{-1}(p)) \cap \mathcal{C}(F)$ に対して，$x < F^{-1}(p)$ より，$F(x) < p$. 一方，$x \in \mathcal{C}(F)$ より，$F_n(x) \to F(x)$ であるから，十分大きな n に対して，$F_n(x) < p$. よって，$F_n^{-1}(p) > x > F^{-1}(p) - \varepsilon$. $n \to \infty$ として，$\liminf_n F_n^{-1}(p) \geq F^{-1}(p) - \varepsilon$.

逆の不等式を示す．$\forall p' > p, \exists x \in \mathcal{C}(F)$ s.t. $F^{-1}(p') < x < F^{-1}(p') + \varepsilon$. 左側の不等式より，$F(x) \geq p' > p$. $x \in \mathcal{C}(F)$ より，$F_n(x) \to F(x)$ であるから，十分大きな n に対して，$F_n(x) \geq p$. よって, $F_n^{-1}(p) \leq x < F^{-1}(p') + \varepsilon$. $n \to \infty$ として，$\limsup_n F_n^{-1}(p) \leq F^{-1}(p') + \varepsilon$. $p \in \mathcal{C}(F^{-1})$ より，$p' \downarrow p$ として，$\limsup_n F^{-1}(p) \leq F^{-1}(p) + \varepsilon$ を得る． □

Proof of Theorem 11.2. 確率空間を

$$(\Omega, \mathcal{F}, P) = ((0,1), \mathcal{B}_{(0,1)}, U(0,1))$$

とする．また，$U(\omega) = \omega, \omega \in \Omega$ とする．このとき，$U \sim U(0,1)$ であるから，

$$Y_n := F_n^{-1}(U) \sim F_n, \; Y := F^{-1}(U) \sim F.$$

F^{-1} の不連続点は高々可算個しかないので，$F_n^{-1}(U) \to F^{-1}(U)$ a.s. □

以下，弱収束・分布収束の性質を調べて行く．

$$C_b(\mathbb{R}) = \{f : f : \mathbb{R} \to \mathbb{R},\ f \text{ は有界連続 }\}$$

とおく．

Theorem 11.4. $X_n \xrightarrow{d} X \Leftrightarrow \lim_n E[g(X_n)] = E[g(X)]$ $(\forall g \in C_b(\mathbb{R}))$.

Proof. $\Rightarrow$. Skorohod の表現定理より明らか．

$\Leftarrow$. $x \in \mathcal{C}(F)$, $\varepsilon > 0$ を任意に固定する．$g_{x,\varepsilon} : \mathbb{R} \to [0,1]$ を連続で，

$$g_{x,\varepsilon}(y) = \begin{cases} 1 & \text{if } y \le x \\ 0 & \text{if } y > x + \varepsilon \end{cases}$$

をみたすものとすると，$1_{(-\infty,x]} \le g_{x,\varepsilon} \le 1_{(-\infty,x+\varepsilon]}$ より，

$$\limsup_n F_n(x) \le \lim_n E[g_{x,\varepsilon}(X_n)] = E[g_{x,\varepsilon}(X)] \le F(x+\varepsilon).$$

一方，$1_{(-\infty,x-\varepsilon]} \le f_{x-\varepsilon,\varepsilon} \le 1_{(-\infty,x]}$ であるから，

$$\liminf_n F_n(x) \ge \lim_n E[g_{x-\varepsilon,\varepsilon}(X_n)] = E[g_{x-\varepsilon,\varepsilon}(X)] \ge F(x-\varepsilon).$$

$x \in \mathcal{C}(F)$ より，$\varepsilon \downarrow 0$ のとき，$F(x-\varepsilon), F(x+\varepsilon) \to F(x)$. □

この定理より，次の系が直ちに従う (“確率収束” $\Rightarrow$ “分布収束” の部分だけが新しい)．

Corollary 11.1. “a.s. 収束” or “L^p 収束” $\Rightarrow$ “確率収束” $\Rightarrow$ “分布収束”.

一般に，“分布収束” $\not\Rightarrow$ “確率収束” だが，分布収束先が定数の場合は逆も成り立つ．

Lemma 11.1 (Durrett 3.2.12)**.** $X_n \xrightarrow{d} c$: 定数なら，$X_n \xrightarrow{P} c$.

次の 2 つの結果は統計学においてよく用いられる．

Theorem 11.5 (連続写像定理)**.** $g : \mathbb{R} \to \mathbb{R}$ を可測関数とし，$D_g = \{x :\ g \text{ は } x \text{ で不連続 }\}$ とおく．このとき，$X_n \xrightarrow{d} X$ & $P(X \in D_g) = 0$ なら，$g(X_n) \xrightarrow{d} g(X)$.

Proof. Skorohod の表現定理より，$\exists Y_n \overset{d}{=} X_n$, $\exists Y \overset{d}{=} X$ s.t. $Y_n \to Y$ a.s. $f \in C_b(\mathbb{R})$ を任意にとると，$D_{f\circ g} \subset D_g$ より，$P(Y \in D_{f\circ g}) = 0$. よって，$f(g(Y_n)) \to f(g(Y))$ a.s. DCT より，$E[f(g(X_n))] = E[f(g(Y_n))] \to E[f(g(Y))] = E[f(g(X))]$. □

Lemma 11.2 (Slutsky の補題. Durrett 3.2.13, 3.2.14)**.** $X_n \xrightarrow{d} X$, $Y_n \xrightarrow{P} c$ なら，$X_n + Y_n \xrightarrow{d} X + c$, $Y_n X_n \xrightarrow{d} cX$.

分布収束はいくつか同値な条件が知られている．それらをまとめて述べておく．

Theorem 11.6 (Portmanteau)**.** 次の (i)–(iv) は同値である．

(i) $X_n \xrightarrow{d} X$.

(ii) $\forall G$: open, $\liminf_n P(X_n \in G) \geq P(X \in G)$.

(iii) $\forall K$: closed, $\limsup_n P(X_n \in K) \leq P(X \in K)$.

(iv) $\forall A \in \mathcal{B}$ s.t. $P(X \in \partial A) = 0$, $\lim_n P(X_n \in A) = P(X \in A)$.

Remark 11.3. “Portmanteau” は人名ではない (Google 検索してみよ)．要は，分布収束のいろいろな同値条件をひとまとめに述べているためこんな名前がついたようである．最初に証明したのは，Aleksandr Aleksandrov であるが，“Portmanteau” の命名者は Billingsley (1968) のようである (Resnick, 1998, p.264)．なお，Billingsley (1968) の第二版は，Jean Pierre Portmanteau (1915, “Espoir pour l’ensemble vide”, Annals of the University of Felletin) を引用しているが，これはジョークである (“Espoir pour l’ensemble vide” を訳すと “空集合を期待して”)．

Proof. (i) $\Rightarrow$ (ii) $\Leftrightarrow$ (iii), (ii)+(iii) $\Rightarrow$ (iv) $\Rightarrow$ (i) を示す．

(i) $\Rightarrow$ (ii). Skorohod の表現定理より，$\exists Y_n \overset{d}{=} X_n$, $\exists Y \overset{d}{=} X$ s.t. $Y_n \to Y$ a.s. G は open なので，

$$1_G(Y) \leq \liminf_n 1_G(Y_n) \quad a.s.$$

従って，Fatou の補題より，

$$P(X \in G) = P(Y \in G) \leq \liminf_n P(Y_n \in G) = \liminf_n P(X_n \in G).$$

(ii) $\Leftrightarrow$ (iii). 明らか．

(ii) + (iii) $\Rightarrow$ (iv). $K = \bar{A}, G = A^\circ$ とすれば，$\partial A = K \setminus G, P(X \in \partial A) = 0$ より，

$$P(X \in K) = P(X \in A) = P(X \in G).$$

従って，(ii) と (iii) より，

$$\begin{aligned}
&\limsup_n P(X_n \in A) \leq \limsup_n P(X_n \in K) \leq P(X \in K) = P(X \in A), \\
&\liminf_n P(X_n \in A) \geq \liminf_n P(X_n \in G) \geq P(X \in G) = P(X \in A).
\end{aligned}$$

(iv) $\Rightarrow$ (i). $x \in \mathcal{C}(F), A = (-\infty, x]$ とすると，$P(X \in \partial A) = P(X = x) = 0$. □

Remark 11.4. 弱収束は距離化可能である．すなわち，$\mathcal{D}$ を d.f.'s の全体とすると，$\mathcal{D}$ 上の適当な距離 ρ に対して，

$$F_n \xrightarrow{w} F \Leftrightarrow \rho(F_n, F) \to 0 \qquad (*)$$

が成り立つ．このような距離 ρ はいくつかある．例えば，$F, G \in \mathcal{D}$ に対して，

$$\rho(F,G) := \inf\{\varepsilon > 0 : F(x-\varepsilon) - \varepsilon \le G(x) \le F(x+\varepsilon) + \varepsilon,\ \forall x \in \mathbb{R}\}$$

と定義すると，ρ は $(*)$ をみたす $\mathcal{D}$ 上の距離である (Durrett 3.2.6)．この距離を Lévy 距離と呼ぶ．また，Lévy 距離の定める距離位相を 弱位相 (weak topology) と呼ぶ．

d.f.'s $\{F_n\}$ が 一様にタイト (uniformly tight) であるとは，$\forall \varepsilon > 0, \exists M_\varepsilon > 0$ s.t.

$$1 - F_n(M_\varepsilon) + F_n(-M_\varepsilon) \le \varepsilon,\ \forall n \in \mathbb{N}$$

となることを言う．

Theorem 11.7 (Prohorov)**.** d.f.'s $\{F_n\}$ の任意の部分列が弱収束するさらなる部分列をもつための必要十分条件は，$\{F_n\}$ が一様にタイトとなることである．

明らかに，F_n が弱収束していれば，$\{F_n\}$ は一様にタイトである．Prohorov の定理の証明の前に，それが導く次の重要な事実を確認しておく．

Corollary 11.2 (重要！). d.f.'s $\{F_n\}$ が弱収束するための必要十分条件は，$\{F_n\}$ が一様にタイトかつ，あらゆる弱収束部分列が共通の極限をもつことである．

Proof of Corollary 11.2. 必要性は明らか．十分性を示す．$\{F_n\}$ の弱収束部分列の共通の極限を F とおく．$g \in C_b(\mathbb{R})$ を任意に固定し，$a_n = \int g dF_n, a = \int g dF$ とおく．$a_n \to a$ を示せばよい．a_n の任意の部分列 $a_{n(m)}$ に対して，$\{F_n\}$ の一様タイト性より，さらなる部分列 $a_{n(m_k)}$ が存在して，$a_{n(m_k)} \to a$ となる．これは，$a_n \to a$ を意味する． □

Proof of Theorem 11.7. 十分性．$\{F_n\}$ が弱収束する部分列をもつことを示せばよい．対角線論法を用いる．$\mathbb{Q}$ は可算なので，$\mathbb{Q} = \{q_1, q_2, \dots\}$ と番号を付ける．まず，$F_n(q_1) \in [0,1]$ であるから，Bolzano-Weierstrass の定理より，$\{F_n\}$ のある部分列 $\{F_{n_1(k)}\}_{k=1}^\infty$ に対して，$\lim_k F_{n_1(k)}(q_1)$ が存在する．同様に，$\{F_{n_1(k)}\}$ のある部分列 $\{F_{n_2(k)}\}_{k=1}^\infty$ に対して，$\lim_k F_{n_2(k)}(q_2)$ が存在する．$\{F_{n_2(k)}\}$ は $\{F_{n_1(k)}\}$ の部分列なので，$\lim_k F_{n_2(k)}(q_1)$ も存在する．この操作を繰り返すと，m 番目の部分列 $\{F_{n_m(k)}\}_{k=1}^\infty$ に対しては，$\lim_k F_{n_m(k)}(q_\ell)$ $(\ell = 1, \dots, m)$ が存在する．そこで，$F_{n(k)} := F_{n_k(k)}$ $(k \in \mathbb{N})$ とおくと，すべての $q \in \mathbb{Q}$ に対して，$\lim_k F_{n(k)}(q) =: F(q)$ が存在する．

$$\begin{array}{cccccc}
F_{n_1(1)} & F_{n_1(2)} & \cdots & F_{n_1(m)} & \cdots \\
F_{n_2(1)} & F_{n_2(2)} & \cdots & F_{n_m(m)} & \cdots \\
\vdots & & \ddots & \vdots & \\
F_{n_m(1)} & \cdots & \cdots & F_{n_m(m)} & \\
\vdots & & & & \ddots
\end{array}$$

F は $\mathbb{Q}$ 上でしか定義されていないが,

$$\tilde{F}(x) := \inf_{q>x, q\in\mathbb{Q}} F(q),\ x \in \mathbb{R}$$

とおけば，$\tilde{F}$ は $\mathbb{R}$ 上に定義された非減少関数である．$\tilde{F}$ がさらに右連続であることを示す．$x_0 \in \mathbb{R}$ を任意に固定する．$\tilde{F}$ の定義より，任意の $\varepsilon > 0$ に対して，$q > x_0, F(q) < \tilde{F}(x_0)+\varepsilon$ をみたす $q \in \mathbb{Q}$ が存在する．$q > x$ より，$\tilde{F}(x_0) \le F(q)$ でもあるから，$F(q)-\varepsilon < \tilde{F}(x_0) \le F(q)$ となる．このとき，$x_0 < x < q$ なる x に対して,

$$F(q) - \varepsilon < \tilde{F}(x_0) \le \tilde{F}(x) \le F(q)$$

となるから，$\tilde{F}(x) - \tilde{F}(x_0) \le \varepsilon$. 従って，$\tilde{F}$ は右連続であり，特に，$\tilde{F}(q) = F(q)$ $(\forall q \in \mathbb{Q})$ を得る．次に，$\lim_{x\to\infty} \tilde{F}(x) = 1$, $\lim_{x\to-\infty} \tilde{F}(x) = 0$ を示す．定義より，$0 \le \tilde{F} \le 1$ なのはよい．$\{F_n\}$ は一様にタイトなので，任意の $\varepsilon > 0$ に対して，M を十分大きな有理数とすると,

$$1 - F_{n(k)}(M) + F_{n(k)}(-M) \le \varepsilon\ \forall k$$

となる．$k \to \infty$ として，$1-\tilde{F}(M)+\tilde{F}(-M) \le \varepsilon$ を得る．特に，$\tilde{F}(M) \ge 1-\varepsilon, \tilde{F}(-M) \le \varepsilon$ であり，$\tilde{F}$ の単調性より，$\lim_{x\to\infty} \tilde{F}(x) = 1, \lim_{x\to-\infty} \tilde{F}(x) = 0$ を得る．以上より，$\tilde{F}$ が d.f. であることが示された．

$F_{n(k)} \xrightarrow{w} \tilde{F}$ を示そう．$x \in \mathcal{C}(\tilde{F})$ として，$q, q' \in \mathbb{Q}$ を $q < x < q'$ となるように選ぶと,

$$\tilde{F}(q) = \lim_k F_{n(k)}(q) \le \liminf_k F_{n(k)}(x) \le \limsup_k F_{n(k)}(x) \le \lim_k F_{n(k)}(q') = \tilde{F}(q')$$

となる．x は $\tilde{F}$ の連続点なので，$q \uparrow x, q' \downarrow x$ として，$\tilde{F}(x) = \lim_k F_{n(k)}(x)$ を得る．

必要性． $\{F_n\}$ が一様にタイトでないとする．このとき，$\exists\varepsilon > 0,\ \exists n(k) \uparrow \infty$ s.t.

$$1 - F_{n(k)}(k) + F_{n(k)}(-k) \ge \varepsilon,\ \forall k.$$

仮定より，$F_{n(k)}$ は弱収束するさらなる部分列 $F_{n(k_j)}$ をもつ．$F_{n(k_j)} \xrightarrow{w} F$ とする．$r < 0 < s$ を F の連続点とすると,

$$\begin{aligned} 1 - F(s) + F(r) &= \lim_j \{1 - F_{n(k_j)}(r) + F_{n(k_j)}(s)\} \\ &\ge \liminf_j \{1 - F_{n(k_j)}(k_j) + F_{n(k_j)}(k_j)\} \ge \varepsilon. \end{aligned}$$

$r \to -\infty,\ s \to \infty$ として，$0 = 1 - F(\infty) + F(-\infty) \ge \varepsilon$. 矛盾． □

Remark 11.5. 証明からわかるように，$\{F_n\}$ が一様にタイトでなくても，適当な部分列 $F_{n(k)}$ と右連続非減少関数 F が存在して，$\lim_k F_{n(k)}(x) = F(x)$ $(\forall x \in \mathcal{C}(F))$ が言える (この部分は Helly の定理 と呼ばれる)．ただし，F は d.f. とは限らない．例えば，G を d.f. と

し, $a, b, c > 0$ を $a+b+c = 1$ となるように選び，$F_n(x) = a1_{[n,\infty)}(x)+b1_{[-n,\infty)}(x)+cG(x)$ とおくと，$\lim_n F_n(x) = b + cG(x) =: F(x)$ であるが，

$$\lim_{x\to-\infty} F(x) = b, \ \lim_{x\to\infty} F(x) = b + c = 1 - a,$$

と，$\pm\infty$ で測度が逃げる．一様タイト性はそのようなことが起こらないための条件である．

Remark 11.6 (Prohorov の定理の意味)．$(\mathbb{R}, \mathcal{B})$ 上の p.m.'s の集合 Π が一様にタイトであるとは，任意の $\varepsilon > 0$ に対して，$M = M_\varepsilon > 0$ を十分大きくとれば，$\mu([-M, M]^c) \leq \varepsilon$ $(\forall \mu \in \Pi)$ となることを言う．Prohorov の定理は，次の (i) と (ii) が同値であることを述べている．

(i) Π が一様にタイトである．

(ii) Π 内の任意の p.m.'s の列が弱収束する部分列をもつ．

$(\mathbb{R}, \mathcal{B})$ 上の p.m.'s の弱収束は，例えば Lévy 距離 ρ によって距離化可能であり，性質 (ii) は Π が ρ に関して相対コンパクトであることを意味する．すなわち，Prohorov の定理は，$(\mathbb{R}, \mathcal{B})$ 上の p.m.'s の集合に対して，一様タイト性と弱位相に関する相対コンパクト性が同値であることを述べている．この同値性はもっと一般に Poland 空間上の p.m.'s の集合に対しても成り立つ．詳細に関しては Part VII を参照せよ．

演習問題

Exercise 11.1 (Schéffe)．f_n, f を確率密度関数とし，$f_n \to f$ λ-a.e. とする．このとき，

$$\sup_{B\in\mathcal{B}} \left| \int_B f_n(x)dx - \int_B f(x)dx \right| = \frac{1}{2} \int_{\mathbb{R}} |f_n(x) - f(x)|dx \to 0$$

を示せ．特に，密度関数の収束は分布収束を含意する．

Exercise 11.2. $U_1, U_2, \cdots \sim U(0, c)$ i.i.d. とする．ただし，$c > 0$ であり，$U(0, c)$ は $(0, c)$ 上の一様分布を表す．いま，$M_n = \sum_{k=1}^n \prod_{i=1}^k U_i, W_n = \sum_{k=1}^n \prod_{i=k}^n U_i$ とおく．

(a) M_n は n に関して非減少なので，$M_\infty = \lim_{n\to\infty} M_n$ は $[0, \infty]$ の範囲で存在する．このとき，$M_\infty < \infty$ a.s. となるための必要十分条件は $c < e$ であることを示せ．

(b) $c < e$ のとき，$W_n \xrightarrow{d} M_\infty$ であるが，W_n は a.s. 収束しないことを示せ．

Exercise 11.3. $X_{m,n}, X_n$, $m, n = 1, 2, \ldots$ を r.v.'s とし，各 m に対して，$X_{m,n} \xrightarrow{d} Z_m$ $(n \to \infty)$ とする．さらに，$Z_m \xrightarrow{d} X$ $(m \to \infty)$ とする．このとき，任意の $\varepsilon > 0$ に対して，

$$\lim_{m\to\infty} \limsup_{n\to\infty} P(|X_{m,n} - X_n| > \varepsilon) = 0$$

ならば，$X_n \xrightarrow{d} X$ となることを示せ．

Exercise 11.4 (Wasserstein距離)**.** $1 \leq p < \infty$ に対して，$\mathcal{D}_p = \{F : F \text{ は d.f.}, \int |x|^p dF(x) < \infty\}$ とおく．いま，$F, G \in \mathcal{D}_p$ に対して，

$$d_p(F,G) := \left\{ \int_0^1 |F^{-1}(u) - G^{-1}(u)|^p du \right\}^{1/p}$$

と定義すると，d_p は $\mathcal{D}_p$ 上の完備な距離になることを示せ．さらに，$F, F_n \in \mathcal{D}_p$ に対して，

$$d_p(F_n, F) \to 0 \Leftrightarrow F_n \xrightarrow{w} F \ \& \ \int |x|^p dF_n(x) \to \int |x|^p dF(x)$$

を示せ．この距離 d_p は p 次 Wasserstein 距離 と呼ばれる．d_2 は Mallows 距離 とも呼ばれる．

Exercise 11.5 (Rubin)**.** $X_n \xrightarrow{d} X$ とし，$g, g_n : \mathbb{R} \to \mathbb{R}, n \in \mathbb{N}$ を，任意の $x \in \mathbb{R}$ と $x_n \to x$ なる任意の $\{x_n\}$ に対して，$\lim_n g_n(x_n) = g(x)$ をみたす可測関数とする．このとき，$g_n(X_n) \xrightarrow{d} g(X)$ を示せ．

Exercise 11.6. r.v.'s $\{X_n : n \in \mathbb{N}\}$ に対して，$\{\mathcal{L}(X_n) : n \in \mathbb{N}\}$ が一様にタイトであるためには，$c_n \to 0$ なる任意の定数列 c_n に対して，$c_n X_n \xrightarrow{P} 0$ となることが必要十分であることを示せ．

12 Lindeberg-Feller の CLT

独立確率変数の正規化した和が標準正規分布に分布収束することを述べる定理 (群) を中心極限定理 (central limit theorem, CLT) と呼ぶ．本節の主目標は Lindeberg の CLT の証明を与えることである．Lindeberg の CLT の証明はいくつかバージョンがあり，Durrett (2010) を含む多くの教科書は (連続性定理を示したあとに)，特性関数の展開にもとづく証明を与えている．本講義ノートはそのルートはとらず，Lindeberg のもともとのアイデアに基づく (個人的にはより単純だと思われる) 証明を与える．

Theorem 12.1 (Lindeberg)**.** 各 $n \in \mathbb{N}$ に対して，$X_{n,m}, 1 \le m \le n$ は独立な r.v.'s であって，$E[X_{n,m}] = 0$ $(1 \le \forall m \le n)$, $\sum_{m=1}^n E[X_{n,m}^2] = 1$ をみたすとする．さらに

$$\forall \varepsilon > 0,\ g_n(\varepsilon) := \sum_{m=1}^n E[X_{n,m}^2 1_{\{|X_{n,m}|>\varepsilon\}}] \to 0,\ n \to \infty \qquad (*)$$

を仮定する．このとき，$S_n = \sum_{m=1}^n X_{n,m} \xrightarrow{d} N(0,1)$.

$(*)$ は Lindeberg 条件 と呼ばれる．追加的な仮定の下で，Lindeberg 条件は CLT が成り立つための必要条件でもある．$r_n^2 := \max_{1\le m\le n} E[X_{n,m}^2]$ とおく．

Theorem 12.2 (Feller)**.** $\{X_{n,m}\}$ を Theorem 12.1 の r.v.'s とし (Lindeberg 条件は仮定しない)，$r_n \to 0$ を仮定する．このとき，$S_n \xrightarrow{d} N(0,1)$ なら，Lindeberg 条件が成り立つ．

Theorem 12.1 と Theorem 12.2 はまとめて，Lindeberg-Feller の CLT と呼ばれる [9]．Lindeberg 条件は $r_n \to 0$ を含意する．実際，$r_n^2 \le \varepsilon^2 + g_n(\varepsilon)$ であるから，Lindeberg 条件が成り立つなら，$\limsup_n r_n \le \varepsilon$ $(\forall \varepsilon > 0)$ となって，$r_n \to 0$ を得る．Feller の定理の証明は次節に与える．次の系は最も基本的な CLT である．

Corollary 12.1. $\{X_m : m \in \mathbb{N}\}$ を i.i.d. r.v.'s とし，$E[X_1] = 0, E[X_1^2] = \sigma^2 \in (0,\infty)$ とする．このとき，$\sum_{m=1}^n X_m/(\sigma\sqrt{n}) \xrightarrow{d} N(0,1)$.

Proof. $X_{n,m} = X_m/(\sigma\sqrt{n})$ に対して，Theorem 12.1 を適用すればよい． □

Lindeberg の CLT の証明に移る．本質的には次の定理から直ちに従う．

Theorem 12.3. $\{X_{n,m}\}$ を Theorem 12.1 の r.v.'s とし，$Z \sim N(0,1)$ とする．このとき，$\forall \varepsilon > 0,\ \forall h \in C^3(\mathbb{R})$ s.t. $\|h''\|_u \vee \|h'''\|_u < \infty$,

$$|E[h(S_n)] - E[h(Z)]| \le \left(\frac{\varepsilon}{6} + \frac{r_n}{2}\right)\|h'''\|_u + g_n(\varepsilon)\|h''\|_u.$$

[9] CLT の歴史に関しては，Le Cam (1986) が興味深い．

Proof. 記号の簡単のために，$X_{n,m} = X_m$ と書く．まず，Taylor の定理より，

$$|h(x)| \le |h(0)| + |h'(0)||x| + \|h''\|_u \frac{x^2}{2}$$

であるから，$E[h(S_n)]$ と $E[h(Z)]$ は有限である．$Y_1, \dots, Y_n$ を独立な r.v.'s であって，各 m に対して，$Y_m \sim N(0, E[X_m^2])$ とする．また，$(Y_1, \dots, Y_n)$ は $(X_n, \dots, X_n)$ と独立とする．$T_n = \sum_{m=1}^n Y_m$ とおくと，$T_n \sim N(0,1)$ より，

$$\Delta := |E[h(S_n)] - E[h(Z)]| = |E[h(S_n)] - E[h(T_n)]|.$$

さらに，$U_m = \sum_{k=1}^{m-1} Y_k + \sum_{k=m+1}^n X_k$ とおくと，

$$\Delta \le \sum_{m=1}^n \underbrace{|E[h(U_m + X_m)] - E[h(U_m + Y_m)]|}_{=:\Delta_m} = \sum_{m=1}^n \Delta_m.$$

Δ_m を評価する．

$$R_m(x) = h(U_m + x) - h(U_m) - xh'(U_m) - \frac{x^2}{2} h''(U_m)$$

とおくと，

$$\begin{aligned} E[h(U_m + X_m)] &= E\left[h(U_m) + X_m h'(U_m) + \frac{X_m^2}{2} h''(U_m) + R_m(X_m)\right] \\ &= E[h(U_m)] + \frac{E[X_m^2]}{2} E[h''(U_m)] + E[R_m(X_m)], \quad (\because X_m \perp\!\!\!\perp U_m,\ E[X_m] = 0) \\ E[h(U_m + Y_m)] &= E[h(U_m)] + \frac{E[Y_m^2]}{2} E[h''(U_m)] + E[R_m(Y_m)]. \end{aligned}$$

ここで，$E[Y_m^2] = E[X_m^2]$ より，

$$\Delta_m = |E[R_m(X_m)] - E[R_m(Y_m)]| \le E[|R_m(X_m)|] + E[|R_m(Y_m)|]$$

となる．次の評価を使う．

Lemma 12.1. $f \in C^{n+1}(\mathbb{R})$ に対して，

$$\left| f(x) - \sum_{m=0}^n \frac{f^{(m)}(0)}{m!} x^m \right| \le \min\left\{ \|f^{(n+1)}\|_u \frac{|x|^{n+1}}{(n+1)!}, \|f^{(n)}\|_u \frac{2|x|^n}{n!} \right\}.$$

Proof of Lemma 12.1. Taylor の定理より，

$$f(x) - \sum_{m=0}^n \frac{f^{(m)}(0)}{m!} x^m = \frac{1}{n!} \int_0^x (x-s)^n f^{(n+1)}(s) ds.$$

また部分積分より，

$$\int_0^x (x-s)^n f^{(n+1)}(s)ds = -x^n f^{(n)}(0) + n\int_0^x (x-s)^{n-1} f^{(n)}(s)ds$$
$$= n\int_0^x (x-s)^{n-1}\{f^{(n)}(s) - f^{(n)}(0)\}ds.$$

この 2 つの評価を合わせて，補題の結論を得る． □

この評価を使うと，

$$|R_m(x)| \le \min\left\{\|h'''\|_u \frac{|x|^3}{6}, \|h''\|_u x^2\right\}$$

であるから，

$$\sum_{m=1}^n E[|R_m(X_m)|] \le \frac{\|h'''\|_u}{6}\sum_{m=1}^n E[|X_m|^3 1_{\{|X_m|\le\varepsilon\}}] + \|h''\|_u \sum_{m=1}^n E[X_m^2 1_{\{|X_m|>\varepsilon\}}]$$
$$\le \frac{\varepsilon\|h'''\|_u}{6} + \|h''\|_u g_n(\varepsilon).$$

同様に，

$$\sum_{m=1}^n E[|R_m(Y_m)|] \le \frac{\|h'''\|_u}{6}\sum_{m=1}^n E[|Y_m|^3]$$
$$= \frac{2\|h'''\|_u}{3\sqrt{2\pi}}\sum_{m=1}^n (E[|Y_m|^2])^{3/2} \quad (\because E[|Z|^3] = 4/\sqrt{2\pi} \text{ for } Z\sim N(0,1))$$
$$\le \frac{r_n\|h'''\|_u}{2}.$$

以上の評価を合わせて，求める結果を得る． □

Proof of Theorem 12.1. $x\in\mathbb{R}, \varepsilon>0$ を任意に固定する．$h_{x,\varepsilon}:\mathbb{R}\to[0,1]$ を C^3 級で

$$h_{x,\varepsilon}(y) = \begin{cases} 1 & \text{if } y\le x \\ 0 & \text{if } y > x+\varepsilon \end{cases}$$

をみたすものとする．$1_{(-\infty,x]} \le h_{x,\varepsilon} \le 1_{(-\infty,x+\varepsilon]}$ より，

$$P(S_n\le x) \le E[h_{x,\varepsilon}(S_n)] = E[h_{x,\varepsilon}(Z)] + o(1) \le P(Z\le x+\varepsilon) + o(1).$$

同様に，$P(S_n\le x) \ge E[h_{x-\varepsilon,x}(S_n)] \ge P(Z\le x-\varepsilon) - o(1)$. □

Theorem 12.1 において，$N(0,1)$ の d.f. を Φ とおくと，Φ は連続なので，Pólya の定理 (Theorem 11.1) より，

$$\sup_{x\in\mathbb{R}} |P(S_n \le x) - \Phi(x)| \to 0$$

を得る．Theorem 12.1 の証明を検討していくと，次のバウンドを導出することもできる (cf Exercise 12.9)：

$$\sup_{x\in\mathbb{R}} |P(S_n \le x) - \Phi(x)| \le A \left(\sum_{m=1}^{n} E[|X_{m,n}|^3] \right)^{1/4}.$$

ここで，$A > 0$ は絶対定数である．だたし，もっと精緻な評価が成り立つことが知られている．

Theorem 12.4 (Berry-Esseen の定理)**.** $\{X_{n,m}\}$ を Theorem 12.1 の r.v.'s とすると，次のバウンドが成り立つ：

$$\sup_{x\in\mathbb{R}} |P(S_n \le x) - \Phi(x)| \le A \sum_{m=1}^{n} E[|X_{m,n}|^3].$$

ただし，A は絶対定数である．

Remark 12.1. 特に，$\{X_m\}$ が i.i.d. r.v.'s で，$E[X_1] = 0, E[X_1^2] = \sigma^2 > 0, E[|X_1|^3] < \infty$ であるならば，$X_{m,n} = X_m/\sigma\sqrt{n}$ とおくと，$\sum_{m=1}^{n} E[|X_{m,n}|^3] = n^{-1/2} E[|X_1|^3]/\sigma^3$ となる．よって，$\check{S}_n = \sum_{m=1}^{n} X_m/(\sigma\sqrt{n})$ とおくと，

$$\sup_{x\in\mathbb{R}} \left| P\left(\check{S}_n \le x\right) - \Phi(x) \right| \le \frac{AE[|X_1|^3]}{\sigma^3\sqrt{n}} = O(n^{-1/2})$$

となる．さらに，$O(n^{-1/2})$ のオーダーは一様には改善できない．すなわち，上の仮定をみたす i.i.d. r.v.'s であり，適当な定数 $c > 0$ に対して，

$$\sup_{x\in\mathbb{R}} |P(\check{S}_n \le x) - \Phi(x)| \ge cn^{-1/2}$$

となるものが存在する (cf. Exercise 12.3).

Berry-Esseen の定理の証明は Chung (2001, Section 7.4)，または Stroock (2011, Section 2.2) を参照せよ．Stroock はいわゆる "Stein の方法" に基づく証明を与えている．なお，著者の Errata によると，Durrett (2010) に載っている証明にはギャップがあるようである．

Example 12.1 (分散が発散する場合)**.** CLT は分散が発散する場合でも成り立つことがある．$\{X_m\}$ を i.i.d. r.v.'s とし, その共通分布を $P(X_1 > x) = P(-X_1 > x) = x^{-2}/2$ $(x \ge 1)$ とする．このとき，

$$E[X_1^2] = \int_0^\infty 2xP(|X_1| > x)dx = \infty$$

であるから，直接 CLT を適用することはできない．ただし，適当に正規化すれば $S_n = \sum_{m=1}^n X_m$ は $N(0,1)$ に分布収束する．まず，

$$Y_{n,m} = X_m 1_{\{|X_m| \le c_n\}},\ c_n = n^{1/2}\log\log n,\ S'_n = \sum_{m=1}^n Y_{n,m}$$

とおくと，$P(S_n \neq S'_n) \le \sum_{m=1}^n P(Y_{n,m} \neq X_m) \le nP(|X_1| > c_n) \to 0$.

ここで，簡単な計算から，$\mathrm{Var}(S'_n) \sim n\log n$ となるから，$X_{n,m} = Y_{n,m}/\sqrt{\mathrm{Var}(S'_n)}$ とおくと，$X_{n,m}$ は Lindeberg 条件をみたす．よって，

$$\frac{S_n}{\sqrt{n\log n}} = \underbrace{\frac{S_n - S'_n}{\sqrt{n\log n}}}_{\xrightarrow{P}0} + \frac{S'_n}{\sqrt{\mathrm{Var}(S'_n)}}\underbrace{\frac{\sqrt{\mathrm{Var}(S'_n)}}{\sqrt{n\log n}}}_{\to 1} \xrightarrow{d} N(0,1).$$

この例の一般化に関しては，安定分布の節を参照せよ．なお，この例だと，分散が有限な場合と比べて，正規化定数を $\sqrt{n\log n}$ と少し大きめにとらなければならない．じつは分散が発散する場合，正規化定数を $\sqrt{n}$ のオーダーにとることは出来ない．すなわち，i.i.d. r.v.'s $\{X_m\}$ に対して，$S_n/\sqrt{n} \xrightarrow{d} N(0,1)$ ならば，$E[X_1] = 0, E[X_1^2] = 1$ となる．証明は Theorem 14.2 を参照せよ．

最後に CLT の応用として，Kolmogorov の 3 級数定理を示す．

Theorem 12.5 (Kolmogorov の 3 級数定理)**.** $\{X_m\}$ を独立な r.v.'s とし，$A > 0$ を任意の定数とする．このとき，$Y_m = X_m 1_{\{|X_m| \le A\}}$ とおくと，a.s. に $\sum_{n=1}^\infty X_n$ が収束するための必要十分条件は

$$\text{(i)}\ \sum_n P(|X_n| > A) < \infty,\ \text{(ii)}\ \sum_n E[Y_n] \text{ が収束する},\ \text{(iii)}\ \sum_n \mathrm{Var}(Y_n) < \infty$$

の 3 条件がすべて成り立つことである．

Proof. 十分性．(iii) と Theorem 10.2 より，$\sum_n (Y_n - E[Y_n])$ は a.s. に収束する．(ii) より $\sum_n Y_n$ も a.s. に収束する．最後に (i) と Borel-Cantelli の補題より，$P(X_n \neq Y_n \text{ i.o.}) = 0$ となるから，$\sum_n X_n$ の収束が従う．

必要性．$\sum_n X_n$ が収束するなら，$X_n \to 0$ となるから，$P(|X_n| > A \text{ i.o.}) = 0$．よって，Borel-Cantelli の補題より，$\sum_n P(|X_n| > A) < \infty$. 次に，(iii) の和が発散すると仮定して矛盾を導く．$c_n = \sum_{m=1}^n \mathrm{Var}(Y_m)$, $X_{n,m} = (Y_m - E[Y_m])/c_n^{1/2}$ とおくと，$|X_{n,m}| \le 2A/c_n^{1/2}$ と $c_n \to \infty$ より，$X_{n,m}$ は Lindeberg 条件をみたすので，$S_n = \sum_{m=1}^n X_{n,m} \xrightarrow{d} N(0,1)$. また (i) と Borel-Cantelli の補題より，$P(X_n \neq Y_n \text{ i.o.}) = 0$ となるから，$\sum_{n=1}^\infty X_n$ が a.s. に収束することと $\sum_{n=1}^\infty Y_n$ が a.s. に収束することは同値である．いま $\sum_{n=1}^\infty Y_n$ は a.s. に収束するので，$T_n = \sum_{m=1}^n Y_m/c_n^{1/2} \xrightarrow{P} 0$. 従って Slutsky の補題より，$S_n - T_n \xrightarrow{d} N(0,1)$ だが，$S_n - T_n$ は非確率的であるから，矛盾．最後に (ii) を示す．(iii) と Theorem 10.2 より，$\sum_{n=1}^\infty (Y_n - E[Y_n])$ は a.s. に収束する．いま $\sum_{n=1}^\infty Y_n$ も a.s. に収束するので，$\sum_{m=1}^n E[Y_m] = \sum_{m=1}^n Y_n + \sum_{m=1}^n (E[Y_n] - Y_n)$ は $n \to \infty$ で a.s. に収束する． □

演習問題

Exercise 12.1. $\{X_m\}$ を i.i.d. r.v.'s とし，$E[X_m] = 0, E[X_m^2] = 1$ とする．$S_n = \sum_{m=1}^n X_m$ とおき，N_n を正整数に値をとる r.v.'s であって，

$$\frac{N_n}{n} \to c > 0 \quad a.s.$$

とする．ただし，c は定数である．このとき，$S_{N_n}/\sqrt{N_n} \xrightarrow{d} N(0,1)$ を示せ．

Exercise 12.2. $x_1, x_2, \cdots \in \mathbb{R}$ を $n^{-1}\sum_{m=1}^n x_m \to 0$, $n^{-1}\sum_{m=1}^n x_m^2 \to 1$ をみたす数列とする．各 $n \in \mathbb{N}$ に対して，$F_n(x) = n^{-1}\sum_{m=1}^n 1_{[x_m,\infty)}(x)$ とし，$X_{n,1}, \ldots, X_{n,n} \sim F_n$ i.i.d. とする．いま，$\bar{X}_n = n^{-1}\sum_{m=1}^n X_{n,m}$, $\theta_n = n^{-1}\sum_{m=1}^n x_m$ とおいたとき，$\sqrt{n}(\bar{X}_n - \theta_n) \xrightarrow{d} N(0,1)$ を示せ．

Exercise 12.3. $\{X_m\}$ を i.i.d. r.v.'s とし，$P(X_m = \pm 1) = 1/2$ とする．$S_n = \sum_{m=1}^n X_m$ とおいたとき，

$$\liminf_n \sup_{x\in\mathbb{R}} \sqrt{n}|P(n^{-1/2}S_n \le x) - \Phi(x)| > 0$$

を示せ．

ヒント：各整数 j に対して，

$$j_n = \begin{cases} j & \text{if } n+j \text{ が偶数} \\ j+1 & \text{if } n+j \text{ が奇数} \end{cases}$$

とすると，$\sqrt{n}P(S_n = j_n) \to 1/\sqrt{2\pi}$ となることを示せ．

Exercise 12.4. $X_n \xrightarrow{d} X \sim N(0,1)$ のとき，すべての $A \in \mathcal{B}$ に対して，$P(X_n \in A) \to P(X \in A)$ となるか．

ヒント:∂A の Lebesgue 測度が 0 なら，Portmanteau 定理より，$P(X_n \in A) \to P(X \in A)$ となる．∂A の Lebesgue 測度が正のときはどうか．cf. Exercise 2.9.

Exercise 12.5. $\{\varepsilon_n\}$ を i.i.d. r.v.'s とし，$P(\varepsilon_n = \pm 1) = 1/2$ とする．$\alpha > 0$ に対して，$\sum_{n=1}^\infty n^{-\alpha}\varepsilon_n$ が a.s. に収束する α の範囲を求めよ．

以下，Theorem 12.1 の記号を用いる．

Exercise 12.6. $g : \mathbb{R} \to \mathbb{R}$ を

$$\sup_{x\in\mathbb{R}} \frac{|g(x)|}{1+|x|^2} < \infty$$

をみたす連続関数とすると，$\lim_n E[g(S_n)] = E[g(Z)]$ を示せ．

Exercise 12.7. 次の条件 (Lyapunov 条件 と呼ばれる) は Lindeberg 条件の十分条件となっていることを示せ.

$$\exists \delta > 0 \text{ s.t. } \lim_n \sum_{m=1}^{n} E[|X_{n,m}|^{2+\delta}] = 0.$$

Exercise 12.8. Lindeberg 条件はみたすが Lyapunov 条件はみたさない r.v.'s $\{X_{n,m}\}$ を構成せよ.

Exercise 12.9. Theorem 12.3 の議論を修正して，次のバウンドを導出せよ : $\forall h \in C^3(\mathbb{R})$ s.t. $\|h''\|_u \vee \|h'''\|_u < \infty$,

$$|E[h(S_n)] - E[h(Z)]| \leq \frac{2\|h'''\|_u}{3} \sum_{m=1}^{n} E[|X_{m,n}|^3]. \qquad (*)$$

Exercise 12.10. (a) $\mathbb{R}$ から $[0,1]$ への C^3 級関数 g であって，

$$g(x) = \begin{cases} 1 & \text{if } x \leq 0 \\ 0 & \text{if } x \geq 1 \end{cases}$$

をみたすものを構成せよ.

(b) 与えられた $x \in \mathbb{R}, \varepsilon > 0$ に対して，$h_{x,\varepsilon}(y) = g((y-x)/\varepsilon),\ y \in \mathbb{R}$ とおく．この $h_{x,\varepsilon}$ に $(*)$ を適用して，次のバウンドを導出せよ：

$$\sup_{x \in \mathbb{R}} |P(S_n \leq x) - \Phi(x)| \leq A \left(\sum_{m=1}^{n} E[|X_{m,n}|^3] \right)^{1/4}.$$

ただし，$A > 0$ は絶対定数である (A の値を具体的に求める必要はない).

13 特性関数

r.v. X の 特性関数 (characteristic function, ch.f.) $\varphi(t)$ を

$$\varphi(t) := E[e^{itX}] = E[\cos(tX)] + iE[\sin(tX)],\ t \in \mathbb{R},\ i := \sqrt{-1}$$

と定義する．φ を分布 $\mathcal{L}(X)$ の ch.f. とも言う．ch.f. は次の性質をみたす：(i) $\varphi(0) = 1$. (ii) $\varphi(-t) = \overline{\varphi(t)}$. (iii) $|\varphi(t)| \leq 1$. (iv) $\varphi(t)$ は一様連続．(v) X_1, X_2 が独立なら，$E[e^{it(X_1+X_2)}] = E[e^{itX_1}]E[e^{itX_2}]$.

Example 13.1. (ch.f. の例).

- $N(0,1)$. $X \sim N(0,1)$ のとき，$\phi(x) = (2\pi)^{-1/2}e^{-x^2/2}$ を $N(0,1)$ の密度関数とすると，
$$\varphi(t) = \int e^{itx}\phi(x)dx = \int \cos(tx)\phi(x)dx.$$
ここで，両辺を微分すると，
$$\varphi'(t) = \int -x\sin(tx)\phi(x)dx = -t\int \cos(tx)\phi(x)dx = -t\varphi(t)$$
となるので，$\{\varphi(t)e^{t^2/2}\}' = 0$. よって，$\varphi(t)e^{t^2/2} = \varphi(0) = 1$ より，$\varphi(t) = e^{-t^2/2}$ を得る.

- $U(a,b)$. $\varphi(t) = (e^{itb} - e^{itb})/it(b-a)$.

- 三角分布．密度関数
$$f(x) = (1 - |x|)^+$$
をもつ分布を三角分布と呼ぶ．X, Y を独立で，$X, Y \sim U(-1/2, 1/2)$ とすると，$X+Y$ の d.f. の密度関数は
$$\int 1_{\{|z-y|\leq 1/2\}}1_{\{|y|\leq 1/2\}}dy = (1 - |z|)^+.$$
よって，三角分布の ch.f. は $\varphi(t) = \{(e^{it/2} - e^{-it/2})/(it)\}^2 = \cdots = 2(1-\cos t)/t^2$.

- 混合分布．$\{\varphi_j\}$ を ch.f.'s の可算集合とし，$\{b_j\}$ を $b_j > 0, \sum_j b_j = 1$ なる集合とすると，$\sum_j b_j\varphi_j$ は ch.f. である．なお，$|\varphi_j| \leq 1$ であるから，$\sum_j b_j\varphi_j$ は一様に絶対収束する．

以下，ch.f. に関する 2 つの重要な結果，反転公式と連続性定理を証明する．

Theorem 13.1 (反転公式)**.** μ を $(\mathbb{R}, \mathcal{B})$ 上の p.m. とし, $\varphi(t)$ をその ch.f. とする：$\varphi(t) = \int e^{itx}d\mu(x)$．このとき，任意の $a < b$ に対して，

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

左辺は広義積分の意味である．$\mu = \delta_0, a = -1, b = 1$ とすると，$\varphi(t) = 1$ であるから，被積分関数は $(2\sin t)/t$ である．このとき，広義積分は収束するが，Lebesgue 積分の意味で，$\int_{\mathbb{R}}\{(\sin t)/t\}dt$ は定義できない．

Proof. $|(e^{-ita} - e^{-itb})/(it)| = |\int_a^b e^{-ity}dy| \le b - a$ であるから，Fubini の定理より，

$$\begin{aligned} I_T &:= \int_{-T}^{T} \frac{e^{-ita} - e^{-itb}}{it}\varphi(t)dt = \int_{-T}^{T}\int \frac{e^{-ita} - e^{-itb}}{it}e^{itx}d\mu(x)dt \\ &= \int\int_{-T}^{T} \frac{e^{-ita} - e^{-itb}}{it}e^{itx}dtd\mu(x) \\ &= \int\left\{\int_{-T}^{T}\frac{\sin(t(x-a))}{t}dt - \int_{-T}^{T}\frac{\sin(t(x-b))}{t}dt\right\}d\mu(x) \\ &= \int\{R(x-a,T) - R(x-b,T)\}d\mu(x). \end{aligned}$$

ここで，

$$R(\theta, T) = \int_{-T}^{T}\frac{\sin(\theta t)}{t}dt$$

とおいた．$S(T) = \int_0^T\{(\sin t)/t\}dt$ とおくと，$\theta > 0$ に対して，

$$R(\theta, T) = 2\int_0^{T\theta}\frac{\sin t}{t}dt = 2S(T\theta).$$

一方，$\theta < 0$ に対して，$R(\theta, T) = -R(|\theta|, T) = -2S(T|\theta|)$. よって，

$$R(\theta, T) = 2\,\mathrm{sgn}(\theta)S(T|\theta|).$$

$T \to \infty$ とすると，$S(T) \to \pi/2$ となるから，$R(\theta, T) \to \pi\,\mathrm{sgn}(\theta)$. 従って，

$$R(x-a,T) - R(x-b,T) \to \begin{cases} 2\pi & \text{if } a < x < b \\ \pi & \text{if } x = a \text{ or } x = b \\ 0 & \text{if } x < a \text{ or } x > b. \end{cases}$$

$|R(\theta, T)| \le 2\sup_y |S(y)| < \infty$ であるから，DCT より，

$$(2\pi)^{-1}I_T \to \mu((a,b)) + \frac{1}{2}\mu(\{a,b\})$$

を得る． □

Corollary 13.1 (Durrett 3.3.2 (i))**.** 任意の $a \in \mathbb{R}$ に対して，

$$\mu(\{a\}) = \lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T}e^{-ita}\varphi(t)dt$$

となる．

反転公式より，分布と ch.f. は 1 対 1 に対応していることがわかる．

Corollary 13.2. μ, ν を $(\mathbb{R}, \mathcal{B})$ 上の p.m.'s とし，それらの ch.f.'s をそれぞれ φ, ψ とおく．このとき，$\mu = \nu \Leftrightarrow \varphi(t) = \psi(t)$ $(\forall t \in \mathbb{R})$.

Proof. $\Rightarrow$ はよい．$\Leftarrow$ を示す．$D = \{x \in \mathbb{R} : \mu(\{x\}) \vee \nu(\{x\}) > 0\}$ とおくと，反転公式より，$\forall a, b \in \mathbb{R} \setminus D, \mu((a,b]) = \mu((a,b)) = \nu((a,b)) = \nu((a,b])$. D は高々可算だから，$a_m \in \mathbb{R} \setminus D, a_m \downarrow -\infty$ なる数列が選べて，$\mu((-\infty, b]) = \lim_m \mu((-a_m, b]) = \lim_m \nu((a_m, b]) = \nu((-\infty, b])$. これは，$\mu$ と ν の d.f.'s が $\mathbb{R} \setminus D$ 上で等しいことを意味する．あとは d.f. の右連続性より，$\mu = \nu$ を得る． □

一般に，ch.f. φ は Lebesgue 測度に関して可積分とは限らないが，可積分のときは次の定理が成り立つ．

Theorem 13.2. $\int |\varphi(t)| dt < \infty$ なら，μ は有界連続な密度

$$f(y) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-ity} \varphi(t) dt$$

をもつ．

Proof. $|(e^{-ita} - e^{-itb})/(it)| \leq |a - b|$ より，$\int |\varphi(t)| dt < \infty$ のとき，反転公式に現れる積分は絶対収束する．よって，

$$\mu((a,b)) + \frac{1}{2}\mu(\{a,b\}) = \frac{1}{2\pi} \int \frac{e^{-ita} - e^{-itb}}{it} \varphi(t) dt \leq \frac{b-a}{2\pi} \int |\varphi(t)| dt.$$

$b \downarrow a$ として，$\mu(\{a\}) = 0$. さらに，μ の d.f. を F とすると，Fubini の定理より，

$$\begin{aligned} F(b) - F(a) = \mu((a,b]) &= \frac{1}{2\pi} \int \frac{e^{-ita} - e^{-itb}}{it} \varphi(t) dt = \frac{1}{2\pi} \int \left(\int_a^b e^{-ity} dy \right) \varphi(t) dt \\ &= \int_a^b \underbrace{\left(\frac{1}{2\pi} \int e^{-ity} \varphi(t) dt \right)}_{=: f(y)} dy = \int_a^b f(y) dy. \end{aligned}$$

DCT より，f は有界連続であり，従って，F は C^1 級で，$F' = f$. □

Example 13.2. Pólya 分布. 密度関数

$$f(x) = \frac{1 - \cos x}{\pi x^2}$$

をもつ分布を Pólya 分布と呼ぶ．三角分布の ch.f. の計算より，

$$\int (1 - |x|)^+ e^{itx} dx = \frac{2(1 - \cos t)}{t^2}$$

だったから，反転公式より，

$$\int \frac{1}{2\pi}\frac{2(1-\cos t)}{t^2}e^{-ity}dt = (1-|y|)^+.$$

変数を入れ替えて，

$$\underbrace{\int \frac{1-\cos x}{\pi x^2}e^{itx}dx}_{=\varphi(t)} = (1-|t|)^+$$

を得る．

Theorem 13.3 (連続性定理)．μ_n を $(\mathbb{R},\mathcal{B})$ 上の p.m.'s とし，μ_n の ch.f. を φ_n とする．(i) $\mu_n \xrightarrow{w} \mu$ なら，すべての $t\in\mathbb{R}$ に対して，$\varphi_n(t)\to\varphi(t) := \int e^{itx}d\mu(x)$. (ii) 逆に，すべての $t\in\mathbb{R}$ に対して，$\lim_n \varphi_n(t) =: \varphi(t)$ が存在して，$\varphi(t)$ が $t=0$ で連続なら，φ は ch.f. であって，μ を φ を ch.f. にもつ p.m. とすると，$\mu_n \xrightarrow{w} \mu$.

(ii) において，$\varphi(t)$ の原点での連続性は本質的な仮定である．例えば，$X_n \sim N(0,n)$ とすると，$\varphi_n(t) = E[e^{itX_n}] = e^{-nt^2/2} \to 0\ (\forall t\neq 0), \varphi_n(0)=1$ だが，X_n は分布収束しない．実際，$Z\sim N(0,1)$ に対して，$X_n \overset{d}{=} \sqrt{n}Z$ だから，

$$P(X_n\le x) = P(Z\le x/\sqrt{n}) \to P(Z\le 0) = 1/2.$$

仮に，X_n が分布収束するなら，その極限分布の d.f. F は高々可算個の x を除いて $F(x) = 1/2$ となるが，これは d.f. の定義に矛盾する．

Proof. (i) は明らか．(ii) を示す．まず，$\{\mu_n\}$ が一様にタイトであることを示す．

$$\int_{-u}^{u}(1-e^{itx})dt = 2u - \int_{-u}^{u}\cos(tx)dt = 2u - \frac{2\sin(ux)}{x}.$$

両辺を u で割って $d\mu_n$ に関して積分すると，

$$u^{-1}\int_{-u}^{u}(1-\varphi_n(t))dt = 2\int\left(1-\frac{\sin(ux)}{ux}\right)d\mu_n(x).$$

$|\sin x|\le|x|$ より，$1-\sin(ux)/(ux)\ge 0$ であるから，

$$\text{右辺} \ge 2\int_{|x|\ge 2/u}\left(1-\frac{1}{|ux|}\right)d\mu_n(x) \ge \mu_n(\{x:|x|\ge 2/u\}).$$

一方，$\varphi(t)\to 1\ (t\to 0)$ であるから，

$$u^{-1}\int_{-u}^{u}(1-\varphi(t))dt \to 0,\ u\downarrow 0.$$

よって，$\forall \varepsilon > 0, \exists u > 0$ s.t. 左辺 $\leq \varepsilon$. DCT より，十分大きな n に対して，

$$u^{-1}\int_{-u}^{u}(1-\varphi_n(t))dt \leq u^{-1}\int_{-u}^{u}(1-\varphi(t))dt + \varepsilon \leq 2\varepsilon.$$

従って，$\mu_n(\{x : |x| \geq 2/u\}) \leq 2\varepsilon$. ゆえに $\{\mu_n\}$ は一様にタイトである．

$\mu_{n(m)} \xrightarrow{w} \mu$ とすると，(i) より，$\varphi(t) = \int e^{itx}d\mu(x)$ であり，従って，φ は ch.f. さらに分布と ch.f. は 1 対 1 に対応するので，μ は部分列のとり方によらず一意に決まる．従って，Prohorov の定理の系 (Corollary 11.2) より，$\mu_n \xrightarrow{w} \mu$ を得る． □

Corollary 13.3 (Durrett 3.3.16)**.** X_n, X の ch,f.'s をそれぞれ φ_n, φ とすると，$X_n \xrightarrow{d} X$ なら，$\forall T > 0$, $\sup_{|t|\leq T}|\varphi_n(t) - \varphi(t)| \to 0$.

連続性定理の応用として，Poisson 少数の法則 (Poisson's law of small numbers) を示す．

Theorem 13.4 (Poisson 少数の法則)**.** 各 $n \in \mathbb{N}$ に対して，$X_{n,m}, 1 \leq m \leq n$ は $\{0,1\}$ に値をとる独立な r.v.'s であって，$P(X_{n,m} = 1) = p_{n,m}$, $P(X_{n,m} = 0) = 1 - p_{n,m}$ とする．さらに，(i) $\sum_{m=1}^{n} p_{n,m} \to \lambda \in (0,\infty)$, (ii) $\max_{1\leq m\leq n} p_{n,m} \to 0$ を仮定する．このとき，$S_n = \sum_{m=1}^{n} X_{n,m} \xrightarrow{d} Po(\lambda)$.

補題を 2 つ用意する．

Lemma 13.1. $z_1, \ldots, z_n, w_1, \ldots, w_n \in \mathbb{C}$, $|z_m|, |w_m| \leq \theta$ $(1 \leq \forall m \leq n)$ に対して，

$$\left|\prod_{m=1}^{n} z_m - \prod_{m=1}^{n} w_m\right| \leq \theta^{n-1}\sum_{m=1}^{n}|z_m - w_m|.$$

Proof. $n = 1$ のときは明らか．$n \geq 2$ のとき，

$$\begin{aligned}\left|\prod_{m=1}^{n} z_m - \prod_{m=1}^{n} w_m\right| &\leq \left|z_1\prod_{m=2}^{n} z_m - z_1\prod_{m=2}^{n} w_m\right| + \left|z_1\prod_{m=2}^{n} w_m - w_1\prod_{m=2}^{n} w_m\right| \\ &\leq \theta\left|\prod_{m=2}^{n} z_m - \prod_{m=2}^{n} w_m\right| + \theta^{n-1}|z_1 - w_1|.\end{aligned}$$

後は帰納法. □

Lemma 13.2. $|e^z - 1 - z| \leq |z|^2$ $(\forall |z| \leq 1, z \in \mathbb{C})$.

Proof.

$$e^z = 1 + z + \sum_{k=2}^{\infty}\frac{z^k}{k!} = 1 + z + z^2\sum_{k=0}^{\infty}\frac{z^k}{(k+2)!}.$$

$(k+2)! \geq 2^{k+1}$ より，$|z| \leq 1$ のとき，

$$|e^z - 1 - z| \leq |z|^2\sum_{k=0}^{\infty}2^{-k-1} = |z|^2.$$

□

Proof of Theorem 13.4. $\max_{1\le m\le n} p_{n,m} \le 1/2$ と仮定してよい．このとき，

$$
\begin{aligned}
&|1 + p_{n,m}(1 - e^{it})| \le 1,\\
&|\exp(p_{n,m}(e^{it} - 1))| = \exp\{p_{n,m}(\cos t - 1)\} \le 1,\\
&p_{n,m}|e^{it} - 1| \le 1
\end{aligned}
$$

であるから，

$$
\begin{aligned}
&\left|\prod_{m=1}^{n}\{1 + p_{n,m}(e^{it} - 1)\} - \exp\left(\sum_{m=1}^{n} p_{n,m}(e^{it} - 1)\right)\right|\\
&\le \sum_{m=1}^{n}\left|\{1 + p_{n,m}(e^{it} - 1)\} - \exp(p_{n,m}(e^{it} - 1))\right|\\
&\le \sum_{m=1}^{n} p_{n,m}^2|e^{it} - 1|^2 \le 4(\max_{1\le m\le n} p_{n,m})\sum_{m=1}^{n} p_{n,m} = o(1)
\end{aligned}
$$

となる．以上より，

$$
E[e^{itS_n}] = \prod_{m=1}^{n}\{1 + p_{n,m}(e^{it} - 1)\} = \exp\left(\sum_{m=1}^{n} p_{n,m}(e^{it} - 1)\right) + o(1) \to \exp(\lambda(e^{it} - 1)).
$$

右辺は $Po(\lambda)$ の ch.f. なので，$S_n \xrightarrow{d} Po(\lambda)$ を得る． □

Feller の定理の証明. 以下，Feller の定理 (Theorem 12.2) を証明する．まず，技術的な準備をする．対数関数の領域 $\mathbb{C} \setminus (-\infty, 0]$ 上の主分枝を $\mathrm{Log}\, z := \log r + i\theta, z = re^{i\theta}$ $(r > 0, |\theta| < \pi)$ とおく．$\mathrm{Log}\, z$ は領域 $\mathbb{C} \setminus (-\infty, 0]$ 上で正則であり，$z = 1$ の近傍でべき級数展開

$$
\mathrm{Log}\, z = \sum_{n=1}^{\infty}(-1)^{n+1}\frac{(z-1)^n}{n},\ |z - 1| < 1
$$

をもつ．

Lemma 13.3. (i) $f : [-T, T] \to \mathbb{C}$ を連続で，$f(0) = 1$, $f(t) \ne 0$ $(\forall t \in [-T, T])$ をみたすとする．このとき，一意な連続関数 $\lambda : [-T, T] \to \mathbb{C}$ が存在して，$\lambda(0) = 0, f(t) = e^{\lambda(t)}$ $(\forall t \in [-T, T])$ をみたす．$\log f(t) := \lambda(t)$ と定義する．(ii) f_n, f は (i) の条件をみたし，$\sup_{|t|\le T}|f_n(t) - f(t)| \to 0$ とする．このとき，$\sup_{|t|\le T}|\log f_n(t) - \log f(t)| \to 0$.

なお，Chung (2001) は Lemma 13.3 の意味での $\log f(t)$ を “distinguished logarithm” と呼んでいる．

Proof. (i). f は連続なので，$\varepsilon := \inf_{t\in[-T,T]}|f(t)| > 0$. f は $[-T, T]$ 上で一様連続なので，$\exists\delta > 0$ s.t. $|t - t'| \le \delta, t, t' \in [-T, T] \Rightarrow |f(t) - f(t')| \le \varepsilon/2$. $[-T, T]$ を長さが δ 以下の区間に分割する：

$$
-T = t_{-L} < \cdots < t_{-1} < t_0 = 0 < t_1 < \cdots < t_L = T.
$$

このとき，$|f(t)/f(t_\ell)-1| \le 1/2$ $(\forall t \in [t_\ell, t_{\ell+1}])$ となる．そこで，帰納的に

$$\lambda(t_0) = \lambda(0) := 0,\ \lambda(t) := \lambda(t_\ell) + \mathrm{Log}\,\frac{f(t)}{f(t_\ell)},\ t \in [t_\ell, t_{\ell+1}]$$

と定義する．λ は well-defined な連続関数で，$e^{\lambda(t)} = f(t)$ をみたす．一意性は，$\tilde{\lambda}$ をもう 1 つのそのような関数とすると，$e^{\lambda(t)-\tilde{\lambda}(t)} = 1$ となるから，$\exists m(t) \in \mathbb{Z}$ s.t. $\lambda(t) - \tilde{\lambda}(t) = 2\pi i m(t)$. 左辺は連続なので，$m(t)$ は定数でなくてはならないが，$\lambda(0) = \tilde{\lambda}(0) = 0$ より，$m(t) \equiv 0$ となる．

(ii). $\log(f_n/f) = \log f_n - \log f$ となることから明らか． □

Proof of Theorem 12.2. $T > 0$ とする．$\varphi_{n,m}(t) = E[e^{itX_{n,m}}]$ とおく．まず，$|e^{itx} - 1 - (itx)| \le t^2x^2$ より，$|\varphi_{n,m}(t) - 1| = |E[e^{itX_{n,m}} - 1 - itX_{n,m}]| \le t^2 E[X_{n,m}^2] \le t^2 r_n$ となる．仮定より，$r_n \to 0$ であるから，十分大きな n に対して，

$$\max_{1\le m\le n} |\varphi_{n,m}(t) - 1| \le 1/2,\ \forall |t| \le T \qquad (*)$$

となる．一方，$S_n \xrightarrow{d} N(0,1)$ より，

$$\lim_n \prod_{m=1}^n \varphi_{n,m}(t) = -e^{-t^2/2} \quad \text{unif. in } |t| \le T$$

なので，Lemma 13.3 より，

$$\lim_n \sum_{m=1}^n \log \varphi_{n,m}(t) = -\frac{t^2}{2} \quad \text{unif. in } |t| \le T$$

を得る．$(*)$ より，$\log \varphi_{n,m}(t) = \mathrm{Log}(\varphi_{n,m}(t))$ $(|t| \le T)$ であるから，

$$\sum_{m=1}^n \log \varphi_{n,m}(t) = \sum_{m=1}^n \{\varphi_{n,m}(t) - 1\} + O(1) \sum_{m=1}^n |\varphi_{n,m}(t) - 1|^2,$$

$$\sum_{m=1}^n |\varphi_{n,m}(t) - 1|^2 \le \max_{1\le m\le n} |\varphi_{n,m}(t) - 1| \sum_{m=1}^n |\varphi_{n,m}(t) - 1|. \qquad (**)$$

ここで，$\sum_{m=1}^n |\varphi_{n,m}(t) - 1| \le t^2 \sum_{m=1}^n E[X_{n,m}^2] = t^2$ より，

$$(**) \text{ の右辺} \le \max_{1\le m\le n} |\varphi_{n,m}(t) - 1| \to 0.$$

よって，$\lim_n \sum_{m=1}^n \{\varphi_{n,m}(t) - 1\} = -t^2/2$ を得る．両辺の実部を比較して，

$$\lim_n \sum_{m=1}^n E[\{1 - \cos(tX_{n,m})\}] = \frac{t^2}{2}$$

を得る．期待値を $\{|X_{n,m}| \le \varepsilon\}$ と $\{|X_{n,m}| > \varepsilon\}$ に分割して評価すると，

$$\begin{aligned}
&\limsup_n \left\{ \frac{t^2}{2} - \sum_{m=1}^n E[\{1-\cos(tX_{n,m})\}1_{\{|X_{n,m}|\le\varepsilon\}}] \right\} \\
&= \limsup_n \sum_{m=1}^n E[\{1-\cos(tX_{n,m})\}1_{\{|X_{n,m}|>\varepsilon\}}] \\
&\le \limsup_n 2\sum_{m=1}^n P(|X_{n,m}| > \varepsilon) \\
&\le \limsup_n 2\sum_{m=1}^n \frac{E[X_{n,m}^2]}{\varepsilon^2} = \frac{2}{\varepsilon^2}
\end{aligned}$$

となる．ここで，$0 \le 1-\cos\theta \le \theta^2/2$ $(\forall\theta \in \mathbb{R})$ より，

$$\frac{4}{t^2\varepsilon^2} \ge \limsup_n \left\{ 1 - \sum_{m=1}^n E[X_{n,m}^2 1_{\{|X_{n,m}|\le\varepsilon\}}] \right\} \ge 0,\ 0 < |t| \le T,$$

であり，括弧の中は $\sum_{m=1}^n E[X_{n,m}^2 1_{\{|X_{n,m}|>\varepsilon\}}]$ に等しい．T は任意だったから，$t \to \infty$ として Lindeberg 条件を得る． □

Pólya の基準. 与えられた関数が ch.f. かどうかを判定する基準を考察する．

Theorem 13.5 (Pólya の基準)**.** $\varphi : \mathbb{R} \to \mathbb{R}_+$ を，$\varphi(0) = 1, \varphi(t) = \varphi(-t)$ $(\forall t \in \mathbb{R})$ をみたし, さらに $\mathbb{R}_+$ 上で非増加，連続かつ凸とする．このとき，φ は ch.f. である．

Proof. 予備的な考察から始める．Pólya 分布の ch.f. は $(1-|t|)^+$ であったから，$a > 0$ に対して，$\varphi_a(t) = (1-|t|/a)^+$ は ch,f. である．よって，$0 < a_1 < \cdots < a_n$ と $b_j \ge 0, \sum_{j=1}^n b_j = 1$ なる $\{b_j\}_{j=1}^n$ に対して，$\psi(t) = \sum_{j=1}^n b_j \varphi_{a_j}(t)$ は ch.f. である．このとき，$(a_0, c_0) = (0, 1),\ c_k = \sum_{j=k}^n (1 - a_k/a_j) b_j$ $(k = 1, \ldots, n)$ とおくと，

$$\psi(t) = \begin{cases} (a_j, c_j) \text{ たちをつないだ折れ線} & \text{if } t \in [0, a_n] \\ 0 & \text{if } t > a_n \\ \psi(-t) & \text{if } t < 0 \end{cases} \qquad (*3)$$

となる．逆に，$0 = a_0 < a_1 < \cdots < a_n$ と $1 = c_0 \ge c_1 \ge \cdots \ge c_n = 0$ が与えられたとき，

$$\frac{c_j - c_{j-1}}{a_j - a_{j-1}} \ge \frac{c_{j+1} - c_j}{a_{j+1} - a_j},\ j = 1, \ldots, n-1 \qquad (*4)$$

がみたされるなら，$(*3)$ から構成した ψ は ch.f. である．条件 $(*4)$ は $\{(a_j, c_j)\}_{j=1}^n$ から構成した $\{b_j\}_{j=1}^n$ が非負制約 $b_j \ge 0$ をみたすことを保証している．

定理の証明に移る．$T = \sup\{t \in \mathbb{R}_+ : \varphi(t) > 0\} \leq \infty$ とおき，各 n に対して，$T < \infty$ のときは，$a_{n,j} = Tj/2^n, c_{n,j} = \varphi(a_{n,j})$ $(j = 0, 1, \ldots, 2^n =: k_n)$ と選び，$T = \infty$ のときは，$a_{n,j} = j/2^n, c_{n,j} = \varphi(a_{n,j})$ $(j = 0, 1, \ldots, n2^n - 1 =: k_n - 1)$，$a_{n,k_n} = n, c_{n,k_n} = 0$ と選ぶ．φ に関する仮定より，分点 $\{(a_{n,j}, c_{n,j})\}_{j=1}^{k_n}$ は $1 = c_{n,0} \geq c_{n,1} \geq \cdots \geq c_{n,k_n} = 0$ と条件 $(*4)$ をみたすから，この分点に対して $(*3)$ の操作を適用して得られた関数 ψ_n は ch.f. である．あとは各 t に対して，$\psi_n(t) \to \varphi(t)$ $(n \to \infty)$ となることと，φ は仮定より連続であることから，連続性定理より，φ が ch.f. であることが示された． □

演習問題

Exercise 13.1. Cantor 分布の特性関数を求めよ．

Exercise 13.2. $F_n \xrightarrow{w} F, G_n \xrightarrow{w} G$ なら，$F_n * G_n \xrightarrow{w} F * G$ となることを示せ．

Exercise 13.3. X, X_n を r.v.'s とし，φ, φ_n をその ch.f.'s とする．$X_n \xrightarrow{d} X$ であるが，$\liminf_n \sup_{t \in \mathbb{R}} |\varphi_n(t) - \varphi(t)| > 0$ となるような例を構成せよ．

Exercise 13.4. 与えられた $T > 0$ に対して，$[-T, T]$ 上では一致しているが，$\mathbb{R} \setminus [-T, T]$ 上では異なる 2 つの ch.f.'s を構成せよ．

関数 $f : \mathbb{R} \to \mathbb{C}$ が 正定値 (positive definite) であるとは，あらゆる $t_1, \ldots, t_n \in \mathbb{R}$ と $z_1, \ldots, z_n \in \mathbb{C}$ に対して，

$$\sum_{j,k=1}^{n} f(t_j - t_k) z_j \bar{z}_k \geq 0$$

となることを言う．

Exercise 13.5. ch.f. は正定値であることを示せ．

Remark 13.1 (Bochner の定理)**.** 逆に，与えられた関数 $\varphi : \mathbb{R} \to \mathbb{C}$ が正定値かつ原点で連続であって，さらに $\varphi(0) = 1$ なら，φ は ch.f. となることも知られている．この命題と Exercise 13.5 を合わせて，Bochner の定理と呼ばれる．証明は，Chung (2001, Section 6.5) を参照せよ．

14　モーメント法と Edgeworth 展開

モーメント法

本節は分布のモーメントと ch.f. の関係を考察する．次の補題は基本的である．

Lemma 14.1 (Durrett 3.3.14)**.** $\int |x|^n d\mu(x) < \infty$ なら，μ の ch.f. $\varphi(t)$ は C^n 級で，

$$\varphi^{(m)}(t) = \int (ix)^m e^{itx} d\mu(x),\ m = 1, \ldots, n.$$

$E[|X|^n] < \infty$ なら，

$$\varphi^{(m)}(0) = i^m E[X^m],\ m = 1, \ldots, n,$$

であるから，Taylor の定理より，

$$\varphi(t) = \sum_{m=0}^{n} \frac{E[(itX)^m]}{m!} + o(t^n),\ t \to 0$$

と展開できる．剰余項をより正確に評価してみると，Lemma 12.1 より，

$$\left| e^{itx} - \sum_{m=0}^{n} \frac{(itx)^m}{m!} \right| \le \min\left\{ \frac{|tx|^{n+1}}{(n+1)!}, \frac{2|tx|^n}{n!} \right\}.$$

であるから，$E[|X|^n] < \infty$ なら，

$$\left| \varphi(t) - \sum_{m=0}^{n} \frac{E[(itX)^m]}{m!} \right| \le E\left[\min\left\{ \frac{|tX|^{n+1}}{(n+1)!}, \frac{2|tX|^n}{n!} \right\} \right]$$

となる．

逆に，$\varphi'(0)$ が存在しても，$E[|X|] = \infty$ となることはありうる．cf. Durrett 3.3.17. ただし，次の評価は成り立つ．

Lemma 14.2. $E[X^2] \le \liminf_{h \downarrow 0} \{2\varphi(0) - \varphi(h) - \varphi(-h)\}/h^2$.

Proof. $(e^{ihx} - 2 + e^{-ihx})/h^2 = -2(1 - \cos(hx))/h^2 \le 0,\ 2(1 - \cos(hx))/h^2 \to x^2\ (h \downarrow 0)$ であるから，Fatou の補題より結論を得る． □

Remark 14.1. φ が原点近傍で 2 回微分可能なら，$\lim_{h \downarrow 0} \{\varphi(h) - 2\varphi(0) + \varphi(-h)\}/h^2 = \varphi''(0)$ である．従って，このとき，$E[X^2] < \infty$ である．一般に，k を正整数とし，φ が原点近傍で $(2k)$ 回微分可能なら，$E[X^{2k}] < \infty$ である．証明は演習問題とする．

応用として，次の定理を示す．

Theorem 14.1. $\varphi(t) = e^{-|t|^\alpha}, \alpha > 0$ が ch.f. $\Leftrightarrow 0 < \alpha \le 2$.

$\varphi(t) = e^{-c|t|^\alpha},\ 0 < \alpha \leq 2$ を ch.f. にもつ分布を，指数 α の対称安定分布と呼ぶ．

Proof. $\alpha > 2$ のとき，$\{\varphi(h) - 2\varphi(0) + \varphi(-h)\}/h^2 \to 0$ であるから，仮に φ が ch.f. なら，φ を ch.f にもつ r.v. X に対して，$E[X^2] = 0$, i.e., $\varphi \equiv 1$. 矛盾．

次に $0 < \alpha \leq 2$ のとき φ が ch.f. であることを示す．$\alpha = 2$ のときは e^{-t^2} は $N(0,1)$ の ch.f. だから，$0 < \alpha < 2$ の場合を考えればよい．$\psi(t) = 1 - (1 - \cos t)^{\alpha/2}$ とおく．ψ がある p.m. の ch.f. であることを確認する．$(1 - \cos t)^{\alpha/2}$ に一般化二項展開を適用して，

$$\psi(t) = \sum_{n=1}^{\infty} c_n (\cos t)^n,\ c_n = (-1)^{n+1} \frac{(\alpha/2)(\alpha/2-1)\cdots(\alpha/2-n+1)}{1 \cdot 2 \cdots n}.$$

定義より，$c_n \geq 0$. また $t = 0$ を代入すると，$\sum_{n=1}^{\infty} c_n = 1$ がわかる．$\nu = (1/2)\delta_{-1} + (1/2)\delta_1$ とおくと，ν の ch.f. は $\cos t$ なので，$\mu_n = \nu * \cdots * \nu$ (n 個の畳み込み) の ch.f. は $(\cos t)^n$. 従って，ψ は $\mu = \sum_n c_n \mu_n$ の ch.f. である．

$X_1, X_2, \cdots \sim \mu$ i.i.d. とし，$\sqrt{2}S_n/n^{1/\alpha}$ の ch.f. を φ_n とおくと，

$$\varphi_n(t) = \{\psi(\sqrt{2}t/n^{1/\alpha})\}^n = \{1 - (1 - \psi(\sqrt{2}t/n^{1/\alpha})\}^n.$$

$1 - \cos t \sim t^2/2,\ t \to 0$ より，

$$1 - \psi(\sqrt{2}t/n^{1/\alpha}) \sim n^{-2/\alpha} t^2,\ n \to \infty.$$

以上より，

$$\varphi_n(t) \sim \left(1 - \frac{t^2}{n^{2/\alpha}}\right)^n \sim e^{-|t|^\alpha} = \varphi(t),\ n \to \infty.$$

連続性定理より φ が ch.f. なことがわかる． □

もう 1 つの応用として，次の定理を示す．cf. Example 12.1.

Theorem 14.2. $\{X_m\}$ を i.i.d. r.v.'s とし，$S_n = \sum_{m=1}^n X_m$ とおく．このとき，$S_n/\sqrt{n} \xrightarrow{d} N(0,1)$ なら，$E[X_1] = 0, E[X_1^2] = 1$ である．

Proof. $X_1, Y_1, X_2, Y_2, \ldots$ を i.i.d. とし，$Z_m = (X_m - Y_m)/\sqrt{2}$ とおくと，$\{Z_m\}$ も i.i.d. であり，$T_n = \sum_{m=1}^n Z_m$ とおくと，$T_n/\sqrt{n} \xrightarrow{d} N(0,1)$ である．φ, ψ をそれぞれ X_1, Z_1 の ch.f.'s とすると，$\psi(t) = |\varphi(t/\sqrt{2})|^2 = \psi(-t)$ である．ここで，

$$\limsup_{h \downarrow 0} \frac{1 - \psi(h)}{h^2} \leq \frac{1}{2}$$

を示す[10]．$g(t) = 1 - \psi(t)$ とおくと，$(1 - g(t/\sqrt{n}))^n \to e^{-t^2/2}$ (広義一様)．仮に，ある $\delta > 0$ と $h_k \downarrow 0$ に対して，$g(h_k) > (1/2 + \delta)h_k^2$ となるとする．適当に部分列をとることにより，

[10] $\lim_{h \downarrow 0}\{1 - \psi(h)\}/h^2 = 1/2$ も示せる．

$h_n = |t_n|/\sqrt{n}$, $|t_n| \le 1$ と仮定してよい．このとき，$(1-g(t_n/\sqrt{n}))^n \le \exp\{-(1/2+\delta)t_n^2\}$ となるが，これは矛盾．

よって，

$$E[Z_1^2] \le \liminf_{h\downarrow 0} \frac{2\{1-\psi(h)\}}{h^2} \le \limsup_{h\downarrow 0} \frac{2\{1-\psi(h)\}}{h^2} \le 1.$$

ここで，X_1 の分布を μ とおくと，$\iint (x-y)^2 d\mu(x)d\mu(y) \le 2$. 特に，$\mu$-a.e. y に対して，$\int (x-y)^2 d\mu(x) < \infty$ である．そのような y を 1 つ固定すると，$\alpha \in (0,1)$ に対して，$(x-y)^2 = x^2 - 2xy + y^2 \ge (1-\alpha)x^2 - (\alpha^{-1}-1)y^2$, i.e., $(1-\alpha)x^2 \le (x-y)^2 + (\alpha^{-1}-1)y^2$. 従って，$E[X_1^2] = \int x^2 d\mu(x) < \infty$ を得る．このとき，CLT より，$(S_n - nE[X_1])/\sqrt{n} \xrightarrow{d} N(0, \mathrm{Var}(X_1))$ となるが，$S_n/\sqrt{n} \xrightarrow{d} N(0,1)$ であったから，$E[X_1] = 0, E[X_1^2] = 1$ でなくてはならない． □

$E[|X|^k] < \infty$ $\forall k = 1, 2, \ldots$ のとき，<u>形式的に</u>

$$\varphi(t) = \sum_{k=0}^{\infty} \frac{(it)^k m_k}{k!},\ m_k = E[X^k] \tag{1}$$

と展開できる．直感的には，m_k を k 次モーメントにもつ分布は一意に決まりそうな気もするが，この直感は正しくない！

Example 14.1. 対数正規分布の密度

$$f_0(x) = \frac{1}{\sqrt{2\pi}x} e^{-(\log x)^2/2},\ x > 0$$

と，それを変形した関数

$$f_a(x) = f_0(x)\{1 + a\sin(2\pi \log x)\},\ x > 0, -1 \le a \le 1$$

を考える．f_a は確率密度関数で，その任意次のモーメントは f_0 と同じ．実際，

$$\int_0^{\infty} x^r f_0(x) \sin(2\pi \log x) dx = 0,\ r = 0, 1, 2, \ldots.$$

$\{m_k\}$ をモーメント列にもつ分布が高々1 つに決まるための十分条件を 1 つ与える．

Theorem 14.3. $\limsup_k m_{2k}^{1/(2k)}/k = r < \infty$ なら，$\{m_k\}$ をモーメント列にもつ d.f. は高々1 つである．すなわち，$m_k = \int x^k dF(x)$ $(\forall k \in \mathbb{N})$ をみたす d.f. F は存在したとしても 1 つである．

例えば，X, Y をコンパクトな区間に集中した r.v.'s であって，すべての $k \in \mathbb{N}$ に対して，$E[X^k] = E[Y^k] = m_k$ (say) とすると，$m_{2k}^{1/(2k)}$ は有界なので，$\lim_k m_{2k}^{1/(2k)}/k = 0$ である．ゆえに，$X \overset{d}{=} Y$ となる [11].

[11] もっと直接に，Weierstrass の近似定理を使ってもよい．

Proof. そのような F が存在したとして，$X \sim F$ とする．$\tilde{m}_k = E[|X|^k]$ とおくと，$\tilde{m}_k \leq m_{2k}^{1/2}$ だから，$\limsup_k \tilde{m}_k^{1/k}/k \leq r < \infty$.

$$\left| e^{i\theta x} \left(e^{itx} - \sum_{k=0}^{n-1} \frac{(itx)^k}{k!} \right) \right| \leq \frac{|tx|^n}{n!}$$

なる評価を使うと，

$$\left| \varphi(\theta + t) - \varphi(\theta) - \sum_{k=1}^{n-1} \frac{\varphi^{(k)}(\theta)}{k!} t^k \right| \leq \frac{|t|^n}{n!} \tilde{m}_n.$$

$|t| < 1/(er)$ とすると，$\exists \varepsilon > 0$ s.t. $e(r+\varepsilon)|t| < 1$. 十分大きな n に対して，$\tilde{m}_n \leq (r+\varepsilon)^n n^n \leq (r+\varepsilon)^n e^n n!$ $(\because e^n \geq n^n/n!)$ だから，

$$\frac{|t|^n}{n!} \tilde{m}_n \leq ((r+\varepsilon)e|t|)^n \to 0,\ n \to \infty \text{ より},$$

$$\varphi(\theta + t) = \varphi(\theta) + \sum_{k=1}^{\infty} \frac{\varphi^{(k)}(\theta)}{k!} t^k,\ |t| < 1/(er).$$

$\{m_k\}$ をモーメント列にもつもう 1 つの d.f. G が存在したとして，その ch.f. を ψ と書く．すると，$\varphi(0) = \psi(0) = 1$, $\varphi^{(k)}(0) = \psi^{(k)}(0) = i^k m_k$ $(\forall k \in \mathbb{N})$ だから，$\varphi(t) = \psi(t)$ $(\forall |t| < 1/(er))$. この操作を繰り返すと，$\varphi(t) = \psi(t)$ $(\forall t \in \mathbb{R})$ を得る．よって，$F = G$. □

Theorem 14.4 (モーメント法)**.** 次の条件を仮定する:すべての $k \in \mathbb{N}$ に対して，$\lim_n \int x^k dF_n(x) =: m_k$ が存在して，$\{m_k\}$ は $\limsup_k m_{2k}^{1/2k}/k < \infty$ をみたす．このとき，F_n は弱収束し，その弱収束極限 F は $\int x^k dF(x) = m_k$ $(\forall k \in \mathbb{N})$ をみたす.

Proof. Markov の不等式より，

$$\int_{|x| \geq M} dF_n(x) \leq \frac{1}{M^2} \underbrace{\sup_m \int x^2 dF_m(x)}_{< \infty}$$

であるから，$\{F_n\}$ は一様にタイトである．$F_{n(m)} \xrightarrow{w} F$ とする．次の補題を示す.

Lemma 14.3. $X_n \xrightarrow{d} X$, $\sup_n E[|X_n|^p] < \infty$ $(\exists p > 1)$ なら，$\lim_n E[X_n] = E[X]$.

Proof of Lemma 14.3. Skorohod の表現定理より，$\exists Y_n \overset{d}{=} X_n, \exists Y \overset{d}{=} X$ s.t. $Y_n \to Y$ a.s. さらに，$\sup_n E[|Y_n|^p] = \sup_n E[|X_n|^p] < \infty$ より，$\{Y_n\}$ は一様可積分であるから，Theorem 6.6 より，$E[X_n] = E[Y_n] \to E[Y] = E[X]$. □

$X_n \sim F_n$, $X \sim F$ とすると，$X_{n(m)} \xrightarrow{d} X$. 連続写像定理より，任意の $k \in \mathbb{N}$ に対して，$X_{n(m)}^k \xrightarrow{d} X^k$. ここで，$\sup_m E[X_{n(m)}^{2k}] < \infty$ であることから，

$$m_k = \lim_m \int x^k dF_{n(m)}(x) = \lim_m E[X_{n(m)}^k] = E[X^k] = \int x^k dF(x) \qquad (*)$$

となる．$(*)$ をみたす F は部分列のとり方によらず一意に決まるから，Prohorov の定理の系 (Corollary 11.2) より，$F_n \xrightarrow{w} F$. □

Edgeworth 展開

$\{X_m : m \in \mathbb{N}\}$ を i.i.d. r.v.'s とし，$E[X_m] = 0, E[X_m^2] = 1$ とする．ここで，$F_n(x) = P(n^{-1/2}\sum_{m=1}^n X_m \le x)$ とおき，$N(0,1)$ の d.f. を Φ とおくと，CLT より，$n \to \infty$ のとき，$\sup_{x\in\mathbb{R}} |F_n(x) - \Phi(x)| \to 0$ となる．また，$E[|X_m|^3] < \infty$ なら，Berry-Esseen の定理より，収束のスピードは $O(n^{-1/2})$ である．さらに，適当な正則条件のもとで，F_n は

$$F_n(x) = \Phi(x) + \phi(x)\left\{\frac{p_1(x)}{n^{1/2}} + \frac{p_2(x)}{n} + \frac{p_3(x)}{n^{3/2}} + \cdots\right\} \qquad (*)$$

と展開できる．ここで，ϕ は $N(0,1)$ の密度関数であり，$p_1, p_2, \ldots$ は適当な多項式である．$(*)$ の形の展開を Edgeworth 展開 と呼ぶ．

以下，正則条件にはこだわらず，形式的に Edgeworth 展開を導出してみよう．X を $\{X_m\}$ の共通分布に従う r.v. とし，$\varphi(t)$ を X の ch.f. とする．このとき，φ は連続であり，$\varphi(0) = 1$ であるから，原点の近傍で $\log\varphi(t)$ が定義できる．与えられた正整数 $k \ge 3$ に対して，$E[|X|^k] < \infty$ なら，$\log\varphi(t)$ は原点の近傍で C^k 級であるから，

$$\log\varphi(t) = \sum_{j=1}^k \frac{\kappa_j}{j!}(it)^j + o(t^k),\ t \to 0$$

と展開できる．一方，$\varphi(t)$ は

$$\varphi(t) = 1 + \sum_{j=1}^k \frac{E[X^j]}{j!}(it)^j + o(t^k),\ t \to 0$$

と展開できるので，$\log\varphi(t)$ は，$t \to 0$ のとき，

$$\begin{aligned}\log\varphi(t) &= \sum_{\ell=1}^k \frac{(-1)^{\ell+1}}{\ell!}\{\varphi(t) - 1\}^\ell + o(|\varphi(t) - 1|^k) \\ &= \sum_{\ell=1}^k \frac{(-1)^{\ell+1}}{\ell!}\left\{\sum_{j=1}^k \frac{E[X^j]}{j!}(it)^j\right\}^\ell + o(t^k)\end{aligned}$$

とも展開できる．そこで，t^j の係数を比較して，各 κ_j が実数であることがわかる．κ_j を X の j 次 キュムラント (cumulant) と呼ぶ．簡単な計算から，$\kappa_1 = E[X] = 0, \kappa_2 = E[X^2] - (E[X])^2 = 1$ となることがわかるので，このとき，

$$\log\varphi(t) = -\frac{t^2}{2} + \sum_{j=3}^k \frac{\kappa_j}{j!}(it)^j + o(t^k)$$

と展開できる．

さて，$\check{S}_n := n^{-1/2}\sum_{m=1}^{n} X_m$ の ch.f. を $\varphi_n(t)$ とおくと，

$$\varphi_n(t) = \{\varphi(t/\sqrt{n})\}^n$$

となる．与えられた $T > 0$ に対して，n が十分大きいとき，$|t| \le T$ の範囲で，$\varphi(t/\sqrt{n}) \neq 0$ となるから，t を固定して，$n \to \infty$ とすると，

$$\log \varphi_n(t) = n \log \varphi(t/\sqrt{n}) = -\frac{t^2}{2} + \sum_{j=3}^{k} \frac{\kappa_j}{j!}(it)^j n^{-(j-2)/2} + o(n^{-(k-2)/2})$$

となる．従って，

$$\begin{aligned}
\varphi_n(t) &= \exp\left\{ -\frac{t^2}{2} + \sum_{j=3}^{k} \frac{\kappa_j}{j!}(it)^j n^{-(j-2)/2} + o(n^{-(k-2)/2}) \right\} \\
&= e^{-t^2/2} \exp\left\{ \sum_{j=3}^{k} \frac{\kappa_j}{j!}(it)^j n^{-(j-2)/2} \right\} + o(n^{-(k-2)/2}) \\
&= e^{-t^2/2} \left[1 + \sum_{\ell=1}^{k-2} \frac{1}{\ell!} \left\{ \sum_{j=3}^{k} \frac{\kappa_j}{j!}(it)^j n^{-(j-2)/2} \right\}^{\ell} \right] + o(n^{-(k-2)/2})
\end{aligned}$$

と展開できる．$T > 0$ は任意であるから，この展開は各 t に関して成り立つ．簡単のため，$k = 5$ としよう．このとき，

$$\begin{aligned}
\varphi_n(t) = e^{-t^2/2} \Bigg[1 &+ \frac{1}{n^{1/2}} \cdot \frac{\kappa_3}{6}(it)^3 + \frac{1}{n}\left\{ \frac{\kappa_4}{24}(it)^4 + \frac{\kappa_3^2}{72}(it)^6 \right\} \\
&+ \frac{1}{n^{3/2}}\left\{ \frac{\kappa_5}{120}(it)^5 + \frac{\kappa_3\kappa_4}{144}(it)^7 + \frac{\kappa_3^3}{1296}(it)^9 \right\} \Bigg] + o(n^{-3/2}).
\end{aligned}$$

いま，$\check{S}_n$ は密度関数 f_n をもつとし，φ_n に対して形式的に反転公式を適用すると，

$$f_n(x) = \phi(x)\left\{ 1 + \frac{r_1(x)}{n^{1/2}} + \frac{r_2(x)}{n} + \frac{r_3(x)}{n^{3/2}} \right\} + o(n^{-3/2})$$

という形になる．$r_j(x), j = 1, 2, 3$ を明示的に求めてみよう．

各 $j \in \mathbb{N}$ に対して，j 次 Hermite 多項式 $H_j(x)$ を

$$(-1)^j \frac{d^j}{dx^j}\phi(x) = H_j(x)\phi(x) \qquad (**)$$

と定義する．例えば，最初の 3 項は，$H_1(x) = x, H_2(x) = x^2 - 1, H_3(x) = x^3 - 3x$ である．また，$(**)$ の両辺を微分して，

$$\frac{d}{dx}\{H_j(x)\phi(x)\} = -H_{j+1}(x)\phi(x)$$

を得る．この関係より，$H_j(x)$ が j 次多項式となることが帰納的に確かめることができる．

いま，各 $j \in \mathbb{N}$ に対して，

$$\frac{1}{2\pi}\int_{\mathbb{R}} (it)^j e^{-itx} e^{-t^2/2} dt = \frac{(-1)^j}{2\pi}\frac{d^j}{dx^j}\int_{\mathbb{R}} e^{-itx} e^{-t^2/2} dt = (-1)^j \frac{d^j}{dx^j}\phi(x) = H_j(x)\phi(x)$$

となるから，$r_j(x), j = 1, 2, 3$ はそれぞれ，

$$\begin{aligned} r_1(x) &= \frac{\kappa_3}{6} H_3(x), \\ r_2(x) &= \frac{\kappa_4}{24} H_4(x) + \frac{\kappa_3^2}{72} H_6(x), \\ r_3(x) &= \frac{\kappa_5}{120} H_5(x) + \frac{\kappa_3\kappa_4}{144} H_7(x) + \frac{\kappa_3^3}{1296} H_9(x) \end{aligned}$$

となる．f_n の展開式を形式的に $-\infty$ から x まで積分して，

$$\begin{aligned} F_n(x) = \Phi(x) - \phi(x) \Bigg[&\frac{1}{n^{1/2}} \cdot \frac{\kappa_3}{6} H_2(x) + \frac{1}{n} \left\{ \frac{\kappa_4}{24} H_3(x) + \frac{\kappa_3^2}{72} H_5(x) \right\} \\ &+ \frac{1}{n^{3/2}} \left\{ \frac{\kappa_5}{120} H_4(x) + \frac{\kappa_3\kappa_4}{144} H_6(x) + \frac{\kappa_3^3}{1296} H_8(x) \right\} \Bigg] + o(n^{-3/2}) \end{aligned} \qquad (*3)$$

を得る．

以上の導出は形式的なものであり，厳密には，$o(n^{-3/2})$ の項に反転公式を適用したら $o(n^{-3/2})$ になるか，とか，f_n の展開式の剰余項を $-\infty$ から x まで積分したものが $o(n^{-3/2})$ になるか，といったことを確認する必要がある．とは言え，X が有限な 5 次モーメントをもち，次の Cramér 条件

$$\limsup_{|t|\to\infty} |\varphi(t)| < 1$$

がみたされるなら，$(*3)$ の展開が $x \in \mathbb{R}$ に関して一様に成り立つことが知られている．例えば，X の分布が絶対連続であれば，Riemann-Lebesgue の補題より，Cramér 条件はみたされる．しかし，例えば，$P(X = \pm 1) = 1/2$ のときは，$\varphi(t) = \cos t$ となるから，Cramér 条件はみたされない．後者の例に関しては，$\kappa_3 = E[X^3] = 0$ となるので，仮に $(*3)$ の展開が $x \in \mathbb{R}$ に関して一様に成り立つなら，$\sup_x |F_n(x) - \Phi(x)| = O(n^{-1})$ となるが，これは正しくない (cf. Exercise 12.3)．

Edgeworth 展開の証明は，Gnedenko and Kolmogorov (1968) を参照せよ．

演習問題

Exercise 14.1. X, Y を独立な r.v.'s とし，$X + Y \overset{d}{=} X$ とする．このとき，$Y = 0$ a.s. を示せ．

Exercise 14.2. X を r.v. とし，その ch.f. を φ とする．いま，k を正整数とし，φ は原点近傍で $(2k)$ 回微分可能とする．このとき，$E[X^{2k}] < \infty$ を示せ．

Exercise 14.3. $N(0,1)$ のモーメント列に対して，Theorem 14.3 の条件を確認せよ．

Exercise 14.4. $\{X_m : m \in \mathbb{N}\}$ を平均 0，分散 1 をもつ i.i.d. r.v.'s とし，すべての $\ell \in \mathbb{N}$ に対して，$E[|X_1|^\ell] < \infty$ とする．また，$\check{S}_n = n^{-1/2}\sum_{m=1}^n X_m$ とおく．このとき，あらゆる $\ell \in \mathbb{N}$ に対して，$\lim_n E[\check{S}_n^\ell] = E[Z^\ell]$ を示せ．ただし，$Z \sim N(0,1)$ である．

Exercise 14.5. $\{X_m\}$ を i.i.d. r.v.'s とし，$E[X_m] = \theta, \mathrm{Var}(X_m) = \sigma^2 > 0, E[|X_m|^3] < \infty$ とする．いま，σ^2 は既知とし，n 個のデータ $X_1, \ldots, X_n$ にもとづいて，θ の信頼区間を構成する問題を考える．$n^{-1/2}\sum_{m=1}^n (X_m - \theta)/\sigma$ の d.f. F_n に対して，Edgeworth 展開

$$F_n(x) = \Phi(x) - \frac{\phi(x)}{n^{1/2}} \cdot \frac{\kappa_3}{6} H_2(x) + O(n^{-1})$$

を認める．このとき，$\bar{X}_n = n^{-1}\sum_{m=1}^n X_m$ とおいたとき，与えられた $\alpha \in (0,1)$ に対して，

$$P\Big\{\theta \in [\bar{X}_n \pm \Phi^{-1}(1-\alpha/2)/(\sigma\sqrt{n})]\Big\} = 1 - \alpha + O(n^{-1})$$

を示せ．

Exercise 14.6. r.v. X が 格子点分布 (lattice distribution) をもつとは，ある $a, b \in \mathbb{R}$ が存在して，$P(X \in a + b\mathbb{Z}) = 1$ となることを言う．ここで，$a + b\mathbb{Z} := \{a + bd : d \in \mathbb{Z}\}$ である．いま，X を r.v. とし，その ch.f. を φ とおく．このとき，X が格子点分布をもつためには，ある $t_0 \neq 0$ が存在して，$|\varphi(t_0)| = 1$ となることが必要十分であることを示せ．これより，格子点分布は Cramér 条件をみたさないことがわかる．

十分性のヒント：$\varphi(t_0) = e^{i\theta_0}$ $(\theta_0 \in \mathbb{R})$ とすると，$0 = 1 - \mathrm{Re}\{e^{-i\theta_0}\varphi(t_0)\} = E[1 - \cos(t_0 X - \theta_0)]$ となる．

15　無限分解可能分布

CLT と Poisson 少数の法則は独立な三角列の和が，適当な仮定のもとで，$N(0,1)$ or $Po(\lambda)$ に分布収束することを述べている．そのような和の極限分布は無限分解可能分布として特徴づけられる．なお，無限分解可能分布と，次の節で扱う安定分布に関しては，Breiman (1968) と Gnedenko and Kolmogorov (1968) が詳しい．

r.v. X が 無限分解可能分布 (infinitely divisible distribution) をもつとは，各 $n \in \mathbb{N}$ に対して，

$$X \stackrel{d}{=} X_{n,1} + \cdots + X_{n,n}$$

なる i.i.d. r.v.'s $X_{n,1}, \ldots, X_{n,n}$ が存在することを言う．

Theorem 15.1. 各 $n \in \mathbb{N}$ に対して $X_{n,1}, \ldots, X_{n,n}$ を i.i.d. r.v.'s とし，$S_n = \sum_{m=1}^{n} X_{n,m} \stackrel{d}{\to} X$ $(n \to \infty)$ とする．このとき，X は無限分解可能分布をもつ．

Proof. S_{nm} を

$$S_{mn} = \underbrace{X_{mn,1} + \cdots + X_{mn,n}}_{=Y_{n,1}} + \cdots + \underbrace{X_{mn,(m-1)n+1} + \cdots + X_{mn,mn}}_{=Y_{n,m}}$$

と分割する．$Y_{n,1}, \ldots, Y_{n,m}$ は i.i.d. である．$S_n \stackrel{d}{\to} X$ とすると，

$$\begin{aligned}
&\{P(Y_{n,1} > y)\}^m = P(Y_{n,1} > y, \ldots, Y_{n,m} > y) \le P(S_{mn} > my), \\
&\{P(Y_{n,1} < -y)\}^m \le P(S_{mn} < -my)
\end{aligned}$$

となるから，$\{\mathcal{L}(Y_{n,1}) : n \in \mathbb{N}\}$ は一様にタイトである．従って，Prohorov の定理より，分布収束する部分列 $Y_{n(k),1}$ が存在する．$Y_{n(k),1} \stackrel{d}{\to} Z_{m,1}$ とする．$Z_{m,2}, \ldots, Z_{m,m}$ を $Z_{m,1}$ の独立なコピーとすると，$Y_{n(k),\ell} \stackrel{d}{=} Y_{n(k),1}$ より，各 $2 \le \ell \le m$ に対して，$Y_{n(k),\ell} \stackrel{d}{\to} Z_{m,\ell}$ $(k \to \infty)$ となる．いま，$Y_{n,1}, \ldots, Y_{n,m}$ は独立だから，$k \to \infty$ のとき，

$$\begin{array}{ccc}
S_{mn(k)} & = & Y_{n(k),1} + \cdots + Y_{n(k),m} \\
\downarrow_d & & \downarrow_d \\
X & \stackrel{d}{=} & Z_{m,1} + \cdots + Z_{m,m}
\end{array}$$

となる．これは X の分布が無限分解可能であることを意味する． □

無限分解可能分布の例として，正規分布，Cauchy 分布，Poisson 分布，複合 Poisson 分布などがある．

Example 15.1 (複合 Poisson 分布)**.** $\xi_1, \xi_2, \ldots$ を i.i.d. とし，$N(\lambda)$ をパラメータ λ をもつ Poisson 分布に従う r.v. とする．このとき，$Z = \xi_1 + \cdots + \xi_{N(\lambda)}$ の分布を複合 Poisson

分布と呼ぶ．$X_{n,j} \stackrel{d}{=} \xi_1 + \cdots + \xi_{N(\lambda/n)}$ とすれば，$Z \stackrel{d}{=} X_{n,1} + \cdots + X_{n,n}$ だから，複合 Poisson 分布は無限分解可能である．また，$\varphi(t) = E[\exp(it\xi_1)]$ とおくと，Z の ch.f. は，

$$E[e^{itZ}] = \sum_{n=0}^{\infty} e^{-\lambda} \frac{\lambda^n}{n!} \varphi(t)^n = \exp\{-\lambda(1 - \varphi(t))\}$$

である．$\nu = \mathcal{L}(\xi_1)$ とおけば，右辺は

$$\exp\left\{\lambda \int (e^{itx} - 1) d\nu(x)\right\}$$

と表せる．

無限分解可能分布の ch.f. は Lévy-Khinchin 表現 と呼ばれる特徴づけがある．その表現を導出してみよう．Lemma 13.3 (i) を思い出す．

Lemma 15.1 (Lemma 13.3 (i) の再掲)**.** $f : [-T, T] \to \mathbb{C}$ を連続で，$f(0) = 1$, $f(t) \neq 0\ \forall t \in [-T, T]$ をみたすとする．このとき，一意な連続関数 $\lambda : [-T, T] \to \mathbb{C}$ が存在して，$\lambda(0) = 0, f(t) = e^{\lambda(t)}\ \forall t \in [-T, T]$ をみたす．$\log f(t) := \lambda(t)$ と定義する．

X は無限分解可能分布をもつとする．X の分布を μ, ch.f. を $\varphi(t)$ とおく．

Lemma 15.2 (Durrett 3.8.3)**.** $\varphi(t) \neq 0\ \forall t \in \mathbb{R}$.

X は無限分解可能分布をもつから，ある i.i.d. r.v.'s $X_{n,m}, 1 \leq m \leq n$ が存在して，$X \stackrel{d}{=} X_{n,1} + \cdots + X_{n,n}$ となる．$X_{n,1}$ の分布を μ_n, ch.f. を φ_n とおくと，

$$\mu = \underbrace{\mu_n * \cdots * \mu_n}_{n},\ \varphi(t) = \{\varphi_n(t)\}^n$$

であり，φ, φ_n は $\neq 0$ であるから，$\log \varphi(t) = n \log \varphi_n(t)$. よって，

$$\varphi_n(t) = \exp\{(1/n) \log \varphi(t)\} \to 1 \text{ (広義一様)}$$

となるから，

$$\begin{aligned}
\log \varphi(t) &= n \log\{1 - (1 - \varphi_n(t))\} \\
&= n\{\varphi_n(t) - 1\}\{1 + \varepsilon_n(t)\} \quad (\varepsilon_n(t) \to 0 \text{ (広義一様)}) \\
&= \{1 + \varepsilon_n(t)\} \int (e^{itx} - 1) n d\mu_n(x) \qquad (*) \\
&= \{1 + \varepsilon_n(t)\} \int (e^{itx} - 1) d\tilde{\mu}_n(x) \quad (\tilde{\mu}_n := n\mu_n)
\end{aligned}$$

と展開できる．次の 2 つの補題を示す．

Lemma 15.3.

$$\limsup_n \tilde{\mu}_n(\{x : |x| \geq 2/u\}) \leq -2u^{-1} \int_0^u \mathrm{Re} \log \varphi(t) dt.$$

さらに，右辺 $\to 0$ $(u \downarrow 0)$.

Proof. 連続性定理の証明に現れた評価を使うと，

$$\mu_n(\{x : |x| \geq 2/u\}) \leq u^{-1} \int_{-u}^u \{1 - \varphi_n(t)\} dt = 2u^{-1} \int_0^u \{1 - \mathrm{Re}\varphi_n(t)\} dt.$$

従って，

$$\begin{aligned} \tilde{\mu}_n(\{x : |x| \geq 2/u\}) &\leq 2u^{-1} \int_0^u n\mathrm{Re}\big[1 - \exp\{(1/n) \log \varphi(t)\}\big] dt \\ &\to -2u^{-1} \int_0^u \mathrm{Re} \log \varphi(t) dt. \end{aligned}$$

さらに，$u \downarrow 0$ とすると，右辺 $\to -2\mathrm{Re} \log \varphi(0) = 0$. □

Lemma 15.4.

$$\limsup_n \int_{[-1,1]} x^2 d\tilde{\mu}_n(x) < \infty.$$

Proof.

$$n\{1 - \mathrm{Re}\varphi_n(1)\} = \int (1 - \cos x) d\tilde{\mu}_n(x) \geq \int_{[-1,1]} (1 - \cos x) d\tilde{\mu}_n(x).$$

$\cos x = 1 - (x^2/2)\cos(\theta x)$ $(\exists \theta \in [0,1])$ であるから，$\exists c > 0$ s.t. $1 - \cos x \geq cx^2$ $(\forall |x| \leq 1)$. よって，

$$\limsup_n \int_{[-1,1]} x^2 d\tilde{\mu}_n(x) \leq c^{-1} \lim_n n\{1 - \mathrm{Re}\varphi_n(1)\} = -c^{-1}\mathrm{Re} \log \varphi(1).$$

□

$(\mathbb{R}, \mathcal{B})$ 上の p.m. ν_n を

$$\nu_n(B) := \alpha_n^{-1} \int_B \frac{x^2}{1 + x^2} d\tilde{\mu}_n(x),\ B \in \mathcal{B},\ \alpha_n := \int_{\mathbb{R}} \frac{x^2}{1 + x^2} d\tilde{\mu}_n(x),$$

と定義する．

$$\frac{x^2}{1 + x^2} = \begin{cases} O(1) & |x| \to \infty \\ O(x^2) & x \to 0 \end{cases}$$

であるから，Lemma 15.3 と Lemma 15.4 より，$\limsup_n \alpha_n < \infty$ である．いま，任意に固定した $t \in \mathbb{R}$ に対して，

$$\begin{aligned}\log \varphi(t) &= (1+o(1)) \int (e^{itx}-1) d\tilde{\mu}_n(x) \\ &= (1+o(1)) \Bigg[\alpha_n \int \underbrace{\left(e^{itx} - 1 - \frac{itx}{1+x^2} \right) \frac{1+x^2}{x^2}}_{=:g(x)} d\nu_n(x) \\ &\qquad + it \underbrace{\int \frac{x}{1+x^2} d\tilde{\mu}_n(x)}_{=:\beta_n} \Bigg].\end{aligned}$$

ここで，$e^{itx} = 1 + itx - t^2x^2/2 + O(x^3)$, $1/(1+x^2) = 1 - x^2 + O(x^4)$ $(x \to 0)$ より，

$$\left(e^{itx} - 1 - \frac{itx}{1+x^2} \right) \frac{1+x^2}{x^2} = \left(-\frac{t^2x^2}{2} + O(x^3) \right) \frac{1+x^2}{x^2} = -\frac{t^2}{2} + o(1), \ x \to 0$$

と展開できて，$\lim_{x \to 0, x \neq 0} g(x) = -t^2/2$ となるから，g は $\mathbb{R}$ 上の連続関数に拡張できる．

Case 1. $\liminf_n \alpha_n > 0$. このとき，$\{\nu_n\}$ は一様にタイトだから，$\exists n(k) \uparrow \infty$ s.t. $\nu_{n(k)} \xrightarrow{w} \nu, \alpha_{n(k)} \to \alpha > 0$. ここで，$g$ は有界連続であるから，

$$\begin{aligned}&\alpha_{n(k)} \int g(x) d\nu_{n(k)}(x) \to \alpha \int g(x) d\nu(x) \\ &\quad = -\frac{t^2}{2} \underbrace{\alpha \nu(\{0\})}_{=:\sigma^2} + \int_{\mathbb{R} \setminus \{0\}} \left(e^{itx} - 1 - \frac{itx}{1+x^2} \right) \underbrace{\frac{1+x^2}{x^2} \alpha d\nu(x)}_{=:dM(x)}.\end{aligned}$$

Case 2. $\liminf_n \alpha_n = 0$. このとき，$\exists n(k) \uparrow \infty$ s.t. $\alpha_{n(k)} \to 0$. g は有界なので，

$$\alpha_{n(k)} \int g(x) d\nu_{n(k)}(x) \to 0.$$

いずれの場合も，$\lim_k \beta_{n(k)} =: \beta$ が存在する．$n(k)$ のとり方は t に依存しなかったから，β も t に依存しない．以上の議論をまとめて，次の定理の前半を得る．

Theorem 15.2 (Lévy-Khinchin)**.** X が無限分解可能分布をもつための必要十分条件は，定数 $\beta \in \mathbb{R}, \sigma \geq 0$ と，$(\mathbb{R}, \mathcal{B})$ 上の測度 M で，$M(\{0\}) = 0, \int (x^2 \wedge 1) dM(x) < \infty$ をみたすものが存在して，X の ch.f. $\varphi(t)$ が

$$\varphi(t) = \exp \left\{ i\beta t - \frac{\sigma^2 t^2}{2} + \int_{\mathbb{R}} \left(e^{itx} - 1 - \frac{itx}{1+x^2} \right) dM(x) \right\}, \ t \in \mathbb{R} \qquad (**)$$

と表せることである．

Remark 15.1 (Lévy 測度). $(**)$ の右辺を Lévy-Khinchin 表現と呼ぶ．また，M を Lévy 測度と呼ぶ．なお，

$$b = \beta - \int_{[-1,1]^c} \frac{x}{1+x^2} dM(x) + \int_{[-1,1]} \frac{x^3}{1+x^2} dM(x)$$

とおくと (右辺の積分は有限！)，Lévy-Khinchin 表現は

$$\varphi(t) = \exp\left[ibt - \frac{\sigma^2 t^2}{2} + \int \left\{e^{itx} - 1 - (itx)1_{[-1,1]}(x)\right\} dM(x)\right] \qquad (*3)$$

とも表せる (Lévy 過程 (後述) の本だと，この表現の方をよく見かける)．

Lévy 測度 M は原点の近傍で $+\infty$ の測度をもつ場合がある．次節で扱う安定分布を例にとると，指数 $\alpha \in (0,2)$ の非退化な安定分布の Lévy 測度は，$m_1 + m_2 > 0$ なる適当な定数 $m_1, m_2 \geq 0$ を用いて，

$$dM(x) = m_1 x^{-\alpha-1} 1_{(0,\infty)} dx + m_2 |x|^{-\alpha-1} 1_{(-\infty,0)} dx$$

と表されるので，あらゆる $\varepsilon > 0$ に対して，$M((-\varepsilon, \varepsilon)) = +\infty$ となる．

また，$e^{itx} - 1 = O(x)$ $(x \to 0)$ なので，$\int_{[-1,1]} |x| dM(x) < \infty$ であれば，

$$\gamma = b - \int_{[-1,1]} x dM(x) \qquad (*4)$$

とおくと，

$$\varphi(t) = \exp\left\{i\gamma t - \frac{\sigma^2 t^2}{2} + \int_{\mathbb{R}} (e^{itx} - 1) dM(x)\right\}$$

と表すことができる．特に，$M(\mathbb{R}) < \infty$ であれば，$\int_{\mathbb{R}} (e^{itx} - 1) dM(x)$ は複合 Poisson 分布の ch.f. なので，X の分布は $N(\gamma, \sigma^2)$ と複合 Poisson 分布の畳み込みになっている．

Proof of Theorem 15.2. 必要性はすでに示した．十分性を示す．$(**)$ の右辺を $\varphi_{\beta,\sigma^2,M}(t)$ とおいて，$\varphi_{\beta,\sigma^2,M}$ が常に ch.f. となることを示す．$(\mathbb{R}, \mathcal{B})$ 上の測度 ν を

$$\nu(B) = \int_B \frac{x^2}{1+x^2} dM(x),\ B \in \mathcal{B}$$

と定義すると，ν は有限測度である．

$$g(x,t) = \left(e^{itx} - 1 - \frac{itx}{1+x^2}\right) \frac{1+x^2}{x^2},\ x \neq 0$$

とおくと，$\lim_{x \to 0, x \neq 0} g(x,t) = -t^2/2$ であるから，任意に固定した t に対して，$g(x,t)$ を $x = 0$ まで連続に拡張できる．$g(x,t)$ は x の関数として有界連続だから，

$$x_{n,k} = k/n,\ I_{n,k} = (x_{n,k-1}, x_{n,k}],\ k = -n^2, \ldots, -1, 1, \ldots, n^2$$

とおくと，DCT より，

$$\sum_{|k|\le n^2,k\neq 0}\left(e^{itx_{n,k}}-1-\frac{itx_{n,k}}{1+x_{n,k}^2}\right)\frac{1+x_{n,k}^2}{x_{n,k}^2}\nu(I_{n,k})\to\int g(x,t)d\nu(x),\ n\to\infty.$$

ここで，

$$\beta_n=\beta-\sum_{|k|\le n^2,k\neq 0}\frac{\nu(I_{n,k})}{x_{n,k}},\ \lambda_{n,k}=\nu(I_{n,k})\frac{1+x_{n,k}^2}{x_{n,k}^2}$$

とおくと，

$$\psi_n(t)=\exp\left\{i\beta_n t-\frac{\sigma^2t^2}{2}+\sum_{|k|\le n^2,k\neq 0}\lambda_{n,k}(e^{itx_{n,k}}-1)\right\}$$

は ch.f. である．実際，$\xi_k,k=-n^2,\dots,n^2$ を独立な r.v.'s であって，$\xi_0\sim N(\beta_n,\sigma^2)$, $\xi_k\sim Po(\lambda_{n,k})$ $(|k|\le n^2,k\neq 0)$ とすると，$\psi_n(t)=E[\exp(it\sum_{|k|\le n^2}\xi_k)]$ である．さらに，$n\to\infty$ とすると，各 $t\in\mathbb{R}$ に対して，$\psi_n(t)\to\varphi_{\beta,\sigma^2,M}(t)$. ところで，$g(x,t)$ は $|t|\le 1$ の範囲で (x,t) の関数として有界であるから，DCT より，

$$\lim_{t\to 0}\int g(x,t)d\nu(x)=\int g(x,0)d\nu(x)=0.$$

従って，$\varphi_{\beta,\sigma^2,M}(t)$ は $t=0$ で連続であり，連続性定理より，$\varphi_{\beta,\sigma^2,M}$ は ch.f. である．

あとは $\varphi_{\beta,\sigma^2,M}$ が無限分解可能分布の ch.f. となることを示せばよいが，これは $\varphi_{\beta,\sigma^2,M}(t)=\{\varphi_{\beta/n,\sigma^2/n,M/n}(t)\}^n$ から従う． □

Theorem 15.3. Lévy-Khinchin 表現に現れるパラメータ (β,σ,M) は ch.f. から一意に決まる．

Proof. ch.f. φ は $(**)$ の表現をもつとする．まず，$\lim_{t\to\infty}\{\log\varphi(t)\}/t^2=-\sigma^2/2$ より，σ は φ から一意に決まる．次に，$\psi(t)=\log\varphi(t)+\sigma^2t^2/2$ とおいて，

$$\theta(t)=\psi(t)-\int_0^1\left\{\frac{\psi(t+u)-\psi(t-u)}{2}\right\}du$$

とおくと，$h(x,t)=e^{itx}-1-itx/(1+x^2)$ に対して，

$$\frac{1}{2}\{h(x,t+u)-h(x,t-u)\}=h(x,t)-e^{itx}(1-\cos(ux))$$

であり，$1-\cos x\sim x^2/2$ $(x\to 0)$ より，$dM(x)$ と du の積分順序の交換ができて，

$$\theta(t)=\int_{\mathbb{R}}e^{itx}\left\{\int_0^1(1-\cos(ux))du\right\}dM(x)=\int_{\mathbb{R}}e^{itx}\underbrace{\left(1-\frac{\sin x}{x}\right)dM(x)}_{=d\rho(x)}$$

を得る．ここで，$1-(\sin x)/x=O(x^2)$ $(x\to 0)$ より，ρ は $(\mathbb{R},\mathcal{B})$ 上の有限測度なので，反転公式より，ρ は θ から一意に決まる．これは，φ から M が一意に決まることを意味する．σ と M の一意性から，β の一意性も従う． □

Remark 15.2. Theorem 15.2 の証明中に現れた評価 $(*)$ より，

$$\exp\left\{n\int(e^{itx}-1)d\mu_n(x)\right\}\to\varphi(t),\ \forall t\in\mathbb{R}$$

がわかる．左辺は複合 Poisson 分布の ch.f. だから，無限分解可能分布は複合 Poisson 分布の弱収束極限で尽くされることがわかる．すなわち，$\mathcal{I}$ を無限分解可能分布の全体，$\mathcal{P}$ を複合 Poisson 分布の全体とし，$\mathcal{P}$ の弱位相に関する閉包を $\overline{\mathcal{P}}$ と書くと，$\mathcal{I}\subset\overline{\mathcal{P}}$ が示された．逆に，$\mathcal{P}\subset\mathcal{I}$ であって，$\mathcal{I}$ は弱位相に関して閉じているので，$\mathcal{I}=\overline{\mathcal{P}}$ が成り立つ．

Theorem 15.1 では，$X_{n,1},\dots,X_{n,n}$ は i.i.d. と仮定していたが，単に独立な場合は和 $S_n=\sum_{m=1}^n X_{m,n}$ の可能な極限分布は何であろうか？$X_{n,1},\dots,X_{n,n}$ に何も制約を課さない場合は任意の分布が極限分布になりうる．実際，μ を任意の p.m. とすると，$X_{n,1}\sim\mu, X_{n,2}=X_{n,n}=0$ とすれば，$S_n\sim\mu$ となる．しかし，直感的に言うと，各 $X_{m,n}$ が一様に小さいなら，S_n の極限分布は無限分解可能分布に限る．正確には次の定理が成り立つ．

Theorem 15.4. 各 $n\in\mathbb{N}$ に対して，$X_{n,1},\dots,X_{n,n}$ は独立であって，条件

$$\forall\varepsilon>0,\quad \sup_{1\le m\le n}P(|X_{m,n}|>\varepsilon)\to 0,\ n\to\infty$$

をみたすとする．このとき，$S_n=\sum_{m=1}^n X_{m,n}$ の極限分布は存在すれば無限分解可能である．

Proof. $S_n\xrightarrow{d}X$ として，X の ch.f. を評価していくと，それが Lévy-Khinchin 表現をもつことがわかる．従って，X は無限分解可能分布をもつ． □

Remark 15.3 (Lévy 過程)**.** 本講義ノートでは扱わないが，無限分解可能分布は Lévy 過程と深い関係がある．一般に，空でない集合 T に対して，r.v.'s の集まり $\{X_t:t\in T\}$ を確率過程と呼ぶ．確率過程 $\{X_t:t\in\mathbb{R}_+\}$ が，

(i) $X_0=0$ a.s.;

(ii) (定常独立増分性) 任意の $0\le t_1<\cdots<t_n$ に対して，$X_{t_2}-X_{t_1},\dots,X_{t_n}-X_{t_{n-1}}$ は独立かつ，各 j に対して，$X_{t_j}-X_{t_{j-1}}\overset{d}{=}X_{t_j-t_{j-1}}$;

(iii) (確率連続性) 任意の $t\ge 0$ に対して，$X_{t+h}\xrightarrow{P}X_t, h\to 0$,

をみたすとき，法則の意味での Lévy 過程と言う．法則の意味での Lévy 過程が，確率 1 で右連続かつ左極限が存在するパスをもつとき，それを単に Lévy 過程と呼ぶ (そのような “修正” は必ず存在することが知られている)．最後のパートで扱う Brown 運動は Lévy 過程の代表的な例である．Lévy 過程が重要である 1 つの理由として，それがジャンプをもつ連続時間確率過程の代表的なクラスであるためである．

無限分解可能分布との関連に戻ると，(ii) より，X_t の周辺分布は無限分解可能であることが分かる．逆に，X_t の分布は X_1 の分布から一意に決まり，与えられた無限分解可能分布を X_1 の分布にもつ Lévy 過程を構成できる．

Lévy 過程に関する基本的な定理として，Lévy-Itô 分解というものがあり，Lévy-Khinchin 表現はその副産物として導出できる．従って，Lévy-Khinchin 表現を別個に導出する必要があるのか，という疑問もあるが，Lévy-Itô 分解の証明は結構大変だし，とにかくそう言うものがあるんだ，というのを理解するために，Lévy-Khinchin 表現の導出を入れてみた．Lévy 過程に関する専門書として，佐藤 (1991), Sato (1999), Bertoin (1996), Applebaum (2009) を挙げておく．

演習問題

Exercise 15.1. ガンマ分布が無限分解可能であることを示せ．さらにその Lévy 測度を導出せよ．

以下，X は無限分解可能分布をもつ r.v. とし，その ch.f. $\varphi(t)$ は Lévy-Khinchin 表現 $(**)$ をもつとする．

Exercise 15.2. 次の 3 条件を仮定する：$\sigma = 0; M((-\infty, 0)) = 0; \int_{[0,1]} x dM(x) < \infty$．このとき，$X$ の分布のサポートは左に有界であって，その左端は γ であることを示せ．ただし，γ は $(*4)$ で与えたものである．

Exercise 15.3. 逆に，X のサポートが左に有界なら，前問の 3 条件がみたされることを示せ．

Exercise 15.4. M が有界な区間に集中していたら (i.e., $M([-K, K]^c) = 0$)，φ は無限回微分可能であることを示せ．このとき，あらゆる $r > 0$ に対して，$E[|X|^r] < \infty$ となる．

$(*3)$ の形の Lévy-Khinchin 表現を思い出す．このとき，$M_1(B) = M(B \cap [-1, 1]), M_2(B) = M(B \cap [-1, 1]^c)$ と定義すれば，$M = M_1 + M_2$ と分解できて，M_2 は有限測度である．また，M_1 は Lévy 測度である．そこで，X_1 を

$$\varphi_1(t) = \exp\left\{ibt - \frac{\sigma^2 t^2}{2} + \int (e^{itx} - 1 - itx) dM_1(x)\right\}$$

を ch.f. にもつ r.v. とし，X_2 を X_1 と独立な r.v. であって，

$$\varphi_2(t) = \exp\left\{\int (e^{itx} - 1) dM_2(x)\right\}$$

を ch.f. にもつ r.v. とすると，$\varphi(t) = \varphi_1(t)\varphi_2(t)$ であるから，$X \stackrel{d}{=} X_1 + X_2$ となる．X_2 の分布は複合 Poisson 分布である．

Exercise 15.5. $r > 0$ とする．このとき，$E[|X|^r] < \infty$ となるためには，$\int |x|^r dM_2(x) < \infty$ であることが必要十分であることを示せ．

Exercise 15.6 (Kolmogorov)**.** X を無限分解可能分布に従う r.v. とし，さらに $E[X] = 0,\ E[X^2] < \infty$ とする．このとき，$(\mathbb{R}, \mathcal{B})$ 上の有限測度 ρ が存在して，X の ch.f. が

$$\varphi(t) = \exp\left\{\int_{\mathbb{R}} \frac{e^{itx} - 1 - itx}{x^2} d\rho(x)\right\}$$

と表せることを示せ．また，$E[X^2] = \rho(\mathbb{R})$ を示せ．ただし，被積分関数の $x = 0$ での値は $x \to 0, x \neq 0$ のときの極限とする．

ヒント：前問より，$\int_{\mathbb{R}} x^2 dM(x) < \infty$ である．

16　安定分布

r.v. Y が 安定分布 (stable distribution) をもつとは，任意の正整数 k に対して，ある $a_k > 0, b_k \in \mathbb{R}$ が存在して，Y の独立なコピー $Y_1, \ldots, Y_k$ に対して，

$$\frac{\sum_{j=1}^k Y_j - b_k}{a_k} \stackrel{d}{=} Y, \text{ i.e., } \sum_{j=1}^k Y_j \stackrel{d}{=} a_k Y + b_k$$

となることを言う．定義より，安定分布は無限分解可能である．Y の ch.f. を $\varphi(t)$ とおくと，Y が安定分布をもつための必要十分条件は，任意の正整数 k に対して，ある $a_k > 0, b_k \in \mathbb{R}$ が存在して，$\{\varphi(t)\}^k = e^{itb_k}\varphi(a_k t)$ となることである．正規分布や Cauchy 分布は安定分布である．

Theorem 16.1. $\{X_m : m \in \mathbb{N}\}$ を i.i.d. r.v.'s とする．このとき，ある数列 $a_n > 0, b_n \in \mathbb{R}$ に対して，$(\sum_{m=1}^n X_m - b_n)/a_n \stackrel{d}{\to} Y$ であれば，Y は安定分布をもつ．

Theorem 16.1 は，i.i.d. r.v.'s の正規化した和の極限分布は安定分布に限ることを述べている．Theorem 16.1 の証明に次の定理を用いる．

Theorem 16.2 (タイプ収束定理)**.** ある数列 $\alpha_n > 0, \beta_n \in \mathbb{R}$ に対して，$W_n \stackrel{d}{\to} W, W_n' := \alpha_n W_n + \beta_n \stackrel{d}{\to} W'$ とする．このとき，W, W' の分布が非退化 (1 点分布でない) なら，$\exists\alpha > 0, \exists\beta \in \mathbb{R}$ s.t. $\alpha_n \to \alpha, \beta_n \to \beta, W' \stackrel{d}{=} \alpha W + \beta$ となる．

W_n, W, W' の d.f.'s をそれぞれ F_n, F, G とおくと，W_n' の d.f. は $F_n((x-\beta_n)/\alpha_n)$ である．Theorem 16.2 は，$F_n \stackrel{w}{\to} F$, $F_n((x-\beta_n)/\alpha_n) \stackrel{w}{\to} G$，かつ F と G が非退化なら，$\exists\alpha > 0, \exists\beta \in \mathbb{R}$ s.t. $\alpha_n \to \alpha$, $\beta_n \to \beta$,

$$G(x) = F((x-\beta)/\alpha), \ \forall x \in \mathbb{R} \qquad (*)$$

となることを述べている．$(*)$ のとき，F と G は同じ タイプ (type) をもつと言う．

Proof of Theorem 16.2. W_n' の d.f. を G_n とおくと，

$$F_n^{-1}(p) \to F^{-1}(p) \quad \forall p \in \mathcal{C}(F^{-1}), \quad G_n^{-1}(p) \to G^{-1}(p) \quad \forall p \in \mathcal{C}(G^{-1})$$

となる．いま，F と G は非退化なので，$\exists p_1, p_2 \in \mathcal{C}(F^{-1}) \cap \mathcal{C}(G^{-1})$ s.t.

$$-\infty < F^{-1}(p_1) < F^{-1}(p_2) < \infty, \ -\infty < G^{-1}(p_1) < G^{-1}(p_2) < \infty.$$

ここで，$G_n^{-1}(p) = \alpha_n F_n^{-1}(p) + \beta_n$ であるから，

$$\alpha_n = \frac{G_n^{-1}(p_2) - G_n^{-1}(p_1)}{F_n^{-1}(p_2) - F_n^{-1}(p_1)} \to \frac{G^{-1}(p_2) - G^{-1}(p_1)}{F^{-1}(p_2) - F^{-1}(p_1)} =: \alpha$$

を得る．また，$\beta_n = G_n^{-1}(p_1) - \alpha_n F_n^{-1}(p_1) \to G^{-1}(p_1) - \alpha F^{-1}(p_1) =: \beta$ となる．　□

Proof of Theorem 16.1. Y の分布が退化しているときは定理の結論は明らかなので，Y の分布は非退化と仮定する．k を正整数とし，$Y_1, \dots, Y_k$ を Y の独立なコピーとする．$Z_n = (\sum_{m=1}^n X_m - b_n)/a_n$, $S_n^j = \sum_{m=(j-1)n+1}^{jn} X_m$ とおくと，

$$
\begin{aligned}
&W_n := Z_{nk} = (S_n^1 + \cdots + S_n^k - b_{nk})/a_{nk}, \\
&a_{nk} Z_{nk} = (S_n^1 - b_n) + \cdots + (S_n^k - b_n) + (kb_n - b_{nk}), \\
&W_n' := \underbrace{(a_{nk}/a_n)}_{=:\alpha_n} Z_{nk} + \underbrace{(b_{nk} - kb_n)/a_n}_{=:\beta_n} = (S_n^1 - b_n)/a_n + \cdots + (S_n^k - b_n)/a_n.
\end{aligned}
$$

$n \to \infty$ のとき，$W_n \xrightarrow{d} Y, W_n' = \alpha_n W_n + \beta_n \xrightarrow{d} \sum_{j=1}^k Y_j$ となるから，タイプ収束定理より，$\alpha_n \to \alpha, \beta_n \to \beta, \sum_{j=1}^k Y_j \stackrel{d}{=} \alpha Y + \beta$ を得る． □

次に安定分布の指数を定義する．

Theorem 16.3 (安定分布の指数)**.** X を非退化な安定分布に従う r.v. とし，$X_1, X_2, \dots$ を X の独立なコピーとする．このとき，一意な定数 $\alpha \in (0, 2]$ が存在して，各正整数 k に対して，ある $b_k \in \mathbb{R}$ が存在して，$\sum_{j=1}^k X_j \stackrel{d}{=} k^{1/\alpha} X + b_k$ となる．さらに，X の ch.f. $\varphi(t)$ の Lévy-Khinchin 表現のパラメータを (β, σ, M) とおくと，$\alpha = 2$ のとき，$\sigma > 0, M = 0$ となり，$\alpha \in (0, 2)$ のときは，$\sigma = 0$ であって，Lévy 測度 M は

$$
dM(x) = m_1 x^{-\alpha-1} 1_{(0,\infty)}(x) dx + m_2 |x|^{-\alpha-1} 1_{(-\infty,0)}(x) dx
$$

と表される．ただし，m_1, m_2 は $m_1, m_2 \geq 0, m_1 + m_2 > 0$ なる定数である．

α を安定分布の <u>指数</u> (index) と呼ぶ．

Proof. X は安定分布に従うから，各正整数 k に対して，適当な定数 $a_k > 0, b_k \in \mathbb{R}$ が存在して，$\{\varphi(t)\}^k = e^{itb_k} \varphi(a_k t)$ となる．$M_k(B) = M(\{x : a_k x \in B\})$ $(B \in \mathcal{B})$ とおくと，両辺の Lévy-Khinchin 表現を比較して，

$$
ik\beta t - \frac{k\sigma^2 t^2}{2} + \int_{\mathbb{R}} h(x,t) k dM(x) = i(\beta a_k + b_k + c_k)t - \frac{a_k^2 \sigma^2 t^2}{2} + \int_{\mathbb{R}} h(x,t) dM_k(x)
$$

を得る．ただし，$h(x,t) = e^{itx} - 1 - itx/(1+x^2)$ であって，

$$
c_k = a_k \int_{\mathbb{R}} \left(\frac{x}{1 + a_k^2 x^2} - \frac{x}{1+x^2} \right) dM(x)
$$

である (右辺の積分は有限)．このとき，Lévy-Khinchin 表現のパラメータの一意性より，

$$
k\beta = \beta a_k + b_k + c_k, \ k\sigma^2 = a_k^2 \sigma^2, \ kM = M_k
$$

を得る．σ が 0 か 0 でないかで場合分けする．

Case 1. $\sigma \neq 0$. このとき，$a_k = \sqrt{k}$ である．$\psi(t) = \log \varphi(t)$ とおくと，$k\psi(t) = ib_k t + \psi(\sqrt{k}t)$ となる．$\psi(t)/t^2 \to -\sigma^2/2$ $(|t| \to \infty)$ であるから，$\psi(\sqrt{k}t)/k \to -\sigma^2 t^2/2$ $(k \to \infty)$. よって，$\psi(t) = i\beta t - \sigma^2 t^2/2$ を得る．

Case 2. $\sigma = 0$. $H^+(x) = M([x, \infty))$ $(x > 0)$, $H^-(x) = M((-\infty, x])$ $(x < 0)$ とおくと，X の分布は非退化なので，両方とも $\equiv 0$ となることはありえない．いま，$H^+ \not\equiv 0$ のとき，十分小さな $x_0 > 0$ に対して，$H^+(x_0) > 0$ となる．このとき，$kH^+(x_0) = H^+(x_0/a_k)$ より，$a_k \uparrow \infty$ であり，また，安定分布の定義から，$a_{nk} = a_n a_k$ となるから，$a_{k^j} = (a_k)^j$ を得る．$k > 1$ を任意に固定し，$n > k$ に対して，$k^j \leq n \leq k^{j+1}$ となるように j を選ぶと，$a_{k^j} \leq a_n \leq a_{k^{j+1}}$ より，$j \log a_k \leq \log a_n \leq (j+1) \log a_k$ となる．両辺を $j \log k$ で割って，

$$\frac{\log a_k}{\log k} \leq \frac{\log a_n}{\log n} \cdot \frac{\log n}{j \log k} \leq \frac{j+1}{j} \cdot \frac{\log a_k}{\log k}.$$

$n \to \infty$ とすると，$j \to \infty$ となるが，$(\log n)/(j \log k) \to 1$ であるから，

$$\frac{\log a_k}{\log k} = \lim_n \frac{\log a_n}{\log n} =: \lambda$$

を得る．ここで，λ は k に依存せず，$a_k = k^\lambda$ となる．$\lambda \in (0, \infty)$ である．

従って，$kH^+(x) = H^+(x/k^\lambda)$ $(x > 0, k \geq 1)$ となり，これに $x = (k/n)^\lambda$ を代入して，

$$kH^+((k/n)^\lambda) = H^+(1/n^\lambda),\ k, n \geq 1$$

を得る．ここで，$k = n$ とすると，$nH^+(1) = H^+(1/n^\lambda)$ であるから，

$$H^+((k/n)^\lambda)) = (n/k)H^+(1),\ k, n \geq 1$$

を得る．$\{(k/n)^\lambda : k, n \geq 1\}$ は $(0, \infty)$ で稠密であり，H^+ は左連続だから，$\alpha = 1/\lambda$ とおいて，$H^+(x) = x^{-\alpha} H^+(1)$ $(x > 0)$ を得る．また，同様の議論より，$H^-(x) = |x|^{-\alpha} H^-(-1)$ $(x < 0)$ を得る．従って，Lévy 測度 M は

$$dM(x) = \alpha H^+(1) x^{-\alpha-1} 1_{(0,\infty)} dx + \alpha H^-(-1)|x|^{-\alpha-1} 1_{(-\infty,0)} dx$$

となるが，$\int_{[-1,1]} x^2 dM(x) < \infty$ でなくてはならないので，$\alpha \in (0, 2)$ である． □

$\alpha = 2$ のときは，$M = 0$ なので，ch.f. は $\varphi(t) = e^{i\beta t - \sigma^2 t^2/2}$ となる．従って，$X \sim N(\beta, \sigma^2)$ である．$\alpha \in (0, 2)$ のときも，ch.f. を具体的に計算することができて，次の定理を得る．詳細は，Breiman (1968, Chapter 9)，または Gnedenko and Kolmogorov (1968, Chapter 7) を参照せよ．

Theorem 16.4. X が指数 $\alpha \in (0, 2]$ の非退化な安定分布をもてば，適当な定数 $c \in \mathbb{R}, d > 0, |\theta| \leq 1$ が存在して，

$$\varphi(t) := E[e^{itX}] = \begin{cases} \exp\{itc - d|t|^\alpha \left(1 + i\theta \operatorname{sgn}(t) \tan \frac{\pi}{2}\alpha\right)\} & \text{if } \alpha \in (0, 2] \setminus \{1\} \\ \exp\{itc - d|t| \left(1 + i\theta \operatorname{sgn}(t) \frac{2}{\pi} \log |t|\right)\} & \text{if } \alpha = 1 \end{cases} \qquad (**)$$

と表せる．逆に，あらゆる $\alpha \in (0,2], c \in \mathbb{R}, d > 0, |\theta| \leq 1$ に対して，$(**)$ の右辺は非退化な安定分布の ch.f. である．

X が対称のとき ($X \stackrel{d}{=} -X$)，$\varphi(t) = e^{-d|t|^\alpha}$ であり，その分布を指数 α の対称安定分布と呼ぶ．安定分布は裾の重い分布の代表的なクラスである．実際，X が指数 $\alpha \in (0,2)$ の非退化な安定分布に従うとき，$E[|X|^r] < \infty$ $(0 < \forall r < \alpha)$ だが $E[|X|^\alpha] = \infty$ となる (cf. Exercise 16.1)．一方，指数が $\alpha = 2$ の安定分布は正規分布なので，その場合はすべてのモーメントが存在して有限である．

$X_1, X_2, \cdots \sim F$ i.i.d. として，ある数列 $a_n > 0, b_n \in \mathbb{R}$ に対して，$(\sum_{m=1}^n X_m - b_n)/a_n \stackrel{d}{\to} G$ となるとき，F は G の 吸引域 (domain of attraction) に属すると言う．G は必ず安定分布である．CLT より，2 次モーメントが有限な F は $N(0,1)$ の吸引域に属する．ところで，Example 12.1 より，F が有限な 2 次モーメントをもたない場合でも，$N(0,1)$ の吸引域に属することがある．

Theorem 16.5. $X \sim F$ とし，F は非退化とする．このとき，F が $N(0,1)$ の吸引域に属するための必要十分条件は，

$$\frac{xP(|X| > x)}{E[X^2 1_{\{|X| \leq x\}}]} \to 0, \ x \to \infty$$

となることである．

指数 $\alpha \in (0,2)$ をもつ安定分布の吸引域に関しても，その特徴づけが知られている．詳細は，Breiman (1968, Chapter 9)，または Gnedenko and Kolmogorov (1968, Chapter 7) を参照せよ．

Example 16.1 (指数 $\alpha \in (0,2)$ の対称安定分布の吸引域に属する分布の例)**.** $\alpha \in (0,2)$ とする．$X_1, X_2, \ldots$ を i.i.d. r.v.'s とし，$P(X_1 > x) = P(X_1 < -x) = x^{-\alpha}/2$ $(x \geq 1)$ とする．$\varphi(t) = E[e^{itX_1}]$ とおくと，X_1 の分布の対称性より，$t \neq 0$ のとき，

$$1 - \varphi(t) = E[1 - \cos(tX_1)] = \alpha \int_1^\infty (1 - \cos(tx)) x^{-\alpha-1} dx = |t|^\alpha \alpha \int_t^\infty (1 - \cos u) u^{-\alpha-1} du.$$

ここで，$c = \alpha \int_0^\infty (1 - \cos u) u^{-\alpha-1} du \in (0, \infty)$ とおくと，$1 - \varphi(t) \sim c|t|^\alpha$ $(t \to 0)$ である．従って，$S_n = \sum_{m=1}^n X_m$ とおくと，$n \to \infty$ のとき，

$$E[e^{itS_n/n^{1/\alpha}}] = \{\varphi(t/n^{1/\alpha})\}^n = \{1 - (1 - \varphi(t/n^{1/\alpha}))\}^n \sim \left(1 - \frac{c|t|^\alpha}{n}\right)^n \to e^{-c|t|^\alpha}$$

を得る．

演習問題

Exercise 16.1. X が指数 $\alpha \in (0,2)$ の (非退化な) 安定分布に従うとき，$0 < r < \alpha$ なら $E[|X|^r] < \infty$ となることを示せ．また，$E[|X|^\alpha] = \infty$ となることを示せ．

ヒント：Exercise 15.5.

Exercise 16.2. X が指数 $\alpha \in (0,2)$ の非退化な対称安定分布に従うとき，適当な定数 $d > 0$ が存在して，その ch.f. が $\varphi(t) = e^{-d|t|^\alpha}$ と表せることを示せ.

Exercise 16.3. Theorem 16.4 の後半を示せ.

17　極値分布

$\{X_m : m \in \mathbb{N}\}$ を i.i.d. r.v.'s として，$\max_{1\le m\le n} X_m$ の，適当に正規化した後での，極限分布を考える．次の定理はそのような極限分布は 3 種類しかないことを示している．d.f.'s F, G に対して，ある $\alpha > 0, \beta \in \mathbb{R}$ が存在して，

$$G(x) = F(\alpha x + \beta),\ \forall x \in \mathbb{R}$$

となるとき，F と G は同じタイプをもつと言うのであった．

Theorem 17.1 (Fischer-Tippett-Gnedenko)**.** $\{X_m\}$ を i.i.d. r.v.'s とする．このとき，ある数列 $a_n > 0, b_n \in \mathbb{R}$ が存在して，$(\max_{1\le m\le n} X_m - b_n)/a_n$ が非退化な分布 G に分布収束するならば，G は次の 3 つの分布のどれかと同じタイプである．

$$\text{(i) } \Phi_\alpha(x) = \begin{cases} 0 & \text{if } x < 0 \\ e^{-x^{-\alpha}} & \text{if } x \ge 0 \end{cases},\ \alpha > 0.$$

$$\text{(ii) } \Psi_\alpha(x) = \begin{cases} e^{-(-x)^{\alpha}} & \text{if } x < 0 \\ 1 & \text{if } x \ge 0 \end{cases},\ \alpha > 0.$$

$$\text{(iii) } \Lambda(x) = e^{-e^{-x}}, x \in \mathbb{R}.$$

(i)–(iii) の分布を総称して 極値分布 (extreme value distribution) と呼ぶ．個別には，Φ_α は Fréchet 分布，Ψ_α は Weibull 分布，Λ は Gumbel 分布と呼ばれる．

Proof. $t \in \mathbb{R}$ に対して，$\lfloor t \rfloor$ を t の整数部分とする．X_1 の d.f. を F とおくと，

$$P((\max_{1\le m\le n} X_m - b_n)/a_n \le x) = P(\max_{1\le m\le n} X_m \le a_n x + b_n) = F^n(a_n x + b_n) \xrightarrow{w} G(x)$$

となるから，$t > 0$ に対して，

$$\begin{aligned} &F^{\lfloor nt \rfloor}(a_{\lfloor nt \rfloor} x + b_{\lfloor nt \rfloor}) \xrightarrow{w} G(x), \\ &F^{\lfloor nt \rfloor}(a_n x + b_n) = \{F^n(a_n x + b_n)\}^{\lfloor nt \rfloor / n} \xrightarrow{w} G^t(x). \end{aligned}$$

タイプ収束定理より，$\exists \alpha(t) > 0, \exists \beta(t) \in \mathbb{R}$ s.t.

$$\lim_n a_n / a_{\lfloor nt \rfloor} = \alpha(t),\ \lim_n (b_n - b_{\lfloor nt \rfloor})/a_{\lfloor nt \rfloor} = \beta(t),\ G^t(x) = G(\alpha(t)x + \beta(t)).$$

$t \mapsto \alpha(t), t \mapsto \beta(t)$ は可測である．$t, s > 0$ とすると，

$$\begin{aligned} G^{ts}(x) &= G(\alpha(ts)x + \beta(ts)), \\ G^{ts}(x) &= \{G^s(x)\}^t = \{G(\alpha(s)x + \beta(s))\}^t \\ &= G(\alpha(t)\{\alpha(s)x + \beta(s)\} + \beta(t)) \\ &= G(\alpha(t)\alpha(s)x + \alpha(t)\beta(s) + \beta(t)). \end{aligned}$$

Lemma 17.1. G を非退化な d.f. とし，$a > 0, c > 0, b, d \in \mathbb{R}$ とする．このとき，

$$G(ax + b) = G(cx + d),\ \forall x \in \mathbb{R}$$

なら，$a = c, b = d$.

Proof of Lemma 17.1. $H_1(x) = G(ax + b), H_2(x) = G(cx + d)$ とおくと，$H_1^{-1}(p) = aG^{-1}(p) + b, H_2^{-1}(p) = cG^{-1}(p) + d$. $H_1^{-1} = H_2^{-1}$ だから，

$$(a - c)G^{-1}(p) + (b - d) = 0.$$

G は非退化だから，$\exists p_1, p_2 \in (0, 1)$ s.t. $G^{-1}(p_1) \neq G^{-1}(p_2)$. よって，$a = c, b = d$. □

定理の証明に戻ると，

$$\underbrace{\alpha(ts) = \alpha(t)\alpha(s)}_{(*)},\ \beta(ts) = \alpha(t)\beta(s) + \beta(t) = \alpha(s)\beta(t) + \beta(s),\ t, s > 0.$$

$(*)$ は Hamel 方程式と呼ばれる．次の事実を認める．$(*)$ をみたす非負可測関数は

$$\alpha(t) = t^{-\theta},\ \theta \in \mathbb{R}$$

なる形の関数のみである．次の 3 つのケースを考える．Case 1: $\theta = 0$, Case 2: $\theta > 0$, Case 3: $\theta < 0$.

Case 1. $\theta = 0$. このとき，$\alpha(t) = 1$ であるから，

$$\beta(ts) = \beta(s) + \beta(t),\ \text{i.e.,}\ e^{\beta(ts)} = e^{\beta(s)}e^{\beta(t)}.$$

これは Hamel 方程式だから，$\exists c \in \mathbb{R}$ s.t.

$$e^{\beta(t)} = t^{-c},\ \text{i.e.,}\ \beta(t) = -c\log t.$$

このとき，

$$G^t(x) = G(x - c\log t) \qquad (**)$$

となる．ここで，仮に $c = 0$ なら，$G(x) < 1$ のとき，$G(x) = G^t(x) \to 0$ $(t \to \infty)$ となる．これは $G(x) \in \{0, 1\}$ $(\forall x \in \mathbb{R})$ を意味するが，G は非退化であったから，$c \neq 0$ である．また，$t \mapsto G^t(x)$ は非増加だから，$c > 0$ である．次に，仮に $\exists x_0 \in \mathbb{R}$ s.t. $G(x_0) = 1$ なら，$1 = G(x_0 - c\log t)$ $(\forall t > 0)$, i.e., $G(u) = 1$ $(\forall u \in \mathbb{R})$ となるが，これは G が d.f. であることに反する．従って，$G(x) < 1$ $(\forall x \in \mathbb{R})$. 同様に，$G(y) > 0$ $(\forall y \in \mathbb{R})$.

いま，$(**)$ に $x = 0$ を代入して，$G^t(0) = G(-c\log t)$ を得る．そこで，$u = -c\log t$ と変数変換して，$e^{-e^{-p}} = G(0) \in (0, 1)$ とおくと，

$$G(u) = e^{-te^{-p}} = e^{-e^{-c^{-1}u-p}} = \Lambda(c^{-1}u + p)$$

となる．従って，このとき，G は Λ と同じタイプをもつ．

Case 2. $\theta > 0$.

$$\alpha(t)\beta(s) + \beta(t) = \alpha(s)\beta(t) + \beta(s)$$

であるから，$s \neq 1, t \neq 1$ に対して，

$$\frac{\beta(s)}{1-\alpha(s)} = \frac{\beta(t)}{1-\alpha(t)}, \quad \text{i.e.,} \quad \frac{\beta(t)}{1-\alpha(t)} = c. \text{ (say)}$$

このとき，$t \neq 1$ に対して，$\beta(t) = c(1-\alpha(t)) = c(1-t^{-\theta})$ となるから，$G^t(x) = G(t^{-\theta}x + c(1-t^{-\theta})) = G(t^{-\theta}(x-c)+c)$. 変数変換をして，$G^t(x+c) = G(t^{-\theta}x+c)$. $H(x) = G(x+c)$ とおく．いま，$H^t(x) = H(t^{-\theta}x)$, $t \neq 1$ であるから，$x = 0$ を代入して，$H^t(0) = H(0)$, i.e., $H(0) \in \{0, 1\}$．仮に $H(0) = 1$ だと G の非退化性より，$\exists x < 0$ s.t. $H(x) < 1$ であり，この x に対して，$t \mapsto H^t(x)$ は狭義単調減少であるが，$t \mapsto H(t^{-\theta}x)$ は非減少であるから，矛盾が生じる．よって，$H(0) = 0$.

次に，$H^t(1) = H(t^{-\theta})$ より，$t \to \infty$ とすると，$H^t(1) \to 0$ となる．よって，$H(1) < 1$ である．仮に $H(1) = 0$ なら，$H(t^{-\theta}) = 0$ となるが，このとき $H \equiv 0$ となって，矛盾が生じる．よって，$H(1) \in (0, 1)$. そこで，$\alpha = \theta^{-1}, H(1) = e^{-p^{-\alpha}}, u = t^{-\theta}$ とおくと，

$$H(u) = H^t(1) = e^{-p^{-\alpha}t} = e^{-(pu)^{-\alpha}} = \Phi_\alpha(pu).$$

従って，このとき，G は Φ_α と同じタイプをもつ．

Case 3. $\theta < 0$. Case 2 と同様．このとき，G は Ψ_α と同じタイプをもつ． □

$X_1, X_2, \cdots \sim F$ i.i.d., $M_n = \max_{1 \leq m \leq n} X_m$ とし，ある $a_n > 0, b_n$ に対して，$(M_n - b_n)/a_n \xrightarrow{d} G$ となるとき，F は G の 吸引域 (domain of attraction) に属すると言う．G は非退化なら極値分布のいずれかと同じタイプをもつ．F が極値分布のそれぞれの吸引域に属するための必要十分条件については，Resnick (1987, Chapter 1) を参照せよ．ここでは，$N(0, 1)$ が Gumbel 分布の吸引域に属することを示す．

Theorem 17.2. $X_1, X_2, \cdots \sim N(0, 1)$ i.i.d. とし，$M_n = \max_{1 \leq m \leq n} X_m$ とおく．このとき，

$$a_n = (2\log n)^{-1/2}, \; b_n = (2\log n)^{1/2} - \frac{1}{2}(2\log n)^{-1/2}(\log\log n + \log 4\pi)$$

に対して，$(M_n - b_n)/a_n \xrightarrow{d} \Lambda$.

証明は次の補題による．

Lemma 17.2. $X_1, X_2, \cdots \sim F$ i.i.d. とする．与えらえた定数 $\tau \geq 0$ と数列 u_n に対して，

$$n(1 - F(u_n)) \to \tau \Leftrightarrow P(M_n \leq u_n) \to e^{-\tau}.$$

Proof. $\Rightarrow$.

$$P(M_n \le u_n) = F^n(u_n) = \{1-(1-F(u_n))\}^n = (1-\tau/n+o(n^{-1}))^n \to e^{-\tau}.$$

$\Leftarrow$.

$$n\log\{1-(1-F(u_n))\} = \log P(M_n \le u_n) \to -\tau.$$

$1-F(u_n)\to 0$ だから，左辺 $=-n(1-F(u_n))(1+o(1))$ より，$n(1-F(u_n))\to\tau$. □

Proof of Theorem 17.2. $N(0,1)$ の密度関数と d.f. をそれぞれ ϕ と Φ とおく．すなわち，

$$\phi(x) = \frac{1}{\sqrt{2\pi}}e^{-x^2/2},\ \Phi(x) = \int_{-\infty}^{x}\phi(y)dy.$$

u_n を $n(1-\Phi(u_n)) = e^{-x}$ により定義すれば，$P(M_n \le u_n)\to\Lambda(x)$. u_n を評価していく．$1-\Phi(u)\sim\phi(u)/u$ $(u\to\infty)$ より，$n^{-1}e^{-x}u_n/\phi(u_n)\to 1$. 両辺の対数をとって整理すると，

$$-\log n - x + \log u_n + \frac{1}{2}\log 2\pi + \frac{u_n^2}{2} \to 0.$$

$u_n\to\infty, u_n^2/(2\log n)\to 1$ だから，$2\log u_n - \log 2 - \log\log n \to 0$, i.e.,

$$\log u_n = \frac{1}{2}(\log 2 + \log\log n) + o(1).$$

この評価を使うと，

$$u_n^2 = 2\log n + 2x - \log 2 - \log\log n - \log 2\pi + o(1).$$

これを書き直すと，

$$u_n^2 = (2\log n)\left\{1 + \frac{x - \frac{1}{2}\log 4\pi - \frac{1}{2}\log\log n}{\log n} + o((\log n)^{-1})\right\}.$$

$\sqrt{x} = 1 + x/2 + O(x^2)$ $(x\to 0)$ なる評価を使うと，

$$\begin{aligned} u_n &= (2\log n)^{1/2}\left\{1 + \frac{x - \frac{1}{2}\log 4\pi - \frac{1}{2}\log\log n}{2\log n} + o((\log n)^{-1})\right\} \\ &= a_n x + b_n + o(a_n). \end{aligned}$$

従って，

$$P((M_n - b_n)/a_n \le x + o(1)) = P(M_n \le u_n) \to \Lambda(x).$$

□

Remark 17.1. 極値分布は稀にしか起こらない事象の統計解析において現れる (Coles, 2001)．また，未知関数の信頼バンドをカーネル推定量を使って構成するときに，正規化したカーネル推定量の最大値分布を近似する必要があるが，その最大値の極限分布に Gumbel 分布が現れる (Bickel and Rosenblatt, 1973)．極値理論の教科書として，既に紹介した Resnick (1987) の他に，Leadbetter et al. (1983) を挙げておく．この 2 つの本はカバーしている範囲がかなり異なる．Leadbetter et al. (1983) は連続時間定常ガウス過程の最大値に対する極限定理を扱っている．この種の結果は Bickel-Rosenblatt タイプの結果を導くときに重要な役割を果たす．

18　多次元 CLT

弱収束の定義を距離空間上の測度に拡張する．(S,d) を距離空間とし，$\mathcal{S}$ をその Borel σ-field とする．S 上の連続関数は $\mathcal{S}$ 可測である．$(S,\mathcal{S})$ 上の p.m.'s μ_n,μ に対して，

$$\lim_n \int g d\mu_n = \int g d\mu,\ \forall g \in C_b(S) := \{f : f : S \to \mathbb{R},\ f \text{ は有界連続}\}$$

が成り立つとき，μ_n は μ に弱収束すると言う．$\mu_n \xrightarrow{w} \mu$ と書く．S-valued r.v.'s X, X_n に対して，$\mathcal{L}(X_n) \xrightarrow{w} \mathcal{L}(X)$ のとき，X_n は X に分布収束すると言う．$X_n \xrightarrow{d} X$ と書く．

μ_n の弱収束極限が一意なことは，次の補題からわかる．

Lemma 18.1. $(S,\mathcal{S})$ 上の p.m.'s μ,ν が $\int g d\mu = \int g d\nu$ $(\forall g \in C_b(S))$ をみたすなら，$\mu = \nu$.

Proof. $F \subset S$ を閉集合とする．

$$\eta(t) = \begin{cases} 1 & \text{if } t < 0 \\ 1-t & \text{if } 0 \le t \le 1 \\ 0 & \text{if } t > 1 \end{cases}$$

とおいて，$\varepsilon > 0$ に対して，$g_\varepsilon(x) = \eta(\varepsilon^{-1} d(x,F))$ とおくと $(d(x,F) = \inf_{y\in F} d(x,y)$; $F = \varnothing$ のときは，$d(x,\varnothing) = +\infty$ と約束しておく)，$g_\varepsilon \in C_b(S)$ であり，$\varepsilon \downarrow 0$ のとき $g_\varepsilon \downarrow 1_F$ である．よって，DCT より，

$$\mu(F) = \lim_{\varepsilon\downarrow 0} \int g_\varepsilon d\mu = \lim_{\varepsilon\downarrow 0} \int g_\varepsilon d\nu = \nu(F).$$

閉集合全体は π システムであり，$\mathcal{S}$ を生成するので，$\mu = \nu$ を得る． □

以下，$S = \mathbb{R}^k$ の場合を考える．$\mathbb{R}^k$ の標準内積とノルムを

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle := \sum_{j=1}^k x_j y_j,\ |\boldsymbol{x}| := \langle \boldsymbol{x}, \boldsymbol{x} \rangle^{1/2} = \sqrt{\sum_{j=1}^k x_j^2},\ \boldsymbol{x} = (x_1,\dots,x_k), \boldsymbol{y} = (y_1,\dots,y_k) \in \mathbb{R}^k$$

と書く．$\mathbb{R}^k$-valued r.v. $\boldsymbol{X} = (X_1,\dots,X_k)$ の ch.f. $\varphi(\boldsymbol{t})$ を

$$\varphi(\boldsymbol{t}) := E[e^{i\langle \boldsymbol{t}, \boldsymbol{X}\rangle}],\ \boldsymbol{t} = (t_1,\dots,t_k) \in \mathbb{R}^k$$

と定義する．φ を分布 $\mathcal{L}(\boldsymbol{X})$ の ch.f. とも言う．$k = 1$ のときと同様に，$\mathbb{R}^k$ での分布収束は ch.f.'s の各点収束と同値である．

Theorem 18.1. $\boldsymbol{X}_n, \boldsymbol{X}$ を $\mathbb{R}^k$-valued r.v.'s，その ch.f.'s を φ_n, φ とする．このとき，$\boldsymbol{X}_n \xrightarrow{d} \boldsymbol{X} \Leftrightarrow \varphi_n(\boldsymbol{t}) \to \varphi(\boldsymbol{t})$ $(\forall \boldsymbol{t} \in \mathbb{R}^k)$.

$\mathbb{R}^k$ においても Prohorov の定理と反転公式が成り立ち，Theorem 18.1 はその 2 つを経由して証明することができる．ただし，$k=1$ のときの繰り返しになるのと，Prohorov の定理の証明がやや複雑なので (多次元の場合，集合の境界の取り扱いが複雑になる)，ここでは別の証明を紹介する．

関数 $f \in L^1(\mathbb{R}^k, \mathcal{B}^k, \lambda^k) = L^1(\lambda^k)$ の 逆 Fourier 変換 を

$$\check{f}(\boldsymbol{t}) := \int e^{i\langle \boldsymbol{t}, \boldsymbol{x}\rangle} f(\boldsymbol{x}) d\boldsymbol{x},\ \boldsymbol{t} \in \mathbb{R}^k$$

と定義する．$\check{f}$ は有界連続である．$f, g \in L^1(\lambda^k)$ に対して，$|\check{f}|, |\check{g}| \in L^1(\lambda^k)$ なら，

$$\int f(\boldsymbol{x})\check{g}(\boldsymbol{x})d\boldsymbol{x} = \int g(\boldsymbol{y}) \left\{ \int e^{i\langle \boldsymbol{x}, \boldsymbol{y}\rangle} f(\boldsymbol{x}) d\boldsymbol{x} \right\} d\boldsymbol{y} = \int g(\boldsymbol{y})\check{f}(\boldsymbol{y}) d\boldsymbol{y} \qquad (*)$$

となる．積分順序の交換は Fubini の定理から保証される．

次に，$C(\mathbb{R}^k) := \{f : f : \mathbb{R}^k \to \mathbb{R},\ f \text{ は連続 }\}$ とし，$f \in C(\mathbb{R}^k)$ のサポートを

$$\mathrm{supp}(f) := \overline{\{\boldsymbol{x} \in \mathbb{R}^k : f(\boldsymbol{x}) \neq 0\}}$$

と定義する．コンパクトサポートをもつ $\mathbb{R}^k$ 上の連続関数全体を $C_c(\mathbb{R}^k)$ とおく：

$$C_c(\mathbb{R}^k) := \{f \in C(\mathbb{R}^k) : \mathrm{supp}(f) \text{ はコンパクト }\}.$$

また，$C_c^\infty(\mathbb{R}^k) := C^\infty(\mathbb{R}^k) \cap C_c(\mathbb{R}^k)$ とおく．ここで，$f \in C_c^\infty(\mathbb{R}^k)$ なら，部分積分を繰り返して，任意の正整数 m に対して，$|\check{f}(\boldsymbol{t})| = O((1+|\boldsymbol{t}|)^{-m})$ $(|\boldsymbol{t}| \to \infty)$ となることが示せる (演習問題)．よって，$f \in C_c^\infty(\mathbb{R}^k)$ なら $|\check{f}| \in L^1(\lambda^k)$ である．

Theorem 18.1 の証明に移る．分布収束 $\boldsymbol{X}_n \xrightarrow{d} \boldsymbol{X}$ の定義はすべての $g \in C_b(\mathbb{R}^k)$ に対して，期待値の収束 $\lim_n E[g(\boldsymbol{X}_n)] = E[g(\boldsymbol{X})]$ が成り立つことだったが，すべての $g \in C_c^\infty(\mathbb{R}^k)$ に対して期待値が収束していれば十分である．

Lemma 18.2. $\boldsymbol{X}_n \xrightarrow{d} \boldsymbol{X} \Leftrightarrow \lim_n E[g(\boldsymbol{X}_n)] = E[g(\boldsymbol{X})]$ $(\forall g \in C_c^\infty(\mathbb{R}^k))$.

Proof. $\Rightarrow$ は明らか．$\Leftarrow$ を示す．まず，$\forall g \in C_c(\mathbb{R}^k), \lim_n E[g(\boldsymbol{X}_n)] = E[g(\boldsymbol{X})]$ を示す．このとき，$\forall \varepsilon > 0,\ \exists g_\varepsilon \in C_c^\infty(\mathbb{R}^k)$ s.t. $\|g_\varepsilon - g\|_u \leq \varepsilon$ だから，

$$|E[g(\boldsymbol{X}_n)] - E[g(\boldsymbol{X})]| \leq 2\varepsilon + |E[g_\varepsilon(\boldsymbol{X}_n)] - E[g_\varepsilon(\boldsymbol{X})]|.$$

$n \to \infty$ とすると，右辺第 2 項 $\to 0$ だから，$\limsup_n |E[g(\boldsymbol{X}_n)] - E[g(\boldsymbol{X})]| \leq 2\varepsilon$.

次に，任意の $\varepsilon > 0$ に対して，$M > 1$ を十分大きくとって，$P(\boldsymbol{X} \in [-M+1, M-1]^k) \geq 1-\varepsilon$ とする．$\eta : \mathbb{R}^k \to [0,1]$ を

$$\eta(\boldsymbol{x}) = \begin{cases} 1 & \text{if } \boldsymbol{x} \in [-M+1, M-1]^k \\ 0 & \text{if } \boldsymbol{x} \notin [-M, M]^k \end{cases}$$

なる連続関数とすると，

$$\lim_n E[\eta(\boldsymbol{X}_n)] = E[\eta(\boldsymbol{X})] \geq P(\boldsymbol{X} \in [-M+1, M-1]^k) \geq 1-\varepsilon.$$

$g \in C_b(\mathbb{R}^k)$ に対して，

$$\begin{aligned}|E[g(\boldsymbol{X}_n)] - E[g(\boldsymbol{X})]| &\leq |E[g(\boldsymbol{X}_n)\eta(\boldsymbol{X}_n)] - E[g(\boldsymbol{X}_n)\eta(\boldsymbol{X})]| \\ &\quad + \|g\|_u \underbrace{\left(E[1-\eta(\boldsymbol{X}_n)] + E[1-\eta(\boldsymbol{X})]\right)}_{\leq 2\varepsilon + o(1)}.\end{aligned}$$

$g\eta \in C_c(\mathbb{R}^k)$ だから，$n \to \infty$ として，$\limsup_n |E[g(\boldsymbol{X}_n)] - E[g(\boldsymbol{X})]| \leq 2\|g\|_u\varepsilon$. □

$g \in C_c^\infty(\mathbb{R}^k)$ に対して期待値の収束を示せばよいことが分かった．そこで，期待値 $E[g(\boldsymbol{X})]$ を $\boldsymbol{X}$ の ch.f. を用いて表現することを考える．

Lemma 18.3. $\boldsymbol{X}$ を $\mathbb{R}^k$-valued r.v. とし，その ch.f. を φ とする．このとき，任意の $g \in C_c^\infty(\mathbb{R}^k)$ に対して，

$$E[g(\boldsymbol{X})] = \frac{1}{(2\pi)^k}\int \check{g}(\boldsymbol{t})\overline{\varphi(\boldsymbol{t})}d\boldsymbol{t}.$$

Proof. $N(\boldsymbol{0}, I_k)$ の密度関数を $\phi(\boldsymbol{z}) = (2\pi)^{-k/2}e^{-|\boldsymbol{z}|^2/2}$ とおくと，$\boldsymbol{Z} \sim N(\boldsymbol{0}, I_k)$ に対して，

$$\int e^{i\langle \boldsymbol{t}, \boldsymbol{z}\rangle}\phi(\boldsymbol{z})d\boldsymbol{z} = E[e^{i\langle \boldsymbol{t}, \boldsymbol{Z}\rangle}] = \prod_{j=1}^k E[e^{it_jZ_j}] = e^{-|\boldsymbol{t}|^2/2}$$

となる．従って，$\boldsymbol{x} \in \mathbb{R}^k$ を任意に固定し，$\sigma > 0$ に対して，$\phi_\sigma(\boldsymbol{z}) = \sigma^{-k}\phi(\boldsymbol{z}/\sigma)$ とおくと，$N(\boldsymbol{x}, \sigma I_k)$ の密度関数は，

$$\begin{aligned}\phi_\sigma(\boldsymbol{y}-\boldsymbol{x}) &= \frac{1}{(2\pi\sigma)^k}\int \exp\{i\langle(\boldsymbol{y}-\boldsymbol{x})/\sigma, \boldsymbol{z}\rangle - |\boldsymbol{z}|^2/2\}d\boldsymbol{z} \\ &= \frac{1}{(2\pi)^k}\int e^{i\langle \boldsymbol{y}, \boldsymbol{z}\rangle}\underbrace{\{e^{-i\langle \boldsymbol{x}, \boldsymbol{z}\rangle - \sigma^2|\boldsymbol{z}|^2/2}\}}_{=h(\boldsymbol{z})}d\boldsymbol{z} = \frac{\check{h}(\boldsymbol{y})}{(2\pi)^k}\end{aligned}$$

と表すことができる．この表現と $(*)$ より，

$$\begin{aligned}(2\pi)^k E[g(\boldsymbol{x}+\sigma\boldsymbol{Z})] &= (2\pi)^k\int g(\boldsymbol{y})\phi_\sigma(\boldsymbol{y}-\boldsymbol{x})d\boldsymbol{y} = \int g(\boldsymbol{y})\check{h}(\boldsymbol{y})d\boldsymbol{y} \\ &= \int \check{g}(\boldsymbol{t})h(\boldsymbol{t})d\boldsymbol{t} = \int e^{-\sigma^2|\boldsymbol{t}|^2/2}\{\check{g}(\boldsymbol{t})e^{-i\langle \boldsymbol{t}, \boldsymbol{x}\rangle}\}d\boldsymbol{t}\end{aligned}$$

を得る．ここで，$\sigma \to 0$ とすると，DCT より，反転公式

$$g(\boldsymbol{x}) = \frac{1}{(2\pi)^k}\int \check{g}(\boldsymbol{t})e^{-i\langle \boldsymbol{t}, \boldsymbol{x}\rangle}d\boldsymbol{t}$$

を得る．補題の結論は $\boldsymbol{x}$ に $\boldsymbol{X}$ を代入して期待値をとることにより従う． □

Theorem 18.1 の証明は，上の 2 つの補題から明らかである．Theorem 18.1 から次の 2 つの系が直ちに従う．

Corollary 18.1. $\boldsymbol{X}, \boldsymbol{Y}$ を $\mathbb{R}^k$-valued r.v.'s とし，その ch.f.'s を φ, ψ とする．このとき，$\boldsymbol{X} \stackrel{d}{=} \boldsymbol{Y} \Leftrightarrow \varphi(\boldsymbol{t}) = \psi(\boldsymbol{t})$ $(\forall \boldsymbol{t} \in \mathbb{R}^k)$.

Corollary 18.2 (Cramér-Wold device)**.** $\boldsymbol{X}_n \stackrel{d}{\to} \boldsymbol{X} \Leftrightarrow \langle \boldsymbol{t}, \boldsymbol{X}_n \rangle \stackrel{d}{\to} \langle \boldsymbol{t}, \boldsymbol{X} \rangle$ $(\forall \boldsymbol{t} \in \mathbb{R}^k)$.

Cramér-Wold device と 1 次元の場合の CLT より，多次元 CLT を得る．

Theorem 18.2. $\{\boldsymbol{X}_m\}$ を i.i.d. $\mathbb{R}^k$-valued r.v.'s であって，$E[\boldsymbol{X}_1] = \boldsymbol{0}$, $(E[X_{1j}X_{1\ell}])_{1 \le j,\ell \le k} = (\sigma_{j,k})_{1 \le j,\ell \le k} = \Sigma$ とする．このとき，$\sqrt{n} \sum_{m=1}^n \boldsymbol{X}_m \stackrel{d}{\to} N(\boldsymbol{0}, \Sigma)$.

Remark 18.1. 多次元の CLT を証明するだけなら，Lindeberg の CLT の証明を修正するだけで十分であり，それは多変数の微積分の範囲でできる．しかし，多次元の分布収束の問題は 1 次元の分布収束の問題に落とすことができるという事実 (Cramér-Wold device) をちゃんとカバーしたかったので，本節をこのような構成にした．

演習問題

Exercise 18.1. $(X, Y), (X_n, Y_n), n = 1, 2, \ldots$ を $\mathbb{R}^2$-valued r.v.'s とする．このとき，$(X_n, Y_n) \stackrel{d}{\to} (X, Y)$ となるためには，あらゆる有界連続関数 $g, h : \mathbb{R} \to \mathbb{R}$ に対して，$\lim_n E[g(X_n)h(Y_n)] = E[g(X)h(Y)]$ となることが必要十分であることを示せ．

Exercise 18.2. $(X_n, Y_n), n = 1, 2, \ldots$ を $\mathbb{R}^2$-valued r.v.'s とし，$(X_n, Y_n) \stackrel{d}{\to} (X, Y)$ とする．さらに，Y_n の分布は n に依存しないとする．このとき，あらゆる有界可測関数 $g : \mathbb{R} \to \mathbb{R}$ に対して，$(X_n, g(Y_n)) \stackrel{d}{\to} (X, g(Y))$ となることを示せ．

Exercise 18.3. $f \in C_c^\infty(\mathbb{R}^k)$ なら，任意の正整数 m に対して，$|\check{f}(\boldsymbol{t})| = O((1+|\boldsymbol{t}|)^{-m})$ $(|\boldsymbol{t}| \to \infty)$ となることを示せ．

Exercise 18.4. $\boldsymbol{X} = (X_1, \ldots, X_k), \boldsymbol{Y} = (Y_1, \ldots, Y_k)$ を $\mathbb{R}^k$-valued r.v.'s とし，ある $M > 0$ が存在して，$P(\boldsymbol{X} \in [-M, M]^k) = P(\boldsymbol{Y} \in [-M, M]^k) = 1$ とする．いま，任意の非負整数 $\alpha_1, \ldots, \alpha_k$ に対して，$E[\prod_{j=1}^k X_j^{\alpha_j}] = E[\prod_{j=1}^k Y_j^{\alpha_j}]$ であれば，$\boldsymbol{X} \stackrel{d}{=} \boldsymbol{Y}$ となることを示せ．

ヒント：Theorem 14.3 を用いて，$\forall \boldsymbol{t} \in \mathbb{R}^k$ に対して $\langle \boldsymbol{t}, \boldsymbol{X} \rangle \stackrel{d}{=} \langle \boldsymbol{t}, \boldsymbol{Y} \rangle$ を示す．または Stone-Weierstrass の定理を使う．

Exercise 18.5. $(S, \mathcal{S})$ 上の p.m.'s $\{\mu_n\}$ が一様にタイトであるとは，任意の $\varepsilon > 0$ に対して，$\mu_n(K) > 1 - \varepsilon$ $(\forall n)$ なるコンパクト集合 K が存在することを言う．$\boldsymbol{X}_n = (X_{n,1}, \ldots, X_{n,k})$ を $\mathbb{R}^k$-valued r.v.'s とすると，$\{\mathcal{L}(\boldsymbol{X}_n) : n \in \mathbb{N}\}$ が一様にタイトなことと，各 j に対して $\{\mathcal{L}(X_{n,j}) : n \in \mathbb{N}\}$ が一様にタイトなことは同値であることを示せ．

Exercise 18.6. $f \in C_c(\mathbb{R}^k)$ とする．$\eta \in C_c^\infty(\mathbb{R}^k)$ を $\eta \geq 0,\ \int \eta(\boldsymbol{x})d\boldsymbol{x} = 1$ をみたす関数とする．$\eta_\varepsilon(\boldsymbol{x}) = \varepsilon^{-k}\eta(\varepsilon^{-1}\boldsymbol{x}),\ \varepsilon > 0,\ f_\varepsilon(\boldsymbol{x}) = \int \eta_\varepsilon(\boldsymbol{x} - \boldsymbol{y})f(\boldsymbol{y})d\boldsymbol{y}$ と定義する．(a) η の例を挙げよ．(b) $f_\varepsilon \in C_c^\infty(\mathbb{R}^k)$ を示せ． (c) $\|f_\varepsilon - f\|_u \to 0\ (\varepsilon \downarrow 0)$ を示せ．

Part IV. マルチンゲール

本パートは条件付き期待値を定義したあと，離散時間のマルチンゲールを考察する．マルチンゲールの重要性は強調してもしすぎることはないが，例えば，確率積分や確率微分方程式を包括する確率解析はマルチンゲールに立脚している．従って，将来的にそういった分野を勉強することを考えているならば，マルチンゲールの理解は不可欠である．また，マルチンゲールは確率不等式を導く強力なテクニックを与える．本パートに関しては，Williams (1991) がよい副読本となると思う．また，Neveu (1975) は離散時間のマルチンゲールに特化した教科書であり，多くの応用例をカバーしている．

19 条件付き期待値

$(\Omega, \mathcal{F}, P)$ を確率空間とし，$\mathcal{G}$ を $\mathcal{F}$ の sub σ-field とする．$X \in L^1$ に対して，

$$E[X, A] := \int_A X dP, \ A \in \mathcal{F}$$

と定義する．r.v. Y が $\mathcal{G}$ 可測，i.e., $Y^{-1}(B) \in \mathcal{G}$ $(\forall B \in \mathcal{B})$ をみたすとき，$Y \in \mathcal{G}$ と書く．

$X \in L^1$ に対して，$\mathcal{G}$ を与えたときの X の 条件付き期待値 (conditional expectation) $E[X \mid \mathcal{G}]$ を，

(i) $Y \in \mathcal{G}$,

(ii) $Y \in L^1$,

(iii) $\forall A \in \mathcal{G}, \ E[X, A] = E[Y, A]$

をみたす任意の r.v. Y と定義する．$E[X \mid \mathcal{G}] := Y$. (i)–(iii) をみたす任意の Y を $E[X \mid \mathcal{G}]$ の バージョン (version) と呼ぶ．まず，$E[X \mid \mathcal{G}]$ の一意性を確認する．

Lemma 19.1 (一意性)**.** (i)–(iii) をみたす Y は a.s. に一意である．

Proof. Y' も (i)–(iii) をみたすなら，

$$E[Y, A] = E[Y', A], \ \forall A \in \mathcal{G}.$$

$A = \{Y - Y' \geq \varepsilon\} \in \mathcal{G}$ とおくと，

$$0 = E[X - X, A] = E[Y - Y', A] \geq \varepsilon P(A)$$

であるから，$P(A) = 0$. ε は任意なので，$Y \leq Y'$ a.s. Y と Y' の役割を入れ替えて，$Y = Y'$ a.s. を得る． □

$E[X \mid \mathcal{G}]$ は a.s. にしか一意に決まらないので，$Y = E[X \mid \mathcal{G}]$ a.s. などと書くべきであるが，記号が煩雑になるので，多くの場合，a.s. は省略する．

$E[X \mid \mathcal{G}]$ の存在を示すために，Random-Nikodym の定理を用いる．可測空間 $(S, \mathcal{S})$ 上の測度 μ, ν に対して，ν が μ に対して 絶対連続 (absolutely continuous) であるとは，$\mu(A) = 0, A \in \mathcal{S}$ なら $\nu(A) = 0$ となることを言う．$\nu \ll \mu$ と書く．

Theorem 19.1 (Radon-Nikodym)**.** ν を有限測度，μ を σ-finite な測度とし，$\nu \ll \mu$ とする．このとき，ある $f \in L^1(S, \mathcal{S}, \mu)$ が存在して，

$$\nu(A) = \int_A f d\mu, \ \forall A \in \mathcal{S}$$

と表せる．さらに，そのような f は μ-a.e. に一意である．$f = d\nu/d\mu$ と書く．

Remark 19.1. $f = d\nu/d\mu$ のことを Radon-Nikodym 微分，または ν の μ に関する 密度 (density) と呼ぶ．逆に，Radon-Nikodym 微分が存在すれば，$\nu \ll \mu$ となるのは明らかである．以前，d.f. F が絶対連続であることを，密度の存在により定義したが，これは d.f. に対応する分布が Lebesgue 測度に関して絶対連続であることと同値である．

ここでは，Hilbert 空間における Riesz の表現定理を用いた証明を紹介する．別の方針の証明は，Durrett (2010, Appendix) を参照せよ．まず，Riesz の定理を述べる．完備な内積空間のことを Hilbert 空間と呼ぶのであった．

Theorem 19.2 (Riesz の表現定理)**.** F を実 Hilbert 空間 $(H, \langle \cdot, \cdot \rangle)$ 上の連続線形汎関数とすると，一意な $h \in H$ が存在して，$F(x) = \langle x, h \rangle$ $(\forall x \in H)$ と表せる．

念のために Riesz の表現定理の証明を載せておく．Riesz の表現定理は複素 Hilbert 空間でも成り立つが，ここでは簡単のために実 Hilbert 空間のみ扱う．$(H, \langle \cdot, \cdot \rangle)$ を実 Hilbert 空間とし，$\| \cdot \|$ をそのノルムとする：$\|x\| = \langle x, x \rangle^{1/2}, \ x \in H$. H に対しては，中線定理 (parallelogram law)

$$\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2, \ x, y \in H$$

が成り立つ (ノルムを内積で展開すればよい)．まず次の射影定理を証明する．

Theorem 19.3 (射影定理)**.** V を H の閉部分空間とし，$x \in H$ を given とする．

(i) $V \ni z \mapsto \|x - z\|$ を最小化する一意なベクトル $y \in V$ が存在する：$\|x - y\| = \min_{z \in V} \|x - z\|$. y を x の V への 正射影 (orthogonal projection) と呼ぶ．

(ii) $y \in V$ が x の V への正射影であるためには，$\langle z, x - y \rangle = 0$ $(\forall z \in V)$ となることが必要十分である．

Proof. (i). 最小化解の存在を示す．$D = \inf_{z\in V} \|x-z\|$ とおくと，$0 \le D < \infty$ である．ここで，$\|x-y_n\| \to D$ となる列 $y_n \in V$ をとると，$(y_m+y_n)/2 \in V$ より，$\|x-(y_m+y_n)/2\| \ge D$ である．このとき，中線定理より，

$$\underbrace{\left\|x - \frac{y_m+y_n}{2}\right\|^2}_{\ge D^2} + \frac{1}{4}\|y_m - y_n\|^2 = \frac{1}{2}(\|x-y_n\|^2 + \|x-y_m\|^2)$$

となるが，$m,n \to \infty$ のとき，右辺 $\to D^2$ となるから，$\|y_m - y_n\| \to 0$ となる．ゆえに $\{y_n\}$ は Cauchy 列なので，H の完備性と V が閉であることから，ある $y \in V$ が存在して，$y_n \to y$ となる．写像 $z \mapsto \|x-z\|$ は連続なので，$\|x-y\| = D$ を得る．

次に一意性を示す．y' をもう 1 つの最小化解とし，$y'' = (y+y')/2 \in V$ とおくと，中線定理より，

$$\|x-y''\|^2 + \frac{1}{4}\|y-y'\|^2 = \frac{1}{2}(\|x-y\|^2 + \|x-y'\|^2) = D^2$$

となるが，左辺は右辺より厳密に小さくはならないので，$y = y'$ でなくてならない．

(ii). y を x の V への正射影とする．$z \in V$ を任意にとると，$\theta \in (0,1)$ に対して，$\|x - \{\theta z + (1-\theta)y\}\| \ge \|x-y\|$ であり，両辺の平方をとって整理して，

$$-2\langle z-y, x-y\rangle + \theta\|z-y\|^2 \ge 0$$

を得る．$\theta \to 0$ として，$\langle z-y, x-y\rangle \le 0$ となるが，$z \in V$ は任意であり V は部分空間であるから，$\langle z, x-y\rangle = 0$ $(\forall z \in V)$ を得る．

逆に，$y \in V$ が $\langle z, x-y\rangle = 0$ $(\forall z \in V)$ をみたすなら，あらゆる $z \in V$ に対して，$\|x-z\|^2 = \|x-y+y-z\|^2 = \|x-y\|^2 + \|y-z\|^2 \ge \|x-y\|^2$ となるから，y は x の V への正射影である． □

V を H の閉部分空間，$x \in H$ を given とし，y を x の V への正射影とすると，

$$x = y + (x-y), \quad y \in V,\ x-y \in V^\perp := \{w : \langle z, w\rangle = 0\ \forall z \in V\}$$

なる一意な分解が成り立つ．また，$\Pi_V x$ を x の V への正射影とすると，Π_V は線形であり，$x \in V$ に対して，$\Pi_V x = x$ であるから，Π_V は V への全射である．Π_V を V の上への正射影作用素 (orthogonal projection operator) と呼ぶ．

Proof of Theorem 19.2. 一意性の証明は難しくないので省略する．存在を示す．$F \ne 0$ と仮定してよい．このとき，$V = \{x \in H : F(x) = 0\}$ とおくと，F は連続線形汎関数であったから，V は H の閉部分空間であり，さらに $F \ne 0$ より，$V \ne H$ である． そこで $y \in H \setminus V$ を 1 つとり，y の V への正射影を z とおくと，$z \ne y$ であり，$w = (y-z)/\|y-z\|$ とおくと，$F(w) \ne 0$ である．このとき，任意の $x \in H$ に対して，$x - F(x)w/F(w) \in V$ より，w との内積をとって，$0 = \langle x, w\rangle - F(x)/F(w)$, i.e., $F(x) = \langle x, F(w)w\rangle$ を得る．従って，$h = F(w)w$ とすればよい． □

以上で Radon-Nikodym の定理を証明する準備が整った．

Proof of Theorem 19.1. 存在のみ示す．まず，μ も有限測度と仮定する．$\rho = \mu + \nu$ とおき，ρ-a.e. に等しい関数を同一視すれば，$H = L^2(S, \mathcal{S}, \rho)$ は内積 $\langle f, g \rangle = \int f g d\rho$ に関して実 Hilbert 空間になる．ρ は有限であるから，$L^2(\rho) \subset L^1(\rho) \subset L^1(\nu)$ である (最後の包含関係は $\nu \leq \rho$ から従う)．また，$h_n \to h$ in $L^2(\rho)$ なら，$h_n \to h$ in $L^1(\rho)$ であり，$\nu \leq \rho$ より，$h_n \to h$ in $L^1(\nu)$ でもあるから，

$$F : h \mapsto \int h d\nu,\ H \to \mathbb{R}$$

は連続線形汎関数である．従って，Riesz の定理より，$\exists g \in H$ s.t.

$$\int h d\nu = \int g h d\rho = \int g h d(\mu + \nu),\ \forall h \in H$$

となる．積分を入れ替えて，

$$\int (1 - g) h d\nu = \int g h d\mu,\ \forall h \in H \qquad (*)$$

を得る．$0 \leq g < 1$ μ-a.e. を示す．$h = 1_{\{g<0\}}$ に対して $(*)$ を適用して，$0 \leq \int_{g<0} (1 - g) d\nu = \int_{g<0} g d\mu$ より，$\mu(\{g < 0\}) = 0$ を得る．同様に，$\mu(\{g \geq 1\}) = 0$ であるから，$B = \{0 \leq g < 1\}$ とおくと，$\mu(B^c) = 0$ であり，さらに $\nu \ll \mu$ より，$\nu(B^c) = 0$ を得る．$(*)$ は $h = 1_A, A \in \mathcal{S}$ に対して成り立つので，極限操作により任意の非負可測関数 h に対しても成り立つ．

いま，

$$f(x) = \begin{cases} \frac{g(x)}{1-g(x)} & \text{if } x \in B \\ 0 & \text{if } x \in B^c \end{cases}$$

とおいて，$A \in \mathcal{S}$ に対して，

$$h(x) = \begin{cases} \frac{1_A(x)}{1-g(x)} & \text{if } x \in B \\ 0 & \text{if } x \in B^c \end{cases}$$

とおくと，

$$\nu(A) = \nu(A \cap B) = \int (1 - g) h d\nu = \int g h d\mu = \int_{A \cap B} f d\mu = \int_A f d\mu$$

を得る．$A \in \mathcal{S}$ は任意より，μ が有限測度の場合に存在が示された．

μ が一般に σ-finite なとき，$\bigcup_n E_n = S, E_n \in \mathcal{S}, \mu(E_n) < \infty\ \forall n$ なる排反な集合列 $\{E_n\}$ が存在する．$\mu_n(A) := \mu(A \cap E_n), \nu_n(A) := \nu(A \cap E_n), A \in \mathcal{S}$ とおけば，各 μ_n は有限測

度であり，さらに $\nu_n \ll \mu_n$ より，$f_n := d\nu_n/d\mu_n$ が存在する．そこで，$f := \sum_n f_n 1_{E_n}$ とおけば，各 $A \in \mathcal{S}$ に対して，MCT より，

$$\nu(A) = \sum_n \nu(A \cap E_n) = \sum_n \nu_n(A) = \sum_n \int_A f_n d\mu_n = \sum_n \int_A f_n 1_{E_n} d\mu = \int_A f d\mu.$$

以上より，定理が示された． □

Remark 19.2 (Lebesgue 分解)**.** $(S, \mathcal{S})$ 上の測度 μ, ν に対して，ある集合 $A \in \mathcal{S}$ が存在して，$\mu(A) = \nu(A^c) = 0$ が成り立つとき，μ と ν は特異 (singular) であると言う．このとき，$\mu \perp \nu$ と書く．Radon-Nikodym の定理の証明を少し修正するだけで，次の Lebesgue 分解が示せる．すなわち，μ と ν を $(S, \mathcal{S})$ 上の σ-finite な測度とすると，次をみたす $(S, \mathcal{S})$ 上の一意な測度 ν_{ac}, ν_s が存在する：

$$\nu = \nu_{ac} + \nu_s,\ \nu_{ac} \ll \mu,\ \nu_s \perp \mu.$$

この分解を Lebesgue 分解と呼び，ν_{ac}, ν_s をそれぞれ ν の μ に関する絶対連続部分，特異部分と呼ぶ．簡単のために，μ と ν が有限測度のときに Lebesgue 分解を証明してみると，上の証明において，$(*)$ までは $\nu \ll \mu$ でなくても正しくて，また，$0 \le g < 1$ μ-a.e. と $0 \le g \le 1$ ν-a.e. が示せる．そこで，$A = \{g = 1\}$ とおいて，ν_{ac}, ν_s を

$$\nu_{ac}(E) = \nu(E \cap A^c),\ \nu_s(E) = \nu(E \cap A),\ E \in \mathcal{S}$$

と定義すれば，ν_{ac}, ν_s は Lebesgue 分解を与える．μ, ν が一般に σ-finite である場合と，分解の一意性は演習問題とする．

$E[X \mid \mathcal{G}]$ の存在を示す．最初に $X \ge 0$ の場合を考える．

$$Q(A) = E[X, A] = \int_A X dP,\ A \in \mathcal{G}$$

とおくと，Q は $(\Omega, \mathcal{G})$ 上の有限測度で，$Q \ll P|_{\mathcal{G}}$ である．従って，Radon-Nikodym の定理より，$\exists Y \in L^1(\Omega, \mathcal{G}, P)$ s.t.

$$E[X, A] = Q(A) = \int_A Y dP,\ \forall A \in \mathcal{G}$$

であるから，Y は $E[X \mid \mathcal{G}]$ のバージョンである．

一般の場合は，$X = X^+ - X^-$ と分解して，$Y_1 = E[X^+ \mid \mathcal{G}], Y_2 = E[X^- \mid \mathcal{G}]$ とおくと，$Y_1 - Y_2 \in \mathcal{G}$ で，$\forall A \in \mathcal{G}$ に対して，

$$E[X, A] = E[X^+, A] - E[X^-, A] = E[Y_1, A] - E[Y_2, A] = E[Y_1 - Y_2, A].$$

従って，$Y_1 - Y_2$ は $E[X \mid \mathcal{G}]$ のバージョンである．これで $E[X \mid \mathcal{G}]$ の存在が示せた．

Lemma 19.2. (a) $X \in \mathcal{G}$ なら，$E[X \mid \mathcal{G}] = X$. (b) X と $\mathcal{G}$ が独立なら，$E[X \mid \mathcal{G}] = E[X]$.

Proof. (a). 明らか．(b). $A \in \mathcal{G}$, $E[X, A] = E[X 1_A] = E[X]P(A) = E[E[X], A]$. □

Example 19.1. $\mathcal{G} = \{\varnothing, \Omega\}$ なら，X と $\mathcal{G}$ は独立なので，$E[X \mid \mathcal{G}] = E[X]$ である．

次に，$A \in \mathcal{F}$ に対して，$\mathcal{G}$ を与えたときの A の 条件付き確率 (conditional probability) を

$$P(A \mid \mathcal{G}) := E[1_A \mid \mathcal{G}]$$

と定義する．また，$P(B) > 0$ なる $B \in \mathcal{F}$ が与えられたとき，A の条件付き確率を

$$P(A \mid B) := \frac{P(A \cap B)}{P(B)}$$

と定義する．

$(S, \mathcal{S})$ を可測空間とし，X を S-valued r.v. とする．このとき，($\mathbb{R}$-valued) r.v. $Y \in L^1$ に対して，X を与えたときの Y の条件付き期待値を

$$E[Y \mid X] := E[Y \mid \sigma(X)]$$

と定義する．$E[Y \mid X]$ は X の関数で書ける．

Theorem 19.4. $(S, \mathcal{S})$ を可測空間とし，X を S-valued r.v. とする．このとき，r.v. $T \in \sigma(X)$ に対して，可測関数 $g : S \to \mathbb{R}$ が存在して，$T = g(X)$ と書ける．

Proof. T は $\sigma(X)$ 可測な単関数の pointwise な極限で書ける．一方，$A \in \sigma(X)$ に対して，定義より，$\exists B \in \mathcal{S}$ s.t. $A = X^{-1}(B)$ であるから，$1_A(\omega) = 1_B(X(\omega))$. よって，単関数 $g_n : S \to \mathbb{R}$ が存在して，$T(\omega) = \lim_n g_n(X(\omega))$ $(\forall \omega \in \Omega)$ と表せる．$E = \{$ 有限な $\lim_n g_n$ が存在する $\}$ とおくと，$\{X \in E\} = \Omega$ であり，

$$g(x) = \begin{cases} \lim_n g_n(x) & \text{if } x \in E \\ 0 & \text{if } x \in E^c \end{cases}$$

とおくと，$T(\omega) = g(X(\omega))$ $(\forall \omega \in \Omega)$. □

この定理より，ある可測関数 $g : S \to \mathbb{R}$ が存在して，$E[Y \mid X] = g(X)$ と表せることがわかる．このとき，

$$E[Y \mid X = x] := g(x)$$

と定義する．$E[Y \mid X = x]$ は $P \circ X^{-1}$-a.s. に一意である．X, Y が独立な場合は次の補題が成り立つ．

Lemma 19.3. $(S', \mathcal{S}')$ をもう 1 つの可測空間とする．X を S-valued r.v., Y を S'-valued r.v. とし，両者は独立とする．また，$\varphi : S \times S' \to \mathbb{R}$ を $\mathcal{S} \times \mathcal{S}'$ 可測な関数とし，$E[|\varphi(X, Y)|] < \infty$ を仮定する．このとき，$g(x) = E[\varphi(x, Y)]$ とおくと，$E[\varphi(X, Y) \mid X] = g(X)$.

Proof. $g(X) \in \sigma(X)$ はよい．一方，$A \in \sigma(X)$ に対して，定義より，$\exists B \in \mathcal{S}$ s.t. $A = \{X \in B\}$ と書ける．$\mathcal{L}(X) = \mu, \mathcal{L}(Y) = \nu$ とおくと，X, Y の独立性より，$\mathcal{L}(X, Y) = \mu \times \nu$ であるから，

$$\begin{aligned} E[\varphi(X,Y), A] &= E[\varphi(X,Y)1_B(X)] = \iint \varphi(x,y)1_B(x)d\mu(x)d\nu(y) \\ &= \int \underbrace{\left\{ \int \varphi(x,y)d\nu(y) \right\}}_{=g(x)} 1_B(x)d\mu(x) \quad (\because \text{Fubini の定理}) \\ &= \int g(x)1_B(x)d\mu(x) = E[g(X), A]. \end{aligned}$$

□

条件付き期待値の性質.

Theorem 19.5.

(a) $a, b \in \mathbb{R}, X, Y \in L^1 \Rightarrow E[aX + bY \mid \mathcal{G}] = aE[X \mid \mathcal{G}] + bE[Y \mid \mathcal{G}]$.

(b) $X, Y \in L^1, X \leq Y \Rightarrow E[X \mid \mathcal{G}] \leq E[Y \mid \mathcal{G}]$.

(c) $0 \leq X_n \uparrow X,\ X \in L^1 \Rightarrow E[X_n \mid \mathcal{G}] \uparrow E[X \mid \mathcal{G}]$.

(d) $X_n \downarrow X,\ X_1, X \in L^1 \Rightarrow E[X_n \mid \mathcal{G}] \downarrow E[X \mid \mathcal{G}]$.

(e) $\varphi : \mathbb{R} \to \mathbb{R}$ が凸関数であり，$X, \varphi(X) \in L^1$ なら，$\varphi(E[X \mid \mathcal{G}]) \leq E[\varphi(X) \mid \mathcal{G}]$.

(f) $p \geq 1$ に対して，$X \in L^p$ なら，$E[|E[X \mid \mathcal{G}]|^p] \leq E[|X|^p]$.

(g) $X \in L^1,\ \mathcal{G}_1 \subset \mathcal{G}_2,\ E[X \mid \mathcal{G}_2] \in \mathcal{G}_1 \Rightarrow E[X \mid \mathcal{G}_2] = E[X \mid \mathcal{G}_1]$.

(h) $X \in L^1,\ \mathcal{G}_1 \subset \mathcal{G}_2 \Rightarrow E[E[X \mid \mathcal{G}_1] \mid \mathcal{G}_2] = E[X \mid \mathcal{G}_1],\ E[E[X \mid \mathcal{G}_2] \mid \mathcal{G}_1] = E[X \mid \mathcal{G}_1]$.

(i) $X \in \mathcal{G},\ XY \in L^1, Y \in L^1 \Rightarrow E[XY \mid \mathcal{G}] = XE[Y \mid \mathcal{G}]$.

Proof. (a). 右辺が左辺のバージョンを与えることを示せばよい．右辺が $\mathcal{G}$ 可測なことはよい．次に，$A \in \mathcal{G}$ に対して，

$$\begin{aligned} \int_A \{aE[X \mid \mathcal{G}] + bE[Y \mid \mathcal{G}]\}dP &= a\int_A E[X \mid \mathcal{G}]dP + b\int_A E[Y \mid \mathcal{G}]dP \\ &= a\int_A XdP + b\int_A YdP = \int_A (aX + bY)dP. \end{aligned}$$

(b). 定義より，

$$\int_A E[X \mid \mathcal{G}]dP = \int_A XdP \leq \int_A YdP = \int_A E[Y \mid \mathcal{G}]dP.$$

$m \in \mathbb{N}$ に対して，$A_m = \{E[X \mid \mathcal{G}] - E[Y \mid \mathcal{G}] \geq 1/m\}$ とおくと，左辺 $\geq \int_{A_m} E[Y \mid \mathcal{G}]dP + (1/m)P(A_m)$ であるから，$P(A_m) = 0$. よって，$P(\bigcup_m A_m) \leq \sum_m P(A_m) = 0$ であるから，$E[X \mid \mathcal{G}] \leq E[Y \mid \mathcal{G}]$ a.s.

(c). $Y_n = E[X_n \mid \mathcal{G}]$ とおくと，(b) より，a.s. に Y_n は非減少であるから，$0 \leq Y_n \uparrow \exists Y$ a.s. ここで，Y_n は $\mathcal{G}$ 可測であるから，Y を $\mathcal{G}$ 可測になるように選べる．また，$Y \leq E[X \mid \mathcal{G}]$ より，$Y \in L^1$ である．このとき，MCT より，$A \in \mathcal{G}$ に対して，

$$\begin{array}{ccc} \displaystyle\int_A Y_n dP & = & \displaystyle\int_A X_n dP \\ \downarrow & & \downarrow \\ \displaystyle\int_A Y dP & = & \displaystyle\int_A X dP. \end{array}$$

従って，$Y = E[X \mid \mathcal{G}]$.

(d). $Y_n = X_1 - X_n$ とおくと，$0 \leq Y_n \uparrow X_1 - X =: Y$. $E[Y] < \infty$ であるから，(c) より，$E[X_1 - X_n \mid \mathcal{G}] = E[Y_n \mid \mathcal{G}] \uparrow E[Y \mid \mathcal{G}] = E[X_1 - X \mid \mathcal{G}]$.

(e). φ は凸関数であるから，高々可算個の a_n, b_n が存在して，$\varphi(x) = \sup_n(a_n x + b_n)$ と表せる (cf. Corollary 6.1)．ここで，$\varphi(x) \geq a_n x + b_n$ であるから，

$$E[\varphi(X_n) \mid \mathcal{G}] \geq a_n E[X \mid \mathcal{G}] + b_n \quad a.s. \tag{$**$}$$

となる．$(**)$ の除外集合を N_n とおくと，$\omega \in (\bigcup_n N_n)^c$ に対して，

$$E[\varphi(X_n) \mid \mathcal{G}](\omega) \geq \sup_n(a_n E[X \mid \mathcal{G}](\omega) + b_n) = \varphi(E[X \mid \mathcal{G}](\omega))$$

が成り立つ．$P(\bigcup_n N_n) = 0$ より求める結論を得る．

(f). (e) より，$|E[X \mid \mathcal{G}]|^p \leq E[|X|^p \mid \mathcal{G}]$ であるから，両辺の期待値をとればよい．

(g). $A \in \mathcal{G}_1 \subset \mathcal{G}_2$ に対して，

$$E[E[X \mid \mathcal{G}_2], A] = E[X, A].$$

$E[X \mid \mathcal{G}_2] \in \mathcal{G}_1$ より，$E[X \mid \mathcal{G}_2] = E[X \mid \mathcal{G}_1]$.

(h). 最初の等号は $E[X \mid \mathcal{G}_1]$ が $\mathcal{G}_2$ 可測なことから従う．次の等号は，$A \in \mathcal{G}_1$ に対して，

$$\begin{aligned} E[E[E[X \mid \mathcal{G}_2] \mid \mathcal{G}_1], A] &= E[E[X \mid \mathcal{G}_2], A] \quad (\because E[E[X \mid \mathcal{G}_2] \mid \mathcal{G}_1] \text{ の定義}) \\ &= E[X, A] \quad (\because A \in \mathcal{G}_2 \text{ と } E[X \mid \mathcal{G}_2] \text{ の定義}) \end{aligned}$$

となることから従う．

(i). $X, Y \geq 0$ と仮定する．$X \in \mathcal{G}$ より，$\mathcal{G}$ 可測な非負単関数列で，$\uparrow X$ となるものが存在する．よって，(c) より，$X = 1_A, A \in \mathcal{G}$ なる場合を考えればよい．このとき，$B \in \mathcal{G}$ に対して，

$$E[1_A E[Y \mid \mathcal{G}], B] = E[E[Y \mid \mathcal{G}], A \cap B] = E[Y, A \cap B] = E[1_A Y, B]$$

であるから，$E[1_A Y \mid \mathcal{G}] = 1_A E[Y \mid \mathcal{G}]$. 一般の X, Y に対しては，$XY = (X^+ - X^-)(Y^+ - Y^-) = X^+Y^+ - X^+Y^- - X^-Y^+ + X^-Y^-$ なる分解を利用すればよい． □

L^2 射影としての条件付き期待値.

Theorem 19.6. $X \in L^2(\Omega, \mathcal{F}, P)$ に対して，$E[X \mid \mathcal{G}]$ は最小化問題

$$\min\{E[(X-Y)^2] : Y \in L^2(\Omega, \mathcal{G}, P)\}.$$

の a.s. な意味での一意な最適解である．

Proof. $Y \in L^2(\Omega, \mathcal{G}, P)$ に対して，

$$\begin{aligned} &E[(X-Y)^2] = E[(X - E[X \mid \mathcal{G}] + E[X \mid \mathcal{G}] - Y)^2] \\ &= E[(X - E[X \mid \mathcal{G}])^2] + 2E[(X - E[X \mid \mathcal{G}])(E[X \mid \mathcal{G}] - Y)] + E[(E[X \mid \mathcal{G}] - Y)^2]. \end{aligned}$$

$E[X \mid \mathcal{G}] - Y \in \mathcal{G}$ より，

$$E[(X - E[X \mid \mathcal{G}])(E[X \mid \mathcal{G}] - Y) \mid \mathcal{G}] = (E[X \mid \mathcal{G}] - Y)E[(X - E[X \mid \mathcal{G}]) \mid \mathcal{G}] = 0.$$

従って，

$$\begin{aligned} E[(X-Y)^2] &= E[(X - E[X \mid \mathcal{G}])^2] + E[(E[X \mid \mathcal{G}] - Y)^2] \\ &\geq E[(X - E[X \mid \mathcal{G}])^2] \end{aligned}$$

であり，等号が成立するのは，$Y = E[X \mid \mathcal{G}]$ a.s. のときのみである． □

Remark 19.3. a.s. に等しい r.v.'s を同一視すれば，$H = L^2(\Omega, \mathcal{F}, P)$ は内積 $\langle X, Y \rangle = E[XY]$ に関して Hilbert 空間であり，$V = L^2(\Omega, \mathcal{G}, P)$ は H の閉部分空間である．Theorem 19.6 は，$X \in H$ に対して，条件付き期待値 $E[X \mid \mathcal{G}]$ が X の V への正射影として特徴づけられることを示している．すなわち，

$$E[X \mid \mathcal{G}] = \Pi_V X, \ X \in H$$

である．特に，$Y \in V$ に対して，

$$Y = E[X \mid \mathcal{G}] \Leftrightarrow E[Z(X-Y)] = 0, \ \forall Z \in V$$

である．逆に，$X \in H$ に対して $E[X \mid \mathcal{G}]$ を V へ正射影として定義し，あとは極限操作によって $X \in L^1$ に対して条件付き期待値の定義を与える方法もある．また，この条件付き期待値の定義から出発して，Radon-Nikodym の定理を証明することもできる．詳細は，Williams (1991) を参照せよ．

なお，与えられた $X_1, \ldots, X_n \in H$ に対して，W を $X_1, \ldots, X_n$ の 1 次結合からなる部分空間とすると，特殊な場合を除いて，

$$\Pi_W \neq E[\,\cdot \mid \sigma(X_1, \ldots, X_n)]$$

である．$Y \in H$ に対して，$\Pi_W Y$ は $X_1, \ldots, X_n$ の 1 次結合であるが，$E[Y \mid \sigma(X_1, \ldots, X_n)]$ は $X_1, \ldots, X_n$ の非線形関数でもよいところが違いである．

分布関数の分解．F が劣分布関数 (sub distribution function, s.d.f.) であるとは，F が右連続かつ非減少であって，$F(-\infty)=0, F(\infty)\leq 1$ をみたすことを言う．$F(\infty)=1$ のとき，F は d.f. である．F に対応する Lebesgue-Stieltjes 測度を μ とおく．また，λ を $\mathbb{R}$ 上の Lebesgue 測度とする．F が絶対連続であるとは，ある非負可測関数 $f:\mathbb{R}\to\mathbb{R}_+$ が存在して，$F(x)=\int_{-\infty}^{x} f(y)dy$ と表せることを言う．これは $\mu\ll\lambda$ と同値である．また，F が離散的 (discrete) であるとは，ある可算集合 $\{a_j\}\subset\mathbb{R}$ と $b_j>0, \sum_j b_j\leq 1$ なる $\{b_j\}$ が存在して，$F(x)=\sum_{j:a_j\leq x} b_j$ と表せることを言う．このとき，$\mu=\sum_j b_j\delta_{a_j}$ であるから，$\mu\perp\lambda$ である．F が特異連続 (singular continuos) であるとは，F が連続であって，$\mu\perp\lambda$ となることを言う．例えば，Cantor 分布の d.f. は特異連続である．

あらゆる d.f. は絶対連続，特異連続，離散劣分布関数の和として表せることを示そう．F を d.f. とし，対応する p.m. を μ とおく．μ の λ に関する Lebesgue 分解を

$$\mu=\mu_{ac}+\mu_s,\quad \mu_{ac}\ll\lambda,\ \mu_s\perp\lambda$$

とおく．$G_{ac}(x)=\mu_{ac}((-\infty,x]), G_s(x)=\mu_s((-\infty,x])$ とおく．G_{ac} は絶対連続である．G_s の不連続点の全体を $\{a_j\}$ とし，$b_j=G(a_j)-G(a_j-)$ とおく．$G_d=\sum_{j:a_j\leq x} b_j$ と定義すると，G_d は離散的である．さらに，$G_{sc}=G_s-G_d$ と定義すると，G_{sc} は特異連続である．よって，F は

$$F=G_{ac}+G_{sc}+G_d$$

と分解できる．さらに，この分解は一意である．まず，Lebesgue 分解の一意性より，$F=G_{ac}+G_s$ の分解の一意性が従う．$G_s=G_{sc}+G_d$ の分解の一意性は明らかであろう．このとき，G_{ac}, G_{sc}, G_d をそれぞれ F の絶対連続部分，特異連続部分，離散部分と呼ぶ．

演習問題

Exercise 19.1. $(\Omega,\mathcal{F})$ 上の p.m.'s の族 $\{P_n : n\in\mathbb{N}\}$ に対して，$P_n\ll\mu$ $(\forall n\in\mathbb{N})$ となる $(\Omega,\mathcal{F})$ 上の p.m. μ が存在することを示せ．

Exercise 19.2. Radon-Nikodym の定理は μ が σ-finite でないと成り立たない．例えば，$S=\mathbb{R}$ とし，$\mathcal{S}=\{A\subset\mathbb{R} : A \text{ or } A^c \text{ は可算集合}\}$ とおいて，$(S,\mathcal{S})$ 上に次の 2 つの測度を考える：

$$\mu(A)=\begin{cases}\mathrm{Card}(A) & A\text{ が有限集合のとき}\\ +\infty & A\text{ が無限集合のとき}\end{cases},\quad \nu(A)=\begin{cases}0 & A\text{ が可算集合のとき}\\ 1 & A^c\text{ が可算集合のとき}\end{cases}.$$

$\mathcal{S}$ が σ-field となることと，μ と ν が $(S,\mathcal{S})$ 上の測度となることは認めてよい．このとき，$\nu\ll\mu$ であるが，

$$\nu(A)=\int_A f d\mu,\ \forall A\in\mathcal{S}$$

をみたす $f\in L^1(S,\mathcal{S},\mu)$ は存在しないことを示せ．

Exercise 19.3. $(S, \mathcal{S})$ を可測空間とする．

(a) μ, ν, ρ を $(S, \mathcal{S})$ 上の有限測度とし，$\nu \ll \mu$ と $\mu \ll \rho$ とする．このとき，$\nu \ll \rho$ であり，

$$\frac{d\nu}{d\rho} = \frac{d\nu}{d\mu}\frac{d\mu}{d\rho} \quad \rho\text{-a.e.}$$

となることを示せ．

(b) μ, ν を $(S, \mathcal{S})$ 上の有限測度とし，$\mu \ll \nu$ かつ $\nu \ll \mu$ とする．このとき，$d\mu/d\nu$ は μ-a.e. に正であり，

$$\frac{d\nu}{d\mu} = \left(\frac{d\mu}{d\nu}\right)^{-1} \quad \mu\text{-a.e.}$$

となることを示せ．

Exercise 19.4. $x \in \mathbb{R}$ に対して，δ_x を x の 1 点分布とする．

(a) $\delta_x \ll \mu$ $(\forall x \in \mathbb{R})$ となる $(\mathbb{R}, \mathcal{B})$ 上の σ-finite な測度 μ は存在しないことを示せ．

(b) (a) において σ-finite という条件を外すとどうか．

Exercise 19.5. Lebesgue 分解を証明せよ．すなわち，μ と ν を $(S, \mathcal{S})$ 上の σ-finite な測度とすると，次をみたす $(S, \mathcal{S})$ 上の一意な測度 ν_{ac}, ν_s が存在することを示せ：

$$\nu = \nu_{ac} + \nu_s,\ \nu_{ac} \ll \mu,\ \nu_s \perp \mu.$$

Exercise 19.6. X, Y を同じ分布に従う独立な r.v.'s とし，$X, Y \in L^1$ とする．このとき，$E[X \mid X + Y]$ を求めよ．

Exercise 19.7. $X, Y \in L^1$ が $E[X \mid Y] = Y, E[Y \mid X] = X$ をみたすなら，$X = Y$ a.s. であることを示せ．

Exercise 19.8. X, Y, Y' を r.v.'s とすると，$(X, Y) \overset{d}{=} (X, Y')$ であるためには，$P(Y \in B \mid X) = P(Y' \in B \mid X)$ a.s. $\forall B \in \mathcal{B}$ であることが必要十分であることを示せ．

射影定理と Riesz の表現定理に関連して，L^2 における弱収束を考察しよう．以下，$L^2 = L^2(\Omega, \mathcal{F}, P)$ とし，$X, X_n \in L^2$ に対して，X_n が X に L^2 において 弱収束 するとは，L^2 上のあらゆる連続線形汎関数 φ に対して，$\varphi(X_n) \to \varphi(X)$ となることを言う．このとき，$X_n \to X$ weakly in L^2 と書く．Riesz の表現定理より，これは，任意の $Y \in L^2$ に対して，$E[YX_n] \to E[YX]$ となることと同値である．$X_n \to X, X_n \to X'$ weakly in L^2 であれば，$E[YX] = E[YX']$ $(\forall Y \in L^2)$ であるから，$Y = X - X'$ とおいて，$E[(X - X')^2] = 0$，よって，$X = X'$ a.s. を得る．すなわち，L^2 の弱収束極限は存在すれば一意である．また，$X_n \to X$ in L^2 なら，$X_n \to X$ weakly in L^2 である．逆は成り立たない．

Exercise 19.9. $\{X_m\}$ を独立な r.v.'s とし，$E[X_m] = 0, E[X_m^2] = 1$ $(\forall m)$ とする．このとき，$X_n \to 0$ weakly in L^2 を示せ．
ヒント：$Y \in L^2$ に対して，$E[Y^2] \geq \sum_{m=1}^n (E[YX_m])^2$ を示せ．

Exercise 19.10. $X_n \to X$ weakly in L^2 なら，$E[X^2] \leq \liminf_n E[X_n^2]$ となることを示せ．

また，L^2 の弱収束列は L^2 有界である．これは関数解析における Banach-Steinhaus の定理 (一様有界性原理) から直ちに従う．

Exercise 19.11. $\{X_m\}$ を i.i.d. r.v.'s とし，$E[X_m] = 0, E[X_m^2] = 1$ とする．$\check{S}_n = n^{-1/2} \sum_{m=1}^n X_m$ とおく．

(a) $\check{S}_n \to 0$ weakly in L^2 を示せ．

(b) $\check{S}_n$ は確率収束しないことを示せ．

以前にも注意したように，L^2 空間はほとんどの場合，無限次元であり，L^2 においては有界閉集合であることとコンパクト集合であることは同値でない．特に，r.v.'s が L^2 有界だからといって，L^2 収束する部分列をもつとは限らない．しかし，L^2 有界な r.v.'s は L^2 において弱収束する部分列をもつ．

Exercise 19.12 (Banach-Alaoglu の定理の特殊ケース)**.** $\{X_n\}$ を L^2 有界な r.v.'s とする．このとき，$\{X_n\}$ は L^2 において弱収束する部分列をもつことを示せ．
ヒント：対角線論法より，$\{n\}$ の適当な部分列 $\{n(k)\}$ が存在して，$\forall m$ に対して，有限な $\lim_k E[X_m X_{n(k)}]$ が存在する．このとき，あらゆる $Y \in L^2$ に対して，有限な $F(Y) := \lim_k E[YX_{n(k)}]$ が存在する．あとは Riesz の表現定理を使え．

次の Dunford-Pettis の定理は一様可積分性の興味深い特徴づけを与える．

Exercise 19.13 (Dunford-Pettis)**.** $\{X_n : n \in \mathbb{N}\} \subset L^1$ とする．このとき，$\{X_n\}$ が一様可積分であるためには，ある部分列 $\{X_{n(m)}\}$ と r.v. $X \in L^1$ が存在して，あらゆる $Y \in L^\infty$ に対して，$E[YX_{n(m)}] \to E[YX]$ となることが必要十分であることを示せ．
必要性の証明のヒント：$X_{n,k} = X_n 1_{\{|X_n| \leq k\}}$ とおくと，$\{X_{n,k} : n \in \mathbb{N}\}$ は L^2 有界であるから，Banach-Alaoglu の定理と対角線論法より，$\{n\}$ の適当な部分列 $\{n(m)\}$ が存在して，$\forall k$ に対して，$X_{n(m),k} \to Y_k$ weakly in L^2 $(m \to \infty)$ となる．$\{Y_k\}$ は L^1 の Cauchy 列である．その極限を X とせよ．

Remark 19.4. すなわち，$\{X_n\} \subset L^1$ が一様可積分であるためには，それが L^1 の弱位相に関して相対点列コンパクトであることが必要十分である．Eberlein-Šmulian の定理より，Banach 空間に対しては，弱コンパクト性と弱点列コンパクト性は同値なので，Dunford-Pettis の定理は，一様可積分性と L^1 における弱相対コンパクト性が同値であることを述べている．

ちょっと寄り道となるが，一様可積分性に関連して，次の Komlós の補題を証明しよう．

Exercise 19.14 (Komlós の補題)**.** 本問の目標は次の Komlós の補題の証明である：一様可積分な r.v.'s $\{X_n : n \in \mathbb{N}\}$ に対して，$Y_n \in \mathrm{conv}\{X_n, X_{n+1}, \dots\}, n \in \mathbb{N}$ を L^1 収束するように選ぶことができる．ただし，$\mathrm{conv}\{X_n, X_{n+1}, \dots\}$ は $X_n, X_{n+1}, \dots$ の有限凸結合の全体である．この命題を以下のステップに従い証明せよ．cf. Beiglbock et al. (2012).

(a) $\{X_n\}$ が L^2 有界なら，$Y_n \in \mathrm{conv}\{X_n, X_{n+1}, \dots\}, n \in \mathbb{N}$ を L^2 収束するように選ぶことができることを示せ．

ヒント：$A = \sup_n \inf\{\|Y\|_2 : Y \in \mathrm{conv}\{X_n, X_{n+1}, \dots\}\}$ とおいたとき，$Y_n \in \mathrm{conv}\{X_n, X_{n+1}, \dots\}$ を $\|Y_n\|_2 \le A + 1/n$ & $\|Y_n\|_2 \uparrow A$ となるように選べる．このとき，$\{Y_n\}$ が L^2 の Cauchy 列であることを示せ．

(b) Komlós の補題を示せ．

ヒント：$X_{n,k} = X_n 1_{\{|X_n| \le k\}}$ とおくと，(a) と対角線論法より，各 n に対して，正整数 N_n と $\alpha_{n,n} + \cdots + \alpha_{n,N_n} = 1$ なる $\alpha_{n,n}, \dots, \alpha_{n,N_n} \in \mathbb{R}_+$ を，$\{\alpha_{n,n}X_{n,k} + \cdots + \alpha_{n,N_n}X_{N_n,k} : n \in \mathbb{N}\}$ がすべての k に対して L^2 収束するように選べる．このとき，$\{X_n\}$ の一様可積分性より，$\{\alpha_{n,n}X_n + \cdots + \alpha_{n,N_n}X_{N_n} : n \in \mathbb{N}\}$ が L^1 の Cauchy 列であることを示せ．

じつは Komlós (1967) はもっと強いことを証明している．その証明は省略するが，次の形の Komlós の補題はなぜだか数理ファイナンスと関連があるそうである [12].

Theorem 19.7 (Komlós (1967))**.** L^1 有界な r.v.'s $\{X_n\}$ に対して，ある部分列 $\{X_{n_k}\}$ が存在して，$N^{-1}\sum_{k=1}^{N} X_{n_k}$ が $N \to \infty$ のとき L^1 収束する．

[12] ちなみに，Komlós の補題は尾張圭太氏に教えてもらった．どうもありがとうございます．

20　正則条件付き分布と Kolmogorov の拡張定理

$(S, \mathcal{S})$ を可測空間とし，X を S-valued r.v. とする．写像 $\mu : \Omega \times \mathcal{S} \to [0,1]$ が $\mathcal{G}$ を与えたときの X の 正則条件付き分布 (regular conditional distribution, r.c.d.) であるとは，

(i) $A \in \mathcal{S}$ を任意に固定すると，$\omega \mapsto \mu(\omega, A)$ が $P(X \in A \mid \mathcal{G})$ のバージョンであり，

(ii) $\omega \in \Omega$ を任意に固定すると，$A \mapsto \mu(\omega, A)$ が $(S, \mathcal{S})$ 上の p.m.

となることを言う．$\Omega = S, X(\omega) = \omega$ のとき，μ を 正則条件付き確率 (regular conditional probability, r.c.p.) と呼ぶ．

r.c.d. が存在するとき，条件付き期待値は r.c.d. に関する期待値に帰着する．

Lemma 20.1 (Durrett 3.1.14)**.** $\mathcal{G}$ を与えたときの X の r.c.d. $(\omega, A) \mapsto \mu(\omega, A)$ が存在するとする．$f : S \to \mathbb{R}$ を $E[|f(X)|] < \infty$ なる可測関数とすると，

$$E[f(X) \mid \mathcal{G}](\omega) = \int f(x)\mu(\omega, dx) \ a.s.$$

一般の S-valued r.v. に対して r.c.d. は必ずしも存在するとは限らない．注意すべきは，排反な $A_n \in \mathcal{S}, n \in \mathbb{N}$ に対して，

$$P\left(X \in \bigcup_{n=1}^{\infty} A_n \mid \mathcal{G}\right) = \sum_{n=1}^{\infty} P(X \in A_n \mid \mathcal{G}) \ a.s.$$

となるのはよいが，除外集合が $\{A_n\}$ のとり方に依存しているため，$\{A_n\}$ のとり方が非可算無限個ある場合 ($\mathcal{S}$ が無限集合である限りそうである)，除外集合を集めたものが測度 0 となるとは限らないことである．

Example 20.1 (r.c.d. が存在しない例)**.** $(\Omega, \mathcal{G}) = ([0,1], \mathcal{B}_{[0,1]})$ とし，λ を $[0,1]$ 上の Lebesgue 測度とする．また，λ^*, λ_* をそれぞれ λ の外測度，内測度とする．すなわち，$\lambda^*(A) := \inf\{\lambda(B) : B \in \mathcal{G}, B \supset A\}$, $\lambda_*(A) := 1 - \lambda^*(A^c)$. いま，$D \subset [0,1]$ を $\lambda^*(D) = 1, \lambda_*(D) = 0$ をみたす集合とする (cf. Theorem 7.5). ここで，$\mathcal{F} := \mathcal{G} \vee \sigma(\{D\})$ とおくと，各 $F \in \mathcal{F}$ は $F = (D \cap A) \cup (D^c \cap B)$ $(A, B \in \mathcal{G})$ の形に表すことができて，$\lambda^*(D) = \lambda^*(D^c) = 1$ より，

$$\lambda^*(D \cap F) = \lambda^*(D \cap A) = \lambda(A), \ \lambda^*(D^c \cap F) = \lambda^*(D^c \cap B) = \lambda(B)$$

となる．そこで，$(\Omega, \mathcal{F})$ 上の p.m. P を

$$P(F) := \frac{1}{2}\lambda(A) + \frac{1}{2}\lambda(B)$$

と定義すると，P は well-defined である．また，$F \in \mathcal{G}$ なら，$P(F) = \lambda(F)$ である．

いま，$\mathcal{G}$ を与えたときの r.c.p. が存在しないことを示そう．仮にそのような r.c.p. $\mu : \Omega \times \mathcal{F} \to [0,1]$ が存在したとして，矛盾を導く．$A, B \in \mathcal{G}$ に対して，

$$E[\mu(\cdot, D \cap A), B] = P(D \cap A \cap B) = \frac{1}{2}\lambda(A \cap B) = E[(1/2)1_A, B].$$

よって，

$$\mu(\omega, D \cap A) = \frac{1}{2}1_A(\omega) \quad \text{a.s.}$$

を得る．$\mathcal{G}$ は可算な π システム $\mathcal{I} = \{(-\infty, q] : q \in \mathbb{Q}\}$ によって生成されていて，各 $\omega \in \Omega$ に対して，$A \mapsto \mu(\omega, D \cap A)$, $A \mapsto (1/2)1_A(\omega)$ は測度なので，

$$J := \{\omega : \mu(\omega, D \cap A) = (1/2)1_A(\omega),\ \forall A \in \mathcal{G}\} = \{\omega : \mu(\omega, D \cap A) = (1/2)1_A(\omega)\ \forall A \in \mathcal{I}\}$$

とおくと，$\omega \mapsto \mu(\omega, D \cap A)$ は仮定より $\mathcal{G}$ 可測なので $J \in \mathcal{G}$ であり，また $P(J) = \lambda(J) = 1$ である．ここで，$\omega \in J$ に対して，

$$\mu(\omega, D \cap J) = \frac{1}{2}1_J(\omega) \neq \frac{1}{2}1_{J \setminus \{\omega\}}(\omega) = \mu(\omega, D \cap (J \setminus \{\omega\}))$$

であるから，$D \cap J \neq D \cap (J \setminus \{\omega\})$ である．これは，$\omega \in D$ を意味する．従って，$J \subset D$ となるが，$\lambda(J) = 1$ であったから，$\lambda_*(D) = 1$ でなくてはならないが，これは $\lambda_*(D) = 0$ に反する．

しかしながら，多くの場合，r.c.d. は存在する．

Theorem 20.1. $(S, \mathcal{S})$ が標準 Borel 空間なら，S-valued r.v. X は r.c.d. をもつ．

Proof. $(S, \mathcal{S}) = (\mathbb{R}, \mathcal{B})$ の場合を考えれば十分である．各 $q \in \mathbb{Q}$ に対して，$P(X \leq q \mid \mathcal{G})$ のバージョンを決める．$\mathbb{Q}$ は可算なので，

$$\Omega_0 = \Big\{\omega \in \Omega : P(X \leq q \mid \mathcal{G})(\omega) \leq P(X \leq q' \mid \mathcal{G})(\omega),\ \forall (q, q') \in \mathbb{Q}^2 \text{ s.t. } q < q',$$
$$\sup_{q \in \mathbb{Q}} P(X \leq q \mid \mathcal{G})(\omega) = 1, \inf_{q \in \mathbb{Q}} P(X \leq q \mid \mathcal{G})(\omega) = 0\Big\}$$

とおくと，$P(\Omega_0) = 1$ である．G を任意の d.f. として，

$$F(x \mid \mathcal{G})(\omega) := \begin{cases} \inf_{q > x, q \in \mathbb{Q}} P(X \leq q \mid \mathcal{G})(\omega) & \text{if } \omega \in \Omega_0 \\ G(x) & \text{if } \omega \notin \Omega_0 \end{cases}$$

と定義すると，各 $\omega \in \Omega$ に対して，$x \mapsto F(x \mid \mathcal{G})(\omega)$ は d.f. であり (cf. Prohorov の定理の証明)，各 $x \in \mathbb{R}$ に対して，$F(x \mid \mathcal{G})$ は $P(X \leq x \mid \mathcal{G})$ のバージョンとなる．

各 $\omega \in \Omega$ に対して，$\mu(\omega, \cdot) : \mathcal{B} \to [0,1]$ を d.f. $x \mapsto F(x \mid \mathcal{G})(\omega)$ に対応する p.m. とする．すなわち，$\mu(\omega, \cdot)$ は

$$\mu(\omega, (a, b]) = F(b \mid \mathcal{G})(\omega) - F(a \mid \mathcal{G})(\omega),\ -\infty \leq a < b < \infty$$

をみたす一意な p.m. である．あとは，任意の $B \in \mathcal{B}$ に対して，$\omega \mapsto \mu(\omega, B)$ が $P(X \in B \mid \mathcal{G})$ のバージョンであることを示せばよい．

$$\begin{aligned}\mathcal{C} = \{B \in \mathcal{B} : \omega \mapsto \mu(\omega, B) \text{ は } \mathcal{G} \text{ 可測}, \\ \mu(\omega, B) = P(X \in B \mid \mathcal{G})(\omega) \ a.s.\}, \\ \mathcal{D} = \{(-\infty, x] : x \in \mathbb{R}\}\end{aligned}$$

とおくと，$\mathcal{D}$ は π システムかつ $\mathcal{D} \subset \mathcal{C}$ であり，$\mathcal{C}$ は λ システムであるから，π-λ 定理より，$\mathcal{C} \supset \sigma(\mathcal{D}) = \mathcal{B}$ を得る．以上より，定理の結論を得る． □

$(S, \mathcal{S}), (T, \mathcal{T})$ を可測空間，X, Y をそれぞれ S-valued, T-valued r.v.'s とする．(X, Y) の $\mathcal{S} \times \mathcal{T}$ 上の同時分布を ν とおく：

$$\nu(E) = P((X, Y) \in E), \ E \in \mathcal{S} \times \mathcal{T}.$$

ここで，確率空間を取り替えて，$(\Omega, \mathcal{F}, P) = (S \times T, \mathcal{S} \times \mathcal{T}, \nu), X(x, y) = x, Y(x, y) = y$ とする．$(T, \mathcal{T})$ が標準 Borel 空間なら，$\sigma(X)$ を与えたときの Y の r.c.d. $\Omega \times \mathcal{T} \ni (\omega, B) \mapsto \mu(\omega, B) \in [0, 1]$ が存在する．$B \in \mathcal{T}$ を任意に固定すると，$\omega \mapsto \mu(\omega, B)$ は $P(Y \in B \mid X)$ のバージョンであるから，$\sigma(X)$ 可測である．従って，次をみたす可測関数 $x \mapsto \nu_{2,1}(x, B)$ が存在する：$\mu(\omega, B) = \nu_{2,1}(X(\omega), B)\ \forall \omega \in \Omega$．このとき，変数変換公式より，$\nu_1 := \mathcal{L}(X)$ とおいて，

$$\begin{aligned}\nu(A \times B) &= P(X \in A, Y \in B) = E[P(Y \in B \mid X), X^{-1}(A)] \\ &= E[\nu_{2,1}(X, B), X^{-1}(A)] = \int_{X^{-1}(A)} \nu_{2,1}(X, B) dP \\ &= \int_A \nu_{2,1}(x, B) d\nu_1(x) = \int_A \int_B \nu_{2,1}(x, dy) d\nu_1(x).\end{aligned}$$

が成り立つ．以上の議論より，次の定理を得る．

Theorem 20.2. $(T, \mathcal{T})$ が標準 Borel 空間なら，次の (i)–(iii) をみたす写像 $S \times \mathcal{T} \ni (x, B) \mapsto \nu_{2,1}(x, B) \in [0, 1]$ が存在する．

(i) 各 $B \in \mathcal{T}$ に対して，$x \mapsto \nu_{2,1}(x, B)$ は $\mathcal{S}$ 可測である．

(ii) 各 $x \in S$ に対して，$B \mapsto \nu_{2,1}(x, B)$ は p.m. である．

(iii) 各 $A \in \mathcal{S}, B \in \mathcal{T}$ に対して，

$$\nu(A \times B) = \int_A \int_B \nu_{2,1}(x, dy) d\nu_1(x).$$

(i)–(iii) をみたす写像 $(x, B) \mapsto \nu_{2,1}(x, B)$ を X を与えたときの Y の r.c.d. と呼ぶ．ややインフォーマルな気もするが，

$$P(Y \in dy \mid X = x) = \nu_{2,1}(x, dy)$$

と書いている本もある．

次に，Fubini の定理より，各 $E \in \mathcal{S} \times \mathcal{T}$ に対して，$x \in S$ を任意に固定すると，$y \mapsto 1_E(x, y)$ は $\mathcal{T}$ 可測である．そこで，

$$\mathcal{E} = \left\{ E \in \mathcal{S} \times \mathcal{T} : x \mapsto \int 1_E(x, y)\nu_{2,1}(x, dy) \text{ は } \mathcal{S} \text{ 可測であり}, \right.$$
$$\left. \nu(E) = \iint 1_E(x, y)\nu_{2,1}(x, dy)d\nu_1(x) \right\}$$

とおくと，$\mathcal{E}$ は λ システムであり，π システム $\{A \times B : A \in \mathcal{S}, B \in \mathcal{T}\}$ を含む．従って，π-λ 定理より，次の定理の前半を得る．

Theorem 20.3. 各 $E \in \mathcal{S} \times \mathcal{T}$ に対して，

$$\nu(E) = \iint 1_E(x, y)\nu_{2,1}(x, dy)d\nu_1(x). \tag{$*$}$$

従って，$\mathcal{S} \times \mathcal{T}$ 可測な関数 $h : S \times T \to \overline{\mathbb{R}}$ に対して，$h \geq 0$ or $E[|h(X, Y)|] < \infty$ なら，

$$E[h(X, Y)] = \int h d\nu = \iint h(x, y)\nu_{2,1}(x, dy)d\nu_1(x)$$

が成り立つ (最右辺の積分が定義できることも主張の一部である)．

Proof. 前半の主張はすでに示した．後半の主張の証明は Fubini の定理と同様． □

Remark 20.1. $(*)$ を

$$d\nu(x, y) = \nu_{2,1}(x, dy)d\nu_1(x)$$

と書く場合もある．

Example 20.2 (十分統計量)**.** 可測空間 $(\Omega, \mathcal{F})$ 上に p.m.'s の族 $\mathcal{P}$ が与えられているとする．r.v. T ("統計量") が $\mathcal{P}$ に対して 十分 (sufficient) であるとは，任意の $A \in \mathcal{F}$ に対して，$P \in \mathcal{P}$ に依存しない r.v. X_A が存在して，$P(A \mid T) = X_A$ P-a.s. $(\forall P \in \mathcal{P})$ をみたすことを言う．p.m. $P \in \mathcal{P}$ の下での期待値を $E^P[\cdot]$ と表して，r.v. Y を $E^P[|Y|] < \infty$ $(\forall P \in \mathcal{P})$ とすると，単関数による近似を用いれば，$P \in \mathcal{P}$ に依存しない r.v. X が存在して，$E^P[Y \mid T] = X$ P-a.s. $(\forall P \in \mathcal{P})$ となることが示せる．また，$(\Omega, \mathcal{F})$ が標準 Borel 空間なら，$P \in \mathcal{P}$ に依存しない適当な写像 $\mu : (\omega, A) \mapsto \mu(\omega, A)$ が存在して，各 $P \in \mathcal{P}$ に対して μ は $\sigma(T)$ を与えたときの r.c.p. となる．証明は $P \in \mathcal{P}$ に関する一様性に気をつければ，Theorem 20.1 と同様である．十分統計量に関しては，Lehmann and Romano (2005) が詳しい．

Kolmogorov の拡張定理

Theorem 20.3 の応用として，Kolmogorov の拡張定理を証明する．$\mathbb{R}^{\mathbb{N}} = \mathbb{R} \times \mathbb{R} \times \cdots$ を $\mathbb{R}$ の可算直積とし，$\mathcal{B}^{\mathbb{N}}$ を $\mathbb{R}^{\mathbb{N}}$ のシリンダー σ-field とする．すなわち，

$$\mathcal{B}^{\mathbb{N}} = \sigma(\{B \times \mathbb{R} \times \mathbb{R} \times \cdots : B \in \mathcal{B}^n, n \in \mathbb{N}\})$$

である．

Theorem 20.4 (Kolmogorov の拡張定理)**.** 各 $n \in \mathbb{N}$ に対して，$(\mathbb{R}^n, \mathcal{B}^n)$ 上の p.m. μ_n が与えられていて，$\mu_n, n \in \mathbb{N}$ は次の意味で 整合的 (consistent) と仮定する：

$$\mu_{n+1}(B \times \mathbb{R}) = \mu_n(B),\ \forall B \in \mathcal{B}^n, \forall n \in \mathbb{N}.$$

このとき，次をみたす $(\mathbb{R}^{\mathbb{N}}, \mathcal{B}^{\mathbb{N}})$ 上の p.m. μ が存在する：

$$\mu(B \times \mathbb{R} \times \mathbb{R} \times \cdots) = \mu_n(B),\ \forall B \in \mathcal{B}^n,\ \forall n \in \mathbb{N}. \tag{$**$}$$

そのような μ は一意に決まる．

Proof. 一意性の証明は π-λ 定理を使えばよい．存在を示す．$(\xi_1, \ldots, \xi_n, \xi_{n+1}) \sim \mu_{n+1}$ に対して，$(\xi_1, \ldots, \xi_n)$ を与えたときの ξ_{n+1} の r.c.d. を $\mu_{n+1,n}(x_1, \ldots, x_n, dx_{n+1})$ とおき，その d.f. を

$$F_{n+1}(x_{n+1} \mid x_1, \ldots, x_n) = \mu_{n+1,n}(x_1, \ldots, x_n, (-\infty, x_{n+1}])$$

とおく．ここで，$x_{n+1} \mapsto F_{n+1}(x_{n+1} \mid x_1, \ldots, x_n)$ の分位点関数を

$$F_{n+1}^{-1}(u \mid x_1, \ldots, x_n) := \inf\{x_{n+1} : F_{n+1}(x_{n+1} \mid x_1, \ldots, x_n) \geq u\}$$

と定義する．写像 $(u, x_1, \ldots, x_n) \mapsto F_{n+1}^{-1}(u \mid x_1, \ldots, x_n)$ の可測性を確認する．$a \in \mathbb{R}$ に対して，

$$\begin{aligned}
&\{(u, x_1, \ldots, x_n) : F_{n+1}^{-1}(u \mid x_1, \ldots, x_n) \leq a\} \\
&= \{(u, x_1, \ldots, x_n) : u \leq F_{n+1}(a \mid x_1, \ldots, x_n)\} \\
&= \{(u, x_1, \ldots, x_n) : u - F_{n+1}(a \mid x_1, \ldots, x_n) \leq 0\}.
\end{aligned}$$

であり，$(x_1, \ldots, x_n) \mapsto F_{n+1}(a \mid x_1, \ldots, x_n) = \mu_{n+1,n}(x_1, \ldots, x_n, (-\infty, a])$ は可測であるから，$(u, x_1, \ldots, x_n) \mapsto F_{n+1}^{-1}(u \mid x_1, \ldots, x_n)$ の可測性が示された．

$U_1, U_2, \cdots \sim U(0,1)$ i.i.d. を適当な確率空間上に構成する．μ_1 の d.f. を F_1 として，r.v.'s $\{X_m\}$ を帰納的に

$$X_1 = F_1^{-1}(U_1),\ X_{n+1} = F_{n+1}^{-1}(U_{n+1} \mid X_1, \ldots, X_n), n \in \mathbb{N}$$

と定義する．$X_1 \sim \mu_1$ である．さらに，F_{n+1} の定義より，任意に固定した $x_1, \dots, x_n \in \mathbb{R}$ に対して，

$$P(F_{n+1}^{-1}(U_{n+1} \mid x_1, \dots, x_n) \in B) = \mu_{n+1,n}(x_1, \dots, x_n, B),\ B \in \mathcal{B}.$$

U_{n+1} と $(X_1, \dots, X_n)$ は独立であるから，任意の $A \in \mathcal{B}^n, B \in \mathcal{B}$ に対して，

$$\begin{aligned}
&P((X_1, \dots, X_n) \in A, X_{n+1} \in B) = P((X_1, \dots, X_n) \in A, F_{n+1}^{-1}(U_{n+1} \mid X_1, \dots, X_n) \in B) \\
&\quad = E[P(F_{n+1}^{-1}(U_{n+1} \mid X_1, \dots, X_n) \in B \mid X_1, \dots, X_n), \{(X_1, \dots, X_n) \in A\}] \\
&\quad = E[\mu_{n+1,n}(X_1, \dots, X_n, B), \{(X_1, \dots, X_n) \in A\}].
\end{aligned}$$

これは，$P(X_{n+1} \in B \mid X_1, \dots, X_n) = \mu_{n+1,n}(X_1, \dots, X_n, B)$ a.s. を意味するから，$\mu_{n+1,n}(x_1, \dots, x_n, dx_{n+1})$ は $(X_1, \dots, X_n)$ を与えたときの X_{n+1} の r.c.d. である．従って，$E \in \mathcal{B}^n$ に対して，Theorem 20.3 を繰り返し適用して，

$$\begin{aligned}
&P((X_1, \dots, X_n) \in E) \\
&= \iint \cdots \int 1_E(x_1, \dots, x_n) \mu_{n,n-1}(x_1, \dots, x_{n-1}, dx_n) \cdots \underbrace{\mu_{2,1}(x_1, dx_2) d\mu_1(x_1)}_{=d\mu_2(x_1,x_2)} \\
&= \iint \cdots \int 1_E(x_1, \dots, x_n) \mu_{n,n-1}(x_1, \dots, x_{n-1}, dx_n) \cdots \underbrace{\mu_{3,2}(x_1, x_2, dx_3) d\mu_2(x_1, x_2)}_{=d\mu_3(x_1,x_2,x_3)} \\
&\ \vdots \\
&= \int 1_E(x_1, \dots, x_n) d\mu_n(x_1, \dots, x_n) = \mu_n(E)
\end{aligned}$$

を得る．すなわち，$(X_1, \dots, X_n) \sim \mu_n$. そこで，$X = (X_1, X_2, \dots) : \Omega \to \mathbb{R}^{\mathbb{N}}$ の像測度を $\mu = P \circ X^{-1}$ とおけば，μ は $(**)$ をみたす． □

Remark 20.2. 上の証明では，独立な一様乱数 $(U_1, \dots, U_n)$ から $(X_1, \dots, X_n) \sim \mu_n$ を構成したが，この構成法は Knothe-Rosenblatt 再配分 (Knothe-Rosenblatt rearrangement) とも呼ばれる (cf. Villani, 2008, Chapter 1).

r.v.'s の集まり $\{X_m : m \in \mathbb{N}\}$ を 確率過程 (stochastic process) と呼び，$n \in \mathbb{N}$ に対して，$\mathcal{L}(X_1, \dots, X_n)$ を 有限次元分布 (finite dimensional distribution, f.d.d.) と呼ぶ．Kolmogorov の拡張定理は，整合的な p.m.'s $\mu_n, n \in \mathbb{N}$ を f.d.d.'s にもつ確率過程の存在を保証している．

Corollary 20.1. Theorem 20.4 は $(\mathbb{R}, \mathcal{B})$ を標準 Borel 空間 $(S, \mathcal{S})$ に取り替えても成り立つ．

しかし，直積確率測度の場合とは異なり，Kolmogorov の拡張定理は一般の可測空間 $(S, \mathcal{S})$ に対しては成り立たない．以下，Wegner (1973) に従い，そのような例を構成しよう[13]．Vitali 集合の構成において現れた記号を用いる．$r \in \mathbb{R}$ に対して，$T_r : [0,1) \to [0,1)$ を $T_r(x) = x + r - \lfloor x + r \rfloor$ と定義する．まず次の補題を証明する．

Lemma 20.2. 各 $r \in \mathbb{R}$ に対して，T_r は全単射であり，$T_r^{-1} = T_{-r}$.

Proof. $x, y \in [0,1)$ に対して，$T_r(x) = T_r(y)$ なら，

$$\underbrace{x - y}_{\in(-1,1)} = \underbrace{\lfloor x + r \rfloor - \lfloor y + r \rfloor}_{\text{整数}}$$

であるから，$x = y$ である．ゆえに T_r は単射であり，$T_r([0,1)) = [0,1)$ より，全射でもある．次に，$x \in [0,1), y = T_{-r}(x)$ に対して，

$$T_r(y) = y + r - \lfloor y + r \rfloor = x - \lfloor x - r \rfloor - \lfloor x - \lfloor x - r \rfloor \rfloor.$$

場合分けにより，最右辺は $= x$ であることがわかる． □

$V \subset [0,1)$ を Vitali 集合とし，$S = [0,1) \setminus V$ とする．また，S の σ-field として，$\mathcal{S} = \{B \cap S : B \in \mathcal{B}_{[0,1)}\}$ を入れる．$S^{\mathbb{N}}$ を S の可算直積とし，$\mathcal{S}^{\mathbb{N}}$ を $S^{\mathbb{N}}$ のシリンダー σ-field とする．さらに，各 $n \in \mathbb{N}$ に対して，$(S^n, \mathcal{S}^n)$ 上の p.m. μ_n を次のように構成する．$\mathbb{Q} = \{r_1, r_2, \dots\}$ と番号を付けて，$J_n = \{r_1, \dots, r_n\}$, $T_{J_n}(x) = (T_{r_1}(x), \dots, T_{r_n}(x))$ と定義する．写像 T_{J_n} は $\mathcal{B}_{[0,1)}/\mathcal{B}^n$ 可測である．λ を $[0,1)$ 上の Lebesgue 測度とし，λ^* と λ_* をそれぞれ λ の外測度，内測度とする．次の補題を示す．

Lemma 20.3. $\lambda^*(T_{J_n}^{-1}(S^n)) = 1$.

Proof. $\lambda_*([0,1) \setminus T_{J_n}^{-1}(S^n)) = 0$ を示せばよい．$T_{J_n}^{-1}(S^n) = \bigcap_{i=1}^n T_{r_i}^{-1}(S) = \bigcap_{i=1}^n T_{r_i}^{-1}([0,1) \setminus V) = \bigcap_{i=1}^n ([0,1) \setminus T_{r_i}^{-1}(V)) = [0,1) \setminus \bigcup_{i=1}^n T_{r_i}^{-1}(V)$ より，

$$[0,1) \setminus T_{J_n}^{-1}(S^n) = \bigcup_{i=1}^n T_{r_i}^{-1}(V) = \bigcup_{i=1}^n T_{-r_i}(V)$$

である．よって，$\lambda_*(\bigcup_{i=1}^n T_{-r_i}(V)) = 0$ を示せばよい．そのためには，任意の Borel 集合 $E \subset \bigcup_{i=1}^n T_{-r_i}(V)$ に対して，$\lambda(E) = 0$ を示せばよいが，Steinhaus の定理 (cf. Exercise 7.9) より，$E - E$ が開区間を含まないことを示せば十分である．

$F = \bigcup_{i=1}^n T_{-r_i}(V)$ とおくと，$F - F$ の各要素は，

$$x_i - r_i - (x_j - r_j) - \lfloor x_i - r_i \rfloor + \lfloor x_j - r_j \rfloor \quad (x_i, x_j \in V; i, j = 1, \dots, n)$$

[13] 以下の部分は今泉允聡氏の協力があった．どうもありがとうございます．

の形をしている．これが有理数となるのは，$x_i - x_j \in \mathbb{Q}$ のときのみであるが，V の定義より，$x_i = x_j$ である．従って，

$$(F-F)\cap\mathbb{Q} = \{-r_i + r_j - \lfloor x - r_i \rfloor + \lfloor x - r_j \rfloor : x \in V,\ i,j = 1,\dots,n\}$$

であるが，右辺は有限集合である．従って，$E - E$ $(\subset F - F)$ は開区間を含みえない． □

ここで，$\mathcal{S}^n = \{B \cap S^n : B \in \mathcal{B}^n, B \subset [0,1)^n\}$ と表せるので，写像 $\mu_n : \mathcal{S}^n \to [0,1]$ を，$A = B \cap S^n, B \in \mathcal{B}^n, B \subset [0,1)^n$ に対して，

$$\mu_n(A) = \lambda(T_{J_n}^{-1}(B))$$

と定義する．Lemma 20.3 と Exercise 6.11 より，μ_n は $(S^n, \mathcal{S}^n)$ 上の well-defined な p.m. である．

以上で，目標である次の定理を述べる準備が整った．

Theorem 20.5 (Wegner (1973))**.** $\{\mu_n : n \in \mathbb{N}\}$ は Kolmogorov の整合性条件をみたす p.m.'s の族であるが，

$$\mu(A \times S \times S \times \cdots) = \mu_n(A),\ \forall A \in \mathcal{S}^n, \forall n \in \mathbb{N} \tag{$*3$}$$

をみたす $(S^{\mathbb{N}}, \mathcal{S}^{\mathbb{N}})$ 上の p.m. μ は存在しない．

Proof. $\{\mu_n : n \in \mathbb{N}\}$ が整合性条件をみたすのは明らか．$(*3)$ をみたす $(S^{\mathbb{N}}, \mathcal{S}^{\mathbb{N}})$ 上の p.m. μ が仮に存在したとして，矛盾を導く．ここで，$D_n = T_{J_n}([0,1)) \cap S^n$ とおくと，$T_{J_n}([0,1)) \in \mathcal{B}^n$ より $D_n \in \mathcal{S}^n$ である．$A_n = D_n \times S \times S \times \cdots \subset S^{\mathbb{N}}$ とおくと，$A_n \in \mathcal{S}^{\mathbb{N}}, A_n \supset A_{n+1}$ $(\forall n \in \mathbb{N})$ である．$\bigcap_{n=1}^{\infty} A_n = \varnothing$ を示す．仮に $\bigcap_{n=1}^{\infty} A_n \neq \varnothing$ ならば，ある $x \in [0,1)$ が存在して，$T_{J_n}(x) \in S^n$ $(\forall n \in \mathbb{N})$ となる．これは，$x \in T_r^{-1}(S)$ $(\forall r \in \mathbb{Q})$ を意味するが，$T_r^{-1} = T_{-r}$ であったから，$x \in \bigcap_{r\in\mathbb{Q}} T_r(S)$ である．しかし，

$$\bigcap_{r\in\mathbb{Q}} T_r(S) = \bigcap_{r\in\mathbb{Q}} T_r([0,1) \setminus V) = [0,1) \setminus \bigcup_{r\in\mathbb{Q}} T_r(V) = \varnothing$$

であるから，矛盾が生じる．従って，$A_n \downarrow \varnothing$ である．

いま，μ を $(*3)$ をみたす $(S^{\mathbb{N}}, \mathcal{S}^{\mathbb{N}})$ 上の p.m. とすると，

$$\mu(A_n) = \mu_n(D_n) = \lambda(T_{J_n}^{-1}(T_{J_n}([0,1))) = \lambda([0,1)) = 1$$

となるが，μ が p.m. であることと，$A_n \downarrow \varnothing$ より，これは起こりえない．従って，$(*3)$ をみたす $(S^{\mathbb{N}}, \mathcal{S}^{\mathbb{N}})$ 上の p.m. は存在しない． □

演習問題

Exercise 20.1 (r.c.d. の一意性)**.** $(S,\mathcal{S})$ を標準 Borel 空間とする．X を S-valued r.v. とし，$\mathcal{G}$ を $\mathcal{F}$ の sub σ-field とする．このとき，$\mu,\nu:\Omega\times\mathcal{S}\to[0,1]$ がともに $\mathcal{G}$ を与えたときの X の r.c.d. なら，$P(\{\omega:\mu(\omega,B)=\nu(\omega,B)\ \forall B\in\mathcal{S}\})=1$ となることを示せ．

Exercise 20.2. X,Y を r.v.'s とし，同時密度関数 $f(x,y)$ をもつとする．すなわち，ある可測関数 $f:\mathbb{R}^2\to\mathbb{R}_+$ が存在して，

$$P((X,Y)\in E)=\int_E f(x,y)dxdy,\ E\in\mathcal{B}^2$$

とする．簡単のため，$f(x,y)>0$ $(\forall(x,y)\in\mathbb{R}^2)$ とし，

$$\mu(x,B)=\frac{\int_B f(x,y)dy}{\int_{\mathbb{R}} f(x,y)dy},\ B\in\mathcal{B}$$

とおくと，$\mathbb{R}\times\mathcal{B}\ni(x,B)\mapsto\mu(x,B)$ は X を与えたときの Y の r.c.d. であることを示せ．

Exercise 20.3 (Vorob'ev-Berkes-Philipp)**.** $(S,\mathcal{S}),(T,\mathcal{T}),(U,\mathcal{U})$ を標準 Borel 空間とし，μ,ν をそれぞれ $(S\times T,\mathcal{S}\times\mathcal{T}),(T\times U,\mathcal{T}\times\mathcal{U})$ 上の p.m.'s であって，$\mu(S\times B)=\nu(B\times U)$ $(\forall B\in\mathcal{T})$ をみたすとする．いま，$\pi_{12}:S\times T\times U\to S\times T,\pi_{23}:S\times T\times U\to T\times U$ を $\pi_{12}(x,y,z)=(x,y),\pi_{23}(x,y,z)=(y,z)$ とすると，次をみたす $(S\times T\times U,\mathcal{S}\times\mathcal{T}\times\mathcal{U})$ 上の p.m. ρ が存在することを示せ：$\rho\circ\pi_{12}^{-1}=\mu,\rho\circ\pi_{23}^{-1}=\nu$．

Exercise 20.4 (Kolmogorov の拡張定理の一般の添え字集合への拡張)**.** T を空でない一般の添え字集合とし，$\mathbb{R}^T$ を T から $\mathbb{R}$ への関数全体とする．ここで，相異なる $t_1,\dots,t_n\in T,n\in\mathbb{N}$ に対して，

$$\{x\in\mathbb{R}^T:(x(t_1),\dots,x(t_n))\in B\},\ B\in\mathcal{B}^n$$

の形の集合を $\mathbb{R}^T$ のシリンダー集合と呼ぶ．$\mathbb{R}^T$ のシリンダー集合全体によって生成される σ-field を $\mathcal{B}^T$ とおく．いま，任意の $n\in\mathbb{N}$ と相異なる $t_1,\dots,t_n\in T$ の各組に対して，$(\mathbb{R}^n,\mathcal{B}^n)$ 上の p.m. $\mu_{t_1,\dots,t_n}$ が与えられていて，次の整合性条件がみたされるとする：

- 任意の相異なる $t_1,\dots,t_n\in T$ と $\{1,\dots,n\}$ の任意の置換 π に対して，

$$\mu_{t_1,\dots,t_n}(\{(x_1,\dots,x_n):(x_{\pi(1)},\dots,x_{\pi(n)})\in B\})=\mu_{t_{\pi(1)},\dots,t_{\pi(n)}}(B),\ \forall B\in\mathcal{B}^n.$$

- 任意の相異なる $t_1,\dots,t_n,t_{n+1}\in T$ に対して，

$$\mu_{t_1,\dots,t_n,t_{n+1}}(B\times\mathbb{R})=\mu_{t_1,\dots,t_n}(B),\ \forall B\in\mathcal{B}^n.$$

このとき，次をみたす $(\mathbb{R}^T, \mathcal{B}^T)$ 上の一意な p.m μ が存在する：任意の $n \in \mathbb{N}$ と相異なる $t_1, \dots, t_n \in T$ に対して，

$$\mu(\{x \in \mathbb{R}^T : (x(t_1), \dots, x(t_n)) \in B\}) = \mu_{t_1,\dots,t_n}(B),\ \forall B \in \mathcal{B}^n. \qquad (*)$$

この命題を以下の手順に従って証明する．

(a) 各可算集合 $S = \{t_1, t_2, \dots\} \subset T$ に対して，

$$\mathcal{F}_S = \{x \in \mathbb{R}^T : (x(t_1), x(t_2), \dots) \in B,\ B \in \mathcal{B}^{\mathbb{N}}\}$$

とおいたとき，$\mathcal{F}_S$ は σ-field であって，$\mathcal{B}^T = \bigcup_{S \subset T:\text{可算}} \mathcal{F}_S$ が成り立つことを示せ．

(b) Theorem 20.4 を用いて，各可算集合 $S \subset T$ に対して，$(\mathbb{R}^T, \mathcal{F}_S)$ 上の p.m. μ_S であって，任意の $n \in \mathbb{N}$ と相異なる $t_1, \dots, t_n \in S$ に対して，

$$\mu_S(\{x \in \mathbb{R}^T : (x(t_1), \dots, x(t_n)) \in B\}) = \mu_{t_1,\dots,t_n}(B),\ \forall B \in \mathcal{B}^n$$

をみたすものが存在することを示せ．

(c) 写像 $\mu : \mathcal{B}^T \to [0, 1]$ を

$$\mu(A) = \mu_S(A) \quad \text{if } A \in \mathcal{F}_S,\ S \subset T\text{: 可算}$$

と定義すると，μ は $(\mathbb{R}^T, \mathcal{B}^T)$ 上の well-defined な p.m. であって，$(*)$ をみたすことを示せ．

(d) $(*)$ をみたす $(\mathbb{R}^T, \mathcal{B}^T)$ 上の p.m. は一意であることを示せ．

T を空でない集合とする．r.v.'s の集まり $\{X_t : t \in T\}$ を T によってインデックスされた確率過程と呼ぶ．また，確率過程 $\{X_t : t \in T\}$ が Gauss 過程であるとは，任意の $n \in \mathbb{N}$ と $t_1, \dots, t_n \in T$ に対して，$(X(t_1), \dots, X(t_n))$ が多変量正規分布に従うことを言う．

Exercise 20.5. $k : T \times T \to \mathbb{R}$ を次をみたす関数とする：任意の $n \in \mathbb{N}$ と $t_1, \dots, t_n \in T$ に対して，n 次正方行列 $((k(t_i, t_j))_{1 \le i,j \le n}$ が半正定値対称行列になる．このとき，k を共分散関数にもつ平均 0 の Gauss 過程 $\{X_t : t \in T\}$ が存在することを示せ．すなわち，

$$E[X_t] = 0,\ \forall t \in T, \quad E[X_s X_t] = k(s, t),\ \forall s, t \in T$$

をみたす Gauss 過程 $\{X_t : t \in T\}$ が存在することを示せ．

Exercise 20.6. $C(\mathbb{R}_+)$ を $\mathbb{R}_+$ 上の連続関数全体とすると，$C(\mathbb{R}_+) \notin \mathcal{B}^{\mathbb{R}_+}$ を示せ．

21　マルチンゲールと a.s. 収束

$\mathbb{Z}_+ = \{0, 1, 2, \dots\}$ とする．$\mathcal{F}$ の sub σ-fields の増加列 $\{\mathcal{F}_n : n \in \mathbb{Z}_+\}$：

$$\mathcal{F}_0 \subset \mathcal{F}_1 \subset \cdots$$

を フィルトレーション (filtration) と呼ぶ．確率過程 $\{X_n : n \in \mathbb{Z}_+\}$ が $\{\mathcal{F}_n\}$ に 適合 (adapted) しているとは，$X_n \in \mathcal{F}_n$ $(\forall n \in \mathbb{Z}_+)$ となることを言う．フィルトレーション $\{\mathcal{F}_n\}$ が与えられたとき，$\{X_n\}$ が $\{\mathcal{F}_n\}$ に関する マルチンゲール (martingale) であるとは，次の 3 条件がみたされることを言う：

(i) $E[|X_n|] < \infty$ $\forall n \in \mathbb{Z}_+$.

(ii) $\{X_n\}$ は $\{\mathcal{F}_n\}$ に適合している．

(iii) $E[X_{n+1} \mid \mathcal{F}_n] = X_n$ a.s. $\forall n \in \mathbb{Z}_+$.

(iii) の代わりに，

$$E[X_{n+1} \mid \mathcal{F}_n] \geq X_n \text{ a.s. } \forall n \in \mathbb{Z}_+$$

が成り立つとき，$\{X_n\}$ を $\{\mathcal{F}_n\}$ に関する 劣マルチンゲール (sub-martingale) と呼ぶ．また，

$$E[X_{n+1} \mid \mathcal{F}_n] \leq X_n \text{ a.s. } \forall n \in \mathbb{Z}_+$$

が成り立つとき，$\{X_n\}$ を $\{\mathcal{F}_n\}$ に関する 優マルチンゲール (super-martingale) と呼ぶ．以下，一般論を議論するときは，特に断らない限り，フィルトレーションは $\{\mathcal{F}_n\}$ に固定し，“$\{\mathcal{F}_n\}$ に関する” という部分をしばしば省略する．

$\{X_n\}$ が劣 (優) マルチンゲールなら，$E[X_{n+1}] \geq (\leq)\ E[X_n]$ となる．従って，劣 (優) マルチンゲールは平均的に増加 (減少) していく過程である．

Example 21.1. (マルチンゲールの例).

- $\{\xi_m\}$ を独立な r.v.'s とし，

$$X_0 = x \in \mathbb{R},\ X_n = X_{n-1} + \xi_n = x + \sum_{m=1}^{n} \xi_m,\ \mathcal{F}_0 = \{\varnothing, \Omega\},\ \mathcal{F}_n = \sigma(\xi_1, \dots, \xi_n)$$

とすると，$E[\xi_m]$ $(> 0, < 0)$ $= 0$ $\forall m$ なら $\{X_n\}$ は (劣/優) マルチンゲールである．

- $\{\xi_m\}$ を独立な非負 r.v.'s とし，

$$X_0 = x \in \mathbb{R}_+,\ X_n = X_{n-1}\xi_n = x \prod_{m=1}^{n} \xi_m,\ \mathcal{F}_0 = \{\varnothing, \Omega\},\ \mathcal{F}_n = \sigma(\xi_1, \dots, \xi_n)$$

とすると，$E[\xi_m]$ $(> 1, < 1)$ $= 1$ $\forall m$ なら $\{X_n\}$ は (劣/優) マルチンゲールである．

$n > m \geq 0$ に対して，$\mathcal{F}_m \subset \mathcal{F}_{n-1}$ であるから，$\{X_n\}$ が劣マルチンゲールなら，

$$E[X_n \mid \mathcal{F}_m] = E[E[X_n \mid \mathcal{F}_{n-1}] \mid \mathcal{F}_m] \geq E[X_{n-1} \mid \mathcal{F}_m] \geq \cdots \geq E[X_{m+1} \mid \mathcal{F}_m] \geq X_m$$

となる．他の場合も同様に考えて，次の補題を得る．

Lemma 21.1. $\{X_n\}$ が (劣/優) マルチンゲールなら，$n > m \geq 0$ に対して，それぞれ，

$$E[X_n \mid \mathcal{F}_m] \ (\geq, \leq) \ = X_m \quad a.s.$$

となる．

次の補題は Jensen の不等式から明らかである．

Lemma 21.2. $\varphi : \mathbb{R} \to \mathbb{R}$ を凸関数とし，$\{X_n\}$ を $E[|\varphi(X_n)|] < \infty$ $(\forall n)$ なる確率過程とする．(i) $\{X_n\}$ がマルチンゲールなら，$\{\varphi(X_n) : n \in \mathbb{Z}_+\}$ は劣マルチンゲールである．(ii) $\{X_n\}$ が劣マルチンゲールのときは，φ が非減少なら，$\{\varphi(X_n) : n \in \mathbb{Z}_+\}$ は劣マルチンゲールである．

Example 21.2. (i) $\{X_n\}$ をマルチンゲールとする．このとき，与えられた $1 \leq p < \infty$ に対して，$E[|X_n|^p] < \infty$ $(\forall n \in \mathbb{Z}_+)$ なら，$\{|X_n|^p : n \in \mathbb{Z}_+\}$ は劣マルチンゲールである．(ii) $\{X_n\}$ が劣マルチンゲールなら，$\{(X_n - a)^+ : n \in \mathbb{Z}_+\}$ は劣マルチンゲールである．(iii) $\{X_n\}$ が優マルチンゲールなら，$\{X_n \wedge a : n \in \mathbb{Z}_+\}$ は優マルチンゲールである．

Remark 21.1. $\max\{-x, 0\}$ は凸関数であるが非増加であるから，劣マルチンゲール $\{X_n\}$ に対して，$\{X_n^-\}$ は劣/優マルチンゲールのどちらでもない．$\{X_n\}$ が非正なら $X_n^- = -X_n$ であるから，$\{X_n^-\}$ は優マルチンゲールであるし，$\{X_n\}$ がマルチンゲールなら $\{X_n^-\}$ は劣マルチンゲールである．また，劣マルチンゲール $\{X_n\}$ に対して，$|X_n| = X_n^+ + X_n^-$ は劣/優マルチンゲールのどちらでもない．

フィルトレーション $\{\mathcal{F}_n\}$ に対して，r.v.'s $\{H_n : n \in \mathbb{N}\}$ が 可予測 (predictable) であるとは，$H_n \in \mathcal{F}_{n-1}$ $(\forall n \in \mathbb{N})$ となることを言う．$\{X_n\}$ を (劣/優) マルチンゲールとすると，$\{H_n\}$ の $\{X_n\}$ に関する マルチンゲール変換 (martingale transform) を

$$(H \cdot X)_n = \sum_{m=1}^{n} H_m(X_m - X_{m-1}), \ n \in \mathbb{N}, \ (H \cdot X)_0 = 0$$

と定義する．このとき，$(H \cdot X)_{n+1} - (H \cdot X)_n = H_{n+1}(X_{n+1} - X_n)$, $H_{n+1} \in \mathcal{F}_n$ であることから，次の定理を得る．

Theorem 21.1. $E[|(H \cdot X)_n|] < \infty$ $(\forall n \in \mathbb{N})$ とする．(i) $\{X_n\}$ がマルチンゲールなら，$\{(H \cdot X)_n : n \in \mathbb{Z}_+\}$ もマルチンゲールである． (ii) $H_n \geq 0$ かつ $\{X_n\}$ が劣/優マルチンゲールなら，$\{(H \cdot X)_n : n \in \mathbb{Z}_+\}$ も劣/優マルチンゲールである．

Example 21.3 (コイン投げ)**.** 1 ドルの賭け金に対して，表 ($\xi_n = 1$) が出たら，1 ドルもらえて，裏 ($\xi_n = -1$) が出たら，1 ドル失うゲームを考える．コインは公平とし，$P(\xi_n = \pm 1) = 1/2$ とする．また，ゲームは互いに独立とし，n 回目のゲームにおいて，ギャンブラーはそれまでの結果 $\xi_1, \dots, \xi_{n-1}$ にもとづいて，H_n ドル賭けるか決めるとする．n 回目までの勝ち負けの合計の差を $X_n = \sum_{m=1}^{n} \xi_m$ とおくと ($X_0 = 0$ とおく)，n 回目のゲーム終了時の儲けは

$$\sum_{m=1}^{n} H_m \xi_m = \sum_{m=1}^{n} H_m (X_m - X_{m-1}) = (H \cdot X)_n$$

である．H_n は賭けの戦略と考えることができる．有名な賭けの戦略として，$(n-1)$ 回目に賭けに負けたら n 回目の賭け金を倍にする倍賭け戦略 (doubling strategy) がある：

$$H_1 = 1,\ H_n = \begin{cases} 2H_{n-1} & \text{if } \xi_{n-1} = -1 \\ 1 & \text{if } \xi_{n-1} = 1 \end{cases},\ n \geq 2.$$

この戦略に従うと，最初の k 回のゲームに負けて次のゲームに勝った場合，儲けは

$$-1 - 2 - \cdots - 2^{k-1} + 2^k = 1$$

となるから，必ず儲かりそうな気もするが，$\{(H \cdot X)_n\}$ はマルチンゲールであるから，

$$E[(H \cdot X)_n] = E[(H \cdot X)_0] = 0.$$

つまり，平均的な儲けは 0 である！

関数 $N : \Omega \to \{0, 1, 2, \dots\} \cup \{\infty\}$ が $\{\mathcal{F}_n\}$ に関する 停止時刻 (stopping time) であるとは，$\{N = n\} \in \mathcal{F}_n$ $(0 \leq \forall n < \infty)$ となることを言う．$\mathcal{F}_n \subset \mathcal{F}_{n+1}$ より，

$$\{N = n\} \in \mathcal{F}_n,\ 0 \leq \forall n < \infty \Leftrightarrow \{N \leq n\} \in \mathcal{F}_n,\ 0 \leq \forall n < \infty$$

である．例えば，$\{X_n\}$ を適合過程とすると，任意の $B \in \mathcal{B}$ に対して，

$$N_B = \inf\{n : X_n \in B\}$$

を $\{X_n\}$ の B への 到達時刻 (hitting time) と呼ぶ．ただし，$\inf \varnothing = +\infty$ とする．ここで，

$$\{N_B = n\} = \{X_m \notin B\ \forall m < n,\ X_n \in B\} \in \mathcal{F}_n$$

であるから，N_B は停止時刻である．

いま，停止時刻 N が与えられたとき，$H_n = 1_{\{N \geq n\}}$ とおくと，$\{N \geq n\} = \{N \leq n-1\}^c \in \mathcal{F}_{n-1}$ より，$\{H_n\}$ は可予測である．$\{H_n\}$ のマルチンゲール変換は

$$(H \cdot X)_n = \sum_{m=1}^{N \wedge n} (X_m - X_{m-1}) = X_{N \wedge n} - X_0$$

であるから，次の (重要な！) 系を得る．

Corollary 21.1. N を停止時刻とする．$\{X_n\}$ が (劣/優) マルチンゲールなら，$\{X_{N\wedge n} : n \in \mathbb{Z}_+\}$ も (劣/優) マルチンゲールである．

Remark 21.2. もちろん，定義に従って $\{X_{N\wedge n}\}$ が (劣/優) マルチンゲールであることを示すのも難しくない．$\{X_n\}$ をマルチンゲールとしよう．まず，$|X_{N\wedge n}| \le \max_{0\le m\le n} |X_m|$ であるから，各 $X_{N\wedge n}$ は可積分である．次に，$B \in \mathcal{B}$ に対して，

$$\begin{aligned}\{X_{N\wedge n} \in B\} &= \bigcup_{k\in\mathbb{Z}_+\cup\{\infty\}} (\{X_{k\wedge n} \in B\} \cap \{N = k\}) \\ &= \bigcup_{k=0}^{n-1} (\{X_k \in B\} \cap \{N = k\}) \cup (\{X_n \in B\} \cap \{N \ge n\}) \in \mathcal{F}_n\end{aligned}$$

であるから，$\{X_{N\wedge n}\}$ は適合している．また，$A \in \mathcal{F}_n$ に対して，

$$A \cap \underbrace{\{N \ge n+1\}}_{=\{N\le n\}^c} \in \mathcal{F}_n$$

であるから，

$$\begin{aligned}E[X_{N\wedge(n+1)}, A] &= \sum_{k=0}^{n} E[X_k, A \cap \{N = k\}] + \underbrace{E[X_{n+1}, A \cap \{N \ge n+1\}]}_{=E[X_n, A\cap\{N\ge n+1\}]} \\ &= E[X_{N\wedge n}, A].\end{aligned}$$

よって，$E[X_{N\wedge(n+1)} \mid \mathcal{F}_n] = X_{N\wedge n}$ a.s. を得る．他の場合も同様である．

$\{X_n : n \in \mathbb{Z}_+\}$ を適合過程とする．$a < b, N_0 = -1$ として，$k \in \mathbb{N}$ に対して，

$$N_{2k-1} = \inf\{m > N_{2k-2} : X_m \le a\},\ N_{2k} = \inf\{m > N_{2k-1} : X_m \ge b\}$$

と定義する．帰納的に確かめられるように，各 k に対して，N_{2k-1}, N_{2k} は停止時刻である．

このとき，$\{N_{2k-1} < m \le N_{2k}\} = \{N_{2k-1} \le m-1\} \cap \{N_{2k} \le m-1\}^c \in \mathcal{F}_{m-1}$ より，

$$H_m = \begin{cases} 1 & \text{if } N_{2k-1} < m \le N_{2k}, \exists k \in \mathbb{N} \\ 0 & \text{otherwise} \end{cases}$$

は可予測である．ここで，

$$U_n = \sup\{k : N_{2k} \le n\}$$

とおくと，U_n は $m \le n$ までに X_m が $[a, b]$ を上向きに横断 (upcross) する回数である．

Theorem 21.2 (上向き横断不等式)**.** $\{X_n\}$ が劣マルチンゲールなら，各 n に対して，

$$(b-a)E[U_n] \le E[(X_n - a)^+] - E[(X_0 - a)^+].$$

Proof. $Y_m = a + (X_m - a)^+$ とおくと，$\{Y_m\}$ も劣マルチンゲールである．ここで，

$$\sum_{m=N_{2k-1}+1}^{N_{2k}} (Y_m - Y_{m-1}) = Y_{N_{2k}} - Y_{N_{2k-1}} \geq b - a$$

となるから，

$$(H \cdot Y)_n \geq \sum_{k=1}^{U_n} \sum_{m=N_{2k-1}+1}^{N_{2k}} (Y_m - Y_{m-1}) \geq (b-a)U_n.$$

いま，$K_m = 1 - H_m$ とおくと，$Y_n - Y_0 = (H \cdot Y)_n + (K \cdot Y)_n$ であり，$\{(K \cdot Y)_n\}$ は劣マルチンゲールであるから，$E[(K \cdot Y)_n] \geq E[(K \cdot Y)_0] = 0$ となる．従って，$E[(H \cdot Y)_n] \leq E[Y_n - Y_0]$ を得る． □

上向き横断不等式から，次のマルチンゲール収束定理を得る．

Theorem 21.3 (マルチンゲール収束定理)**.** $\{X_n\}$ を劣マルチンゲールとし，$\sup_n E[X_n^+] < \infty$ とする．このとき，$\exists X_\infty \in L^1$ s.t. $X_n \to X$ a.s.

Proof. 上向き横断不等式より，

$$E[U_n] \leq \frac{|a| + E[X_n^+]}{b - a}.$$

$0 \leq U_n \uparrow U_\infty := \sup_m U_m$ であるから，Fatou の補題より，$E[U_\infty] < \infty$．よって，$U_\infty < \infty$ a.s. であるから，X_n が $[a, b]$ を上向きに横断する回数は a.s. に有限であり，$P(\{X_n < a \text{ i.o.}\} \cap \{X_n > b \text{ i.o.}\}) = 0$ となる．$a < b$ は任意だったから，

$$P\left(\bigcup_{a,b \in \mathbb{Q}, a<b} (\{X_n < a \text{ i.o.}\} \cap \{X_n > b \text{ i.o.}\}) \right) = 0$$

を得る．これは，

$$\limsup_n X_n = \liminf_n X_n \quad a.s.$$

を意味するので，a.s. に $\lim_n X_n$ が $[-\infty, \infty]$ の範囲で存在する．そこで，

$$X_\infty(\omega) = \begin{cases} \lim_n X_n(\omega) & \text{if } \limsup_n X_n(\omega) = \liminf_n X_n(\omega) \\ 0 & \text{otherwise} \end{cases}$$

とおくと，X_∞ は $\overline{\mathbb{R}}$-valued r.v. である．$E[|X_\infty|] < \infty$ を示す．Fatou の補題より，$E[X_\infty^+] \leq \liminf_n E[X_n^+] < \infty$．一方，$\{X_n\}$ は劣マルチンゲールであるから，$E[X_n^-] = E[X_n^+] - E[X_n] \leq E[X_n^+] - E[X_0]$．再び Fatou の補題より，$E[X_\infty^-] \leq \liminf_n E[X_n^+] - E[X_0] < \infty$ を得る． □

Corollary 21.2. $\{X_n\}$ が非負優マルチンゲールなら (i.e., $\{X_n\}$ が優マルチンゲールであって，$X_n \geq 0$ a.s. $\forall n$ なら)，X_n はある r.v. $X_\infty \geq 0$ に a.s. 収束する：$X_n \to X_\infty$ a.s. さらに，$E[X_\infty] \leq E[X_0]$ である．

Proof. $\{-X_n\}$ は劣マルチンゲールであって，$(-X_n)^+ = 0$ であるから，マルチンゲール収束定理より，ある r.v. $X_\infty \geq 0$ が存在して，$X_n \to X_\infty$ a.s. となる．また，Fatou の補題より，$E[X_\infty] \leq \liminf_n E[X_n] \leq E[X_0]$ となる． □

Example 21.4 (重要な例)**.** マルチンゲール収束定理において，$X_n \to X_\infty$ in L^1 は必ずしも成り立たない．$\{S_n : n \in \mathbb{Z}_+\}$ を初期値が 0 の単純ランダム・ウォークとする．すなわち，$\{\xi_m\}$ を i.i.d. r.v.'s とし，$P(\xi_m = \pm 1) = 1/2, \mathcal{F}_0 = \{\varnothing, \Omega\}, \mathcal{F}_n = \sigma(\xi_1, \ldots, \xi_n), S_0 = 1, S_n = S_{n-1} + \xi_n = 1 + \sum_{m=1}^n \xi_m$ とする．$N = \inf\{n : S_n = 0\}$ とおくと，N は停止時刻である．このとき，$X_n := S_{N \wedge n}$ は非負マルチンゲールであるから，X_n は a.s. に収束する．$N = \infty$ なら，$X_n = S_n$ なので，$|X_{n+1} - X_n| = 1$ となり，X_n は収束しない．従って，$N < \infty$ a.s. であるから，$X_n \to 0$ a.s. となる．一方，$E[X_n] = E[X_0] = 1$ より，$X_n \not\to 0$ in L^1 である．なお，$\{X_n\}$ は L^1 有界だが一様可積分でない例でもある．

Theorem 21.4 (Doob 分解)**.** 任意の劣マルチンゲール $\{X_n : n \in \mathbb{Z}_+\}$ に対して，適当なマルチンゲール $\{M_n : n \in \mathbb{Z}_+\}$ と初期値が 0 の可予測な増加過程 $\{A_n : n \in \mathbb{Z}_+\}$ (i.e., $\{A_n : n \in \mathbb{N}\}$ は可予測で，a.s. に $A_0 = 0, A_n \leq A_{n+1}$ $(\forall n \in \mathbb{Z}_+)$ となる) が存在して，a.s. に $X_n = M_n + A_n$ $(\forall n \in \mathbb{Z}_+)$ なる分解が成り立つ．さらに，$X_n = M'_n + A'_n$ をもう 1 つのそのような分解とすると，a.s. に $M'_n = M_n, A'_n = A_n$ $(\forall n \in \mathbb{Z}_+)$ となる．

Proof. 存在は，$A_0 = 0, M_0 = X_0$ とし，A_n, M_n を帰納的に

$$A_n - A_{n-1} = \underbrace{E[X_n \mid \mathcal{F}_{n-1}] - X_{n-1}}_{\geq 0}, \ M_n = X_n - A_n, \ n \in \mathbb{N}$$

と決めればよい．

次に一意性を示す．$X_n = M'_n + A'_n$ をもう 1 つの分解とすると，

$$\begin{aligned} E[X_n \mid \mathcal{F}_{n-1}] - X_{n-1} &= E[X_n - X_{n-1} \mid \mathcal{F}_{n-1}] \\ &= \underbrace{E[M'_n \mid \mathcal{F}_{n-1}]}_{= M'_{n-1}} - M'_{n-1} + A'_n - A'_{n-1} = A'_n - A'_{n-1} \end{aligned}$$

となるから，$A'_n = A_n$ (従って，$M'_n = M_n$) である． □

演習問題

Exercise 21.1. $\{X_n\}$ を劣マルチンゲールとし，$E[X_n] = E[X_0]$ $(\forall n \in \mathbb{N})$ をみたすとする．このとき，$\{X_n\}$ はマルチンゲールであることを示せ．

Exercise 21.2. M, N を停止時刻とすると，$M \wedge N, M \vee N, M + N$ も停止時刻であることを示せ．

Exercise 21.3. 劣マルチンゲール $\{X_n\}$ に対して，$\sup_n E[X_n^+] < \infty \Leftrightarrow \sup_n E[|X_n|] < \infty$ を示せ．

Exercise 21.4 (Azuma の不等式)**.** この問題の目標は，次の Azuma の不等式 を証明することである．

Theorem 21.5 (Azuma (1967))**.** $\{X_n\}$ をマルチンゲールとし，ある数列 $c_n > 0$ が存在して，$|X_n - X_{n-1}| \leq c_n$ a.s. $(\forall n \in \mathbb{N})$ とする．このとき，任意の $r > 0$ に対して，

$$P(X_n - X_0 \geq r) \leq e^{-\frac{r^2}{2\sum_{i=1}^n c_i^2}} \qquad (*)$$

が成り立つ．

以下の手順に従い，定理を証明せよ．

(a) ξ を $\xi \in [-1, 1]$ なる r.v. とし，$E[\xi] = 0$ を仮定する．このとき，任意の $t \in \mathbb{R}$ に対して，$E[e^{t\xi}] \leq e^{t^2/2}$ を示せ．

ヒント：$E[e^{t\xi}] \leq (e^t + e^{-t})/2$ を示せ．

(b) (a) の結果を使って，$E[e^{t(X_n - X_0)} \mid \mathcal{F}_{n-1}] \leq e^{t(X_{n-1} - X_0)} e^{t^2 c_n^2/2}$ を示せ．

(c) (b) の結果を使って，$(*)$ を導出せよ．

なお，$\{X_n - X_{n-1}\}$ が独立なとき，Azuma の不等式は Hoefdding の不等式と呼ばれる．そのため，Azuma の不等式のことを Azuma-Hoeffding の不等式と呼ぶこともある．

Exercise 21.5 (McDiarmid の不等式)**.** 各 $i = 1, \ldots, n$ に対して，X_i を可測空間 $(S_i, \mathcal{S}_i)$ に値をとる r.v. とし，$X_1, \ldots, X_n$ は独立とする．$f : \prod_{i=1}^n S_i \to \mathbb{R}$ を $(\prod_{i=1}^n \mathcal{S}_i)/\mathcal{B}$ 可測な関数であって，ある定数 $c_1, \ldots, c_n > 0$ が存在して，任意の $i = 1, \ldots, n$ と $x_i, x_i' \in S_i$ に対して，

$$\sup_{x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n} |f(x_1, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_n) - f(x_1, \ldots, x_{i-1}, x_i', x_{i+1}, \ldots, x_n)| \leq c_i$$

とする．このとき，$Y = f(X_1, \ldots, X_n)$ は可積分であって，任意の $r > 0$ に対して，

$$P(Y \geq E[Y] + r) \leq e^{-\frac{r^2}{2\sum_{i=1}^n c_i^2}}$$

が成り立つことを示せ．

ヒント：Azuma の不等式．

22　マルチンゲールの L^p 収束

Lemma 22.1 (有界な停止時刻に対する任意停止定理)**.** $\{X_n\}$ を劣マルチンゲールとし, N を $N(\omega) \leq k$ $(\forall \omega \in \Omega)$ なる停止時刻とすると，$E[X_0] \leq E[X_N] \leq E[X_k]$ である．

Proof. $|X_N| \leq \max_{0\leq n\leq k} |X_n|$ より，X_N は可積分である．次に，$\{X_{N\wedge n} : n \in \mathbb{Z}_+\}$ は劣マルチンゲールであるから，$E[X_0] = E[X_{N\wedge 0}] \leq E[X_{N\wedge k}] = E[X_N]$ となる．$H_n = 1_{\{N\leq n-1\}}$ とおくと，$\{H_n\}$ は可予測なので，$(H\cdot X)_n = X_n - X_{N\wedge n}$ は劣マルチンゲールである．従って，$E[X_k] - E[X_N] = E[(H\cdot X)_k] \geq E[(H\cdot X)_0] = 0.$ □

Theorem 22.1 (Doob の不等式)**.** $\{X_n\}$ を劣マルチンゲールとし, $n \in \mathbb{Z}_+$ に対して，$\bar{X}_n := \max_{0\leq m\leq n} X_m$ と定義する．このとき，任意の $n \in \mathbb{Z}_+$ と $\alpha > 0$ に対して，

$$\alpha P(\bar{X}_n \geq \alpha) \leq E\left[X_n 1_{\{\bar{X}_n\geq\alpha\}}\right] \leq E[X_n^+].$$

Proof. $\bar{X}_n \geq \alpha$ のとき $N = \min\{0 \leq m \leq n : X_m \geq \alpha\}$ とし，それ以外のとき $N = n$ と定義すると，N は停止時刻であって，$\alpha P(\bar{X}_n \geq \alpha) = E[\alpha 1_{\{\bar{X}_n\geq\alpha\}}] \leq E[X_N 1_{\{\bar{X}_n\geq\alpha\}}]$ となる．ここで，$\{X_m\}$ は劣マルチンゲールであるから，Lemma 22.1 より，$E[X_N 1_A] = E[X_N] - E[X_N 1_{A^c}] \leq E[X_n] - E[X_n 1_{A^c}] = E[X_n 1_A]$ を得る． □

Doob の不等式は Kolmogorov の最大不等式の別証明を与える．$\{\xi_m\}$ を独立な r.v.'s とし，$E[\xi_m] = 0, E[\xi_m^2] < \infty$ とする．$S_0 = 0, S_n = \sum_{m=1}^n \xi_m, \mathcal{F}_0 = \{\varnothing, \Omega\}, \mathcal{F}_n = \sigma(\xi_1, \ldots, \xi_n)$ とすると，$\{S_n^2 : n \in \mathbb{Z}_+\}$ は劣マルチンゲールであるから，Doob の不等式より，$x > 0$ に対して，

$$P\left(\max_{1\leq m\leq n} |S_m| \geq x\right) = P\left(\max_{1\leq m\leq n} S_m^2 \geq x^2\right) \leq \frac{E[S_n^2]}{x^2}$$

が成り立つ．

また，$\{X_n\}$ が優マルチンゲールなら，Doob の不等式の証明において，$E[X_N 1_{\{\bar{X}_n\geq\alpha\}}] = E[X_N] - E[X_N 1_{\{\bar{X}_n\geq\alpha\}^c}] \leq E[X_0] - E[X_n 1_{\{\bar{X}_n\geq\alpha\}^c}]$ となるから，$\alpha P(\bar{X}_n \geq \alpha) \leq E[X_0] - E[X_n 1_{\{\bar{X}_n\geq\alpha\}^c}] \leq E[X_0] + E[X_n^-]$ を得る．$\{X_n\}$ が劣マルチンゲールなら，$\{-X_n\}$ は優マルチンゲールであるから，次の系を得る．

Corollary 22.1. $\{X_n\}$ が劣マルチンゲールなら，任意の $n \in \mathbb{Z}_+$ と $\alpha > 0$ に対して，

$$\alpha P\left(\max_{0\leq m\leq n} |X_n| \geq \alpha\right) \leq 2E[X_n^+] - E[X_0].$$

Theorem 22.2 (L^p 最大不等式)**.** $1 < p < \infty$ とする．$\{X_n\}$ が非負劣マルチンゲールなら，任意の $n \in \mathbb{Z}_+$ に対して，

$$E[\bar{X}_n^p] \leq \left(\frac{p}{p-1}\right)^p E[X_n^p].$$

特に，$\{Y_n\}$ がマルチンゲールなら，$Y_n^* = \max_{0\le m\le n} |Y_m|$ と定義すると，

$$E[(Y_n^*)^p] \le \left(\frac{p}{p-1}\right)^p E[|Y_n|^p].$$

Proof. $E[\bar{X}_n^p] = \infty$ となる場合も考慮に入れて，$K > 0$ に対して，$E[(\bar{X}_n \wedge K)^p]$ を評価する．

$$\begin{aligned}
E[(\bar{X}_n \wedge K)^p] &= \int_0^\infty px^{p-1} P(\bar{X}_n \wedge K \ge x)dx \\
&= \int_0^K px^{p-1} P(\bar{X}_n \ge x)dx \\
&\le \int_0^K px^{p-1} x^{-1} \int X_n 1_{\{\bar{X}_n \ge x\}} dP dx \quad (\because \text{Doob の不等式}) \\
&= \int X_n \int_0^{\bar{X}_n \wedge K} px^{p-2} dx dP \quad (\because \text{Fubini の定理}) \\
&= \frac{p}{p-1} \int X_n (\bar{X}_n \wedge K)^{p-1} dP.
\end{aligned}$$

$q = p/(p-1)$ とおくと，Hölder の不等式より，$E[(\bar{X}_n \wedge K)^p] \le q(E[X_n^p])^{1/p}(E[(\bar{X}_n \wedge K)^p])^{1/q}$. 両辺を $(E[(\bar{X}_n \wedge K)^p])^{1/q}$ で割って，$(E[(\bar{X}_n \wedge K)^p])^{1/p} \le q(E[X_n^p])^{1/p}$ を得る．あとは $K \to \infty$ とすれば，MCT より，左辺 $\to (E[\bar{X}_n^p])^{1/p}$ となって，求める不等式を得る． □

Theorem 22.3 (L^p 収束定理)**.** $\{X_n\}$ をマルチンゲールまたは非負劣マルチンゲールとし，ある $1 < p < \infty$ に対して，$\sup_n E[|X_n|^p] < \infty$ とする．このとき，$\exists X_\infty \in L^p$ s.t. $X_n \to X_\infty$ a.s. and in L^p.

Proof. $E[X_n^+] \le 1 + E[|X_n|^p]$ であるから，マルチンゲール収束定理より，$X_n \to \exists X_\infty$ a.s. さらに，Fatou の補題より，$X_\infty \in L^p$. 次に，L^p 最大不等式より，

$$E\left[\max_{0\le m\le n} |X_m|^p\right] \le \left(\frac{p}{p-1}\right)^p \sup_m E[|X_m|^p].$$

$n \to \infty$ とすると，MCT より，$\sup_m |X_m| \in L^p$. $|X_n - X_\infty|^p \le 2^p \sup_m |X_m|^p$ であるから，DCT より，$E[|X_n - X_\infty|^p] \to 0$. □

Example 22.1. $\{X_n\}$ を $\sup_n E[X_n] < \infty$ なる非負劣マルチンゲールとすると，

$$E\left[\sup_n X_n\right] = \infty$$

となることは ありうる．$\{X_n\}$ を Example 21.4 に現れた非負マルチンゲールとすると，$X_n \to 0$ a.s. であるから，仮に $E[\sup_n X_n] < \infty$ なら，$X_n \to 0$ in L^1 となるが，これは $E[X_n] = 1$ に反する．

$X_n \in L^2$ $(\forall n)$ なるマルチンゲール $\{X_n\}$ に対して，$m < n$ に対して，$E[X_n - X_m \mid \mathcal{F}_m] = 0$ であるから，任意の $Y \in L^2(\Omega, \mathcal{F}_m, P)$ に対して，$E[(X_n - X_m)Y] = 0$ である．よって，

$$X_n = X_0 + \sum_{k=1}^{n} (X_k - X_{k-1})$$

より，

$$E[X_n^2] = E[X_0^2] + \sum_{k=1}^{n} E[(X_k - X_{k-1})^2]$$

を得る．特に，$\sum_{k=1}^{\infty} E[(X_k - X_{k-1})^2] < \infty$ なら $\sup_n E[X_n^2] < \infty$ であるから，L^2 収束定理より，$X_n \to X_\infty$ a.s. and in L^2 となる．$X_k - X_{k-1}, k \in \mathbb{N}$ が独立なときは，これは Theorem 10.2 に他ならない．

L^1 収束

マルチンゲールの L^1 収束を保証する条件を考える．一様可積分性の定義を復習する．r.v.'s $\{X_i : i \in I\}$ が一様可積分であるとは

$$\lim_{K \to \infty} \underbrace{\sup_{i \in I} E[|X_i|, \{|X_i| > K\}]}_{(*)} = 0$$

となることを言う．このとき，K を十分大きくとると，$(*) \leq 1$ となるから，

$$\sup_{i \in I} E[|X_i|] \leq K + 1 < \infty$$

である．次の定理はこれ以降でしばしば用いる重要なものである．

Theorem 22.4. $X \in L^1$ に対して，$\{E[X \mid \mathcal{G}] : \mathcal{G}$ は $\mathcal{F}$ の sub σ-field$\}$ は一様可積分である．

Proof. $\varepsilon > 0$ を任意に固定する．$\delta > 0$ を，$P(A) \leq \delta \Rightarrow E[|X|, A] \leq \varepsilon$ となるように選ぶ．K を十分大きくとって，$E[|X|]/K \leq \delta$ とすると，任意の sub σ-field $\mathcal{G} \subset \mathcal{F}$ に対して，Markov の不等式より，$P(E[|X| \mid \mathcal{G}] > K) \leq E[|X|]/K \leq \delta$ となるから，

$$\begin{aligned}
&E[|E[X \mid \mathcal{G}]|, \{|E[X \mid \mathcal{G}]| > K\}] \\
&\leq E[E[|X| \mid \mathcal{G}], \{E[|X| \mid \mathcal{G}] > K\}] \quad (\because \text{Jensen の不等式}) \\
&= E[|X|, \{E[|X| \mid \mathcal{G}] > K\}] \leq \varepsilon.
\end{aligned}$$

以上より，定理の結論を得る． □

Theorem 22.5. $\{X_n\}$ を劣マルチンゲールとすると，次の (i)–(iii) は同値である．

(i) $\{X_n\}$ は一様可積分.

(ii) X_n は a.s. にかつ in L^1 で収束する.

(iii) X_n は in L^1 で収束する.

(i)–(iii) のいずれかが成り立てば (従って，すべて成り立てば)，$\exists X_\infty \in L^1$ s.t. $X_n \le E[X_\infty \mid \mathcal{F}_n]$ a.s. $\forall n \in \mathbb{Z}_+$ となる.

Proof. (i) $\Rightarrow$ (ii). $\{X_n\}$ は一様可積分であるから，$\sup_n E[|X_n|] < \infty$. よって，マルチンゲール収束定理より，$\exists X_\infty \in L^1$ s.t. $X_n \to X_\infty$ a.s. ここで，Theorem 6.6 より，$X_n \to X_\infty$ in L^1 を得る. (ii) $\Rightarrow$ (iii). 明らか. (iii) $\Rightarrow$ (i). $X_n \to X_\infty$ in L^1 より，$X_n \xrightarrow{P} X_\infty$. よって，Theorem 6.6 より (i) を得る.

(i)–(iii) が成り立つとき，$m > n, A \in \mathcal{F}_n$ に対して，$E[X_n, A] \le E[X_m, A]$ であって，$m \to \infty$ とすると，$X_m \to X_\infty$ in L^1 であるから，$E[X_m, A] \to E[X_\infty, A]$ である. よって，$E[X_n, A] \le E[X_\infty, A]$ を得る. これは，$X_n \le E[X_\infty \mid \mathcal{F}_n]$ a.s. を意味する. □

劣マルチンゲール $\{X_n\}$ に対して，ある $X_\infty \in L^1$ が存在して，$X_n \le E[X_\infty \mid \mathcal{F}_n]$ であったとしても，$\{X_n\}$ は一様可積分とは限らない. しかし，$X_n = E[X_\infty \mid \mathcal{F}_n]$ なら (この場合，$\{X_n\}$ はマルチンゲールである)，Theorem 22.4 より，$\{X_n\}$ は一様可積分である. よって，次の系を得る.

Corollary 22.2. $\{X_n\}$ をマルチンゲールとすると，次の (i)–(iv) は同値である.

(i) $\{X_n\}$ は一様可積分.

(ii) X_n は a.s. にかつ in L^1 で収束する.

(iii) X_n は in L^1 で収束する.

(iv) $\exists X_\infty \in L^1$ s.t. $X_n = E[X_\infty \mid \mathcal{F}_n]$ a.s. $\forall n \in \mathbb{Z}_+$.

以下，$\mathcal{F}_\infty = \bigvee_n \mathcal{F}_n$ とおく.

Theorem 22.6 (Lévy の上向き定理)**.** 任意の $X \in L^1$ に対して，$E[X \mid \mathcal{F}_n] \to E[X \mid \mathcal{F}_\infty]$ a.s. and in L^1.

Proof. $Y_n = E[X \mid \mathcal{F}_n]$ はマルチンゲールである. Theorem 22.4 より，$\{Y_n\}$ は一様可積分であるから，Theorem 22.2 より，$Y_n \to \exists Y_\infty$ a.s. and in L^1. 各 Y_n は $\mathcal{F}_\infty$ 可測なので，Y_∞ を $\mathcal{F}_\infty$ 可測に選べる. ここで，Y_n の定義と L^1 収束より，$E[X \mid \mathcal{F}_n] = Y_n = E[Y_\infty \mid \mathcal{F}_n]$ であるから，$E[X, A] = E[Y_\infty, A]$ $(\forall A \in \mathcal{F}_n)$ となる. n は任意であるから，等号は $A \in \bigcup_n \mathcal{F}_n$ に対して成り立つ. $\bigcup_n \mathcal{F}_n$ は π システムであって，$\{A \in \mathcal{F} : E[X, A] = E[Y_\infty, A]\}$ は λ システムであるから，π-λ 定理より，$E[X, A] = E[Y_\infty, A]$, $\forall A \in \mathcal{F}_\infty$. $Y_\infty \in \mathcal{F}_\infty$ より，$Y_\infty = E[X \mid \mathcal{F}_\infty]$ を得る. □

Corollary 22.3 (Lévy の 0-1 法則)**.** 任意の $A \in \mathcal{F}_\infty$ に対して，$E[1_A \mid \mathcal{F}_n] \to 1_A$ a.s.

Lévy の 0-1 法則は Kolmogorov の 0-1 法則の別証明を与える．$X_1, X_2, \ldots$ を独立な r.v.'s とし，$\mathcal{F}_n = \sigma(X_1, \ldots, X_n)$ とおく．$\mathcal{T}$ を tail σ-field とすると，$A \in \mathcal{T}$ は $\mathcal{F}_n$ と独立であるから，$E[1_A \mid \mathcal{F}_n] = P(A)$. 一方，左辺 $\to 1_A$ a.s. であるから，$P(A) = 1_A$ a.s. となり，$P(A) \in \{0, 1\}$ を得る．

また，Lévy の上向き定理 (Theorem 22.6) の応用として，次の定理を示す．

Theorem 22.7 (条件付き期待値に対する DCT)**.** $Y_n \to Y_\infty$ a.s., $\exists Z \in L^1$ s.t. $|Y_n| \le Z$ ならば，$E[Y_n \mid \mathcal{F}_n] \to E[Y_\infty \mid \mathcal{F}_\infty]$ a.s.

Proof. $W_N = \sup\{|Y_n - Y_m| : n, m \ge N\}$ とおくと，$W_N \le 2Z$ であるから，$E[W_N] < \infty$. また，$n \ge N$ に対して，$|Y_n - Y_\infty| \le W_N$ a.s. となるから，

$$\limsup_n E[|Y_n - Y_\infty| \mid \mathcal{F}_n] \le \lim_n E[W_N \mid \mathcal{F}_n] = E[W_N \mid \mathcal{F}_\infty].$$

ここで，$W_N \downarrow 0$ as $N \to \infty$ より，右辺 $\to 0$. すなわち，$\limsup_n E[|Y_n - Y_\infty| \mid \mathcal{F}_n] = 0$ a.s. を得る．Jensen の不等式より，$|E[Y_n \mid \mathcal{F}_n] - E[Y_\infty \mid \mathcal{F}_n]| \le E[|Y_n - Y_\infty| \mid \mathcal{F}_n] \to 0$ a.s. 一方，Lévy の上向き定理より，$E[Y_\infty \mid \mathcal{F}_n] \to E[Y_\infty \mid \mathcal{F}_\infty]$ となるから，定理の結論を得る． □

Example 22.2 (Kakutani の定理)**.** $\{\xi_m\}$ を $E[\xi_m] = 1$ なる非負 r.v.'s とし，$X_0 = 1, X_n = X_{n-1}\xi_n = \prod_{m=1}^n \xi_m, \mathcal{F}_0 = \{\varnothing, \Omega\}, \mathcal{F}_n = \sigma(\xi_1, \ldots, \xi_n)$ とすると，$\{X_n\}$ は非負マルチンゲールであるから，マルチンゲール収束定理より，$X_\infty = \lim_n X_n$ が a.s. に存在する．$a_n = E[\xi_n^{1/2}]$ とおく．Jensen の不等式より，$0 < a_n \le (E[\xi_n])^{1/2} = 1$ である．

Theorem 22.8 (Kakutani)**.** 次の (i)–(v) は同値である．(i) $E[X_\infty] = 1$. (ii) $X_n \to X_\infty$ in L^1. (iii) $\{X_n\}$ は一様可積分．(iv) $\prod_{n=1}^\infty a_n > 0$. (v) $\sum_{n=1}^\infty (1 - a_n) < \infty$.

(i)–(v) のいずれかの条件が成り立たない (従って，すべての条件が成り立たない) なら，

$$X_\infty = 0 \quad a.s.$$

Proof. (i)–(iii) の同値性は Theorem 6.6 から従う．(iv) $\Rightarrow$ (iii), (i) $\Rightarrow$ (iv), (iv) $\Leftrightarrow$ (v) の順に証明する．定理の最後の主張は (i) $\Rightarrow$ (iv) の証明から従う．

(iv) $\Rightarrow$ (iii).

$$Y_n = \frac{\xi_1^{1/2}}{a_1} \cdots \frac{\xi_n^{1/2}}{a_n} = \frac{X_n^{1/2}}{\prod_{m=1}^n a_m} \ (\ge X_n^{1/2})$$

とおくと，$\{Y_n\}$ も非負マルチンゲールである．

$$E[Y_n^2] = \frac{1}{(\prod_{m=1}^n a_m)^2} \le \frac{1}{(\prod_{m=1}^\infty a_m)^2} < \infty$$

であるから，L^p 最大不等式より，

$$E\left[\sup_{0\le m\le n} X_n\right] \le E\left[\max_{0\le m\le n} Y_n^2\right] \le 4\sup_n E[Y_n^2] < \infty.$$

左辺において $n\to\infty$ として，Fatou の補題より，$E[\sup_m X_m] < \infty$ を得る．従って，$\{X_n\}$ は一様可積分である．

(i) $\Rightarrow$ (iv). 対偶を証明する．$\prod_{n=1}^\infty a_n = 0$ を仮定して，$E[X_\infty] \ne 1$ を示せばよい．Y_n は非負マルチンゲールであるから，有限な $\lim_n Y_n$ が a.s. に存在する．しかし，$\prod_{n=1}^\infty a_n = 0$ であるから，$X_n \to 0$ a.s., i.e., $X_\infty = 0$ a.s. でなくてはならない．

(iv) $\Leftrightarrow$ (v). $b_n = 1 - a_n \in [0,1)$ とおいて，$\prod_n (1-b_n) = 0 \Leftrightarrow \sum_n b_n = \infty$ を示せばよい．$\limsup_n b_n > 0$ のとき，$\prod_n(1-b_n) = 0$ & $\sum_n b_n = \infty$ であるから，$b_n \to 0$ と仮定してよい．$\Leftarrow$ は $1 - x \le e^{-x}$ より，$\prod_{m=1}^n (1-b_m) \le \exp(-\sum_{m=1}^n b_m) \to 0$．$\Rightarrow$ を示す．$b_n \to 0$ より，十分大な M に対して，$b_n \le 1/2\ \forall n > M$．従って，$\prod_{m=M+1}^n (1-b_m) \ge \exp(-2\sum_{m=M+1}^n b_n)$ ($\because 1 - x \ge e^{-2x}, 0 \le x \le 1/2$)．$n\to\infty$ とすると，左辺 $\to 0$ であるから，$\sum_{n=M+1}^\infty b_n = \infty$. □

Burkholder の不等式

マルチンゲールは確率不等式を導く強力なテクニックを与える．以下，次の Burkholder の不等式を証明する．$\{\mathcal{F}_n : n \in \mathbb{Z}_+\}$ を $\mathcal{F}_0 = \{\varnothing, \Omega\}$ なるフィルトレーションとする．また，確率過程 $\{X_n : n \in \mathbb{Z}_+\}$ に対して，

$$[X]_n = \sum_{m=1}^n (X_m - X_{m-1})^2,\ n \in \mathbb{N}$$

と定義する．$\{[X]_n : n \in \mathbb{N}\}$ は $\{X_n\}$ の 2 次変動過程 (quadratic variation process) と呼ばれる．

Theorem 22.9 (Burkholder (1966))**.** $\{X_n : n \in \mathbb{Z}_+\}$ を $X_0 = 0$ なるマルチンゲールとする． このとき，与えられた $1 < p < \infty$ に対して，p のみに依存する定数 A_p, B_p が存在して，各 $n \in \mathbb{N}$ に対して，

$$A_p(E[[X]_n^{p/2}])^{1/p} \le (E[|X_n|^p])^{1/p} \le B_p(E[[X]_n^{p/2}])^{1/p} \qquad (*)$$

が成り立つ．

$\{X_m - X_{m-1}\}$ が独立な r.v.'s の場合は，Burkholder の不等式は Marcinkiewicz-Zygmund (cf. Exercise 8.5) の不等式に帰着する．定理の証明のために，いくつかの補題を証明する．

Lemma 22.2. $\{X_n : n \in \mathbb{Z}_+\}$ を $X_0 = 0$ なるマルチンゲールまたは非負劣マルチンゲールとする．いま，$\alpha > 0$ に対して，

$$N = \begin{cases} \min\{1 \le m \le n : |X_m| > \alpha\} & \text{if } \max_{1\le m\le n} |X_m| > \alpha \\ n+1 & \text{otherwise} \end{cases}$$

と定義すると，

$$E[[X]_{N-1}] + E[X_{N-1}^2] \le 2\alpha E[|X_n|]$$

が成り立つ．

Proof. $Y_m = X_m, 0 \le m \le n,\ Y_m = X_n, m \ge n+1$ とおき，$e_m = Y_m - Y_{m-1}, m \in \mathbb{N}$ とおく．このとき，

$$[X]_{N-1} + X_{N-1}^2 = \sum_{m=1}^{N-1} e_m^2 + Y_{N-1}^2 = 2Y_{N-1}^2 - 2 \sum_{1\le k<m\le N-1} e_k e_m$$
$$= 2Y_{N-1}(Y_N - e_N) - 2\sum_{m=2}^{N-1} Y_{m-1}e_m = 2Y_N Y_{N-1} - 2\sum_{m=2}^{N} Y_{m-1}e_m$$

となる．ここで，$\{m \le N\} = \{N \le m-1\}^c \in \mathcal{F}_{m-1}$ であるから，

$$E\left[\sum_{m=2}^{N} Y_{m-1}e_m\right] = E\left[\sum_{m=2}^{n} Y_{m-1}E[e_m \mid \mathcal{F}_{m-1}]1_{\{m\le N\}}\right] \le 0.$$

あとは，$|Y_{N-1}| \le \alpha$ より，$E[Y_N Y_{N-1}] \le \alpha E[|Y_N|] \le \alpha E[|Y_{n+1}|] = \alpha E[|X_n|]$ を得る． □

Lemma 22.3. $\{X_m : n \in \mathbb{Z}_+\}$ を $X_0 = 0$ なる非負劣マルチンゲールとし，与えられた $\theta > 0$ に対して，$Y = \max\{\theta[X]_n^{1/2}, \max_{1\le m\le n} X_m\}$ とおく．このとき，あらゆる $\alpha > 0$ に対して，

$$\alpha P(Y > \beta\alpha) \le 3E[X_n 1_{\{Y>\alpha\}}]$$

が成り立つ．ただし，$\beta = (1+2\theta^2)^{1/2}$ である．さらに，あらゆる $1 < p < \infty$ に対して，

$$(E[[X]^{p/2}])^{1/p} \le 9p^{1/2}q(E[X_n^p])^{1/p}$$

が成り立つ．ただし，$q = p/(p-1)$ である．

Proof. $\beta > 1$ より，

$$P(Y > \beta\alpha) \le P\left(\max_{1\le m\le n} X_m > \alpha\right) + P\left\{\theta[X]_n^{1/2} > \beta\alpha, \max_{1\le m\le n} X_m \le \alpha\right\}.$$

Doob の不等式より，右辺第 1 項 $\le \alpha^{-1}E[X_n 1_{\{\max_{1\le m\le n} X_m>\alpha\}}] \le \alpha^{-1}E[X_n 1_{\{Y>\alpha\}}]$ である．次に，$Y_m = X_m 1_{\{\theta[X]_m^{1/2}>\alpha\}}$, $m \in \mathbb{Z}_+$ とおくと，$\{Y_m : n \in \mathbb{Z}_+\}$ は非負劣マルチンゲールである．

$$\left\{\theta[X]_n^{1/2} > \beta\alpha, \max_{1\le m\le n} X_m \le \alpha\right\} \subset \left\{[Y]_n^{1/2} > \alpha, \max_{1\le m\le n} Y_m \le \alpha\right\} \qquad (**)$$

を示そう．

$$M = \begin{cases} \min\left\{1 \le m \le n : \theta[X]_m^{1/2} > \alpha\right\} & \text{if } \{\cdots\} \ne \varnothing \\ n & \text{otherwise} \end{cases}$$

とおくと，$\theta[X]_n^{1/2} > \beta\alpha$ と $\max_{1\le m\le n} X_m \le \alpha$ が成り立てば，

$$\begin{aligned}\beta^2\alpha^2 < \theta^2[X]_n &= \theta^2\left\{\sum_{m=1}^{M-1}(X_m - X_{m-1})^2 + (X_M - X_{M-1})^2 + \sum_{m=M+1}^{n}(X_m - X_{m-1})^2\right\} \\ &\le \alpha^2 + \theta^2\alpha^2 + \theta^2\sum_{m=M+1}^{n}(Y_m - Y_{m-1})^2\end{aligned}$$

より，$[Y]_n \ge \sum_{m=M+1}^{n}(Y_m - Y_{m-1})^2 > \alpha^2$ となる．よって，$(**)$ を得る．

あとは，Markov の不等式と Lemma 22.2 より，

$$\begin{aligned}P\left\{[Y]_n^{1/2} > \alpha, \max_{1\le m\le n} Y_m \le \alpha\right\} &\le \alpha^{-2}E\left[[Y]_n, \{\max_{1\le m\le n} Y_m \le \alpha\}\right] \\ &\le 2\alpha^{-1}E[Y_n] \le 2\alpha^{-1}E[X_n 1_{\{Y>\alpha\}}]\end{aligned}$$

を得る．

次に，L^p 最大不等式の証明と同様の議論により，

$$\theta(E[[X]^{p/2}])^{1/p} \le (E[Y^p])^{1/p} \le 3\beta^p q(E[X_n^p])^{1/p}$$

を得る．$\theta = p^{-1/2}$ とおくと，$\beta^p = (1+2/p)^{p/2} < e < 3$ であるから，2 番目の不等式も得る． □

Proof of Theorem 22.9. Lemma 22.3 を $\{X_n^+\}$ と $\{X_n^-\}$ に適用して，

$$(E[[X]_n^{p/2}])^{1/p} \le 18p^{1/2}q(E[|X_n|^p])^{1/p} \qquad (*3)$$

を得る．よって，$(*)$ の左側の不等式が $A_p = 1/(18p^{1/2}q)$ として成り立つ．

次に，$(*)$ の最右辺を有限と仮定して，右側の不等式を証明する (最右辺が $+\infty$ なら何も証明することはない)．このとき，$E[|X_n|^p] < \infty$ である．さらに，$E[|X_n|^p] > 0$ と仮定してよい．$Y = \mathrm{sgn}(X_n)|X_n|^{p-1}/(E[|X_n|^p])^{(p-1)/p}$ とおくと，$E[|Y|^q] = 1$, $E[X_nY] = (E[|X_n|^p])^{1/p}$ となる．$Y_m = E[Y \mid \mathcal{F}_m]$, $m \in \mathbb{Z}_+$ とおくと，$\{Y_m\}$ はマルチンゲール

であって，(∗3) より，$(E[[Y]_n^{q/2}])^{1/q} \le 18q^{1/2}p(E[|Y_n|^q])^{1/q} = 18q^{1/2}p$ となる．さらに，Cauchy-Schwarz の不等式と Hölder の不等式より，

$$E[X_nY] = E\left[\sum_{m=1}^{n}(X_m - X_{m-1})(Y_m - Y_{m-1})\right] \le E[[X]_n^{1/2}[Y]_n^{1/2}]$$
$$\le (E[[X]_n^{p/2}])^{1/p}(E[[Y]_n^{q/2}])^{1/q} \le 18q^{1/2}p(E[[X]_n^{p/2}])^{1/p}$$

を得る．よって，(∗) の右側の不等式が $B_p = 18q^{1/2}p$ として成り立つ． □

演習問題

Exercise 22.1. M, N を停止時刻とし，$M \le N$ とする．また，

$$H_n = \begin{cases} 1 & \text{if } M < n \le N \\ 0 & \text{otherwise} \end{cases}$$

とおく．このとき，$\{H_n\}$ が可予測であることを示せ．また，$\{X_n\}$ が (劣/優) マルチンゲールなら，$\{X_{N\wedge n} - X_{M\wedge n} : n \in \mathbb{Z}_+\}$ も (劣/優) マルチンゲールであることを示せ．

Exercise 22.2. $\{X_n\}$ を劣マルチンゲールとする．いま，$\exists X_\infty \in L^1$ s.t. $X_n \le E[X_\infty \mid \mathcal{F}_n]$ a.s. $\forall n \in \mathbb{Z}_+$ とする．

(a) $\{X_n\}$ が非負なら，$\{X_n\}$ は一様可積分であることを示せ．

(b) $\{X_n\}$ は一般には一様可積分とは限らない．反例を構成せよ．

Exercise 22.3. $\{X_n\}$ を非負劣マルチンゲールとし，$\bar{X}_n = \max_{0\le m\le n} X_m$ とおく．このとき，

$$E[\bar{X}_n] \le \frac{e}{e-1}\left\{1 + E[X_n \log^+ X_n]\right\}$$

を示せ．ただし，$\log^+ x = \max\{\log x, 0\}$ である．
ヒント：$a, b > 0$ に対して，$a \log b \le a \log^+ a + e^{-1}b$.

この評価より，非負劣マルチンゲール $\{X_n\}$ に対して，$\sup_n E[X_n \log^+ X_n] < \infty$ であれば，$E[\sup_n X_n] < \infty$ となる．特に，$\{X_n\}$ は一様可積分である．じつはこの評価は一様には改善できないことが次の定理からわかる．

Theorem 22.10 (Gundy). $\{X_n\}$ を非負マルチンゲールとし，ある定数 $C > 0$ が存在して，$X_{n+1} \le CX_n$ a.s. $(\forall n)$ とする．このとき，

$$E[X_0 \log^+ X_0] \vee E[\sup_n X_n] < \infty \Rightarrow \sup_n E[X_n \log^+ X_n] < \infty.$$

Proof. 証明は Neveu (1975) による．$\alpha > 0$ に対して，$N_\alpha := \inf\{n : X_n > \alpha\}$ とおくと，

$$E[X_n, \{0 < N_\alpha \le n\}] = \sum_{k=1}^{n} E[X_n, \{N_a = k\}] = \sum_{k=1}^{n} E[E[X_n \mid \mathcal{F}_k], \{N_\alpha = k\}]$$
$$= \sum_{k=1}^{n} E[X_k, \{N_\alpha = k\}] = E[X_{N_\alpha}, \{0 < N_\alpha \le n\}] \le C\alpha P(0 < N_\alpha \le n)$$

である．最後の不等号は $\{0 < N_\alpha \le n\}$ 上で $X_{N_a} \le CX_{N_\alpha - 1} \le C\alpha$ であるから．よって，

$$E[X_n, \{\bar{X}_n > \alpha\}] \le E[X_n, \{X_0 > 0\}] + E[X_n, \{0 < N_\alpha \le n\}]$$
$$\le E[X_0, \{X_0 > \alpha\}] + C\alpha P(0 < N_\alpha \le n).$$

両辺を測度 $1_{[1,\infty)}\alpha^{-1}d\alpha$ に関して積分して，

$$E[X_n \log^+ \bar{X}_n] \le E[X_0 \log^+ X_0] + CE[\bar{X}_n]$$

を得る．左辺は $E[X_n \log^+ X_n]$ で下からバウンドできる．そこで，$n \to \infty$ として，

$$\sup_n E[X_n \log^+ X_n] \le E[X_0 \log^+ X_0] + CE[\sup_n X_n] < \infty$$

を得る． □

Exercise 22.4. $\{X_n\} \subset L^2$ を非負マルチンゲールとし，ある定数 $c \ge 1$ が存在して，$(X_{n+1} - X_n)^2 \le cE[(X_{n+1} - X_n)^2 \mid \mathcal{F}_n]$ a.s. $(\forall n)$ とする．このとき，$X_{n+1} \le (1 + 2c)X_n$ a.s. $(\forall n)$ を示せ．

期待値が 0 に収束する非負優マルチンゲールを ポテンシャル (potential) と呼ぶ．

Exercise 22.5. $\{X_n\}$ を優マルチンゲールとし，$\sup_n E[X_n^-] < \infty$ とする．このとき，ある一様可積分マルチンゲール $\{M_n\}$ とポテンシャル $\{V_n\}$ が存在して，$X_n = M_n + V_n$ $(\forall n)$ a.s. と表せることを示せ．またこの分解は a.s. に一意であることを示せ．

Exercise 22.6 (Riesz 分解)**.** $\{X_n\}$ を非負優マルチンゲールとする．このとき，$Y_n = \lim_m E[X_{n+m} \mid \mathcal{F}_n]$ が a.s. に存在し，$\{Y_n\}$ は非負マルチンゲールであることを示せ．さらに，$Z_n = X_n - Y_n$ とおくと，$\{Z_n\}$ はポテンシャルであることを示せ．$X_n = Y_n + Z_n$ の分解を Riesz 分解 と呼ぶ．

Exercise 22.7. Riesz 分解において，$\{Y_n'\}$ を $Y_n' \le X_n$ $(\forall n)$ a.s. なる劣マルチンゲールとすると，$Y_n' \le Y_n$ $(\forall n)$ a.s. となることを示せ．

Exercise 22.8 (Krickeberg 分解)**.** $\{X_n\}$ を L^1 有界なマルチンゲールとする．このとき，$Y_n = \lim_m E[X_{n+m}^+ \mid \mathcal{F}_n]$ と $Z_n = \lim_m E[X_{n+m}^- \mid \mathcal{F}_n]$ が a.s. に存在することを示せ．また，$\{Y_n\}$ と $\{Z_n\}$ は L^1 有界な非負マルチンゲールであって，$X_n = Y_n - Z_n$ $(\forall n)$ a.s. をみたすことを示せ．この分解を Krickeberg 分解 と呼ぶ．マルチンゲールが L^1 有界であるためには，それが L^1 有界な非負マルチンゲールの差で表されることが必要十分である．

Exercise 22.9. Krickeberg 分解において，$\{Y'_n\}$ と $\{Z'_n\}$ をそれぞれ非負マルチンゲールとし，$X_n = Y'_n - Z'_n$ $(\forall n)$ a.s. をみたすとすると，$Y_n \le Y'_n, Z_n \le Z'_n$ $(\forall n)$ a.s. となることを示せ．

Exercise 22.10. $-\infty < a < b < \infty$ とする．関数 $f : [a,b] \to \mathbb{R}$ が 絶対連続 (absolutely continuous) であるとは，任意の $\varepsilon > 0$ に対して，ある $\delta > 0$ が存在して，

$$a \le a_1 < b_1 \le \cdots \le a_k < b_k \le b,\ k \in \mathbb{N},\ \sum_{\ell=1}^{k}(b_\ell - a_\ell) < \delta \Rightarrow \sum_{\ell=1}^{k} |f(b_\ell) - f(a_\ell)| < \varepsilon$$

となることを言う．本問の目標は次の定理 (の必要性の部分) を証明することである [14]．

Theorem 22.11. 関数 $f : [a,b] \to \mathbb{R}$ が絶対連続であるための必要十分条件は，ある $g \in L^1([a,b], \mathcal{B}_{[a,b]}, \lambda)$ が存在して，

$$f(x) - f(a) = \int_a^x g(y)dy,\ x \in [a,b] \qquad (*)$$

と表せることである．

以下，$a = 0, b = 1$ とし，$f : [0,1] \to \mathbb{R}$ を絶対連続とする．また，確率空間を $(\Omega, \mathcal{F}, P) = ([0,1], \mathcal{B}_{[0,1]}, \lambda)$ とし，$A_{n,i} = [(i-1)2^{-n}, i2^{-n}),\ i = 1, \ldots, 2^n - 1,\ A_{n,2^n} = [1 - 2^{-n}, 1]$ とおいて，$\mathcal{F}_n = \sigma(\{A_{n,i} : i = 1, \ldots, 2^n\})$ と定義する．

(a) $b_{n,i} = 2^n\{f(i2^{-n}) - f((i-1)2^{-n})\}$ とおいて，$X_n(\omega) = \sum_{i=1}^{2^n} b_{n,i} 1_{A_{n,i}}(\omega)$ と定義する．このとき，$\{X_n\}$ は $\{\mathcal{F}_n\}$ に関するマルチンゲールであることを示せ．

(b) $\{X_n\}$ が一様可積分であることを示し，$(*)$ をみたす $g \in L^1([0,1], \mathcal{B}_{[0,1]}, \lambda)$ の存在を示せ．

ヒント：$M := \sup_n E[|X_n|] < \infty$ を示せ．これから，$P(|X_n| > \rho) \le M/\rho$ を得る．

Remark 22.1. あとで証明する Lebesgue の微分定理より，f が $(*)$ の形に表せるなら，f は λ-a.e. $x \in [a,b]$ において微分可能であって，$f' = g$ λ-a.e. $x \in [a,b]$ となる．

[14] 十分性の部分はすでに演習問題に現れた．

23　バックワード・マルチンゲール

$\{\mathcal{F}_n : n \le 0\}$ を $\mathcal{F}$ の sub σ-fields の増加列とする．

$$\mathcal{F}_0 \supset \mathcal{F}_{-1} \supset \mathcal{F}_{-2} \supset \cdots .$$

確率過程 $\{X_n : n \le 0\}$ が バックワード・マルチンゲール (backward martingale) であるとは，次の 3 条件がみたされることを言う．

(i) $E[|X_n|] < \infty \ \forall n \le 0$.

(ii) $X_n \in \mathcal{F}_n \ \forall n \le 0$.

(iii) $E[X_{n+1} \mid \mathcal{F}_n] = X_n$ a.s. $\forall n \le -1$.

ところで，(iii) より

$$X_n = E[X_0 \mid \mathcal{F}_n],\ n \le 0$$

であるから，$\{X_n : n \le 0\}$ は一様可積分である．

Theorem 23.1. $\mathcal{F}_{-\infty} = \bigcap_{n \le 0} \mathcal{F}_n$ とおく．バックワード・マルチンゲール $\{X_n : n \le 0\}$ に対して，$n \to -\infty$ とすると，$X_n \to E[X_0 \mid \mathcal{F}_{-\infty}]$ a.s. and in L^1.

Proof. U_n を $X_{-n}, \ldots, X_0$ が $[a, b]$ を上向きに横断する回数とすると，上向き横断不等式より，$(b-a)E[U_n] \le E[(X_0 - a)^+]$ を得る．$n \to \infty$ とすると，$0 \le U_n \uparrow U_\infty := \sup_m U_m$ であって，Fatou の補題より，$E[U_\infty] < \infty$ を得る．あとはマルチンゲール収束定理の証明と同様にして，$\exists X_{-\infty} \in L^1$ s.t. $X_n \to X_{-\infty}$ a.s. $(n \to -\infty)$ が示される．ここで，$\{X_n : n \le 0\}$ は一様可積分であるから，$X_n \to X_{-\infty}$ in L^1 も従う．

次に，$X_{-\infty} = E[X_0 \mid \mathcal{F}_{-\infty}]$ を示す．まず，任意の $m \le 0$ に対して，$X_n \in \mathcal{F}_m$ $(\forall n \le m)$ であるから，$X_{-\infty} \in \mathcal{F}_m$ である．厳密には，$X_{-\infty}$ のとり方は確率 0 の集合上で任意性があり，その分だけ $\mathcal{F}_m$ の外側にいるかもしれないので，都合のよい $X_{-\infty}$ をとってくる．$E :=$ { 有限な $\lim_{n \to -\infty} X_n$ が存在する } $\in \mathcal{F}_m$ より，

$$\tilde{X}_{-\infty}(\omega) = \begin{cases} \lim_{n \to -\infty} X_n(\omega) & \text{if } \omega \in E \\ 0 & \text{if } \omega \in E^c \end{cases}$$

とおくと，$\tilde{X}_{-\infty} \in \mathcal{F}_m$ であり, $\tilde{X}_{-\infty} = X_{-\infty}$ a.s. そこで $X_{-\infty}$ を $\tilde{X}_{-\infty}$ に取り替えればよい．$m \le 0$ は任意であるから，$X_{-\infty} \in \bigcap_{m \le 0} \mathcal{F}_m = \mathcal{F}_{-\infty}$. 一方，$X_n = E[X_0 \mid \mathcal{F}_n]$ より，$A \in \mathcal{F}_{-\infty} \subset \mathcal{F}_n$ に対して，$E[X_n, A] = E[X_0, A]$ となる．$n \to -\infty$ とすると，$X_n \to X_{-\infty}$ in L^1 より，左辺 $\to E[X_{-\infty}, A]$ となるから，$X_{-\infty} = E[X_0 \mid \mathcal{F}_{-\infty}]$ を得る． □

Corollary 23.1 (Lévy の下向き定理)**.** 任意の $Y \in L^1$ に対して，$n \to -\infty$ とすると，$E[Y \mid \mathcal{F}_n] \to E[Y \mid \mathcal{F}_{-\infty}]$ a.s. and in L^1.

Proof. $X_n = E[Y \mid \mathcal{F}_n]$ はバックワード・マルチンゲールであるから，$n \to -\infty$ のとき，$X_n \to E[X_0 \mid \mathcal{F}_{-\infty}] = E[E[Y \mid \mathcal{F}_0] \mid \mathcal{F}_{-\infty}] = E[Y \mid \mathcal{F}_{-\infty}]$ a.s. and in L^1 となる． □

$\{X_n : n \leq 0\}$ がバックワード・劣マルチンゲールであるとは，$E[|X_n|] < \infty, X_n \in \mathcal{F}_n$ $(\forall n \leq 0)$，$E[X_{n+1} \mid \mathcal{F}_n] \geq X_n$ a.s. $(\forall n \leq -1)$ の 3 条件がみたされることを言う．

Theorem 23.2 (Doob 分解)**.** $\{X_n : n \leq 0\}$ をバックワード・劣マルチンゲールとし，$\inf_{n\leq 0} E[X_n] > -\infty$ とする．このとき，適当なバックワード・マルチンゲール $\{M_n\}$ と可予測な増加過程 $\{A_n\}$ (i.e., $A_n \in \mathcal{F}_{n-1}$ $(\forall n \leq 0)$ かつ a.s. に $A_{n+1} \geq A_n$ $(\forall n \leq -1)$) であって，$A_n \downarrow 0$ a.s. $(n \to -\infty)$ となるものが存在して，a.s. に $X_n = M_n + A_n$ $(\forall n \leq 0)$ なる分解が成り立つ．さらに，$n \to -\infty$ のとき, X_n は a.s. にかつ in L^1 で収束する．

特に，期待値が下に有界なバックワード・劣マルチンゲールは一様可積分である．

Proof. $D_n = E[X_n \mid \mathcal{F}_{n-1}] - X_{n-1},\ n = 0, -1, \ldots$ とおくと，$D_n \geq 0$ a.s. であり，$K = \inf_{n\leq 0} E[X_n] > -\infty$ とおくと，$E[X_n] \downarrow K$ $(n \to -\infty)$ より，$E[\sum_{n\leq 0} D_n] = \sum_{n\leq 0} E[D_n] = \sum_{n\leq 0} E[X_n - X_{n-1}] = E[X_0] - K$ であるから，$\sum_{n\leq 0} D_n$ は可積分な r.v. に収束する．いま，$A_n = \sum_{k\leq n} D_k,\ n = 0, -1, \ldots$ とおくと，$A_n \in \mathcal{F}_{n-1}$ であり，a.s. に $A_0 \geq A_{-1} \geq \ldots \downarrow 0$ となる．また，$M_n = X_n - A_n$ とおくと，

$$E[M_{n+1} \mid \mathcal{F}_n] = X_n + D_{n+1} - A_{n+1} = X_n - A_n = M_n$$

であるから，$\{M_n\}$ はバックワード・マルチンゲールである．さらに，$n \to -\infty$ のとき，$A_n \to 0$ a.s. and in L^1 であるから，

$$X_n = M_n + A_n \to E[M_0 \mid \mathcal{F}_{-\infty}] \quad \text{a.s. and in } L^1$$

を得る． □

Hewitt-Savage の 0-1 法則と de Finetti の定理

バックワード・マルチンゲールに対する収束定理 (Theorem 23.1) と Lévy の下向き定理 (Corollary 23.1) の応用として，Hewitt-Savage の 0-1 法則と de Finetti の定理を証明する．ここでは，可測空間 $(\Omega, \mathcal{F})$ と r.v.'s $X_n, n \in \mathbb{N}$ は

$$\begin{aligned}
\Omega &= S^{\mathbb{N}} = S \times S \times \cdots, \\
\mathcal{F} &= \mathcal{S}^{\mathbb{N}} = \mathcal{S} \times \mathcal{S} \times \cdots, \\
X_n(\omega) &= X_n(\omega_1, \omega_2, \ldots) = \omega_n
\end{aligned}$$

とする．$\mathbb{N}$ の <u>有限置換</u> π とは，$\mathbb{N}$ から $\mathbb{N}$ への全単射であって，有限個の i を除いて $\pi(i) = i$ となるものを指す．Σ を $\mathbb{N}$ の有限置換の全体とする．$\pi \in \Sigma, \omega \in \Omega$ に対して，$(\pi\omega)_i = \omega_{\pi(i)}$ と定義して，$A \subset \Omega$ に対して，$\pi^{-1}A = \{\omega : \pi\omega \in A\}$ とすると，

$$\mathcal{E} := \{A \in \mathcal{F} : \pi^{-1}A = A,\ \forall \pi \in \Sigma\}$$

は σ-field である．$\mathcal{E}$ を exchangeable σ-field と呼ぶ．直感的には，$\mathcal{E}$ に属する事象は有限個の X_i を入れ替えても不変な事象と理解できる．

$n \in \mathbb{N}$ に対して，$\Sigma_n = \{\pi \in \Sigma : \pi(i) = i, i \geq n+1\}$ とおいて，さらに

$$\mathcal{E}_n = \{A \in \mathcal{F} : \pi^{-1}A = A,\ \forall \pi \in \Sigma_n\}$$

と定義すると，$\mathcal{E}_n$ も σ-field で，$\mathcal{E} = \bigcap_n \mathcal{E}_n$ となる．直感的には，$\mathcal{E}_n$ に属する事象は最初の $X_1, \dots, X_n$ を入れ替えても不変な事象と理解できる．

r.v.'s $X_1, X_2, \dots$ が exchangeable であるとは，任意の $\pi \in \Sigma$ に対して，

$$(X_{\pi(1)}, \dots, X_{\pi(n)}, \dots) \overset{d}{=} (X_1, \dots, X_n, \dots) \qquad (*)$$

となることを言う．$(*)$ の意味は，左辺と右辺の $(S^{\mathbb{N}}, \mathcal{S}^{\mathbb{N}})$ 上の分布が等しいことである．$(S^{\mathbb{N}}, \mathcal{S}^{\mathbb{N}})$ 上の 2 つの p.m.'s が等しいことは，その f.d.d.'s がすべて等しいことと同値であるから，$X_1, X_2, \dots$ が exchangeable であるための必要十分条件は，各 $n \in \mathbb{N}$ と $\{1, \dots, n\}$ のあらゆる置換 π に対して，

$$(X_{\pi(1)}, \dots, X_{\pi(n)}) \overset{d}{=} (X_1, \dots, X_n)$$

となることである．

Lemma 23.1. $X_1, X_2, \dots$ が exchangeable なら，$E[|\varphi(X_1, \dots, X_k)|] < \infty$ となる任意の可測関数 $\varphi : S^k \to \mathbb{R}$ に対して，$n \to \infty$ とすると，

$$A_n(\varphi) := \frac{1}{(n)_k} \sum_{\substack{1 \leq i_1, \dots, i_k \leq n \\ i_1, \dots, i_k \text{は相異なる}}} \varphi(X_{i_1}, \dots, X_{i_k}) \to E[\varphi(X_1, \dots, X_k) \mid \mathcal{E}] \ a.s.$$

が成り立つ．ただし，$(n)_k = n(n-1)\cdots(n-k+1)$ である．

Proof. $A_n(\varphi) \in \mathcal{E}_n$ と $X_1, X_2, \dots$ の exchangeability より，

$$A_n(\varphi) = E[A_n(\varphi) \mid \mathcal{E}_n] = \frac{1}{(n)_k} \sum E[\varphi(X_{i_1}, \dots, X_{i_k}) \mid \mathcal{E}_n] = E[\varphi(X_1, \dots, X_k) \mid \mathcal{E}_n].$$

$Y = \varphi(X_1, \dots, X_k), \mathcal{F}_n = \mathcal{E}_{-n}$ とおくと，$A_{-n}(\varphi) = E[Y \mid \mathcal{F}_n], n \leq -1$ であるから，$n \to -\infty$ として，$A_{-n}(\varphi) \to E[Y \mid \mathcal{F}_{-\infty}] = E[Y \mid \mathcal{E}]$ a.s. □

Theorem 23.3 (Hewitt-Savage の 0-1 法則)**.** $X_1, X_2, \dots$ が i.i.d. なら，あらゆる $A \in \mathcal{E}$ に対して，$P(A) \in \{0, 1\}$ である．

Proof. $\varphi : S^k \to \mathbb{R}$ を有界可測とすると，$E[\varphi(X_1, \dots, X_k) \mid \mathcal{E}] = \lim_n A_n(\varphi)$ a.s.

$$A_n(\varphi) = \frac{1}{(n)_k} \sum_{1 \notin \{i_1, \dots i_k\}} + \frac{1}{(n)_k} \sum_{1 \in \{i_1, \dots, i_k\}}$$

と分割すると，

$$|(\text{第 2 項})| \le \frac{k(n-1)_{k-1}}{(n)_k}\|\varphi\|_u \to 0.$$

第 1 項は X_1 に依存しないので，

$$E[\varphi(X_1,\dots,X_k) \mid \mathcal{E}] \in \sigma(X_2, X_3, \dots).$$

より正確に述べると，$E[\varphi(X_1,\dots,X_k) \mid \mathcal{E}]$ のバージョンで，$\sigma(X_2, X_3, \dots)$ 可測なものが存在する．この操作を繰り返すと，

$$E[\varphi(X_1,\dots,X_k) \mid \mathcal{E}] \in \sigma(X_{k+1}, X_{k+2}, \dots).$$

$\varphi(X_1,\dots,X_k)$ は $\sigma(X_{k+1}, X_{k+2}, \dots)$ と独立であるから，直感的には左辺は $E[\varphi(X_1,\dots,X_k)]$ に等しいことが予想される．実際，次の補題が成り立つ．

Lemma 23.2. $E[X^2] < \infty, E[X \mid \mathcal{G}] \in \mathcal{H}$ とする．このとき，X と $\mathcal{H}$ が独立ならば，$E[X \mid \mathcal{G}] = E[X]$.

Proof of Lemma 23.2. $Y = E[X \mid \mathcal{G}]$ とおくと，$E[Y^2] \le E[X^2] < \infty$. $Y \in \mathcal{H}$ と X と $\mathcal{H}$ の独立性より，$E[XY] = E[X]E[Y] = (E[Y])^2$. 一方，条件付き期待値の性質より，$E[(X-Y)Y] = 0$, i.e., $E[Y^2] = (E[Y])^2$. 従って，$\mathrm{Var}(Y) = 0$ であるから，$Y = E[Y]$. □

この補題より，

$$E[\varphi(X_1,\dots,X_k) \mid \mathcal{E}] = E[\varphi(X_1,\dots,X_k)].$$

特に，$\varphi = 1_B, B \in \mathcal{S}^k$ とすると，$A \in \mathcal{E}$ に対して，

$$\begin{aligned} P(A \cap \{(X_1,\dots,X_k) \in B\}) &= E[1_B(X_1,\dots,X_k), A] \\ &= E[E[1_B(X_1,\dots,X_k) \mid \mathcal{E}], A] = P(A)P((X_1,\dots,X_k) \in B) \end{aligned}$$

であるから，$\mathcal{E}$ と $\sigma(X_1,\dots,X_k) =: \mathcal{G}_k$ は独立である．k は任意であって，$\bigcup_k \mathcal{G}_k$ は Ω を含む π システムであるから，$\mathcal{E}$ と $\bigvee_k \mathcal{G}_k$ は独立である．特に，$A \in \mathcal{E}$ に対して，$\mathcal{E} \subset \bigvee_k \mathcal{G}_k$ より $A \in \bigvee_k \mathcal{G}_k$ と考えると，A は A 自身と独立であり，$P(A) = P(A \cap A) = P(A)^2$ より，$P(A) \in \{0, 1\}$ を得る. □

Remark 23.1 (SLLN の別証明)**.** $X_1, X_2, \dots$ を i.i.d. ($\mathbb{R}$-valued) r.v.' とし，$E[|X_1|] < \infty$ とする．このとき，Lemma 23.1 より，$n^{-1}\sum_{i=1}^n X_i \to E[X_1 \mid \mathcal{E}]$ a.s. ここで，Hewitt-Savage の 0-1 法則より，$E[X_1 \mid \mathcal{E}] = E[X_1]$ となるから，SLLN が示された．

Theorem 23.4 (de Finetti)**.** $X_1, X_2, \dots$ がexchangeable なら，任意の $k \in \mathbb{N}$ と $A_1, \dots, A_k \in \mathcal{S}$ に対して，

$$P(X_1 \in A_1, \dots, X_k \in A_k \mid \mathcal{E}) = P(X_1 \in A_1 \mid \mathcal{E}) \cdots P(X_1 \in A_k \mid \mathcal{E}).$$

Proof. $f : S^{k-1} \to \mathbb{R}, g : S \to \mathbb{R}$ を有界可測関数とし，$I_{n,k}$ を相異なる k 個の整数の組 $1 \le i_1, \dots, i_k \le n$ の全体とする．このとき，

$$(n)_{k-1} A_n(f) n A_n(g) = \sum_{I_{n,k-1}} f(X_{i_1}, \dots, X_{i_{k-1}}) \sum_m g(X_m)$$

$$= \sum_{I_{n,k}} f(X_{i_1}, \dots, X_{i_{k-1}}) g(X_{i_k}) + \sum_{I_{n,k-1}} \sum_{j=1}^{k-1} f(X_{i_1}, \dots, X_{i_{k-1}}) g(X_{i_j}).$$

$\varphi(x_1, \dots, x_k) = f(x_1, \dots, x_{k-1}) g(x_k), \varphi_j(x_1, \dots, x_{k-1}) = f(x_1, \dots, x_{k-1}) g(x_j)$ とおくと，

$$\text{右辺} = (n)_k A_n(\varphi) + (n)_{k-1} \sum_{j=1}^{k-1} A_n(\varphi_j).$$

ここで，

$$\frac{(n)_{k-1} n}{(n)_k} = \frac{n}{n-k+1}, \ \frac{(n)_{k-1}}{(n)_k} = \frac{1}{n-k+1}$$

より，

$$A_n(\varphi) = \frac{n}{n-k+1} A_n(f) A_n(g) - \frac{1}{n-k+1} \sum_{j=1}^{k-1} A_n(\varphi_j).$$

$n \to \infty$ として，

$$E[f(X_1, \dots, X_{k-1}) g(X_k) \mid \mathcal{E}] = E[f(X_1, \dots, X_{k-1}) \mid \mathcal{E}] E[g(X_k) \mid \mathcal{E}].$$

この操作を繰り返すと，任意の有界可測関数 $f_j : S \to \mathbb{R}$ に対して，

$$E\left[\prod_{j=1}^{k} f_j(X_j) \mid \mathcal{E} \right] = \prod_{j=1}^{k} E[f_j(X_j) \mid \mathcal{E}] = \prod_{j=1}^{k} E[f_j(X_1) \mid \mathcal{E}]$$

となることが示される． □

S が標準 Borel 空間なら，$\mathcal{E}$ を与えたときの $(X_1, X_2, \dots)$ の r.c.d. が存在するので，de Finetti の定理は，$\mathcal{E}$ を条件付けたとき，$X_1, X_2, \dots$ が i.i.d. になることを述べている．特に，$X_i \in \{0, 1\}$ のとき，$P(X_1 = 1 \mid \mathcal{E})$ の d.f. を F とおくと，次の定理を得る．

Theorem 23.5. $X_1, X_2, \dots$ が exchangeable で，$\{0, 1\}$ に値をとるなら，$[0, 1]$ に集中した d.f. F が存在して，

$$P(X_1 = 1, \dots, X_k = 1, X_{k+1} = 0, \dots, X_n = 0) = \int_{[0,1]} \theta^k (1-\theta)^{n-k} dF(\theta)$$

となる．

Remark 23.2. 本節の最後に扱った，exchangeability や de Finitti の定理に関しては，Aldous (1985) が参考になる．

演習問題

Exercise 23.1. $X_1, X_2, \ldots$ を exchangeable な ($\mathbb{R}$-valued) r.v.'s とし，$E[X_1^2] = 1$ とする．

(a) $E[X_i X_j]$ $(i \neq j)$ の値は $i \neq j$ の選び方によらないことを示し，$E[X_1 X_2] \geq 0$ を示せ．

(b) 与えられた $\rho \in [0,1]$ に対して，$\rho = E[X_1 X_2]$ となるような $X_1, X_2, \ldots$ の例を与えよ．

Dirichlet 分布の定義．$Ga(\alpha, \beta)$ をシェイプパラメータ $\alpha \geq 0$，逆スケールパラメータ $\beta > 0$ をもつガンマ分布とする．ただし，$Ga(0, \beta)$ は 0 に退化した分布である．$\alpha > 0$ のとき，$Ga(\alpha, \beta)$ は密度関数 $\{\beta^\alpha/\Gamma(\alpha)\} x^{\alpha-1} e^{-\beta x} 1_{(0,\infty)}(x)$ をもつ．$\Gamma(\alpha)$ はガンマ関数である：$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx$.

いま，$\alpha_1, \ldots, \alpha_k \geq 0$ の少なくとも 1 つは正であるとし，$G_1, \ldots, G_k$ を独立な r.v.'s であって，$G_j \sim Ga(\alpha_j, 1)$ とする．ここで，$Y_j = G_j / \sum_{i=1}^k G_i$ $(j = 1, \ldots, k)$ とおいたとき，$(Y_1, \ldots, Y_k)$ の同時分布をパラメータ $(\alpha_1, \ldots, \alpha_k)$ をもつ Dirichlet 分布 と呼び，$D(\alpha_1, \ldots, \alpha_k)$ と表す．

Exercise 23.2. (Blackwell-MacQueen の壺) 壺のなかに k 色のボールが入っているとし，色を $1, \ldots, k$ でラベルする．色 j のボールは α_j 個あるとすると，壺のなかには最初 $\alpha_1 + \cdots + \alpha_k =: \alpha$ 個のボールが入っている．まず壺のなかのボールをランダムに 1 つ選び，その色の番号を X_1 とする．次に最初にとったボールを壺に戻し，さらに X_1 と同じ色のボールを 1 つ壺に追加し，再び壺のなかからランダムにボールを 1 つ選び，その色の番号を X_2 とする．この操作を繰り返して，r.v.'s $X_1, X_2, \ldots$ を得る．$n+1$ 回目の操作の直前に，壺に入っている色 j のボールの個数は $\alpha_j + \sum_{i=1}^n \delta_{X_i}(j)$ である．ただし，$\delta_{X_i}(j) = 1$ if $X_i = j$, $\delta_{X_i}(j) = 0$ if $X_i \neq j$ である．

(a) $(X_1, \ldots, X_n)$ の同時分布は

$$P(X_1 = j_1, \ldots, X_n = j_n) = \frac{1}{\alpha^n} \prod_{j=1}^k (\alpha_j)^{[n(j)]}, \quad 1 \leq j_1, \ldots, j_n \leq k$$

と表せることを示せ．ただし，$n(j) = \mathrm{Card}(\{i \in \{1, \ldots, n\} : j_i = j\}$, $(a)^{[m]} = a(a+1)\cdots(a+m-1)$ である．

(b) $X_1, X_2, \ldots$ が exchangeable であることを示せ．

(c) (b) と de Finetti の定理より，exchangeable σ-field $\mathcal{E}$ を条件付けたとき，$X_1, X_2, \ldots$ は i.i.d. である．そこで，$Y_j = P(X_1 = j \mid \mathcal{E})$ $(j = 1, \ldots, k)$ とおくと，$Y_1, \ldots, Y_k$ は $Y_1 + \cdots + Y_k = 1$ a.s. となる非負 r.v.'s であり，

$$P(X_1 = j_1, \ldots, X_n = j_n \mid \mathcal{E}) = Y_{j_1} \cdots Y_{j_n} = \prod_{j=1}^k Y_j^{n(j)}$$

をみたす．このとき，$(Y_1, \ldots, Y_k) \sim D(\alpha_1, \ldots, \alpha_k)$ を示せ．

ヒント：$(Z_1, \ldots, Z_k) \sim D(\alpha_1, \ldots, \alpha_k)$ とすると，$\forall \beta_1, \ldots, \beta_k \in \mathbb{Z}_+$, $E[\prod_{j=1}^k Y_k^{\beta_j}] = E[\prod_{j=1}^k Z_k^{\beta_j}]$ を示せ．

(d) $n \to \infty$ のとき，

$$\left(\frac{\alpha_1 + \sum_{i=1}^n \delta_{X_i}(1)}{\alpha + n}, \ldots, \frac{\alpha_k + \sum_{i=1}^n \delta_{X_i}(k)}{\alpha + n} \right) \to (Y_1, \ldots, Y_k) \quad a.s.$$

を示せ．

Exercise 23.3. 本節後半の設定に従うとし，$X_1, X_2, \ldots$ を i.i.d. とする．$\mathcal{A}$ を可測関数 $f : S \to \mathbb{R}$ からなる可算集合とし，$F(x) := \sup_{f \in \mathcal{A}} |f(x)|$ は $E[F(X_1)] < \infty$ をみたすとする．ここで，$Y_n = \sup_{f \in \mathcal{A}} |n^{-1} \sum_{i=1}^n \{f(X_i) - E[f(X_i)]\}|$ とおいたとき，$\{Y_{-n} : n \leq -1\}$ は $\{\mathcal{E}_{-n} : n \leq -1\}$ に関するバックワード・劣マルチンゲールであることを示せ．

24　任意停止定理

確率過程 $\{X_n : n \in \mathbb{Z}_+\}$ に対して，有限な $\lim_n X_n$ が a.s. に存在するときは，

$$X_\infty(\omega) = \begin{cases} \lim_n X_n(\omega) & \text{if 有限な } \lim_n X_n(\omega) \text{ が存在する} \\ 0 & \text{otherwise} \end{cases}$$

と定義しておく．$\{X_n\}$ が適合過程なら，X_∞ は存在すれば $\mathcal{F}_\infty (:= \bigvee_n \mathcal{F}_n)$ 可測な r.v. である．このとき，任意の停止時刻 N に対して，$N(\omega) = \infty$ のときは，$X_{N(\omega)}(\omega) = X_\infty(\omega)$ と定義しておけば，$X_{N(\omega)}(\omega)$ がすべての $\omega \in \Omega$ に対して定義される．

停止時刻 N に対して，

$$\mathcal{F}_N := \{A \in \mathcal{F}_\infty : A \cap \{N = n\} \in \mathcal{F}_n \ \forall n \in \mathbb{Z}_+\}$$

と定義する．$\mathcal{F}_N$ は σ-field である ($A \in \mathcal{F}_N \Rightarrow A^c \in \mathcal{F}_N$ だけ確認してみると，$A^c \cap \{N = n\} = \{N = n\} \setminus (A \cap \{N = n\}) \in \mathcal{F}_n$)．容易に確認できるように，$N \equiv n \in \mathbb{Z}_+ \cup \{+\infty\}$ なら，$\mathcal{F}_N = \mathcal{F}_n$ である．また，$\{N \le n\} \in \mathcal{F}_N$ であるから，N は $\mathcal{F}_N$ 可測である．

Lemma 24.1. M, N を停止時刻とする．(i) 適合過程 $\{X_n : n \in \mathbb{Z}_+\}$ に対して，$N < \infty$ a.s. であるか X_∞ が存在すれば，$X_N \in \mathcal{F}_N$．(ii) $M(\omega) \le N(\omega)$ $(\forall \omega \in \Omega)$ なら，$\mathcal{F}_M \subset \mathcal{F}_N$．(iii) $\mathcal{F}_M \cap \mathcal{F}_N = \mathcal{F}_{M \wedge N}$.

ところで，$N < \infty$ a.s. であっても $N(\omega) = \infty$ となる ω は存在するかもしれない．X_∞ が定義されない場合，$N(\omega) = \omega$ となる ω に対しては，$X_{N(\omega)}(\omega)$ は定義されないことになるが，その場合は X_N を a.s. に定義された r.v. とみなす．

Proof. 演習問題とする． □

本節の目標は任意停止定理 (optional stopping theorem) を証明することである．任意停止定理とは，おおまかに言って，適当な仮定のもとで，マルチンゲール性が確定的な時刻をランダムな停止時刻に置き換えても保存されることを主張するものである．以下，停止時刻 M, N に対して，$M \le N$ と書いたら，$M(\omega) \le N(\omega)$ $(\forall \omega \in \Omega)$ を意味するものとする ($M \le N$ a.s. ではない)．

Theorem 24.1 (任意停止定理)**.** $\{X_n : n \in \mathbb{Z}_+\}$ を確率過程とする．

(i) N を停止時刻とし，$\{X_{N \wedge n} : n \in \mathbb{Z}_+\}$ は一様可積分な劣マルチンゲールであるとする．このとき，$M \le N$ なる任意の停止時刻 M に対して，X_M は可積分であり，$X_M \le E[X_N \mid \mathcal{F}_M]$ a.s. が成り立つ．

(ii) $\{X_n\}$ が非負優マルチンゲールなら，任意の停止時刻 N に対して，X_N は可積分であり，任意の停止時刻 $M \le N$ に対して，$X_M \ge E[X_N \mid \mathcal{F}_M]$ a.s. が成り立つ．

(i) において，$M = N$ とすれば，X_N の可積分性も保証されている．もちろん，(i) において，$\{X_{N\wedge n}\}$ が一様可積分なマルチンゲールなら，$X_M = E[X_N \mid \mathcal{F}_M]$ a.s. が成り立つ ($\{X_n\}$ と $\{-X_n\}$ に対して，定理を適用すればよい)．しかし，(ii) において，$\{X_n\}$ が非負マルチンゲールであっても，$X_M = E[X_N \mid \mathcal{F}_M]$ となるとは限らない．例えば，$\{X_n\}$ を Example 21.4 に現れた非負マルチンゲールとすると，$X_\infty = 0$ であるが，$E[X_n] = E[X_0] = 1$ である ($M = 0, N = \infty$ とせよ)．

定理の証明の前に，$\{X_{N\wedge n}\}$ が一様可積分な劣マルチンゲールとなるための十分条件を述べておく．

Lemma 24.2. 劣マルチンゲール $\{X_n\}$ と停止時刻 N に対して，$\{X_{N\wedge n} : n \in \mathbb{Z}_+\}$ が一様可積分であるためには，次の (a)–(c) のいずれかの条件がみたされれば十分である．

(a) N が有界．

(b) $\{X_n\}$ が一様可積分．

(c) $N \in L^1$ であって，ある定数 B が存在して，$E[|X_{n+1} - X_n| \mid \mathcal{F}_n] \le B$ a.s. $\forall n$.

Proof. (a). $N \le k$ とすると，$|X_{N\wedge n}| \le \max_{0\le m\le k} |X_m|$ であり，右辺は可積分であるから，$\{X_{N\wedge n}\}$ は一様可積分である．

(b). $\{X_n^+\}$ は劣マルチンゲールであって，$N \wedge n$ は有界な停止時刻であるから，補題 22.1 より，$E[X_{N\wedge n}^+] \le E[X_n^+]$ である．右辺は n に関して有界であるから，$\{X_{N\wedge n}\}$ に対してマルチンゲール収束定理を適用して，$X_N = \lim_n X_{N\wedge n} \in L^1$ を得る．次に，

$$\begin{aligned}
&E[|X_{N\wedge n}|, \{|X_{N\wedge n}| > K\}] \\
&= E[|X_{N\wedge n}|, \{|X_{N\wedge n}| > K\} \cap \{N > n\}] + E[|X_{N\wedge n}|, \{|X_{N\wedge n}| > K\} \cap \{N \le n\}] \\
&\le E[|X_N|, \{|X_N| > K\}] + \sup_m E[|X_m|, \{|X_m| > K\}]
\end{aligned}$$

であり，$K \to \infty$ のとき，右辺 $\to 0$ となるから，求める結論を得る．

(c).

$$|X_{N\wedge n}| \le |X_0| + \sum_{m=0}^{N-1} |X_{m+1} - X_m| = |X_0| + \underbrace{\sum_{m=0}^{\infty} |X_{m+1} - X_m| 1_{\{N>m\}}}_{=:Y}.$$

$Y \in L^1$ を示せばよい．$\{N > m\} = \{N \le m\}^c \in \mathcal{F}_m$ より，

$$E[|X_{m+1} - X_m| 1_{\{N>m\}}] = E[E[|X_{m+1} - X_m| \mid \mathcal{F}_m], \{N > m\}] \le BP(N > m).$$

従って，$E[Y] \le B \sum_{m=0}^{\infty} P(N > m) = BE[N] < \infty$. □

特に，$\{X_n\}$ が一様可積分な (劣) マルチンゲールなら，あらゆる停止時刻 N に対して，$X_N = (\le)\ E[X_\infty \mid \mathcal{F}_N]$ a.s. となる．

任意停止定理の証明に移る．定理の証明は次の補題による．

Lemma 24.3. $\{X_n\}$ を適合過程とし，N を停止時刻とする．いま，$M \leq N$ なる任意の停止時刻 M に対して，X_M は (少なくとも a.s. に定義された) 可積分 r.v. であって，$E[X_M] \leq E[X_N]$ となると仮定する．このとき，$M \leq N$ なる任意の停止時刻 M に対して，$X_M \leq E[X_N \mid \mathcal{F}_M]$ a.s. となる．

Proof. $A \in \mathcal{F}_M$ に対して，$L = M1_A + N1_{A^c}$ とおくと，L は停止時刻である (Exercise 24.2)．仮定より，X_L, X_M, X_N は可積分であり，$E[X_N] \geq E[X_L] = E[X_M, A] + E[X_N, A^c]$ であるから，$E[X_M, A] \leq E[X_N, A]$ を得る．よって，$X_M \leq E[X_N \mid \mathcal{F}_M]$ a.s. □

Proof of Theorem 24.1. (i). $Y_n = X_{N\wedge n}$ とおく．$Y_M = X_M, Y_N = X_N$ であるから，Lemma 24.3 より，Y_M が可積分であり (よって，Y_N も可積分であり)，$E[Y_M] \leq E[Y_N]$ となることを示せばよい．$\{Y_n\}$ は一様可積分な劣マルチンゲールであるから，Lemma 24.2 より，$\{Y_{M\wedge n} : n \in \mathbb{Z}_+\}$ は一様可積分である．特に，Y_M は可積分である．また，Lemma 22.1 より，$E[Y_{M\wedge n}] \leq E[Y_n]$ であり，$n \to \infty$ のとき，$E[Y_{M\wedge n}] \to E[Y_M], E[Y_n] \to E[Y_\infty] = E[Y_N]$ となるから，$E[Y_M] \leq E[Y_N]$ を得る．

(ii). $\{X_n\}$ が非負優マルチンゲールなら，任意の停止時刻 N に対して，$E[X_{N\wedge n}] \leq E[X_0]$ であるから，Fatou の補題より，$E[X_N] \leq \liminf_n E[X_{N\wedge n}] \leq E[X_0]$ を得る．次に，$M \leq N$ を停止時刻とすると，$Y_n = X_{N\wedge n}$ は優マルチンゲールであるから，Lemma 22.1 より，$E[X_{M\wedge n}] = E[Y_{M\wedge n}] \geq E[Y_n] = E[X_{N\wedge n}]$. 両辺から $E[X_n 1_{\{M\geq n\}}]$ を引いて，

$$E[X_M 1_{\{M<n\}}] \geq E[X_N 1_{\{N<n\}}] + E[X_n 1_{\{M<n\leq N\}}]$$

を得る．ここで，Fatou の補題より，

$$\liminf_n E[X_n 1_{\{M<n\leq N\}}] \geq E[X_\infty 1_{\{M<\infty, N=\infty\}}]$$

であり，MCT より，$E[X_M 1_{\{M<n\}}] \uparrow E[X_M 1_{\{M<\infty\}}], E[X_N 1_{\{N<n\}}] \uparrow E[X_N 1_{\{N<\infty\}}]$ であるから，

$$E[X_M 1_{\{M<\infty\}}] \geq E[X_N 1_{\{N<\infty\}}] + E[X_\infty 1_{\{M<\infty, N=\infty\}}]$$

を得る．両辺に $E[X_M 1_{\{M=\infty\}}] = E[X_N 1_{\{M=\infty\}}]$ を加えて ($M \leq N$ であるから，$M = \infty$ なら $N = \infty$ である)，$E[X_M] \geq E[X_N]$ を得る．あとは Lemma 24.3 より，$X_M \geq E[X_N \mid \mathcal{F}_M]$ a.s. を得る． □

次の系の証明は演習問題とする．

Corollary 24.1. $\{X_n\}$ を劣マルチンゲールとし，ある可積分な r.v. Y が存在して，$X_n \leq E[Y \mid \mathcal{F}_n]$ とする．このとき，X_∞ が存在して，任意の停止時刻 $M \leq N$ に対して，X_M と X_N は可積分であって，$X_M \leq E[X_N \mid \mathcal{F}_M]$ a.s. が成り立つ．

Example 24.1. $Y \in L^1$ に対して，$X_n = E[Y \mid \mathcal{F}_n]$ とすると，停止時刻 N に対して，

$$X_N = E[Y \mid \mathcal{F}_N] \quad a.s.$$

である．なぜなら，$\{X_n\}$ は一様可積分なマルチンゲールであり，Lévy の上向き定理より，$X_\infty = E[Y \mid \mathcal{F}_\infty]$ であるから，任意停止定理より，$X_N = E[E[Y \mid \mathcal{F}_\infty] \mid \mathcal{F}_N] = E[Y \mid \mathcal{F}_N]$ a.s. である．

次に，停止時刻 M, N に対して，

$$E[E[Y \mid \mathcal{F}_N] \mid \mathcal{F}_M] = E[Y \mid \mathcal{F}_{M\wedge N}] \quad a.s.$$

を示そう．$X_{N\wedge n}$ は一様可積分なマルチンゲールであり，$\lim_n X_{N\wedge n} = X_N$ であるから，任意停止定理より，$X_{N\wedge M} = E[X_N \mid \mathcal{F}_M] = E[E[Y \mid \mathcal{F}_N] \mid \mathcal{F}_M]$ a.s. を得る．

任意停止定理の応用として，Wald の等式を証明する．

Theorem 24.2 (Wald の等式)**.** $\{\xi_m\}$ を i.i.d. r.v.'s とし，$E[|\xi_m|] < \infty$ とする．また，$S_0 = 0, S_n = \sum_{m=1}^n \xi_m, \mathcal{F}_0 = \{\varnothing, \Omega\}, \mathcal{F}_n = \sigma(\xi_1, \ldots, \xi_n)$ とおく．このとき，任意の可積分な停止時刻 N に対して，S_N は可積分であって，$E[S_N] = E[\xi_1]E[N]$ となる．

Proof. $X_n = S_n - nE[\xi_1]$ とおくと，$\{X_n\}$ はマルチンゲールであり，$E[|X_{n+1} - X_n| \mid \mathcal{F}_n] = E[|\xi_{n+1} - E[\xi_1]|] \le 2E[|\xi_1|]$ であるから，Lemma 24.2 と任意停止定理より，S_N は可積分であって，$0 = E[X_0] = E[X_N] = E[S_N] - E[N]E[\xi_1]$ を得る． □

Theorem 24.3. $\{\xi_m\}$ を i.i.d. r.v.'s とし，$P(\xi_m = \pm 1) = 1/2$ とする．$S_0 = 0, S_n = \sum_{m=1}^n \xi_n$ とおいて，整数 $a < 0 < b$ に対して，$N = \inf\{n \ge 1 : S_n \in \{a, b\}\}$ と定義すると，$E[N] < \infty$ であり，

$$P(S_N = b) = \frac{-a}{b-a},\ P(S_N = a) = \frac{b}{b-a}$$

となる．

Proof. $x \in (a, b)$ から出発して $(b-a)$ 回 $+1$ 進めば必ず (a, b) から外に出るので，

$$P(x + S_{b-a} \notin (a, b)) \ge P(\xi_1 = \cdots = \xi_{b-a}) = 2^{-(b-a)}$$

となる．この評価を使うと，

$$\begin{aligned}
P(N > n(b-a)) &= P(S_m \in (a, b), 1 \le \forall m \le n(b-a)) \\
&\le P(S_{m(b-a)} \in (a, b), 1 \le \forall m \le n) \\
&= E[P(S_{n(b-a)} \in (a, b) \mid \mathcal{F}_{(n-1)(b-a)}), \{S_{m(b-a)} \in (a, b), 1 \le \forall m \le n-1\}] \\
&\le (1 - 2^{-(b-a)}) P(S_{m(b-a)} \in (a, b), 1 \le \forall m \le n-1)) \\
&\ \ \vdots \\
&\le (1 - 2^{-(b-a)})^n.
\end{aligned}$$

従って，$E[N] < \infty$. 一方，Wald の等式より，

$$0 = E[\xi_1]E[N] = E[S_N] = bP(S_N = b) + aP(S_N = a).$$

これと $P(S_N = a) + P(S_N = b) = 1$ を連立させて，最後の主張を得る． □

演習問題

Exercise 24.1. Lemma 24.1 を示せ.

Exercise 24.2. $M \leq N$ を停止時刻とし，任意の $A \in \mathcal{F}_M$ に対して，$L = M1_A + N1_{A^c}$ とおくと，L は停止時刻であることを示せ.

Exercise 24.3. Corollary 24.1 を示せ.

Exercise 24.4. 一様可積分な劣マルチンゲールに対する任意停止定理は Doob 分解を経由しても証明できる．$\{X_n\}$ を一様可積分な劣マルチンゲールとし，$X_n = M_n + A_n$ を Doob 分解とする．$\{A_n\}$ は増加過程であるから，a.s. に $A_\infty = \lim_n A_n$ が存在する.

(a) $E[A_\infty] < \infty$ と $\{M_n\}$ が一様可積分であることを示せ.

(b) N を停止時刻として，$F \in \mathcal{F}_N$ に対して，

$$E[X_N, F] = E[M_\infty, F] + E[A_N, F]$$

を示せ．これから，

$$X_N = E[M_\infty \mid \mathcal{F}_N] + A_N \quad a.s.$$

を得る.

(b) より，停止時刻 $M \leq N$ に対して，$\mathcal{F}_M \subset \mathcal{F}_N$ より，$E[X_N \mid \mathcal{F}_M] = E[M_\infty \mid \mathcal{F}_M] + E[A_N \mid \mathcal{F}_M] \geq E[M_\infty \mid \mathcal{F}_M] + A_M = X_M$ a.s. を得る.

Exercise 24.5. N を停止時刻とし，ある $m \in \mathbb{N}$ と $\varepsilon \in (0,1)$ が存在して，$P(N \leq n+m \mid \mathcal{F}_n) > \varepsilon$ a.s. $(\forall n \in \mathbb{Z}_+)$ とする.

(a) $P(N > km) \leq (1-\varepsilon)^k$ $(k = 1, 2, \dots)$ を示せ．ヒント：帰納法.

(b) $E[N] < \infty$ を示せ.

Exercise 24.6 (Wald の第 2 等式)**.** $\{\xi_m\}$ を i.i.d. r.v.'s とし，$E[\xi_m] = 0,\ E[\xi_m^2] < \infty$ とする．$S_0 = 0, S_n = \sum_{m=1}^n \xi_n, \mathcal{F}_0 = \{\varnothing, \Omega\}, \mathcal{F}_n = \sigma(\xi_1, \dots, \xi_n)$ とすると，任意の可積分な停止時刻 N に対して，$\mathrm{Var}(S_N) = \mathrm{Var}(\xi_1)E[N]$ を示せ.

25　マルチンゲールに関するその他の話題

マルチンゲール CLT

各 $n \in \mathbb{N}$ に対して，$\mathcal{F}_{n,m} \subset \mathcal{F}, 0 \le m \le n$ を σ-fields の増加列とする．$X_{n,m}, 1 \le m \le n$ が $\mathcal{F}_{n,m}$ に関する マルチンゲール差分アレイ (martingale difference array) であるとは，$X_{n,m} \in L^1, X_{n,m} \in \mathcal{F}_{n,m}, E[X_{n,m} \mid \mathcal{F}_{n,m-1}] = 0, 1 \le m \le n$ となることを言う．

Theorem 25.1 (McLeish (1974))**.** $X_{n,m}, 1 \le m \le n$ を $\mathcal{F}_{n,m}$ に関するマルチンゲール差分アレイとし，次の 3 条件を仮定する．

(i) $\sup_n E[\max_{1\le m\le n} X_{n,m}^2] < \infty$.

(ii) $\sum_{m=1}^n X_{n,m}^2 \xrightarrow{P} 1$.

(iii) $\max_{1\le m\le n} |X_{n,m}| \xrightarrow{P} 0$.

このとき，$S_n = \sum_{m=1}^n X_{n,m} \xrightarrow{d} N(0,1)$.

Theorem 25.1 の証明は次の補題による

Lemma 25.1. $X_{n,m}, 1 \le m \le n$ を (マルチンゲール差分とは限らない) 三角列とし，$T_n(t) = \prod_{m=1}^n (1 + itX_{n,m}),\ t \in \mathbb{R}$ とおき，次の 3 条件を仮定する．

(a) 任意の $t \in \mathbb{R}$ に対して，$E[T_n(t)] = 1$, かつ $\{T_n(t) : n \in \mathbb{N}\}$ は一様可積分．

(b) $\sum_{m=1}^n X_{n,m}^2 \xrightarrow{P} 1$.

(c) $\max_{1\le m\le n} |X_{n,m}| \xrightarrow{P} 0$.

このとき，$S_n = \sum_{m=1}^n X_{n,m} \xrightarrow{d} N(0,1)$.

Proof. $t \ne 0$ を任意に固定し，$T_n(t)$ を T_n と書く．$I_n = e^{itS_n}$ とおく．$E[I_n] \to e^{-t^2/2}$ を示せばよい．まず，$r(x) = ix + x^2/2 - \mathrm{Log}(1+ix)$ $(x \in \mathbb{R})$ とおくと，$r(x) = O(|x|^3)$ $(x \to 0)$ である．両辺の指数をとって整理して，$e^{ix} = (1+ix)\exp\{-x^2/2 + r(x)\}$ を得る．$U_n = \exp\{-t^2 \sum_{m=1}^n X_{n,m}^2/2 + \sum_{m=1}^n r(tX_{n,m})\}$ とおくと，$I_n = T_n U_n$ となる．ここで，$\max_{1\le m\le n} |X_{n,m}| \le 1/(2|t|)$ のとき，

$$\sum_{m=1}^n |r(tX_{n,m})| \le O(1)|t|^3 \sum_{m=1}^n |X_{n,m}|^3 \le O(1)|t|^3 (\max_{1\le m\le n} |X_{n,m}|) \sum_{m=1}^n X_{n,m}^2 \xrightarrow{P} 0.$$

よって，$U_n \xrightarrow{P} e^{-t^2/2}$ となる．$\{T_n\}$ は一様可積分なので，特に $\sup_n E[|T_n|] < \infty$ であるから，$I_n - T_n e^{-t^2/2} = T_n(U_n - e^{-t^2/2}) \xrightarrow{P} 0$．あとは左辺が一様可積分であることを示せば，$E[I_n - T_n e^{-t^2/2}] = E[I_n] - e^{t^2/2} \to 0$ となって，補題の結論を得る．しかし，$\{I_n - T_n e^{-t^2/2} : n \in \mathbb{N}\}$ が一様可積分であることは，$|I_n| = 1, e^{-t^2/2} \le 1$ と，$\{T_n\}$ が一様可積分であることから従う．　□

Proof of Theorem 25.1. $Z_{n,m} = X_{n,m}1_{\{\sum_{k=1}^{m-1} X_{n,k}^2 \le 2\}}$ とおくと，$Z_{n,m}, 1 \le m \le n$ も $\mathcal{F}_{n,m}$ に関するマルチンゲール差分アレイであり，$P(Z_{n,m} \ne X_{n,m}, 1 \le \exists m \le n) \le P(\sum_{m=1}^n X_{n,m}^2 > 2) \to 0$ であるから，$Z_{n,m}$ も (i)–(iii) をみたし，$P(\sum_{m=1}^n Z_{n,m} \ne S_n) \to 0$ となる. 従って，$\sum_{m=1}^n Z_{n,m} \xrightarrow{d} N(0,1)$ を示せばよい．$Z_{n,m}$ に対して，Lemma 25.1 の条件 (a)–(c) を確認する．(b) と (c) は仮定からみたされるので，(a) のみ確認すればよい．$t \in \mathbb{R}$ を任意に固定して，$T_n = \prod_{m=1}^n (1 + itZ_{n,m})$ とおくと，

$$E[T_n] = E[E[T_n \mid \mathcal{F}_{n,n-1}]] = E\left[\prod_{m=1}^{n-1}(1 + itZ_{n,m})\right] = \cdots = 1$$

となる. 次に，

$$J_n = \begin{cases} \min\{1 \le m \le n : \sum_{k=1}^m X_{n,k}^2 > 2\} & \text{if } \sum_{m=1}^n X_{n,m}^2 > 2 \\ n & \text{otherwise} \end{cases}$$

とおくと，

$$\begin{aligned}
E[|T_n|^2] &\le 2E\left[\prod_{m=1}^{n}(1 + t^2 Z_{n,m}^2)\right] \quad (\because (a+b)^2 \le 2a^2 + 2b^2) \\
&= 2E\left[\prod_{m=1}^{J_n}(1 + t^2 Z_{n,m}^2)\right] = 2E\left[\prod_{m=1}^{J_n}(1 + t^2 X_{n,m}^2)\right] \\
&= 2E\left[\prod_{m=1}^{J_n-1}(1 + t^2 X_{n,m}^2)(1 + t^2 X_{n,J_n}^2)\right] \\
&\le 2E\left[\exp\left(t^2 \sum_{m=1}^{J_n-1} X_{n,m}^2\right)(1 + t^2 X_{n,J_n}^2)\right] \quad (\because 1 + x \le e^x) \\
&\le 2e^{2t^2}\left(1 + t^2 E[X_{n,J_n}^2]\right) \le 2e^{2t^2}\left(1 + t^2 E[\max_{1\le m\le n} X_{n,m}^2]\right) = O(1).
\end{aligned}$$

従って，$\{T_n\}$ は一様可積分である． □

Remark 25.1. マルチンゲール CLT を含む，マルチンゲールに対する極限定理とその応用に関しては，Hall and Heyde (1980) が詳しい．

Hardy-Littlewood の最大不等式と Lebesgue の微分定理

本節では (劣) マルチンゲールを σ-finite な測度にまで拡張して，その応用として，Hardy-Littlewood の最大不等式のマルチンゲールを用いた証明を与える．また，Hardy-Littlewood の最大不等式から Lebesgue の微分定理を導く．

$(\Omega, \mathcal{F}, \mu)$ を $0 < \mu(\Omega) \leq \infty$ なる測度空間とし，$X \in L^1(\mu)$ に対して，

$$E^\mu[X] = \int X d\mu$$

と書く．μ が p.m. の場合と同様にして，$A \in \mathcal{F}$ に対して，

$$E^\mu[X, A] = \int_A X d\mu$$

とも書く．$\mathcal{G}$ を $\mathcal{F}$ の sub σ-field であって，$\mu|_{\mathcal{G}}$ は σ-finite とする．このとき，$Y \in \mathcal{G}, Y \in L^1(\mu), E^\mu[Y, A] = E^\mu[X, A]$ $(\forall A \in \mathcal{G})$ をみたす任意の r.v. Y を，$\mathcal{G}$ を与えたときの μ 条件付き期待値と呼び，$E^\mu[X \mid \mathcal{G}]$ と書く．μ が p.m. の場合と全く同様の議論により，μ 条件付き期待値の存在と μ-a.e. の意味での一意性が従う．

$\{\mathcal{F}_n : n \in \mathbb{Z}_+\}$ をフィルトレーションとし，$\mu|_{\mathcal{F}_0}$ は σ-finite とする．このとき，r.v.'s $\{X_n : n \in \mathbb{Z}_+\}$ が $\{\mathcal{F}_n\}$ に関する μ マルチンゲールであるとは，$\{X_n\}$ が $\{\mathcal{F}_n\}$ に適合していて，$X_n \in L^1(\mu)$ $(\forall n \in \mathbb{Z}_+)$, かつ $E^\mu[X_{n+1} \mid \mathcal{F}_n] = X_n$ μ-a.e. $(\forall n \in \mathbb{Z}_+)$ となることを言う．μ マルチンゲールの最後の条件において，代わりに $E^\mu[X_{n+1} \mid \mathcal{G}] \geq X_n$ μ-a.e. $(\forall n \in \mathbb{Z}_+)$ が成り立つとき，$\{X_n\}$ を μ 劣マルチンゲールと言う．

p.m. に対する (劣) マルチンゲールに対して成立する多くの性質は，μ(劣) マルチンゲールに対しても成り立つ．例えば，次の Doob の不等式を証明してみよう．

Theorem 25.2. $\{X_n\}$ を μ 劣マルチンゲールとし，$\bar{X}_n = \max_{0 \leq m \leq n} X_m$ とおく．このとき，任意の $n \in \mathbb{Z}_+$ と $\alpha > 0$ に対して，

$$\alpha\mu(\bar{X}_n \geq \alpha) \leq E^\mu[X_n, \{\bar{X}_n \geq \alpha\}] \leq E^\mu[X_n^+].$$

Proof. 問題となるのは $\mu(\Omega) = \infty$ の場合のみである．$\Omega' \in \mathcal{F}_0$ を $0 < \mu(\Omega') < \infty$ とし，

$$\mathcal{F}' = \{E \cap \Omega' : E \in \mathcal{F}\},\ \mathcal{F}'_n = \{E \cap \Omega' : E \in \mathcal{F}_n\},\ X'_n = X_n|_{\Omega'},\ P = \frac{\mu|_{\mathcal{F}'}}{\mu(\Omega')}$$

とおくと，$\{X'_n\}$ は確率空間 $(\Omega', \mathcal{F}', P)$ 上に定義された $\{\mathcal{F}'_n\}$ に関する劣マルチンゲールである．ゆえに，Doob の不等式 (Theorem 22.1) より，

$$\alpha P\left(\max_{0 \leq m \leq n} X'_n \geq \alpha\right) \leq E^P\left[X'_n, \{\max_{0 \leq m \leq n} X'_m \geq \alpha\}\right]$$

を得る．これを書き直して，$\alpha\mu(\{\bar{X}_n \geq \alpha\} \cap \Omega') \leq E^\mu[X_n, \{\bar{X}_n \geq \alpha\} \cap \Omega']$ を得る．そこで，集合列 $\Omega_\ell \in \mathcal{F}_0$ を $\mu(\Omega_\ell) < \infty$ $(\forall \ell)$, $\Omega_\ell \uparrow \Omega$ となるように選ぶと，$\alpha\mu(\{\bar{X}_n \geq \alpha\} \cap \Omega_\ell) \leq E^\mu[X_n, \{\bar{X}_n \geq \alpha\} \cap \Omega_\ell]$ であり，あとは $\ell \to \infty$ として求める結論を得る． □

以下，$(\Omega, \mathcal{F}, \mu) = (\mathbb{R}^k, \mathcal{B}^k, \lambda^k)$ とし，$L^1 = L^1(\mathbb{R}^k, \mathcal{B}^k, \lambda^k)$ とする．

$$Q = \prod_{j=1}^k [a_j, a_j + r),\ \boldsymbol{a} = (a_1, \ldots, a_k) \in \mathbb{R}^k, r > 0$$

なる形の集合を $\mathbb{R}^k$ の 立方体 (cube) と呼ぶ．このとき，$f \in L^1$ に対して，

$$Mf(\boldsymbol{x}) := \sup_{Q \ni \boldsymbol{x}} \frac{1}{\lambda^k(Q)} \int_Q |f(\boldsymbol{y})| d\boldsymbol{y},\ \boldsymbol{x} \in \mathbb{R}^k$$

と定義する．ただし，$\sup_{Q \ni \boldsymbol{x}}$ は $\boldsymbol{x}$ を含む $\mathbb{R}^k$ の立方体 Q に関してとるものとする．簡単にわかるように，$\mathbb{R}^k \ni \boldsymbol{x} \mapsto Mf(\boldsymbol{x}) \in [0, \infty]$ は下半連続であるから，特に可測である．

Theorem 25.3 (Hardy-Littlewood の最大不等式)**.** $f \in L^1$ に対して，

$$\lambda^k(Mf > \alpha) \leq \frac{(12)^k}{\alpha} E^{\lambda^k}[|f|],\ \forall \alpha > 0.$$

Proof. $n \in \mathbb{Z}, \boldsymbol{\ell} = (\ell_1, \ldots, \ell_k) \in \mathbb{Z}^k$ に対して，

$$C_n(\boldsymbol{\ell}) = \prod_{j=1}^k \left[\frac{\ell_j}{2^n}, \frac{\ell_j + 1}{2^n} \right)$$

とおいて，各 $\boldsymbol{\eta} \in \{0,1\}^k$ に対して，$\mathbb{R}^k$ の分割

$$\mathcal{P}_n(\boldsymbol{\eta}) = \left\{ \frac{(-1)^n \boldsymbol{\eta}}{3 \cdot 2^n} + C_n(\boldsymbol{\ell}) : \boldsymbol{\ell} \in \mathbb{Z}^k \right\}$$

を考える．このとき，$\mathcal{P}_{n+1}(\boldsymbol{\eta})$ は $\mathcal{P}_n(\boldsymbol{\eta})$ の細分になっている．すなわち，$\mathcal{F}_n^{(\boldsymbol{\eta})} = \sigma(\mathcal{P}_n(\boldsymbol{\eta}))$ とおくと，$\mathcal{F}_n^{(\boldsymbol{\eta})} \subset \mathcal{F}_{n+1}^{(\boldsymbol{\eta})}$ となる．そこで，

$$X_n^{(\boldsymbol{\eta})} = E^{\lambda^k}[|f| \mid \mathcal{F}_n^{(\boldsymbol{\eta})}]$$

とおくと，任意に固定した $m \in \mathbb{Z}$ に対して，$\{X_{n+m}^{(\boldsymbol{\eta})} : n \in \mathbb{Z}_+\}$ は $\{\mathcal{F}_{n+m}^{(\boldsymbol{\eta})} : n \in \mathbb{Z}_+\}$ に関する λ^k マルチンゲールである．従って，Theorem 25.2 より，任意の $N \geq m$ に対して，

$$\lambda^k(\max_{m \leq n \leq N} X_n^{(\boldsymbol{\eta})} > \alpha) \leq \alpha^{-1} E^{\lambda^k}[|f|] \qquad (*)$$

を得る．ここで，$\boldsymbol{x} \in Q \in \mathcal{P}_n(\boldsymbol{\eta})$ に対して，

$$X_n^{(\boldsymbol{\eta})}(\boldsymbol{x}) = \frac{1}{\lambda^k(Q)} \int_Q |f(\boldsymbol{y})| d\boldsymbol{y}$$

であるから，

$$\max_{m \leq n \leq N} X_n^{(\boldsymbol{\eta})}(\boldsymbol{x}) = \max \left\{ \frac{1}{\lambda^k(Q)} \int_Q |f(\boldsymbol{y})| d\boldsymbol{y} : Q \ni \boldsymbol{x}, Q \in \bigcup_{m \leq n \leq N} \mathcal{P}_n(\boldsymbol{\eta}) \right\}.$$

ゆえに，

$$M^{(\boldsymbol{\eta})} f(\boldsymbol{x}) := \sup \left\{ \frac{1}{\lambda^k(Q)} \int_Q |f(\boldsymbol{y})| d\boldsymbol{y} : Q \ni \boldsymbol{x}, Q \in \bigcup_{n \in \mathbb{Z}} \mathcal{P}_n(\boldsymbol{\eta}) \right\}$$

とおくと，$(*)$ において $m \to -\infty, N \to \infty$ として，

$$\lambda^k(M^{(\boldsymbol{\eta})} f > \alpha) \le \alpha^{-1} E^{\lambda^k}[|f|]$$

を得る．

いま，立方体 $Q = \prod_{j=1}^k [a_j, a_j + r)$ に対して，$1/(3 \cdot 2^{n+1}) < r \le 1/(3 \cdot 2^n)$ であれば，適当な $\boldsymbol{\eta} \in \{0,1\}^k$ と $C \in \mathcal{P}_n(\boldsymbol{\eta})$ が存在して，$Q \subset C$ となる ($k = 1$ の場合を考えてみればよい)．この Q と C に対して，

$$\frac{\lambda^k(C)}{\lambda^k(Q)} \le \frac{(3 \cdot 2^{n+1})^k}{2^{kn}} = 6^k$$

であるから，

$$\max_{\boldsymbol{\eta} \in \{0,1\}^k} M^{(\boldsymbol{\eta})} f \le Mf \le 6^k \max_{\boldsymbol{\eta} \in \{0,1\}^k} M^{(\boldsymbol{\eta})} f$$

を得る．以上より，定理の結論を得る． □

Theorem 25.4 (Lebesgue の微分定理)．$f \in L^1$ に対して，

$$\lim_{B \downarrow \{\boldsymbol{x}\}} \frac{1}{\lambda^k(B)} \int_B |f(\boldsymbol{y}) - f(\boldsymbol{x})| d\boldsymbol{y} = 0 \quad \lambda^k\text{-a.e. } \boldsymbol{x} \in \mathbb{R}^k. \tag{$**$}$$

ただし，この意味は，λ^k-a.e. $\boldsymbol{x} \in \mathbb{R}^k$ と任意の $\varepsilon > 0$ に対して，ある $\delta > 0$ が存在して，B が $\boldsymbol{x}$ を含む $\mathbb{R}^k$ の球であって，B の半径が δ より小さいなら，$(1/\lambda^k(B)) \int_B |f(\boldsymbol{y}) - f(\boldsymbol{x})| d\boldsymbol{y} < \varepsilon$ となることである．特に，λ^k-a.e. $\boldsymbol{x} \in \mathbb{R}^k$ に対して，

$$f(\boldsymbol{x}) = \lim_{B \downarrow \{\boldsymbol{x}\}} \frac{1}{\lambda^k(B)} \int_B f(\boldsymbol{y}) d\boldsymbol{y}.$$

Proof. まず，

$$\tilde{M} f(\boldsymbol{x}) := \sup_{B \ni \boldsymbol{x}} \frac{1}{\lambda^k(B)} \int_B |f(\boldsymbol{y})| d\boldsymbol{y}, \ \boldsymbol{x} \in \mathbb{R}^k$$

とおく．ただし，$\sup_{B \ni \boldsymbol{x}}$ は $\boldsymbol{x}$ を含む $\mathbb{R}^k$ の球に関してとるものとする．ここで，κ_k を $\mathbb{R}^k$ の単位球の体積として，$c_k = 2^k/\kappa_k$ とおくと，

$$\tilde{M} f(\boldsymbol{x}) \le c_k M f(\boldsymbol{x}), \ \boldsymbol{x} \in \mathbb{R}^k.$$

次に，$f \in C_c(\mathbb{R}^k)$ に対して，$(**)$ はすべての $\boldsymbol{x} \in \mathbb{R}^k$ に対して成り立つ．そこで，$\{f_n\} \subset C_c(\mathbb{R}^k)$ を $E^{\lambda^k}[|f_n - f|] \to 0$ となるように選ぶと，任意の $\varepsilon > 0$ に対して，

$$E := \left\{ \boldsymbol{x} : \exists B_m \ni \boldsymbol{x} \text{ s.t. } \lim_m \lambda^k(B_m) = 0 \ \& \ \limsup_m \frac{1}{\lambda^k(B_m)} \int_{B_m} |f(\boldsymbol{y}) - f(\boldsymbol{x})| d\boldsymbol{y} \ge \varepsilon \right\}$$
$$\subset \left\{ \boldsymbol{x} : \tilde{M}(f - f_n)(\boldsymbol{x}) \ge \frac{\varepsilon}{3} \right\} \bigcup \left\{ \boldsymbol{x} : |f_n(\boldsymbol{x}) - f(\boldsymbol{x})| \ge \frac{\varepsilon}{3} \right\}.$$

右辺の Lebesgue 測度は

$$\frac{3}{\varepsilon}(1 + (12)^k c_k) E^{\lambda^k}[|f_n - f|]$$

でバウンドできるので，$\lambda^k(E) = 0$ を得る．以上より，定理の結論を得る． □

(∗∗) が成り立つ $\boldsymbol{x} \in \mathbb{R}^k$ の集合を f の Lebesgue 集合 と呼ぶ.

Remark 25.2. 上で紹介した Hardy-Littlewood の最大不等式の証明は Stroock (2011, Section 6.1) を参考にした．より標準的な Vitali の被覆補題を用いる証明は Folland (1999, Section 3.4) などにある.

$F : \mathbb{R} \to \mathbb{R}$ を非減少関数とし，$G(x) := F(x+) := \lim_{y \downarrow x} F(y), x \in \mathbb{R}$ と定義すれば，G は右連続かつ非減少である．G とその Lebesgue-Stieltjes 測度を同一視して，G の Lebesgue 分解を $G = G_{ac} + G_s$ とする．ここで，G_{ac} と G_s は G の絶対連続部分と特異部分である．Radon-Nikodym の定理より，ある可測関数 $f : \mathbb{R} \to \mathbb{R}_+$ が存在して，任意の $-\infty < a < b < \infty$ に対して，$G_{ac}(b) - G_{ac}(a) = \int_a^b f(x)dx$ となる．Lebesgue の微分定理より，G_{ac} は a.e. に微分可能であり，$G'_{ac} = f$ a.e. となる．さらに，G_s も a.e. に微分可能であり，$G'_s = 0$ a.e. となることが示される (Folland, 1999, Section 3.4 を参照せよ)．よって，G は a.e. に微分可能であり，$G' = f$ a.e. である．F は高々可算個の点を除いて連続であるから，それらの点を除いて，$F = G$ である．よって，F も a.e. に微分可能であり，$F' = f$ a.e. である．すなわち，次の定理を得る.

Theorem 25.5. $F : \mathbb{R} \to \mathbb{R}$ を非減少関数とすると，F は a.e. に微分可能である．さらに，$F(\bullet+)$ の絶対連続部分の Lebesgue 測度に関する密度を f とおくと，$F' = f$ a.e. である.

演習問題

Exercise 25.1. $\boldsymbol{x} \in \mathbb{R}^k, r > 0$ に対して，$B(\boldsymbol{x}, r)$ を $\boldsymbol{x}$ を中心とする半径 r の開球とする．$E \in \mathcal{B}^k$ に対して，

$$D_E(\boldsymbol{x}) := \lim_{r \downarrow 0} \frac{\lambda^k(E \cap B(\boldsymbol{x}, r))}{\lambda^k(B(\boldsymbol{x}, r))}$$

が存在するとき，$D_E(\boldsymbol{x})$ を E の $\boldsymbol{x}$ における密度と呼ぶ．$D_E(\boldsymbol{x}) = 1$ a.e. $\boldsymbol{x} \in E, D_E(\boldsymbol{x}) = 0$ a.e. $x \in E^c$ を示せ.

Exercise 25.2. $1 < p < \infty$ に対して，

$$(E^{\lambda^k}[(Mf)^p])^{1/p} \leq \frac{(12)^k p}{p-1} (E^{\lambda^k}[|f|^p])^{1/p}$$

を示せ.

Exercise 25.3. 前問の不等式は $p = 1$ の場合には拡張できない．$k = 1$ とし，$f : \mathbb{R} \to \mathbb{R}$ を a.e. に 0 とはならない可測関数とする．このとき，$\liminf_{|x| \to \infty} |x| Mf(x) > 0$ を示せ．よって，$E^{\lambda}[Mf] = \infty$ となる.

Exercise 25.4 (Fubini の微分定理)**.** $F_n : \mathbb{R} \to \mathbb{R}, n \in \mathbb{N}$ を非減少とし，$G(x) := \sum_{n=1}^{\infty} F_n(x)$ はすべての $x \in \mathbb{R}$ に対して有限の値に収束しているとする．このとき，$G' = \sum_{n=1}^{\infty} F'_n$ a.e. を示せ.

ヒント：まず，$F_n, n \in \mathbb{N}$ を右連続かつ $F_n(-\infty) = 0$ と仮定せよ．このとき，G も右連続であり，$G_{ac} = \sum_n F_{n,ac}$ である．さらに，$G'_{ac} = \sum_n F'_{n,ac}$ a.e. である．

Exercise 25.5. F を Cantor 分布の d.f. とする．$\mathbb{Q} = \{q_1, q_2, \dots\}$ とすると，

$$G(x) = \sum_{n=1}^{\infty} \frac{1}{2^n} F(q_n + x),\ x \in \mathbb{R}$$

は連続な d.f. であり，そのサポートは $\mathbb{R}$ であることを示せ．また，$G' = 0$ a.e. を示せ．

Part V. マルコフ・チェイン

本パートは離散時間の確率過程の代表的なクラスである，マルコフ・チェインをカバーする．統計学においては，マルコフ連鎖モンテカルロ法との関連が特に重要であろう．本パートでは，次節のあと，可算な状態空間をもつマルコフ・チェインとその極限定理を考察する．一般の状態空間をもつマルコフ・チェインとその極限定理に関しては，Meyn and Tweedie (2009) が詳しい．

26　マルコフ・チェイン

$(S, \mathcal{S})$ を標準 Borel 空間とする．関数 $p : S \times \mathcal{S} \to [0, 1]$ が次の 2 条件をみたすとき，p を マルコフ・カーネル (Markov kernel) と呼ぶ：

(i) 各 $x \in S$ に対して，$B \mapsto p(x, B)$ は $(S, \mathcal{S})$ 上の p.m.

(ii) 各 $B \in \mathcal{S}$ に対して，$x \mapsto p(x, B)$ は可測．

マルコフ・カーネルを単にカーネルと呼んだりもする．また，マルコフ・カーネルは遷移確率 (transition probability) とも呼ばれる．

確率過程 $\{X_n : n \in \mathbb{Z}_+\}$ がカーネル p をもつ マルコフ・チェイン[15] (Markov chain) であるとは，任意の $B \in \mathcal{S}$ と $n \in \mathbb{Z}_+$ に対して，

$$P(X_{n+1} \in B \mid X_0, X_1, \ldots, X_n) = p(X_n, B) \quad a.s. \tag{$*$}$$

をみたすことを言う．除外集合は B と n に依存してよい．$(S, \mathcal{S})$ をマルコフ・チェインの 状態空間 (state space) と呼ぶ．以下，$\mathcal{F}_n = \sigma(X_0, X_1, \ldots, X_n)$ とおく．

マルコフ・チェインの有限次元分布．　マルコフ･チェインに対して，X_0 の分布を 初期分布 (intial distribution) と呼ぶ．カーネルと初期分布に対応して，マルコフ･チェインの f.d.d.'s が決まる．条件 $(*)$ は写像

$$(x_0, x_1, \ldots, x_n, B) \mapsto p(x_n, B),\ S^{n+1} \times \mathcal{S} \to [0, 1]$$

が，$(X_0, X_1, \ldots, X_n)$ を与えたときの X_{n+1} の r.c.d. となることを意味するから，Theorem 20.3 を繰り返し適用して，次の定理を得る．

Theorem 26.1 (マルコフ・チェインの有限次元分布)**.** $\{X_n : n \in \mathbb{Z}_+\}$ を初期分布 μ, カーネル p をもつマルコフ・チェインとすると，$E \in \mathcal{S}^{n+1}, n \in \mathbb{Z}_+$ に対して，

$$\begin{aligned} &P((X_0, X_1, \cdots, X_n) \in E) \\ &= \iint \cdots \int 1_E(x_0, x_1, \ldots, x_n) p(x_{n-1}, dx_n) \cdots p(x_0, dx_1) d\mu(x_0) \end{aligned} \tag{$**$}$$

となる．

[15] 正確には，斉時的 (homogeneous) なマルコフ・チェインと言うべきである．

与えられた初期分布とカーネルをもつマルコフ・チェインの構成. いままでは，マルコフ・チェインの存在を仮定していたが，逆に初期分布 μ とカーネル p が与えられたとき，それらに対応するマルコフ・チェインの構成 (存在) を示す．まず，h を $\mathcal{S}^{m+1}$ 可測な非負関数とすると，

$$(x_0, x_1, \dots, x_{m-1}) \mapsto \int h(x_0, x_1, \dots, x_{m-1}, x_m) p(x_{m-1}, dx_m)$$

が $\mathcal{S}^m$ 可測になることを確認する [16]． Fubini の定理より，右辺の積分は定義できる．$h = 1_{A\times B}, A \in \mathcal{S}^m, B \in \mathcal{S}$ のときは可測性は明らかなので，π-λ 定理を使って，可測性が成立する範囲を $h = 1_E, E \in \mathcal{S}^{m+1}$ にまで拡張する．一般の非負可測関数 h に対しては，非負単関数による近似を使えばよい．

このとき，任意の $E \in \mathcal{S}^{n+1}$ に対して，帰納的に $(**)$ の右辺の積分が定義できることが示される．それを $\mu_n(E)$ とおくと，μ_n は $(S^{n+1}, \mathcal{S}^{n+1})$ 上の p.m. となる．$\mu_0 = \mu$ とおくと，p.m.'s $\mu_n, n \in \mathbb{Z}_+$ は整合的であるから，Kolmogorov の拡張定理より，それらを f.d.d.'s にもつ確率過程 $\{X_n : n \in \mathbb{Z}_+\}$ が存在する．

Theorem 26.2. このように構成した $\{X_n : n \in \mathbb{Z}_+\}$ は μ を初期分布，p をカーネルにもつマルコフ・チェインとなる．

Proof. $A \in \mathcal{S}^n, B \in \mathcal{S}$ に対して，$(**)$ の右辺の積分に $E = A \times B$ を代入すると，

$$\begin{aligned}
&P((X_0, \cdots, X_{n-1}) \in A, X_n \in B) \\
&= \int_A \Big\{ \underbrace{\int_B p(x_{n-1}, dx_n)}_{=p(x_{n-1},B)} \Big\} p(x_{n-2}, dx_{n-1}) \cdots p(x_0, dx_1) d\mu(x_0) \\
&= E[p(X_{n-1}, B), \{(X_0, \cdots, X_{n-1}) \in A\}].
\end{aligned}$$

これは，$P(X_n \in B \mid \mathcal{F}_{n-1}) = p(X_{n-1}, B)$ を意味するから，定理の主張を得る． □

$X = (X_0, X_1, \dots)$ の分布を $P_{\mu,p}$ とおくと，各 $\tilde{X}_n : S^{\mathbb{Z}_+} \to S$ を座標写像

$$\tilde{X}_n(\omega) = \omega_n, \ \omega = (\omega_0, \omega_1, \dots) \in S^{\mathbb{Z}_+}$$

とすると，$P_{\mu,p}$ の下で $\{\tilde{X}_n : n \in \mathbb{Z}_+\}$ も初期分布 μ とカーネル p をもつマルコフ・チェインとなる．そこで，以降ははじめから，

$$(\Omega, \mathcal{F}) = (S^{\mathbb{Z}_+}, \mathcal{S}^{\mathbb{Z}_+})$$

とし，X_n を座標写像 $X_n(\omega) = \omega_n, n \in \mathbb{Z}_+$ とする．また，p は省略して，$P_{\mu,p} = P_\mu$ と書く．このとき，$P_\mu(X_0 \in B) = \mu(B)$ となる．

[16] p がマルコフ・チェインのカーネルなら，$p(x_m, dx_{m+1})$ は $(X_0, X_1, \dots, X_m)$ を与えたときの X_{n+1} の r.c.d. となるので，可測性は Theorem 20.3 から (implicit に) 保証されている．いまは逆に p からマルコフ・チェインを構成しようとしているので，可測性を別個に確認する必要がある．

Remark 26.1. $\mu = \delta_x, x \in S$ のときは，$P_{\delta_x} = P_x$ と書く．このとき，

$$P_\mu(A) = \int P_x(A) d\mu(x),\ A \in \mathcal{S}^{\mathbb{Z}_+}$$

となる．$x \mapsto P_x(A)$ の可測性は，$A = B \times S \times S \times \cdots, B \in \mathcal{S}^n, n \in \mathbb{N}$ のときは明らか．あとは，π-λ 定理を使って，可測性が成立する範囲を $\mathcal{S}^{\mathbb{Z}_+}$ にまで拡張すればよい．

マルコフ性と強マルコフ性. $\theta_n : \Omega \to \Omega$ をシフト作用素

$$\theta_n(\omega) = (\omega_n, \omega_{n+1}, \dots)$$

とする．$E_\mu[\cdot]$ を P_μ に関する期待値とする ($E_x[\cdot] = E_{\delta_x}[\cdot]$)．

Theorem 26.3 (マルコフ性)**.** $Y : \Omega \to \mathbb{R}$ を有界可測とすると，P_μ の下で，

$$E_\mu[Y \circ \theta_n \mid \mathcal{F}_n] = E_{X_n}[Y].$$

Remark 26.2. 右辺の正確な意味は，$\varphi(x) = E_x[Y]$ とおいたとき，$\varphi(X_n)$ のことである．$x \mapsto E_x[Y]$ の可測性は，$x \mapsto P_x(A), A \in \mathcal{S}^{\mathbb{Z}_+}$ の可測性と極限操作から確認できる．

Proof. まず，$Y(\omega) = 1_A(\omega_0, \dots, \omega_k), A \in \mathcal{S}^{k+1}$ のときを考える．このとき，任意の $B \in \mathcal{S}^{n+1}$ に対して，

$$\begin{aligned}
&E_\mu[Y \circ \theta_n 1_B(\omega_0, \dots, \omega_n)] = E_\mu[1_A(\omega_n, \dots, \omega_{n+k}) 1_B(\omega_0, \dots, \omega_n)] \\
&= \int \cdots \int \left\{ \int \cdots \int 1_A(x_n, \dots, x_{n+k}) p(x_{n+k-1}, dx_{n+k}) \cdots p(x_n, dx_{n+1}) \right\} \\
&\qquad\qquad 1_B(x_0, \dots, x_n) p(x_{n-1}, dx_n) \cdots d\mu(x_0) \\
&= \int \cdots \int E_{x_n}[Y] 1_B(x_0, \dots, x_n) dp(x_n, dx_{n-1}) \cdots d\mu(x_0) \\
&= E_\mu[E_{X_n}[Y] 1_B(X_0, \dots, X_n)].
\end{aligned}$$

従って，$E_\mu[Y \circ \theta_n \mid \mathcal{F}_n] = E_{X_n}[Y]$．次に，$\pi$-$\lambda$ 定理を用いて，等号が成立する範囲を $Y = 1_A, A \in \mathcal{S}^{\mathbb{Z}_+}$ にまで拡張する．単関数による近似を用いると，Ω 上の非負可測関数に対して等号が成立することがわかる．一般の有界可測関数 Y に対しては，$Y = Y^+ - Y^-$ なる分解を使えばよい． □

Corollary 26.1 (Chapman-Kolmogorov 方程式)**.** S を可算集合とし，$\mathcal{S} = 2^S$ とする．このとき，任意の $x, y \in S$ と非負整数 m, n に対して，

$$P_x(X_{m+n} = y) = \sum_{z \in S} P_x(X_m = z) P_z(X_n = y).$$

Proof. $Y(\omega) = 1_{\{\omega_n = y\}}$ とすれば，

$$\begin{aligned} &P_x(X_{m+n} = y) = E_x[Y \circ \theta_m] = E_x[E_{X_m}[Y]] \\ &= \sum_{z \in S} P_x(X_m = z) E_z[Y] = \sum_{z \in S} P_x(X_m = z) P_z(X_n = y). \end{aligned}$$

□

Theorem 26.3 の n を ($\{\mathcal{F}_n\}$ に関する) 停止時刻 N に置き換えることを考える．定義を復習する．r.v. $N : \Omega \to \{0, 1, \dots\} \cup \{\infty\}$ が $\{\mathcal{F}_n\}$ に関する停止時刻であるとは，任意の $0 \le n < \infty$ に対して，$\{N = n\} \in \mathcal{F}_n$ となることを言う．また，停止時刻 N に対して，σ-field $\mathcal{F}_N$ を

$$\mathcal{F}_N = \{A \in \mathcal{F}_\infty : A \cap \{N = n\} \in \mathcal{F}_n, 0 \le \forall n < \infty\}$$

と定義する．ただし，$\mathcal{F}_\infty = \bigvee_n \mathcal{F}_n$ である．

Theorem 26.4 (強マルコフ性)**.** 各 n に対して，$Y_n : \Omega \to \mathbb{R}$ は可測であって，定数 M が存在して，すべての n に対して，$|Y_n| \le M$ が成り立つとする．このとき，P_μ の下で，

$$E_\mu[Y_N \circ \theta_N \mid \mathcal{F}_N] = E_x[Y_n]|_{x=X_N, n=N} \quad \text{on } \{N < \infty\}.$$

Proof. $A \in \mathcal{F}_N$ に対して，

$$E_\mu[(Y_N \circ \theta_N) 1_{A \cap \{N < \infty\}}] = \sum_{n=0}^{\infty} E_\mu[(Y_n \circ \theta_n) 1_{A \cap \{N = n\}}].$$

$A \cap \{N = n\} \in \mathcal{F}_n$ とマルコフ性より，

$$= \sum_{n=0}^{\infty} E_\mu[E_{X_n}[Y_n] 1_{A \cap \{N = n\}}] = E_\mu[(E_x[Y_n]|_{x=X_N, n=N}) 1_{A \cap \{N < \infty\}}].$$

□

Theorem 26.5 (反射原理)**.** $\xi_1, \xi_2, \dots$ を i.i.d. とし，共通分布は原点対称とする．$S_n = \xi_1 + \cdots + \xi_n$ とおくと，任意の $a > 0$ に対して，

$$P(\max_{1 \le m \le n} S_m > a) \le 2P(S_n > a).$$

Proof. 演習問題とする． □

27　マルコフ・チェインの基本概念

再帰性・推移性・既約性

本節以降，S は可算とし，$\mathcal{S} = 2^S$ とする．記号の簡単のために，$p(x, \{y\}) = p(x, y)$ と書く．また，

$$p^{(n)}(x, y) = P_x(X_n = y)$$

とおく．前節の $(**)$ と，S が可算なことから，

$$p^{(n)}(x, y) = \sum_{x_1, x_2, \ldots, x_{n-1} \in S} p(x, x_1)p(x_1, x_2) \cdots p(x_{n-1}, y)$$

となる．

$y \in S$ に対して，$T_y^0, T_y^1, \ldots$ を帰納的に，

$$T_y^0 = 0,\ T_y^k = \inf\{n > T_y^{k-1} : X_n = y\}.$$

と定義する．T_y^k は k 回目に y に戻ってくる時刻である (初期値はカウントに入れない)．

$$T_y = T_y^1,\ \rho_{xy} = P_x(T_y < \infty)$$

とおく．ρ_{xy} は x から出発して有限時間以内に y に到達する確率である．

$x \neq y$ に対して，$\rho_{xy} > 0$ のとき，y は x から 到達可能 (accessible) と言って，$x \to y$ と書く．$\rho_{xy} > 0$ かつ $\rho_{yx} > 0$ のとき，x と y は 互いに移りえる (intercommunicate) と言って，$x \leftrightarrow y$ と書く．

Lemma 27.1. $x \to y, y \to z$ なら，$x \to z$.

Proof. $x \to y, y \to z$ より，ある K, L に対して，$P_x(X_K = y) > 0, P_y(X_L = z) > 0$ である．Chapman-Kolmogorov 方程式より，$P_x(X_{K+L} = z) = \sum_w P_x(X_K = w)P_w(X_L = z) \geq P_x(X_K = y)P_y(X_L = z) > 0$ であるから，$x \to z$. □

状態 $y \in S$ が 再帰的 (reccurent) であるとは，$\rho_{yy} = 1$ となることを言う．つまり，y から出発したら，有限時間以内に必ず戻ってくるような状態を再帰的と呼ぶ．y が 推移的 (transient) であるとは，$\rho_{yy} < 1$ となることを言う．つまり，y から出発したら，有限時間以内に戻ってこない確率が正であるような状態を推移的と呼ぶ．

Theorem 27.1. (i) $P_x(T_y^k < \infty) = \rho_{xy}\rho_{yy}^{k-1}$. (ii) $N(y) = \sum_{n=1}^{\infty} 1_{\{X_n = y\}}$ を y に到達する回数とすると，

$$E_x[N(y)] = \begin{cases} \frac{\rho_{xy}}{1-\rho_{yy}} & \text{if } \rho_{xy} > 0 \text{ and } \rho_{yy} < 1 \\ \infty & \text{if } \rho_{xy} > 0 \text{ and } \rho_{yy} = 1 \\ 0 & \text{otherwise} \end{cases}$$

Proof. (i). $N = T_y^{k-1}, Y = 1_{\{T_y<\infty\}}$ として強マルコフ性を使う．$Y \circ \theta_N = 1_{\{T_y^k<\infty\}}$ と，$N < \infty$ のとき $X_N = y$ となることから，$E_{X_N}[Y] = P_y(T_y < \infty) = \rho_{yy}$. 従って，

$$\begin{aligned} P_x(T_y^k < \infty) &= E_x[(Y \circ \theta_N)1_{\{N<\infty\}}] \\ &= E_x[E_x[Y \circ \theta_N \mid \mathcal{F}_N]1_{\{N<\infty\}}] \\ &= E_x[\rho_{yy}1_{\{N<\infty\}}] \\ &= \rho_{yy}P_x(T_y^{k-1} < \infty) \\ &\vdots \\ &= \rho_{yy}^{k-1}P_x(T_y < \infty) = \rho_{yy}^{k-1}\rho_{xy}. \end{aligned}$$

(ii). $\{T_y^k < \infty\} = \{N(y) \geq k\}$ であるから，

$$\begin{aligned} E_x[N(y)] &= \sum_{k=1}^{\infty} P_x(N(y) \geq k) = \sum_{k=1}^{\infty} P_x(T_y^k < \infty) \\ &= \sum_{k=1}^{\infty} \rho_{xy}\rho_{yy}^{k-1}. \end{aligned}$$

□

Corollary 27.1. (i) y が再帰的 $\Leftrightarrow$ (ii) $P_y(X_n = y \text{ i.o.}) = 1 \Leftrightarrow$ (iii) $E_y[N(y)] = \sum_{n=1}^{\infty} p^{(n)}(y, y) = \infty$.

Proof. (i) $\Rightarrow$ (ii). Theorem 27.1 (i) より，$P_y(T_y^k < \infty) = 1$ であるから，$P_y(\bigcap_{k=1}^{\infty}\{T_y^k < \infty\}) = 1$. これは，$P_y(X_n = y \text{ i.o.}) = 1$ を意味する．(ii) $\Rightarrow$ (iii). (ii) は，$P_y(N(y) = \infty) = 1$ を意味するから，特に，$E_y[N(y)] = \infty$ である．(iii) $\Rightarrow$ (i). Theorem 27.1 (ii) より明らか. □

$C \subset S$ が閉じている (closed) とは，任意の $x \in C, y \notin C$ に対して，$\rho_{xy} = 0$ となることを言う．つまり，C の外側の状態は C 内の状態から到達可能でないとき，C は閉じていると言う．$C \subset S$ が既約 (irreducible) であるとは，任意の $x, y \in C, x \neq y$ に対して，$\rho_{xy} > 0$ となることを言う．つまり，C 内の任意の異なる 2 つの状態が互いに移りえるとき，C を既約と言う．

Remark 27.1. 既約であって閉じていない (またはその逆) ことはありうる．例えば，$p(x, y) > 0\ \forall x, y \in S$ とすると，任意の $z \in S$ に対して，$C = S \setminus \{z\}$ は既約だが，閉じていない．逆に，$p(x, y) = 0\ \forall x \in S \setminus \{y\}$ となるような y があるとき，S は閉じているが，既約でない．

Theorem 27.2 (分解定理)**.** S は推移的な状態からなる集合 T と，再帰的な状態からなる，互いに排反な，既約かつ閉じた集合 $R_1, R_2, \ldots$ に分解できる．

$$S = T \cup R_1 \cup R_2 \cup \cdots .$$

さらに，各 i に対して，$x, y \in R_i$ なら $\rho_{xy} = 1$ である．

Proof. $x \leftrightarrow y$ なら，ある K, L に対して，$P_x(X_K = y) > 0, P_y(X_L = x) > 0$ となるから，Chapman-Kolmogorov 方程式より，

$$\begin{aligned} P_x(X_{K+n+L} = x) &= \sum_{z,v} P_x(X_K = z) P_z(X_n = v) P_v(X_L = x) \\ &\geq P_x(X_K = y) P_y(X_n = y) P_y(X_L = x). \end{aligned}$$

$E_y[N(y)] = \sum_n P_y(X_n = y)$ であるから，$c = P_x(X_K = y) P_y(X_L = x) > 0$ とおくと，$E_x[N(x)] \geq c E_y[N(y)]$. 対称性より，$x \leftrightarrow y$ なら，x, y はともに推移的か，ともに再帰的かのいずれかである．

そこで，$T = \{x : \rho_{xx} < 1\}, R = \{x : \rho_{xx} = 1\}$ とおくと，T の状態と R の状態は互いに移りえない．さらに，$x \in R$ に対して，

$$C_x = \{y \in R : x \leftrightarrow y\}$$

とおいて，$C_x, x \in R$ のうち異なるものを $R_1, R_2, \dots$ とおく．

$$\{C_x : x \in R\} = \{R_1, R_2, \dots\}.$$

各 R_i は既約である．

最後に，x が再帰的で，$x \to y$ なら，$\rho_{yx} = 1$ となることを示す．各 R_i が閉じていることと，$x, y \in R_i$ に対して，$\rho_{xy} = 1$ となることは，この事実から従う．x を再帰的とし，$x \to y$ とすると，$L = \inf\{n \geq 1 : P_x(X_n = y) > 0\}$ は有限である．L の定義より，ある $y_0 = x, y_L = y, y_i \neq x, 1 \leq i \leq L-1$ が存在して，$\prod_{k=1}^{L} p(y_{k-1}, y_k) > 0$ となる．従って，仮に $\rho_{yx} < 1$ なら，

$$\begin{aligned} P_x(T_x = \infty) &\geq P_x(\{X_1 = y_1\} \cap \cdots \cap \{X_L = y_L\} \cap \{\inf\{n \geq 1 : X_{n+L} = x\} = \infty\}) \\ &= \prod_{k=1}^{L} p(y_{k-1}, y_k) P_x(\inf\{n \geq 1 : X_{n+L} = x\} = \infty \mid \mathcal{F}_L) \\ &= \prod_{k=1}^{L} p(y_{k-1}, y_k) P_y(T_x = \infty) = \prod_{k=1}^{K} p(y_{k-1}, y_k)(1 - \rho_{yx}) > 0 \end{aligned}$$

となって，x が再帰的であるという仮定に反する．従って，$\rho_{yx} = 1$. □

すべての状態が推移的なとき，マルコフ・チェイン or カーネル p は推移的であると言う．すべての状態が再帰的なとき，マルコフ・チェイン or カーネル p は再帰的であると言う．S が既約なとき，マルコフ・チェイン or カーネル p は既約であると言う．

Remark 27.2. 分解定理より，$x \leftrightarrow y$ なら，

$$x \text{ が再帰的 (推移的)} \Leftrightarrow y \text{ が再帰的 (推移的)}$$

が成り立つ．ある性質 (A としよう) について，$x \leftrightarrow y$ のとき，x, y の片方が A をみたすならもう片方も必ず A をみたすとき，性質 A を クラスの性質 (class property) と呼ぶ．再帰性と推移性はクラスの性質である．特に，マルコフ・チェインが既約なとき，推移的か再帰的かのいずれかである．後者の場合，任意の $x, y \in S$ に対して，$\rho_{xy} = 1$ である．

$X_0 \in R_i$ なら，R_i は閉じているので，マルコフ・チェインは R_i にずっととどまったままである．そこで，そのような場合は，状態空間を R_i に制限すれば，制限されたマルコフ・チェインは既約かつ再帰的になる ($p(x, R_i) = 1, x \in R_i$ であるから)．$X_0 \in T$ のときはもっと複雑で，T にとどまったままか，R_i のいずれかに移るかのどちらかである．次の補題は，T が有限なとき，前者は起こりえないことを示している．

Lemma 27.2. F を推移的な状態からなる有限集合とすると，任意の $x \in S$ に対して，$P_x(X_n \in F \text{ i.o.}) = 0$. 従って，任意の閉じた有限集合 C は少なくとも 1 つの再帰的な状態をもち，C がさらに既約なら，C のすべての状態は再帰的である．

Proof. $N(F) = \sum_{y \in F} N(y)$ とおくと，$y \in F$ は推移的であるから，Theorem 27.1 より，$E_x[N(F)] = \sum_{y \in F} E_x[N(y)] < \infty$. これは，$P_x(X_n \in F \text{ i.o.}) = 0$ を意味する．C が閉じた有限集合なら，$x \in C$ に対して，$P_x(X_n \in C\ \forall n) = 1$ であるから，C は少なくとも 1 つの再帰的な状態をもつ．また，分解定理より，既約な集合が再帰的な状態をもてば，その集合のすべての状態は再帰的になるから，最後の主張を得る． □

定常測度

引き続き S は可算で，$\mathcal{S} = 2^S$ とする．$(S, \mathcal{S})$ 上の測度 μ であって，

$$0 < \mu(S) \leq \infty, \quad \mu(x) < \infty\ \forall x \in S, \quad \sum_{x \in S} \mu(x) p(x, y) = \mu(y)\ \forall y \in S \qquad (*)$$

をみたすものを 定常測度 (stationary measure) と呼ぶ ($\mu(\{x\}) = \mu(x)$ と書く)．μ が定常測度なら，定数 $a > 0$ に対して，$a\mu$ も定常測度であるから，定常測度は定数倍の任意性がある．$(*)$ の第 3 式は $\sum_{x \in S} \mu(x) P_x(X_1 = y) = \mu(y)$ と同値である．Chapman-Kolmogorov 方程式より，$P_x(X_{n+1} = y) = \sum_{z \in S} P_x(X_1 = z) P_z(X_n = y)$ であるから，

$$\sum_{x \in S} \mu(x) P_x(X_{n+1} = y) = \sum_{z \in S} \underbrace{\sum_{x \in S} \mu(x) P_x(X_1 = z)}_{= \mu(z)} P_z(X_n = y).$$

よって，帰納的に，

$$\sum_{x \in S} \mu(x) P_x(X_{n+1} = y) = \sum_{x \in S} \mu(x) P_x(X_n = y) = \cdots = \mu(y). \qquad (**)$$

μ が p.m. のとき，それを 定常分布 (stationary distribution) と呼ぶ．このとき，

$$(**) \Leftrightarrow P_\mu(X_n = y) = \mu(y).$$

μ が p.m. でないときも，$P_\mu = \sum_{x\in S} \mu(x) P_x$ と書くことにする．

Example 27.1 (可逆測度)**.**

$$0 < \mu(S) \le \infty, \quad \mu(x) < \infty \ \forall x \in S, \quad \mu(x)p(x,y) = \mu(y)p(y,x) \ \forall x, y \in S$$

をみたす測度 μ を 可逆測度 (reversible measure) と呼ぶ．第 3 式は 詳細釣り合い条件 (detailed balanced condition) と呼ばれる．可逆測度は定常測度である．実際，

$$\sum_{x\in S} \mu(x)p(x,y) = \mu(y) \underbrace{\sum_{x\in S} p(y,x)}_{=1} = \mu(y)$$

となるからである．

定常測度は多くの場合存在する．

Theorem 27.3. $x \in S$ を再帰的な状態とすると，

$$\mu_x(y) = E_x\left[\sum_{n=0}^{T_x-1} 1_{\{X_n=y\}}\right] = \sum_{n=0}^{\infty} P_x(X_n = y, T_x > n)$$

は定常測度である．

Proof. $\mu_x(x) = P_x(T_x > 0) = 1$ であるから，$\mu_x(S) > 0$ はよい．次に，$\bar{p}_n(x,y) = P_x(X_n = y, T_x > n)$ とおくと，

$$\sum_y \mu_x(y)p(y,z) = \sum_{n=0}^{\infty}\sum_y \bar{p}_n(x,y)p(y,z).$$

z が x に等しいか等しくないかで場合分けする．

Case 1. $z \neq x$. $\{T_x > n\} = \{T_x \le n\}^c \in \mathcal{F}_n$ より，

$$\begin{aligned}\sum_y \bar{p}_n(x,y)p(y,z) &= \sum_y P_x(X_n = y, T_x > n, X_{n+1} = z)\\ &= P_x(T_x > n, X_{n+1} = z) = P_x(T_x > n+1, X_{n+1} = z) = \bar{p}_{n+1}(x,z).\end{aligned}$$

3 番目の等号は，$z \neq x$ による．従って，$\sum_{n=0}^{\infty}\sum_y \bar{p}_n(x,y)p(y,z) = \sum_{n=0}^{\infty} \bar{p}_{n+1}(x,z) = \mu_x(z)$. 最後の等号は，$\bar{p}_0(x,z) \le P_x(X_0 = z) = 0$ による．

Case 2. $z = x$.

$$\sum_y \bar{p}_n(x,y)p(y,x) = \sum_y P_x(X_n = y, T_x > n, X_{n+1} = x) = P_x(T_x = n+1)$$

であるから，$\sum_{n=0}^{\infty}\sum_y \bar{p}_n(x,y)p(y,x) = \sum_{n=0}^{\infty} P_x(T_x = n+1) = P_x(T_x < \infty) = 1 = \mu_x(x)$.

最後に，$\mu_x(y) < \infty$ $(\forall y \in S)$ を確認する．μ_x は $\sum_y \mu_x(y)p^{(n)}(y,x) = \mu_x(x) = 1$ をみたすから，$p^{(n)}(y,x) > 0$ なら $\mu_x(y) < \infty$ でなくてはならない．n は任意であるから，$\rho_{yx} > 0$ なら $\mu_x(y) < \infty$ である．x は再帰的であるから，分解定理より，$\rho_{xy} > 0$ なら $\rho_{yx} > 0$ である．一方，$\rho_{xy} = 0$ のとき，$\bar{p}_n(x,y) = 0$ であるから，$\mu_x(y) = 0$ である． □

Theorem 27.4. 既約かつ再帰的なマルコフ・チェインに対して，定常測度は定数倍の任意性を除いて一意に決まる．

Proof. ν を定常測度とすると，$a \in S$ に対して，

$$\nu(z) = \sum_y \nu(y)p(y,z) = \nu(a)p(a,z) + \sum_{y\neq a} \nu(y)p(y,z).$$

$\nu(y) = \nu(a)p(a,y) + \sum_{x\neq a} \nu(x)p(x,y)$ を代入すると，

$$\begin{aligned}\nu(z) &= \nu(a)p(a,z) + \sum_{y\neq a} \nu(a)p(a,y)p(y,z) + \sum_{x\neq a}\sum_{y\neq a} \nu(x)p(x,y)p(y,z) \\ &= \nu(a)P_a(X_1 = z) + \nu(a)P_a(X_1 \neq a, X_2 = z) + P_\nu(X_0 \neq a, X_1 \neq a, X_2 = z).\end{aligned}$$

この操作を繰り返すと，

$$\begin{aligned}\nu(z) &= \nu(a)\sum_{m=1}^{n} P_a(X_k \neq a, 1 \leq \forall k < m, X_m = z) + P_\nu(X_j \neq a, 0 \leq \forall j < n, X_n = z) \\ &\geq \nu(a)\sum_{m=1}^{n} \underbrace{P_a(X_k \neq a, 1 \leq \forall k < m, X_m = z)}_{=P_a(T_a > m-1, X_m = z)}.\end{aligned}$$

$n \to \infty$ として，

$$\nu(z) \geq \nu(a)\sum_{m=1}^{\infty} P_a(T_a > m-1, X_m = z)$$

を得る．$z \neq a$ のとき，$\{T_a > m-1, X_m = z\} = \{T_a > m, X_m = z\}$ であり，$P_a(T_a > 0, X_0 = z) = 0$ であるから，

$$\nu(z) \geq \nu(a)\mu_a(z).$$

一方，$\mu_a(a) = 1$ であるから，不等号は $z = a$ のときも成り立つ．

このとき，

$$\nu(a) = \sum_x \nu(x)p^{(n)}(x,a) \geq \nu(a)\sum_x \mu_a(x)p^{(n)}(x,a) = \nu(a)\mu_a(a) = \nu(a)$$

より，$\sum_x \nu(x)p^{(n)}(x,a) = \nu(a)\sum_x \mu_a(x)p^{(n)}(x,a)$ である．ここで，$\nu(x) \geq \nu(a)\mu_a(x)$ であるから，$p^{(n)}(x,a) > 0$ なる x に対しては，$\nu(x) = \nu(a)\mu_a(x)$ でなくてはならない．既約性より，任意の $x \in S$ に対して，$p^{(n)}(x,a) > 0$ なる n が存在するので，$\nu(x) = \nu(a)\mu_a(x)$ $(\forall x \in S)$ が示された． □

定常測度が p.m. であるとき，それを定常分布と呼んだ．既約なマルコフ・チェインに対して，定常分布は存在すれば一意である．定常分布が存在するための条件を考える．$x \in S$ が $E_x[T_x] < \infty$ をみたすとき，x を 正再帰的 (positive recurrent) と呼ぶ．再帰的だか正再帰的でない状態を 零再帰的 (null recurrent) と呼ぶ．

Theorem 27.5. マルコフ・チェインが既約なら，次の (i)–(iii) は同値である．

(i) 正再帰的な状態 x が存在する．

(ii) 定常分布が存在する．

(iii) すべての状態が正再帰的である．

(i)–(iii) のいずれかが成り立てば (すなわち，(i)–(iii) がすべて成り立てば)，定常分布 π は

$$\pi(x) = 1/E_x[T_x],\ x \in S$$

と表すことができる．

まず次の補題を示す．

Lemma 27.3. 定常分布 π が存在するとき，$\pi(y) > 0$ なら y は再帰的である．

Proof. $\sum_x \pi(x)p^{(n)}(x,y) = \pi(y)$ と Fubini の定理より，

$$\sum_x \pi(x) \underbrace{\sum_{n=1}^{\infty} p^{(n)}(x,y)}_{=E_x[N(y)]} = \sum_{n=1}^{\infty} \sum_x \pi(x)p^{(n)}(x,y) = \sum_{n=1}^{\infty} \pi(y) = \infty.$$

一方，Theorem 27.1 より，

$$\text{左辺} = \sum_x \pi(x) \frac{\rho_{xy}}{1-\rho_{yy}} \le \frac{1}{1-\rho_{yy}}$$

であるから，$\rho_{yy} = 1$ を得る． □

Proof of Theorem 27.5. (i) $\Rightarrow$ (ii). x が正再帰的なら，

$$\pi(y) = \sum_{n=0}^{\infty} P_x(X_n = y, T_x > n)/E_x[T_x]$$

は定常分布になる．

(ii) $\Rightarrow$ (iii). 少なくとも 1 つの y に対して，$\pi(y) > 0$ となるから，この y は再帰的である．このとき，既約性より，すべての状態は再帰的である．Theorem 27.3 と Theorem 27.4 よ

り，任意の $x \in S$ に対して，π は μ_x と定数倍の違いを除いて等しいから，特に $\mu_x(S) < \infty$ となるが，

$$\mu_x(S) = \sum_y \mu_x(y) = \sum_{n=0}^{\infty} \sum_y P_x(X_n = y, T_x > n) = \sum_{n=0}^{\infty} P_x(T_x > n) = E_x[T_x]$$

より，x は正再帰的であり，$\pi(x) = \mu_x(x)/\mu_x(S) = 1/E_x[T_x]$ となる．

(iii) $\Rightarrow$ (i). 明らか． □

(i) $\Rightarrow$ (iii) (と分解定理) より，正再帰性 (or 零再帰性) はクラスの性質であることがわかる．すべての状態が正再帰的なら，マルコフ・チェイン or カーネル p は正再帰的であると言う．また，すべての状態が零再帰的なら，マルコフ・チェイン or カーネル p は零再帰的であると言う．マルコフ・チェインが既約なら，

推移的 or 正再帰的 or 零再帰的

のいずれかである．

演習問題

Exercise 27.1. μ を既約かつ再帰的なマルコフ・チェインの定常測度とする．このとき，$\mu(x) > 0$ $(\forall x \in S)$ を示せ．

Exercise 27.2. 有限な状態空間をもつマルコフ・チェインは少なくとも 1 つの定常分布をもつことを示せ．

Exercise 27.3. $\{X_n\}$ を既約かつ正再帰的なマルコフ・チェインで，そのカーネル $p(x, y)$ は $p(x, x) = 0$ $(\forall x \in S)$ をみたすとする．π を $\{X_n\}$ の一意な定常分布とし，$r : S \to (0, 1)$ を与えられた関数とする．いま，$q(x, x) = 1 - r(x), q(x, y) = r(x)p(x, y)$ $(x \neq y)$ とし，$\{Y_n\}$ を $q(x, y)$ をカーネルにもつマルコフ・チェインとすると，$\{Y_n\}$ は既約かつ再帰的であって，$\mu(x) = \pi(x)/r(x)$ を定常測度にもつことを示せ．さらに，$\{Y_n\}$ が正再帰的であるための必要十分条件は $\sum_{x \in S} \pi(x)/r(x) < \infty$ であることを示せ．

28　極限定理

本節では，$n \to \infty$ のときの，$p^{(n)}(x,y)$ の収束を考察する．

$$N_n(y) = \sum_{k=1}^{n} 1_{\{X_k = y\}}$$

とおく．

Theorem 28.1. 任意の $x, y \in S$ に対して，$n \to \infty$ のとき，

$$\frac{N_n(y)}{n} \to \frac{1}{E_y[T_y]} 1_{\{T_y < \infty\}} \quad P_x - a.s. \tag{$*$}$$

従って，

$$\frac{1}{n} \sum_{k=1}^{n} p^{(k)}(x,y) \to \rho_{xy} / E_y[T_y].$$

y が推移的なら，

$$p^{(n)}(x,y) \to 0.$$

Remark 28.1. y が正再帰的なら，$N_n(y)/n \to 1/E_y[T_y] > 0$ P_y-a.s. 一方，y が零再帰的なら，$N_n(y)/n \to 0$ P_y-a.s. この事実が，正再帰的と零再帰的の名前の由来である．

Proof. y が推移的なときは，$E_x[N(y)] < \infty$ であるから，$N(y)$ は P_x-a.s. に有限である．よって，$N_n(y)/n \le N(y)/n \to 0$ P_x-a.s. 一方，$P_y(T_y = \infty) > 0$ であるから，$E_y[T_y] = \infty$ より，$(*)$ の右辺は 0 である．

次に，y が再帰的な場合を考える．P_y の下での $T_y = T_y^1$ の分布を ν_y とおく．

Lemma 28.1. y を再帰的な状態とすると，P_y の下で，

$$T_y^1, T_y^2 - T_y^1, T_y^3 - T_y^2, \cdots \sim \nu_y \text{ i.i.d.}$$

Proof of Lemma 28.1. $P = P_y$ とする．y は再帰的であるから，Theorem 27.1 より，$T_y^k < \infty$ P_y-a.s. である．$X_n^{(2)} = X_{T_y^1 + n}$ とおくと，$X_{T_y^1} = y$ なので，強マルコフ性より，$X^{(2)} = \{X_n^{(2)} : n \in \mathbb{Z}_+\}$ は $\mathcal{F}_{T_y^1}$ と独立に P_y に従う．よって，$T_y^2 - T_y^1 = \inf\{n \ge 1 : X_n^{(2)} = y\}$ は T_y^1 と独立に ν_y に従う．次に，$X_n^{(3)} = X_{T_y^2 + n}$ とおくと，$X^{(3)} = \{X_n^{(3)} : n \in \mathbb{Z}_+\}$ は $\mathcal{F}_{T_y^2}$ と独立に P_y に従うから，$T_y^3 - T_y^2 = \inf\{n \ge 1 : X_n^{(3)} = y\}$ は $(T_y^1, T_y^2 - T_y^1)$ と独立に ν_y に従う．同様の操作を繰り返して，補題の結論を得る． □

強マルコフ性より，$B \in \mathcal{S}^{\mathbb{Z}_+}, A \in \mathcal{F}_{T_y}$ に対して，

$$P_x(\{(X_{T_y + \bullet}) \in B\} \cap A \cap \{T_y < \infty\}) = P_y(B) P_x(A \cap \{T_y < \infty\})$$

となる．次の 2 つの場合を考える．

Case 1. $P_x(T_y < \infty) = 0$ のとき．$N_n(y) = 0$ P_x-a.s.

Case 2. $P_x(T_y < \infty) > 0$ のとき．$P = P_x(\,\cdot \mid T_y < \infty)$ の下で，$\{X_{T_y+n} : n \in \mathbb{Z}_+\}$ は $\mathcal{F}_{T_y}$ と独立に P_y に従うから，Lemma 28.1 より，$T_y^2 - T_y^1, T_y^3 - T_y^2, \cdots \sim \nu_y$ i.i.d. あとは，$N_n(y) = \sup\{k \geq 1 : T_y^k \leq n\}$ より，Theorem 10.6 の証明と同様にして，

$$P_x\left\{\lim_{n\to\infty} \frac{N_n(y)}{n} = \frac{1}{E_y[T_y]} \;\middle|\; T_y < \infty\right\} = 1$$

を得る．従って，P_x-a.e. $\omega \in \{T_y < \infty\}$ に対して，$N_n(y,\omega)/n \to 1/E_y[T_y]$ となる．一方，$T_y(\omega) = \infty$ のときは，$N_n(y,\omega) = 0$ であるから，$(*)$ を得る．

後半の主張は，$0 \leq N_n(y)/n \leq 1$ であるから，DCT より，

$$\frac{1}{n}\sum_{k=1}^{n} p^{(k)}(x,y) = \frac{E_x[N_n(y)]}{n} \to \rho_{xy}/E_y[T_y].$$

y が推移的なら，$\sum_n p^{(n)}(x,y) = E_x[N(y)] < \infty$ より，$p^{(n)}(x,y) \to 0$. □

Corollary 28.1. マルコフ・チェインが既約かつ正再帰的なら，一意な定常分布 π に対して，

$$\forall x, y \in S,\ \lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} p^{(k)}(x,y) = \pi(y).$$

Proof. 分解定理より，$\rho_{xy} = 1$ である．一方，$\pi(y) = 1/E_y[T_y]$ であるから，$\rho_{xy}/E_y[T_y] = \pi(y)$. □

Theorem 28.1 より，任意の $x, y \in S$ に対して，$p^{(n)}(x,y)$ は Cesaro 和の意味で収束する．この結果は，(可算な状態空間をもつ) 一般のマルコフ・チェインに対して成り立つ．また，一般に，y が推移的であれば，$p^{(n)}(x,y) \to 0$ となる．

マルコフ・チェインが既約かつ正再帰的なら，初期値 x に関わらず，$p^{(n)}(x,y)$ は Cesaro 和の意味で $\pi(y)$ に収束する．しかし，$p^{(n)}(x,y)$ は通常の意味で必ずしも収束するとは限らない．

Example 28.1. $S = \{0,1\}$ とし，カーネルは (行列表記して)

$$p = (p(i,j)) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},\ p^{(2)} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

とする．このチェインは既約で定常分布 $\pi = (1/2, 1/2)$ をもつ (従って，正再帰的)．しかし，n が奇数のとき，$p^{(n)} = p$ であり，n が偶数のとき，$p^{(n)} = p^{(2)}$ であるから，$p^{(n)}(x,y)$ は収束しない．

そこで，このような “周期性” を排除する．まず，周期性を正式に定義する．状態 x に対して，$I_x = \{n \geq 1 : p^{(n)}(x,x) > 0\}$ とおき，I_x の最大公約数 (g.c.d.) を d_x とおく．$I_x = \varnothing$ のときは，$d_x = 0$ とする．d_x を x の 周期 (period) と呼ぶ．$d_x \leq 1$ のとき，x を 非周期的 (aperiodic) と呼び，すべての $x \in S$ が非周期的なとき，マルコフ・チェイン or カーネル p を非周期的と呼ぶ．

Remark 28.2. x が再帰的なら，$\infty = E_x[N(x)] = \sum_n p^{(n)}(x,x)$ であるから，$p^{(n)}(x,x) > 0$ なる n は無限個ある．従って，$d_x = 0$ となるのは，x が推移的な場合のみである．

Lemma 28.2. (i) $x \leftrightarrow y$ なら，$d_x = d_y$. (ii) x が再帰的かつ $d_x = 1$ なら，$\exists m_0 \in \mathbb{N}$ s.t. $p^{(m)}(x,x) > 0\ \forall m \geq m_0$.

Remark 28.3. (i) は周期性がクラスの性質であることを意味している．

Proof. (i) $x \leftrightarrow y$ より，$\exists K, L$ s.t. $p^{(K)}(x,y) > 0, p^{(L)}(y,x) > 0$. このとき，

$$p^{(K+L)}(y,y) \geq p^{(L)}(y,x)p^{(K)}(x,y) > 0$$

であるから，d_y は $K+L$ を割り切る．同様に，$p^{(K+L)}(x,x) > 0$ であり，特に $I_x \neq \varnothing$. $n \in I_x$ とすると，

$$p^{(K+n+L)}(y,y) \geq p^{(L)}(y,x)p^{(n)}(x,x)p^{(K)}(x,y)$$

より，d_y は $K+n+L$ を割り切るから，n を割り切る．$n \in I_x$ は任意であるから，$d_y \leq d_x$. 対称性より，$d_x \leq d_y$ でもあるから，$d_x = d_y$ を得る．

(ii)　まず，I_x が少なくとも一組の隣り合う整数を含むことを示す．$m, n \in I_x$ 対して，$p^{(m+n)}(x,x) \geq p^{(m)}(x,x)p^{(n)}(x,x) > 0$ であるから，I_x は和について閉じている．また，x は再帰的であるから，I_x は無限集合である．$n_0, n_0 + k \in I_x$ とすると，$k = 1$ なら $n_0, n_0 + 1 \in I_x$ であるから，$k > 1$ とする．I_x の g.c.d. が $d_x = 1$ であるから，k で割り切れない $n_1 \in I_x$ が存在する．$n_1 = mk + r, 0 < r < k$ と書くと，I_x は和について閉じているから，

$$(m+1)(n_0+k) = \underbrace{(n_0+k) + \cdots + (n_0+k)}_{m} \in I_x,\ (m+1)n_0 + n_1 \in I_x.$$

このとき，$(m+1)(n_0+k) > (m+1)n_0 + n_1$ であり，その差は，

$$k(m+1) - n_1 = k - r < k.$$

そこで，$n_0' = (m+1)n_0 + n_1, k' = k - r$ とおくと，$n_0', n_0' + k' \in I_x$. $k' = 1$ なら終了し，$k' > 1$ なら同様の操作を繰り返すと，$n_0'', n_0'' + k'' \in I_x, k'' < k$ なる n_0'', k'' を見つけることができる．この操作は必ず有限回で終了するから，隣り合う整数 $N, N+1 \in I_x$ の存在が示された．

いま，$m \geq N^2$ なら，$m - N^2 = kN + r, 0 \leq r < N$ と書くと，$m = r + N^2 + kN = r(1+N) + (N - r + k)N \in I_x$ であるから，結論を得る．　□

カーネル p に対して，p をカーネルにもつ独立な 2 つのマルコフ・チェイン $\{X_n : n \in \mathbb{Z}_+\}$, $\{Y_n : n \in \mathbb{Z}_+\}$ を構成する．初期分布は異なっていてもよい．このとき，$Z_n = (X_n, Y_n)$, $n \in \mathbb{Z}_+$ は状態空間 $(S^2, \mathcal{S}^2)$ 上のマルコフ・チェインであり，そのカーネルは

$$\bar{p}((x,y),(x',y')) = p(x,x')p(y,y'),\ (x,y),(x',y') \in S^2$$

である．$\{Z_n : n \in \mathbb{Z}_+\}$ をカーネル p に対する カップリング (coupling) と呼ぶ．ここで，

$$\tau = \inf\{n \geq 0 : X_n = Y_n\}$$

とおいて，それを カップリング時間 (coupling time) と呼ぶ．さらに，S 上の p.m.'s μ, ν に対して，μ と ν の 全変動距離 (total variation distance) を

$$\|\mu - \nu\|_{TV} = \frac{1}{2}\sum_y |\mu(y) - \nu(y)|$$

と定義する (右辺の和を 2 で割っているのは理由がある).

Lemma 28.3. 各 $n \in \mathbb{N}$ に対して，

$$\|P(X_n \in \cdot) - P(Y_n \in \cdot)\|_{TV} \leq P(\tau > n) \qquad (**)$$

が成り立つ.

Proof. マルコフ性より，

$$\begin{aligned}
P(X_n = y, \tau \leq n) &= \sum_{m=0}^{n}\sum_x P(\tau = m, X_m = x, X_n = y) \\
&= \sum_{m=0}^{n}\sum_x P(\tau = m, X_m = x)P(X_n = y \mid X_m = x) \\
&= \sum_{m=0}^{n}\sum_x P(\tau = m, Y_m = x)P(Y_n = y \mid Y_m = x) \\
&= P(Y_n = y, \tau \leq n)
\end{aligned}$$

であるから，

$$\begin{aligned}
P(X_n = y) &= P(X_n = y, \tau \leq n) + P(X_n = y, \tau > n) \\
&= P(Y_n = y, \tau \leq n) + P(X_n = y, \tau > n) \\
&\leq P(Y_n = y) + P(X_n = y, \tau > n).
\end{aligned}$$

対称性より，$P(Y_n = y) \leq P(X_n = y) + P(Y_n = y, \tau > n)$ でもあるから，

$$|P(X_n = y) - P(Y_n = y)| \leq P(X_n = y, \tau > n) + P(Y_n = y, \tau > n).$$

y に関して和をとって，

$$\|P(X_n \in \cdot) - P(Y_n \in \cdot)\|_{TV} \leq P(\tau > n)$$

を得る． □

p が既約かつ正再帰的なら，一意な定常分布 π が存在する．このとき，$X_0 = x, Y_0 \sim \pi$ とすれば，$P(X_n = y) = p^{(n)}(x, y), P(Y_n = y) = \pi(y)$ であるから，

$$(**) \text{ の左辺} = \|p^{(n)}(x, \cdot) - \pi\|_{TV}.$$

従って，$\|p^{(n)}(x, \cdot) - \pi\|_{TV} \to 0$ を示すには，$P(\tau < \infty) = 1$ を保証する条件を見つけてくればよい．この部分で非周期性が重要な役割を果たす．

Theorem 28.2. p が既約かつ非周期的かつ正再帰的なら，初期分布 $\mathcal{L}(X_0, Y_0)$ に関わらず，$P(\tau < \infty) = 1$ となる．従って，p の一意な定常分布 π に対して，

$$\forall x \in S, \ \lim_{n\to\infty} \|p^{(n)}(x, \cdot) - \pi\|_{TV} = 0.$$

Proof. 前半を示せば十分である．まず，$\{Z_n\}$ が既約であることを示す．これを示すために p の非周期性を用いる．$z = (x, y), z' = (x', y') \in S^2$ を任意に固定する．まず，p の既約性より，$\exists K, L$ s.t. $p^{(K)}(x, x') > 0, p^{(L)}(y, y') > 0$. 次に，$p$ の非周期性より，十分大きな M に対して，$p^{(L+M)}(x', x') > 0, p^{(K+M)}(y, y) > 0$ となるから，Chapman-Kolmogorov 方程式より，

$$\begin{aligned}
\bar{p}^{(K+L+M)}(z, z') &= \sum_{v,v' \in S^2} \bar{p}^{(K)}(z, v)\bar{p}^{(M)}(v, v')\bar{p}^{(L)}(v', z') \\
&\geq \sum_{x'',y'' \in S} \bar{p}^{(K)}((x, y), (x', y''))\bar{p}^{(M)}((x', y''), (x'', y))\bar{p}^{(L)}((x'', y)(x', y')) \\
&= \sum_{x'',y'' \in S} p^{(K)}(x, x')p^{(K)}(y, y'')p^{(M)}(x', x'')p^{(M)}(y'', y)p^{(L)}(x'', x')p^{(L)}(y, y') \\
&= p^{(K)}(x, x')p^{(L)}(y, y') \sum_{y'' \in S} p^{(K)}(y, y'')p^{(M)}(y'', y) \sum_{x'' \in S} p^{(M)}(x', x'')p^{(L)}(x'', x') \\
&= p^{(K)}(x, x')p^{(L)}(y, y')p^{(K+M)}(y, y)p^{(L+M)}(x', x') > 0.
\end{aligned}$$

また，$\bar{\pi}(x, y) = \pi(x)\pi(y)$ とおけば，$\bar{\pi}$ は $\{Z_n\}$ の定常分布であるから，$\{Z_n\}$ は正再帰的である．$\tau_{(x,x)} = \inf\{n \geq 0 : Z_n = (x, x)\}$ とおくと，$\{Z_n\}$ は既約かつ再帰的であるから，分解定理より，初期分布に関わらず，$P(\tau_{(x,x)} < \infty) = 1$ となる．よって，$P(\tau < \infty) \geq P(\tau_{(x,x)} < \infty) = 1$ を得る． □

Remark 28.4. 既約なマルコフ・チェインは，推移的 or 正再帰的 or 零再帰的のいずれかである．(i) 推移的な場合は，任意の $x, y \in S$ に対して，$p^{(n)}(x, y) \to 0$ となる．(ii) 正再帰的かつ非周期性な場合は，$\|p^{(n)}(x, \cdot) - \pi\|_{TV} \to 0$ となる．(iii) 零再帰的かつ非周期的な場合は，$p^{(n)}(x, y) \to 0$ となる (Billinsgley, 1995, Theorem 8.7).

S が有限集合のとき. S が有限集合のときは，議論が特に簡単になる．次の補題は有限な状態空間をもつ既約なマルコフ・チェインは必ず正再帰的になることを示している．

Lemma 28.4. S が有限集合のとき，マルコフ・チェインは正再帰的な状態を必ずもち，逆に零再帰的な状態をもたない．

Proof. Theorem 28.1 より，任意の $x \in S$ に対して，

$$1 = \lim_{n\to\infty} \frac{1}{n} \sum_{y\in S} N_n(y) = \sum_{y\in S} \frac{1}{E_y[T_y]} 1_{\{T_y<\infty\}} \quad P_x - a.s.$$

従って，少なくとも 1 つの y に対して，$E_y[T_y] < \infty$ となる. また，分解定理より，チェインを各 R_i に制限すると，各 R_i は必ず正再帰的な状態をもつから，R_i のすべての状態は正再帰的である．従って，零再帰的な状態は存在しない． □

従って，S が有限のときは，マルコフ・チェインが既約かつ非周期的なら，一意な定常分布 π が存在して，$\forall x \in S, \lim_{n\to\infty} \|p^{(n)}(x,\cdot) - \pi\|_{TV} = 0$ となる．収束レートも評価できて，ある $0 < r < 1$ が存在して，$\|p^{(n)}(x,\cdot) - \pi\|_{TV} = O(r^n)$ $(n \to \infty)$ となる (Durrett 6.6.2)．なお，有限状態空間上のマルコフ・チェインの収束に関しては，Levin et al. (2008) が詳しい．

大数の法則. マルコフ・チェインが既約かつ正再帰的なとき，Theorem 28.1 の最初の結果は，

$$\frac{1}{n} \sum_{k=1}^{n} 1_{\{X_k=y\}} \to \frac{1}{E_y[T_y]} = \pi(y) \quad P_x - a.s.$$

を意味する．$f(x) = 1_{\{y\}}(x)$ とおくと，

$$\frac{1}{n} \sum_{k=1}^{n} f(X_k) \to \pi(y) = \int f d\pi \quad P_x - a.s.$$

と書き直すことが出来る．これはある種の大数の法則である．もっと一般に次の結果が成り立つ．

Theorem 28.3 (マルコフ・チェインに対する大数の法則)**.** マルコフ・チェインは既約かつ正再帰的とする．このとき，一意な定常分布を π とすると，初期分布 μ に関わらず，任意の $f \in L^1(\pi)$ に対して，

$$\frac{1}{n} \sum_{k=1}^{n} f(X_k) \to \int f d\pi \quad P_\mu - a.s.,\ n \to \infty$$

となる．

Proof. Durrett 6.6.4. 本質的に Theorem 28.1 の証明と同様. □

Remark 28.5 (MCMC との関係)**.** Theorem 28.3 は次のような応用がある．いま，与えられた p.m. π に関する積分 $\int f d\pi$ を計算したいが，単純な数値積分は難しいとする．このとき，π を定常分布にもつような既約なマルコフ・チェインを構成し，そこから標本 $X_1, \ldots, X_n$ を発生させれば，初期値に関わらず，

$$\frac{1}{n}\sum_{k=1}^{n} f(X_k) \approx \int f d\pi$$

となることが (n が十分大きければ) 期待できる．これが マルコフ連鎖モンテカルロ法 (Markov Chain Monte Carlo, MCMC) の基本的なアイデアである．MCMC に関しては，Robert and Casella (2005) を参照せよ．

Part VI. Brown 運動と連続時間マルチンゲール

本パートは連続時間の代表的な確率過程である Brown 運動と，連続時間のマルチンゲールに関するごく基本的を結果をカバーする．Brown 運動に関する最近の専門書として Mörters and Peres (2010) を挙げておく．本講義ノートでは確率積分は扱わないが，確率積分に関しては，舟木 (2005), Karatzas and Shreve (1997), Protter (2005) などが参考になる．舟木 (2005) は Brown 運動に関する確率積分しか扱っていないが，後半部分は発展的な話題をカバーしている．Karatzas and Shreve (1997) は詳細な教科書であるが，連続なパスをもつマルチンゲールに関する確率積分に限定している．Protter (2005) はジャンプをもつマルチンゲールに関する確率積分をカバーしている．ただし，Protter (2005) のアプローチは標準的なものとは異なっている．

29 Brown 運動

T を空でない集合とすると，r.v.'s の集まり $\{X_t : t \in T\}$ を T によってインデックスされた 確率過程 (stochastic process) と呼ぶ．X_t を $X(t)$ と書く場合もある．$X_t(\omega)$ は t と ω の関数であり，$\omega \in \Omega$ を固定すると，t の関数 $t \mapsto X_t(\omega)$ を サンプルパス (sample path) と呼ぶ．有限部分集合 $F \subset T$ に対して，$\mathcal{L}(X_t, t \in F)$ を 有限次元分布 (f.d.d.) と呼ぶ．すべての f.d.d. が多変量正規分布である確率過程を Gauss 過程 と呼ぶ．Gauss 過程 $\{X_t : t \in T\}$ に対しては，平均関数 $t \mapsto E[X_t]$ と共分散関数 $(s,t) \mapsto \mathrm{Cov}(X_s, X_t)$ が決まれば，その f.d.d.'s が一意に決まる．もっと一般に，与えられた可測空間 S に値をとる r.v.'s の集まりのことを，S 値確率過程と呼ぶ．本講義ノートでは，単に確率過程と言ったら，それは $\mathbb{R}$ 値確率過程のことを意味する．

以下では，$T = \mathbb{R}_+$ の場合を扱う．確率過程 $\{X_t : t \in \mathbb{R}_+\}$ が与えられたとき，すべての $\omega \in \Omega$ に対して，パス $t \mapsto X_t(\omega)$ が連続であるとき，$\{X_t\}$ は連続なパスをもつと言う．右連続または左連続なパスをもつ場合も同様に定義する．

確率過程 $\{B_t : t \in \mathbb{R}_+\}$ が標準 Brown 運動 (Brownian motion, BM) であるとは，次の 4 条件がみたされることを言う：

(i) $B_0 = 0$ a.s.

(ii) 任意の $0 < t_1 < t_2 < \cdots < t_n$ に対して，$B_{t_1}, B_{t_2} - B_{t_1}, \ldots, B_{t_n} - B_{t_{n-1}}$ は独立．

(iii) 任意の $t \geq 0$ と $h > 0$ に対して，$B_{t+h} - B_t \sim N(0, h)$.

(iv) $\{B_t\}$ は連続なパスをもつ．

また，$x \in \mathbb{R}$ に対して，$X_t = x + B_t, t \in \mathbb{R}_+$ を x を初期値にもつ Brown 運動と呼ぶ．

(ii) と (iii) は

$$(B_{t_1}, B_{t_2} - B_{t_1}, \ldots, B_{t_n} - B_{t_{n-1}}) \sim N\left(\mathbf{0}, \begin{pmatrix} t_1 & 0 & \cdots & 0 \\ 0 & t_2 - t_1 & \cdots & 0 \\ \vdots & & \ddots & \\ 0 & 0 & \cdots & t_n - t_{n-1} \end{pmatrix}\right) \qquad (*)$$

と同値である．ここで，線形変換

$$(x_1, x_2 - x_1, \ldots, x_n - x_{n-1}) \mapsto (x_1, x_2, \ldots, x_n)$$

は正則なので，

$$(*) \Leftrightarrow (B_{t_1}, B_{t_2}, \ldots, B_{t_n}) \sim N\left(\mathbf{0}, \begin{pmatrix} t_1 & t_1 & \cdots & t_1 \\ t_1 & t_2 & \cdots & t_2 \\ \vdots & & \ddots & \\ t_1 & t_2 & \cdots & t_n \end{pmatrix}\right)$$

である．従って，標準 BM の同値な定義は，次の条件をみたす確率過程 $\{B_t : t \in \mathbb{R}_+\}$ である：

(i') $\{B_t : t \in \mathbb{R}_+\}$ は Gauss 過程.

(ii') $E[B_t] = 0, E[B_s B_t] = s \wedge t$.

(iii') $\{B_t\}$ は連続なパスをもつ.

((ii') より，$E[B_0^2] = 0$ であるから，$B_0 = 0$ a.s. となる．)

Wiener 測度

.

$$W := C(\mathbb{R}_+) := \{f : f : \mathbb{R}_+ \to \mathbb{R},\ f \text{ は連続 }\}$$

とおくと，標準 BM $\{B_t : t \in \mathbb{R}_+\}$ が与えられたとき，

$$B : \omega \mapsto B(\cdot, \omega)$$

は Ω から W への写像である．従って，B を W に値をとる r.v. と考えるのが自然である．そのためには，W に適切な σ-field を入れる必要がある．W のシリンダー集合とは

$$\{f \in W : (f(t_1), \ldots, f(t_n)) \in A\},\ A \in \mathcal{B}^n, t_1, \ldots, t_n \in \mathbb{R}_+, n \in \mathbb{N}$$

なる形の集合であり，W のシリンダー集合全体から生成される σ-field $\mathcal{W}$ を W のシリンダー σ-field と呼ぶ．このとき，B は Ω から $(W, \mathcal{W})$ への可測写像となる．

W にはもう 1 つ自然な σ-field を考えることができる．いま，W に距離

$$d(f,g) = \sum_{n=1}^{\infty} \frac{1}{2^n}(1 \wedge \sup_{t\in[0,n]} |f(t) - g(t)|)$$

を入れて，d に関する Borel σ-field を $\mathcal{B}(W)$ とおく．この 2 つの σ-fields は一致する．

Lemma 29.1. $\mathcal{W} = \mathcal{B}(W)$.

まず，次の補題を示す．

Lemma 29.2. (W, d) は Poland 空間である (i.e., 完備かつ可分である).

Proof of Lemma 29.2. $d(f_n, f) \to 0 \Leftrightarrow f_n \to f$ (広義一様) である．完備性を示す．$\{f_n\}$ を d に関する Cauchy 列とすると，各 $t \in \mathbb{R}_+$ に対して，$\{f_n(t)\}_{n=1}^{\infty}$ は $\mathbb{R}$ の Cauchy 列なので，有限な $\lim_n f_n(t) =: f(t)$ が存在する．あとは $d(f_n, f) \to 0$ を示せばよいが，これは任意の $k \in \mathbb{N}$ に対して，

$$\sup_{t\in[0,k]} |f_n(t) - f(t)| \le \sup_{m\ge n} \sup_{t\in[0,k]} |f_n(t) - f_m(t)| \to 0,\ n \to \infty$$

となることから従う．

次に可分性を示す．$\mathcal{P}$ を有理係数をもつ多項式全体とすると，$\mathcal{P}$ は可算であって，任意の $f \in W, n \in \mathbb{N}, \varepsilon > 0$ に対して，$\sup_{t\in[0,n]} |f(t) - p(t)| < \varepsilon$ となる $p = p_{f,n,\varepsilon} \in \mathcal{P}$ が存在する (cf. Weierstrass の近似定理)．よって，W は可分である． □

Proof of Lemma 29.1. 任意の $t_1, \dots, t_n \in \mathbb{R}_+$ に対して，射影

$$\pi_{t_1,\dots,t_n} : f \mapsto (f(t_1), \dots, f(t_n)),\ W \to \mathbb{R}^n$$

は連続なので，$\mathcal{B}(W)$ 可測である．よって，$A \in \mathcal{B}^n$ に対して，

$$\{f \in W : (f(t_1), \dots, f(t_n)) \in A\} = \pi^{-1}_{t_1,\dots,t_n}(A) \in \mathcal{B}(W)$$

より，$\mathcal{W} \subset \mathcal{B}(W)$ である．次に，$f_0 \in W$ を任意に固定し，

$$F(f) = d(f, f_0) = \sum_{n=1}^{\infty} \frac{1}{2^n}(1 \wedge \sup_{t\in[0,n]} |f(t) - f_0(t)|),\ f \in W$$

とおくと，f, f_0 の連続性より，

$$\begin{aligned} F(f) &= \sum_{n=1}^{\infty} \frac{1}{2^n}(1 \wedge \sup_{t\in[0,n]\cap\mathbb{Q}} |f(t) - f_0(t)|) \\ &= \sum_{n=1}^{\infty} \frac{1}{2^n}(1 \wedge \sup_{t\in[0,n]\cap\mathbb{Q}} |\pi_t(f) - \pi_t(f_0)|). \end{aligned}$$

各 $\pi_t : W \to \mathbb{R}$ は $\mathcal{W}$ 可測であるから，$F : W \to \mathbb{R}$ も $\mathcal{W}$ 可測である．従って，

$$\{f \in W : d(f, f_0) < r\} = \{f \in W : F(f) < r\} \in \mathcal{W}.$$

よって，W の任意の開球は $\mathcal{W}$ 可測である．W の可分性より，W の任意の開集合は開球の可算和で表せるので，$\mathcal{W}$ 可測である．従って，$\mathcal{B}(W) \subset \mathcal{W}$ を得る． □

$\mathcal{B}(W)$ 上に定義される B の像測度 $P \circ B^{-1}$ を Wiener 測度 と呼ぶ．$\mathcal{W} = \mathcal{B}(W)$ と π-λ 定理より，Wiener 測度は f.d.d.'s から一意に決まる．

標準 BM の存在

いままでの議論では標準 BM の存在を仮定していたが，その存在を証明する．

Theorem 29.1 (Wiener)**.** 適当な確率空間上に標準 BM は存在する．

標準 BM の存在は自明ではない．注意すべきは，Kolmogorov の拡張定理 (cf. Exercise 20.4) は整合的な p.m.'s を f.d.d.'s にもつ確率過程の存在を保証し，標準 BM の f.d.d.'s の整合性を示すのは難しくないが，Kolmogorov の拡張定理はパスの連続性までは保証しないことである．

Proof. まず，$[0,1]$ 上の Brown 運動を構成する．証明をいくつかのステップに分ける．$\mathbb{Q}_2 = \{m2^{-n} : m \in \mathbb{Z}, n \in \mathbb{Z}_+\}$ を 2 進有理数の全体とする．

Step 1. 離散時間の Gauss 過程 $\{B_t : t \in [0,1] \cap \mathbb{Q}_2\}$ で，$E[B_t] = 0$, $E[B_s B_t] = s \wedge t$ をみたすものを構成する．$[0,1] \cap \mathbb{Q}_2 = \{q_1, q_2, \dots\}$ とする．各 $n \in \mathbb{N}$ に対して，

$$\mu_n = N(\mathbf{0}, (q_j \wedge q_k)_{1 \le j,k \le n})$$

とおくと，$\mu_n, n \in \mathbb{N}$ は整合的な p.m.'s であるから，Kolmogorov の拡張定理より，それらを f.d.d.'s にもつ確率過程 $\{X_m : m \in \mathbb{N}\}$ を適当な確率空間上に構成できる．あとは $B_{q_m} = X_m, m \in \mathbb{N}$ とおけばよい．

Step 2. Step 1 で構成した $\{B_t : t \in [0,1] \cap \mathbb{Q}_2\}$ に対して，a.s. にパス $t \mapsto B_t, t \in [0,1] \cap \mathbb{Q}_2$ が一様連続であることを示す．次の補題を示す．

Lemma 29.3 (Kolmogorov-Chentsov)**.** $m \in \mathbb{N}$ を given とし，$D_m = \{i/2^m : i = 0, 1, \dots, 2^m\}$ とおく．$\{X_t : t \in D_m\}$ を r.v.'s の集まりとし，ある $c_1 > 0, \varepsilon \in (0,1], 1 \le p < \infty$ が存在して，すべての $s, t \in D_m$ に対して，

$$(E[|X_t - X_s|^p])^{1/p} \le c_1 |s - t|^{(1/p)+\varepsilon} \qquad (**)$$

をみたすとする．このとき，任意の $\alpha \in (0, \varepsilon)$ に対して，

$$\left\{ E\left[\left(\max_{s \ne t, s,t \in D_m} \frac{|X_t - X_s|}{|t - s|^\alpha} \right)^p \right] \right\}^{1/p} \le \frac{2c_1}{(1 - 2^{-\varepsilon})(1 - 2^{-(\varepsilon - \alpha)})} =: c_2(c_1, \varepsilon, \alpha)$$

が成り立つ．

c_2 は m に依存しないので，$m \to \infty$ として次の系を得る．

Corollary 29.1. $\{X_t : t \in [0,1] \cap \mathbb{Q}_2\}$ を r.v.'s の集まりとし，ある $c_1 > 0, \varepsilon \in (0,1], 1 \le p < \infty$ が存在して，すべての $s,t \in [0,1] \cap \mathbb{Q}_2$ に対して $(**)$ をみたすとする．このとき，任意の $\alpha \in (0,\varepsilon)$ に対して，

$$\left\{ E\left[\left(\sup_{s \neq t, s,t, \in [0,1] \cap \mathbb{Q}_2} \frac{|X_t - X_s|}{|t-s|^\alpha} \right)^p \right] \right\}^{1/p} \le c_2(c_1, \varepsilon, \alpha)$$

が成り立つ．特に，

$$\sup_{s \neq t, s,t, \in [0,1] \cap \mathbb{Q}_2} \frac{|X_t - X_s|}{|t-s|^\alpha} < \infty \quad a.s.$$

である．

Proof. $M_k = \max_{1 \le i \le 2^k} |X(i/2^k) - X((i-1)/2^k)|$ とおくと，

$$(E[M_k^p])^{1/p} \le \left(\sum_{i=1}^{2^k} E[|X(i/2^k) - X((i-1)/2^k)|^p] \right)^{1/p} \le (c_1^p 2^k 2^{-(1+p\varepsilon)k})^{1/p} = c_1 2^{-\varepsilon k}.$$

$2^{-k} \le t - s < 2^{-k+1}$ のとき，$s,t \in D_m$ は

$$s = \frac{(i-1)}{2^k} - \frac{*}{2^{k+1}} - \cdots - \frac{*}{2^m},$$
$$t = \frac{i}{2^k} + \frac{*}{2^{k+1}} + \cdots + \frac{*}{2^m} \quad (* \text{ は } 0 \text{ or } 1)$$

の形に表せるから，

$$\begin{aligned} &|X_t - X_s| \\ &\le |X(t) - X(i/2^k)| + |X(i/2^k) - X((i-1)/2^k)| + |X((i-1)/2^k) - X(s)| \\ &\le \sum_{\ell=k+1}^{m} M_\ell + M_k + \sum_{\ell=k+1}^{m} M_\ell \le 2 \sum_{\ell=k+1}^{m} M_\ell + 2^k (t-s) M_k. \end{aligned}$$

ゆえに，$2^{-k} \le t - s < 2^{-k+1}, s,t \in D_m$ ならば，

$$\frac{|X_t - X_s|}{(t-s)^\alpha} \le 2 \cdot 2^{\alpha k} \sum_{\ell=k+1}^{m} M_\ell + 2^k \cdot 2^{(-k+1)(1-\alpha)} M_k \le 2 \cdot 2^{\alpha k} \sum_{\ell=k}^{m} M_\ell.$$

ここで，$(E[(\sum_{\ell=k}^m M_\ell)^p])^{1/p} \le \sum_{\ell=k}^m (E[M_\ell^p])^{1/p} \le c_1 \sum_{\ell=k}^\infty 2^{-\varepsilon\ell} = c_1 2^{-\varepsilon k}/(1-2^{-\varepsilon})$. $s,t \in D_m$ に対して，$s \neq t$ なる限り，$|s-t| \ge 2^{-m}$ であるから，

$$\left\{ E\left[\left(\max_{s \neq t, s,t \in D_m} \frac{|X_t - X_s|}{|t-s|^\alpha} \right)^p \right] \right\}^{1/p} \le \frac{2c_1}{1-2^{-\varepsilon}} \sum_{k=0}^{m} 2^{-(\varepsilon-\alpha)k} \le \frac{2c_1}{(1-2^{-\varepsilon})(1-2^{-(\varepsilon-\alpha)})}.$$

よって，補題が示された． □

いま，$B_s - B_t \sim N(0, |s-t|)$ であるから，$E[(B_s - B_t)^4] = 3|s-t|^2$ である．そこで，Corollary 29.3 を $p = 4, \varepsilon = 1/4, \alpha = 1/8$ として適用すれば，$P(\Omega_0) = 1$ なる事象 Ω_0 が存在して，各 $\omega \in \Omega_0$ に対して，パス $t \mapsto B_t(\omega), t \in [0,1] \cap \mathbb{Q}_2$ が一様連続であることがわかる．

Step 3. $\{B_t : t \in [0,1] \cap \mathbb{Q}_2\}$ を $[0,1]$ に連続拡張する．次の補題を示す．

Lemma 29.4. D を $[0,1]$ の稠密な部分集合とし，$f : D \to \mathbb{R}$ は D 上一様連続とする．このとき，D 上で f に一致する一意な連続関数 $\tilde{f} : [0,1] \to \mathbb{R}$ が存在する : $\tilde{f}(x) = f(x), x \in D$.

Proof of Lemma 29.4. 一意性は明らか．存在を示す．D は $[0,1]$ で稠密であるから，任意の $t \in [0,1]$ に対して，$t_n \in D, t_n \to t$ となるものが存在する．このとき，$\{f(t_n)\}_{n=1}^{\infty}$ は $\mathbb{R}$ の Cauchy 列なので，有限な $\lim_n f(t_n)$ が存在する．この極限が t_n の選び方に依存しないことを示す．$s_n \in D, s_n \to t$ をもう 1 つ選ぶ．このとき，$s_n - t_n \to 0$ と f の一様連続性より，$f(s_n) - f(t_n) \to 0$ となるから，$\lim_n f(t_n) = \lim_n f(s_n)$ を得る．そこで，$\tilde{f}(t) := \lim_n f(t_n)$ と定義すれば，$\tilde{f}$ が求める連続関数である． □

そこで，各 $t \in [0,1]$ に対して，

$$\tilde{B}_t(\omega) = \begin{cases} \lim_{s \to t, s \in [0,1] \cap \mathbb{Q}_2} B_s(\omega) & \text{if } \omega \in \Omega_0 \\ 0 & \text{if } \omega \notin \Omega_0 \end{cases}$$

と定義する．各 $\omega \in \Omega$ に対して，$t \mapsto \tilde{B}_t(\omega), t \in [0,1]$ は連続である．あとは，$\{\tilde{B}_t : t \in [0,1]\}$ の各 f.d.d. が標準 BM のそれと一致することを確認すればよい．$t_1, \ldots, t_m \in [0,1]$ に対して，$t_k(n) \in [0,1] \cap \mathbb{Q}_2$ を $t_k(n) \to t_k$ $(n \to \infty)$ となるように選ぶと，$B(t_k(n)) \to B(t_k)$ a.s. $(n \to \infty)$ であるから，

$$(B(t_1(n)), \ldots, B(t_m(n))) \xrightarrow{d} (\tilde{B}(t_1), \ldots, \tilde{B}(t_m)),\ n \to \infty$$

である．左辺は平均 $\mathbf{0}$ の多変量正規分布に従い，$E[B(t_k(n))B(t_\ell(n))] = t_k(n) \wedge t_\ell(n) \to t_k \wedge t_\ell$ となる．Cramér-Wold device と合わせると，次の補題を示せば十分である．

Lemma 29.5. $X_n \sim N(0, \sigma_n^2)$ とし，$n \to \infty$ のとき，$\sigma_n^2 \to \sigma^2$ とする．このとき，$X_n \xrightarrow{d} N(0, \sigma^2)$.

Proof of Lemma 29.5. $E[e^{itX_n}] = \varphi_n(t) = e^{-t^2\sigma_n^2/2} \to e^{-t^2\sigma^2/2}$. □

以上より，$\{\tilde{B}_t : t \in [0,1]\}$ が $[0,1]$ 上の標準 BM であることが示された．あとは，

$$\hat{B}_t = (1+t)\tilde{B}_{t/(1+t)} - t\tilde{B}_1,\ t \in \mathbb{R}_+$$

とおくと，$\{\hat{B}_t : t \in \mathbb{R}_+\}$ は標準 BM である． □

証明を検証していくと，任意の $\alpha \in (0,1/2)$ に対して，BM のサンプルパスは a.s. に局所 α-Hölder 連続であることがわかる (cf. Exercise 29.9)．しかし，標準 BM のサンプルパスは a.s. にいたるところ微分不可能である．これは次の定理から従う．

Theorem 29.2 (Paley-Wiener-Zygmund)**.** $(\Omega, \mathcal{F}, P)$ を完備とし，$\{B_t : t \in \mathbb{R}_+\}$ を標準 BM とする．このとき，任意の $\alpha > 1/2$ に対して，

$$P\left(\left\{\omega : \exists s \in \mathbb{R}_+ \text{ s.t. } \limsup_{t \downarrow s} \frac{|B_t(\omega) - B_s(\omega)|}{(t-s)^\alpha} < \infty\right\}\right) = 0.$$

Proof. $f : \mathbb{R}_+ \to \mathbb{R}$ に対して，

$$\delta_s(f) = \limsup_{t \downarrow s} |f(t) - f(s)|/(t-s)^\alpha$$

と定義する．このとき，

$$\{\exists s \in \mathbb{R}_+ \text{ s.t. } \delta_s(B) < \infty\} = \bigcup_{n=0}^{\infty} \{\exists s \in [0,1) \text{ s.t. } \delta_s(B_{n+\bullet} - B_n) < \infty\}$$

であって，各 n に対して，$\{B_{n+t} - B_n : t \in \mathbb{R}_+\}$ は標準 BM なので，

$$P(\{\exists s \in [0,1) \text{ s.t. } \delta_s(B) < \infty\}) = 0$$

を示せばよい．ところで，あらゆる $L \in \mathbb{N}$ に対して，

$$\begin{aligned}
&\{\exists s \in [0,1) \text{ s.t. } \delta_s(B) < \infty\} \\
&\quad \subset \bigcup_{M=1}^{\infty} \bigcup_{N=1}^{\infty} \bigcap_{n=N}^{\infty} \bigcup_{m=0}^{n} \bigcap_{\ell=0}^{L-1} \left\{|B((m+\ell+1)/n) - B((m+\ell)/n)| \le M/n^\alpha\right\}
\end{aligned}$$

であるから，適当な L に対して，

$$\lim_n nP\left(|B((m+\ell+1)/n) - B((m+\ell)/n)| \le M/n^\alpha, 0 \le \forall \ell \le L-1\right) = 0$$

となることを示せばよい．しかし，

$$\begin{aligned}
&P\left(|B((m+\ell+1)/n) - B((m+\ell)/n)| \le M/n^\alpha, 0 \le \forall \ell \le L-1\right) \\
&\quad = \{P(|B_{1/n}| \le M/n^\alpha)\}^L = \{P(|B_1| \le Mn^{1/2-\alpha})\}^L = O(n^{(1/2-\alpha)L})
\end{aligned}$$

であるから，$(1/2-\alpha)L < -1$ となるように L を選べばよい． □

多次元の Brown 運動. $\{B_t^{(\ell)} : t \in \mathbb{R}_+\}, \ell = 1, \ldots, d$ を独立な標準 BM とすると，$\mathbb{R}^d$ に値をとる確率過程

$$\boldsymbol{B}_t = (B_t^{(1)}, \ldots, B_t^{(d)}),\ t \in \mathbb{R}_+$$

を d 次元標準 Brown 運動と呼ぶ．ただし，以下では $d = 1$ の場合のみ扱う．

可分修正

やや寄り道となるが，確率過程の可分修正を考察する．一般に 2 つの確率過程 $\{X_t : t \in T\}, \{Y_t : t \in T\}$ が互いの 修正 (modification) であるとは，$P(X_t = Y_t) = 1$ $(\forall t \in T)$ となることを言う．以下，$(\Omega, \mathcal{F}, P)$ を完備な確率空間，$\mathcal{N} = \{N \in \mathcal{F} : P(N) = 0\}$, T を可分な距離空間とし，$\{X_t : t \in T\}$ を確率過程とする．このとき，$\{X_t\}$ が 可分 (separable) であるとは，次をみたす零集合 $N \in \mathcal{N}$ と可算集合 $D \subset T$ が存在することを言う：任意の $\omega \in \Omega \setminus N$ と $t \in T$ に対して，$t_n \to t$ なる $t_n \in D$ が存在して，$X(t_n, \omega) \to X(t, \omega)$ となる．もっと一般に距離空間に値をとる確率過程に対しても，可分性を同様に定義する．

例えば，$\{X_t\}$ が連続なパスをもてばそれは可分である．また，$T = \mathbb{R}_+$ のとき，$\{X_t\}$ が右連続なパスをもてばそれは可分である．T が非可算であるとき，例えば $\sup_{t \in T} X_t$ は非可算集合に関して sup をとっているので，その可測性は明らかでない．しかしながら，$\{X_t\}$ が可分なら，$\sup_{t \in T} X_t$ は可測である．もっと一般に，$\{X_t\}$ が可分なら，任意の開集合 $G \subset T$ に対して，

$$\sup_{t \in G} X_t = \sup_{t \in G \cap D} X_t \quad a.s.$$

であるから，左辺の可測性が保証される．$\sup_t X_t$ を $\sup_t |X_t|$ としても同様である．可分性は確率過程の専門書においてしばしば現れる概念である．ここでの目標は，$\{X_t\}$ は必ず可分な修正をもつということを証明することである (ただし，その修正は $\overline{\mathbb{R}}$ 値の確率過程になる)．従って，パスの性質ではなく，$\sup_{t \in G} X_t$ のような量の評価 (期待値や裾確率) に興味がある場合，インデックス集合 T を可算と仮定してしまってもそれほど一般性を失うわけではない，と言える [17].

Theorem 29.3. T を可分な距離空間とし，$\{X_t : t \in T\}$ を確率過程とする．このとき，次をみたす $\overline{\mathbb{R}}$ 値の確率過程 $\{\tilde{X}_t : t \in T\}$ が存在する：$\{\tilde{X}_t\}$ は可分であって，$\{X_t\}$ の修正になっている：$P(\tilde{X}_t = X_t) = 1$ $\forall t \in T$.

$\{\tilde{X}_t\}$ は $\overline{\mathbb{R}}$ 値であるが，固定した $t \in T$ に対して，$\tilde{X}_t$ が $\pm\infty$ をとる確率は 0 である．

Proof. 証明を 3 つのステップに分割する．

Step 1. 任意の $S \subset T, S \neq \varnothing$ と $B \in \mathcal{B}$ に対して，次をみたす $t_n = t_n^{(S,B)} \in S, n = 1, 2, \ldots$ が存在することを示す：任意の $t \in S$ に対して，

$$N(t, B) := \{X_t \in B\} \bigcap \left(\bigcap_{n=1}^{\infty} \{X_{t_n} \notin B\} \right)$$

の確率は 0 である．

[17] とは言っても，X_t に f.d.d.'s 以外のなんらかの構造があって，もともとの X_t に対して $\sup_t X_t$ の評価をしたい場合は，可分修正をとることは無意味である．

$S \subset T, S \neq \varnothing$ と $B \in \mathcal{B}$ を given とし，$\{t_n\}$ を次のように帰納的に決める．$t_1 \in T$ を任意にとり，t_n まで決まったら，

$$M_n := \bigcap_{k=1}^{n} \{X_{t_k} \notin B\},\ L_n(t) := M_n \cap \{X_t \in B\},\ p_n = \sup_{t \in S} P\{L_n(t)\}$$

と定義する．$p_n = 0$ なら，$p_{n+1} = p_n$ とすればよい．$p_n > 0$ のときは，$P\{L_n(t)\} \geq p_n/2$ をみたす $t \in S$ を 1 つ選んでそれを t_{n+1} とする．

このとき，$L_n(t_{n+1}), n = 1, 2, \dots$ は排反であって，

$$1 \geq P\left\{\bigcup_n L_n(t_{n+1})\right\} = \sum_n P\{L_n(t_{n+1})\} \geq \frac{1}{2}\sum_n p_n.$$

よって，$p_n \to 0$ である．いま，各 $t \in S$ に対して，$L_n(t) \downarrow N(t, B)$ であるから，$P\{N(t,B)\} = \lim_n P\{L_n(t)\} = 0$ を得る．

Step 2. $\mathcal{I}$ を有理点を端点にもつ開区間の全体とする．また，T_0 を T の可算稠密集合とし，$\mathcal{U}$ を T_0 の点を中心にもち，有理数の半径をもつ T の開球の全体とする．さらに，

$$D = \{t_n^{(U,I)} : U \in \mathcal{U}, I \in \mathcal{I}, n = 1, 2, \dots\},$$
$$N(t) = \bigcup_{U \in \mathcal{U}} \bigcup_{I \in \mathcal{I}} \left(\{X_t \in I\} \bigcap \left(\bigcap_{n=1}^{\infty} X_{t_n^{(U,I)}} \notin I\}\right)\right), \quad t \in T$$

とおく．Step 1 より，$P\{N(t)\} = 0$ である．このとき，次を示す：各 $t \in T, \omega \in \Omega \setminus N(t)$ に対して，$s_n \to t, s_n \in D$ が存在して，$X(s_n, \omega) \to X(t, \omega)$ となる．

$t \in T, \omega \in \Omega \setminus N(t)$ を given とし，$n = 1, 2, \dots$ に対して，直径が n^{-1} 以下の $\mathcal{U}$ の開球 $U \ni t$ と，直径が n^{-1} 以下の $\mathcal{I}$ の区間 $I \ni X(t, \omega)$ を選ぶ．ここで，$N(t)$ の定義より，ある $s_n \in \{t_k^{(U,I)} : k = 1, 2, \dots\}$ が存在して，$X(s_n, \omega) \in I$ となる．この構成から，$s_n \to t, X(s_n, \omega) \to X(t, \omega)$ となる．

Step 3. 定理の結論を示す．Step 2 の s_n を $s_n^{(t)}(\omega)$ とする．このとき，

$$\tilde{X}(t, \omega) := \begin{cases} X(t, \omega) & \text{if } t \in D \text{ or } \omega \in \Omega \setminus N(t) \\ \limsup_n X(s_n^{(t)}(\omega), \omega) & \text{if } t \notin D \text{ and } \omega \in N(t) \end{cases}$$

と定義する．各 $t \in T$ に対して，$\tilde{X}_t(\omega) = X_t(\omega)\ \forall \omega \in \Omega \setminus N(t)$ であるから，各 $\tilde{X}_t$ は $\overline{\mathbb{R}}$-valued r.v. であって，$\{\tilde{X}_t\}$ は $\{X_t\}$ の修正になる．$\{\tilde{X}_t\}$ が可分であることを示そう．$N := \bigcup_{t \in D} N(t)$ とおくと，$P(N) = 0$ である．$t \in T, \omega \in \Omega \setminus N$ を given とする．問題となるのは，$t \notin D\ \&\ \omega \in N(t)$ のときのみである．しかし，このとき，$\tilde{X}_t$ の定義から，$\{s_n^{(t)}(\omega)\}$ の部分列 $\{s_{n_k}^{(t)}(\omega)\} \subset D$ が存在して，$s_{n_k}^{(t)}(\omega) \to t\ \&\ X(s_{n_k}^{(t)}(\omega), \omega) \to X(t, \omega)$ となる． □

演習問題

Exercise 29.1. 標準 BM の修正であるが，パスが不連続な確率過程を構成せよ．

Exercise 29.2. 2 つの確率過程 $\{X_t : t \in T\}, \{Y_t : t \in T\}$ が 区別できない (indistinguishable) とは，ある $N \in \mathcal{F}, P(N) = 0$ が存在して，あらゆる $\omega \in N^c$ に対して，$X_t(\omega) = Y_t(\omega)$ $(\forall t \in T)$ となることを言う．$T = \mathbb{R}_+$ のとき，右連続なパスをもつ 2 つの確率過程が互いの修正であれば，それらは区別できないことを示せ．

Exercise 29.3. $\{X_t : t \in \mathbb{R}_+\}$ を右連続なパスをもつ確率過程とすると，

$$(t, \omega) \mapsto X_t(\omega),\ \mathbb{R}_+ \times \Omega \to \mathbb{R}$$

は $\mathcal{B}_{\mathbb{R}_+} \times \mathcal{F}/\mathcal{B}$ 可測であることを示せ．

Exercise 29.4. 次の 3 条件をすべてみたす確率過程 $\{X_t : t \in [0,1]\}$ は 存在しない ことを示せ．(1) 相異なる $s, t \in [0,1]$ に対して，X_s と X_t は独立．(2) 各 X_t は同一分布に従い，分散 1 をもつ．(3) パス $t \mapsto X_t$ は連続．

Exercise 29.5. $t \in \mathbb{R}_+$ に対して，

$$X_t(\omega) = \int_0^t B_s(\omega) ds$$

とおく．$\{X_t\}$ が Gauss 過程であることを示し，その平均関数と共分散関数を計算せよ．

Exercise 29.6. $B_t^\circ = B_t - tB_1, t \in [0,1]$ を Brown 橋 (Brownian bridge) と呼ぶ．

(a) $\{B_t^\circ\}$ が Gauss 過程であることを示し，その共分散関数を計算せよ．

(b) $0 < t_1 < \cdots < t_k < 1$ を任意に固定し，$\varepsilon > 0$ に対して，$\mu_\varepsilon(A) := P((B_{t_1}, \ldots, B_{t_k}) \in A \mid |B_1| \le \varepsilon), A \in \mathcal{B}^k$ と定義する．このとき，任意の $\varepsilon_n \downarrow 0$ に対して，

$$\mu_{\varepsilon_n} \xrightarrow{w} \mathcal{L}(B_{t_1}^\circ, \ldots, B_{t_k}^\circ)$$

を示せ．

Exercise 29.7. Lemma 29.4 において，f が D 上で一様連続でなくて，単に連続な場合は結論は成り立つか．

Exercise 29.8 (時間反転)**.** $X_t = tB_{1/t}$ $(t \neq 0), X_0 = 0$ とおくと，$\{X_t\}$ は標準 BM であることを示せ．これから，$\lim_{t\to\infty} B_t/t = 0$ a.s. を導け．

Exercise 29.9. 任意の $0 < \alpha < 1/2$ に対して，

$$P\left\{\sup_{s,t\in[0,1], s\neq t} \frac{|B_t - B_s|}{|t-s|^\alpha} < \infty\right\} = 1$$

を示せ．

Exercise 29.10. $t > 0$ を固定し，$\Delta_{m,n} = |B(tm2^{-n}) - B(t(m-1)2^{-n})|$ とおく．このとき，$\sum_{m=1}^{2^n} \Delta_{m,n}^2 \to t$ a.s. $(n \to \infty)$ を示せ．

Exercise 29.11. $-\infty < a < b < \infty$ とし，関数 $F : [a,b] \to \mathbb{R}$ に対して，F の 全変動関数 (total variation function) を

$$V_F(x) = \sup \left\{ \sum_{i=1}^{n} |F(x_i) - F(x_{i-1})| : a = x_0 < x_1 < \cdots < x_n = x, n \in \mathbb{N} \right\},\ x \in [a,b]$$

と定義する．$V_F(x)$ は非減少であって，$V_F(b) < \infty$ のとき F は 有界変動 (bounded variation) であると言う．

(a) F が 2 つの非減少関数の差で表せるなら，F は有界変動であることを示せ．

(b) $G(x) = V_F(x) + F(x), H(x) = V_F(x) - F(x)$ とおくと，G, H はそれぞれ非減少であることを示せ．

(c) F が右連続なら，$V_F(x)$ も右連続であることを示せ．

Remark 29.1. F が右連続な有界変動関数であって，$F(a) = 0$ なら，$F_1(x) = (V_F(x) + F(x))/2, F_2(x) = (V_F(x) - F(x))/2$ とおくと，F_1, F_2 は右連続かつ非減少で，$F_1(a) = F_2(a) = 0, F = F_1 - F_2$ をみたす．形式的に，$F_i(x) = F_i(b)$ (for $x > b, i = 1,2$), $F_1(x) = F_2(x) = 0$ (for $x < a$) として，F_1, F_2 を $\mathbb{R}$ 上の関数に拡張すれば，F_1, F_2 は $\mathbb{R}$ 上で右連続かつ非減少であって，それぞれに対応して Lebesgue-Stieltjes 積分を構成できる．従って，$[a,b]$ 上の有界可測関数 f に対して，

$$\int_{[a,b]} f dF := \int_{[a,b]} f dF_1 - \int_{[a,b]} f dF_2$$

として F に関する積分を定義できる．

Exercise 29.12. 標準 BM のパスの $[0,1]$ 上での全変動は a.s. に $+\infty$ であることを示せ．従って，B_t に関する積分をパスごとに Lebesgue-Stieltjes 積分として構成することはできないことがわかる．この認識が確率積分の出発点である．

ヒント：Exercise 29.10.

30　連続時間のマルチンゲール

本節は Brown 運動に関する考察を一旦中断し，連続時間のマルチンゲールに関するごく基本的な結果をカバーする．$\mathcal{F}$ の sub σ-fields の増加列 $\{\mathcal{F}_t : t \in \mathbb{R}_+\}$:

$$\mathcal{F}_s \subset \mathcal{F}_t,\ s < t,$$

を フィルトレーション (filtration) と呼ぶ．技術的な問題を避けるために，$(\Omega, \mathcal{F}, P)$ は完備とし，各 $\mathcal{F}_t$ は P 零集合の全体を含むとする．すなわち，

$$\mathcal{N} := \{N \in \mathcal{F} : P(N) = 0\} \subset \mathcal{F}_t \tag{$*$}$$

と仮定する．さらに，$\{\mathcal{F}_t\}$ は次の意味で 右連続 と仮定する．

$$\mathcal{F}_t = \mathcal{F}_{t+} := \bigcap_{s>t} \mathcal{F}_s,\ t \in \mathbb{R}_+. \tag{$**$}$$

$\{\mathcal{F}_t\}$ が条件 $(*)$ と $(**)$ をみたすとき，通常の仮定をみたす，と言う．

一般にフィルトレーションは通常の仮定をみたすとは限らない．例えば，$\mathcal{F}_0 = \{\varnothing, \Omega\}, \mathcal{F}_t = \mathcal{F} \neq \{\varnothing, \Omega\},\ t > 0$ とすると，$\{\mathcal{F}_t\}$ はフィルトレーションであるが，$\mathcal{F}_{0+} = \mathcal{F} \neq \mathcal{F}_0$ である．しかし，通常の仮定をみたすとは限らないフィルトレーション $\{\mathcal{F}_t^{00}\}$ が与えられたとき，$\mathcal{F}_t^0 := \sigma(\mathcal{F}_t^{00} \cup \mathcal{N}), \mathcal{F}_t := \mathcal{F}_{t+}^0$ としてフィルトレーションを拡大すれば，$\{\mathcal{F}_t\}$ は通常の仮定をみたすフィルトレーションである．以下では，$\{\mathcal{F}_t\}$ は通常の仮定をみたすとする．

関数 $T : \Omega \to [0, \infty]$ が 停止時刻 (stopping time) であるとは，各 $t \in \mathbb{R}_+$ に対して，$\{T \leq t\} \in \mathcal{F}_t$ となることを言う．$T < \infty$ a.s. のとき，T を有限な停止時刻と呼び，ある定数 K に対して，$T \leq K$ a.s. となるとき，T を有界な停止時刻と呼ぶ．$\{\mathcal{F}_t\}$ の右連続性より，次の補題が成り立つ．

Lemma 30.1. T が停止時刻 $\Leftrightarrow \{T < t\} \in \mathcal{F}_t\ (\forall t \in (0, \infty))$.

Proof. $\Rightarrow$. $\{T < t\} = \bigcup_{n=1}^{\infty} \{T \leq t - 1/n\} \in \mathcal{F}_t$.

$\Leftarrow$. $\{T \leq t\} = \bigcap_{m=n}^{\infty} \{T < t + 1/m\} \in \mathcal{F}_{t+1/n}$. n は任意であるから，$\{T \leq t\} \in \bigcap_{n=1}^{\infty} \mathcal{F}_{t+1/n} = \mathcal{F}_{t+} = \mathcal{F}_t$ を得る. □

S, T が停止時刻なら，$S \vee T, S \wedge T, S + T$ も停止時刻である．特に，定数 $S \equiv t \geq 0$ は停止時刻なので，$T + t$ は停止時刻である．また，T_n が停止時刻の列ならば，

$$\sup_n T_n,\ \inf_n T_n,\ \limsup_n T_n,\ \liminf_n T_n$$

はすべて停止時刻である．

確率過程 $\{X_t : t \in \mathbb{R}_+\}$ が $\{\mathcal{F}_t\}$ に 適合 (adapted) しているとは，各 $t \in \mathbb{R}_+$ に対して，$X_t \in \mathcal{F}_t$ となることを言う．適合過程 $\{X_t\}$ と $A \in \mathcal{B}$ に対して，

$$T_A = \inf\{t \geq 0 : X_t \in A\}, \quad \inf \varnothing = +\infty$$

を $\{X_t\}$ の A への 到達時刻 (hitting time) と呼ぶ．

Lemma 30.2. 次のいずれかの条件がみたされるなら，T_A は停止時刻である．

(a) A が開集合で，$\{X_t\}$ が右連続 (or 左連続) なパスをもつ．

(b) A が閉集合で，$\{X_t\}$ が連続なパスをもつ．

Proof. (a). $\{T_A < t\} = \bigcup_{q\in[0,t)\bigcap\mathbb{Q}}\{X_q \in A\} \in \mathcal{F}_t$.
(b). $d(x,A) = \inf_{y\in A}|x-y|$ に対して，$x \in A \Leftrightarrow d(x,A) = 0$ である．ここで，写像 $x \mapsto d(x,A)$ は連続であるから，$\{T_A \le t\} = \{\inf_{q\in[0,t]\cap\mathbb{Q}} d(X_q,A) = 0\} \in \mathcal{F}_t$ を得る． □

Remark 30.1. 到達時刻 T_A はもっと弱い仮定のもとで停止時刻になることが知られている：S を距離空間とし，$\{X_t\}$ を右連続か左連続なパスをもつ S 値適合過程とすると，任意の Borel 集合 $A \subset S$ に対して，T_A は停止時刻になる．さらに，パスに関する制約も，発展的可測性というもっと弱い仮定に取り替えることができる．この一見すると単純なステートメントの証明には，深い測度論の議論を必要とする．第 39 節を参照せよ．ただし，本講義ノートの範囲では上の補題で十分である．

Lemma 30.3 (Durrett 8.3.1)**.** 有限な停止時刻 T が与えられたとき，各 $n \in \mathbb{N}$ に対して，$T_n(\omega) := \sum_{k=1}^{\infty}(k/2^n)1_{[(k-1)/2^n,k/2^n)}(T(\omega))$ は停止時刻であって，$T_n \downarrow T$ a.s. $(n \to \infty)$ をみたす．T_n を T の離散近似と呼ぶことにする．

確率過程 $\{X_t\}$ が 可測 であるとは，$(t,\omega) \mapsto X_t(\omega)$ が $\mathcal{B}_{\mathbb{R}_+} \times \mathcal{F}/\mathcal{B}$ 可測になることを言う．$\{X_t\}$ が可測なら，有限な停止時刻 T に対して，$X_{T(\omega)}(\omega)$ は a.s. に定義された r.v. である．Exercise 29.3 より，$\{X_t\}$ が右連続なパスをもてばそれは可測である．
停止時刻 T に対して，

$$\mathcal{F}_T = \{A \in \mathcal{F}_\infty : A \cap \{T \le t\} \in \mathcal{F}_t \ \forall t \in \mathbb{R}_+\}$$

と定義する．ただし，$\mathcal{F}_\infty := \bigvee_{t\in\mathbb{R}_+}\mathcal{F}_t$ である．$T \equiv t \in \mathbb{R}_+ \cup \{+\infty\}$ なら，$\mathcal{F}_T = \mathcal{F}_t$ である．また，$\{\mathcal{F}_t\}$ の右連続性より，

$$\mathcal{F}_T = \{A \in \mathcal{F}_\infty : A \cap \{T < t\} \in \mathcal{F}_t \ \forall t \in (0,\infty)\}$$

とも表せる (Durrett 8.3.4).

Lemma 30.4. (i) $\mathcal{F}_T$ は σ-field である．(ii) $S \le T$ なら，$\mathcal{F}_S \subset \mathcal{F}_T$. (iii) $\mathcal{F}_T = \bigcap_{\varepsilon>0}\mathcal{F}_{T+\varepsilon}$. (iv) $\{X_t\}$ が右連続なパスをもつ適合過程とし，T を有限な停止時刻とすると，$X_T \in \mathcal{F}_T$.

Proof. (i) はよい．(ii). $A \in \mathcal{F}_S$ なら，$A\cap\{T \le t\} = (A\cap\{S \le t\})\cap\{T \le t\} \in \mathcal{F}_t$. (iii). (ii) より，$\mathcal{F}_T \subset \mathcal{F}_{T+\varepsilon}$ であるから，$\mathcal{F}_T \subset \bigcap_{\varepsilon>0}\mathcal{F}_{T+\varepsilon}$. 逆の包含関係を示す．$A \in \bigcap_{\varepsilon>0}\mathcal{F}_{T+\varepsilon}$ に対して，$A\bigcap\{T+1/n \le t\} \in \mathcal{F}_t$. $n \to \infty$ として，$\{T+1/n \le t\} = \{T \le t-1/n\} \uparrow \{T < t\}$ となるから，$A \cap \{T < t\} \in \mathcal{F}_t$. (iv). T_n を T の離散近似とすると，$A \in \mathcal{B}$ に対して，

$$\{X_{T_n} \in A\} \cap \{T_n = k/2^n\} = \{X_{k/2^n} \in A\} \cap \{T_n = k/2^n\} \in \mathcal{F}_{k/2^n}$$

である．T_n は $\{k/2^n : k \in \mathbb{N}\}$ にしか値をとらないから，

$$\{X_{T_n} \in A\} \cap \{T_n \le t\} = \bigcup_{k:k/2^n \le t} (\{X_{T_n} \in A\} \cap \{T_n = k/2^n\}) \in \mathcal{F}_t$$

である．よって，$X_{T_n} \in \mathcal{F}_{T_n} \subset \mathcal{F}_{T+1/2^n}$．$n \ge m$ に対して，$X_{T_n} \in \mathcal{F}_{T+1/2^n} \subset \mathcal{F}_{T+1/2^m}$ であって，X_t の右連続性より，$X_{T_n} \to X_T$ a.s. $(n \to \infty)$ となるから，$X_T \in \mathcal{F}_{T+1/2^m}$．$m$ は任意であるから，$X_T \in \bigcap_m \mathcal{F}_{T+1/2^m} = \mathcal{F}_T$ を得る． □

次にマルチンゲールを定義しよう．確率過程 $\{X_t : t \in \mathbb{R}_+\}$ が，次の 3 条件をみたすとき，それを $\{\mathcal{F}_t\}$ に関するマルチンゲールと呼ぶ：

(i) $X_t \in L^1$ $\forall t \in \mathbb{R}_+$.

(ii) $\{X_t\}$ は $\{\mathcal{F}_t\}$ に適合している．

(iii) $E[X_t \mid \mathcal{F}_s] = X_s$ a.s. for $0 \le s < t$.

(iii) の代わりに，$E[X_t \mid \mathcal{F}_s] \ge X_s$ (or $\le X_s$) a.s. for all $0 \le s < t$ が成り立つとき，$\{X_t\}$ を劣 (or 優) マルチンゲールと呼ぶ．

Remark 30.2. 停止時刻，適合性，およびマルチンゲールは，通常の仮定をみたさないフィルトレーションに対してもそのまま定義される．しかし，以下の議論では，引き続き，$\{\mathcal{F}_t\}$ は通常の仮定をみたすとする．この仮定のもとで，$\mathcal{N} \subset \mathcal{F}_t$ であるから，$\{X_t\}$ が適合過程 (マルチンゲール) であって，$\{\tilde{X}_t\}$ が $\{X_t\}$ の修正なら，$\{\tilde{X}_t\}$ も適合過程 (マルチンゲール) である．

離散時間のマルチンゲールに対して成り立つ多くの性質は連続時間のマルチンゲールに対しても自然に拡張される．まず，最大不等式を示す．

Theorem 30.1 (Doob の不等式と L^p 最大不等式)**.** $\{X_t\}$ を右連続なパスをもつ劣マルチンゲールとし，$t, \alpha > 0$ とする．このとき，

$$\alpha P\left(\sup_{0 \le s \le t} X_s \ge \alpha\right) \le E\left[X_t, \{\sup_{0 \le s \le t} X_s \ge \alpha\}\right] \le E[X_t^+]$$

が成り立つ．さらに，$X_s \ge 0$ a.s. $(\forall s \in \mathbb{R}_+)$ であれば，$1 < p < \infty$ に対して，

$$E\left[\sup_{0 \le s \le t} X_s^p\right] \le \left(\frac{p}{p-1}\right)^p E[X_t^p]$$

が成り立つ．

Proof. Doob の不等式のみ示せば十分である．$\{X_{tk/2^n} : k \in \mathbb{Z}_+\}$ は $\{\mathcal{F}_{tk/2^n} : k \in \mathbb{Z}_+\}$ に関する劣マルチンゲールであるから，Theorem 22.1 より，

$$\alpha P\left(\max_{0\le k\le 2^n} X_{tk/2^n} > \alpha\right) \le E\left[X_t, \{\max_{0\le k\le 2^n} X_{tk/2^n} \ge \alpha\}\right] \le E\left[X_t, \{\sup_{0\le s\le t} X_t \ge \alpha\}\right]$$

が成り立つ．$n \to \infty$ のとき，パスの右連続性より，

$$\left\{\max_{0\le k\le 2^n} X_{tk/2^n} > \alpha\right\} \uparrow \left\{\sup_{0\le s\le t} X_s > \alpha\right\}$$

となるから，求める結論を得る． □

次に収束定理を示す．

Theorem 30.2 (収束定理)**.** $\{X_t\}$ を右連続なパスをもつ劣マルチンゲールとし，

$$\sup_{t\in\mathbb{R}_+} E[X_t^+] < \infty \tag{$*3$}$$

とする．このとき，$\exists X_\infty \in L^1$ s.t. $\lim_{t\to\infty} X_t = X_\infty$ a.s. となる．さらに，$\{X_t\}$ が一様可積分なら，$X_t \to X_\infty$ in L^1 $(t \to \infty)$ となり，$\{X_n\}$ が $\sup_{t\in\mathbb{R}_+} E[|X_t|^p] < \infty$ $(1 < \exists p < \infty)$ をみたすマルチンゲールか非負劣マルチンゲールなら，$X_t \to X_\infty$ in L^p $(t \to \infty)$ となる．

$\{X_t\}$ が劣マルチンゲールなら，$E[|X_t|] = 2E[X_t^+] - E[X_t] \le 2E[X_t^+] - E[X_0]$. よって，$(*3)$ は $\{X_t\}$ が L^1 有界であることと同値である．

Proof. 前半のみ示せば十分である．$a < b$ とし，各 $n \in \mathbb{N}$ に対して，$X_{tk/2^n}$, $k = 0, 1, \dots, 2^n$ が $[a, b]$ を上向きに横断する回数を $U_{t,n}[a, b]$ とおくと，Theorem 21.2 より，

$$(b - a)E[U_{t,n}[a, b]] \le E[(X_t - a)^+]$$

が成り立つ．$U_{t,n}[a, b]$ は n に関して非減少なので，$U_{t,n}[a, b] \uparrow U_t[a, b] := \sup_m U_{t,m}[a, b]$ $(n \to \infty)$ となる．さらに，Fatou の補題より，

$$E[U_t[a, b]] \le \frac{E[(X_t - a)^+]}{b - a}$$

を得る．右辺は t に関して有界であって，$U_n[a, b] \uparrow U_\infty[a, b] := \sup_m U_m[a, b]$ $(n \to \infty)$ であるから，Fatou の補題より，$E[U_\infty[a, b]] < \infty$ を得る．特に，$U_\infty[a, b] < \infty$ a.s. であるから，

$$F_{a,b} = \left\{\omega : \liminf_{t\to\infty, t\in\mathbb{Q}_2} X_t(\omega) < a < b < \limsup_{t\to\infty, t\in\mathbb{Q}_2} X_t(\omega)\right\}$$

とおいたとき，$P(F_{a,b}) = 0$ である．ただし，$\mathbb{Q}_2$ は 2 進有理数の全体である．あとは離散時間の場合と同様にして，$\exists X_\infty \in L^1$ s.t. $\lim_{t\to\infty, t\in\mathbb{Q}_2} X_t = X_\infty$ a.s. となることが示される．パスの右連続性より，これは $\lim_{t\to\infty} X_t = X_\infty$ a.s. を意味する． □

$\{X_t\}$ が右連続なパスをもつ L^1 有界な劣マルチンゲールなら，a.s. に一意な r.v. $X_\infty \in L^1$ が存在して，$\lim_{t\to\infty} X_t = X_\infty$ a.s. となる ($\mathcal{N} \subset \mathcal{F}_\infty$ であるから，X_∞ はどのように選んでも $\mathcal{F}_\infty$ 可測である)．そこで，そのような場合は，与えられた停止時刻 T に対して，$T(\omega) = \infty$ のときは，$X_{T(\omega)}(\omega) = X_\infty(\omega)$ と定義しておけば，$X_{T(\omega)}(\omega)$ がすべての $\omega \in \Omega$ に対して定義される．このとき，$X_{T\wedge n}$ は $\mathcal{F}_{T\wedge n}$ ($\subset \mathcal{F}_T$) 可測な r.v. であって，$X_T = \lim_n X_{T\wedge n}$ であるから，X_T は $\mathcal{F}_T$ 可測な r.v. である．

Theorem 30.3 (任意停止定理)**.** $\{X_t\}$ を右連続なパスをもつ劣マルチンゲールとし，$S \le T$ を停止時刻とする．このとき，$\{X_t\}$ が一様可積分であるか，または T が有界であれば，X_S と X_T は可積分であって，$X_S \le E[X_T \mid \mathcal{F}_S]$ a.s. が成り立つ．

もちろん，$\{X_t\}$ が右連続なパスをもつマルチンゲールであって，$S \le T$ が停止時刻であるとき，$\{X_t\}$ が一様可積分であるか，または T が有界であれば，$X_S = E[X_T \mid \mathcal{F}_S]$ a.s. が成り立つ ($\{X_t\}$ と $\{-X_t\}$ に対して定理を適用すればよい)．

Proof. S_n, T_n をそれぞれ S, T の離散近似とする．このとき，$\{X_{k/2^n} : k \in \mathbb{Z}_+\}$ は $\{\mathcal{F}_{k/2^n} : k \in \mathbb{Z}_+\}$ に関する劣マルチンゲールであって，$S_n \le T_n$ は $\{k/2^n : k \in \mathbb{Z}_+\}$ に値をとる．ここで，$\{X_t\}$ が一様可積分であるか T が有界であれば，$\{X_{T_n\wedge(k/2^n)} : k \in \mathbb{Z}_+\}$ は一様可積分であるから，Theorem 24.1 より，X_{S_n} と X_{T_n} は可積分であって，$X_{S_n} \le E[X_{T_n} \mid \mathcal{F}_{S_n}]$ が成り立つ．よって，$A \in \mathcal{F}_S \subset \mathcal{F}_{S_n}$ に対して，$E[X_{S_n}, A] \le E[X_{T_n}, A]$ を得る．一方，$T_{n+1} \le T_n$ より，$X_{T_{n+1}} \le E[X_{T_n} \mid \mathcal{F}_{T_{n+1}}]$ であるから，$\{X_{T_{-n}} : n \le -1\}$ は $\{\mathcal{F}_{T_{-n}} : n \le -1\}$ に関するバックワード・劣マルチンゲールである．ここで，$\inf_n E[X_{T_n}] \ge E[X_0] > -\infty$ であるから，Theorem 23.2 とパスの右連続性より，X_T は可積分であって，$X_{T_n} \to X_T$ a.s. and in L^1 となる．同様に，X_S は可積分であって，$X_{S_n} \to X_S$ a.s. and in L^1 となる．従って，$A \in \mathcal{F}_S$ に対して，$E[X_S, A] \le E[X_T, A]$ となるから，$X_S \le E[X_T \mid \mathcal{F}_S]$ a.s. を得る． □

Theorem 30.4. $\{X_t : t \in \mathbb{R}_+\}$ を右連続なパスをもつ (劣) マルチンゲールとし，T を停止時刻とする．このとき，$\{X_{T\wedge t} : t \in \mathbb{R}_+\}$ も (劣) マルチンゲールである．

Proof. 劣マルチンゲールの場合のみ考えればよい．$X_{T\wedge t} \in \mathcal{F}_{T\wedge t} \subset \mathcal{F}_t$ より，$\{X_{T\wedge t} : t \in \mathbb{R}_+\}$ は $\{\mathcal{F}_t\}$ に適合している．$t > s \ge 0$ とし，$A \in \mathcal{F}_s$ とする．このとき，$A \cap \{T > s\} \in \mathcal{F}_{T\wedge s}$ である．また，Theorem 30.3 より，$X_{T\wedge s}$ と $X_{T\wedge t}$ は可積分であって，

$$\begin{aligned} E[X_{T\wedge t}.A] &= E[X_T, A \cap \{T \le s\}] + E[X_{T\wedge t}, A \cap \{T > s\}] \\ &\ge E[X_T, A \cap \{T \le s\}] + E[X_{T\wedge s}, A \cap \{T > s\}] = E[X_{T\wedge s}, A] \end{aligned}$$

が成り立つ．よって，$X_{T\wedge s} \le E[X_{T\wedge t} \mid \mathcal{F}_s]$ を得る． □

Corollary 30.1. $\{X_t\}$ を右連続なパスをもつ一様可積分なマルチンゲール (非負劣マルチンゲール) とし，T を停止時刻とする．このとき，$\{X_{T\wedge t} : t \in \mathbb{R}_+\}$ も一様可積分なマルチンゲール (非負劣マルチンゲール) である．

Proof. $|X_{T\wedge t}| \le E[|X_T| \mid \mathcal{F}_{T\wedge t}]$ より，$\{X_{T\wedge t} : t \in \mathbb{R}_+\}$ は一様可積分である． □

連続時間において Doob 分解に対応する結果として，Doob-Meyer 分解があり，これは確率積分の基礎となる重要な結果である．ただ，そのステートメントと証明は多くのページ数を要するのと，本講義ノートでは確率積分を扱わないので，Doob-Meyer 分解は扱わない．Doob-Meyer 分解の証明は Karatzas and Shreve (1997) にある．また，もっと短い証明が Beiglbock et al. (2012) にある．

Doob の正則化定理

さて，これまでの議論において，$\{\mathcal{F}_t\}$ は通常の仮定をみたし，(劣) マルチンゲール $\{X_t\}$ は右連続なパスをもつと仮定してきたが，その妥当性を検証しよう．通常の仮定をみたすとは限らないフィルトレーション $\{\mathcal{F}_t^{00}\}$ が与えられたとしても，$\mathcal{F}_t^0 := \sigma(\mathcal{F}_t^{00} \cup \mathcal{N}), \mathcal{F}_t := \mathcal{F}_{t+}^0$ としてフィルトレーションを拡大すれば，$\{\mathcal{F}_t\}$ は通常の仮定をみたす．フィルトレーションの拡大は停止時刻の条件を緩和するので，T が $\{\mathcal{F}_t^{00}\}$ に関する停止時刻であれば，$\{\mathcal{F}_t\}$ に関する停止時刻でもある．しかし，マルチンゲール性が保たれるかどうかは明らかではない．しかし，じつは $\{X_t\}$ が $\{\mathcal{F}_t^{00}\}$ に関する劣マルチンゲールなら，緩やかな条件のもとで，$\{X_t\}$ の修正 $\{\tilde{X}_t\}$ であって，右連続なパスをもち，$\{\mathcal{F}_t\}$ に関する劣マルチンゲールとなるものが存在する．2 つの確率過程 $\{X_t\}$ と $\{\tilde{X}_t\}$ が互いの修正であるとは，$P(X_t = \tilde{X}_t) = 1$ $(\forall t \in \mathbb{R}_+)$ となることを言うのであった．関数 $f : \mathbb{R}_+ \to \mathbb{R}$ が càdlàg (“continue à droite, limite à gauche”) であるとは，それが右連続であって，あらゆる $t > 0$ に対して有限な左極限 $\lim_{s\uparrow t} f(s)$ をもつことを言う．

Theorem 30.5 (Doob の正則化定理)**.** $\{\mathcal{F}_t^{00}\}$ を通常の仮定をみたすとは限らないフィルトレーションとし，$\mathcal{F}_t^0 = \sigma(\mathcal{F}_t^{00} \cup \mathcal{N})$, $\mathcal{F}_t = \mathcal{F}_{t+}^0$ とおく．また，$\{X_t\}$ を $\{\mathcal{F}_t^{00}\}$ に関する劣マルチンゲールとする．このとき，$P(\Omega_0) = 1$ なる事象 Ω_0 が存在して，あらゆる $\omega \in \Omega_0$ と $t \in \mathbb{R}_+$ に対して，有限な

$$X_{t+}(\omega) := \lim_{s\downarrow t, s\in\mathbb{Q}_2} X_s(\omega)$$

が存在する．そこで，$\omega \in \Omega_0^c$ に対しては，$X_{t+}(\omega) = 0$ $(\forall t \in \mathbb{R}_+)$ と定義すると，$\{X_{t+}\}$ は càdlàg なパスをもつ $\{\mathcal{F}_t\}$ に関する劣マルチンゲールである．さらに，$\{X_{t+}\}$ が $\{X_t\}$ の修正であるためには，$t \mapsto X_t$ が L^1 への写像として右連続であること，すなわち，あらゆる $t \in \mathbb{R}_+$ に対して，

$$\lim_{s\downarrow t} E[|X_s - X_t|] = 0$$

となることが必要十分である．

Proof. $t > s \ge 0$ に対して，$X_s \le E[X_t \mid \mathcal{F}_s^{00}] = E[X_t \mid \mathcal{F}_s^0]$ a.s. であるから，$\{X_t\}$ は $\{\mathcal{F}_t^0\}$ に関する劣マルチンゲールでもある (cf. Exercise 30.1)．また，$t \in \mathbb{R}_+$ と $t_n \downarrow t$ な

る数列 t_n に対して，$\{X_{t_{-n}} : n \leq -1\}$ は $\{\mathcal{F}^0_{t_{-n}} : n \leq -1\}$ に関するバックワード・劣マルチンゲールであるから，$\{X_{t_n}\}$ は一様可積分である．

$a < b$ とし，各 $n \in \mathbb{N}$ に対して，$X_{tk/2^n},\ k = 0, 1, \ldots, 2^n$ が $[a, b]$ を上向きに横断する回数を $U_{t,n}[a, b]$ とおく．また，$U_n[a, b] = \sup_m U_{m,n}[a, b]$ とおくと，Theorem 30.2 の証明より，$E[U_n[a, b]] \leq (b - a)^{-1}E[(X_n - a)^+] < \infty$ であるから，

$$F := \bigcup_{n\in\mathbb{N}} \bigcup_{\substack{a<b\\ a,b\in\mathbb{Q}}} \{U_n[a, b] = \infty\}$$

とおくと，$P(F) = 0$ である．ここで，仮にある $\omega \in \Omega$ と $0 \leq t < n$ において，

$$\liminf_{s\downarrow t, s\in\mathbb{Q}_2} X_s(\omega) < \limsup_{s\downarrow t, s\in\mathbb{Q}_2} X_s(\omega)$$

であるならば，ある $a, b \in \mathbb{Q}, a < b$ と数列 $s_k \in \mathbb{Q}_2$ が存在して，$X_{s_{2k-1}}(\omega) < a < b < X_{s_{2k}}(\omega)$ となるから，$U_n[a, b](\omega) = \infty$ となる．よって，$\omega \in F^c$ に対して，

$$\liminf_{s\downarrow t, s\in\mathbb{Q}_2} X_s(\omega) = \limsup_{s\downarrow t, s\in\mathbb{Q}_2} X_s(\omega) = \lim_{s\downarrow t, s\in\mathbb{Q}_2} X_s(\omega)$$

が $[-\infty, \infty]$ の範囲で存在する．さらに，Theorem 30.1 の証明の議論に Corollary 22.1 を適用して，各 $n \in \mathbb{N}$ と $\alpha > 0$ に対して，

$$\alpha P\left(\sup_{s\in\mathbb{Q}_2\cap[0,n]} |X_s| > \alpha\right) \leq 2E[X_n^+] - E[X_0]$$

となることが示されるので，

$$G = \bigcup_n \left\{\sup_{s\in\mathbb{Q}_2\cap[0,n]} |X_s| = \infty\right\}$$

とおくと，$P(G) = 0$ である．よって，$\Omega_0 = F^c \cap G^c$ とおくと，$P(\Omega_0) = 1$ であって，あらゆる $\omega \in \Omega_0$ と $t \in \mathbb{R}_+$ に対して，有限な $\lim_{s\downarrow t, s\in\mathbb{Q}_2} X_s(\omega)$ が存在する．ここで，その構成から，$\{X_{t+}\}$ は càdlàg なパスをもつ．

次に，$\{X_{t+}\}$ が $\{\mathcal{F}_t\}$ に関する劣マルチンゲールであることを示す．$\{X_{t+}\}$ が $\{\mathcal{F}_t\}$ に適合していることは明らか．また，$t \in \mathbb{R}_+$ に対して，$t_n \in \mathbb{Q}_2$ を $t_n \downarrow t$ なる数列とすると，$\{X_{t_n}\}$ は一様可積分であるから，$X_{t+} = \lim_n X_{t_n}$ は可積分である．さらに，$t > s \geq 0$ とし，$s_n \in \mathbb{Q}_2$ を $s_n \downarrow s$ なる数列とすると，

$$X_{s_n} \leq E[X_t \mid \mathcal{F}^0_{s_n}] \quad a.s.$$

である．ここで，$X_{s_n} \to X_{s+}$ であって，一方，Lévy の下向き定理より，$E[X_t \mid \mathcal{F}^0_{s_n}] \to E[X_t \mid \mathcal{F}^0_{s+}] = E[X_t \mid \mathcal{F}_s]$ であるから，

$$X_{s+} \leq E[X_t \mid \mathcal{F}_s] \quad a.s.$$

を得る．t_n を先ほどと同様に選ぶと，$X_{s+} \le E[X_{t_n} \mid \mathcal{F}_s]$ であって，$\{X_{t_n}\}$ は一様可積分であったから，$n \to \infty$ として，

$$X_{s+} \le E[X_{t+} \mid \mathcal{F}_s] \quad a.s.$$

を得る．よって，$\{X_{t+}\}$ は $\{\mathcal{F}_t\}$ に関する劣マルチンゲールである．

さて，$\{X_{t+}\}$ が $\{X_t\}$ の修正であるなら，$t \in \mathbb{R}_+$ と $s_n \downarrow t$ なる s_n に対して，$\{X_{s_n}\}$ は一様可積分であって，$X_{s_n} \to X_{t+} = X_t$ a.s. であるから，$E[|X_{s_n} - X_t|] \to 0$ を得る．逆に，$t \mapsto X_t$ が L^1 への写像として右連続なら，$t \in \mathbb{R}_+$ と $s_n \downarrow t$ なる s_n に対して，Fatou の補題より，

$$E[|X_{t+} - X_t|] \le \liminf_n E[|X_{s_n} - X_t|] = 0$$

であるから，$X_{t+} = X_t$ a.s. を得る． □

Corollary 30.2. $\{\mathcal{F}_t\}$ を通常の仮定をみたすフィルトレーションとし，$\{X_t\}$ を $\{\mathcal{F}_t\}$ に関する劣マルチンゲールとする．このとき，$\{X_t\}$ に対して càdlàg なパスをもつ修正 $\{\check{X}_t\}$ が存在するためには，$t \mapsto E[X_t]$ が右連続であることが必要十分である．

Proof. 必要性は明らか．十分性を示す．$t \in \mathbb{R}_+, s_n \in \mathbb{Q}_2, s_n \downarrow t$ とすると，$X_t \le E[X_{s_n} \mid \mathcal{F}_t]$ である．$\{X_{s_n}\}$ は一様可積分であって，$X_{s_n} \to X_{t+}$ であるから，$X_t \le E[X_{t+} \mid \mathcal{F}_t] = X_{t+}$ a.s. を得る．一方，$E[X_t] = \lim_n E[X_{s_n}] = E[X_{t+}]$ より，$X_t = X_{t+}$ a.s. を得る． □

Example 30.1. $Y \in L^1$ とし，$X_t = E[Y \mid \mathcal{F}_t]$ a.s. $t \in \mathbb{R}_+$ とする．$\{X_t\}$ は一様可積分なマルチンゲールである．さらに，$\{X_t\}$ は右連続なパスをもつ修正に選ぶ．$t \mapsto E[X_t] = E[X_0]$ は一定であるから，そのような修正は存在する．このとき，Lévy の上向き定理より，$X_\infty = E[Y \mid \mathcal{F}_\infty]$ a.s. であって，任意停止定理より，あらゆる停止時刻 T に対して，$X_T = E[X_\infty \mid \mathcal{F}_T] = E[Y \mid \mathcal{F}_T]$ a.s. が成り立つ．

演習問題

Exercise 30.1. $\mathcal{G}$ を $\mathcal{F}$ の sub σ-field とすると，

$$\sigma(\mathcal{G} \cup \mathcal{N}) = \{A : \exists \check{A} \in \mathcal{G} \text{ s.t. } A \Delta \check{A} \in \mathcal{G}\}$$

を示せ．さらに，$Y \in L^1$ に対して，$E[Y \mid \sigma(\mathcal{G} \cup \mathcal{N})] = E[Y \mid \mathcal{G}]$ a.s. を示せ．

Exercise 30.2. S を $\{\mathcal{F}_t\}$ に関する停止時刻とし，T を $\{\mathcal{F}_{S+t} : t \in \mathbb{R}_+\}$ に関する停止時刻とする．このとき，$S + T$ は $\{\mathcal{F}_t\}$ に関する停止時刻であることを示せ．

Exercise 30.3. $\{X_t\}$ を右連続なパスをもつ劣マルチンゲールとし，ある可積分 r.v. Y が存在して，$X_t \le E[Y \mid \mathcal{F}_t]$ a.s. $(\forall t \in \mathbb{R}_+)$ とする．このとき，$\exists X_\infty \in L^1$ s.t. $\lim_{t\to\infty} X_t = X$ a.s. を示せ．また，任意の停止時刻 $S \le T$ に対して，X_S と X_T は可積分であって，$X_S \le E[X_T \mid \mathcal{F}_S]$ a.s. が成り立つことを示せ．

Exercise 30.4. S, T を停止時刻とする.

(a) $\mathcal{F}_S \cap \mathcal{F}_T = \mathcal{F}_{S \wedge T}$ を示せ.

(b) $Y \in L^1$ に対して，$E[E[Y \mid \mathcal{F}_S] \mid \mathcal{F}_T] = E[Y \mid \mathcal{F}_{S \wedge T}]$ a.s. を示せ.

31 Brown 運動のマルコフ性とマルチンゲール性

$(\Omega, \mathcal{F}, P)$ を完備な確率空間とし，標準 BM$\{B_t : t \in \mathbb{R}_+\}$ に対して，

$$\mathcal{F}_t^0 := \sigma(B_s, s \le t) := \bigvee_{s \le t} \sigma(B_s),\ t \in \mathbb{R}_+$$

を $\{B_t\}$ の自然なフィルトレーションと呼ぶ．ただし，このままだと $\{\mathcal{F}_t^0\}$ は通常の仮定をみたさないので，

$$\mathcal{F}_t := \sigma(\mathcal{F}_t^0 \cup \mathcal{N})$$

とフィルトレーションを拡大する．ここで，

$$\mathcal{F}_t = \{A : \exists \check{A} \in \mathcal{F}_t^0 \text{ s.t. } A \Delta \check{A} \in \mathcal{N}\}$$

であって，$Y \in L^1$ に対して，$E[Y \mid \mathcal{F}_t] = E[Y \mid \mathcal{F}_t^0]$ a.s. である (Exercise 30.1)．じつは $\{\mathcal{F}_t\}$ は自動的に右連続になる．

Theorem 31.1. (i) (マルコフ性) $s > 0$ を任意に固定する．$\{B_{t+s} - B_s : t \in \mathbb{R}_+\}$ は $\mathcal{F}_{s+}$ と独立な標準 BM となる．(ii) 各 $t \in \mathbb{R}_+$ に対して，$\mathcal{F}_{t+} = \mathcal{F}_t$.

Proof. (i). $\{B_{t+s} - B_t : t \in \mathbb{R}_+\}$ と $\mathcal{F}_{s+}$ が独立なことを示せば十分である．

Step 1. $0 < t_1 < \cdots < t_n$ と $0 < s_1 < \cdots < s_m \le s$ に対して，確率ベクトル $(B_{t_k+s} - B_s)_{1 \le k \le n}$ と $(B_{s_\ell})_{1 \le \ell \le m}$ の独立性を示す．これは，$t_0 = s_0 = 0$ とおくと，$(B_{t_k+s} - B_{t_{k-1}+s})_{1 \le k \le n}$ と $(B_{s_\ell} - B_{s_{\ell-1}})_{1 \le \ell \le m}$ の独立性から従う．

Step 2. 次に，$\mathcal{C}$ を

$$\{(B_{s_1}, \ldots, B_{s_m}) \in A\},\ A \in \mathcal{B}^m, 0 \le s_1, \ldots, s_m \le s, m \in \mathbb{N}$$

なる形の集合の全体とすると，$\mathcal{C}$ は Ω を含む π システムであって，$\sigma(\mathcal{C}) = \mathcal{F}_s^0$ であるから，Step 1 より，$(B_{t_k+s} - B_s)_{1 \le k \le n}$ と $\mathcal{F}_s^0$ (従って，$\mathcal{F}_s$) は独立である．

Step 3. Y を $\mathcal{F}_{s+}$ 可測な任意の r.v. とすると，Step 2 より，$\varepsilon > 0$ に対して，$(B_{t_k+s+\varepsilon} - B_{s+\varepsilon})_{1 \le k \le n}$ と Y は独立であるから，任意の $u_1, \ldots, u_n, v \in \mathbb{R}$ に対して，

$$E\left[e^{i\sum_{k=1}^n u_k (B_{t_k+s+\varepsilon} - B_{s+\varepsilon}) + ivY}\right] = E\left[e^{i\sum_{k=1}^n u_k (B_{t_k+s+\varepsilon} - B_{s+\varepsilon})}\right] E\left[e^{ivY}\right]$$

が成り立つ．ここで，$\varepsilon \downarrow 0$ として，

$$E\left[e^{i\sum_{k=1}^n u_k (B_{t_k+s} - B_s) + ivY}\right] = E\left[e^{i\sum_{k=1}^n u_k (B_{t_k+s} - B_s)}\right] E\left[e^{ivY}\right]$$

を得る．よって，$(B_{t_k+s} - B_s)_{1 \le k \le n}$ と $\mathcal{F}_{s+}$ は独立である．あとは，Step 2 と同様の議論により，$\{B_{t+s} - B_s : t \in \mathbb{R}_+\}$ と $\mathcal{F}_{s+}$ の独立性が従う．

(ii) **Step 1**. $0 \le t_1 < \cdots < t_n, t_i = t$ とする．$A \in \mathcal{B}^n$ に対して，変換

$$(x_1, \ldots, x_i, x_{i+1}, \ldots, x_n) \mapsto (x_1, \ldots, x_i, x_{i+1} - x_i, \ldots, x_n - x_i)$$

による像を A' とおくと，

$$\{(B_{t_1},\dots,B_{t_n})\in A\}=\{(B_{t_1},\dots,B_{t_i},B_{t_{i+1}}-B_{t_i},\dots,B_{t_n}-B_{t_i})\in A'\}.$$

(i) より，$(B_{t_{i+1}}-B_{t_i},\dots,B_{t_n}-B_{t_i})$ と $\mathcal{F}_{t+}$ は独立であるから，

$$\varphi(x_1,\dots,x_i)=P((x_1,\dots,x_i,B_{t_{i+1}}-B_{t_i},\dots,B_{t_n}-B_{t_i})\in A')$$

とおくと，

$$P((B_{t_1},\dots,B_{t_n})\in A\mid\mathcal{F}_{t+})=\varphi(B_{t_1},\dots,B_{t_i})\ a.s.$$

従って，$P((B_{t_1},\cdots,B_{t_n})\in A\mid\mathcal{F}_{t+})\in\mathcal{F}_t$.

Step 2. 次に，$\mathcal{P}$ を

$$\{(B_{t_1},\dots,B_{t_n})\in A\},\ A\in\mathcal{B}^n,t_1,\dots,t_n\in\mathbb{R}_+,n\in\mathbb{N}$$

なる形の集合の全体とし，$\mathcal{L}=\{A\in\mathcal{F}:P(A\mid\mathcal{F}_{t+})\in\mathcal{F}_t\}$ とおくと，$\mathcal{P}$ は π システムであって, $\mathcal{L}$ は $\mathcal{P}$ を含む λ システムであるから，π-λ 定理より，$\mathcal{L}\supset\sigma(\mathcal{P})=\bigvee_{t\in\mathbb{R}_+}\mathcal{F}_t^0$ を得る．

Step 3. $A\in\mathcal{F}_{t+}$ に対して，$A\in\mathcal{F}_{t'}$ $(\forall t'>t)$. また，$\exists\check{A}\in\mathcal{F}_{t'}^0$ s.t. $A\Delta\check{A}\in\mathcal{N}$ であるから，$1_A=P(A\mid\mathcal{F}_{t+})=P(\check{A}\mid\mathcal{F}_{t+})$ a.s. Step 2 より，最右辺は $\mathcal{F}_t$ 可測であって，$\mathcal{F}_t$ はすべての零集合を含むから，$1_A\in\mathcal{F}_t$, i.e., $A\in\mathcal{F}_t$. □

Remark 31.1. マルコフ性は次のように言い換えることができる．いま，$W\times\mathbb{R}\ni(f,x)\mapsto(f(t)+x)_{t\in\mathbb{R}_+}\in W$ は連続であるから，$\mathcal{B}(W)\times\mathcal{B}$ 可測である．よって，$x+B_\bullet$ の分布を μ_x とおくと，$A\in\mathcal{B}(W)$ に対して，$\mathbb{R}\ni x\mapsto\mu_x(A)=P(x+B_\bullet\in A)$ は可測であって，$\mathcal{F}_s$ を与えたときの $B_{s+\bullet}$ の r.c.d. は $\Omega\times\mathcal{B}(W)\ni(\omega,A)\mapsto\mu_{B_s(\omega)}(A)$ である．r.c.d. の存在は W が Poland 空間であることから保証される．

BM のマルコフ性より，次の系を得る．

Corollary 31.1. (i) $\{B_t\}$ は $\{\mathcal{F}_t\}$ に関するマルチンゲールである．(ii) $\{B_t^2-t:t\in\mathbb{R}_+\}$ は $\{\mathcal{F}_t\}$ に関するマルチンゲールである．(iii) $u\in\mathbb{R}$ を given とすると，$M_t^u=\exp\{uB_t-u^2t/2\},t\in\mathbb{R}_+$ は $\{\mathcal{F}_t\}$ に関するマルチンゲールである．

Proof. (i). $t>s$ とすると，B_t-B_s は $\mathcal{F}_s$ と独立であるから，$E[B_t-B_s\mid\mathcal{F}_s]=E[B_t-B_s]=0$. 一方，左辺 $=E[B_t\mid\mathcal{F}_s]-B_s$ であるから，$E[B_t\mid\mathcal{F}_s]=B_s$ を得る．

(ii). $t>s$ とすると，$B_t^2=(B_s+(B_t-B_s))^2=B_s^2+2B_s(B_t-B_s)+(B_t-B_s)^2$. B_t-B_s は $\mathcal{F}_s$ と独立であるから，

$$E[B_t^2\mid\mathcal{F}_s]=B_s^2+2B_s\underbrace{E[B_t-B_s]}_{=0}+\underbrace{E[(B_t-B_s)^2]}_{=t-s}=B_s^2+t-s.$$

項を入れ替えて，$E[B_t^2-t\mid\mathcal{F}_s]=B_s^2-s$ を得る．

(iii). $t>s$ とすると，$E[e^{uB_t}\mid\mathcal{F}_s]=e^{uB_s}E[e^{u(B_t-B_s)}\mid\mathcal{F}_s]=e^{uB_s}E[e^{u(B_t-B_s)}]=e^{uB_s+u^2(t-s)/2}$ であるから，$E[M_t^u\mid\mathcal{F}_s]=M_s^u$ を得る． □

(iii) の例にもとづいて，B_t の多項式からなるマルチンゲールのクラスを構成しよう．$n \in \mathbb{Z}_+$ に対して，$H_n(x)$ を n 次 Hermite 多項式とする．すなわち，

$$H_n(x) = e^{x^2/2}(-1)^n \frac{d^n}{dx^n}(e^{-x^2/2})$$

である．ただし，$H_0(x) = 1$ である．$H_n(x)$ は x の n 次多項式であって，x^n の係数は 1 である．ここで，

$$e^{-(u-x)^2/2} = \sum_{n=0}^{\infty} \frac{u^n}{n!} \frac{d^n}{du^n} e^{-(u-x)^2/2}\Big|_{u=0} = \sum_{n=0}^{\infty} \frac{u^n}{n!}(-1)^n \frac{d^n}{dx^n}(e^{-x^2/2})$$

であるから，

$$e^{ux-u^2/2} = \sum_{n=0}^{\infty} \frac{u^n}{n!} H_n(x)$$

を得る．いま，$a > 0$ を given とすると，

$$e^{ux-au^2/2} = \exp\left\{u\sqrt{a}\left(\frac{x}{\sqrt{a}}\right) - \left(\frac{u\sqrt{a}}{2}\right)^2\right\} = \sum_{n=0}^{\infty} \frac{u^n}{n!} \underbrace{a^{n/2} H_n(x/\sqrt{a})}_{=\tilde{H}_n(x,a)}$$

を得る．形式的に，$\tilde{H}_n(x,0) = x^n$ とおく．このとき，$L_t^{(n)} = \tilde{H}_n(B_t, t)$ とおくと，$M_t^u = \sum_{n=0}^{\infty}(u^n/n!)L_t^{(n)}$ となる．

Theorem 31.2. 各 $n \in \mathbb{N}$ に対して，$\{L_t^{(n)} : t \in \mathbb{R}_+\}$ は $\{\mathcal{F}_t\}$ に関するマルチンゲールである．

Proof. Hermite 多項式の定義より，$Z \sim N(0,1)$ に対して，

$$E[H_n(Z)H_m(Z)] = \begin{cases} n! & \text{if } n = m \\ 0 & \text{if } n \neq m \end{cases}$$

となる (演習問題)．いま，$B_t/\sqrt{t} \sim N(0,1)$ であるから，

$$E\left[\sum_n |(u^n/n!)L_t^{(n)}|\right] = \sum_n |u^n/n!| E[L_t^{(n)}] \leq \sum_n |u|^n/\sqrt{n!} < \infty$$

である．よって，$t > s$ に対して，

$$\sum_{n=0}^{\infty} \frac{u^n}{n!} L_s^{(n)} = M_s^u = E[M_t^u \mid \mathcal{F}_s] = \sum_{n=0}^{\infty} \frac{u^n}{n!} E[L_t^{(n)} \mid \mathcal{F}_s] \quad a.s.$$

を得る．除外集合は u に依存するが，$\delta > 0$ を given とすると，最右辺と最左辺は a.s. に $|u| \le \delta$ において一様に絶対収束するので，u の関数として連続である．よって，$P(\Omega_0) = 1$ なる事象 Ω_0 が存在して，各 $\omega \in \Omega_0$ に対して，

$$\sum_{n=0}^{\infty} \frac{u^n}{n!} L_s^{(n)}(\omega) = \sum_{n=0}^{\infty} \frac{u^n}{n!} E[L_t^{(n)} \mid \mathcal{F}_s](\omega), \quad \forall |u| \le \delta$$

が成り立つ．これから，$L_s^{(n)} = E[L_t^{(n)} \mid \mathcal{F}_s]$ a.s. を得る． □

Corollary 31.1 より，$\{B_t : t \in \mathbb{R}_+\}$ と $\{B_t^2 - t : t \in \mathbb{R}_+\}$ はともにマルチンゲールであるが，じつはこの逆も成り立つ．

Theorem 31.3 (Lévy)**.** $\{\mathcal{G}_t : t \in \mathbb{R}_+\}$ を通常の仮定をみたすフィルトレーション，$\{X_t : t \in \mathbb{R}_+\}$ を連続なパスをもつ $\{\mathcal{G}_t\}$ に関するマルチンゲールとし，$\{X_t^2 - t : t \in \mathbb{R}_+\}$ も $\{\mathcal{G}_t\}$ に関するマルチンゲールとする．このとき，$\{X_t - X_0 : t \in \mathbb{R}_+\}$ は標準 BM である．

この定理の証明は確率積分に対する伊藤の公式の典型的な応用例である．証明は，Karatzas and Shreve (1997, Theorem 3.16) を参照せよ．

また，BM のマルチンゲール性より，次の定理を得る．この定理は後述する Skorohod の埋め込み定理のプロトタイプである．

Theorem 31.4. 任意に固定した $a < 0 < b$ に対して，$T = \inf\{t \ge 0 : B_t \in \{a, b\}\}$ とおくと，T は有限な停止時刻であって，

$$P(B_T = a) = \frac{b}{b-a},\ P(B_T = b) = \frac{-a}{b-a}$$

となる．さらに，$E[T] = E[B_T^2] = -ab$ である．

Proof. $\{B_t^2 - t : t \in \mathbb{R}_+\}$ はマルチンゲールであるから，Theorem 30.3 より，$E[B_{T\wedge u}^2] = E[T \wedge u]$. $u \to \infty$ として，MCT より，右辺 $\to E[T]$ を得る．一方，$|B_{T\wedge u}| \le |b-a|$ であるから，DCT より，$E[B_T^2] = \lim_{u\to\infty} E[B_{T\wedge u}^2] \le (b-a)^2$. 従って，$E[T] = E[B_T^2] < \infty$ を得る．特に，$T < \infty$ a.s. であるから，

$$1 = P(B_T = a) + P(B_T = b). \tag{$*$}$$

一方，$\{B_t\}$ はマルチンゲールであるから，再び Theorem 30.3 より，$E[B_{T\wedge u}] = E[B_0] = 0$. ここで，$u \to \infty$ として，DCT より，$E[B_T] = 0$ を得る．従って，

$$0 = aP(B_T = a) + bP(B_T = b). \tag{$**$}$$

$(*)$ と $(**)$ を解いて，$P(B_T = a) = 1 - P(B_T = b) = b/(b-a)$ を得る．また，

$$E[T] = E[B_T^2] = a^2 \frac{b}{b-a} + b^2 \frac{-a}{b-a} = -ab$$

である． □

次に BM の強マルコフ性と，その応用として反射原理を証明する．

Theorem 31.5 (強マルコフ性)**.** 任意の有限な停止時刻 T に対して，$\{B_{T+t}-B_T : t \in \mathbb{R}_+\}$ は $\mathcal{F}_T$ と独立な標準 BM である．

Proof. 証明を 2 つのステップに分ける．

Step 1. $0 < t_1 < \cdots < t_M$ に対して，$(B_{T+t_m} - B_T)_{1\le m\le M}$ は $\mathcal{F}_T$ と独立に分布 $\mathcal{L}(B_{t_1}, \ldots, B_{t_M})$ に従うことを示す．

T_n を T の離散近似とすると，任意の $u_1, \ldots, u_M \in \mathbb{R}$ と $A \in \mathcal{F}_T \subset \mathcal{F}_{T_n}$ に対して，

$$\begin{aligned} &E\left[e^{i\sum_{m=1}^M u_m(B_{T_n+t_m}-B_{T_n})}, A\right] \\ &= \sum_{k=1}^\infty E\left[e^{i\sum_{m=1}^M u_m(B_{T_n+t_m}-B_{T_n})}, A\cap\{T_n = k/2^n\}\right] \\ &= \sum_{k=1}^\infty E\left[e^{i\sum_{m=1}^M u_m(B_{k/2^n+t_m}-B_{k/2^n})}, A\cap\{T_n = k/2^n\}\right] \end{aligned}$$

ここで，$T_n = k/2^n \Leftrightarrow (k-1)/2^n \le T < k/2^n$ であるから，

$$A\cap\{T_n = k/2^n\} = A\cap(\{T < k/2^n\}\setminus\{T < (k-1)/2^n\}) \in \mathcal{F}_{k/2^n}.$$

BM のマルコフ性より，$(B_{k/2^n+t_m} - B_{k/2^n})_{1\le m\le M}$ は $\mathcal{F}_{k/2^n}$ と独立であるから，

$$\begin{aligned} E\left[e^{i\sum_{m=1}^M u_m(B_{T_n+t_m}-B_{T_n})}, A\right] &= E\left[e^{i\sum_{m=1}^M u_m B_{t_m}}\right]\sum_{k=1}^\infty P(A\cap\{T_n = k/2^n\}) \\ &= E\left[e^{i\sum_{m=1}^M u_m B_{t_m}}\right]P(A). \end{aligned}$$

$n\to\infty$ として，

$$E\left[e^{i\sum_{m=1}^M u_m(B_{T+t_m}-B_T)}, A\right] = E\left[e^{i\sum_{m=1}^M u_m B_{t_m}}\right]P(A)$$

を得る．$A \in \mathcal{F}_T$ は任意だったから，

$$E\left[e^{i\sum_{m=1}^M u_m(B_{T+t_m}-B_T)} \mid \mathcal{F}_T\right] = E\left[e^{i\sum_{m=1}^M u_m B_{t_m}}\right] \quad a.s.$$

を得る．これは，

$$\mathcal{F}_T \text{ を与えたときの } (B_{T+t_m} - B_T)_{1\le m\le M} \text{ の r.c.d.} = \mathcal{L}(B_{t_1}, \ldots, B_{t_M})$$

を意味するから，求める主張を得る．

Step 2. Step 1 より，$\{B_{T+t} - B_T : t \in \mathbb{R}_+\}$ の f.d.d.'s は標準 BM のそれと一致するので，$\{B_{T+t} - B_T : t \in \mathbb{R}_+\}$ が標準 BM となることはよい．あとは，$\{B_{T+t} - B_T : t \in \mathbb{R}_+\}$ が $\mathcal{F}_T$ と独立なことを示せばよい．$\mathcal{P}$ を

$$\{(B_{T+t_1} - B_T, \ldots, B_{T+t_M} - B_T) \in A\},\ A \in \mathcal{B}^M, t_1, \ldots, t_M \in \mathbb{R}_+, M \in \mathbb{N}$$

なる形の集合の全体とし，$\mathcal{L} = \{A \in \mathcal{F} : P(A \mid \mathcal{F}_T) = P(A)\ a.s.\}$ とおくと，$\mathcal{P}$ は π システムであって，$\mathcal{L}$ は $\mathcal{P}$ を含む λ システムであるから，π-λ 定理より，$\mathcal{L} \supset \sigma(\mathcal{P}) = \sigma(B_{T+t} - B_T, t \in \mathbb{R}_+)$ を得る． □

Remark 31.2. マルコフ性と同様に，強マルコフ性も次のように言い換えることができる．μ_x を $x + B_\bullet$ の分布とすると，$\mathcal{F}_T$ を与えたときの $B_{T+\bullet}$ の r.c.d. は $(\omega, A) \mapsto \mu_{B_{T(\omega)}(\omega)}(A)$ である．

Theorem 31.6 (反射原理)**.** $t > 0$ を given とし，$M_t = \sup_{0 \le s \le t} B_s$ とおく．このとき，$b > 0$ に対して，$P(M_t \ge b) = 2P(B_t \ge b) = P(|B_t| \ge b)$ となる．

Proof. $T = \inf\{t \ge 0 : B_t = b\}$ とおく．ここで，$\limsup_{t\to\infty} B_t/\sqrt{t} = \infty$ a.s. であるから (演習問題)，$T < \infty$ a.s. である．$X_t = B_{T+t} - B_T = B_{T+t} - b$ とおくと，$\{X_t : t \in \mathbb{R}_+\}$ は $\mathcal{F}_T$ と独立な標準 BM である．特に，$\{X_t\}$ は T と独立である．

ここで，$M_t \ge b \Leftrightarrow T \le t$ より，$P(T \le t) = 2P(B_t \ge b)$ を示せばよい．$\{B_t \ge b\} \subset \{T \le t\}$ より，$\{B_t \ge b\} = \{B_t \ge b, T \le t\} = \{X_{t-T} \ge 0, T \le t\}$ であるから，

$$P(B_t \ge b) = P(B_t \ge b, T \le t) = P(X_{t-T} \ge 0, T \le t)$$

となる．$\varphi(X, T) = 1_{\{X_{t-T} \ge 0, T \le t\}}$ とおくと，X と T の独立性より，$E[\varphi(X,T) \mid T] = E[\varphi(X,s)]|_{s=T}$．$s < t$ のとき，$E[\varphi(X,s)] = P(X_{t-s} \ge 0) = 1/2$．ここで，$\{T = t\} \subset \{B_t = b\}$ より，$P(T = t) \le P(B_t = b) = 0$ であるから，

$$E[\varphi(X,T) \mid T] = 1/2 \quad a.s.\ \text{on}\ \{T \le t\}$$

を得る．以上より，$P(B_t \ge b) = E[E[\varphi(X,T) \mid T], \{T \le t\}] = (1/2)P(T \le t)$ を得る． □

演習問題

Exercise 31.1 (Blumenthal の 0-1 法則)**.** あらゆる $A \in \mathcal{F}_0$ に対して，$P(A) \in \{0, 1\}$ となることを示せ．

Exercise 31.2. $T = \inf\{t \ge 0 : B_t > 0\}$, $T' = \inf\{t > 0 : B_t = 0\}$ とおいたとき，$P(T = 0) = P(T' = 0) = 1$ を示せ．

Exercise 31.3. $\limsup_{t\downarrow 0} B_t/\sqrt{t} = \infty$, $\limsup_{t\to\infty} B_t/\sqrt{t} = \infty$ a.s. を示せ．

Exercise 31.4. $b > 0$ とし，$T = \inf\{t : B_t = b\}$ とおく．このとき，$E[T] = +\infty$ を示せ．

Exercise 31.5. $H_n(x)$ を n 次 Hermite 多項式とし，$Z \sim N(0,1)$ とする．このとき，

$$E[H_n(Z)H_m(Z)] = \begin{cases} n! & \text{if } n = m \\ 0 & \text{if } n \ne m \end{cases}$$

を示せ．

Exercise 31.6. T を可積分な停止時刻とすると，$E[B_T] = 0, E[B_T^2] = E[T]$ を示せ．

32 Skorohod の埋め込み定理とその応用

Skorohod の埋め込み定理

$[0,\infty]$ に値をとる r.v. を 確率時刻 (random time) と呼ぶ．Theorem 31.4 より，2 点分布に従う平均 0 の r.v. X に対して，適当な確率時刻 T が存在して，$B_T \overset{d}{=} X, E[T] = E[X^2]$ となる．なぜなら，$E[X] = 0$ であるから，ある $a < 0 < b$ が存在して，$P(X = a) + P(X = b) = 1, aP(X = a) + bP(X = b) = 0$ となり，これを解いて $P(X = a) = b/(b-a), P(X = b) = -a/(b-a)$ を得る．よって，$T = \inf\{t : B_t \in \{a, b\}\}$ が求める確率時刻である．じつは 2 点分布でなくても，X が平均 0 で分散が有限な r.v. であれば，そのような確率時刻が存在する．

Theorem 32.1 (Skorohod の埋め込み定理)**.** X を $E[X] = 0, E[X^2] < \infty$ をみたす r.v. とすると，次をみたす有限な確率時刻 T が存在する：$X \overset{d}{=} B_T$ & $E[T] = E[X^2]$.

$\{B_t\}$ は可測過程なので，有限な確率時刻 T に対して，$B_{T(\omega)}(\omega)$ は a.s. に定義された r.v. である．

Proof. キーとなる事実は，X の d.f. を F とすると，F が 2 点分布の混合として表せることである．$X = 0$ a.s. なら $T = 0$ とすればよいので，F は非退化と仮定する．ここで，

$$c = \int_{(0,\infty)} v dF(v) = -\int_{(-\infty,0)} u dF(u) = \frac{1}{2}\int_{\mathbb{R}} |x| dF(x) > 0$$

とおくと，有界可測関数 $\varphi : \mathbb{R} \to \mathbb{R}$ に対して，

$$\begin{aligned} c\int_{\mathbb{R}\setminus\{0\}} \varphi dF &= \left\{\int_{(-\infty,0)} \varphi(u) dF(u)\right\}\left\{\int_{(0,\infty)} v dF(v)\right\} \\ &\quad - \left\{\int_{(0,\infty)} \varphi(v) dF(v)\right\}\left\{\int_{(-\infty,0)} u dF(u)\right\} \\ &= \int_{(0,\infty)}\int_{(-\infty,0)} (v-u)\left\{\frac{v\varphi(u)}{v-u} - \frac{u\varphi(v)}{v-u}\right\} dF(u)dF(v). \end{aligned}$$

$u < 0 < v$ に対して，

$$\mu_{u,v} = \frac{v}{v-u}\delta_u + \frac{-u}{v-u}\delta_v$$

とおくと，

$$\int_{\mathbb{R}} \varphi dF = \varphi(0)\{F(0) - F(0-)\} + c^{-1}\int_{(0,\infty)}\int_{(-\infty,0)} \left\{\int_{\mathbb{R}} \varphi(x) d\mu_{u,v}(x)\right\} (v-u) dF(u)dF(v)$$

を得る．そこで，$E := \{(0,0)\} \cup ((-\infty, 0) \times (0,\infty))$ 上の測度 ν を

$$\nu(A) = \{F(0) - F(0-)\}\delta_{(0,0)}(A) + c^{-1}\iint_{A\setminus\{(0,0)\}} (v-u) dF(u)dF(v),\ A \in \mathcal{B}_E^2$$

と定義すると，ν は p.m. であって，$\mu_{0,0} = \delta_0$ とおけば，

$$\int_{\mathbb{R}} \varphi dF = \int_E \left\{ \int_{\mathbb{R}} \varphi(x) d\mu_{u,v}(x) \right\} d\nu(u,v)$$

となる．

いま，$\{B_t\}$ と独立な確率ベクトル $(U,V) \sim \nu$ を発生させる．また，$(u,v) \in E$ に対して，$T_{u,v} := \inf\{t : B_t \in \{u,v\}\}$ とおくと，$E \times \Omega \ni (u,v,\omega) \mapsto T_{u,v}(\omega) \in [0,\infty]$ は可測であって (演習問題)，$B_{T_{u,v}} \sim \mu_{u,v}$．従って，$T(\omega) := T_{U(\omega),V(\omega)}(\omega)$ とおけば，T は有限であり，有界可測関数 $\varphi : \mathbb{R} \to \mathbb{R}$ に対して，

$$E[\varphi(B_T)] = E[E[\varphi(B_{T_{U,V}}) \mid U,V]] = E[\varphi(X)]$$

となる．以上より，$B_T \stackrel{d}{=} X$ を得る．

最後に，

$$\begin{aligned} E[T] &= E[E[T \mid U,V]] = E[-UV] = c^{-1} \int_{(0,\infty)} \int_{(-\infty,0)} (-uv)(v-u) dF(u) dF(v) \\ &= \int_{(-\infty,0)} u^2 dF(u) + \int_{(0,\infty)} v^2 dF(v) = E[X^2] \end{aligned}$$

より，定理の結論を得る． □

Remark 32.1. 上の証明で構成した確率時刻 T は $\{B_t\}$ と独立な確率ベクトル (U,V) に依存しているので，$\{\mathcal{F}_t\}$ に関する停止時刻とはならない．しかし，適当に拡大したフィルトレーションに関して停止時刻となる．以下の応用では，T が停止時刻であることは本質的ではないので，T を単に確率時刻とした．Skorohod の埋め込み定理に関してはいくつかの証明法があり，T を $\{\mathcal{F}_t\}$ に関する停止時刻となるように選ぶこともできる (しかし，紹介したものより複雑な証明を要する)．Skorohod の埋め込み定理に関しては，Obłój (2004) によるサーベイ論文が詳しい．

Corollary 32.1. X を $E[X] = 0, E[X^2] < \infty$ をみたす r.v. とすると，次をみたす有限な確率時刻の非減少列 $0 = T_0 \le T_1 \le T_2 \le \cdots$ が存在する：

(a) $B_{T_n} - B_{T_{n-1}}, n \in \mathbb{N}$ は X の独立なコピーである．

(b) $T_n - T_{n-1}, n \in \mathbb{N}$ は i.i.d. であって，$E[T_n - T_{n-1}] = E[X^2]$.

Proof. $(U_1,V_1),(U_2,V_2),\ldots$ を (U,V) の独立なコピーとし，$\{B_t\}$ と独立とする．このとき，$T_0 = 0$ として，あとは帰納的に，

$$T_n = \inf\{t \ge T_{n-1} : B_t - B_{T_{n-1}} \in \{U_n, V_n\}\}$$

と決めれば，$\{T_n\}$ は (a) と (b) をみたす．詳細は演習問題とする． □

Donsker の不変原理

Skorohod の埋め込み定理を用いて，Donsker の不変原理 (invariance principle) を証明しよう．$\{X_m : m \in \mathbb{N}\}$ を $E[X_m] = 0, E[X_m^2] = 1$ なる i.i.d. r.v.'s とし，$S_0 = 0$, $S_k = \sum_{m=1}^k X_m, k \in \mathbb{N}$ とおいたとき，$t > s \geq 0$ に対して，簡単な計算から，

$$(S_{\lfloor ns \rfloor}/\sqrt{n}, (S_{\lfloor nt \rfloor} - S_{\lfloor ns \rfloor})/\sqrt{n}) \xrightarrow{d} (B_s, B_t - B_s) \sim N\left(\mathbf{0}, \begin{pmatrix} s & 0 \\ 0 & t-s \end{pmatrix}\right),\ n \to \infty$$

となることが示せる．そこで，$n^{-1/2}S_k, k = 0, 1, \ldots, n$ を線形につないで，

$$S^{(n)}(t) = \frac{1}{\sqrt{n}} S_{\lfloor nt \rfloor} + (nt - \lfloor nt \rfloor) X_{\lfloor nt \rfloor + 1},\ t \in [0, 1]$$

なる確率過程を構成すると，$\{S^{(n)}(t) : t \in [0, 1]\}$ は連続なパスをもち，各 $t_1, \ldots, t_k \in [0, 1]$ に対して，$n \to \infty$ のとき，

$$(S^{(n)}(t_1), \ldots, S^{(n)}(t_k)) \xrightarrow{d} (B_{t_1}, \ldots, B_{t_k})$$

となる．すなわち，$S^{(n)}$ の各 f.d.d. は $\{B_t : t \in [0, 1]\}$ のそれに弱収束する．ここまでを示すのは難しくない．Donsker の不変原理は，$S^{(n)}$ が $\{B_t : t \in [0, 1]\}$ に (直後に定義する) C 空間において分布収束することを述べる定理である．C 空間における分布収束は有限次元分布の収束よりもはるかに強い結果である．

以下，必要な定義を与える．

$$C = C([0, 1]) = \{f : f : [0, 1] \to \mathbb{R},\ f \text{ は連続 }\}$$

とおき，C には一様距離

$$d_{\sup}(f, g) = \sup_{t \in [0,1]} |f(t) - g(t)|$$

を入れる．$(C, d_{\sup})$ は Poland 空間である．また，C には $d_{\sup}$ に関する Borel σ-field $\mathcal{B}(C)$ を入れる．じつは $\mathcal{B}(C)$ は C のシリンダー σ-field に一致する．C のシリンダー集合とは，

$$\{f \in C : (f(t_1), \ldots, f(t_m)) \in A\},\ A \in \mathcal{B}^m, t_1, \ldots, t_m \in [0, 1], m \in \mathbb{N}$$

なる形の集合であって，C のシリンダー集合全体の生成する σ-field を C のシリンダー σ-field と呼ぶ．Borel σ-field $\mathcal{B}(C)$ がシリンダー σ-field に一致することの証明は Lemma 29.1 と同様なので省略する．このことから，連続なパスをもつ確率過程 $\{Y_t : t \in [0, 1]\}$ に対して，

$$Y : \omega \mapsto Y(\cdot, \omega),\ \Omega \to C$$

は自動的に C-valued r.v. (i.e., $\mathcal{F}/\mathcal{B}(C)$ 可測) になる．

C-valued r.v.'s の分布収束の定義を復習する．C-valued r.v.'s $Y, Y^{(n)}, n \in \mathbb{N}$ に対して，

$$\lim_{n\to\infty} E[H(Y^{(n)})] = E[H(Y)],\ \forall H \in C_b(C)$$

が成り立つとき，$Y^{(n)}$ は Y に分布収束すると言って，$Y^{(n)} \xrightarrow{d} Y$ と書く．定義から明らかなように，$Y^{(n)} \xrightarrow{d} Y$ なら，任意の連続写像 $\psi : C \to \mathbb{R}^k$ に対して，$\psi(Y^{(n)}) \xrightarrow{d} \psi(Y)$ in $\mathbb{R}^k$ となる．例えば，

$$\psi_1(f) = \sup_{t\in[0,1]} |f(t)|,\ \psi_2(t) = \int_0^1 f(t)dt$$

などは C から $\mathbb{R}$ への連続写像である．

Example 32.1. $t_1, \ldots, t_k \in [0,1]$ に対して，射影 $\pi_{t_1,\ldots,t_k} : C \to \mathbb{R}^k$ を $\pi_{t_1,\ldots,t_k}(f) = (f(t_1), \ldots, f(t_k))$ と定義すると，$\pi_{t_1,\ldots,t_k}$ は連続なので，$Y^{(n)} \xrightarrow{d} Y$ in C なら，

$$(Y^{(n)}(t_1), \ldots, Y^{(n)}(t_k)) = \pi_{t_1,\ldots,t_k}(Y^{(n)}) \xrightarrow{d} \pi_{t_1,\ldots,t_k}(Y) = (Y(t_1), \ldots, Y(t_k)) \text{ in } \mathbb{R}^k$$

となる．従って，$Y^{(n)}$ の各 f.d.d. は Y のそれに弱収束する．ただし，逆は成り立たない．例えば，

$$f_n(t) = nt1_{[0,n^{-1}]}(t) + \{1 - n(t - n^{-1})\}1_{[n^{-1},2n^{-1}]}(t),\ t \in [0,1]$$

とすると，$f_n(t) \to 0$ $(n \to \infty, \forall t \in [0,1])$ であるが，$d_{\sup}(f_n, 0) = 1$ である．$Y^{(n)} \equiv f_n$ とすると，$(Y^{(n)}(t_1), \ldots, Y^{(n)}(t_k)) \xrightarrow{d} (0, \ldots, 0)$ であるが，C において $Y^{(n)}$ は 0 に分布収束しない．すなわち，C において分布収束を示すためには，f.d.d.'s の弱収束を示すだけは不十分である．

B の $[0,1]$ への制限を $B_{[0,1]}$ と書く：$B_{[0,1]}(t) = B(t), t \in [0,1]$.

Theorem 32.2 (Donsker の不変原理)**.** $n \to \infty$ のとき，$S^{(n)} \xrightarrow{d} B_{[0,1]}$ in C.

Proof. $T_k, k \in \mathbb{Z}_+$ を $X = X_1$ に対して Corollary 32.1 を適用して構成した確率時刻の非減少列とすると，

$$(X_1, \ldots, X_n) \overset{d}{=} (B_{T_1} - B_{T_0}, \ldots, B_{T_n} - B_{T_{n-1}})$$

となる．$S^{(n)}(t)$ の定義において，$X_i, 1 \le i \le n$ の代わりに $B_{T_i} - B_{T_{i-1}}, 1 \le i \le n$ を用いたものを $\tilde{S}^{(n)}(t)$ とおく．このとき，

$$\sup_{t\in[0,1]} |B(nt)/\sqrt{n} - \tilde{S}^{(n)}(t)| \xrightarrow{P} 0 \qquad (*)$$

を示す．ここで，$Y^{(n)}(t) = B(nt)/\sqrt{n}$ とおくと，$Y^{(n)}$ も標準 BM である．$\varepsilon > 0$ を任意に固定する．パスの連続性より，$P(\sup_{t\in[0,1)} |Y^{(n)}(t) - \tilde{S}^{(n)}(t)| > \varepsilon) \to 0$ を示せばよい．$A_n = \{\sup_{t\in[0,1)} |Y^{(n)}(t) - \tilde{S}^{(n)}(t)| > \varepsilon\}$ とおく．$t \in [0,1)$ に対して，$k(t)$ を $(k-1)/n \le$

$t < k/n$ をみたす一意な k と定義する．このとき，$\tilde{S}^{(n)}(t)$ は $B(T_k)/\sqrt{n} = Y^{(n)}(T_k/n)$ を線形につなげたものであるから，

$$\begin{aligned} A_n &\subset \left\{ \sup_{t\in[0,1)} |Y^{(n)}(t) - Y^{(n)}(T_{k(t)}/n)| > \varepsilon \right\} \\ &\qquad \bigcup \left\{ \sup_{t\in[0,1)} |Y^{(n)}(t) - Y^{(n)}(T_{k(t)-1}/n)| > \varepsilon \right\} \\ &=: A_n^*. \end{aligned}$$

任意の $0 < \delta < 1$ に対して，

$$\begin{aligned} A_n^* &\subset \left\{ \exists s,t \in [0,2] \text{ s.t. } |s-t| < \delta \ \& \ |Y^{(n)}(s) - Y^{(n)}(t)| > \varepsilon \right\} \\ &\qquad \bigcup \left\{ \exists t \in [0,1) \text{ s.t. } |T_{k(t)}/n - t| \vee |T_{k(t)-1}/n - t| \geq \delta \right\} \\ &=: A_{n1}^* \cup A_{n2}^*. \end{aligned}$$

ここで，$Y^{(n)}$ は標準 BM であるから，

$$P(A_{n1}^*) = P(\exists s.t \in [0,2] \text{ s.t. } |s-t| < \delta \text{ and } |B(s) - B(t)| > \varepsilon).$$

$B(t)$ は $[0,2]$ 上で一様連続なパスをもつから，$\delta = 1/m \downarrow 0, m \to \infty$ のとき，右辺 $\to 0$. 従って，任意に固定した $0 < \delta < 1$ に対して，$P(A_{n2}^*) \to 0, n \to \infty$ を示せばよい．まず，$T_k - T_{k-1}$ は i.i.d. で，$E[T_k - T_{k-1}] = E[X_k^2] = 1$ であるから，SLLN より，

$$\frac{T_n}{n} = \frac{1}{n}\sum_{k=1}^{n}(T_k - T_{k-1}) \to 1 \quad a.s.$$

次の補題を用いる．証明は演習問題とする．

Lemma 32.1. 実数列 $\{a_n\}$ に対して，

$$\lim_{n\to\infty} \frac{a_n}{n} = 1 \Rightarrow \lim_{n\to\infty} \max_{1\leq k\leq n} \frac{|a_k - k|}{n} = 0.$$

この補題より，

$$\max_{1\leq k\leq n} \frac{|T_k - k|}{n} \to 0 \quad a.s.$$

を得る．$k(t)$ は定義より，$(k(t)-1)/n \leq t < k(t)/n$ をみたすから，

$$\begin{aligned} &|T_{k(t)}/n - t| \leq |T_{k(t)} - k(t)|/n + 1/n \leq \max_{1\leq k\leq n} |T_k - k|/n + 1/n, \\ &|T_{k(t)-1}/n - t| \leq \max_{1\leq k\leq n} |T_{k-1} - (k-1)|/n + 1/n \end{aligned}$$

となる．従って，$P(A^*_{n2}) \to 0$ を得る．以上より，$(*)$ が示された．

あとは，$H \in C_b(C), \varepsilon, \delta > 0$ に対して，

$$F_{\varepsilon,\delta} = \{f \in C : \forall g, h \in C, d_{\sup}(f,g) < \delta, d_{\sup}(f,h) < \delta \Rightarrow |H(g) - H(h)| < \varepsilon\}$$

とおくと，$F_{\varepsilon,\delta}$ は C の閉集合であって，$S^{(n)} \stackrel{d}{=} \tilde{S}^{(n)}, Y^{(n)}_{[0,1]} \stackrel{d}{=} B_{[0,1]}$ より，

$$\begin{aligned}
&|E[H(S^{(n)})] - E[H(B_{[0,1]})]| = |E[H(\tilde{S}^{(n)})] - E[H(Y^{(n)}_{[0,1]})]| \\
&\leq \varepsilon + 2\|H\|_u \{P(B_{[0,1]} \in F^c_{\varepsilon,\delta}) + P(\sup_{t\in[0,1]} |\tilde{S}^{(n)}(t) - Y^{(n)}(t)| \geq \delta)\}.
\end{aligned}$$

$n \to \infty$ として，$\limsup_n |E[H(S^{(n)})] - E[H(B_{[0,1]})]| \leq \varepsilon + 2\|H\|_u P(B_{[0,1]} \in F^c_{\varepsilon,\delta})$. $\delta = 1/m \downarrow 0, m \to \infty$ のとき，$F_{\varepsilon,\delta} \uparrow C$ となるから，$P(B_{[0,1]} \in F^c_{\varepsilon,\delta}) \downarrow 0$. $\varepsilon > 0$ は任意だったから，$\lim_n E[H(S^{(n)})] = E[H(B_{[0,1]})]$ を得る．従って，定理の結論を得る． □

Remark 32.2. 本節で紹介した Skorohod の埋め込み定理にもとづく Donsker の不変原理の証明は，ややアドホックではある．他の証明法として，有限次元分布の収束と一様タイト性を示す方法があり，この証明法は一般性がある．詳細は，次のパートを参照せよ．

重複対数の法則

本節の目標は，Hartman-Winter の重複対数の法則の証明である．そのステートメントを思い出す．記号の簡単のために，$\log_{(2)} x = \log\log x$ とおく．

Theorem 32.3 (Hartman-Winter)**.** $X_1, X_2, \ldots$ を i.i.d. r.v.'s とし，$E[X_1] = 0, E[X_1^2] = 1$ とする．また，$S_n = \sum_{m=1}^n X_m$ とおく．このとき，a.s. に

$$\limsup_n \frac{S_n}{\sqrt{2n\log_{(2)} n}} = 1, \quad \liminf_n \frac{S_n}{\sqrt{2n\log_{(2)} n}} = -1$$

が成り立つ．

明らかにどちらか一方を示せば，もう一方も従う．ここでは，前者を証明する．まず，標準 BM に対して重複対数の法則を証明する．

Theorem 32.4 (標準 BM に対する重複対数の法則)**.** a.s. に

$$\limsup_{t\to\infty} \frac{B_t}{\sqrt{2t\log_{(2)} t}} = 1$$

が成り立つ．

Proof. 時間反転 (cf. Exercise 29.8) より，

$$\limsup_{t\downarrow 0} \frac{B_t}{\sqrt{2t\log_{(2)}(1/t)}} = 1 \quad a.s.$$

を示せばよい．$h(t) = \sqrt{2t\log_{(2)}(1/t)}$ とおき，$\theta, \delta \in (0,1)$ を given とする．さらに，

$$\alpha_n = (1+\delta)\theta^{-n}h(\theta^n),\ \beta_n = h(\theta^n)/2$$

とおく．このとき，

$$P\left\{\sup_{0\le s\le 1}(B_s - \alpha_n s/2) \ge \beta_n\right\} = P\left\{\sup_{0\le s\le 1}\exp(\alpha B_s - \alpha_n^2 s/2) \ge e^{\alpha_n\beta_n}\right\}$$

であって，$\{e^{\alpha_n B_s - \alpha_n^2 s/2} : s \in \mathbb{R}_+\}$ はマルチンゲールであるから，Doob の不等式より，右辺は

$$e^{-\alpha_n\beta_n} = O(1)n^{-1-\delta}$$

でバウンドできる．従って，Borel-Cantelli の補題より，

$$\Omega_0 = \left\{\sup_{0\le s\le 1}(B_s - \alpha_n s/2) < \beta_n \text{ ev.}\right\}$$

なる事象は確率 1 をもつ．このとき，各 $\omega \in \Omega_0$ に対して，ある $n_0(\omega) \in \mathbb{N}$ が存在して，$n > n_0(\omega)$ のとき，あらゆる $s \in [0, \theta^{n-1})$ に対して，

$$B_s(\omega) \le \frac{\alpha_n s}{2} + \beta_n \le \frac{\alpha_n\theta^{n-1}}{2} + \beta_n = \left(\frac{1+\delta}{2\theta} + \frac{1}{2}\right)h(\theta^n)$$

が成り立つ．h は原点の近傍で非減少なので，十分大きな n に対して，

$$B_s(\omega) \le \left(\frac{1+\delta}{2\theta} + \frac{1}{2}\right)h(s),\ \forall s \in [\theta^n, \theta^{n-1})$$

となる．ゆえに，

$$\limsup_{s\downarrow 0}\frac{B_s}{h(s)} \le \frac{1+\delta}{2\theta} + \frac{1}{2} \quad a.s.$$

を得る．$\theta, \delta \in (0,1)$ は任意であったから，$\limsup_{s\downarrow 0} B_s/h(s) \le 1$ a.s. を得る．

次に，$\theta \in (0,1)$ を given とし，

$$A_n = \left\{B_{\theta^n} - B_{\theta^{n+1}} \ge (1-\sqrt{\theta})h(\theta^n)\right\}$$

とおくと，$A_n, n \in \mathbb{N}$ は独立であって，$B_{\theta^n} - B_{\theta^{n+1}} \sim N(0, \theta^n - \theta^{n+1})$ であるから，$Z \sim N(0,1)$ とすると，

$$P(A_n) = P\left\{Z \ge \frac{(1-\sqrt{\theta})h(\theta^n)}{\sqrt{\theta^n(1-\theta)}}\right\}$$

となる．ここで，

$$\frac{(1-\sqrt{\theta})h(\theta^n)}{\sqrt{\theta^n(1-\theta)}} = \frac{(1-\sqrt{\theta})}{\sqrt{1-\theta}}\sqrt{2\log_{(2)}\theta^{-n}} = x_n$$

とおくと，$x_n \to \infty$ であって，

$$P(A_n) \sim x_n^{-1}\phi(x_n)$$

となる．ただし，$\phi(x)$ は $N(0,1)$ の密度関数である．このとき，$\sum_n x_n^{-1}\phi(x_n) = +\infty$ であるから，Borel-Cantelli の補題より，

$$\Omega_1 = \left\{B_{\theta_n} \geq (1-\sqrt{\theta})h(\theta^n) + B_{\theta^{n+1}} \text{ i.o.}\right\}$$

なる事象は確率 1 をもつ．また，前半の議論より，

$$\Omega_2 = \left\{-B_{\theta^{n+1}} < 2h(\theta^{n+1}) \text{ ev.}\right\}$$

なる事象は確率 1 をもつ．ここで，十分大きな n に対して，$h(\theta^{n+1}) \leq 2\sqrt{\theta}h(\theta^n)$ となるから，各 $\omega \in \Omega_1 \cap \Omega_2$ に対して，

$$B_{\theta^n}(\omega) \geq h(\theta^n)(1-5\sqrt{\theta})$$

となる n は無限個ある．従って，

$$\limsup_{t\downarrow 0} B_t/h(t) \geq \limsup_n B_{\theta^n}/h(\theta^n) \geq 1-5\sqrt{\theta} \quad a.s.$$

を得る． □

Proof of Theorem 32.3. $T_k, k \in \mathbb{Z}_+$ を $X = X_1$ に対して Corollary 32.1 を適用して構成した確率時刻の非減少列とすると，$S_n \stackrel{d}{=} B_{T_n}$ である．従って，

$$\limsup_{t\to\infty} \frac{|B_t - B_{T_{\lfloor t \rfloor}}|}{\sqrt{t\log_{(2)} t}} = 0 \quad a.s.$$

を示せば十分である．いま，$T_n/n \to 1$ a.s. であって，$\varepsilon > 0$ を given とすると，各 $\omega \in \{T_n/n \to 1\}$ に対して，$t_0 = t_0(\omega)$ を十分大きく選べば，

$$t \geq t_0 \Rightarrow \frac{1}{1+\varepsilon} \leq \frac{T_{\lfloor t \rfloor}}{t} \leq 1+\varepsilon$$

となる．そこで，$M_t = \sup_{t/(1+\varepsilon)\leq s\leq t(1+\varepsilon)} |B_s - B_t|$ とおいたとき，

$$\limsup_{t\to\infty} \frac{M_t}{\sqrt{t\log_{(2)} t}} = 0 \quad a.s.$$

を示せばよい．いま，$t_n = (1+\varepsilon)^n, n \in \mathbb{N}$, $M'_n = \sup_{t_{n-1} \le s \le t_{n+2}} |B_s - B_{t_{n-1}}|$ とおくと，$t \in [t_n, t_{n+1}]$ に対して，$M_t \le 2M'_n$ となる．$\delta = (1+\varepsilon)^3 - 1$ とおくと，$t_{n+2} - t_{n-1} = \delta t_{n-1}$ であって，$M'_n \stackrel{d}{=} \sqrt{\delta t_{n-1}} \sup_{0 \le t \le 1} |B_t|$ であるから，反射原理より，

$$
\begin{aligned}
&P\left\{M'_n > \sqrt{3\delta t_{n-1} \log_{(2)} t_{n-1}}\right\} = P\left\{\sup_{0 \le t \le 1} |B_t| > \sqrt{3 \log_{(2)} t_{n-1}}\right\} \\
&\quad \le 2P\left\{\sup_{0 \le t \le 1} B_t > \sqrt{3 \log_{(2)} t_{n-1}}\right\} = 4P\left\{B_1 > \sqrt{3 \log_{(2)} t_{n-1}}\right\} = O(n^{-3/2})
\end{aligned}
$$

を得る．よって，Borel-Cantelli の補題より，

$$
\limsup_{t \to \infty} \frac{M_t}{\sqrt{t \log_{(2)} t}} \le \limsup_n \frac{2M'_n}{\sqrt{t_{n-1} \log_{(2)} t_{n-1}}} \le 2\sqrt{3}\delta \quad a.s.
$$

を得る． □

演習問題

Exercise 32.1. Skorohod の埋め込み定理において，$B_T \stackrel{d}{=} X$ という条件をみたす確率時刻を構成するだけなら，次のようなトリビアルな例がある．Φ を $N(0,1)$ の d.f., F を X の d.f. とし，$H = F^{-1} \circ \Phi$ とおくと，$H(B_1) \sim F$ である．$T = \inf\{t \ge 2 : B_t = H(B_1)\}$ とおく．

(a) $T < \infty$ a.s. を示せ．

(b) (a) より，$B_T = H(B_1)$ a.s. となるから，$B_T \sim F$ である．しかし，$E[T] = +\infty$ となることを示せ．

Exercise 32.2. Theorem 32.1 の証明において，$E \times \Omega \ni (u, v, \omega) \mapsto T_{u,v}(\omega) \in [0, \infty]$ の可測性を示せ．

Exercise 32.3. Corollary 32.1 の証明の詳細を埋めよ．

Exercise 32.4. Lemma 32.1 を示せ．

Exercise 32.5. $\{X_m\}$ を i.i.d. r.v.'s とし，$E[X_m] = 0, E[X_m^2] = 1$ とする．また，$S_n = \sum_{m=1}^n X_m$ とおく．ただし，$S_0 = 0$ とする．

(a) $\ell = 1, 2, \ldots$ に対して，$n^{-(1+\ell/2)} \sum_{m=1}^n S_m^\ell \stackrel{d}{\to} \int_0^1 B_t^\ell dt$ $(n \to \infty)$ を示せ．

(b) $n^{-1} \sum_{m=1}^n S_{m-1}(S_m - S_{m-1}) \stackrel{d}{\to} (1/2)(B_1^2 - 1)$ $(n \to \infty)$ を示せ．

Exercise 32.6. 次のような時系列回帰モデルを考える:$Y_t = \alpha Y_{t-1} + U_t$ $(t = 1, 2, \dots)$, $Y_0 = 0$. ただし，$\{U_t\}$ は i.i.d. であって，$E[U_t] = 0, E[U_t^2] = \sigma^2 \in (0, \infty)$ とする．$Y_0, Y_1, \dots, Y_n$ が与えられたとき，α の最小二乗推定量は

$$\hat{\alpha} = \frac{\sum_{t=1}^n Y_t Y_{t-1}}{\sum_{t=1}^n Y_{t-1}^2}$$

である．$\alpha = 1$ のとき，$n(\hat{\alpha} - 1)$ の極限分布 $(n \to \infty)$ を求めよ．

Exercise 32.7. Exercise 32.5 において，$P(X_m = \pm 1) = 1/2$ とし，

$$R_n = \mathrm{Card}(\{S_0, S_1, \dots, S_n\})$$

とおくと，$n \to \infty$ のとき，$n^{-1/2} R_n \xrightarrow{d} \sup_{t \in [0,1]} B_t - \inf_{t \in [0,1]} B_t$ を示せ．

Exercise 32.8 (Heyde and Brown (1970))**.** 有限な確率時刻 T と $0 < \varepsilon < 1$ に対して，

$$\sup_{x \in \mathbb{R}} |P(B_T \le x) - \Phi(x)| \le A\varepsilon^{1/2} + P(|T - 1| > \varepsilon)$$

を示せ．ただし，A は絶対定数であって，Φ は $N(0, 1)$ の d.f. である．

ヒント：

$$P(B_T \le x) \le P(B_T \le x, |T - 1| \le \varepsilon) + P(|T - 1| > \varepsilon)$$

と分割すると，第 1 項 $\le P\{B_{1-\varepsilon} + \inf_{|t-1| \le \varepsilon}(B_t - B_{1-\varepsilon}) \le x\}$ となる．さらに，$B_{1-\varepsilon}$ と $\{B_t - B_{1-\varepsilon} : |t - 1| \le \varepsilon\}$ は独立であって，$\inf_{|t-1| \le \varepsilon} |B_t - B_{1-\varepsilon}| \overset{d}{=} \inf_{|t| \le 2\varepsilon} B_t$ である．

Part VII. 距離空間上の確率測度の弱収束

本パートは距離空間上の確率測度の弱収束に関する基本的な結果と，C 空間および D 空間 (有界区間上の càdlàg な関数からなる空間) における弱収束を考察する．本パートの内容はおもに Billingsley (1968) の Chapter 1 から Chapter 3 にもとづいている．その他に関連する文献として，Dudley (2002, Chapter 9) と Stroock (2011, Chapter 9) を挙げておく．

33　距離空間上の確率測度の弱収束

弱収束の定義を復習する．(S,d) を距離空間とし，$\mathcal{B}(S)$ を S の Borel σ-field とする．$(S,\mathcal{B}(S))$ 上の p.m.'s の全体を $\mathcal{P}(S)$ と書く．また，

$$C(S)=\{f\in\mathbb{R}^S : f \text{ は連続}\},\ C_b(S)=\{f\in\mathbb{R}^S : f \text{ は有界連続}\}$$

とおく．$\mu,\mu_n\in\mathcal{P}(S)$ に対して，

$$\lim_n \int f d\mu_n = \int f d\mu,\ \forall f\in C_b(S)$$

となるとき，μ_n は μ に弱収束すると言って，$\mu_n \xrightarrow{w} \mu$ と書くのであった．すでに示したように，弱収束極限は存在すれば一意である．

次の定理は弱収束の同値な条件を与える．

$$U_d(S)=\{f\in\mathbb{R}^S : f \text{ は有界かつ } d \text{ に関して一様連続}\}.$$

とおく． 以下，$A\subset S$ に対して，$A^\circ,\overline{A},\partial A$ をそれぞれ A の内部，閉包，境界とする．

Theorem 33.1 (Portmanteau)**.** $\mu,\mu_n\in\mathcal{P}(S)$ に対して，次の (i)–(iv) は同値である．

(i) $\mu_n \xrightarrow{w} \mu$.

(ii) $\int f d\mu_n \to \int f d\mu,\ \forall f\in U_d(S)$.

(iii) $\forall F\subset S$: closed, $\limsup_n \mu_n(F)\le \mu(F)$.

(iv) $\forall G\subset S$: open, $\liminf_n \mu_n(G)\ge \mu(G)$.

(v) $\forall A\in\mathcal{B}(S)$ s.t. $\mu(\partial A)=0$, $\lim_n \mu_n(A)=\mu(A)$.

Proof. (i) $\Rightarrow$ (ii) $\Rightarrow$ (iii) $\Leftrightarrow$ (iv), (iii)+(iv) $\Rightarrow$ (v) $\Rightarrow$ (i) を示す．

(i) $\Rightarrow$ (ii) は明らか．(iii) $\Leftrightarrow$ (iv) は互いに補集合をとればよい．(iii)+(iv) $\Rightarrow$ (v) は，

$$\limsup_n \mu_n(A)\le \limsup_n \mu_n(\overline{A})\le \mu(\overline{A})=\mu(A^\circ)\le \liminf_n \mu_n(A^\circ)\le \liminf_n \mu_n(A)$$

より従う．

(ii) ⇒ (iii). $F \subset S$ を閉集合とする．

$$\eta(t) = \begin{cases} 1 & \text{if } t < 0 \\ 1 - t & \text{if } 0 \le t \le 1 \\ 0 & \text{if } t > 1 \end{cases}$$

とおいて，$\varepsilon > 0$ に対して，$f_\varepsilon(x) = \eta(\varepsilon^{-1} d(x, F))$ とおくと，$f_\varepsilon \in U_d(S)$ である．ここで，$F^\varepsilon = \{x \in S : d(x, F) \le \varepsilon\}$ とおくと，$1_F \le f_\varepsilon \le 1_{F^\varepsilon}$ であるから，

$$\limsup_n \mu_n(F) \le \lim_n \int f_\varepsilon d\mu_n = \int f_\varepsilon d\mu \le \mu(F^\varepsilon).$$

F は閉集合であるから，$\varepsilon = 1/m \downarrow 0$ とすると，右辺 $\downarrow \mu(F)$.

(v) ⇒ (i). $f \in C_b(S)$ を任意にとる．$\int f d\mu_n \to \int f d\mu$ を示せばよいが，一般性を失うことなく，$0 \le f \le 1$ と仮定してよい．このとき，

$$\int f d\mu = \int_0^\infty \mu(f > t) dt = \int_0^1 \mu(f > t) dt, \ \int f d\mu_n = \int_0^1 \mu_n(f > t) dt.$$

ここで，$\partial\{f > t\} \subset \{f = t\}$ であって，$\mu(f = t) > 0$ となる t は高々可算個しかないから (単調関数の不連続点は高々可算個しかない！)，(v) の仮定より，高々可算個の $t \in [0, 1]$ を除いて，$\mu_n(f > t) \to \mu(f > t)$ となる．あとは DCT より，$\int f d\mu_n \to \int f d\mu$ を得る． □

Remark 33.1. (i) ⇔ (ii) より，$\int f d\mu = \int f d\nu$ $(\forall f \in U_d(S))$ なら，$\mu = \nu$ である．

次に S が可分なら，弱収束は距離化可能であることを示そう．

Theorem 33.2. (S, d) が可分なら，次をみたす $\mathcal{P}(S)$ 上の距離 Δ が存在する：任意の $\mu, \mu_n \in \mathcal{P}(S)$ に対して，$\mu_n \xrightarrow{w} \mu \Leftrightarrow \Delta(\mu_n, \mu) \to 0$. このとき，$\Delta$ の距離位相を 弱位相 (weak topology) と呼ぶ．

まず，いくつかの準備をする．$x \in S$ を中心とする半径 $r > 0$ の開球を

$$B(x, r) = \{y \in S : d(y, x) < r\}$$

と書く．$A \subset S$ が コンパクト (compact) であるとは，$A \subset \bigcup_{i \in I} G_i$ なる任意の開集合の族 $G_i, i \in I$ に対して，有限個の $i_1, \ldots, i_n \in I$ が存在して，$A \subset \bigcup_{j=1}^n G_{i_j}$ となることを言う．$A \subset S$ が 全有界 (totally bounded) であるとは，任意の $\varepsilon > 0$ に対して，ある $x_1, \ldots, x_n \in S$ が存在して，$A \subset \bigcup_{j=1}^n B(x_j, \varepsilon)$ となることを言う．

次の Theorem 33.3 と Theorem 33.4 は位相空間論における基本的な定理なので，証明は省略する．例えば，Dudley (2002, Chapter 2) を参照せよ．

Theorem 33.3. $A \subset S$ に対して，次の (i)–(iii) は同値である．

(i) A はコンパクト.

(ii) A は点列コンパクト．すなわち，A 内の任意の点列は A の点に収束する部分列をもつ.

(iii) A は全有界かつ完備.

全有界な距離空間は可分であるから，コンパクトな距離空間は Poland 空間である.

Corollary 33.1. $A \subset S$ に対して，次の (i)–(iii) は同値である.

(i) A は相対コンパクト．すなわち，$\overline{A}$ はコンパクトである.

(ii) A 内の任意の点列は収束部分列をもつ.

(iii) A は全有界であって，$\overline{A}$ は完備である.

距離空間 $(S,d),(T,e)$ に対して，写像 $\varphi : S \to T$ が 等長 (isometric) であるとは，

$$e(\varphi(x),\varphi(y)) = d(x,y),\ \forall x,y \in S$$

となることを言う．等長写像は必ず単射である.

Theorem 33.4 (距離空間の完備化)**.** 距離空間 (S,d) に対して，完備な距離空間 $(T,\bar{d})$ と等長写像 $\varphi : S \to T$ が存在して，$\varphi(S)$ は T で稠密となる.

$(T,\bar{d})$ を (S,d) の 完備化 (completion) と呼ぶ．(S,d) と $(\varphi(S),\bar{d})$ は距離空間として全く同じ構造をもつから，S と $\varphi(S)$ を同一視して，S を T の部分集合とみなす.

Theorem 33.2 の証明に移る．まず，次の補題を証明する.

Lemma 33.1. (i) (S,d) が可分なら，d と同値な距離 $\hat{d}$ が存在して，$(S,\hat{d})$ は全有界になる [18]．(ii) $U_{\hat{d}}(S)$ は一様距離に関して可分である.

Proof. (i). $\{x_m\}$ を (S,d) の稠密な部分集合とし，$[0,1]^{\mathbb{N}}$ に直積位相を入れたとき，

$$h(x) = \left(\frac{d(x,x_m)}{1+d(x,x_m)}\right)_{m\in\mathbb{N}},\ x \in S \qquad (*)$$

は S から $h(S) =: E$ への同相写像である．$[0,1]^{\mathbb{N}}$ の直積位相は距離

$$e((a_m),(b_m)) = \sum_{m=1}^{\infty} \frac{|a_m - b_m|}{2^m},\ (a_m),(b_m) \in [0,1]^{\mathbb{N}}$$

によって距離化可能である．Tychonoff の定理より，$[0,1]^{\mathbb{N}}$ はコンパクトであるから (直接証明することも難しくない)，E は e に関して全有界である．従って，

$$\hat{d}(x,y) := e(h(x),h(y)),\ x,y \in S$$

[18] $\hat{d}$ が d と同値であるというのは，$\hat{d}$ が d と同じ位相を与えるということである．$\hat{d}$ と d が同値であるためには，任意の $x, x_n \in S$ に対して，$d(x_n,x) \to 0 \Leftrightarrow \hat{d}(x_n,x) \to 0$ となることが必要十分である.

が求める距離である．

(ii)．まず，S がコンパクトなら，$C(S)$ が一様距離に関して可分であることを示す．S と $E \subset [0,1]^{\mathbb{N}}$ は同相なので，$C(E)$ が一様距離に関して可分であることを示せばよい．$\mathcal{A} \subset C(E)$ を多項式全体，すなわち，単項式

$$(y_1, y_2, \dots) \mapsto y_1^{\alpha_1} y_2^{\alpha_2} \cdots y_n^{\alpha_n},\ \alpha_1, \dots, \alpha_n \in \mathbb{Z}_+, n \in \mathbb{N}$$

の有限 1 次結合の全体とすると，$\mathcal{A}$ は Stone-Weierstrass の定理の条件をすべてみたすので，$\mathcal{A}$ は $C(E)$ で稠密である．いま，$\mathcal{A}_0$ を $\mathcal{A}$ の多項式のうち有理係数をもつもの全体とすると，$\mathcal{A}_0$ は可算であって，$C(E)$ で稠密である．よって，$C(S)$ は可分である．

(ii) の証明に戻ると，$(S, \hat{d})$ の完備化を T とすると，$(S, \hat{d})$ が全有界なことから，T はコンパクトである．このとき，$C(T)$ は一様距離に関して可分であって，

$$C(T) \ni f \mapsto f|_S \in U_{\hat{d}}(S)$$

は等長同型 (等長かつ全射) であるから，$U_{\hat{d}}(S)$ も可分である． □

Proof of Theorem 33.2. $\{f_m\}$ を $U_{\hat{d}}(S)$ の一様距離に関する稠密な部分集合とし，

$$\Delta(\mu, \nu) := \sum_{m=1}^{\infty} \frac{1}{2^m} \left\{ 1 \wedge \left| \int f_m d\mu - \int f_m d\nu \right| \right\},\ \mu, \nu \in \mathcal{P}(S)$$

と定義する．このとき，Δ が対称性と三角不等式をみたすのは明らかである．さらに，Δ の定義と Portmanteau の定理より，

$$\begin{aligned} \lim_n \Delta(\mu_n, \mu) = 0 &\Leftrightarrow \lim_n \int f_m d\mu_n = \int f d\mu\ (\forall m) \\ &\Leftrightarrow \lim_n \int f d\mu_n = \int f d\mu\ (\forall f \in U_{\hat{d}}(S)) \\ &\Leftrightarrow \mu_n \xrightarrow{w} \mu \end{aligned}$$

であり，これから Δ が距離であることも従う． □

Theorem 33.5 (Varadarajan の定理)**.** (S, d) を可分とし，$\mu \in \mathcal{P}(S)$ とする．$X_1, X_2, \cdots \sim \mu$ i.i.d. に対して，経験測度を

$$\mu_n(A, \omega) := \frac{1}{n} \sum_{i=1}^{n} 1_A(X_i(\omega)),\ A \in \mathcal{B}(S), \omega \in \Omega$$

と定義する．このとき，$P(\{\omega \in \Omega : \mu_n(\cdot, \omega) \xrightarrow{w} \mu\}) = 1$ である．

Proof. $f \in U_{\hat{d}}(S)$ に対して，SLLN より，$n \to \infty$ のとき，$\int f d\mu_n = n^{-1} \sum_{i=1}^n f(X_i) \to \int f d\mu$ a.s. ゆえに，$\Omega_0 = \bigcap_m \{\lim_n \int f_m d\mu_n = \int f_m d\mu\}$ とおくと，$P(\Omega_0) = 1$ であって，$\omega \in \Omega_0$ に対して，$\lim_n \Delta(\mu_n(\cdot, \omega), \mu) = 0$ となる． □

Varadarajan の定理より，$\mathcal{P}(S)$ の可分性が従う．

Theorem 33.6. (S,d) が可分なら，$\mathcal{P}(S)$ は弱位相に関して可分である．

Proof. $\{x_m\}$ を S の稠密な部分集合とすると，

$$\left\{\sum_{i=1}^{n}\alpha_i\delta_{x_i} : \alpha_1,\ldots,\alpha_n \in \mathbb{Q}_+, \sum_{i=1}^{n}\alpha_i = 1, n\in\mathbb{N}\right\}$$

は $\mathcal{P}(S)$ で稠密である． □

分布収束

距離空間 S に対して，$\mathcal{F}/\mathcal{B}(S)$ 可測写像のことを S-valued r.v. と呼ぶ．S-valued r.v.'s $X, X_n, n\in\mathbb{N}$ に対して，$\mathcal{L}(X_n)\xrightarrow{w}\mathcal{L}(X)$ となるとき，X_n は X にすると言って，$X_n\xrightarrow{d}X$ と書くのであった．

$$X_n\xrightarrow{d}X \Leftrightarrow \lim_n E[f(X_n)] = E[f(X)]\ (\forall f\in C_b(S))$$

である．Portmanteau 定理を言い換えると，

$$\begin{aligned}
&\text{(i) } X_n\xrightarrow{d}X\\
\Leftrightarrow\ &\text{(ii) } \lim_n E[f(X_n)] = E[f(X)],\ \forall f\in U_d(S)\\
\Leftrightarrow\ &\text{(iii) } \forall F\subset S\text{: closed, } \limsup_n P(X_n\in F)\le P(X\in F)\\
\Leftrightarrow\ &\text{(iv) } \forall G\subset S\text{: open, } \liminf_n P(X_n\in G)\ge P(X\in G)\\
\Leftrightarrow\ &\text{(v) } \forall A\in\mathcal{B}(S)\text{ s.t. } P(X\in\partial A)=0,\ \lim_n P(X_n\in A)=P(X\in A).
\end{aligned}$$

(S',d') をもう 1 つの距離空間とし，$\mathcal{B}(S')$ を S' の Borel σ-field とする．$g:S\to S'$ を連続写像とすると，$f\in C_b(S')$ に対して，$f\circ g\in C_b(S)$ であるから，$X_n\xrightarrow{d}X$ なら，

$$\lim_n E[f\circ g(X_n)] = E[f\circ g(X)]$$

である．従って，$g(X_n)\xrightarrow{d}g(X)$ となる．より一般に次の定理が成り立つ．

Theorem 33.7 (連続写像定理)**.** $g:S\to S'$ を $\mathcal{B}(S)/\mathcal{B}(S')$ 可測とし，g の不連続点全体を D_g とおく．このとき，$D_g\in\mathcal{B}(S)$ である．また，$X_n\xrightarrow{d}X$ & $P(X\in D_g)=0$ なら，$g(X_n)\xrightarrow{d}g(X)$ となる．

Proof. $D_g \in \mathcal{B}(S)$ であることの証明は省略する (cf. Exercise 3.16)．$F \subset S$ に対して，

$$D_g^c \cap \overline{g^{-1}(F)} \subset g^{-1}(\overline{F})$$

を示す．$x \in D_g^c \cap \overline{g^{-1}(F)}$ に対して，$\exists x_n \in g^{-1}(F)$ s.t. $x_n \to x$. このとき，$g(x_n) \in F$ であって，x は g の連続点であるから，$g(x) \in \overline{F}$，i.e.，$x \in g^{-1}(\overline{F})$ である．

従って，任意の閉集合 $F \subset S$ に対して，

$$\begin{aligned}\limsup_n P(g(X_n) \in F) &\le \limsup_n P(X_n \in \overline{g^{-1}(F)}) \le P(X \in \overline{g^{-1}(F)}) \\ &= P(X \in D_g^c \cap \overline{g^{-1}(F)}) \le P(X \in g^{-1}(F)) = P(g(X) \in F).\end{aligned}$$

よって，Portmanteau 定理より，$g(X_n) \xrightarrow{d} g(X)$ が示された． □

以下，(S, d) は可分と仮定する．このとき，$S^2 = S \times S$ の直積位相に関する Borel σ-field $\mathcal{B}(S^2)$ と，$\mathcal{B}(S)$ と $\mathcal{B}(S)$ の直積 σ-field は一致する：$\mathcal{B}(S^2) = \mathcal{B}(S) \times \mathcal{B}(S)$. よって，$S$-valued r.v.'s X, Y に対して，(X, Y) は $\mathcal{F}/\mathcal{B}(S^2)$ 可測である．$S^2 \ni (x, y) \mapsto d(x, y)$ は連続であるから，$\mathcal{B}(S^2)$ 可測であり，よって，$d(X, Y)$ は r.v. である．

そこで，S-valued r.v.'s X, X_n が与えられたとき，任意の $\varepsilon > 0$ に対して，

$$P(d(X_n, X) > \varepsilon) \to 0,\ n \to \infty$$

となるとき，X_n は X に 確率収束する (converge in probability) と言って，$X_n \xrightarrow{P} X$ と書く．次の補題の証明は演習問題とする．

Lemma 33.2. (i) $X_n \xrightarrow{P} X$ ならば，$X_n \xrightarrow{d} X$. (ii) $a \in S$ に対して，$X_n \xrightarrow{d} a$ ならば，$X_n \xrightarrow{P} a$.

次の定理は，しばしば “principle of accompanying laws” と呼ばれる．

Theorem 33.8. $X_{m,n}, X_n, m, n \in \mathbb{N}$ を S-valued r.v.'s とし，各 $m \in \mathbb{N}$ に対して，$X_{m,n} \xrightarrow{d} Z_m$ $(n \to \infty)$ とする．さらに，$Z_m \xrightarrow{d} X$ $(m \to \infty)$ とする．このとき，任意の $\varepsilon > 0$ に対して，

$$\lim_{m\to\infty} \limsup_{n\to\infty} P(d(X_{m,n}, X_n) > \varepsilon) = 0$$

ならば，$X_n \xrightarrow{d} X$ となる．

Proof. $F \subset S$ を閉集合とし，$F^\varepsilon = \{x \in S : d(x, F) \le \varepsilon\}$ とおくと，

$$\begin{aligned}P(X_n \in F) &= P(X_n \in F, d(X_n, X_{m,n}) \le \varepsilon) + P(X_n \in F, d(X_n, X_{m,n}) > \varepsilon) \\ &\le P(X_{m,n} \in F^\varepsilon) + P(d(X_n, X_{m,n}) > \varepsilon).\end{aligned}$$

$n \to \infty$ のとき，$X_{m,n} \xrightarrow{d} Z_m$ であって，F^ε は閉集合であるから，Portmanteau 定理より，

$$\limsup_n P(X_n \in F) \le P(Z_m \in F^\varepsilon) + \limsup_n P(d(X_n, X_{m,n}) > \varepsilon).$$

$Z_m \xrightarrow{d} X$ $(m \to \infty)$ であるから，再び Portmanteau 定理より，

$$\limsup_n P(X_n \in F) \le P(X \in F^\varepsilon).$$

$\varepsilon = 1/k \downarrow 0$ とすると，右辺 $\downarrow P(X \in F)$ より，$X_n \xrightarrow{d} X$ を得る． □

Skorohod の表現定理

λ を $[0,1]$ 上の Lebesgue 測度 (一様分布) とする

Theorem 33.9 (Skorohod の表現定理)．(S,d) を Poland 空間とし，$\mu_n \in \mathcal{P}(S), 1 \le n \le \infty$ とする．$\mu_n \xrightarrow{w} \mu_\infty$ なら，次をみたす $(\Omega, \mathcal{F}, P) = ([0,1], \mathcal{B}_{[0,1]}, \lambda)$ 上の S-valued r.v.'s $X_n, 1 \le n \le \infty$ が存在する：$X_n \sim \mu_n, 1 \le n \le \infty$ & $d(X_n, X_\infty) \to 0$ a.s. $(n \to \infty)$.

Proof. 各 $k = 1, 2, \dots$ に対して，$\{A_{k,i}\}_{i\in\mathbb{N}}$ を可算個の排反な Borel 集合からなる S の被覆とし，次の性質をみたすとする：$\mathrm{diam}(A_{k,i}) < 1/k$, $\mu_\infty(\partial A_{k,i}) = 0$，かつ $\{A_{k,i}\}_{i\in\mathbb{N}}$ は $\{A_{k-1,i}\}_{i\in\mathbb{N}}$ の細分となる，すなわち，$A_{k-1,i} = \bigcup_{i' : A_{k,i'} \subset A_{k-1,i}} A_{k,i'}$．$\{A_{k,i}\}_{i\in\mathbb{N}}$ は次のように帰納的に構成できる．S は可分であるから，直径が 1 より小さい可算個の開球 $\{B_i\}_{i\in\mathbb{N}}$ で S を被覆できる．各 $x_0 \in S$ に対して，$\mu_\infty(\partial B(x_0, r)) > 0$ なる r は高々可算個しかないので，$\mu_\infty(\partial B_i) = 0$ と仮定してよい．そこで，

$$A_{1,1} = B_1,\ A_{1,i} = B_i \setminus \bigcup_{j=1}^{i-1} B_j,\ i \ge 2$$

とおく．次に，直径が $1/2$ より小さい開球たちからなる同様の被覆 $\{A'_{2,i}\}_{i\in\mathbb{N}}$ を構成し，$\{A'_{2,i} \cap A_{1,j}\}_{i,j\in\mathbb{N}}$ を $\{A_{2,i}\}_{i\in\mathbb{N}}$ とする．同様の操作を続けて，$\{A_{k,i}\}_{i\in\mathbb{N}}$ を得る．

また，各 $n = 1, 2, \dots, \infty$ に対して，$[0,1]$ の部分区間の族 $\{I^n_{k,i}\}_{i\in\mathbb{N}}$ を次のように構成する．まず，$\alpha^n_0 = 0$,

$$I^n_{1,1} = [0, \underbrace{\mu_n(A_{1,1})}_{=\alpha^n_1}],\ I^n_{1,2} = [\mu_n(A_{1,1}), \underbrace{\mu_n(A_{1,1}) + \mu_n(A_{1,2})}_{=\alpha^n_2}], \cdots$$

とし，次に，$A_{1,i} = \bigcup_\ell A_{2,i_\ell}$ $(i_1 < i_2 < \cdots)$ のとき，

$$\begin{aligned} I^n_{1,i} &= [\alpha^n_{i-1}, \alpha^n_{i-1} + \mu_n(A_{2,i_1})] \cup [\alpha^n_{i-1} + \mu_n(A_{2,i_1}), \alpha^n_{i-1} + \mu_n(A_{2,i_1}) + \mu_n(A_{2,i_2})] \cup \cdots \\ &=: I^n_{2,i_1} \cup I^n_{2,i_2} \cup \cdots \end{aligned}$$

とおいて，$\{I^n_{2,i}\}_{i\in\mathbb{N}}$ を決める．この操作を繰り返して，$\{I^n_{k,i}\}_{i\in\mathbb{N}}$ を得る．ここで，$\{I^n_{k,i}\}_{i\in\mathbb{N}}$ は $\{I^n_{k-1,i}\}_{i\in\mathbb{N}}$ の細分であって，$P(I^n_{k,i}) = \lambda(I^n_{k,i}) = \mu_n(A_{k,i})$ をみたす．

各 k, i に対して，$x_{k,i} \in A_{k,i}$ を任意に 1 つ選び，$X^k_n : \Omega \to S$ を

$$X^k_n(\omega) = x_{k,i}, \quad \omega \in (I^n_{k,i})^\circ$$

と定義する．$P(\partial I^n_{k,i}) = 0$ より，$I^n_{k,i}$ の端点では X^k_n はどう選んでもよい．例えば，$x_0 \in S$ を任意にとり，$\omega \in \bigcup_i \partial I^n_{k,i}$ に対して $X^k_n(\omega) = x_0$ と決めておく．$N_n = \bigcup_k \bigcup_i \partial I^n_{k,i}$ とおく．$\{A_{k,i}\}$ の構成より，$\omega \in N^c_n$ に対して，

$$\ell \geq k \Rightarrow d(X^k_n(\omega), X^\ell_n(\omega)) < 1/k$$

であるから，$\{X^k_n(\omega)\}_{k=1}^\infty$ は Cauchy 列である．従って，S の完備性より，$X^k_n(\omega)$ は $k \to \infty$ のとき収束する．また，$\omega \in N$ のときは，十分大きな k に対して，$X^k_n(\omega) = x_0$ となるから，いずれにせよ $X^k_n(\omega)$ は $k \to \infty$ のとき収束する．そこで，$X_n(\omega) := \lim_k X^k_n(\omega), \omega \in \Omega$ とおくと，X_n は $\mathcal{F}/\mathcal{B}(S)$ 可測である (cf. Exercise 33.3)．

$X_n \sim \mu_n$ を示す．$F \subset S$ を閉集合とし，$F_k = \{x \in S : d(x, F) \leq 1/k\}, J_k = \{i : A_{k,i} \cap F \neq \varnothing\}$ とおくと，

$$\begin{aligned} P(X^k_n \in F) &\leq P\left(X^k_n \in \bigcup_{i\in J_k} A_{k,i}\right) \leq \sum_{i\in J_k} P(X^k_n \in A_{k,i}) = \sum_{i\in J_k} P(I^n_{k,i}) \quad (\because P(N_n) = 0) \\ &= \sum_{i\in J_k} \mu_n(A_{k,i}) \leq \mu_n(F_k). \end{aligned}$$

$k \to \infty$ のとき，$\mu_n(F_k) \downarrow \mu_n(F)$ であるから，Portmanteau 定理より，$\mathcal{L}(X^k_n) \xrightarrow{w} \mu_n$ を得る．よって，$X_n \sim \mu_n$ が示された．

最後に，$d(X_n, X_\infty) \to 0$ a.s. $(n \to \infty)$ を示す．$1 \leq n, k < \infty$ に対して，$K_{k,n} = \{i : \mu_\infty(A_{k,i}) > \mu_n(A_{k,i})\}$ とおくと，

$$\sum_i \{\mu_\infty(A_{k,i}) - \mu_n(A_{k,i})\} = 1 - 1 = 0$$

より，$\sum_{i\notin K_{k,n}} \{\mu_\infty(A_{k,i}) - \mu_n(A_{k,i})\} = -\sum_{i\in K_{k,n}} \{\mu_\infty(A_{k,i}) - \mu_n(A_{k,i})\}$. 従って，

$$\begin{aligned} \sum_i |P(I^\infty_{k,i}) - P(I^n_{k,i})| &= \sum_i |\mu_\infty(A_{k,i}) - \mu_n(A_{k,i})| = \sum_{i\in K_{k,n}} \{\cdots\} - \sum_{i\notin K_{k,n}} \{\cdots\} \\ &= 2 \sum_{i\in K_{k,n}} \{\cdots\} = 2 \sum_i \{\mu_\infty(A_{k,i}) - \mu_n(A_{k,i})\}^+. \end{aligned}$$

ここで，$\mu_n \xrightarrow{w} \mu_\infty$ と $\mu_\infty(\partial A_{k,i}) = 0$ より，$\mu_n(A_{k,i}) \to \mu_\infty(A_{k,i})$ $(n \to \infty)$ である (Portmanteau 定理)．さらに，$\{\cdots\}^+ \leq \mu_\infty(A_{k,i}), \sum_i \mu_\infty(A_{k,i}) = 1$ であるから，DCT より，$\sum_i |P(I^\infty_{k,i}) - P(I^n_{k,i})| \to 0$ $(n \to \infty)$ を得る．

k, i を任意に固定し，$1 \leq n \leq \infty$ に対して，β_n を $I^n_{k,i}$ の左端点とし，$L_n = \{j : I^n_{k,j}$ は $I^n_{k,i}$ の左側にある $\}$ とおくと，L_n は n に依存しない：$L_n \equiv L_\infty$. さらに，

$$\beta_\infty = \sum_{j \in L_\infty} P(I^\infty_{k,j}) = \lim_{n\to\infty} \sum_{j \in L_\infty} P(I^n_{k,j}) = \lim_{n\to\infty} \beta_n.$$

同様に，$I^n_{k,i}$ の右端点も $I^\infty_{k,i}$ のそれに収束する．

以上より，$\omega \in (I^\infty_{k,i})^\circ$ のとき，十分大きな n に対して，$\omega \in (I^n_{k,i})^\circ$ となるから，

$$d(X_n(\omega), X_\infty(\omega)) < 1/k.$$

従って，$\lim_n d(X_n(\omega), X_\infty(\omega)) = 0$ $(\forall \omega \in N^c_\infty)$ を得る． □

Theorem 33.9 において，S の完備性は外すことができる．この拡張は Dudley (1968) による．さらに，Wichura (1970) は S が単に距離空間であっても (つまり完備性や可分性をみたさなくても)，極限分布 μ_∞ が可分な Borel 集合に集中していたら，Theorem 33.9 の結論が成り立つことを示した．以上の経緯から，最後に述べた最も一般的なバージョンを，Skorohod-Dudley-Wichura の表現定理と呼ぶ．

$\mathbb{R}^{\mathbb{N}}$ における弱収束

$\mathbb{R}^{\mathbb{N}}$ を $\mathbb{R}$ の可算直積とし，直積位相を入れる．直積位相は距離

$$d(x, y) = \sum_{j=1}^{\infty} \frac{1}{2^j}(1 \wedge |x_j - y_j|), \quad x = (x_j), y = (y_j) \in \mathbb{R}^{\mathbb{N}}$$

によって距離化可能であり，この距離に関して $\mathbb{R}^{\mathbb{N}}$ は完備かつ可分である．また，$\mathbb{R}^{\mathbb{N}}$ の Borel σ-field $\mathcal{B}(\mathbb{R}^{\mathbb{N}})$ はシリンダー σ-field に一致する (cf. Exercise 7.6)：$\mathcal{B}(\mathbb{R}^{\mathbb{N}}) = \mathcal{B}^{\mathbb{N}}$. 各 $k \in \mathbb{R}^k$ に対して，射影 $\pi_k : \mathbb{R}^{\mathbb{N}} \to \mathbb{R}^k$ を $\pi(x_1, x_2, \dots) = (x_1, \dots, x_k)$ と定義する．$\mu \in \mathcal{P}(\mathbb{R}^{\mathbb{N}})$ に対して，$\mu \circ \pi_k^{-1}$ の形の分布を μ の有限次元分布 (f.d.d.) と呼ぶのであった．$\mathbb{R}^{\mathbb{N}}$ における弱収束は f.d.d.'s の収束と同値であることを示そう (C 空間との違いに気づくべし！)．

Theorem 33.10. $\mu_n, \mu \in \mathcal{P}(\mathbb{R}^{\mathbb{N}})$ に対して，$\mu_n \xrightarrow{w} \mu$ in $\mathbb{R}^{\mathbb{N}}$ $(n \to \infty) \Leftrightarrow \forall k \in \mathbb{N}$; $\mu_n \circ \pi_k^{-1} \xrightarrow{d} \mu \circ \pi_k^{-1}$ in $\mathbb{R}^k$ $(n \to \infty)$.

証明は次の補題による．

Lemma 33.3. S を可分な距離空間とし，$\mathcal{A} \subset 2^S$ を次の条件をみたす π システムとする：各 $x \in S$ と $\varepsilon > 0$ に対して，$\mathcal{A}_{x,\varepsilon} = \{A \in \mathcal{A} : x \in A^\circ \subset A \subset B(x, \varepsilon)\}$ とおいたとき，$\partial\mathcal{A}_{x,\varepsilon} := \{\partial A : A \in \mathcal{A}_{x,\varepsilon}\}$ は $\varnothing$ を含むか，非可算無限個の排反な集合を含む．$\mu, \mu_n \in \mathcal{P}(S)$ を given とする．このとき，$\mu(\partial A) = 0$ なる任意の $A \in \mathcal{A}$ に対して，$\mu_n(A) \to \mu(A)$ なら，$\mu_n \xrightarrow{w} \mu$ となる．

Proof. $\mathcal{A}_\mu = \{A \in \mathcal{A} : \mu(\partial A) = 0\}$ とおく．$\mathcal{A}_\mu$ は π システムである．さらに，$A_1, \dots, A_r \in \mathcal{A}_\mu$ に対して，包除原理より，

$$\mu_n\left(\bigcup_i A_i\right) = \sum_i \mu_n(A_i) - \sum_{i<j} \mu_n(A_i \cap A_j) + \cdots + (-1)^{r-1}\mu_n\left(\bigcap_i A_i\right)$$
$$\to \sum_i \mu(A_i) - \sum_{i<j} \mu(A_i \cap A_j) + \cdots + (-1)^{r-1}\mu\left(\bigcap_i A_i\right) = \mu\left(\bigcup_i A_i\right)$$

となる．

いま，$\partial\mathcal{A}_{x,\varepsilon}$ が $\varnothing$ を含まないなら，$\partial\mathcal{A}_{x,\varepsilon}$ は非可算無限個の排反な集合を含むから，そのうちの少なくとも 1 つは μ 測度が 0 である．よって，開集合 $G \subset S$ を given とすると，各 $x \in G$ に対して，次をみたす $A_x \in \mathcal{A}_\mu$ が存在する：$x \in A_x^\circ \subset A_x \subset G$. S は可分であるから，可算個の $x_i, i = 1, 2, \dots$ が存在して，$G = \bigcup_i A_{x_i}^\circ$ とできる．そこで，任意の $\delta > 0$ に対して，$r \in \mathbb{N}$ を十分大きく選んで，$\mu(\bigcup_{i=1}^r A_{x_i}^\circ) \geq \mu(G) - \delta$ とすると，

$$\mu(G) - \delta \leq \mu\left(\bigcup_{i=1}^r A_{x_i}^\circ\right) = \lim_n \mu_n\left(\bigcup_{i=1}^r A_{x_i}\right) \leq \liminf_n \mu_n(G).$$

よって，Portmanteau 定理より，$\mu_n \xrightarrow{w} \mu$ を得る． □

Proof of Theorem 33.10. $\pi_k : \mathbb{R}^{\mathbb{N}} \to \mathbb{R}^k$ は連続であるから，連続写像定理より，$\mu_n \xrightarrow{w} \mu$ なら，$\mu_n \circ \pi_k^{-1} \xrightarrow{w} \mu \circ \pi_k^{-1}$ となる．逆を示そう．$\mathcal{A} = \{\{y = (y_j) \in \mathbb{R}^{\mathbb{N}} : |y_j - x_j| < \eta, 1 \leq j \leq k\} : x = (x_j) \in \mathbb{R}^{\mathbb{N}}, \eta > 0, k \in \mathbb{N}\}$ とおく．いま，$\forall k \in \mathbb{N}; \mu_n \circ \pi_k^{-1} \xrightarrow{w} \mu \circ \pi_k^{-1}$ なら，Portmanteau 定理より，$\mu(\partial A) = 0$ なる任意の $A \in \mathcal{A}$ に対して，$\mu_n(A) \to \mu(A)$ となる．よって，あとは $\partial\mathcal{A}_{x,\varepsilon}$ が非可算無限個の排反な集合を含むことを示せばよいが，これは難しくない． □

Corollary 33.2. $X^n = (X_1^n, X_2^n, \dots), n \in \mathbb{N}$ を $\mathbb{R}^{\mathbb{N}}$-valued r.v.'s とすると，X^n が分布収束するためには，X^n の各 f.d.d. が弱収束することが必要十分である．さらに，各 $k \in \mathbb{N}$ に対して，$\mathcal{L}(X_1^n, \dots, X_k^n) \xrightarrow{w} \mu_k$ in $\mathbb{R}^k$ $(n \to \infty)$ となるなら，$\mu_k, k \in \mathbb{N}$ を f.d.d.'s にもつ $\mathbb{R}^{\mathbb{N}}$-valued r.v. X が存在して，$X^n \xrightarrow{d} X$ となる．

Proof. 必要性は明らか．各 $k \in \mathbb{N}$ に対して，$\mathcal{L}(X_1^n, \dots, X_k^n) \xrightarrow{w} \mu_k$ in $\mathbb{R}^k$ $(n \to \infty)$ とすると，$\mu_k, k \in \mathbb{N}$ は整合性条件をみたすので，それらを f.d.d.'s にもつ $\mathbb{R}^{\mathbb{N}}$-valued r.v. X が存在する．ここで，各 k に対して，$(X_1^n, \dots, X_k^n) \xrightarrow{d} (X_1, \dots, X_k)$ であるから，$X^n \xrightarrow{d} X$ を得る． □

Vitali-Hahn-Sacks の定理

Portmanteau 定理より，距離空間 S 上の Borel p.m.'s μ, μ_n に対して，μ_n が μ に弱収束することは，$\mu(\partial A) = 0$ なる任意の $A \in \mathcal{B}(S)$ に対して，$\mu_n(A) \to \mu(A)$ となることが必

要十分であった．では，μ_n が与えられて，あらゆる $A \in \mathcal{B}(S)$ に対して，$\lim_n \mu_n(A)$ が存在するとき，$\mu(A) := \lim_n \mu_n(A), A \in \mathcal{B}(S)$ は p.m. となるであろうか？じつはもっと一般に次の定理が成り立つ．

Theorem 33.11 (Vitali-Hahn-Sacks)**.** $(\Omega, \mathcal{F})$ を可測空間，P_n を $(\Omega, \mathcal{F})$ 上の p.m.'s とし，あらゆる $A \in \mathcal{F}$ に対して，$\lim_n P_n(A) =: P(A)$ が存在するとする．このとき，P は $(\Omega, \mathcal{F})$ 上の p.m. である．

もちろん P が $\mathcal{F}$ 上で有限加法的であることはすぐわかる．従って，$A_n \downarrow \varnothing, A_n \in \mathcal{F}$ に対して，$\lim_n P(A_n) = 0$ を示せばよいが，この部分が難しい．まず，次の定理を示す．位相空間の部分集合が 疎 (nowhere dense) であるとは，その閉包の内部が空であることを言う．例えば，開集合の境界は疎である．しかし，$\mathbb{Q}$ は $\mathbb{R}$ の部分集合としての閉包が $\mathbb{R}$ であるから，疎ではない ($\mathbb{Q}$ 自身の内部は空であるが)．

Theorem 33.12 (Baire のカテゴリー定理)**.** 完備距離空間は疎な閉集合の可算和にならない．

疎な閉集合の可算和として表せる位相空間を 第 1 カテゴリー に属すると言う．第 1 カテゴリーに属さない位相空間は 第 2 カテゴリー に属すると言う．Baire のカテゴリー定理を言い換えると，完備距離空間は第 2 カテゴリーに属するということである．

Proof. S を完備距離空間とし，仮に疎な閉集合 $F_n, n = 1, 2, \ldots$ が存在して，$S = \bigcup_n F_n$ と表せるとする．このとき，F_1^c は空でない開集合であるから，ある $x_1 \in S$ と $\varepsilon_1 > 0$ が存在して，$B(x_1, \varepsilon_1) \subset F_1^c$ となる．F_2 は疎であるから，

$$U_2 := B(x_1, \varepsilon_1/2) \cap F_2^c$$

は空でない．よって，ある $x_2 \in S$ と $0 < \varepsilon_2 < \varepsilon_1/2$ が存在して，

$$B(x_2, \varepsilon_2) \subset U_2$$

となる．これを繰り返して，次をみたす $x_n \in S, \varepsilon_n > 0$ を得る：

$$\varepsilon_{n+1} < \varepsilon_n/2,\ B(x_{n+1}, \varepsilon_{n+1}) \subset U_{n+1} := B(x_n, \varepsilon_n/2) \cap F_{n+1}^c.$$

$d(x_n, x_{n+1}) < \varepsilon_n/2$ より，$\{x_n\}$ は Cauchy 列であるから，S の完備性より，$x_n \to x \in S$ となる．このとき，

$$x \in \bigcap_n B(x_n, \varepsilon_n) \subset \bigcap_n F_n^c$$

となるが，これは $S = \bigcup_n F_n$ に反する． □

Proof of Theorem 33.11. $Q(A) = \sum_{n=1}^{\infty} P(A)/2^n, A \in \mathcal{F}$ とおくと，Q は $(\Omega, \mathcal{F})$ 上の p.m. であり，$P_n \ll Q, \forall n$ となる．いま，$d(A,B) = Q(A\Delta B), A, B \in \mathcal{F}$ と定義すると，$Q(A\Delta B) = 0$ なる A, B を同一視すれば，d は $\mathcal{F}$ 上の完備な距離となる．もっと正確に述べると，$A \in \mathcal{F}$ に対して，$[A] = \{\check{A} \in \mathcal{F} : Q(A\Delta\check{A}) = 0\}$ とおき，$[\mathcal{F}] = \{[A] : A \in \mathcal{F}\}$ とおくと，$[\mathcal{F}]$ は $\tilde{d}([A],[B]) = d(A,B), A, B \in \mathcal{F}$ に関して完備な距離空間になる．ここで，各 P_n は $P_n \ll Q$ であるから，$P_n([A]) := P_n(A), A \in \mathcal{F}$ と定義すれば，P_n は $[\mathcal{F}]$ 上の well-defined な連続関数である．以下，$\mathcal{F}$ と $[\mathcal{F}]$ を同一視する．

与えられた $\varepsilon > 0$ に対して，

$$F_k(\varepsilon) = \{A \in \mathcal{F} : \sup_n |P_k(A) - P_{k+n}(A)| \le \varepsilon\}$$

は $\mathcal{F}$ の閉集合であり，仮定より，$\mathcal{F} = \bigcup_{k=1}^{\infty} F_k(\varepsilon)$ である．$\mathcal{F}$ は完備であるから，Baire のカテゴリー定理より，少なくとも 1 つの $F_{k_0}(\varepsilon)$ は空でない内部をもつ．これは，ある $A_0 \in \mathcal{F}$ と $\eta > 0$ が存在して，

$$d(A, A_0) < \eta, A \in \mathcal{F} \Rightarrow \sup_n |P_{k_0}(A) - P_{k_0+n}(A)| \le \varepsilon$$

となることを意味する．一方，$Q(A) \le \eta$ なる任意の $A \in \mathcal{F}$ に対して，$A_1 = A \cup A_0, A_2 = A_0 \setminus (A \cap A_0)$ とおくと，$A = A_1 \setminus A_2, d(A_1, A_0) \le \eta, d(A_2, A_0) \le \eta$ をみたす．よって，$Q(A) \le \eta, k \ge k_0$ なら，

$$\begin{aligned} P_k(A) &\le P_{k_0}(A) + |P_{k_0}(A) - P_k(A)| \\ &\le P_{k_0}(A) + |P_{k_0}(A_1) - P_k(A_1)| + |P_k(A_2) - P_{k_0}(A_2)| \le P_{k_0}(A) + 2\varepsilon \end{aligned}$$

となる．この評価と $P_{k_0} \ll Q$ より，$Q(A) \to 0$ なら n に関して一様に $P_n(A) \to 0$ となることがわかる．これから，P の可算加法性が従う． □

演習問題

Exercise 33.1. 完備性は距離に依存した概念である．ある距離に関して完備な空間が，別の同値な距離に関して完備でないことはありうる．そのような例を構成せよ．

Exercise 33.2. 同値ではない 2 つの距離が同じ Borel σ-field を誘導することはありうる．そのような例を構成せよ．

Exercise 33.3. S を距離空間，$X_n, n \in \mathbb{N}$ を S-valued r.v.'s とし，各 $\omega \in \Omega$ に対して，$n \to \infty$ のとき，$X_n(\omega) \to X(\omega)$ in S とする．このとき，X も S-valued r.v. であることを示せ．

Remark 33.2. Exercise 33.3 の結論は S が単に位相空間のときは成り立たない．$S = [0,1]^{[0,1]} = \{f : f : [0,1] \to [0,1]\}$ とし，S には各点収束位相を入れる．$(\Omega, \mathcal{F}) = ([0,1], \mathcal{B}_{[0,1]})$ とし，

$$X_n(\omega)(t) = \max\{1 - n|\omega - t|, 0\},\ \omega \in \Omega, t \in [0,1]$$

とおくと，$\Omega \ni \omega \mapsto X_n(\omega) \in S$ は連続であるから，X_n は $\mathcal{F}/\mathcal{B}(S)$ 可測であり，各 $\omega \in \Omega$ に対して，$X_n(\omega) \to X(\omega) = 1_{\{\omega\}}(\cdot)$ となる．しかし，X は $\mathcal{F}/\mathcal{B}(S)$ 可測でない．実際，任意の $C \subset \Omega$ に対して，$U_C = \cup_{t \in C}\{x \in S : x(t) > 0\}$ は S の開集合であるが，C を $[0,1]$ の Lebesgue 非可測集合とすれば，$X^{-1}(U_C) = C \notin \mathcal{F}$ となる．

Exercise 33.4. $\mu \in \mathcal{P}(S)$ に対して，$\mathcal{A} = \{A \in \mathcal{B}(S) : \mu(\partial A) = 0\}$ とおいたとき，$\mathcal{A}$ が field であることを示せ．また，$\mathcal{A}$ は σ-field とは限らない．そのような例を構成せよ．

Exercise 33.5. Theorem 33.3 から Corollary 33.1 を導け．

Exercise 33.6. Lemma 33.1 の証明において，(i) $[0,1]^{\mathbb{N}}$ が直積位相に関してコンパクトであることを示せ．(ii) h が S から $h(S)$ への同相写像になることを示せ．

Exercise 33.7. Lemma 33.2 を示せ．

Exercise 33.8. $C_b(S)$ は一様距離に関して完備であることを示せ．また，(S,d) が全有界であっても，$C_b(S)$ は可分とは限らない．そのような例を構成せよ．

Exercise 33.9. S を可分な距離空間とし，$A \in \mathcal{B}(S)$ に対して，射影 $\pi_A : \mathcal{P}(S) \to [0,1]$ を $\pi_A(\mu) = \mu(A)$ と定義する．このとき，$\mathcal{P}(S)$ の弱位相に関する Borel σ-field は，すべての射影 $\pi_A, A \in \mathcal{B}(S)$ が可測になる最小の σ-field に一致することを示せ．

S を可分な距離空間とし，$\mathcal{P}(S)$ の弱位相に関する Borel σ-field を $\mathcal{B}(\mathcal{P}(S))$ とおく．写像 $\hat{\mu} : \Omega \times \mathcal{B}(S) \to [0,1]$ が $\mathcal{P}(S)$-valued r.v. であるとは，$\omega \mapsto \hat{\mu}(\omega, \cdot)$ が $\Omega/\mathcal{B}(\mathcal{P}(S))$ 可測になることを言う．前問より，$\hat{\mu}$ が $\mathcal{P}(S)$-valued r.v. であるための必要十分条件は，各 $A \in \mathcal{B}(S)$ に対して，$\omega \mapsto \hat{\mu}(\omega, A)$ が r.v. であり，各 $\omega \in \Omega$ に対して，$\mathcal{B}(S) \ni A \mapsto \hat{\mu}(\omega, A)$ が p.m. になることである．

Exercise 33.10. S を可分な距離空間とし，$\hat{\mu}, \hat{\mu}_n, n = 1, 2, \ldots$ を $\mathcal{P}(S)$-valued r.v.'s とする．いま，各 $f \in C_b(S)$ に対して，

$$\int f d\hat{\mu}_n \xrightarrow{P} \int f d\hat{\mu}$$

とする．ここで，$f \in C_b(S)$ に対して，$\mu \mapsto \int f d\mu$ は連続であるから，両辺は r.v.'s である．このとき，弱位相を与える任意の距離 ρ に対して，$\rho(\hat{\mu}_n, \hat{\mu})$ は r.v. であり，$\rho(\hat{\mu}_n, \hat{\mu}) \xrightarrow{P} 0$ となることを示せ．

Exercise 33.11 (Rubin)**.** S を Poland 空間，S' を可分な距離空間とし，$X_n, 1 \le n \le \infty$ を次をみたす S-valued r.v.'s とする：$X_n \xrightarrow{d} X_\infty$ $(n \to \infty)$．また，$g_n : S \to S', 1 \le n \le \infty$ を任意の $x_\infty \in S$ と $x_n \to x_\infty$ $(n \to \infty)$ なる任意の $\{x_n\} \subset S$ に対して $\lim_n g_n(x_n) = g_\infty(x_\infty)$ をみたす可測写像とする．このとき，$g_n(X_n) \xrightarrow{d} g_\infty(X_\infty)$ を示せ．

Exercise 33.12. Corollary 33.2 は $\mathbb{R}^{\mathbb{N}}$ を Poland 空間 S の可算直積 $S^{\mathbb{N}}$ に取り替えても成り立つ．このことを示せ．

Exercise 33.13. $1 \le p < q \le \infty$ とし，$L^q = L^q(\Omega, \mathcal{F}, P)$ は無限次元とする．このとき，$(L^q, \|\cdot\|_p)$ は第 1 カテゴリーに属することを示せ．

ヒント：$L^q = \bigcup_n \{X \in L^q : \|X\|_q \le n\} =: \bigcup_n F_n$ である．各 F_n が $\|\cdot\|_p$ に関して疎な閉集合であることを示せ．

34 Prohorov の定理と Prohorov 距離

Prohorov の定理

$(S, \mathcal{B}(S))$ 上の p.m.'s の族 $\Pi \subset \mathcal{P}(S)$ が 一様にタイト (uniformly tight) であるとは，任意の $\varepsilon > 0$ に対して，あるコンパクト集合 $K = K_\varepsilon \subset S$ が存在して，$\mu(K) \geq 1-\varepsilon$ $(\forall \mu \in \Pi)$ となることを言う．

Theorem 34.1 (Prohorov)**.** (S, d) を可分とし，$\Pi \subset \mathcal{P}(S)$ とする．このとき，(i) Π が一様にタイトなら，Π は (弱位相に関して) 相対コンパクトである．(ii) 逆に，(S, d) が Poland 空間のとき，Π が相対コンパクトなら，Π は一様にタイトである．

可分な距離空間 S に対して，弱位相は距離位相であったから，コンパクト性と点列コンパクト性は等価である．従って，Π が相対コンパクトであることは，Π に属する任意の p.m.'s の列が弱収束する部分列をもつことと同値である．

Corollary 34.1. (S, d) が Poland 空間のとき，$\{\mu_n\} \subset \mathcal{P}(S)$ が弱収束するための必要十分条件は，$\{\mu_n\}$ が一様にタイトであり，かつその任意の弱収束部分列が一意な極限をもつことである．

1 つの p.m. $\mu \in \mathcal{P}(S)$ が タイト (tight) であるとは，任意の $\varepsilon > 0$ に対して，あるコンパクト集合 $K = K_\varepsilon \subset S$ が存在して，$\mu(K) \geq 1 - \varepsilon$ となることを言う．$\Pi = \{\mu\}$ (1 点集合) はコンパクトなので，(S, d) が Poland 空間なら，μ はタイトである．従って，次の系を得る．

Corollary 34.2 (Ulam の定理)**.** Poland 空間上の任意の Borel p.m. はタイトである．

Theorem 34.1 の定理の証明に，次の Riesz の表現定理を用いる．証明は本節の最後を参照せよ．

Theorem 34.2 (Riesz の表現定理)**.** (S, d) をコンパクトとし，L を $C(S)$ 上の非負値線形汎関数とする．すなわち，$L : C(S) \to \mathbb{R}$ は線形であり，

$$f \in C(S), f \geq 0 \Rightarrow L(f) \geq 0$$

をみたすと仮定する．このとき，S 上の一意な有限 Borel 測度 μ が存在して，

$$L(f) = \int f d\mu, \ \forall f \in C(S)$$

と表すことができる．

Lemma 34.1. $S_0 \subset S$ を S の部分集合とし，$\mathcal{B}(S_0)$ を S_0 の相対位相に関する Borel σ-field とすると，$\mathcal{B}(S_0)$ は $\mathcal{B}(S)$ の S_0 への制限に一致する：$\mathcal{B}(S_0) = \{S_0 \cap A : A \in \mathcal{B}(S)\}$.

Proof. cf. Exercise 2.8. □

Proof of Theorem 34.1. (i). まず (S,d) をコンパクトと仮定する．このとき，$C(S)$ は一様距離に関して可分であるから，$\{f_m\} \subset C(S)$ を次をみたすように選べる：各 $m \in \mathbb{N}$ に対して $0 \le f_m \le 1$ であり，$f_1, f_2, \dots$ の有限 1 次結合の全体 $\mathrm{span}\{f_1, f_2, \dots\}$ は $C(S)$ で稠密となる．このとき，任意の $\{\mu_n\} \subset \mathcal{P}(S)$ に対して，対角線論法より，ある部分列 $\{\mu_{n_k}\}$ が存在して，すべての m に対して有限な $\lim_k \int f_m d\mu_{n_k}$ が存在する．ここで，$\mathrm{span}\{f_1, f_2, \dots\}$ は $C(S)$ で稠密であるから，すべての $f \in C(S)$ に対して，有限な

$$\lim_k \int f d\mu_{n_k} =: L(f)$$

が存在する．L は $C(S)$ 上の非負値線形汎関数であるから，Riesz の表現定理より，S 上の一意な有限 Borel 測度 μ が存在して，

$$L(f) = \int f d\mu, \ \forall f \in C(S)$$

と表せる．さらに，$L(1) = 1$ より μ は p.m. であり，L の構成より，$\mu_{n_k} \xrightarrow{w} \mu$ $(k \to \infty)$ となる．以上より，(S,d) がコンパクトなら $\mathcal{P}(S)$ がコンパクトになることが示された．

一般に，(S,d) がコンパクトとは限らないときは，d と同値な距離 $\hat{d}$ であって，$(S,\hat{d})$ が全有界になるものをとる．$(S,\hat{d})$ の完備化を T とすると，T はコンパクトである．$\{\mu_n\} \subset \Pi$ を任意にとり，各 μ_n に対して，$\mathcal{B}(T)$ 上の p.m. $\bar{\mu}_n$ を

$$\bar{\mu}_n(A) = \mu_n(S \cap A), \ A \in \mathcal{B}(T)$$

と定義する (cf. Lemma 34.1)．このとき，$\mathcal{P}(T)$ のコンパクト性より，適当な部分列 $\{\bar{\mu}_{n_k}\}$ と p.m. $\bar{\mu} \in \mathcal{P}(T)$ が存在して，$\bar{\mu}_{n_k} \to \bar{\mu}$ in $\mathcal{P}(T)$ となる．あとは，$\bar{\mu}$ の S への制限が well-defined な p.m. になることを確認する．この部分に Π の一様なタイトネスを用いる．

いま，Π が一様にタイトであることから，S のコンパクト集合の列 $K_\ell, \ell \in \mathbb{N}$ であって，各 ℓ に対して，$\mu_{n_k}(K_\ell) \ge 1 - 1/\ell$ $(\forall k \in \mathbb{N})$ をみたすものが存在する．各 K_ℓ は T のコンパクト集合でもあるから，Portmanteau 定理より，$\bar{\mu}(K_\ell) \ge \limsup_k \bar{\mu}_{n_k}(K_\ell) = \limsup_k \mu_{n_k}(K_\ell) \ge 1 - 1/\ell$. よって，$\bar{\mu}(\bigcup_{\ell=1}^{\infty} K_\ell) = 1$ となるから，$\bar{\mu}$ の外測度を $\bar{\mu}^*$ とおくと，$\bar{\mu}^*(S) = 1$ となる [19]．そこで，$\mu : \mathcal{B}(S) \to [0,1]$ を

$$\mu(S \cap A) = \bar{\mu}(A), \ A \in \mathcal{B}(T)$$

と定義すれば (cf. Lemma 34.1)，$\bar{\mu}^*(S) = 1$ より μ は $(S, \mathcal{B}(S))$ 上の well-defined な p.m. になる．あとは，$\mu_{n_k} \to \mu$ in $\mathcal{P}(S)$ を示せばよい．各 $f \in U_{\hat{d}}(S)$ は T 上の連続関数 $\bar{f}$ に一意に拡張できる．ここで，$\bar{\mu}_{n_k} \to \bar{\mu}$ in $\mathcal{P}(T)$ より，

$$\int f d\mu_{n_k} = \int \bar{f} d\bar{\mu}_{n_k} \to \int \bar{f} d\bar{\mu} = \int f d\mu, \ k \to \infty$$

[19] $S \in \mathcal{B}(T)$ とは限らない．例えば，S を $\mathbb{R}$ の Lebesgue 非可測集合とすると，S の $\mathbb{R}$ での閉包 $T = \overline{S}$ は S の完備化であるが，$S \notin \mathcal{B}(T)$ である．

を得る．従って，Portmanteau 定理より，$\mu_{n_k} \to \mu$ in $\mathcal{P}(S)$ を得る．

(ii). $G_n \uparrow S$ なる任意の開集合列 G_n と任意の $\varepsilon > 0$ に対して，$\mu(G_n) > 1-\varepsilon$ $(\forall \mu \in \Pi)$ となる n が存在することを示す．仮にそのような n が存在しないなら，各 n に対して，$\mu_n(G_n) \leq 1-\varepsilon$ となる $\mu_n \in \Pi$ が存在するが，Π の相対コンパクト性より，$\exists\{\mu_{n_k}\}$ s.t. $\mu_{n_k} \overset{w}{\to} \mu$. しかし，Portmanteau 定理より，$\mu(G_n) \leq \liminf_k \mu_{n_k}(G_n) \leq 1-\varepsilon$ であるから，$n \to \infty$ とすると，$\mu(S) \leq 1-\varepsilon$ となって，μ が p.m. であることに反する．

$\{x_m\}$ を S の稠密な部分集合とし，各 $n, k \in \mathbb{N}$ に対して，

$$G_n^k = \bigcup_{m=1}^{n} B(x_m, 1/k)$$

とおくと，$\bigcup_{n=1}^{\infty} G_n^k = S$. よって，$\forall \varepsilon > 0, \forall k \in \mathbb{N}, \exists n_k \in \mathbb{N}$ s.t. $\mu(G_{n_k}^k) > 1-\varepsilon/2^k$ $(\forall \mu \in \Pi)$. $K = \bigcap_{k=1}^{\infty} \overline{G_{n_k}^k}$ とおくと，$K^c = \bigcup_{k=1}^{\infty} (\overline{G_{n_k}^k})^c \subset \bigcup_{k=1}^{\infty} (G_{n_k}^k)^c$ であるから，$\mu(K^c) \leq \sum_{k=1}^{\infty} \mu((G_{n_k}^k)^c) \leq \varepsilon$ $(\forall \mu \in \Pi)$. 従って，$\inf_{\mu \in \Pi} \mu(K) \geq 1-\varepsilon$ である．あとは K がコンパクトであることを示せばよい．K は閉集合であるから，S の完備性より K も完備である．さらに，各 k に対して，

$$K \subset \bigcup_{m=1}^{n_k} B(x_m, 2/k)$$

であるから，K は全有界である．以上より，K がコンパクトであることが示された． □

Remark 34.1. Ulam の定理において，S の完備性は本質的な仮定である．λ を $[0,1]$ 上の Lebesgue 測度とし，$S \subset [0,1]$ を $\lambda_*(S) = 0$, $\lambda^*(S) = 1$ をみたす集合とする (S は Lebesgue 非可測集合である)．S に通常の距離を入れると，S は完備でない．いま，$\mu \in \mathcal{P}(S)$ を

$$\mu(S \cap A) = \lambda(A), \ A \in \mathcal{B}_{[0,1]}$$

と定義すると，μ は well-defined な p.m. である．このとき，S のコンパクト集合 K は $[0,1]$ のコンパクト集合でもある．ところが，$\lambda_*(S) = 0$ より，$\mu(K) = \lambda(K) = 0$ でなくてはならないので，μ はタイトになりえない．

Remark 34.2. Prohorov の定理を適用するためには，S のコンパクト集合がどのような集合であるかわかっていないといけない．$S = \mathbb{R}^k$ のときは，コンパクト集合であることは有界閉集合であることと同値なので，$\mathbb{R}^k$-valued r.v.'s $\{\boldsymbol{X}_n\}$ に対して，$\{\mathcal{L}(\boldsymbol{X}_n)\}$ が一様にタイトであるためには，$\lim_{M\to\infty} \sup_n P(|\boldsymbol{X}_n| > M) = 0$ となることが必要十分である．S が (ベクトル空間であることは仮定して) 無限次元の場合は，一様タイト性の特徴づけはもっと複雑である．注意しなくてはならないのは，無限次元の (実/複素) ノルム空間 $(S, \|\cdot\|)$ では，コンパクト集合であることと有界閉集合であることは同値ではないので，S-valued r.v.'s $\{X_n\}$ に対して，$\lim_{M\to\infty} \sup_n P(\|X_n\| > M) = 0$ であっても，$\{\mathcal{L}(X_n)\}$ が一様にタイトであるとは限らない (逆は正しい)．

測度の正則性と Radon 測度. Ulam の定理のついでに，測度の正則性を議論する．(S,d) を距離空間とし，μ を $(S,\mathcal{B}(S))$ 上の測度とする (p.m. とは限らない)．$A\in\mathcal{B}(S)$ に対して，μ が

$$\mu(A)=\inf\{\mu(U):U\supset A,\ U:\text{open}\}$$

をみたすとき，μ は A 上で外正則 (inner regular) であると言う．また，μ が

$$\mu(A)=\sup\{\mu(K):K\subset A,\ K:\text{compact}\}$$

をみたすとき，μ は A 上で内正則 (inner regular) であると言う．μ がすべての Borel 集合上で外正則 (内正則) であるとき，μ は外正則 (内正則) であると言う．μ が外正則かつ内正則であるとき，μ は正則 (regular) であると言う．μ が局所有限 であり (すなわち，S の各点が μ 測度が有限な開近傍をもち)，さらに内正則であるとき，μ を Radon 測度 と呼ぶ[20]．局所有限な測度は各コンパクト集合に対して有限の値をとる．S が局所コンパクトなら (すなわち，S の各点が相対コンパクトな開近傍をもてば)，逆も正しい．有限測度 μ が S 上で内正則であるとき，μ はタイトであると言う．この定義は，p.m. のタイト性と整合的である．

Theorem 34.3. μ が有限なら，μ は外正則である．さらに，μ がタイトなら，μ は正則 (従って，Radon) である．

Proof. $\mathcal{A}=\{A\in\mathcal{B}(S):\mu$ は A と A^c 上で外正則 $\}$ とおく．A が開集合なら，μ は A 上で外正則であり，$U_n=\{x:d(x,A^c)<1/n\}$ とおくと，$U_n\supset A^c$ は開集合であり，$U_n\downarrow A^c$ となる．よって，$\mu(A^c)=\lim_n\mu(U_n)$ であるから，μ は A^c 上で外正則である．すなわち，$\mathcal{A}$ は開集合全体を含むので，$\mathcal{A}$ が σ-field であることを示せば，$\mathcal{A}=\mathcal{B}(S)$ が従う．$S\in\mathcal{A}$ と $A\in\mathcal{A}\Rightarrow A^c\in\mathcal{A}$ はよい．$A_n\in\mathcal{A}$ $(n=1,2,\dots)$ として，与えられた $\varepsilon>0$ と各 n に対して，開集合 $U_n\supset A_n$ を $\mu(U_n\setminus A_n)\le\varepsilon/2^n$ をみたすように選ぶ．ここで，$U=\bigcup_n U_n$ とおくと，$U\setminus A\subset\bigcup_n(U_n\setminus A_n)$ より，$\mu(U\setminus A)\le\varepsilon$ を得る．よって，$A\in\mathcal{A}$ であるから，$\mathcal{A}$ は σ-field である．

さらに，μ がタイトなら，与えられた $\varepsilon>0$ に対して，適当なコンパクト集合 K が存在して，$\mu(K^c)\le\varepsilon$ となる．次に，$A\in\mathcal{A}$ に対して，外正則性より，適当な閉集合 $F\subset A$ が存在して，$\mu(A\setminus F)\le\varepsilon$ となる．$L=K\cap F$ とおくと，L はコンパクトであり，$\mu(A\setminus L)\le 2\varepsilon$ となる．よって，μ は正則である． □

従って，有限 Borel 測度に対しては，(a) タイト性，(b) 正則性，(c) Radon 性，はすべて等価である．さらに Ulam の定理より，Poland 空間上のあらゆる有限 Borel 測度はタイトである．では，$(\mathbb{R}^k,\mathcal{B}^k)$ 上の局所有限測度，例えば Lebesgue 測度，は正則 (従って，Radon) であろうか．答えは Yes である．もう少し一般に，次の定理が成り立つ．S が σ コンパクト であるとは，$K_n\uparrow S$ なるコンパクト集合列 $\{K_n\}$ が存在することを言う．例えば，S が可分かつ局所コンパクトなら，S は σ コンパクトである．

[20] Radon 測度の定義は本によってちょっとずつ違っているので，注意が必要である．

Theorem 34.4. S を局所コンパクトかつ σ コンパクトとし，μ を局所有限とする．このとき，μ は正則 (従って，Radon) である．

Proof. 証明を 2 つのステップに分割する．

Step 1. μ の内正則性を示す．$\{K_n\}$ を $K_n \uparrow S$ なるコンパクト集合列とし，$n = 1, 2, \dots$ に対して，$\mu_n(A) := \mu(A \cap (K_n \setminus K_{n-1})), A \in \mathcal{B}(S)$ と定義する．ただし，$K_0 = \varnothing$ である．$A \in \mathcal{B}(S)$ を given とする．各 μ_n は有限であるから，任意の $\varepsilon > 0$ と $n \in \mathbb{N}$ に対して，あるコンパクト集合 $L_n \subset A$ が存在して，$\mu_n(L_n) \geq \mu_n(A) - \varepsilon/2^n$ をみたす．$M_N = \bigcup_{n=1}^N L_n$ とおくと，M_N は A に含まれるコンパクト集合であり，$\mu(M_N) = \sum_{n=1}^\infty \mu_n(M_N) \geq \sum_{n=1}^N \mu_n(L_n) \geq \sum_{n=1}^N \mu_n(A) - \varepsilon$ をみたす．よって，

$$\sup_N \mu(M_N) \geq \sum_{n=1}^\infty \mu_n(A) - \varepsilon = \mu(A) - \varepsilon$$

より，μ の内正則性が示された．

Step 2. 次に μ の外正則性を示す．まず，$A \subset S$ が相対コンパクトな Borel 集合の場合を考える．S の局所コンパクト性より，各 $x \in S$ に対して，相対コンパクトな開近傍 U_x が存在する．いま，A の閉包 $\overline{A}$ はコンパクトであるから，有限個の $x_1, \dots, x_k \in \overline{A}$ が存在して，$A \subset \overline{A} \subset \bigcup_{j=1}^k U_{x_j} =: U$ となる．$L = \overline{U}$ とおくと，L はコンパクトであるから，$\mu(L \setminus A) < \infty$ である．よって，内正則性より，任意の $\varepsilon > 0$ に対して，あるコンパクト集合 $K \subset L \setminus A$ が存在して，$\mu(K) \geq \mu(L \setminus A) - \varepsilon$ をみたす．$V = U \setminus K = U \cap K^c$ とおくと，V は A を含む開集合であり，

$$\mu(V) \leq \mu(L \cap K^c) \leq \mu(L) - \mu(K) \leq \mu(L) - \mu(L \setminus A) + \varepsilon = \mu(A) + \varepsilon$$

をみたす．よって，μ は A 上で外正則である．

$A \subset S$ を一般の Borel 集合とし，$A = \bigcup_n A_n$ と分割する．ただし，$\{A_n\}$ は排反な相対コンパクト集合列である．S の σ コンパクト性より，このような分割は存在する．$\mu(A) = +\infty$ なら A 上での外正則性は明らかなので，$\mu(A) < +\infty$ と仮定する．前段落の結果より，任意の $\varepsilon > 0$ と $n \in \mathbb{N}$ に対して，ある開集合 $U_n \supset A_n$ が存在して，$\mu(U_n) \leq \mu(A_n) + \varepsilon/2^n$ をみたす．ここで，$U = \bigcup_n U_n$ とおくと，U は A を含む開集合であり，

$$\mu(U) \leq \sum_n \mu(U_n) \leq \sum_n \mu(A_n) + \varepsilon = \mu(A) + \varepsilon$$

をみたす．よって，μ は A 上で外正則である． □

Prohorov 距離

S が可分なら $\mathcal{P}(S)$ は弱位相に関して可分であることを示したが，じつは S が Poland 空間なら，弱位相は完備な距離によって距離化可能である．(S, d) を可分な距離空間とし，$\delta >$

$0, A \subset S$ に対して，$A^\delta = \{x \in S : d(x, A) \le \delta\}$ とおく．ただし，$d(x, A) = \inf_{y \in A} d(x, y)$ である．$\mu, \nu \in \mathcal{P}(S)$ に対して，

$$\rho(\mu, \nu) := \inf\{\varepsilon > 0 : \mu(A) \le \nu(A^\varepsilon) + \varepsilon,\ \forall A \in \mathcal{B}(S)\}$$

と定義する．$(\overline{A})^\varepsilon = A^\varepsilon$ であるから，右辺の "$\forall A \in \mathcal{B}(S)$" は "$\forall A \subset S$: closed" に変えてもよい．

Theorem 34.5. ρ は弱位相の距離を与える．

Proof. まず，ρ が距離であることを示す．$\rho(\mu, \nu) \ge 0$ はよい．対称性を示す．$\rho(\mu, \nu) > \varepsilon$ なら，$\exists A \in \mathcal{B}(S)$ s.t. $\mu(A) > \nu(A^\varepsilon) + \varepsilon$. ここで，$((A^\varepsilon)^c)^\varepsilon \subset A^c$ である ($d(x, y) \le \varepsilon$, $\exists y \in (A^\varepsilon)^c$ なら，$x \in A^c$ である．仮に $x \in A$ なら $y \in A^\varepsilon$ となって矛盾が起こる)．ゆえに，$\nu((A^\varepsilon)^c) > \mu(A^c) + \varepsilon \ge \mu(((A^\varepsilon)^c)^\varepsilon) + \varepsilon$. すなわち，$\rho(\nu, \mu) > \varepsilon$. $\varepsilon \uparrow \rho(\mu, \nu)$ として，$\rho(\nu, \mu) \ge \rho(\mu, \nu)$. μ と ν の役割を入れ替えて，$\rho(\mu, \nu) = \rho(\nu, \mu)$ を得る．

次に，$\rho(\mu, \mu) = 0$ はよい．いま，$\rho(\mu, \nu) = 0$ なら，任意の閉集合 $F \subset S$ に対して，$\mu(F) \le \nu(F^{1/n}) + n^{-1}$. $n \to \infty$ とすると，$F^{1/n} \downarrow F$ より，$\mu(F) \le \nu(F)$. μ と ν を入れ替えて，$\mu(F) = \nu(F)$ を得る．ゆえに，$\mu = \nu$ である．

三角不等式は明らかである．これで ρ が距離であることが示された．

次に，

$$\mu_n \xrightarrow{w} \mu \Leftrightarrow \rho(\mu_n, \mu) \to 0$$

を示す．まず，$\rho(\mu_n, \mu) \to 0$ なら，任意の $\varepsilon > 0$ と閉集合 $F \subset S$ に対して，$\limsup_n \mu_n(F) \le \mu(F^\varepsilon) + \varepsilon$. $\varepsilon = 1/k \downarrow 0$ として，$\limsup_n \mu_n(F) \le \mu(F)$. ゆえに，Portmanteau 定理より，$\mu_n \xrightarrow{w} \mu$ を得る．

逆に，$\mu_n \xrightarrow{w} \mu$ とする．S は可分であるから，与えられた $\varepsilon > 0$ に対して，S の可測分割 $\{A_i\}$ であって，各 A_i の直径が ε 以下のものが存在する．k を十分大きく選んで，$\mu(\bigcup_{i=k+1}^\infty A_i) \le \varepsilon$ とする．いま，

$$\mathcal{G} = \{((A_{i_1} \cup \cdots \cup A_{i_m})^\varepsilon)^\circ : 1 \le i_1 < \cdots < i_m \le k\}$$

とおくと，$\mathcal{G}$ は有限であるから，Portmanteau 定理より，$\exists n_0 \in \mathbb{N}$ s.t. $n \ge n_0 \Rightarrow \mu_n(G) > \mu(G) - \varepsilon$ ($\forall G \in \mathcal{G}$). $A \in \mathcal{B}(S)$ に対して，$A_0 = \bigcup_{i : A_i \cap A \ne \varnothing} A_i$ とおくと，$(A_0^\varepsilon)^\circ \in \mathcal{G}$ であるから，$n \ge n_0$ のとき，

$$\begin{aligned}\mu(A) &\le \mu(A_0) + \mu\left(\bigcup_{i=k+1}^\infty A_i\right) \le \mu(A_0) + \varepsilon \le \mu((A_0^\varepsilon)^\circ) + \varepsilon \\ &< \mu_n((A_0^\varepsilon)^\circ) + 2\varepsilon \le \mu_n(A^{2\varepsilon}) + 2\varepsilon.\end{aligned}$$

ゆえに，$\limsup_n \rho(\mu_n, \mu) \le 2\varepsilon$ が示された． □

ρ を Prohorov 距離 と呼ぶ．

Theorem 34.6. (S,d) が Poland 空間なら，$(\mathcal{P}(S),\rho)$ は Poland 空間である．

Proof. $(\mathcal{P}(S),\rho)$ が完備なことを示せばよい．$\{\mu_n\}\subset\mathcal{P}(S)$ を ρ に関する Cauchy 列とする．S は Poland 空間であるから，Ulam の定理より，各 μ_n はタイトである．ゆえに，$\varepsilon>0$ を given とすると，各 $\ell\in\mathbb{N}$ に対して，正整数 m_ℓ とコンパクト集合 $K_\ell\subset S$ を

$$\rho(\mu_n,\mu_{m_\ell})\le\varepsilon/2^{\ell+1},\forall n\ge m_\ell,\quad \mu_n(K_\ell^c)\le\varepsilon/2^{\ell+1},1\le\forall n\le m_\ell$$

をみたすように選べる．$\varepsilon_\ell=\varepsilon/2^\ell$ とおくと，各 $\ell\in\mathbb{N}$ に対して，$\mu_n((K_\ell^{\varepsilon_\ell})^c)\le\varepsilon_\ell$ $(\forall n)$. そこで，$K=\bigcap_{\ell=1}^\infty K_\ell^{\varepsilon_\ell}$ とおくと，$\mu_n(K)\ge1-\varepsilon$ $(\forall n)$. S は完備であり，K は閉集合であるから，K は完備である．また，各 K_ℓ はコンパクトであるから，K は全有界である．従って，K はコンパクトであるから，$\{\mu_n\}$ は一様にタイトである．ゆえに，$\{\mu_n\}$ は弱収束する部分列をもつが，$\{\mu_n\}$ は Cauchy 列だったので，$\{\mu_n\}$ 自身が弱収束する． □

S-valued r.v.'s X,Y に対して，

$$\alpha(X,Y):=\inf\{\varepsilon>0:P(d(X,Y)>\varepsilon)\le\varepsilon\}\qquad(*)$$

と定義する．

Theorem 34.7. (i) $(*)$ の右辺の inf は min に取り替えることができる．

(ii) a.s. に等しい S-valued r.v.'s を同一視すると，α は S-valued r.v.'s の全体のなす空間上の距離を与える．

(iii) $X_n\xrightarrow{P}X\Leftrightarrow\alpha(X_n,X)\to0$.

Proof. (i). $\varepsilon_m\downarrow\alpha=\alpha(X,Y)$ とすると，$\{d(X,Y)>\varepsilon_m\}\uparrow\{d(X,Y)>\alpha\}$ であるから，$P(d(X,Y)>\alpha)=\lim_m P(d(X,Y)>\varepsilon_m)\le\lim_m\varepsilon_m=\alpha$.
(ii) と (iii) の証明は難しくないので省略する． □

α は Ky Fan 距離 と呼ばれる．ところで，$X_n\xrightarrow{P}X$ なら $X_n\xrightarrow{d}X$ であるが，Prohorov 距離と Ky Fan 距離には次の関係がある．

Theorem 34.8. S-valued r.v.'s X,Y に対して，$\rho(\mathcal{L}(X),\mathcal{L}(Y))\le\alpha(X,Y)$.

Proof. $\alpha=\alpha(X,Y)$ とおくと，

$$\begin{aligned}P(X\in A)&=P(\{X\in A\}\cap\{d(X,Y)\le\alpha\})+P(\{X\in A\}\cap\{d(X,Y)>\alpha\})\\&\le P(Y\in A^\alpha)+\alpha.\end{aligned}$$

よって，$\rho(\mathcal{L}(X),\mathcal{L}(Y))\le\alpha$. □

Prohorov 距離と Ky Fan 距離の間には，さらに次のような興味深い関係がある．すなわち，Prohorov 距離の意味で近い 2 つの p.m.'s に対して，それぞれを分布にもつ r.v.'s であって，Ky Fan 距離の意味で近いものを構成できる．より正確には，次の定理が成り立つ．

Theorem 34.9 (Strassen-Dudley)**.** (S,d) を可分とし，$\mu,\nu\in\mathcal{P}(S)$ をタイトとする．また，$\varepsilon,\delta>0$ を given とする．このとき，次の (i) と (ii) は同値である．

(i) 次をみたす S-valued r.v.'s X,Y が存在する：$\mathcal{L}(X)=\mu,\mathcal{L}(Y)=\nu,\ \&\ P(d(X,Y)>\delta)\le\varepsilon$.

(ii) $\mu(A)\le\nu(A^{\delta})+\varepsilon\ (\forall A\in\mathcal{B}(S))$.

Proof. (i) $\Rightarrow$ (ii) は明らか．(ii) $\Rightarrow$ (i) の証明は，Dudley (2002, Theorem 11.6.2) を参照せよ． □

Riesz の表現定理の証明

Riesz の表現定理 (Theorem 34.2) を証明する．Riesz の表現定理は S がコンパクトな Hausdorff 空間でも成り立つが，ここでは距離空間としておく．なお，以下の証明は，Tao (2010) を参考にした．以下，(S,d) はコンパクトな距離空間とする．

Lemma 34.2 (Urysohn の補題の特殊ケース)**.** $K\neq\varnothing$ を S の閉集合とし，$U\supset K$ を開集合とする．このとき，次をみたす連続関数 $f:S\to[0,1]$ が存在する：$1_K\le f\le 1_U$.

Proof. S はコンパクトであるから，K,U^c はコンパクトである．よって，ある $\varepsilon>0$ が存在して，$d(x,y)\ge 2\varepsilon,\forall x\in K,\forall y\in U^c$ をみたす．そこで，

$$\eta(t)=\begin{cases}1 & \text{if } t<0\\ 1-t & \text{if } 0\le t\le 1\\ 0 & \text{if } t>1\end{cases}$$

とおき，$f(x)=\eta(\varepsilon^{-1}d(x,K))$ とおくと，f が求める関数である． □

Lemma 34.3 (1 の分割)**.** S の開被覆 $U_1,\dots,U_k$ に対して次をみたす連続関数 $\psi_j:S\to[0,1],1\le j\le k$ が存在する：$\psi_j=0,\forall x\in U_j^c,1\le j\le k\ \&\ \sum_{j=1}^k\psi_j(x)=1,\forall x\in S$.

Proof. S はコンパクトであるから，各 $x\in S$ に対して，ある $1\le j\le k$ とコンパクトな近傍 $B_x\subset U_j$ が存在する．このとき，$B_x^{\circ},x\in S$ は S の開被覆であるから，有限個の $x_1,\dots,x_m\in S$ が存在して，$S=\bigcup_{i=1}^m B_{x_i}^{\circ}=\bigcup_{i=1}^m B_{x_i}$ となる．そこで，$K_j=\bigcup_{i:B_{x_i}\subset U_j}B_{x_i}$ とおくと，各 K_j は U_j のコンパクトな部分集合である．よって，Lemma 34.2 より，ある連続関数 $f_j:S\to[0,1]$ が存在して，$1_{K_j}\le f_j\le 1_{U_j}$ となる．$\sum_{j=1}^k f_j(x)\ge\sum_{j=1}^k 1_{K_j}(x)\ge 1\ (\forall x\in S)$ より，$\psi_j(x)=f_j(x)/\sum_{\ell=1}^k f_\ell(x),x\in S$ とおくと，$\psi_j,1\le j\le k$ は補題の条件をみたす． □

Proof of Theorem 34.2. 一意性は明らか．μ の存在を示す．L は $C(S)$ 上で次の性質をみたす：$f \le g \Rightarrow L(f) \le L(g)$; $L(f) \le L(1)\|f\|_u$. 次に，

$$B^+_{lsc}(S) = \{f : f : S \to \mathbb{R}_+,\ f \text{ は有界かつ下半連続 }\}$$

とおく[21]．このとき，$f \in B^+_{lsc}(S)$ に対して，

$$L(f) := \sup\{L(g) : 0 \le g \le f, g \in C(S)\}$$

として，L を $B^+_{lsc}(S)$ 上に拡張する．拡張された L は $C(S)$ 上ではもともとの L に一致している．さらに，次の性質をみたす．

(a) $f, g \in B^+_{lsc}(S), f \le g \Rightarrow L(f) \le L(g)$.

(b) $f \in B^+_{lsc}(S), c > 0$ に対して，$L(cf) = cL(f)$.

(c) $f, f' \in B^+_{lsc}(S)$ に対して，$L(f) + L(f') \le L(f + f')$.

(d) $f, f_n \in B^+_{lsc}(S)$ $(n = 1, 2, \dots), f \le \sum_{n=1}^{\infty} f_n \Rightarrow L(f) \le \sum_{n=1}^{\infty} L(f_n)$.

(a) と (b) は明らか．(c) は $g, g' \in C(S)$ を $0 \le g \le f, 0 \le g' \le f'$ とすると，$0 \le g + g' \le f + f'$ より，$L(g) + L(g') = L(g + g') \le L(f + f')$. g, g' に関して sup をとって，$L(f) + L(f') \le L(f + f')$ を得る．

(d) を示す．$0 \le g \le f$ なる任意の $g \in C(S)$ に対して，$L(g) \le \sum_{n=1}^{\infty} L(f_n)$ を示せばよい．$0 < \varepsilon \le 1/4$ を given とし，$c > 0$ は ε と独立な定数とする．c の値はそれが現れる場所ごとに異なっていてよい．g は連続であるから，各 $x \in S$ に対して，適当な開近傍 U_x が存在して，$|g(y) - g(x)| \le \varepsilon, \forall y \in U_x$ となる．また，ある $N_x \in \mathbb{N}$ が存在して，

$$\sum_{n=1}^{N_x} f_n(x) \ge g(x) - \varepsilon \qquad (*)$$

となる．ここで，各 f_n は下半連続であるから，必要なら U_x を小さくとり直して，

$$f_n(y) \ge f_n(x) - \varepsilon/2^n,\ \forall y \in U_x, 1 \le \forall n \le N_x \qquad (**)$$

とできる．

いま，$U_x, x \in S$ は S の開被覆であり，S はコンパクトであるから，有限開被覆 $U_{x_1}, \dots, U_{x_k}$ が存在する．Lemma 34.3 より，次をみたす連続関数 $\psi_j : S \to [0, 1], 1 \le j \le k$ が存在する：$\psi_j(x) = 0, \forall x \in U_j^c, 1 \le j \le k$ & $\sum_{j=1}^{k} \psi(x) = 1, \forall x \in S$.

[21] コンパクト集合上の実数値下半連続関数は必ずしも有界ではない．例えば，$f : [0, 1] \to \mathbb{R}$ を $f(x) = 1/x, x \ne 0, f(0) = 0$ とおくと，f は下半連続であるが，有界でない．

$x \in S$ を $g(x) \geq \sqrt{\varepsilon}$ なる点とし，j を $x \in U_{x_j}$ をみたす任意の j とすると，$|g(x)-g(x_j)| \leq \varepsilon$ となる．このとき，$g(x_j) \geq g(x)-\varepsilon \geq \sqrt{\varepsilon}-\varepsilon \geq \sqrt{\varepsilon}/2$ となる．$f \geq g$ より，$f(x_j) \geq \sqrt{\varepsilon}/2$ であるから，$(*)$ の両辺を $f(x_j)$ で割って，

$$1 \leq \frac{g(x_j)}{f(x_j)} \leq \sum_{n=1}^{N_{x_j}} \frac{f_n(x_j)}{f(x_j)} + c\sqrt{\varepsilon}$$

を得る．また，

$$g(x) = \sum_{j:x\in U_{x_j}} g(x)\psi_j(x) \leq \sum_{j:x\in U_{x_j}} g(x_j)\psi_j(x) + \varepsilon$$

となる．以上の評価と，g が有界であることから，

$$\begin{aligned} g(x) &\leq \sum_{j:x\in U_j} \sum_{n=1}^{N_{x_j}} \frac{f_n(x_j)}{f(x_j)} g(x_j)\psi_j(x) + c\sqrt{\varepsilon} \\ &\leq \sum_{n=1}^{\infty} \sum_{j:f(x_j)\geq\sqrt{\varepsilon}/2, N_{x_j}\geq n} \frac{f_n(x_j)}{f(x_j)} g(x_j)\psi_j(x) + c\sqrt{\varepsilon} \end{aligned}$$

を得る．$c \geq 1$ としておけば，この評価は $g(x) < \sqrt{\varepsilon}$ なる x に対しても成り立つ．右辺の n に関する和のうち，ゼロでないものは有限個しかないので，

$$L(g) \leq \sum_{n=1}^{\infty} L\Big(\sum_{j:f(x_j)\geq\sqrt{\varepsilon}/2, N_{x_j}\geq n} \frac{f_n(x_j)}{f(x_j)} g(x_j)\psi_j \Big) + c\sqrt{\varepsilon}$$

を得る．ここで，

$$\sum_{j:f(x_j)\geq\sqrt{\varepsilon}/2, N_{x_j}\geq n} \frac{f_n(x_j)}{f(x_j)} g(x_j)\psi_j(x) \leq \sum_{j=1}^{k} f_n(x_j)\psi_j(x)$$

であり，さらに $(**)$ より，右辺は $f_n(x) + \varepsilon/2^n$ でバウンドできる．よって，

$$L(g) \leq \sum_{n=1}^{\infty} L(f_n) + c\sqrt{\varepsilon}$$

を得る．これで (d) が示された．

(c) と (d) より，L は有限加法的である．すなわち，$f, g \in B^+_{lsc}(S)$ に対して，$L(f+g) = L(f) + L(g)$ となる．

いま，開集合 $U \subset S$ に対して，1_U は下半連続であるから，

$$\mu(U) := L(1_U)$$

と定義する．さらに，閉集合 $K \subset S$ に対しては，

$$\mu(K) := \mu(S) - \mu(S \setminus K)$$

と定義する．L の有限加法性より，μ は well-defined であり，また (a) より $\mu \geq 0$ である．

そこで，任意の $A \subset S$ に対して，外測度を

$$\mu^+(A) := \inf\{\mu(U) : U \supset A, U : \text{open}\}$$

と定義し，内測度を

$$\mu^-(A) := \sup\{\mu(K) : K \subset A, K : \text{closed}\}$$

と定義する．ここで，$0 \leq \mu^-(A) \leq \mu^+(A) \leq \mu(S) = L(1) < \infty$ である．いま，

$$\mathcal{M}_\mu := \{A \subset S : \mu^+(A) = \mu^-(A)\}$$

とおくと，$\mathcal{M}_\mu$ は開集合全体を含む．なぜなら，開集合 U に対して，明らかに $\mu^+(U) = \mu(U)$ であり，また μ の定義より，$\mu(U) = \sup\{L(g) : 0 \leq g \leq 1_U, g \in C(S)\}$ であるが，$0 \leq g \leq 1_U$ なる $g \in C(S)$ に対して，そのサポートを $K := \overline{\{g > 0\}}$ とおくと，$1_{K^c} \leq 1 - g$ であるから，$\mu(K^c) = L(1_{K^c}) \leq L(1-g) = L(1) - L(g)$ より，$L(g) \leq \mu(K)$ を得る．よって，$\mu(U) \leq \mu^-(U) \leq \mu^+(U) = \mu(U)$ であるから，$U \in \mathcal{M}_\mu$ を得る．そこで，$A \in \mathcal{M}_\mu$ に対して，

$$\mu(A) := \mu^+(A) = \mu^-(A)$$

と定義する．この定義は以前の定義と整合的である．

$\mathcal{M}_\mu$ が σ-field であることと，μ が測度であることを示そう．$S \in \mathcal{M}_\mu$ と $A \in \mathcal{M}_\mu \Rightarrow A^c \in \mathcal{M}_\mu$ はよい．$A_n \in \mathcal{M}_\mu$ $(n = 1, 2, \dots)$ と与えられた $\varepsilon > 0$ に対して，次をみたす開集合 $U_n \supset A_n$ と閉集合 $K_n \subset A_n$ が存在する：$\mu(A_n) \leq \mu(U_n) \leq \mu(A_n) + \varepsilon/2^n$ & $\mu(A_n) - \varepsilon/2^n \leq \mu(K_n) \leq \mu(A_n)$．$L$ の性質 (d) より，$\mu(\bigcup_{n=1}^\infty U_n) \leq \sum_{n=1}^\infty \mu(U_n) \leq \sum_{n=1}^\infty \mu(A_n) + \varepsilon$ である．よって，$\mu^+(\bigcup_{n=1}^\infty A_n) \leq \sum_{n=1}^\infty \mu(A_n)$．また，$L$ の有限加法性より，$\mu(\bigcup_{n=1}^N K_n) = \sum_{n=1}^N \mu(K_n) \geq \sum_{n=1}^N \mu(A_n) - \varepsilon$ であるから，$\mu^-(\bigcup_{n=1}^\infty A_n) \geq \sum_{n=1}^N \mu(A_n)$．$N \to \infty$ として，$\mu^-(\bigcup_{n=1}^\infty A_n) \geq \sum_{n=1}^\infty \mu(A_n)$ を得る．以上より，$\mathcal{M}_\mu$ は σ-field であり，μ は測度であることが示された．$\mathcal{M}_\mu$ は開集合全体を含むので，$\mathcal{M}_\mu \supset \mathcal{B}(S)$ であり，μ の $\mathcal{B}(S)$ への制限を再び μ と書くことにすれば，μ は有限 Borel 測度である．

あとは，$f \in C(S)$ に対して，$L(f) = \int f d\mu =: L_\mu(f)$ を示せばよい．$L(f)$ と $L_\mu(f)$ の線形性から，$0 \leq f \leq 1$ と仮定してよい．$K_0 = S$ とし，$j = 1, \dots, N$ に対して，$K_j = \{f \geq jN^{-1}\}$ とおき，次に，$f_1, \dots, f_N \in C(S)$ を

$$f_j = \min\left\{\max\left\{f - \frac{j-1}{N}, 0\right\}, \frac{1}{N}\right\}$$

とおく．このとき，$N^{-1}1_{K_j} \leq f_j \leq N^{-1}1_{K_{j-1}}$ であるから，$N^{-1}\mu(N_j) \leq L_\mu(f_j) \leq N^{-1}\mu(K_{j-1})$．また，開集合 $U \supset K_{j-1}$ に対して，$1_{K_j} \leq N f_j \leq 1_U$ であるから，$N^{-1}\mu(K_k) \leq$

$L(f_j) \le N^{-1}\mu(U)$. U に関する inf をとって，$N^{-1}\mu(K_j) \le L(f_j) \le N^{-1}\mu(K_{j-1})$ を得る．いま，$f = \sum_{j=1}^{N} f_j$ であるから，$N^{-1}\sum_{j=1}^{N}\mu(K_j) \le L(f) \wedge L_\mu(f_j) \le L(f) \vee L_\mu(f) \le N^{-1}\sum_{j=0}^{N-1}\mu(K_j)$. よって，$|L(f)-L_\mu(f)| \le (\mu(K_0)+\mu(K_N))/N \le 2\mu(S)/N$ を得る． □

演習問題

Exercise 34.1. S が可分でなくても，$\Pi \subset \mathcal{P}(S)$ が一様にタイトなら，任意の $\{\mu_n\} \subset \Pi$ は弱収束する部分列をもつことを示せ．

Exercise 34.2. S, S' を可分な距離空間とし，A を一般の添字集合とする．各 $\alpha \in A$ に対して，X_α, Y_α をそれぞれ S-valued および S'-valued r.v.'s とする．このとき，$\{\mathcal{L}(X_\alpha, Y_\alpha) : \alpha \in A\}$ が一様にタイトであるためには，$\{\mathcal{L}(X_\alpha) : \alpha \in A\}$ と $\{\mathcal{L}(Y_\alpha) : \alpha \in A\}$ がともに一様にタイトであることが必要十分であることを示せ．

Exercise 34.3 (Fernique)**.** H を可分な実 Hilbert 空間とし，$\{x_n\}$ を H の完全正規直交系とする．H に (ノルム位相ではなく) 弱位相を入れたとき，p.m.'s

$$\mu_n = n^{-3}\sum_{m=1}^{n^3}\delta_{nx_m},\ n \in \mathbb{N}$$

は一様にタイトでないが，$\mu_n \xrightarrow{w} \delta_0$ となることを示せ．cf. Topsøe (1974).

以下，$(B, \|\cdot\|)$ を可分な実 Banach 空間とし，B^* をその双対空間とする．$\mathcal{B}(B)$ を B の Borel σ-field とする．また，B_1^* を B^* の閉単位球とし，$\mathcal{F}(B)$ をすべての $f \in B_1^*$ が可測になる B 上の最小の σ-field とする．

Exercise 34.4. $\mathcal{B}(B) = \mathcal{F}(B)$ を示せ．特に，μ, ν を $(B, \mathcal{B}(B))$ 上の p.m.'s とすると，$\mu \circ f^{-1} = \nu \circ f^{-1}, \forall f \in B_1^*$ なら $\mu = \nu$ であることを確認せよ．

Exercise 34.5. $K \subset B$ が相対コンパクトであるためには，K が有界であり，かつ任意の $\varepsilon > 0$ に対して，適当な有限次元部分空間 F が存在して，$d(x, F) < \varepsilon, \forall x \in K$ となることが必要十分である．ただし，$d(x, F) = \inf_{y \in F}\|x - y\|$ である．このことを示せ．

Exercise 34.6. Π を $(B, \mathcal{B}(B))$ 上の p.m.'s の族とする．Π が一様にタイトであるためには，次の 2 条件がみたされることが必要十分であることを示せ．

(a) あらゆる $f \in B_1^*$ に対して，$\{\mu \circ f^{-1} : \mu \in \Pi\}$ が一様にタイトである．

(b) 任意の $\varepsilon > 0$ に対して，ある有限次元部分空間 F が存在して，

$$\inf_{\mu \in \Pi}\mu(\{x : d(x, F) < \varepsilon\}) \ge 1 - \varepsilon$$

となる．

ヒント：問題なのは十分性の証明である．(a) の代わりに次の条件を考える：

(c) 任意の $\varepsilon > 0$ に対して，ある有界閉集合 L が存在して，$\inf_{\mu\in\Pi} \mu(L) \geq 1-\varepsilon$ となる．

まず，(c) と (b) がみたされるなら，Π が一様にタイトであることを示せ．次に，(a) と (b) がみたされるなら，(c) がみたされることを示せ．

Exercise 34.7 (Hilbert 空間における CLT)**.** $(H, \langle\cdot,\cdot\rangle)$ を可分な実 Hilbert 空間とし，$\mathcal{B}(H)$ を H の Borel σ-field とする．$X, X_1, X_2, \ldots$ を $\mathcal{F}/\mathcal{B}(H)$ 可測な i.i.d. H-valued r.v.'s とし，$E[\langle X, h\rangle] = 0, \forall h \in H, E[\|X\|^2] < \infty$ とする．また，$\check{S}_n = n^{-1/2}\sum_{i=1}^n X_i$ とおく．このとき，CLT が成り立つこと示せ．すなわち，$\check{S}_n \xrightarrow{d} G$ であり，G の分布は $\langle G, h\rangle \sim N(0, E[\langle X, h\rangle^2]), h \in H$ から (一意に) 決まる．

ヒント：任意の $\varepsilon > 0$ に対して，H の有限次元部分空間に値をとる r.v. Y が存在して，$E[\|X - Y\|^2] \leq \varepsilon$ とできる．このことから，$\{\mathcal{L}(\check{S}_n) : n \in \mathbb{N}\}$ の一様タイト性が従う．

Exercise 34.7 の結果はそのまま可分な Banach 空間に拡張できるわけではない．$X_1, \ldots, X_n$ を独立な B-valued r.v.'s とし，$E[f(X_i)] = 0, \forall f \in B_1^*, E[\|X_i\|^2] < \infty$ とすると，等式

$$E\left[\left\|\sum_{i=1}^n X_i\right\|^2\right] = \sum_{i=1}^n E[\|X_i\|^2]$$

は B が Hilbert 空間の場合は成り立つが，それ以外のときは成り立たない．さらに，そもそも，$E[f(X)] = 0, \forall f \in B_1^*, E[\|X\|^2] < \infty$ なる B-valued r.v. X を given とすると，$f(G) \sim N(0, E[f(X)^2]), \forall f \in B_1^*$ なる B-valued r.v. G が存在するとは限らない．

Exercise 34.8. $c_0 = \{(x_k) \in \mathbb{R}^{\mathbb{N}} : \lim_k x_k = 0\}$ とし，$(x_k) \in c_0$ に対して，$\|(x_k)\| = \sup_k |x_k|$ と定義する．

(a) $(c_0, \|\cdot\|)$ は可分な Banach 空間であることを示せ．

(b) $\varepsilon_1, \varepsilon_2, \ldots$ を独立な Rademacher r.v.'s として，

$$X = (X_k)_{k\in\mathbb{N}} = \left(\frac{\varepsilon_k}{(2\log(k+1))^{1/2}}\right)_{k\in\mathbb{N}}$$

とおく．このとき，次をみたす c_0-valued r.v. $G = (G_k)_{k\in\mathbb{N}}$ は存在しない ことを示せ: $(G_k)_{k\in\mathbb{N}}$ は Gauss 過程であり，$E[G_k G_\ell] = E[X_k X_\ell]$ をみたす．さらに，$X^1, X^2, \ldots$ を X の独立なコピーとして，$\check{S}^n = n^{-1/2}\sum_{i=1}^n X^i$ とおいたとき，$\check{S}^n$ は c_0 において分布収束しないことを確認せよ．

Banach 空間上の確率論に関しては，Ledoux and Talagrand (1991) が詳しい．

35 C 空間

Arzelà-Ascoli の定理

C 空間における弱収束を特徴づけるために，C のコンパクト集合の特徴づけが必要になる．Arzelà-Ascoli の定理はそのような特徴づけを与える．

(K,e) をコンパクトな距離空間とし，$C(K)=\{f\in\mathbb{R}^K : f \text{ は連続}\}$ とおく．$C(K)$ には一様距離

$$d(f,g)=\sup_{x\in K}|f(x)-g(x)|,\ f,g\in C(K)$$

を入れる．$(C(K),d)$ は完備である．

$F\subset\mathbb{R}^K$ が $x\in K$ において 同程度連続 (equicontinuous) であるとは，

$$\forall\varepsilon>0,\exists\delta>0,\forall y\in K, e(y,x)<\delta\Rightarrow|f(x)-f(y)|<\varepsilon\ \forall f\in F$$

となることを言う．F がすべての $x\in K$ において同程度連続であるとき，F を同程度連続と呼ぶ．また，

$$\forall\varepsilon>0,\exists\delta>0,\forall x,y\in K, e(x,y)<\delta\Rightarrow|f(x)-f(y)|<\varepsilon\ \forall f\in F$$

となるとき，F は 一様に同程度連続 (uniformly equicontinuous) であると言う．(K,e) がコンパクトなときは，両者は同値である．

Lemma 35.1. (K,e) をコンパクトとし，$F\subset\mathbb{R}^K$ とする．このとき，F が同程度連続であることと，F が一様に同程度連続であることは同値である．

Proof. F が同程度連続であると仮定して，一様に同程度連続になることを示せばよい．仮に F が一様に同程度連続でないなら，ある $\varepsilon>0, x_n,y_n\in K$ と $f_n\in F$ が存在して，$e(x_n,y_n)\to 0$ & $\|f_n(x_n)-f_n(y_n)|\geq\varepsilon\ (\forall n)$ となる．K はコンパクトなので，$x_n\to x\in K$ と仮定してよい．このとき，$e(x_n,y_n)\to 0$ であるから，$y_n\to x$ でもある．しかし，F は同程度連続なので，十分大きな n に対して，$|f_n(x_n)-f_n(x)|<\varepsilon/2$ & $|f_n(y_n)-f_n(x)|<\varepsilon/2$ となるが，このとき，$|f_n(x_n)-f_n(y_n)|<\varepsilon$ となって矛盾が起こる． □

Theorem 35.1 (Arzelà-Ascoli)**.** (K,e) をコンパクトとし，$F\subset C(K)$ とする．このとき，F が $C(K)$ において相対コンパクトであるための必要十分条件は，F が一様有界かつ同程度連続であることである．

Remark 35.1. $C(K)$ は完備であるから，$F\subset C(K)$ に対して，次の (i)–(iii) は同値である．

(i) F は相対コンパクト．

(ii) F 内の任意の関数列は一様収束する部分列をもつ．

(iii) F は全有界.

Proof. 必要性．F は全有界であるから，$C(K)$ の部分集合として有界である．すなわち，$\exists f_0 \in F, \exists M > 0$ s.t. $F \subset \{f \in C(K) : d(f, f_0) \leq M\}$. 従って，$\sup_{x\in K} |f(x)| \leq M + \sup_{x\in K} |f_0(x)|$ $(\forall f \in F)$ であるから，F は一様有界である．一方，F が全有界なことから，$\forall \varepsilon > 0, \exists f_1, \ldots, f_n \in F$ s.t. $\forall f \in F, \exists j = j(f) \in \{1, \ldots, n\}$ with $d(f, f_j) < \varepsilon/3$. 各 f_j は一様連続であるから，$\{f_1, \ldots, f_n\}$ は一様に同程度連続である．よって，$\exists \delta > 0$ s.t. $\forall x, y \in K, e(x, y) < \delta \Rightarrow |f_j(x) - f_j(y)| < \varepsilon/3$ $(1 \leq \forall j \leq n)$. 従って，$e(x, y) < \delta$ なら，任意の $f \in F$ に対して，$j = j(f)$ として

$$\begin{aligned} |f(x) - f(y)| &\leq \underbrace{|f(x) - f_j(x)|}_{\leq d(f, f_j)} + |f_j(x) - f_j(y)| + \underbrace{|f_j(y) - f(y)|}_{\leq d(f_j, f)} \\ &< \varepsilon/3 + \varepsilon/3 + \varepsilon/3 = \varepsilon. \end{aligned}$$

以上より，F は一様に同程度連続である．

十分性．F は一様に同程度連続とする．任意の $\{f_n\} \subset F$ が一様収束する部分列をもつことを示せばよい．K はコンパクトであるから，各 $m \in \mathbb{N}$ に対して，K を半径 $1/m$ の有限個の開球たちで覆うことができる．そのような開球の中心を集めた集合を D_m とおくと，各 D_m は有限なので，$D := \bigcup_{m=1}^{\infty} D_m$ は可算である．このとき，F が一様有界であることと，対角線論法より，$\{f_n\}$ の部分列 $\{f_{n_k}\}$ を，各 $x \in D$ に対して $\{f_{n_k}(x)\}_{k=1}^{\infty}$ が収束列となるように選ぶことができる．$\{f_{n_k}\}$ が一様収束することを示す.

いま，F は一様に同程度連続性であるから，$\forall \varepsilon > 0, \exists \delta > 0$ s.t. $e(x, y) < \delta \Rightarrow |f_{n_k}(x) - f_{n_k}(y)| < \varepsilon/3$ $(\forall k)$. $m > 1/\delta$ とする．D_m は有限であるから，十分大きな N に対して，$k, \ell \geq N \Rightarrow |f_{n_k}(x) - f_{n_\ell}(x)| < \varepsilon/3$ $(\forall x \in D_m)$. このとき，任意の $x \in K$ に対して，$e(x, x') < 1/m < \delta$ なる $x' \in D_m$ が存在するから，

$$\begin{aligned} |f_{n_k}(x) - f_{n_\ell}(x)| &\leq |f_{n_k}(x) - f_{n_k}(x')| + |f_{n_k}(x') - f_{n_\ell}(x')| + |f_{n_\ell}(x') - f_{n_\ell}(x)| \\ &< \varepsilon/3 + \varepsilon/3 + \varepsilon/3 = \varepsilon. \end{aligned}$$

従って，$\{f_{n_k}\}$ は $C(K)$ の Cauchy 列であり，$C(K)$ の完備性より，それは一様収束する． □

C 空間における弱収束

以下，$C = C([0, 1])$ とし，C には一様距離

$$d(x, y) = \sup_{t\in[0,1]} |x(t) - y(t)|, \ x, y \in C$$

を入れる．(C, d) は Poland 空間なので，Prohorov の定理の系より，$\{\mu_n\} \subset \mathcal{P}(C)$ が弱収束するための必要十分条件は，$\{\mu_n\}$ が一様にタイトであり，かつ $\{\mu_n\}$ の任意の弱収束部分列の極限が一意に決まることである．C の Borel σ-field $\mathcal{B}(C)$ はシリンダー σ-field に一

致するから，π-λ 定理より，$(C, \mathcal{B}(C))$ 上の各 p.m. はその f.d.d.'s によって一意に決まる．よって，次の定理を得る．

Theorem 35.2. $\mu_n \in \mathcal{P}(C)$ とする．μ_n が弱収束するための必要十分条件は，μ_n の各 f.d.d. が弱収束して，かつ $\{\mu_n\}$ が一様にタイトなことである．

有限次元分布の収束を示すのは容易なことが多いので，以下では，一様なタイトネスを保証する条件を考える．$x \in C, \delta \in (0,1]$ に対して，x の 連続度 (modulus of continuity) を

$$w_x(\delta) = w(x,\delta) = \sup\{|x(s) - x(t)| : |s-t| \le \delta, s, t \in [0,1]\}$$

と定義する．このとき，$A \subset C$ が同程度連続であることは，

$$\lim_{\delta \downarrow 0} \sup_{x \in A} w_x(\delta) = 0$$

と同値である．また，A が同程度連続であるとき，ある $t \in [0,1]$ に対して $\sup_{x\in A} |x(t)| < \infty$ であれば，A は一様有界である．従って，Arzelà-Ascoli の定理より，

$$A \subset C \text{ が相対コンパクト } \Leftrightarrow \sup_{x\in A} |x(0)| < \infty \ \& \ \lim_{\delta \downarrow 0} \sup_{x \in A} w_x(\delta) = 0$$

である．

Theorem 35.3. $\{\mu_n\} \subset \mathcal{P}(C)$ が一様にタイトであるための必要十分条件は，次の (i) と (ii) が成り立つことである．

(i)
$$\lim_{\alpha \uparrow \infty} \inf_n \mu_n(\{x \in C : |x(0)| \le \alpha) = 1.$$

(ii) 任意の $\varepsilon > 0$ に対して，
$$\lim_{\delta \downarrow 0} \limsup_{n \to \infty} \mu_n(\{x \in C : w_x(\delta) > \varepsilon\}) = 0.$$

Remark 35.2. $C \ni x \mapsto x(0)$ と $C \ni x \mapsto w_x(\delta)$ は連続であるから，ともに $\mathcal{B}(C)$ 可測である．

Proof. 必要性．$\{\mu_n\}$ が一様にタイトなら，任意の $\eta > 0$ に対して，コンパクト集合 K が存在して，$\mu_n(K) \ge 1 - \eta$ $(\forall n \in \mathbb{N})$ となる．Arzelà-Ascoli の定理より，十分大きな $\alpha > 0$ に対して，$K \subset \{x : |x(0)| \le \alpha\}$ であり，また，任意の $\varepsilon > 0$ に対して，$\delta > 0$ を十分小さく選べば，$K \subset \{x : w_x(\delta) \le \varepsilon\}$ となる．従って，

$$\inf_n \mu_n(\{x : |x(0)| \le \alpha\}) \ge \inf_n \mu_n(K) \ge 1 - \eta,$$
$$\liminf_n \mu_n(\{x : w_x(\delta) \le \varepsilon\}) \ge \liminf_n \mu_n(K) \ge 1 - \eta.$$

十分性．まず，任意に固定した $\varepsilon > 0$ に対して，

$$\lim_{\delta \downarrow 0} \sup_n \underbrace{\mu_n(\{x : w_x(\delta) > \varepsilon\})}_{=\psi_n(\delta)} = 0$$

を示す．$\eta > 0$ を given とすると，(ii) より，$\exists \delta(\eta) \in (0,1], \exists n_0(\eta) \in \mathbb{N}$ s.t. $n \geq n_0(\eta) \Rightarrow \psi_n(\delta(\eta)) \leq \eta$．一方，$n$ を固定すると，$\psi_n(\delta) \to 0$ $(\delta \downarrow 0)$ となるから，$\exists \delta_1(\eta) \in (0,1]$ s.t. $\delta \leq \delta_1(\eta) \Rightarrow \psi_n(\eta) \leq \eta$ $(\forall n \leq n_1(\eta))$．従って，$\delta_2(\eta) = \delta(\eta) \vee \delta_1(\eta)$ とおけば，$\delta \leq \delta_2(\eta) \Rightarrow \psi_n(\delta) \leq \eta$ $(\forall n \in \mathbb{N})$．これは，$\lim_{\delta \downarrow 0} \sup_n \psi_n(\delta) = 0$ を意味する．

次に，$\eta > 0$ を given とすると，各 $k \in \mathbb{N}$ に対して，$\exists \delta_k = \delta_k(\eta) \in (0,1]$ s.t.

$$\inf_n \mu_n(\{x : w_x(\delta_k) \leq 1/k\}) \geq 1 - \eta/2^{k+1}.$$

また，(i) より，$\alpha > 0$ を十分大きく選べば，

$$\inf_n \mu_n(\{x : |x(0)| \leq \alpha\}) \geq 1 - \eta/2$$

となる．そこで，

$$K = K_\eta = \left(\bigcap_k \{x : w_x(\delta_k) \leq 1/k\} \right) \cap \{x : |x(0)| \leq \alpha\}$$

とおけば，$\mu_n(K) \geq 1 - \eta$ $(\forall n \in \mathbb{N})$ となる．いま，Arzelà-Ascoli の定理より K は相対コンパクトであり，$x \mapsto x(0), x \mapsto w_x(\delta)$ が連続なことより，K は閉集合であるから，K はコンパクトである．ゆえに，$\{\mu_n\}$ は一様にタイトである． □

Theorem 35.3 の (i) は，$\mu_n \circ \pi_0^{-1}$ が一様にタイトであればみたされるので，$\mu_n \circ \pi_0^{-1}$ が弱収束していればみたされる．従って，次の系を得る．

Corollary 35.1. $\mu_n \in \mathcal{P}(C)$ が弱収束するための必要十分条件は，μ_n の各 f.d.d. が弱収束して，かつ Theorem 35.3 の (ii) が成り立つことである．

この系を分布収束に言い換えて次の系を得る．

Corollary 35.2. C-valued r.v.'s $X_n, n \in \mathbb{N}$ が分布収束するための必要十分条件は，X_n の各 f.d.d. が弱収束して，かつ

$$\lim_{\delta \downarrow 0} \limsup_{n \to \infty} P(w(X_n, \delta) > \varepsilon) = 0, \ \forall \varepsilon > 0 \qquad (*)$$

が成り立つことである．

$(*)$ が成り立つとき，X_n は 漸近的な同程度連続性 (asymptotic equicontinuity) をみたすと言う．

Donsker の不変原理

Theorem 35.3 にもとづいて，Donsker の不変原理を証明しよう．$\{B_t : t \in [0,1]\}$ を $[0,1]$ 上の標準 BM とする．$\{\xi_m\}$ を i.i.d. r.v.'s とし，$E[\xi_m] = 0, E[\xi_m^2] = 1$ とする．$S_0 = 0, S_k = \sum_{m=1}^k \xi_m, k \in \mathbb{N}$ とおき，

$$X_n(t) = \frac{1}{\sqrt{n}} S_{\lfloor nt \rfloor} + (nt - \lfloor nt \rfloor) \frac{\xi_{\lfloor nt \rfloor + 1}}{\sqrt{n}},\ t \in [0,1]$$

と定義する．$X_n(t)$ は $S_k/\sqrt{n}, k = 0, \ldots, n$ を線形につなげたものである．X_n のパス $t \mapsto X_n(t)$ は連続であるから，X_n は C-valued r.v. である．

Theorem 35.4 (Donsker の不変原理)**.** $X_n \xrightarrow{d} \{B_t : t \in [0,1]\}$ in C.

Proof. **Step 1**. まず，任意の $t_1, \ldots, t_k \in [0,1]$ に対して，

$$(X_n(t_1), \ldots, X_n(t_k)) \xrightarrow{d} (B_{t_1}, \ldots, B_{t_k}),\ n \to \infty$$

を示す．各 $t \in [0,1]$ に対して，

$$(nt - \lfloor nt \rfloor) \frac{\xi_{\lfloor nt \rfloor + 1}}{\sqrt{n}} \xrightarrow{P} 0,\ n \to \infty$$

であるから，X_n の各 f.d.d. が多変量正規分布に弱収束することは，多変量 CLT から従う．従って，各 $0 \le s < t \le 1$ に対して，

$$(n^{-1/2} S_{\lfloor ns \rfloor}, n^{-1/2}(S_{\lfloor nt \rfloor} - S_{\lfloor ns \rfloor})) \xrightarrow{d} N\left(\mathbf{0}, \begin{pmatrix} s & 0 \\ 0 & t-s \end{pmatrix}\right)$$

を示せばよいが，これは難しくない．

Step 2. 次に，$E[\xi_1^4] < \infty$ のとき，X_n が漸近的な同程度連続性

$$\lim_{\delta \downarrow 0} \limsup_{n \to \infty} P\{w(X_n, \delta) > \varepsilon\} = 0,\ \forall \varepsilon > 0$$

をみたすことを示す．$\delta \in (0,1)$ を given とすると，$0 = t_0 < t_1 < \cdots < t_v = 1$ を $\min_{1<i<v}(t_i - t_{i-1}) \ge \delta$ をみたすように選べば，

$$w_x(\delta) \le 3 \max_{1 \le i \le v} \sup_{t_{i-1} \le s \le t_i} |x(s) - x(t_{i-1})|$$

であるから，

$$P\{w(X_n, \delta) > 3\varepsilon\} \le \sum_{i=1}^v P\left\{ \sup_{t_{i-1} \le s \le t_i} |X_n(s) - X_n(t_{i-1})| > \varepsilon \right\}.$$

そこで，$m_0, m_1, \ldots, m_v$ を

$$0 = m_0 < m_1 < \cdots < m_v = n,\ (m_i - m_{i-1})/n \geq \delta,\ 1 < i < v \qquad (*)$$

をみたす整数とし，$t_i = m_i/n$ とおけば，$X_n(t)$ は $S_k/\sqrt{n}, k = 0, \ldots, n$ を線形につなげたものだったから，各 $1 \leq i \leq v$ に対して，

$$\sup_{t_{i-1} \leq s \leq t_i} |X_n(s) - X_n(t_{i-1})| \leq \max_{m_{i-1} \leq k \leq m_i} |S_k - S_{m_{i-1}}|/\sqrt{n}$$

となる．ここで，$\{\xi_m\}$ は i.i.d. より，

$$(S_{m_{i-1}+1} - S_{m_{i-1}}, \ldots, S_{m_i} - S_{m_{i-1}}) \stackrel{d}{=} (S_1, \ldots, S_{m_i - m_{i-1}})$$

であるから，

$$P\{w(X_n, \delta) > 3\varepsilon\} \leq \sum_{i=1}^{v} P\left(\max_{0 \leq k \leq m_i - m_{i-1}} |S_k| > \varepsilon\sqrt{n} \right)$$

を得る．いま，$n \geq 2/\delta$ として，

$$m_i = im,\ 0 \leq i < v,\ m_v = n,\ m = \lceil n\delta \rceil,\ v = \lceil n/m \rceil$$

と選べば，$m_0, m_1, \ldots, m_v$ は $(*)$ をみたし，また，$m_v - m_{v-1} \leq m$ である．

$$\begin{aligned} &m \geq n\delta \Rightarrow v \leq \lceil 1/\delta \rceil \leq 1/\delta + 1 \leq 2/\delta, \\ &m \leq n\delta + 1 \Rightarrow n \geq (m-1)/\delta \geq m/(2\delta) \end{aligned}$$

なる評価より，

$$P\{w(X_n, \delta) > 3\varepsilon\} \leq vP\left(\max_{1 \leq k \leq m} |S_k| > \varepsilon\sqrt{n} \right) \leq \frac{2}{\delta} P\left\{ \max_{1 \leq k \leq m} |S_k| > \varepsilon\sqrt{m/(2\delta)} \right\}$$

を得る．$n \to \infty$ のとき，$m \to \infty$ であるから，

$$\limsup_n P\{w(X_n, \delta) > 3\varepsilon\} \leq \frac{2}{\delta} \limsup_n P\left\{ \max_{1 \leq k \leq n} |S_k| > \varepsilon\sqrt{n/(2\delta)} \right\}.$$

右辺の評価に，次の Etemedi の不等式を用いる．

Lemma 35.2 (Etemedi)**.** 任意の $\alpha > 0$ に対して，

$$P\left(\max_{1 \leq k \leq n} |S_k| > \alpha \right) \leq 3 \max_{1 \leq k \leq n} P\left(|S_k| > \alpha/3 \right).$$

Etemedi の不等式の証明は定理の証明のあとに与える．Etemedi の不等式より，

$$P\left\{\max_{1\le k\le n}|S_k|>\varepsilon\sqrt{n/(2\delta)}\right\}\le 3\max_{1\le k\le n}P\left\{|S_k|>(\varepsilon\sqrt{n/(2\delta)})/3\right\}.$$

いま，$E[\xi_1^4]<\infty$ と仮定していたから，ある定数 M が存在して，$E[S_k^4]\le Mk^2$ $(\forall k\in\mathbb{N})$ となる．よって，Markov の不等式より，

$$(\text{右辺})\le 972M\varepsilon^{-4}\delta^2=O(\delta^2),\ \delta\downarrow 0$$

を得る．以上より，求める結論を得る．

Step 3. Step 2 までの結果より，$E[\xi_1^4]<\infty$ のとき，$X_n\xrightarrow{d}\{B_t:t\in[0,1]\}$ in C が示された．$E[\xi_1^4]<\infty$ なる条件を取り除く．正整数 k に対して，$\xi_{k,i}=\xi_i I(|\xi_i|\le k)$ とおき，ξ_i の代わりに $\bar{\xi}_{k,i}=(\xi_{k,i}-E[\xi_{k,i}])/\sqrt{\mathrm{Var}(\xi_{k,i})}$ を代入した X_n を $X_{k,n}$ とおく．$\mathrm{Var}(\xi_{k,i})=\mathrm{Var}(\xi_{k,1})\to\mathrm{Var}(\xi_1)=1$ $(k\to\infty)$ であるから，十分大きな k に対して，$\sqrt{\mathrm{Var}(\xi_{k,1})}\ge 1/2$ となる．よって，$|\bar{\xi}_{k,i}|\le 4k$ であるから，$n\to\infty$ のとき，$X_{k,n}\xrightarrow{d}\{B_t:t\in[0,1]\}$ in C. あとは，任意の $\varepsilon>0$ に対して，

$$\lim_{k\to\infty}\limsup_{n\to\infty}P\{d(X_{k,n},X_n)>\varepsilon\}=0$$

を示せばよい．$S_{k,n}=\sum_{i=1}^n\bar{\xi}_{k,i}$ とおくと，

$$d(X_{k,n},X_n)=\sup_{t\in[0,1]}|X_{n,k}(t)-X_n(t)|=\max_{m\le n}|S_{k,m}-S_m|/\sqrt{n}$$

であり，Etemedi の不等式より，

$$P\left(\max_{1\le m\le n}|S_{k,m}-S_m|>\varepsilon\sqrt{n}\right)\le 3\max_{1\le m\le n}P(|S_{k,m}-S_m|>\varepsilon\sqrt{n}/3).$$

ここで，Markov の不等式より，

$$(\text{右辺})\le 27\varepsilon^{-2}E[(\bar{\xi}_{k,1}-\xi_1)^2]\to 0,\ k\to\infty$$

であるから，求める主張を得る． □

Etemedi の不等式の証明を与える．

Proof of Lemma 35.2. $A_k = \{|S_k| > 3\alpha, |S_j| \le 3\alpha, \forall j < k\}$ とおくと，

$$\begin{aligned}
P\left(\max_{1\le k\le n} |S_k| > 3\alpha\right) &= \sum_{k=1}^{n} P(A_k) \le P(|S_n| > \alpha) + \sum_{k=1}^{n} P(A_k \cap \{|S_n| \le \alpha\}) \\
&= P(|S_n| > \alpha) + \sum_{k=1}^{n} P(A_k \cap \{|S_n - S_k| > 2\alpha\}) \\
&= P(|S_n| > \alpha) + \sum_{k=1}^{n} P(A_k) P(|S_n - S_k| > 2\alpha) \\
&= P(|S_n| > \alpha) + \max_{1\le k\le n} P(|S_n - S_k| > 2\alpha) \\
&\le P(|S_n| > \alpha) + \max_{1\le k\le n} \{P(|S_n| > \alpha) + P(|S_k| > \alpha)\} \\
&\le 3 \max_{1\le k\le n} P(|S_k| > \alpha).
\end{aligned}$$

□

Theorem 35.4 の証明において，標準 BM のアプリオリな存在は必要でないことがわかる．従って，この証明は同時に $[0,1]$ 上の標準 BM が存在することの証明にもなっている．

演習問題

Exercise 35.1. (S,d) を可分な距離空間とする．$f : S \to \mathbb{R}$ に対して，

$$\|f\|_{BL} = \sup_{x\in S} |f(x)| + \sup_{x\ne y} \frac{|f(x) - f(y)|}{d(x,y)}$$

と定義する．また，$\mu, \nu \in \mathcal{P}(S)$ に対して，

$$\beta(\mu,\nu) = \sup_{\|f\|_{BL}\le 1} \left| \int f d\mu - \int f d\nu \right|$$

と定義する．β は bounded Lipschitz 距離または，Fortet-Mourier 距離と呼ばれる．

(i) β が $\mathcal{P}(S)$ 上の距離をなすことを示せ．

(ii) $\mu, \mu_n \in \mathcal{P}(S)$ に対して，$\beta(\mu_n, \mu) \to 0 \Rightarrow \mu_n \xrightarrow{w} \mu$ を示せ．

(iii) 逆に，$\mu_n \xrightarrow{w} \mu \Rightarrow \beta(\mu_n, \mu) \to 0$ を示せ．

コメント：(iii) は Arzelà-Ascoli の定理の応用問題である．

36　D 空間と Skorohod 位相

本節と次節はコンパクトな区間上の càdlàg な関数からなる空間 (D 空間) の上の弱収束を考察する．その主たる応用として，正規化された経験分布関数が Brown 橋に弱収束すること (Donsker の定理) を証明する．D 空間は一様距離に関して可分にならないので，Skorohod 位相と呼ばれる特殊な位相を考える．

D 空間

$$D = \{x \in \mathbb{R}^{[0,1]} : x \text{ は右連続かつ左極限が存在する } \}$$

と定義する．

$$x \in D \Leftrightarrow \begin{cases} \forall t \in [0,1),\ x(t+) := \lim_{s \downarrow t} x(s) = x(t) \\ \forall t \in (0,1],\ x(t-) := \lim_{s \uparrow t} x(s) \text{ が存在する} \end{cases}$$

である．$x \in D, T_0 \subset [0,1]$ に対して，

$$w_x(T_0) = \sup\{|x(s) - x(t)| : s, t \in T_0\}$$

と定義する．x の連続度は，

$$w_x(\delta) = \sup_{0 \le t \le 1-\delta} w_x[t, t+\delta]$$

と書ける．

Lemma 36.1. $x \in D$ であるための必要十分条件は，任意の $\varepsilon > 0$ に対して，

$$0 = t_0 < t_1 < \cdots < t_r = 1$$

なる $t_0, \ldots, t_r$ が存在して，

$$w_x[t_{i-1}, t_i) < \varepsilon,\ 1 \le \forall i \le r \qquad (*)$$

が成り立つことである．

Proof. 必要性．

$$\tau = \sup\{t \in [0,1] : [0,t) \text{ は } (*) \text{ をみたす有限個の部分区間 } [t_{i-1}, t_i) \text{ に分割できる } \}$$

とおくと， $x(0) = x(0+)$ より，$\tau > 0$ である．さらに，$x(\tau-)$ が存在するから，$[0, \tau)$ 自身も $(*)$ をみたす有限個の部分区間に分割できる．仮に $\tau < 1$ なら，$x(\tau) = x(\tau+)$ より，

十分小さな δ に対して，$w_x[\tau, \tau+\delta) < \varepsilon$ となるから，$[0, \tau+\delta)$ も $(*)$ をみたす有限個の部分区間に分割できる．しかしこれは τ の定義に反するから，$\tau = 1$ である．

十分性．x がある $t \in [0,1]$ で右連続でないか左極限が存在しないなら，

$$\varepsilon := \begin{cases} \max\{|x(0) - \liminf_{s\downarrow 0} x(s)|, |x(0) - \limsup_{s\downarrow 0} x(s)|\} & \text{if } t = 0 \\ \max\Big\{|x(t) - \liminf_{s\downarrow t} x(s)|, |x(t) - \limsup_{s\downarrow t} x(s)|, \\ \qquad \limsup_{s\uparrow t} x(s) - \liminf_{s\uparrow t} x(s)\Big\} & \text{if } t \in (0,1) \\ \limsup_{s\uparrow 1} x(s) - \liminf_{s\uparrow 1} x(s) & \text{if } t = 1 \end{cases}$$

とおくと，$\varepsilon > 0$ であり，このときどのように $t_0, \dots, t_r$ を選んできても，

$$\max_{1\le i\le r} w_x[t_{i-1}, t_i) \ge \varepsilon$$

となる． □

Remark 36.1. $x \in D$ とする．$(*)$ をみたす $0 = t_0 < t_1 < \cdots < t_r = 1$ を選ぶ．

- $\mathrm{Card}\{t : |x(t) - x(t-)| \ge \varepsilon\} \le r < \infty$ より，x の不連続点は高々可算個 である．
- $\sup_{t\in[0,1]} |x(t)| \le \max\{|x(t_0)|, \dots, |x(t_r)|\} + \varepsilon < \infty$ であるから，x は有界 である．
- $x_\varepsilon = \sum_{i=1}^r x(t_{i-1}) 1_{[t_{i-1}, t_i)} + x(1) 1_{\{1\}}$ とすると，$\sup_t |x(t) - x_\varepsilon(t)| < \varepsilon$ であるから，x は単関数により一様に近似できる．特に，x は可測関数 である．

いま，$\delta \in (0,1)$ に対して，

$$w'_x(\delta) = w'(x, \delta) = \inf_{\{t_i\}} \max_{1\le i\le r} w_x[t_{i-1}, t_i)$$

と定義する．ただし，$\inf_{\{t_i\}}$ は

$$\begin{cases} 0 = t_0 < t_1 < \cdots < t_r = 1, \\ t_i - t_{i-1} > \delta,\ 1 \le i \le r \end{cases}$$

をみたす $\{t_i\}$ に関して下限をとることを意味する．このとき，

$$x \in D \Leftrightarrow \lim_{\delta\downarrow 0} w'_x(\delta) = 0.$$

次に，$w'_x(\delta)$ と $w_x(\delta)$ との関係を調べる．$\delta \in (0, 1/2)$ に対して，$[0,1)$ は $\delta < t_i - t_{i-1} \le 2\delta$ なる区間 $[t_{i-1}, t_i)$ に分割できる．このとき，

$$w_x[t_{i-1}, t_i) \le \sup_{0\le t\le 1-2\delta} w_x[t, t+2\delta] = w_x(2\delta)$$

であるから，$w'_x(\delta) \le w_x(2\delta)$ である．

$x \in C$ のときは．逆の不等式も成り立つ．$\{t_i\}$ を

$$\max_i w_x[t_{i-1}, t_i) < w'_x(\delta) + \varepsilon$$

をみたすように選ぶ．$s \le t, t-s < \delta$ とすると，(i) $s,t \in [t_{i-1}, t_i)$ or (ii) $s \in [t_{i-1}, t_i), t \in [t_i, t_{i+1})$ ($i = r$ のときは，$t = 1$) の 2 通りの可能性がある．(i) のとき，

$$\begin{aligned} |x(s) - x(t)| &\le |x(s) - x(t_{i-1})| + |x(t_{i-1}) - x(t)| \\ &\le 2w_x[t_{i-1}, t_i) < 2w'_x(\delta) + 2\varepsilon. \end{aligned}$$

一方，(ii) のとき，

$$\begin{aligned} |x(s) - x(t)| &\le |x(s) - x(t_i)| + |x(t_i) - x(t)| \\ &\le w_x[t_{i-1}, t_i] + w_x[t_i, t_{i+1}) \\ &= w_x[t_{i-1}, t_i) + w_x[t_i, t_{i+1}) \quad (\because x \in C \Rightarrow w_x[t_{i-1}, t_i] = w_x[t_{i-1}, t_i)) \\ &< 2w'_x(\delta) + 2\varepsilon. \end{aligned}$$

従って，$x \in C \Rightarrow w_x(\delta) \le 2w'_x(\delta)$.

Skorohod 位相

D の距離を考察しよう．各 $x \in D$ は有界であるから，D には一様距離

$$d_{\sup}(x, y) = \sup_{t \in [0,1]} |x(t) - y(t)|$$

を入れることができる．$(D, d_{\sup})$ は完備である (演習問題).

しかし，$d_{\sup}$ は D において “よい” 距離とは言えない．まず，$(D, d_{\sup})$ は完備であるが，可分ではない．例えば，$0 < \theta < 1$ に対して，

$$x_\theta(t) = \begin{cases} 0 & \text{if } 0 \le t < \theta \\ 1 & \text{if } \theta \le t \le 1 \end{cases}$$

とおくと，$d_{\sup}(x_\theta, x_{\theta'}) = 1,\ \theta \ne \theta'$ であるから，$(D, d_{\sup})$ は可分でない．さらに，$\mathcal{U}$ を一様距離に関する D の Borel σ-field とすると，経験分布関数が Ω から D への写像として一般に $\mathcal{F}/\mathcal{U}$ 可測とならない．

Example 36.1. $\xi_1, \ldots, \xi_n \sim U[0,1] =: \lambda$ i.i.d. として，経験分布関数を

$$F_n(t, \omega) = \frac{1}{n} \sum_{i=1}^{n} 1_{[0,t]}(\xi_i(\omega)),\ t \in [0,1], \omega \in \Omega$$

と定義する．確率空間を $(\Omega, \mathcal{F}, P) = ([0,1]^n, \mathcal{B}([0,1]^n), \lambda^n)$ とし，ξ_i を座標写像 $\xi_i(\omega) = \omega_i, \omega = (\omega_1, \ldots, \omega_n)$ にとる．このとき，$\omega \mapsto F_n(\cdot, \omega)$, $\Omega \to D$ は $\mathcal{F}/\mathcal{U}$ 可測でない．簡単のために，$n = 1$ とし，$X(t, \omega) = 1_{[0,t]}(\xi_1(\omega)) = 1_{[\omega,1]}(t)$ とおく．

$$H_\theta = \left\{ x \in D : d_{\sup}(x, 1_{[\theta,1]}) < 1/2 \right\}$$

とおくと，あらゆる $A \subset [0,1]$ に対して，$\bigcup_{\theta \in A} H_\theta$ は $d_{\sup}$ の開集合である．しかし，$X(\cdot, \omega) \in H_\theta \Leftrightarrow \omega = \theta$ であるから，A を $[0,1]$ の Lebesgue 非可測集合とすると，

$$\left\{ \omega : X(\cdot, \omega) \in \bigcup_{\theta \in A} H_\theta \right\} = A \notin \mathcal{F}$$

となる．さらに，$\mathcal{F}$ を P に関して完備化しても，すなわち，$\mathcal{F}$ を $[0,1]$ 上の Lebesgue 可測集合の全体としても，左辺は $\mathcal{F}$ に属さない．

確率空間をもっと広くとったとしても，仮に $\omega \mapsto X(\cdot, \omega)$ が $\mathcal{F}/\mathcal{U}$ 可測になるなら，$\mu = P \circ \xi_1^{-1}$ は $2^{[0,1]}$ 上に定義された測度であって，$\mu((a,b]) = b - a$ なる性質をみたす必要があるが，連続体仮説を仮定すると，そのような測度は存在しないことが知られている (詳しくは，(Dudley, 2002, Appendix C) を参照せよ)．

以上の考察から，D 上には，一様距離とは異なる，都合のよい距離を考える必要がある．

$$\Lambda = \{ \lambda : [0,1] \to [0,1] : \lambda \text{ は狭義増加，連続かつ全射 } \}$$

とおく．$\lambda \in \Lambda$ なら，$\lambda(0) = 0, \lambda(1) = 1$ である．$x, y \in D$ に対して，

$$d(x,y) = \inf\{ \varepsilon > 0 : \exists \lambda \in \Lambda \text{ s.t. } \sup_t |\lambda(t) - t| \le \varepsilon \ \&\ \sup_t |x(t) - y \circ \lambda(t)| \le \varepsilon \}$$

と定義する．$\lambda(t) = t$ とすれば，

$$d(x,y) \le d_{\sup}(x,y)$$

であり，特に $d(x,y)$ は有限である．次に，d が距離であることをチェックする．

Lemma 36.2. *d は D 上の距離である．*

Proof. $d \ge 0$ はよい．$d(x,y) = 0$ ならば，$\exists \lambda_n \in \Lambda$ s.t. $\sup_t |\lambda_n(t) - t| \to 0$ & $\sup_t |x(t) - y \circ \lambda_n(t)| \to 0$. y の不連続点は高々可算個しかないことと，x, y の右連続性より，$x = y$ を得る．次に，$\lambda \in \Lambda \Rightarrow \lambda^{-1} \in \Lambda$ であることと，$\sup_t |\lambda^{-1}(t) - t| = \sup_t |\lambda(t) - t|, \sup_t |x \circ \lambda^{-1}(t) - y(t)| = \sup_t |x(t) - y \circ \lambda(t)|$ より，$d(x,y) = d(y,x)$ を得る．最後に，$\lambda_1, \lambda_2 \in \Lambda \Rightarrow \lambda_2 \circ \lambda_1 \in \Lambda$ であることと，

$$\sup_t |\lambda_2 \circ \lambda_1(t) - t| \le \underbrace{\sup_t |\lambda_2 \circ \lambda_1(t) - \lambda_1(t)|}_{= \sup_t |\lambda_2(t) - t|} + \sup_t |\lambda_1(t) - t|,$$

$$\sup_t |x(t) - z \circ \lambda_2 \circ \lambda_1(t)| \le \sup_t |x(t) - y \circ \lambda_1(t)| + \underbrace{\sup_t |y \circ \lambda_1(t) - z \circ \lambda_2 \circ \lambda_1(t)|}_{= \sup_t |y(t) - z \circ \lambda_2(t)|}$$

より，$d(x,z) \le d(x,y) + d(y,z)$ を得る．以上より，d は距離である． □

この d を Skorohod 距離 と呼び，d の定める位相を Skorohod 位相 と呼ぶ.

$$d(x_n, x) \to 0 \Leftrightarrow \exists \lambda_n \in \Lambda \text{ s.t. } \lambda_n(t) \to t,\ x_n \circ \lambda_n(t) \to x(t) \text{ unif. in } t$$

である．特に，$x_n \in C$ なら，$x_n \circ \lambda_n \in C$ であるから，C は Skorohod 位相に関して閉じている．以下，d の性質を考察していく．

(a) $d_{\sup}(x_n, x) \to 0 \Rightarrow d(x_n, x) \to 0$ であるが，逆は正しくない．$x_n = 1_{[0,1/2+1/n)}$ として，λ_n を $[0, 1/2]$ 上で 0 から $1/2 + 1/n$ まで線形に増加し，$[1/2, 1]$ 上で $1/2 + 1/n$ から 1 まで線形に増加する関数とすると，

$$x_n \circ \lambda_n = 1_{[0,1/2)},\ |\lambda_n(t) - t| \le 1/n$$

であるから，$d(x_n, 1_{[0,1/2)}) \to 0$ であるが，$x_n(1/2) \to 1 \neq 0 = 1_{[0,1/2)}(1/2)$.

(b) $d(x_n, x) \to 0$ なら，x のすべての連続点 t に対して，$x_n(t) \to x(t)$. また，$d(x_n, x) \to 0,\ x \in C$ なら，$d_{\sup}(x_n, x) \to 0$. 特に，Skorohod 位相から決まる C の相対位相は一様距離に関する位相に一致する．

Proof. $d(x_n, x) \to 0$ より，$\exists \lambda_n \in \Lambda$ s.t. $\lambda_n(t) \to t, x_n \circ \lambda_n(t) \to x$ unif. in $t \in [0, 1]$ である．よって，t が x の連続点なら，

$$|x_n(t) - x(t)| \le |x_n(t) - x \circ \lambda_n^{-1}(t)| + |x \circ \lambda_n^{-1}(t) - x(t)| \to 0.$$

$x \in C$ なら，右辺の収束は $t \in [0, 1]$ に関して一様であるから，$d_{\sup}(x_n, x) \to 0$ となる． □

(c) (D, d) は可分である.

Proof. まず，次の補題を証明する.

Lemma 36.3. $0 = s_0 < s_1 < \cdots < s_k = 1$ なる $\sigma = \{s_u\}$ に対して，$A_\sigma : D \to D$ を

$$A_\sigma : x \mapsto \sum_{u=1}^{k} x(s_{u-1}) 1_{[s_{u-1}, s_u)} + x(1) 1_{\{1\}}$$

と定める．このとき，$\max_u (s_u - s_{u-1}) \le \delta \Rightarrow d(A_\sigma(x), x) \le \delta \vee w'_x(\delta)$.

Proof of Lemma 36.3. $t \in [0, 1]$ に対して，$\zeta(t) = s_{u-1}$ for $t \in [s_{u-1}, s_u),\ \zeta(1) = 1$ と定義すると，$A_\sigma(x)(t) = \zeta(t)$. $x \in D$ のとき，$\exists \{t_i\}$ s.t. $\min_i (t_i - t_{i-1}) > \delta,\ \max_i w_x[t_{i-1}, t_i) < w'_x(\delta) + \varepsilon$. $\lambda(t_i) = s_v$ for $t_i \in (s_{v-1}, s_v],\ \lambda(t_0) = s_0 = 0$ と選べば，$t_i - t_{i-1} > \delta \ge s_v - s_{v-1}$ より，t_i, t_{i+1} は同じ区間 $(s_{v-1}, s_v]$ に入ることはないので，$\lambda(t_i)$ は i に関して狭義単調増加である．$\lambda(t_i)$ を線形につなげた $[0, 1]$ から $[0, 1]$ への関数を λ とすると，$\sup_t |\lambda(t) - t| \le \delta$. あとは，$|A_\sigma(x)(t) - x \circ \lambda^{-1}(t)| = |x \circ \zeta(t) - x \circ \lambda^{-1}(t)| \le w'_x(\delta) + \varepsilon$ を示せばよい．$t = 0$ or 1 なら明らかなので，$0 < t < 1$ とする．このとき，$\zeta(t)$ と $\lambda^{-1}(t)$ が同じ区間 $[t_{i-1}, t_i)$ に入ることを示せばよいが，これをチェックするのは容易である． □

(D,d) の可分性の証明に戻ると，この補題より，

$$\left\{\sum_{i=1}^{k} q_i 1_{[(i-1)/k,i/k)} + q_{k+1}1_{\{1\}} : q_1,\dots,q_{k+1} \in \mathbb{Q}, k \in \mathbb{N}\right\}$$

は (D,d) で稠密である． □

(d) $\mathcal{B}(D)$ を D の Skorohod 位相に関する Borel σ-field とする．また，各 $t_1,\dots,t_k \in [0,1]$ に対して，射影を

$$\pi_{t_1,\dots,t_k} : x \mapsto (x(t_1),\dots,x(t_k)),\ D \to \mathbb{R}^k$$

と定義する．$\mathcal{D}$ を各 $\pi_{t_1,\dots,t_k}$ が可測になる D 上の最小の σ-field とする：

$$\mathcal{D} = \sigma(\{\pi^{-1}_{t_1,\dots,t_k}(A) : A \in \mathcal{B}^k, t_1,\dots,t_k \in [0,1], k \in \mathbb{N}\}).$$

このとき，$\mathcal{B}(D) = \mathcal{D}$ である．

Proof. $\mathcal{D} = \sigma(\pi_t, t \in [0,1])$ である．まず，$\mathcal{D} \subset \mathcal{B}(D)$ を示す．各 π_t が $\mathcal{D}$ 可測であることを示せばよい．$d(x_n,x) \to 0$ のとき，

$$|x_n(t) - x(t)| \le |x_n(t) - x \circ \lambda_n^{-1}(t)| + |x \circ \lambda_n^{-1}(t) - t|$$

より，t が x の連続点であれば，$x_n(t) \to x(t)$. x の不連続点は高々可算個しかないから，$x_n(t) \to x(t)$ a.e. $t \in [0,1]$. また，十分大きな n に対して，$\sup_t |x_n(t)| \le \sup_t |x(t)| + 1$ であるから，$\{x_n\}$ は一様有界である．従って，各 $t \in [0,1)$ と $\varepsilon \in (0, 1-t]$ に対して，

$$\frac{1}{\varepsilon}\int_t^{t+\varepsilon} x_n(s)ds \to \frac{1}{\varepsilon}\int_t^{t+\varepsilon} x(s)ds,\ n \to \infty$$

であるから，$x \mapsto \varepsilon^{-1}\int_t^{t+\varepsilon} x(s)ds$ は連続であり，特に $\mathcal{B}(D)$ 可測である．各 $x \in D$ は右連続であるから，$\varepsilon = \varepsilon_m \downarrow 0$ として，$\pi_t(x) = x(t) = \lim_m \varepsilon_m^{-1}\int_t^{t+\varepsilon_m} x(s)ds$ を得る．ゆえに，π_t は $\mathcal{B}(D)$ 可測である．また，π_1 は連続であるから，$\mathcal{B}(D)$ 可測である．

次に，$\mathcal{B}(D) \subset \mathcal{D}$ を示す．$m \in \mathbb{N}$ に対して，$0 = s_0 < s_1 < \cdots < s_k = 1$ を $\max_u(s_u - s_{u-1}) < m^{-1}$ となるように選び，$\sigma = \{s_u\}$ とする．このとき，

$$E_m : (\alpha_0,\alpha_1,\dots,\alpha_k) \mapsto \sum_{u=1}^{k} \alpha_{u-1}1_{[s_{u-1},s_u)} + \alpha_k 1_{\{1\}},\ \mathbb{R}^{k+1} \to D$$

は連続であるから，$\mathcal{B}^{k+1}/\mathcal{B}(D)$ 可測である．また，

$$\pi_{s_0,\dots,s_k} : x \mapsto (x(s_0),\dots,x(s_k)),\ D \to \mathbb{R}^{k+1}$$

は $\mathcal{D}/\mathcal{B}^{k+1}$ 可測であるから，

$$E_m \circ \pi_{s_0,\dots,s_k} : x \to \sum_{u=1}^{k} x(s_{u-1})1_{[s_{u-1},s_u)} + x(s_k)1_{\{1\}} = A_\sigma(x),\ D \to D$$

は $\mathcal{D}/\mathcal{B}(D)$ 可測である．$m \to \infty$ のとき，各 $x \in D$ に対して，$d(A_\sigma(x), x) \to 0$ となるから，恒等写像 $x \mapsto x$ は $\mathcal{D}/\mathcal{B}(D)$ 可測である．これは，$\mathcal{B}(D) \subset \mathcal{D}$ を意味する． □

Remark 36.2. ($\mathcal{B}(D)$ に関する注意).

- $X : \Omega \to D$ が D-valued r.v. (すなわち，$\mathcal{F}/\mathcal{B}(D)$ 可測) であるための必要十分条件は，各 $X(t)$ が r.v. になることである．
- $(D, \mathcal{B}(D))$ 上の各 p.m. はその f.d.d. たちによって一意に決まる．
- $x \in D$ に対して，$w_x(\delta) = \sup\{|x(s) - x(t)| : |s - t| \le \delta, s, t \in \mathbb{Q} \cap [0, 1]\}$ と表せるので [22]，$x \mapsto w_x(\delta)$ は $\mathcal{B}(D)$ 可測である．
- Example 36.1 の議論より，$\mathcal{D} \subsetneq \mathcal{U}$ であるから，$\mathcal{B}(D)$ は $\mathcal{U}$ より真に小さい．

(e) しかし，(D, d) は (距離空間として) 完備でない．$x_n = 1_{[0,2^{-n})}$ とすると，$d(x_n, x_{n+1}) \le 2^{-n}$ より，x_n は Cauchy 列である．一方，$t \ne 0$ に対して，$x_n(t) \to 0$ であるから，x_n が仮に収束するとしたら，収束先は $x \equiv 0$ だが，$d(x_n, x) = 1$ である．

Skorohod 位相の位相的完備性

(D, d) は完備ではないが，d と同値な距離 d_0 であって，d_0 の下で D が完備となるようなものが存在する．そのような距離 d_0 を構成する．なお，弱収束は距離に位相を通してしか依存しないので，例えば，Prohorov の定理の (ii) は $(D, \mathcal{B}(D))$ 上の p.m.'s に対しても成り立つ．

$\lambda \in \Lambda$ に対して，

$$\|\lambda\| = \sup_{s \ne t, s, t \in [0,1]} \left| \log \frac{\lambda(t) - \lambda(s)}{t - s} \right|$$

と定義する．このとき，

$$d_0(x, y) = \inf\{\varepsilon > 0 : \exists \lambda \in \Lambda \text{ s.t. } \|\lambda\| \le \varepsilon \ \&\ \sup_t |x(t) - y \circ \lambda(t)| \le \varepsilon\}$$

と定義する．いま，$|u - 1| \le e^{|\log u|} - 1$ $(u > 0)$ より，

$$\sup_t |\lambda(t) - t| = \sup_t t \left| \frac{\lambda(t) - \lambda(0)}{t - 0} - 1 \right| \le e^{\|\lambda\|} - 1.$$

ここで，$u \le e^u - 1$ より，

$$d(x, y) \le e^{d_0(x,y)} - 1 \qquad (*)$$

[22] $\{(s, t) \in [0, 1]^2 : |s - t| \le \delta\}$ の図を描いてみよ．

を得る．特に，$d_0(x,y)=0$ なら，$d(x,y)=0$ であるから，$x=y$ である．また，

$$\|\lambda^{-1}\| = \sup_{s\neq t}\left|\log\frac{t-s}{\lambda(t)-\lambda(s)}\right| = \sup_{s\neq t}\left|-\log\frac{\lambda(t)-\lambda(s)}{t-s}\right| = \|\lambda\|,$$
$$\|\lambda_2\circ\lambda_1\| \le \sup_{s\neq t}\left|\log\frac{\lambda_2\circ\lambda_1(t)-\lambda_2\circ\lambda_1(s)}{\lambda_1(t)-\lambda_1(s)}\right| + \sup_{s\neq t}\left|\log\frac{\lambda_1(t)-\lambda_1(s)}{t-s}\right| = \|\lambda_2\|+\|\lambda_1\|$$

より，d_0 は対称性と三角不等式もみたすので，d_0 が D 上の距離であることが示された．次に d_0 と d が同値であることを示す．

$$d(x_n,x)\to 0 \Leftrightarrow d_0(x_n,x)\to 0$$

を示せばよいが，$(*)$ より，$d_0(x_n,x)\to 0 \Rightarrow d(x_n,x)\to 0$ は直ちに従う．

Lemma 36.4. $d(x,y)<\delta^2, 0<\delta<1/4 \Rightarrow d_0(x,y)\le 4\delta + w'_x(\delta)$.

Proof. $\{t_i\}$ を $\min_i(t_i-t_{i-1})>\delta$, $\max_i w_x[t_{i-1},t_i) < w'_x(\delta)+\delta$ をみたすように選ぶ．次に，$d(x,y)<\delta^2$ より，$\exists\lambda\in\Lambda$ s.t. $\sup_t|\lambda(t)-t|<\delta^2$, $\sup_t|x(t)-y\circ\lambda(t)|<\delta^2$．このとき，$\tilde\lambda\in\Lambda$ を，$\tilde\lambda(t_i)=\lambda(t_i)$ をみたし，あとは線形につないだものとする．ここで，$\lambda^{-1}\circ\tilde\lambda(t_i)=t_i$ より，t と $\lambda^{-1}\circ\tilde\lambda(t)$ は同じ区間 $[t_{i-1},t_i)$ に入る．従って，

$$\begin{aligned}|x(t)-y\circ\tilde\lambda(t)| &\le \underbrace{|x(t)-x\circ\lambda^{-1}\circ\tilde\lambda(t)|}_{\le w_x[t_{i-1},t_i)} + |x\circ\lambda^{-1}\circ\tilde\lambda(t)-y\circ\tilde\lambda(t)|\\ &< w'_x(\delta)+\delta+\delta^2 < 4\delta+w'_x(\delta).\end{aligned}$$

あとは，$\|\lambda\|<4\delta$ を示せば十分である．まず，$\tilde\lambda(t_i)=\lambda(t_i)$ であるから，

$$|\tilde\lambda(t_i)-\tilde\lambda(t_{i-1})-(t_i-t_{i-1})|<2\delta^2<2\delta(t_i-t_{i-1}) \quad (\because\ t_i-t_{i-1}>\delta).$$

$\tilde\lambda$ は区分的に線形なので，任意の $s,t\in[0,1]$ に対して，

$$|\lambda(t)-\lambda(t)-(t-s)|\le 2\delta|t-s|$$

となる．ゆえに，$s\neq t$ ならば，

$$\left|\frac{\lambda(t)-\lambda(s)}{t-s}-1\right|\le 2\delta<1/2$$

であるから，$\|\lambda\|\le\max\{-\log(1-2\delta),\log(1+2\delta)\}\le 2\delta+4\delta^2<4\delta$. □

従って，次の結論を得る．

Theorem 36.1. d_0 と d は同値な距離である．

最後に，(D,d_0) が完備であることを示す．

Theorem 36.2. (D, d_0) は完備である．

Proof. $\{x_n\}$ を d_0 に関する Cauchy 列とすると，$\exists n_j \in \mathbb{N}$ s.t. $d_0(x_{n_j}, x_{n_{j+1}}) < 2^{-j}$. d_0 の定義より，各 j に対して，$\exists \lambda_j \in \Lambda$ s.t. $\sup_t |x_{n_j}(t) - x_{n_{j+1}} \circ \lambda_j(t)| \leq 2^{-j}$, $\|\lambda_j\| \leq 2^{-j}$. $(*)$ より，$|\lambda_j(t) - t| \leq e^{\|\lambda_j\|} - 1 \leq 2^{-j+1}$ であるから，

$$\sup_t |\lambda_{n+m+1} \circ \cdots \circ \lambda_n(t) - \lambda_{n+m} \circ \cdots \circ \lambda_n(t)| = \sup_t |\lambda_{n+m+1}(t) - t| \leq 2^{-(m+n)}.$$

従って，各 n に対して，$\{\lambda_{m+n} \circ \cdots \circ \lambda_n\}_{m=1}^{\infty}$ は C の Cauchy 列であるから，C の完備性より，$\exists \nu_n \in C$ s.t. $\lambda_{m+n} \circ \cdots \circ \lambda_n(t) \to \nu_n(t)$ unif. in $t \in [0,1]$. $\nu_n(0) = 0, \nu_n(1) = 1$ であり，ν_n は非減少である．また，

$$\|\lambda_{m+n} \circ \cdots \circ \lambda_n\| \leq \|\lambda_{m+n}\| + \cdots + \|\lambda_n\| \leq 2^{-n+1}.$$

ゆえに，$\nu_n \in \Lambda$ & $\|\nu_n\| \leq 2^{-n+1}$ となる．

いま，$\nu_n = \nu_{n+1} \circ \lambda_n$ であるから，

$$\sup_t |x_{n_j} \circ \nu_j^{-1}(t) - x_{n_{j+1}} \circ \nu_{j+1}^{-1}(t)| = \sup_t |x_{n_j}(t) - x_{n_{j+1}} \circ \lambda_j(t)| \leq 2^{-j}.$$

ゆえに，$\{x_{n_j} \circ \nu_j^{-1}\}$ は $d_{\sup}$ に関する Cauchy 列であるから，$(D, d_{\sup})$ の完備性より，$\exists x \in D$ s.t. $\sup_t |x_{n_j} \circ \nu_j^{-1}(t) - x(t)| = \sup_t |x_{n_j}(t) - x \circ \nu_j(t)| \to 0$. 一方，$\|\nu_j\| \to 0$ であるから，$d_0(x_{n_j}, x) \to 0$ を得る．$\{x_n\}$ は d_0 に関する Cauchy 列であったから，これは $d_0(x_n, x) \to 0$ を意味する． □

演習問題

Exercise 36.1. $(D, d_{\sup})$ が完備であることを示せ．

Exercise 36.2. X, Y を D-valued r.v.'s とし，T を 1 を含む $[0,1]$ の稠密な部分集合とする．このとき，各 $t_1, \ldots, t_k \in T$ に対して，

$$(X(t_1), \ldots, X(t_k)) \stackrel{d}{=} (Y(t_1), \ldots, Y(t_k))$$

ならば，$X \stackrel{d}{=} Y$ となることを示せ．

37　D 空間における弱収束

$(D, \mathcal{B}(D))$ 上の p.m.'s の全体を $\mathcal{P}(D)$ と書く．$\mu \in \mathcal{P}(D)$ に対して，

$$T_\mu = \{t \in [0,1] : D \ni x \mapsto \pi_t(x) \text{ は } \mu\text{-a.e. に連続 }\}$$

とおく．π_0 と π_1 は常に連続なので，$\{0,1\} \subset T_\mu$ である．一方，$0 < t < 1$ に対して，π_t が x で連続であるための必要十分条件は，x が t で連続なことであるから，t をジャンプ点にもつ $x \in D$ の全体を $J_t = \{x \in D : x(t) \neq x(t-)\}$ とおくと，

$$t \in T_\mu \Leftrightarrow \mu(J_t) = 0.$$

$\mu(J_t) > 0$ なる t が高々可算個しかないことを示す．まず，任意に固定した $\varepsilon, \delta > 0$ に対して，$J_t(\varepsilon) = \{x \in D : |x(t) - x(t-)| > \varepsilon\}$ とおくと，$\mu(J_t(\varepsilon)) \geq \delta$ なる t は有限個しかない．実際，仮にそのような t が無限個 $\{t_n\}$ あるならば，$\mu(J_{t_n}(\varepsilon) \text{ i.o.}) = \lim_n \mu(\bigcup_{m=n}^\infty J_{t_m}(\varepsilon)) \geq \limsup_n \mu(J_{t_n}(\varepsilon)) \geq \delta$ となるから，ある $x \in D$ は ε より大きいジャンプを無限個もつことになるが，これは起こりえない．$\varepsilon \downarrow 0$ とすると，$\mu(J_t(\varepsilon)) \uparrow \mu(J_t)$ となるから，

$$\begin{aligned}\{t : \mu(J_t) > 0\} &= \bigcup_{\delta > 0, \delta \in \mathbb{Q}} \{t : \mu(J_t) > \delta\} \\ &= \bigcup_{\delta > 0, \delta \in \mathbb{Q}} \bigcup_{\varepsilon > 0, \varepsilon \in \mathbb{Q}} \{t : \mu(J_t(\varepsilon)) > \delta\}.\end{aligned}$$

よって，$\{t : \mu(J_t) > 0\}$ は高々可算である．以上の議論をまとめて，次の補題を得る．

Lemma 37.1. $\mu \in \mathcal{P}(D)$ に対して，T_μ は 0 と 1 を含み，$[0,1] \setminus T_\mu$ は高々可算である．

次の定理は，Prohorov の定理の系から従う．

Theorem 37.1. $\mu, \mu_n \in \mathcal{P}(D)$ とする．$\mu_n \xrightarrow{w} \mu$ となるための必要十分条件は，任意の $t_1, \ldots, t_k \in T_\mu, k \in \mathbb{N}$ に対して，$\mu_n \circ \pi_{t_1,\ldots,t_k}^{-1} \xrightarrow{w} \mu \circ \pi_{t_1,\ldots,t_k}^{-1}$ となることと，$\{\mu_n\}$ が一様にタイトであることである．

Proof. 必要性．任意の $t_1, \ldots, t_k \in T_\mu$ に対して，$\mu(\{x : \pi_{t_1,\ldots,t_k} \text{ は } x \text{ で不連続 }\}) = 0$ であるから，連続写像定理より，$\mu_n \circ \pi_{t_1,\ldots,t_k}^{-1} \xrightarrow{w} \mu \circ \pi_{t_1,\ldots,t_k}^{-1}$.

十分性．$\mu_{n_m} \xrightarrow{w} \nu$, $m \to \infty$ ならば，$\nu = \mu$ となることを示せばよい．いま，任意の $t_1, \ldots, t_k \in T_\mu \cap T_\nu$ に対して，$\mu \circ \pi_{t_1,\ldots,t_k}^{-1} = \nu \circ \pi_{t_1,\ldots,t_k}^{-1}$. $T_\mu \cap T_\nu$ は 1 を含む $[0,1]$ の稠密な部分集合であるから，$\mu = \nu$ である．　□

Remark 37.1. μ が C に集中している $(\mu(C) = 1)$ なら，$T_\mu = [0,1]$ であるから，$\mu_n \xrightarrow{w} \mu$ となるための必要十分条件は，μ_n の各 f.d.d. が μ のそれに弱収束し，かつ $\{\mu_n\}$ が一様にタイトなことである．

次に，D において一様なタイトネスを保証する条件を考える．まず，Arzelà-Ascoli の定理に対応する結果を証明する．

Theorem 37.2. $A \subset D$ が Skorohod 位相に関して相対コンパクトであるための必要十分条件は，

$$\sup_{x\in A}\sup_{t\in[0,1]}|x(t)| < \infty$$

と

$$\lim_{\delta\downarrow 0}\sup_{x\in A} w'_x(\delta) = 0$$

が成り立つことである．

Proof. (十分性). (D, d_0) は完備であるから，A が全有界であることを示せばよい．$\delta > 0$ を given とし，k を十分大きく選んで，$1/k < \delta$, $w'_x(1/k) < \delta$ $(\forall x \in A)$ とする．$M \in \mathbb{N}$ を $M \geq \sup_{x\in A}\sup_t |x(t)|$ となるように選び，

$$\mathcal{N}_\delta = \left\{ \sum_{i=1}^{k} \alpha_{i-1} 1_{[(i-1)/k, i/k)} + \alpha_k 1_{\{1\}} : \alpha_0, \alpha_1, \ldots, \alpha_k \in \{-M + j/k : j = 0, 1, \ldots, 2kM\} \right\}$$

とおく．ここで，任意の $x \in A$ に対して，$d(x, A_\sigma x) \leq (1/k) \vee w'_x(1/k) < \delta$ である．ただし，$\sigma = \{i/k : i = 0, 1, \ldots, k\}$．さらに，$\mathcal{N}_\delta$ の構成より，$\exists y \in \mathcal{N}_\delta$ s.t. $d(A_\sigma x, y) \leq d_{\sup}(A_\sigma x, y) < \delta$. ゆえに，$d(x, y) \leq d(x, A_\sigma x) + d(A_\sigma x, y) < 2\delta$ (つまり $\mathcal{N}_\delta$ は A の d に関する (2δ) ネットになっている)．

いま，任意の $\varepsilon > 0$ に対して，$\delta = \delta_\varepsilon \in (0, \varepsilon)$ を十分小さく選んで，$w'_x(\delta) < \varepsilon$ $(\forall x \in A)$ とする．このとき，$\forall x \in A, \exists y \in \mathcal{N}_{\delta^2/2}$ s.t. $d(x, y) < \delta^2$. ゆえに，この y に対して，$d_0(x, y) \leq 4\delta + w'_x(\delta) < 5\varepsilon$. □

必要性の証明に，次の Dini の定理を用いる．(S, d) を距離空間とすると，$f : S \to \overline{\mathbb{R}}$ が上半連続であるとは，任意の $x \in S$ と $x_n \to x$ なる任意の $\{x_n\}$ に対して，$\limsup_n f(x_n) \leq f(x)$ となることを言う．f が上半連続であるためには，任意の $\alpha \in \mathbb{R}$ に対して，$\{f < \alpha\}$ が開集合であることが必要十分である．

Theorem 37.3 (Dini)**.** $f_n : S \to \overline{\mathbb{R}}, n \in \mathbb{N}$ を上半連続とし，$f : S \to \mathbb{R}$ を連続とする．このとき，$f_n(x) \downarrow f(x)$ $(n \to \infty, \forall x \in S)$ なら，f_n は f に広義一様収束する，i.e., 任意のコンパクト集合 $K \subset S$ に対して，$\sup_{x\in K}\{f_n(x) - f(x)\} \to 0$.

Proof. $f_n - f$ は上半連続であるから，$f \equiv 0$ としてよい．各 $\varepsilon > 0$ に対して，$G_n = \{x \in S : f_n(x) < \varepsilon\}$ は開集合であり，$f_n \geq f_{n+1}$ より，$G_n \subset G_{n+1}$．また，$f_n(x) \to 0$ $(n \to \infty, \forall x \in S)$ より，

$$S = \bigcup_{N=1}^{\infty} \underbrace{\bigcap_{n=N}^{\infty} G_n}_{=G_N} = \bigcup_{N=1}^{\infty} G_N.$$

ゆえに，$K \subset S$ をコンパクトとすると，$\exists N \in \mathbb{N}$ s.t. $K \subset \bigcup_{n=1}^{N} G_n = G_N$, i.e., $f_n(x) < \varepsilon$ $(\forall x \in K, \forall n \geq N)$. これは，$\sup_{x \in K} f_n(x) \to 0$ を意味する． □

Lemma 37.2. $\delta > 0$ を given とすると，$x \mapsto w'_x(\delta),\ D \to \mathbb{R}$ は上半連続である．

Remark 37.2. 特に，$x \mapsto w'_x(\delta)$ は $\mathcal{B}(D)/\mathcal{B}$ 可測である．

Proof. $\varepsilon > 0$ を given とする．$0 = t_0 < t_1 < \cdots < t_r = 1$ を，$t_i - t_{i-1} > \delta,\ w_x[t_{i-1}, t_i) < w_x(\delta) + \varepsilon,\ 1 \leq \forall i \leq r$ をみたすように選ぶ．$\eta > 0$ を十分小さく選んで，$\eta < \varepsilon,\ \delta + 2\eta < \min_i(t_i - t_{i-1})$ が成り立つとする．ここで，$d(x, y) < \eta$ に対して，$\exists \lambda \in \Lambda$ s.t. $\sup_t |y(t) - x \circ \lambda(t)| < \eta,\ \sup_t |\lambda^{-1}(t) - t| < \eta$. $s_i = \lambda^{-1}(t_i)$ とおくと，$s_i - s_{i-1} > t_i - t_{i-1} - 2\eta > \delta$ であり，$s, t \in [s_{i-1}, s_i)$ なら，$\lambda(s), \lambda(t) \in [t_{i-1}, t_i)$ であるから，$|y(s) - y(t)| < |x \circ \lambda(s) - x \circ \lambda(t)| + 2\eta < w'_x(\delta) + 3\varepsilon$. ゆえに，$d(x, y) < \eta$ なら，$w'_y(\delta) < w'_x(\delta) + 3\varepsilon$. これは，$x \mapsto w'_x(\delta)$ が上半連続であることを意味する． □

Proof of Theorem 37.2. (必要性). $\overline{A}$ がコンパクトなら，A は d に関して有界であり，$d(x, 0) = \sup_t |x(t)|$ より，$\sup_{x \in A} \sup_t |x(t)| < \infty$. 次に，各 $x \in D$ に対して，$w'_x(\delta) \downarrow 0$ $(\delta \downarrow 0)$ であり，$x \mapsto w'_x(\delta)$ は上半連続であるから，Dini の定理より，$\delta = \delta_k \downarrow 0$ とすると，$\sup_{x \in \overline{A}} w'_x(\delta) \to 0$. □

次の定理の証明は，C 空間の場合と同様である．

Theorem 37.4. $\{\mu_n\} \subset \mathcal{P}(D)$ が一様にタイトであるための必要十分条件は，次の (i) と (ii) が成り立つことである．

(i)
$$\lim_{\alpha \uparrow \infty} \inf_n \mu_n(\{x : \sup_t |x(t)| \leq \alpha\}) = 1.$$

(ii) 任意の $\varepsilon > 0$ に対して，
$$\lim_{\delta \downarrow 0} \limsup_{n \to \infty} \mu_n(\{x : w'_x(\delta) > \varepsilon\}) = 0.$$

次に極限分布が C に集中する場合を考える．

Theorem 37.5. $\mu_n, \mu \in \mathcal{P}(D)$ とする．$\mu_n \xrightarrow{w} \mu,\ \mu(C) = 1$ となるための必要十分条件は，μ_n の各 f.d.d. が μ のそれに弱収束し，かつ任意の $\varepsilon > 0$ に対して，
$$\lim_{\delta \downarrow 0} \limsup_{n \to \infty} \mu_n(\{x : w_x(\delta) > \varepsilon\}) = 0 \qquad (*)$$
が成り立つことである．

Proof. 必要性．$(*)$ のみ示せばよい．$x \in C$ なら，$d(x_n, x) \to 0 \Leftrightarrow d_{\sup}(x_n, x) \to 0$ であるから，$x \mapsto w_x(\delta)$ は $x \in C$ で連続である．従って，連続写像定理と Portmanteau 定理より，$\limsup_n \mu_n(w_x(\delta) \geq \varepsilon) \leq \mu(w_x(\delta) \geq \varepsilon)$. $\mu(C) = 1$ より，$\delta \downarrow 0$ として，$\mu(w_x(\delta) \geq \varepsilon) \to 0$ を得る．

十分性．$w'_x(\delta) \leq w_x(2\delta)$, $\sup_t |x(t)| \leq |x(0)| + k w_x(k^{-1})$ であるから，Theorem 37.4 の (i) と (ii) がみたされる．ゆえに，$\mu_n \xrightarrow{w} \mu$. あとは $\mu(C) = 1$ を示せばよい．$x \in D$ の最大ジャンプ・サイズを $j(x) = \sup_{0<t\leq 1} |x(t) - x(t-)|$ とおく ($|x(t) - x(t-)| \geq \varepsilon$ なる t は有限個しかないので，sup は max に取り替えることができる)．$d(x, y) < \varepsilon$ なら，$\exists \lambda \in \Lambda$ s.t. $\sup_t |x(t) - y \circ \lambda(t)| < \varepsilon$. このとき，$j(y) = j(y \circ \lambda)$ であるから，$|j(x) - j(y)| = |j(x) - j(y \circ \lambda)| < 2\varepsilon$. ゆえに，$j : D \to \mathbb{R}$ は連続であるから，連続写像定理より，$\mu_n \circ j^{-1} \xrightarrow{w} \mu \circ j^{-1}$. しかし，$j(x) \leq w_x(\delta)$ より，$\mu(j = 0) = 1$，すなわち，$\mu(C) = 1$ である． □

Example 37.1 (Donsker の不変原理)．$\{\xi_m\}$ を i.i.d. r.v.'s とし，$E[\xi_m] = 0, E[\xi_m^2] = 1$ とする．いま，$S_0 = 0, S_n = \sum_{i=1}^k \xi_i, k \in \mathbb{N}$ とおき，

$$X_n(t) = \frac{1}{\sqrt{n}} S_{\lfloor nt \rfloor}, \ t \in [0, 1]$$

と定義すると，X_n は D-valued r.v. である．$\{B(t) : t \in [0, 1]\}$ を標準 BM とすると，$B : \omega \mapsto B(\omega, \cdot)$ は D への写像として $\mathcal{F}/\mathcal{B}(D)$ 可測である．このとき，

$$X_n \xrightarrow{d} B \quad \text{in } D$$

を示す．X_n の各 f.d.d. が B のそれに弱収束することはよい．次に，

$$X'_n(t) = X_n(t) + (nt - \lfloor nt \rfloor) \frac{\xi_{\lfloor nt \rfloor + 1}}{\sqrt{n}}$$

とおくと，C における Donker の不変原理より，X'_n は漸近的に同程度連続である．いま，

$$\sup_t |X_n(t) - X'_n(t)| \leq \max_{1 \leq i \leq n} |\xi_i| / \sqrt{n}$$

であり，$\{\xi_m\}$ が i.i.d. であることと，$E[\xi_1^2] < \infty$ より，右辺 $\to 0$ a.s. 以上より，

$$\lim_{\delta \downarrow 0} \limsup_{n \to \infty} P(w(X_n, \delta) > \varepsilon) = 0$$

を得る．

Donsker の定理

確率過程 $\{B^\circ(t) : t \in [0, 1]\}$ が Brown 橋 (Brownian bridge) であるとは，それが連続なパスをもつ Gauss 過程であって，$E[B^\circ(t)] = 0$, $E[B^\circ(s)B^\circ(t)] = s \wedge t - st$, $\forall s, t \in [0, 1]$

をみたすことである．$\{B(t) : t \in [0,1]\}$ を標準 BM とすれば，$X(t) = B(t) - tB(1)$ は Brown 橋の定義をみたすから，Brown 橋は存在する．$B^\circ : \omega \mapsto B^\circ(\cdot, \omega)$ は C-valued r.v. であり (D への写像とみなせば，D-valued r.v. でもある)，その f.d.d.'s は一意に決まるから，B° の分布 $P \circ (B^\circ)^{-1}$ は ($\mathcal{B}(C)$ 上でも $\mathcal{B}(D)$ 上でも) 一意に決まる．

$\xi_1, \xi_2, \cdots \sim U[0,1]$ i.i.d. とし，経験分布関数を

$$F_n(t, \omega) = \frac{1}{n} \sum_{i=1}^{n} 1_{[0,t]}(\xi_i(\omega)),\ t \in [0,1]$$

と定義する．F_n は D-valued r.v. である．

$$X_n(t) = \sqrt{n}(F_n(t) - t),\ t \in [0,1]$$

とおく．

Theorem 37.6 (Donsker)**.** $X_n \xrightarrow{d} B^\circ$ in D.

次の評価を用いる．

Lemma 37.3. 次をみたす絶対定数 $c > 0$ が存在する：$S_n \sim Bin(n,p)$ に対して，

$$E[(S_n - E[S_n])^4] \le c(np + n^2p^2),\ E[S_n^4] \le c(np + n^4p^4).$$

Proof. $\eta_1, \ldots, \eta_n \sim Bin(1,p)$ i.i.d. とすると，$S_n \overset{d}{=} \sum_{i=1}^n \eta_i$. よって，

$$\begin{aligned} E[(S_n - E[S_n])^4] &= nE[(\eta_1 - p)^4] + 3(n^2 - n)(E[(\eta_1 - p)^2])^2 \\ &= n\{p(1-p)^4 + (1-p)p^4\} + 3(n^2 - n)p^2(1-p)^2 \le 2np + 3n^2p^2. \end{aligned}$$

次に，$S_n^4 \le 8\{(S_n - E[S_n])^4 + (E[S_n])^4\}$ より，

$$E[S_n^4] \le 16np + 24n^2p^2 + 8n^4p^4 \le 40np + 32n^4p^4.$$

2 番目の不等号は $x \le \max\{\sqrt{x}, x^2\}, x \ge 0$ による． □

Proof of Theorem 37.6. 多変量 CLT より，各 $t_1, \ldots, t_k \in [0,1]$ に対して，

$$(X_n(t_1), \ldots, X_n(t_k)) \xrightarrow{d} (B^\circ(t_1), \ldots, B^\circ(t_k)),\ n \to \infty.$$

B° のパスは連続であるから，あとは，任意の $\varepsilon > 0$ に対して，

$$\lim_{\delta \downarrow 0} \limsup_{n \to \infty} P(w(X_n, \delta) > \varepsilon) = 0$$

を示せばよい．

以下，$c_1, c_2, \ldots$ は適当な絶対定数とする．$m = m_n$ を $n \le 2^m < 2n$ となるように選び，$T_m = \{k/2^m : k = 0, 1, \ldots, 2^m\}$ とおく．また，$t \in [0,1]$ に対して，$\pi_m(t) = \max\{s \in T_m :$

$s \le t\}$ とおく．$s,t \in T_m, s<t$ に対して，$X_n(t)-X_n(s) = n^{-1/2}\sum_{i=1}^n\{1_{(s,t]}(\xi_i)-(t-s)\}$ であり，$\sum_{i=1}^n 1_{(s,t]}(\xi_i) \sim Bin(n, t-s)$ であるから，$p=t-s$ とおくと，Lemma 37.3 より，

$$E[(X_n(t)-X_n(s))^4] \le cn^{-2}(np+n^2p^2) = c(p/n+p^2) \le 3cp^2 = 3c(t-s)^2$$

を得る．よって，Lemma 29.3 より，任意の $\delta \in (0,1]$ に対して，

$$E\left[\max_{s,t\in T_m, |t-s|\le\delta}|X_n(t)-X_n(s)|\right] \le c_1\delta^{1/8}.$$

次に，$Y_n(t) = \sum_{i=1}^n I_{[0,t]}(\xi_i)$ とおくと，$t \mapsto Y_n(t)$ は非減少であるから，

$$\sup_{t\in[0,1]}\{Y_n(t)-Y_n(\pi_m(t))\} \le \max_{1\le k\le 2^m}\underbrace{\{Y_n(k/2^m)-Y_n((k-1)/2^m)\}}_{=:M_k}.$$

ここで，$E[M_k^4] \le c\{n2^{-m}+(n2^{-m})^4\} \le cn2^{-m+1}$ $(\because n2^{-m}\le 1)$ より，

$$E\left[\max_{1\le k\le 2^m} M_k\right] \le \left(E\left[\max_{1\le k\le 2^m} M_k^4\right]\right)^{1/4} \le \left(\sum_{k=1}^{2^m} E[M_k^4]\right)^{1/4} \le c_2 n^{1/4}.$$

また，$E[Y_n(t)-Y_n(\pi_m(t))] = n(t-\pi_m(t)) \le n2^{-m} \le 1$. 以上より，

$$E\left[\sup_{t\in[0,1]}|X_n(t)-X_n(\pi_m(t))|\right] \le c_3 n^{-1/4}.$$

いま，$s,t\in[0,1]$ に対して，

$$|X_n(t)-X_n(s)| \le |X_n(t)-X_n(\pi_m(t))|+|X_n(\pi_m(t))-X_n(\pi_m(s))|+|X_n(\pi_m(s))-X_n(s)|$$

であり，$|t-s|\le\delta$ なら，$|\pi_m(t)-\pi_m(s)| \le \delta + 2^{-m+1} \le \delta + 2n^{-1}$ であるから，

$$E[w(X_n,\delta)] \le c_1(\delta+2n^{-1})^{1/8} + 2c_3 n^{-1/4}.$$

よって，$\limsup_n E[w(X_n,\delta)] \le c_1\delta^{1/8} = o(1),\ \delta\downarrow 0.$ □

Donsker の定理の精緻化として，次の Komlós-Major-Tusnády の強近似定理が知られている．

Theorem 37.7 (Komlós et al. (1975))**.** ある絶対定数 $a,b,c>0$ が存在して，各 $n\in\mathbb{N}$ に対して，次の不等式をみたす Brown 橋 $\{B_n^\circ(t): t\in[0,1]\}$ が存在する：

$$P\left\{\sup_{t\in[0,1]}\sqrt{n}|X_n(t)-B_n^\circ(t)| > ax+b\log n\right\} \le ce^{-x},\ \forall x>0.$$

証明は，(かなり難しいが) Bretagnolle and Massart (1989) を参照せよ．Pollard (2002) の Chapter 10 にも証明がある．強近似定理に関しては，Csörgo and Horváth (1993) が詳しい．

Example 37.2 (適合度検定への応用)．F を $\mathbb{R}$ 上の d.f. とし，$\xi_1, \xi_2, \cdots \sim F$ i.i.d. とする．n 個のデータ $\xi_1, \ldots, \xi_n$ が与えられたとき，共通分布 F が与えられた分布 F_0 に等しいかどうか検定したいとする．すなわち，次の検定問題を考える：

$$H_0 : F = F_0 \quad \text{vs.} \quad H_1 : F \neq F_0.$$

F_0 は連続だと仮定すると，$F = F_0 \Leftrightarrow F_0(\xi_1) \sim U[0,1]$ である．そこで，$u_i = F_0(\xi_i), i = 1, \ldots, n$ と変換して，$u_1, \ldots, u_n$ の経験分布関数 $G_n(t) = n^{-1} \sum_{i=1}^n 1_{[0,t]}(u_i)$ が $U[0,1]$ の d.f. $G(t) = t$ から乖離していたら，H_0 を棄却する検定方法を考える．代表的な検定統計量として，Kolmogorov-Smirnov 統計量

$$KS_n = \sup_{t \in [0,1]} \sqrt{n} |G_n(t) - t|,$$

または，Cramér-von Mises 統計量

$$CM_n = \int_0^1 n\{G_n(t) - t\}^2 dt$$

などがある．D から $\mathbb{R}$ への写像

$$x \mapsto \sup_t |x(t)|, \ x \mapsto \int_0^1 \{x(t)\}^2 dt$$

はいずれも (Skorohod 位相に関して) 連続であるから (演習問題)，連続写像定理より，KS_n と CM_n の帰無分布はそれぞれ

$$KS_n \xrightarrow{d} \sup_t |B^\circ(t)|, \ CM_n \xrightarrow{d} \int_0^1 \{B^\circ(t)\}^2 dt$$

である．

演習問題

Exercise 37.1. Theorem 37.4 を示せ．

Exercise 37.2. D から $\mathbb{R}$ への写像

$$x \mapsto \sup_t |x(t)|, \ x \mapsto \int_0^1 \{x(t)\}^2 dt$$

はいずれも Skorohod 位相に関して連続であることを示せ．

Part VIII. 解析集合

本パートでは解析集合と射影定理/可測選択定理を考察する．本パートの内容はやや専門的ではあるが，統計学においてしばしば現れる問題に関連している．例えば，次のような問題を考えてみる．

- 与えられた確率過程の sup は可測になるか.
- 確率過程の極値点として定義される推定量を可測に選ぶことができるか.

こうした問題に対して一般的な十分条件を考察するために，解析集合なる集合が重要な役割を果たす.

38 解析集合

解析集合の定義をしよう．位相空間 S に対して，$\mathcal{B}(S)$ を S の Borel σ-field とする．位相空間 S,T に対して，写像 $f : S \to T$ が $\mathcal{B}(S)/\mathcal{B}(T)$ 可測なとき，f を Borel 可測であると言う．完備な距離によって距離化可能な可分位相空間のことを Poland 空間と呼ぶのであった．Poland 空間の可算直積は直積位相に対して Poland 空間になる [23]．Poland 空間 T に対して，空でない $A \subset T$ が 解析集合 (analytic set) であるとは，Poland 空間 S と Borel 可測な写像 $f : S \to T$ が存在して，$A = f(S)$ と表せることを言う．解析集合のことを Suslin 集合 とも言う．便宜的に，空集合も解析集合と定義しておこう．Poland 空間の Borel 集合が解析的であることを示そう.

Theorem 38.1. Poland 空間 S の任意の空でない Borel 集合 A に対して，$\mathbb{N}^{\mathbb{N}}$ から A の上への連続写像が存在する．ただし，$\mathbb{N}^{\mathbb{N}} = \{(n_k)_{k=1}^{\infty} : n_k \in \mathbb{N}, k = 1, 2, \dots\}$ には直積位相を入れる ($\mathbb{N}$ には離散位相を入れる).

まず次の補題を示そう.

Lemma 38.1. S を Poland 空間とし，$\varnothing \neq A \subset S$ を閉集合か開集合とする．このとき，A は相対位相に対して Poland 空間になる.

Remark 38.1. d を S の完備な距離とすると，S の開集合 $A \neq \varnothing$ に対して，(A, d) は完備とは限らない．しかし，A 上で d と同値な距離 $\hat{d}$ が存在して，$(A, \hat{d})$ は完備になる．また，Poland 空間の任意の空でない部分集合が相対位相に対して Poland 空間になるわけではない．例えば，$\mathbb{Q}$ は $\mathbb{R}$ の相対位相に関して Poland 空間にならない (演習問題).

[23] 完備かつ可分な距離空間 (S, d) に対して，$S^{\mathbb{N}}$ の直積位相は距離 $e((x_n), (y_n)) = \sum_{n=1}^{\infty} 2^{-n}\{d(x_n, y_n) \wedge 1\}$ によって距離化可能である．e の完備性は明らか．また，$C \subset S$ を S の可算稠密集合とし，$x_0 \in S$ を任意に固定した点とすると，$\{(x_1, \dots, x_k, x_0, x_0, \dots) : x_1, \dots, x_k \in C, k = 1, 2, \dots\}$ は $S^{\mathbb{N}}$ において稠密である.

Proof. A が閉集合のときは補題の結論は明らかであるから，A は開集合としよう．d を S の完備な距離とし，$B \subset S$ に対して，$d(x,B) := \inf_{y\in B} d(x,y)$ と定義する．このとき，A^c は閉集合であるから，$x \in A^c \Leftrightarrow d(x,A^c) = 0$ である．そこで，$x, y \in A$ に対して，

$$\hat{d}(x,y) := d(x,y) + \left| \frac{1}{d(x,A^c)} - \frac{1}{d(y,A^c)} \right|$$

と定義すれば，$\hat{d}$ は A 上で d と同値な距離であり，$(A,\hat{d})$ は完備になる． □

Proof of Theorem 38.1. 証明を 2 つのステップに分割する．

Step 1. $\mathbb{N}^{\mathbb{N}}$ から S の上への連続写像が存在することを示す．d を S の完備な距離とする．$\varnothing \neq B \subset S$ に対して，$\mathrm{diam}(B) := \{d(x,y) : x, y \in B\}$ を B の直径と呼ぶ．このとき，直径が 1 以下の空でない閉集合 $A_n, n = 1,2,\dots$ が存在して，$S = \bigcup_n A_n$ と表せる．ここで，$A(n) := A_n$ とし，各 $A(n)$ に対して，直径 $1/2$ 以下の空でない閉集合 $A(n,\ell), \ell = 1,2,\dots$ を選び，$A(n) = \bigcup_\ell A(n,\ell)$ と表す．同様の操作を繰り返して，各 $k = 1,2,\dots$ に対して，直径 $1/k$ 以下の空でない閉集合 $A(n_1,\dots,n_k), (n_1,\dots,n_k) \in \mathbb{N}^k$ が存在して，各 $(n_1,\dots,n_k) \in \mathbb{N}^k$ に対して，

$$A(n_1,\dots,n_k) = \bigcup_\ell A(n_1,\dots,n_k,\ell)$$

と表せる．このとき，$(n_k)_{k=1}^\infty \in \mathbb{N}$ に対して，$F_k = A(n_1,\dots,n_k), k = 1,2,\dots$ と定義すると，F_k は空でない閉集合であり，$F_k \supset F_{k+1}$ となる．F_k の直径は $1/k$ 以下であるから，任意の点列 $x_k \in F_k$ に対して，x_k は Cauchy 列であり，$k \to \infty$ のとき，$x_k \to x \in \bigcap_\ell F_\ell$ となる．ここで，$\mathrm{diam}(F_k) \downarrow 0$ であるから，x は点列の選び方によらず一意に決まる．そこで，$g((n_k)_{k=1}^\infty) := x$ と定義する．g は全射である．

g が連続であることを示そう．$(n_k^m) \to (n_k)$ in $\mathbb{N}^{\mathbb{N}}$ とする．このとき，任意の K に対して，m を十分大きく選べば，$(n_1^m,\dots,n_K^m) = (n_1,\dots,n_K)$ となる．ここで，g の構成より，$x_m := g((n_k^m))$ と $x := g((n_k))$ はともに $A(n_1,\dots,n_K)$ に属し，$A(n_1,\dots,n_K)$ の直径は $1/K$ 以下であるから，$d(x_m,x) \le 1/K$ である．よって，$d(x_m,x) \to 0$ $(m \to \infty)$ を得る．

Step 2. 定理の結論を示そう．

$$\mathcal{A}_0 := \{\varnothing\} \cup \{B \subset S : \exists f : \mathbb{N}^{\mathbb{N}} \to S \text{ 連続 s.t. } B = f(\mathbb{N}^{\mathbb{N}})\},$$
$$\mathcal{A} := \{B \subset S : B \in \mathcal{A}_0 \ \&\ B^c \in \mathcal{A}_0\}$$

とおく．S の空でない閉集合と開集合は相対位相に関して Poland 空間になるから，$\mathcal{A}_0$ に属する．よって，S の開集合は $\mathcal{A}$ に属する．あとは $\mathcal{A}$ が σ-field であることを示せば，$\mathcal{A} \supset \mathcal{B}(S)$ より，定理の結論が従う．$S \in \mathcal{A}, B \in \mathcal{A} \Rightarrow B^c \in \mathcal{A}$ は明らかである．そこで，$(\bigcup_n B_n)^c = \bigcap_n B_n^c$ より，

$$B_n \in \mathcal{A}_0, n = 1,2,\cdots \Rightarrow \bigcup_n B_n, \bigcap_n B_n \in \mathcal{A}_0$$

を示せばよい．$B_n \neq \varnothing$ $(\forall n = 1, 2, \dots)$ と仮定してよい．このとき，各 n に対して連続写像 $f_n : \mathbb{N}^{\mathbb{N}} \to S$ が存在して，$B_n = f_n(\mathbb{N}^{\mathbb{N}})$ と表せる．そこで，各 $(n_k)_{k=1}^{\infty} \in \mathbb{N}^{\mathbb{N}}$ に対して，$f((n_k)_{k=1}^{\infty}) := f_{n_1}((n_{k+1})_{k=1}^{\infty})$ と定義すれば，$f(\mathbb{N}^{\mathbb{N}}) = \bigcup_n B_n$ であり，$(n_k^m) \to (n_k)$ in $\mathbb{N}^{\mathbb{N}}$ に対して，m が十分大きいとき $n_1^m = n_1$ となるから，f は連続である．よって，$\bigcup_n B_n \in \mathcal{A}_0$ を得る．次に，$\bigcap_n B_n \in \mathcal{A}_0$ を示す．$\bigcap_n B_n \neq \varnothing$ と仮定してよい．$F = \{(\alpha^m)_{m=1}^{\infty} \in (\mathbb{N}^{\mathbb{N}})^{\mathbb{N}} : f_1(\alpha^1) = f_m(\alpha^m)\ \forall m = 2, 3, \dots\}$ とおくと，F は閉集合の可算積なので閉集合であり，$\bigcap_n B_n \neq \varnothing$ より $F \neq \varnothing$ である．そこで，$h : F \to S$ を $h((\alpha^m)_{m=1}^{\infty}) = f_1(\alpha^1)$ と定義すれば，h は $\bigcap_n B_n$ の上への連続写像である．ここで，$(\mathbb{N}^{\mathbb{N}})^{\mathbb{N}}$ は直積位相に対して Poland 空間であり，その閉集合 F は相対位相に対して Poland 空間になる．よって，Step 1 より，上への連続写像 $\iota : \mathbb{N}^{\mathbb{N}} \to F$ が存在するが，このとき，$\iota \circ h$ は $\mathbb{N}^{\mathbb{N}}$ から $\bigcap_n B_n$ の上への連続写像になる．以上より，$\bigcap_n B_n \in \mathcal{A}_0$ が示された． □

$\mathbb{N}$ は Poland 空間であるから，$\mathbb{N}^{\mathbb{N}}$ は Poland 空間である．よって，Poland 空間の Borel 集合は解析的である．さらに次の定理が成り立つ．

Theorem 38.2. T を Poland 空間とし，$\varnothing \neq A \subset T$ とする．このとき，A が解析的であるためには，$\mathbb{N}^{\mathbb{N}}$ から A の上への連続写像が存在することが必要十分である．

Proof. 十分性は明らか．必要性を証明しよう．A が解析集合なら Poland 空間 S と Borel 可測写像 $g : S \to T$ が存在して，$A = g(S)$ と表せる．$S \times T$ は直積位相に対して Poland 空間になる．また，T の位相を距離化する距離 e に対して，$\varphi(x, y) = e(y, g(x)), (x, y) \in S \times T$ と定義すれば，$(x, y) \mapsto y$ と $(x, y) \mapsto g(x)$ は $S \times T$ から T への Borel 可測写像であり，T が可分なことから，$(x, y) \mapsto (y, g(x))$ は $S \times T$ から $T \times T$ への Borel 可測写像になる．よって，φ は $S \times T$ から $\mathbb{R}$ への Borel 可測写像であり，g のグラフは

$$\mathrm{graph}(g) := \{(x, g(x)) : x \in S\} = \varphi^{-1}(\{0\})$$

と表せるから，$\mathrm{graph}(g)$ は $S \times T$ の Borel 集合である．よって，Theorem 38.1 より，$\mathbb{N}^{\mathbb{N}}$ から $\mathrm{graph}(g)$ の上への連続写像 h が存在する．$\pi : S \times T \to T$ を $\pi(x, y) = y$ と定義すれば，$A = g(S) = \pi(\mathrm{graph}(g)) = (\pi \circ h)(\mathbb{N}^{\mathbb{N}})$ となる．ここで，π は連続であるから，$\pi \circ h$ は $\mathbb{N}^{\mathbb{N}}$ から A の上への連続写像である． □

Remark 38.2. Theorem 38.2 と Theorem 38.1 の証明から次のこともわかる：$A_n, n = 1, 2, \dots$ を Poland 空間 S の解析集合とすると，$\bigcup_n A_n, \bigcap_n A_n$ も S の解析集合である．

解析集合は必ずしも Borel 可測でない．また，Poland 空間 S の解析集合の全体は σ-field になるとは限らない．なぜなら，$A \subset S$ に対して，A と A^c がともに解析集合なら，A は Borel 可測になるためである．従って，Borel 非可測な解析集合の存在から，解析集合の補集合は解析的とは限らないことがわかる．詳細は，Cohn (2013, Chapter 8) を参照せよ．

解析集合は Borel 可測とは限らないが，$\mathcal{B}(S)$ 上の任意の p.m. に対して，解析集合は $\mathcal{B}(S)$ の完備化に対して可測になる．可測空間 $(\Omega, \mathcal{F})$ に対して，$A \subset \Omega$ が ユニバーサルに可測

(universally measurable, u.m.) であるとは，$\mathcal{F}$ 上の任意の p.m. P に対して，A が $\mathcal{F}$ の P に関する完備化に対して可測であることを言う．$\mathcal{F}$ の P に関する完備化を $\mathcal{F}^P$ と書くとき，その P に関する共通部分

$$\overline{\mathcal{F}} = \bigcap_{P\text{: p.m. on } \mathcal{F}} \mathcal{F}^P$$

は $\mathcal{F}$ を含む σ-field である．$\overline{\mathcal{F}}$ を $\mathcal{F}$ の ユニバーサルな完備化 (universal completion) と呼ぶ．$\overline{\mathcal{F}}$ は Ω のユバーサルな可測集合の全体である．

Theorem 38.3. Poland 空間 S に対して，$A \subset S$ が解析的なら，$A \in \overline{\mathcal{B}(S)}$ である．

まず次の補題を示す．

Lemma 38.2. $(\Omega, \mathcal{F}, P)$ を確率空間とし，P^* を P の外測度とする：$P^*(A) = \inf\{P(B) : B \in \mathcal{F}, B \supset A\}, A \subset \Omega$. このとき，$A_n \subset \Omega, n = 1, 2, \ldots, A_n \uparrow A$ に対して，$P^*(A_n) \uparrow P^*(A)$.

Proof of Lemma 38.2. $P^*(A_n)$ は n に関して非減少である．$c = \lim_n P(A_n) \in [0, 1]$ とおくと，$c \leq P^*(A)$ である．$\varepsilon > 0$ を given とする．このとき，P^* の定義より，各 A_n に対して，次をみたす $B_n \in \mathcal{F}$ が存在する：$B_n \supset A_n$ & $P(B_n) \leq P^*(A_n) + \varepsilon$. そこで，$C_n = \bigcap_{m \geq n} B_m$ とおくと，$m \geq n$ に対して，$A_n \subset A_m \subset B_m$ より，$A_n \subset C_n \subset B_n$ であり，$C_n \in \mathcal{F}_n$ であるから，$P(C_n) \leq P^*(A_n) + \varepsilon$. さらに，$C_n \subset C_{n+1}$ であるから，$C = \bigcup_n C_n \in \mathcal{F}$ とおくと，$P(C_n) \uparrow P(C)$ となる．よって，$P(C) \leq c + \varepsilon$. C は A を含む $\mathcal{F}$ 可測集合であるから，$P^*(A) \leq c + \varepsilon$ を得る． □

Proof of Theorem 38.3. μ を $\mathcal{B}(S)$ 上の p.m. とし，μ^* と μ_* をその外測度と内測度とする．A は解析集合であるから，$\mathbb{N}^{\mathbb{N}}$ から A の上への連続写像 f が存在する．ここで，$m_1, m_2, \cdots \in \mathbb{N}$ に対して，

$$N(m_1, \ldots, m_k) := \{(n_j)_{j=1}^{\infty} \in \mathbb{N}^{\mathbb{N}} : n_j \leq m_j, j = 1, \ldots, k\},\ k = 1, 2, \ldots$$

と定義すると，$k = 1$ に対して，$m \to \infty$ のとき，$f(N(m)) \uparrow f(\mathbb{N}^{\mathbb{N}}) = A$ となるから，任意の $\varepsilon > 0$ に対して，m_1 を十分大きく選べば，

$$\mu^*(f(N(m_1))) > \mu^*(A) - \varepsilon$$

となる．同様に，$m \to \infty$ のとき，$f(N(m_1, m)) \uparrow f(N(m_1))$ となるから，m_2 を十分大きく選べば，

$$\mu^*(f(N(m_1, m_2))) > \mu^*(A) - \varepsilon$$

となる．この操作を繰り返して，次をみたす $(m_j)_{j=1}^{\infty} \in \mathbb{N}^{\mathbb{N}}$ を得る：各 k に対して，

$$\mu^*(f(N(m_1, \ldots, m_k))) > \mu^*(A) - \varepsilon.$$

そこで，$F = \bigcap_k N(m_1, \dots, m_k) = \{(n_j)_{j=1}^\infty : n_j \le m_j \ \forall j = 1, 2, \dots\}$ とおくと，F はコンパクト集合 $\{1, \dots, m_j\}$ の可算直積であるから，F もコンパクトである．よって，$K := f(F)$ は S のコンパクト集合である．

$N_k = N(m_1, \dots, m_k)$ とおいて，$K = \bigcap_k \overline{f(N_k)}$ を示そう．定義より，$K \subset \bigcap_k \overline{f(N_k)}$ は明らかなので，逆の包含関係を示す．d を S の位相を距離化する距離とし，$x \in \bigcap_k \overline{f(N_k)}$ とする．このとき，各 k に対して，$d(f((n_j^k)_{j=1}^\infty), x) \le 1/k$ をみたす $(n_j^k)_{j=1}^\infty \in M$ が存在する．ここで，$n_j^k \le m_j, j = 1, \dots, k$ であるから，対角線論法より，$\{(n_j^k) : k = 1, 2, \dots\}$ は $\mathbb{N}^\mathbb{N}$ において収束する部分列をもつ．この部分列の極限を (n_j) とおくと，$(n_j) \in \bigcap_k N_k = F$ である．一方，$f((n_j)) = x$ であるから，$\bigcap_k \overline{f(N_k)} \subset K$ を得る．

いままでの結果より，

$$\mu\left(\overline{f(N_k)}\right) \ge \mu^*(f(N_k)) > \mu^*(A) - \varepsilon$$

であり，$\overline{f(N_k)} \downarrow K$ であるから，

$$\mu(K) \ge \lim_k \mu\left(\overline{f(N_k)}\right) \ge \mu^*(A) - \varepsilon$$

を得る．ここで，$K \subset A$ であるから，$\mu_*(A) \ge \mu^*(A) - \varepsilon$ を得る．以上より，$\mu^*(A) = \mu_*(A)$ であるから，$A \in \mathcal{B}(S)^\mu$ が示された． □

Example 38.1. S, T を Poland 空間とし，$f : S \to T$ を Borel 可測とする．このとき，$f(S)$ は T の解析集合であるから，$f(S) \in \overline{\mathcal{B}(T)}$ である．例えば，A, B を $\mathbb{R}^k$ の Borel 集合とすると，$A+B := \{x+y : x \in A, y \in B\}$ は $\mathbb{R}^{2k}$ の Borel 集合 $A \times B$ の写像 $\mathbb{R}^{2k} \ni (x, y) \mapsto x+y$ による像であるから，ユニバーサルに可測である．じつは，$A + B$ は必ずしも Borel 可測にならない．Erdös and Stone (1970) を参照せよ．

演習問題

Exercise 38.1. Lemma 38.1 の証明の詳細を埋めよ．

Exercise 38.2. $\mathbb{Q}$ は $\mathbb{R}$ の相対位相に対して Poland 空間にならないことを示せ．
ヒント：Baire のカテゴリー定理．

空でない集合 S, T に対して，$A \subset S \times T$ の T への射影とは，$\{y \in T : (x, y) \in A \ \exists x \in S\}$ のことである．$\pi : S \times T \to T$ を $\pi(x, y) = y$ と定義すれば，A の T への射影は $\pi(A)$ である．

Exercise 38.3. S を Poland 空間とし，$A \subset S$ とする．このとき，A が解析集合であるためには，A が $\mathbb{N}^\mathbb{N} \times S$ の閉集合の S への射影になることが必要十分であることを示せ．

39　射影定理と可測選択定理

$(S,\mathcal{S}),(T,\mathcal{T})$ を可測空間とすると，$A\in\mathcal{S}\times\mathcal{T}$ に対して，A の T への射影 C は $\mathcal{T}$ 可測になるとは限らない．これは，Exercise 38.3 と Borel 可測でない解析集合の存在からわかる．C の可測性に関する定理はまとめて射影定理と呼ばれる．また，C は可測であるとして，$(g(y),y)\in C$ をみたす可測写像 $g:C\to S$ の存在に関する定理はまとめて可測選択定理と呼ばれる．

Theorem 39.1 (Jankov-von Neumann)**.** S,T を Poland 空間とし，$A\subset S\times T$ を空でない解析集合とする．このとき，$C:=\{y\in T:(x,y)\in A\ \exists x\in S\}$ は解析的である．さらに，次をみたす写像 $g:C\to S$ が存在する：g は $\overline{\mathcal{B}(T)}_C/\mathcal{B}(S)$ 可測であり，$(g(y),y)\in A\ \forall y\in C$ をみたす [24]．

証明に次の補題を使う．

Lemma 39.1. $\mathbb{N}^{\mathbb{N}}$ に辞書式順序を入れる．すなわち，

$$(m_k)_{k=1}^{\infty}<(n_k)_{k=1}^{\infty}\Leftrightarrow\exists\ell\in\mathbb{N}\text{ s.t. }m_k=n_k\ \forall k<\ell\ \&\ m_\ell<n_\ell$$

と定義する．このとき，$\mathbb{N}^{\mathbb{N}}$ の任意の空でない閉集合は最小元をもつ．

Proof of Lemma 39.1. $\varnothing\neq F\subset\mathbb{N}^{\mathbb{N}}$ を閉集合とする．F の第 1 座標の最小値を n_1 とし，n_1 を第 1 座標にもつ F の元のうち第 2 座標の最小値を n_2 とする．同様の操作を続けると，$(n_1,\dots,n_{k-1})$ を given とすると，n_k は $(n_1,\dots,n_{k-1})$ を $(k-1)$ までの座標にもつ F の元のうち第 k 座標の最小値である．このとき，$(n_j)_{j=1}^{\infty}$ は $(n_1,\dots,n_k)$ を k までの座標にもつ F の元の $k\to\infty$ のときの極限であるが，F は閉集合であるから，$(n_j)_{j=1}^{\infty}\in F$ であり，これが F の最小元を与える．　□

Proof of Theorem 39.1. $\pi_1:S\times T\to S,\pi_2:S\times T\to T$ を $\pi_1(x,y)=x,\pi_2(x,y)=y$ と定義しよう．A は解析集合であるから，$\mathbb{N}^{\mathbb{N}}$ から A の上への連続写像 f が存在する．ここで，$\gamma:=\pi_2\circ f$ は $\mathbb{N}^{\mathbb{N}}$ から C の上への連続写像であるから，C は解析的である．

次に，各 $y\in C$ に対して，$\gamma^{-1}(\{y\})$ は $\mathbb{N}^{\mathbb{N}}$ の閉集合であるから，Lemma 39.1 より，その最小元 $h(y)$ が存在する．ここで，h が $\overline{\mathcal{B}(T)}_C/\mathcal{B}(\mathbb{N}^{\mathbb{N}})$ 可測であれば，$g=\pi_1\circ f\circ h$ が求める写像である．

そこで，h の可測性を示そう．各 n に対して，

$$C(n)=\{y\in C:\gamma((m_j)_{j=1}^{\infty})=y\ \exists(m_j)_{j=1}^{\infty}\in\mathbb{N}^{\mathbb{N}}\text{ s.t. }m_1=n\}$$

とおくと，$C(n)$ は Poland 空間 $\{n\}\times\mathbb{N}^{\mathbb{N}}$ の連続像なので，$C(n)$ は解析的である．$y\in C$ に対して，$h_1(y):=\min\{n:y\in C(n)\}$ と定義すると，$h_1(y)=n\Leftrightarrow y\in C(n)\setminus\bigcup_{j<n}C(j)$

[24] $\overline{\mathcal{B}(T)}$ は $\mathcal{B}(T)$ のユニバーサルな完備化であり，$\overline{\mathcal{B}(T)}_C$ は $\overline{\mathcal{B}(T)}$ の C への制限である：$\overline{\mathcal{B}(T)}_C=\{B\cap C:B\in\overline{\mathcal{B}(T)}\}$

であるから，h_1 は $\overline{\mathcal{B}(T)}_C$ 可測である．これを繰り返して，$h_1, \dots, h_k$ が与えられたとして，

$$h_{k+1}(y) := \min\{n : \gamma((m_j)_{j=1}^{\infty}) = y \ \exists (m_j)_{j=1}^{\infty} \in \mathbb{N}^{\mathbb{N}} \\ \text{s.t. } m_1 = h_1(y), \dots, m_k = h_k(y), m_{k+1} = n\}$$

と定義すると，h_{k+1} も $\overline{\mathcal{B}(T)}_C$ 可測である．以上より，$h(y) = (h_1(y), h_2(y), \dots)$ の $\overline{\mathcal{B}(T)}_C/\mathcal{B}(\mathbb{N}^{\mathbb{N}})$ の可測性が従う（$\mathcal{B}(\mathbb{N}) = 2^{\mathbb{N}}$ のシリンダー σ-field が $\mathcal{B}(\mathbb{N}^{\mathbb{N}})$ に一致することに注意）．□

Theorem 39.1 において，S と T のいずれかが単に可測空間であっても，類似の結論が成り立つ．Poland 空間に含まれる解析集合に，相対位相に関する Borel σ-field を入れた可測空間を Suslin 可測空間 と呼ぶ．

Theorem 39.2 (Sainte-Beuve (1974))．$(\Omega, \mathcal{F})$ を可測空間，$\overline{\mathcal{F}}$ を $\mathcal{F}$ のユニバーサルな完備化とし，$(S, \mathcal{S})$ を Suslin 可測空間とする．このとき，空でない $A \in \mathcal{S} \times \mathcal{F}$ に対して，$C := \{\omega \in \Omega : (x, \omega) \in A \ \exists x \in S\}$ は $\overline{\mathcal{F}}$ 可測である．さらに，次をみたす写像 $H : C \to S$ が存在する：H は $\overline{\mathcal{F}}_C/\mathcal{S}$ 可測であり，$(H(\omega), \omega) \in A \ \forall \omega \in C$ をみたす．

Theorem 39.2 の証明の前に，いくつか応用例を紹介しよう [25]．

Example 39.1 (確率過程の sup の可測性)．$(\Omega, \mathcal{F})$ を可測空間とし，$(T, \mathcal{T})$ を Suslin 可測空間とする．$X(t, \omega), (t, \omega) \in T \times \Omega$ を T によってインデックスされた確率過程とし，X は $\mathcal{T} \times \mathcal{F}$ 可測とする．このとき，$x \in \mathbb{R}$ に対して，$\{\omega : \sup_{t \in T} X(t, \omega) > x\} = \{\omega : X(t, \omega) > x \ \exists t \in T\}$ であるから，$\sup_{t \in T} X(t)$ は $\overline{\mathcal{F}}$ 可測である．

Example 39.2 (可測な M 推定量の存在)．$(\Omega, \mathcal{F})$ を可測空間とし，$(\Theta, \mathcal{C})$ を Suslin 可測空間とする．また，$F : \Omega \times \Theta \to \mathbb{R} \cup \{+\infty\}$ を $\mathcal{F} \times \mathcal{C}$ 可測とする．いま，すべての $\omega \in \Omega$ に対して，$\inf_{\theta \in \Theta} F(\omega, \theta) < \infty$ であり，inf を達成する $\theta \in \Theta$ が存在すると仮定する．このとき，次をみたす $\hat{\theta} : \Omega \to \Theta$ が存在する：$\hat{\theta}$ は $\overline{\mathcal{F}}/\mathcal{C}$ 可測であり，$F(\omega, \hat{\theta}(\omega)) = \inf_{\theta \in \Theta} F(\omega, \theta) \ \forall \omega \in \Omega$．実際，$\inf_{\theta \in \Theta} F(\omega, \theta)$ は ω の関数として $\overline{\mathcal{F}}$ 可測であるから，$G(\omega, \theta) := F(\omega, \theta) - \inf_{\vartheta \in \Theta} F(\omega, \vartheta)$ は (ω, θ) の関数として $\overline{\mathcal{F}} \times \mathcal{C}$ 可測である．いま，仮定より，

$$\{\omega : G(\omega, \theta) = 0 \ \exists \theta \in \Theta\} = \Omega$$

であるから，次をみたす $\overline{\mathcal{F}}/\mathcal{C}$ 可測写像 $\hat{\theta} : \Omega \to \Theta$ が存在する：$G(\omega, \hat{\theta}(\omega)) = 0 \ \forall \omega \in \Omega$.

Example 39.3 (到達時刻の可測性)．$(S, \mathcal{S})$ を可測空間とし，$\{X_t : t \in \mathbb{R}_+\}$ を S 値確率過程とする．さらに，$(\Omega, \mathcal{F}, P)$ を完備とし，$\{\mathcal{F}_t : t \in \mathbb{R}_+\}$ を通常の仮定をみたすフィルトレーションとする．いま，各 $t > 0$ に対して，写像

$$[0, t] \times \Omega \ni (s, \omega) \mapsto X_s(\omega)$$

[25] 可測選択定理の数理経済学への応用に関しては，Stinchcomb and White (1992) が参考になる．

は $(\mathcal{B}_{[0,t]} \times \mathcal{F}_t)/\mathcal{S}$ 可測と仮定する (このとき，$\{X_t\}$ は発展的可測であると言う)．このとき，任意の $A \in \mathcal{S}$ に対して，到達時刻

$$T_A = \inf\{t \in \mathbb{R}_+ : X_t \in A\}$$

は停止時刻になる．すなわち，任意の $t \in \mathbb{R}_+$ に対して，$\{T_A \le t\} \in \mathcal{F}_t$ となる．このことを証明してみよう．$\{\mathcal{F}_t\}$ は通常の仮定をみたすので，任意の $t > 0$ に対して，$\{T_A < t\} \in \mathcal{F}_t$ を示せばよいが，$B = \{(s,\omega) \in [0,t] \times \Omega : X_s(\omega) \in A\}$ とおくと，

$$\{T_A < t\} = \{X_s \in A \ \exists s \in [0,t)\} = \{\omega : (s,\omega) \in B \cap ([0,t) \times \Omega) \ \exists s \in [0,t]\}.$$

ここで，$B \cap ([0,t) \times \Omega) \in \mathcal{B}_{[0,t]} \times \mathcal{F}_t$ であり，$[0,t]$ は通常の位相に対して Poland 空間であることと，$\mathcal{F}_t$ がすべての P 零集合を含むことから，$\{T_A < t\} \in \mathcal{F}_t$ を得る．

さて，Theorem 39.2 の証明に移ろう．次の補題を使う．

Lemma 39.2. $(\Omega,\mathcal{F}),(S,\mathcal{S})$ を可測空間とし，$g : \Omega \to S$ を $\mathcal{F}/\mathcal{S}$ 可測とする．このとき，$E \in \overline{\mathcal{S}}$ に対して，$g^{-1}(E) \in \overline{\mathcal{F}}$ となる．

Proof of Lemma 39.2. P を $\mathcal{F}$ 上の p.m. とし，$\mu = P \circ g^{-1}$ とおく．このとき，$E \in \mathcal{S}^\mu$ であるから，次をみたす集合 $A, B \in \mathcal{S}$ が存在する：$A \subset E \subset B$ & $\mu(B \setminus A) = 0$. ここで，$g^{-1}(A), g^{-1}(B) \in \mathcal{F}, g^{-1}(A) \subset g^{-1}(E) \subset g^{-1}(B), P(g^{-1}(B) \setminus g^{-1}(A)) = \mu(B \setminus A) = 0$ であるから，$g^{-1}(E) \in \mathcal{F}^P$ を得る． □

Proof of Theorem 39.2. S は解析集合なので，$\mathbb{N}^{\mathbb{N}} =: T$ から S の上への連続写像 f が存在する．このとき，$B := \{(y,\omega) : (f(y),\omega) \in A\}$ は $\mathcal{B}(T) \times \mathcal{F}$ 可測であり，$C = \{\omega : (y,\omega) \in B \ \exists y \in T\}$ と表せる．さらに，可算個の $E_n \in \mathcal{B}(T), F_n \in \mathcal{F}, n = 1,2,\dots$ が存在して，$B \in \sigma(\{E_n \times F_n\}_{n=1}^\infty)$ となる．いま，

$$g(\omega) := \sum_{n=1}^\infty 2 \cdot 1_{F_n}(\omega)/3^n, \ X(y,\omega) := (y, g(\omega)) \in T \times \mathbb{R}$$

と定義すると [26]，$B \in \sigma(X)$ であるから，ある $D \in \mathcal{B}(T) \times \mathcal{B}$ が存在して，$B = \{(y,\omega) : (y, g(\omega)) \in D\}$ と表せる．D は $T \times \mathbb{R}$ の Borel 集合であるから，Theorem 39.1 より，$\pi(D) := \{z \in \mathbb{R} : (y,z) \in D \ \exists y \in T\}$ は $\mathbb{R}$ の解析集合であり，よって，$\overline{\mathcal{B}}$ 可測であるから，Lemma 39.2 より，$C = g^{-1}(\pi(D)) \in \overline{\mathcal{F}}$ を得る．同じ定理より，$(h(z),z) \in D \ \forall z \in \pi(D)$ をみたす $\overline{\mathcal{B}}_{\pi(D)}/\mathcal{B}(T)$ 可測写像 h が存在する．$H := f \circ h \circ (g|_C)$ とおくと，$U \in \mathcal{S}$ に対して，$h^{-1}(f^{-1}(U)) \in \overline{\mathcal{B}}$ であるから，Lemma 39.2 より，$H^{-1}(U) = g^{-1}(h^{-1}(f^{-1}(U)) \in \overline{\mathcal{F}}$ を得る．よって，H は $\overline{\mathcal{F}}_C/\mathcal{S}$ 可測であり，$(H(\omega),\omega) \in A \ \forall \omega \in C$ をみたす． □

Theorem 39.2 において，$(S,\mathcal{S})$ を単に可測空間とすることはできない．例えば，$(\Omega,\mathcal{F}) = ([0,1],\mathcal{B}_{[0,1]})$ とし，S を $[0,1]$ の Lebesgue 非可測集合とする．このとき，$A = \{(x,\omega) \in \Omega \times S : x = \omega\}$ とおくと，$(x,\omega) \mapsto x - \omega$ の連続性より，A は $\mathcal{F} \times \mathcal{B}(S)$ 可測であるが，A の Ω への射影は S であるから，$\overline{\mathcal{F}}$ 可測でない (cf. Exercise 39.1).

[26] この関数 g は Marczewski 関数とも呼ばれる．

演習問題

Exercise 39.1. $\mathbb{R}$ の Lebesgue 可測集合の全体を $\mathcal{L}$ とおいたとき，$\overline{\mathcal{B}} \subsetneq \mathcal{L}$ を示せ.

S, T を空でない集合とすると，S から T への集合値関数とは，S の点から T の部分集合を対応させる関数のことを指す．S から T への集合値関数 φ を $\varphi : S \twoheadrightarrow T$ と表す.

Exercise 39.2. $(\Omega, \mathcal{F})$ を可測集合，$(S, \mathcal{S})$ を Suslin 可測空間とし，$\varphi : \Omega \twoheadrightarrow S$ を集合値関数とする．また，φ のグラフ $\mathrm{graph}(\varphi) := \{(\omega, x) : x \in \varphi(\omega)\}$ は $\mathcal{F} \times \mathcal{S}$ 可測とする．このとき，$B \in \mathcal{S}$ に対して，

$$\varphi^u(B) := \{\omega : \varphi(\omega) \subset B\},\ \varphi^\ell(B) := \{\omega : \varphi(\omega) \cap B \neq \varnothing\}$$

はともにユニバーサルに可測であることを示せ.

Exercise 39.3. $(\Omega, \mathcal{F})$ を可測空間とし，$(\Theta, \mathcal{C})$ を Suslin 可測空間とする．また，$\varphi : \Omega \twoheadrightarrow \Theta$ を $\mathrm{graph}(\varphi) \in \mathcal{F} \times \mathcal{C}$ なる集合値関数とし，$F : \Omega \times \Theta \to \mathbb{R} \cup \{+\infty\}$ を $\mathcal{F} \times \mathcal{C}$ 可測とする．いま，すべての $\omega \in \Omega$ に対して，$\varphi(\omega) \neq \varnothing$ & $\inf_{\theta \in \varphi(\omega)} F(\omega, \theta) < \infty$ であり，inf を達成する θ が存在すると仮定する．このとき，次をみたす $\mathcal{F}/\mathcal{C}$ 可測写像 $\hat{\theta} : \Omega \to \Theta$ が存在することを示せ：$\hat{\theta}(\omega) \in \varphi(\omega)$ & $F(\omega, \hat{\theta}(\omega)) = \inf_{\theta \in \varphi(\omega)} F(\omega, \theta)\ \forall \omega \in \Omega$.

索 引

アトム, 59
Arzelà-Ascoli の定理, 289
安定分布, 133
安定分布の吸引域, 136
安定分布の指数, 134

一様可積分, 55
一様にタイト, 96, 276
一様に同程度連続, 289
1 点分布, 9

Wiener 測度, 230
上向き横断不等式, 174
Ulam の定理, 276
Urysohn の部分列原理, 77

a.e., 33
a.e. に定義された可測関数, 34
a.s., 48
a.s. 収束, 54
exchangeable, 191
exchangeable σ-field, 191
Edgeworth 展開, 120
Etemedi の不等式, 294
L^p 最大不等式, 178
L^p 収束, 54
L^p 収束定理, 179
L^p ノルム, 37
Hermite 多項式, 121

外正則, 279
解析集合, 313
外測度, 10
Ky Fan 距離, 282
Gauss 過程, 227
可逆測度, 216
Kakutani の定理, 182
確率過程, 166, 227
確率空間, 48
確率時刻, 253
確率収束, 54, 268
確率測度 (p.m.), 48
確率ベクトル, 48
確率変数 (r.v.), 48
確率密度関数, 50
可算加法性, 9
可算生成, 22
可積分, 32
可測過程, 239
可測空間, 8
可測写像, 24
可測集合, 8
カップリング, 223
カップリング時間, 223
càdlàg, 243
下半連続, 29
可分, 234
可予測, 172
Carathéodory の拡張定理, 10
Cantor 関数, 21
Cantor 集合, 20
Cantor 分布, 50
完備, 12
完備化, 13

期待値, 50
既約, 213
逆 Fourier 変換, 144
キュムラント, 120
共分散, 51
強マルコフ性, 211, 251
局所コンパクト, 279
局所有限, 279

極値分布, 138
極値分布の吸引域, 140

Knothe-Rosenblatt 再配分, 166
区別できない, 236
クラスの性質, 215
Cramér-Wold device, 146
Cramér 条件, 122
Krickeberg 分解, 187
Glivenko-Cantelli の定理, 89
Kronecker の補題, 86

経験分布関数, 89
計数測度, 9

格子点分布, 123
Cauchy-Schwarz の不等式, 34
Komlós の補題, 160
Kolmogorov の拡張定理, 165
Kolmogorov の最大不等式, 85
Kolmogorov の 3 級数定理, 104
Kolmogorov の 0-1 法則, 82
Kolmogorov の定理, 132

再帰的, 212
再生理論, 89
Suslin 可測空間, 319
Suslin 集合, 313
座標写像, 64
サポート, 58
Sainte-Beuve の定理, 319
サンプルパス, 227

Schéffe の定理, 98
Jensen の不等式, 51
σ コンパクト, 280
σ-finite, 9
σ-field, 7
事象, 48
自明な σ-field, 82

射影定理, 149
弱位相, 96, 264
弱収束, 92, 143
弱収束 (in L^2), 158
Jankov-von Neumann の定理, 318
周期, 222
修正, 234
十分統計量, 164
条件付き確率, 153
条件付き期待値, 148
詳細釣り合い条件, 216
状態空間, 208
上半連続, 29
初期分布, 208
シリンダー σ-field, 67
シリンダー集合, 67

推移的, 212
Skorohod の埋め込み定理, 253
Steinhaus の定理, 69
Strassen-Dudley の定理, 283
Slutsky の補題, 94

制限, 9
整合的, 165
正再帰的, 218
正射影, 149
正射影作用素, 150
生成される, 8
正則, 279
正則条件付き確率 (r.c.p.), 161
正則条件付き分布 (r.c.d.), 161
正定値, 115
絶対連続, 50, 149, 188
零再帰的, 218
全変動, 237
全変動距離, 223

疎, 273

像測度, 33
測度, 8
測度空間, 9
Skorohod 距離, 301
Skorohod 位相, 301
Skorohod の表現定理, 93

第 1 カテゴリー, 273
対称化不等式, 73
大数の強法則, 70, 85
大数の弱法則, 70
第 2 カテゴリー, 273
タイプ, 133
タイプ収束定理, 133
互いに移りえる, 212
畳み込み, 63
単関数, 30
単調収束定理 (MCT), 35
Dunford-Pettis の定理, 159

Chebyshev の不等式, 53
Chapman-Kolmogorov 方程式, 210
中心極限定理 (CLT), 100
中線定理, 149
直積 σ-field, 19
直積測度, 42

通常の仮定, 238

停止時刻, 173, 238
定常測度, 215
定常分布, 216
Dini の定理, 307
Dirac 測度, 9
Dirichlet 分布, 194
tail event, 82
tail σ-field, 82
適合, 171, 238
de Finetti の定理, 192

Doob の正則化定理, 243
Doob の不等式, 178
Doob 分解, 176
同型, 65
同型写像, 65
同時分布, 48
到達可能, 212
到達時刻, 173, 238
同程度連続, 289
特異, 152
特異連続, 157
特性関数 (ch.f.), 107
独立, 61
閉じている, 213
凸関数, 51
凸集合, 53
Donsker の定理, 310
Donsker の不変原理, 256, 293, 309

内正則, 279
内測度, 15

2 次変動過程, 183
任意停止定理, 196

Burkholder の不等式, 183
バージョン, 148
Hardy-Littlewood の最大不等式, 204
Hartman-Winter の重複対数の法則, 87
π システム, 11
π-λ 定理, 11
バックワード・マルチンゲール, 189
Banach-Alaoglu の定理, 159
Banach 空間, 79
反射原理, 211, 252
反転公式, 107

非減少再配分, 56
非周期的, 222
Vitali 集合, 19

Vitali-Hahn-Sacks の定理, 272
微分同相, 45
Hewitt-Savage の 0-1 法則, 191
標準 Borel 空間, 65
表面測度, 46
Hilbert 空間, 79
Khinchin の不等式, 74

Fatou の補題, 35
field, 7
Fischer-Tippett-Gnedenko の定理, 138
フィルトレーション, 171, 238
Feller の定理, 100
Fubini の定理, 42
Brown 運動 (BM), 227
Brown 橋, 236
Blumenthal の 0-1 法則, 252
Proholov 距離, 282
Proholov の定理, 96, 276
分位点関数, 49
分解定理, 213
分散, 51
分布, 48
分布関数, 48
分布収束, 92, 143

Paley-Zygmund の不等式, 60
Baire σ-field, 29
Baire のカテゴリー定理, 273
Hoeffiding の不等式, 74
Berry-Esseen の定理, 103
Helly の定理, 97
Hölder の不等式, 34
Bernstein 多項式, 70
変数変換公式, 33

Portmanteau 定理, 95
Poland 空間, 66
Bochner の定理, 115

ポテンシャル, 187
Pólya の基準, 114
Pólya の定理, 92
Borel-Cantelli の補題, 76
Borel σ-field, 18
Borel 集合, 18
Borel 測度, 18
Poisson 少数の法則, 111

マルコフ・カーネル, 208
マルコフ・チェイン, 208
マルコフ性, 210, 247
マルコフ連鎖モンテカルロ法, 226
Markov の不等式, 53
Marcinkiewicz-Zygmund の定理, 88
Marcinkiewicz-Zygmund の不等式, 74
マルチンゲール, 171, 240
マルチンゲール差分アレイ, 201
マルチンゲール CLT, 201
マルチンゲール収束定理, 175
マルチンゲール変換, 172
Mallows 距離, 99

右連続 (フィルトレーション), 238
密度, 50, 149
Minkowski の不等式, 37

無限分解可能分布, 124

モーメント法, 119

有界変動, 237
有限次元分布 (f.d.d.), 166, 227
有限測度, 9
有限置換, 190
優収束定理 (DCT), 35
優マルチンゲール, 171, 240
ユニバーサルな完備化, 316
ユニバーサルに可測, 315

Radon-Nikodym の定理, 149

Radon-Nikodym 微分, 149
Radon 測度, 279
Laplace 変換, 73
λ システム, 11

Lyapunov 条件, 105
Riesz の表現定理, 149, 276
Riesz-Fischer の定理, 84
Riesz 分解, 187
離散的, 157
離散分布, 59
立方体, 204
Littlewood の 3 原理, 40
Littlewood の不等式, 60
Lindeberg-Feller の CLT, 100
Lindeberg 条件, 100
Lindeberg の定理, 100

Lusin の定理, 40
Lebesgue 可測集合, 19
Lebesgue 集合, 206
Lebesgue-Stieltjes 測度, 16
Lebesgue 測度, 16
Lebesgue の微分定理, 205
Lebesgue 分解, 152

劣分布関数, 157
劣マルチンゲール, 171, 240
Lévy-Khinchin 表現, 125
Lévy 過程, 130
Lévy 距離, 96
Lévy 測度, 128
Lévy の上向き定理, 181
Lévy の下向き定理, 189
Lévy の 0-1 法則, 182
連続写像定理, 94
連続性定理, 110
連続度, 291

Weierstrass の近似定理, 70
Wasserstein 距離, 99
Wald の等式, 199

参考文献

Aldous, D. (1985). Exchangeability and related topics. In: *École d'Été de Probabilités de Saint-Flour XIII* (Lecture Notes in Mathematics **1117**), Springer, pp. 1-198.

Applebaum, D. (2009). *Lévy Processes and Stochastic Calculus (2nd Edition)*. Cambridge University Press.

Azuma, K. (1967). Weighted sums of certain dependent random variables. *Tohoku Math. J.* **19** 357-367.

Beiglbock, M., Schachermayer, W., and Veliyev, B. (2012). A short proof of the Doob-Meyer theorem. *Stoch. Proc. Appl.* **122** 1204-1209.

Bertoin, J. (1996). *Lévy Processes*. Cambridge University Press.

Bickel, P.J. and Rosenblatt, M. (1973). On some global measures of the deviations of the kernel density estimates. *Ann. Statist.* **1** 1071-1095.

Billinsgley, P. (1968). *Convergence of Probability Measures*. Wiley.

Billinsgley, P. (1995). *Probability and Measures (3rd Edition)*. Wiley.

Bogachev, V. (2007). *Measure Theory. Volume I.* Springer.

Breiman, L. (1968). *Probability*. Addison-Wesley.

Bretagnolle, J. and Massart, P. (1989). Hungarian construction from the nonasymptotic viewpoint. *Ann. Probab.* **17** 239-256.

Burkholder, D.L. (1966). Martingale transforms. *Ann. Math. Statist.* **37** 1494-1504.

Chernozhukov, V., Fernández-Val, I., and Galichon, A. (2009). Improving point and interval estimators of monotone functions by rearrangement. *Biometrika* **96** 559-575.

Chung, K.-L. (2001). *A Course in Probability Theory (3rd Edition)*. Academic Press.

Cohn, D.L. (2013). *Measure Theory (2nd Edition)*. Birkhauser.

Coles, S. (2001). *An Introduction to Statistical Modeling of Extreme Values*. Springer.

Csörgo, M. and Horváth, L. (1993). *Weighted Approximations in Probability and Statistics*. Wiley.

Dembo, A. (2015). *Probability Theory*. Lecture notes available from the author's webpage.

Doob, J.L. (1953). *Stochastic Processes*. Wiley.

Dudley, R.M. (1968). Distances of probability measures and random variables. *Ann. Math. Statist.* **39** 1563-1572.

Dudley, R.M. (2002). *Real Analysis and Probability*. Cambridge University Press.

Durrett, R. (2010). *Probability: Theory and Examples (4th Edition)*. Cambridge University Press.

Etemadi, N. (1981). An elementary proof of the strong law of large numbers. *Z. Warsch. verw. Gebiete.* **55** 119-122.

Erdös, P. and Stone, A.H. (1970). On the sum of two Borel sets. *Proc. Amer. Math. Soc.* **25** 304-306.

Folland, G.B. (1999). *Real Analysis (2nd Edition)*. Wiley.

Gnedenko, B.V. and Kolmogorov, A.N. (1968). *Limit Distributions for Sums of Independent Random Variables (2nd translated edition)*. Addison-Wesley.

Hall, P. and Heyde, C.C. (1980). *Martingale Limit Theory and Its Application*. Academic Press.

Heyde, C.C. and Brown, B.M. (1970). On the departure from normality of a certain class of martingales. *Ann. Math. Statist.* **41** 2161-2165.

Jenrich, R.I. (1969). Asymptotic properties of nonlinear least squares estimators. *Ann. Math. Statist.* **40** 633-643.

Karatzas, I. and Shreve, S.E. (1997). *Browninan Motion and Stochastic Calculus*. Springer.

Kolmogorov, A.N. (1933). *Grundbegriffe der Wahrscheinlichkeitsrechnung*. Springer-Verlag.

Komlós, J. (1967). A generalization of a problem of Steinhaus. *Acta Math. Acad. Sci. Hungar.* **18** 217-229.

Komlós, J., Major, P., and Tusnády, G. (1975). An approximation for partial sums of independent rv's and the sample df I. *Z. Warhsch. Verw. Gabiete* **32** 111-131.

Leadbetter, M.R., Lindgren, G., and Rootzén, H. (1983). *Extremes and Related Properties of Random Sequences and Processes*. Springer.

Le Cam, L. (1986). The Central Limit Theorem around 1935. *Statist. Science.* **1** 78-96.

Ledoux, M. and Talagrand, M. (1991). *Probability in Banach Spaces.* Springer.

Lehmann, E.L. and Romano, J. (2005). *Testing Statistical Hypotheses (3rd edition).* Springer.

Levin, D.A., Peres, Y., and Wilmer, E.L. (2008). *Markov Chains and Mixing Times.* American Mathematical Society.

McLeish, D.L. (1974). Dependent central limit theorems and invariance principles. *Ann. Probab.* **2** 620-628.

Meyn, S. and Tweedie, R.L. (2009). *Markov Chains and Stochastic Stability (2nd Edition).* Cambridge University Press.

Mörters, P. and Peres, Y. (2010). *Brownian Motion.* Cambridge University Press.

Neveu, J. (1975). *Dicrete-Parameter Martingales.* North-Holland.

Obłój, J. (2004). The Skorokhod embedding problem and its offspring. *Probab. Surveys* **1** 321-392.

Okamoto, M. (1973). Distinctiveness of the eigenvalues of a quadratic form in a multivariate sample. *Ann. Statist.* **1** 763-765.

Pollard, D. (2002). *User's Guide to Measure Theoretic Probability.* Cambridge University Press.

Protter, P.E. (2005). *Stochastic Integration and Differential Equations (2nd Edition).* Springer.

Reed, M. and Simon, B. (1980). *Functional Analysis (Vol. I of Method of Modern Physics).* Academic Press.

Resnick, S. (1987). *Extreme Values, Regular Variations, and Point Processes.* Springer.

Resnick, S. (1998). *A Probability Path.* Birkhauser.

Robert, C. and Casella, G. (2005). *Monte Carlo Statistical Methods (2nd Edition).* Springer.

Sainte-Beuve, M.-F. (1974). On the extension of the von Neumann-Aumann's theorem. *J. Functional Anal.* **17** 112-129.

Sato, K.-I. (1999). *Lévy Processes and Infinite Divisibility.* Cambridge University Press.

Stinchcomb, M.B. and White, H. (1992). Some measurability results for extrema of random functions over random sets. *Rev. Econ. Stud.* **59** 495-512.

Stroock, D.W. (2011). *Probability Theory: An Analytic View (2nd Edition)*. Cambridge University Press.

Tao, T. (2010). *An Epsilon of Room, I: Real Analysis*. American Mathematical Society.

Topsøe, F. (1974). Compactness and tightness in a space of measures with the topology of weak convergence. *Math. Scand.* **34** 187-210.

Villani, C. (2008). *Optimal Transport: Old and New*. Springer.

Wegner, H. (1973). On consistency of probability measures. *Z. Warsch. verw. Gebiete.* 27 335-338.

Wichura, M.J. (1970). On the construction of almost uniformly convergent random variables with given weakly convergent image laws. *Ann. Math. Statist.* **41** 284-291.

Williams, D. (1991). *Probability with Martingales*. Cambridge University Press.

黒田成俊. (1980). 「関数解析」共立出版.

佐藤健一. (1991). 「加法過程」紀伊国屋書店.

竹村彰通. (1991). 「現代数理統計学」創文社.

藤田宏・伊藤清三・黒田成俊.「関数解析」岩波書店.

舟木直久. (2004).「確率論」朝倉書店.

舟木直久. (2005). 「確率微分方程式」岩波書店.

吉田伸生. (2006). 「ルベーグ積分入門」遊星社.