測度論と確率論

広島大学理学部数学科確率統計 C 講義ノート

岩田耕一郎

2007 年 7 月 4 日

目 次

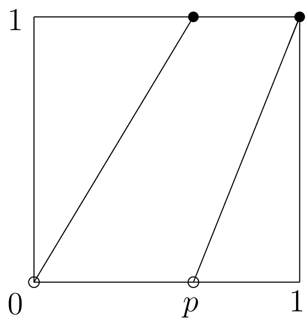
1 導入–あるモデル

記号

$\mathbb{Z}$ 整数全体、$\mathbb{N}$ 正の整数全体、$\mathbb{Q}$ 有理数全体、$\mathbb{R}$ 実数全体、$\mathbb{C}$ 複素数全体
$\mathbb{R}_{\geq 0} := \{x \in \mathbb{R} : x \geq 0\}$, $\mathbb{R}_{>0} := \{x \in \mathbb{R} : x > 0\}$, $\overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty, +\infty\}$.

$0 < p < 1$ なる実数 p を一つ固定して次のような写像 $\varphi : (0, 1] \to (0, 1]$ を与える。

$$\varphi(\omega) = \begin{cases} \omega/p & 0 < \omega \leq p \\ (\omega - p)/(1-p) & p < \omega \leq 1 \end{cases}$$

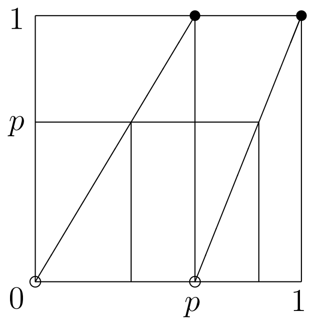


写像 φ のグラフ

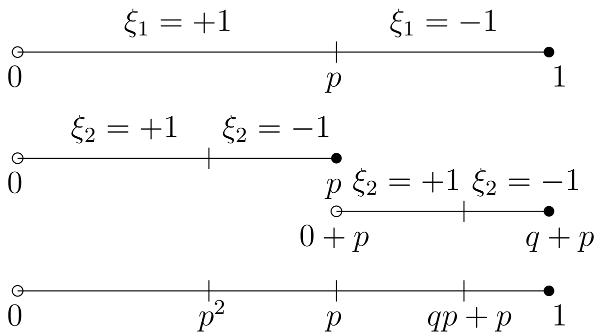
次に関数 $(0, 1] \to \mathbb{R}$ の列 ξ_k $k \in \mathbb{N}$ を下のように定義する。

$$\xi_1(\omega) := \begin{cases} 1 & 0 < \omega \leq p \\ -1 & p < \omega \leq 1 \end{cases}, \quad \xi_k(\omega) := \xi_{k-1}(\varphi(\omega)) \ k = 2, 3, \ldots$$

ξ_1 は区間を $p : 1-p$ の比に内分して左側区間では 1 右側区間では -1 を返すような関数である。次に、下左のグラフから見て取れるように $\xi_1 \circ \varphi$ は ξ_1 による分割区分のそれぞれを $p : 1-p$ の比に内分して左側区間では 1 右側区間では -1 を返すような関数となる。



$\xi_1 \circ \varphi$ による $(0, 1]$ の分割



ただし、$q = 1 - p$ と書いた。以後しばらくは q は $1 - p$ を表すものとする。$n \in \mathbb{N}$ を一つ固定してやると $(\xi_1(\omega), \xi_2(\omega), \ldots, \xi_n(\omega))$ により $\omega \in (0, 1]$ の分類ができる。

n	分割			
1	$(0, p]$ $+$			
2	$(0, p^2]$ $++$		$(p^2, p]$ $+-$	
3	$(0, p^3]$ $+++$	$(p^3, p^2]$ $++-$	$(p^2, p(qp+p)]$ $+-+$	$(p(qp+p), p]$ $+--$

<table>
<tr><th>n</th><th colspan="4">分割</th></tr>
<tr><td>1</td><td colspan="4">$(p,1]$ $-$</td></tr>
<tr><td>2</td><td colspan="2">$(p,qp+p]$ $-+$</td><td colspan="2">$(qp+p,1]$ $--$</td></tr>
<tr><td>3</td><td>$(p,qp^2+p]$
$-++$</td><td>$(qp^2+p,qp+p]$
$-+-$</td><td>$(qp+p,q(qp+p)+p]$
$--+$</td><td>$(q(qp+p)+p,1]$
$---$</td></tr>
</table>

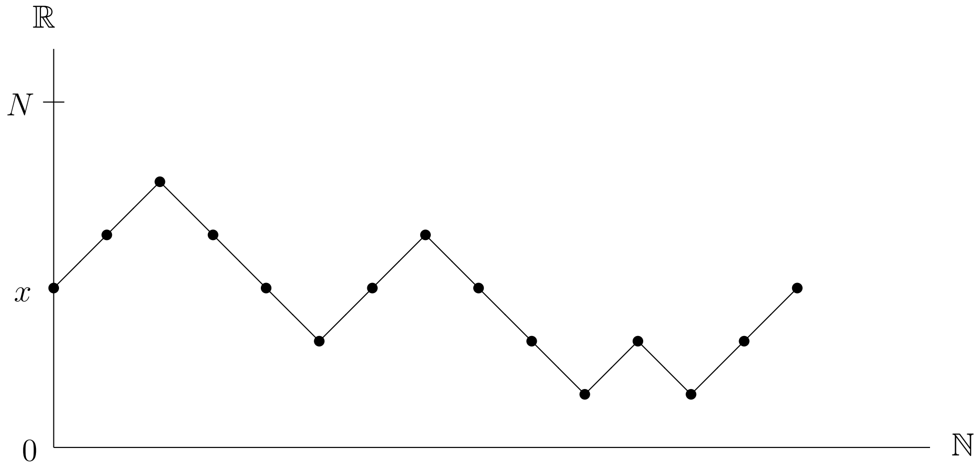
以上から各 $k\in\mathbb{N}$ に対して ξ_k が値 1 を返すような分割区分のそれぞれの長さの合計は p であり ξ_k が値 -1 を返すような分割区分のそれぞれの長さの合計は $1-p$ であることが予想できるであろう。実際、次が成り立つが、その厳密な証明は後で行う。

> $\varepsilon(1),\varepsilon(2),\dots,\varepsilon(n)$ を 1 または -1 からなる有限数列とする。このとき $\xi_k(\omega)=\varepsilon(k)$ $\forall k\in\{1,2,\dots,n\}$ であるような $\omega\in(0,1]$ 全体のなす集合の長さは次に等しい。
>
> $$p^{\#\{k:\varepsilon(k)=1\}}(1-p)^{\#\{k:\varepsilon(k)=-1\}}$$

さらに関数 $\mathbb{R}\times(0,1]\to\mathbb{R}$ の列 η_k $k\in\mathbb{N}\cup\{0\}$ を以下で定義する。

$$\eta_0(x,\omega):=x,\ \eta_k(x,\omega):=\eta_{k-1}(x,\omega)+\xi_k(\omega)\ k=1,2,\dots$$

$x\in\mathbb{R}.$ $\omega\in(0,1]$ を固定して数列 $\eta_k(x,\omega)$ の挙動を追跡する。



k を時間の経過を表現するパラメータと解釈するのが自然な考えである。そのとき量 $\eta_k(x,\omega)$ が最初に 0 になる時間などに関心が向かうことが多い。そこで次の関数を考察する。

$$\tau(x,y,\omega):=\min\{k\in\mathbb{N}\cup\{0\}:\eta_k(x,\omega)=y\}$$

ここで $\min\emptyset=+\infty$ と約束すれば定義が破綻することはない。素朴な問題としては、$N\in\mathbb{N}$ および $0<x<N$ を一つ固定するとき

> $\tau(x,0,\omega)<\tau(x,N,\omega)$ となるような $\omega\in(0,1]$ 全体のなす集合の長さはいくらか？

というのがある。だがこれは初等確率論の守備範囲ではない。$\tau(x,0,\omega) < \tau(x,N,\omega)$ であるかどうかをすべての ω について判定するには $\eta_k(x,\omega)$ $k \in \mathbb{N}$ を知る必要があるからである。すなわち真に無限が関わる数学の問題なのである。そのような視点に立つとき、測度論に基づく設定が自然なものであるというのが現在のコンセンサスとなっている。次節以降はその解説にあてていくことにする。

2　確率空間と確率変数

数学としての確率論は、可能性すべてを網羅したものを表す集合を設定し、確率をはかる対象としての事象はその部分集合でしかるべき条件を満たすものとして把握する。多くの場合、集合の元そのものが直に見えるわけではなく、観測にかかる量はなにかが介在していると考えるのが自然である。そのような介在物が確率変数であり、個々の事象も対応する確率変数によって決定されていると考えるわけである。

2.1 定義. 三つ組 $(\Omega, \mathcal{F}, P)$ が**確率空間**(probability space) であるとは、

(i) Ω は**標本空間**(sample space) と呼ばれるある集合。
(ii) $\mathcal{F}$ は**事象の族**(family of events) と呼ばれる Ω 上のある σ 加法族。
(iii) P は**確率測度**(probability measure) と呼ばれる可測空間 $(\Omega, \mathcal{F})$ 上のある測度であって $P(\Omega) = 1$ を満たすもの。

各 $\omega \in \Omega$ を**標本**(sample) あるいは見本、$A \in \mathcal{F}$ を**事象**(event)、また $P(A)$ を**事象 A の確率**(probability) という。

以後、確率空間を表す専用の記号として $(\Omega, \mathcal{F}, P)$ を登録しておく。

記号

集合 A, B に対して $B \setminus A := \{x \in B : x \notin A\}$

復習

集合 S が与えられ、その部分集合の族 $\mathcal{B}$ が次の条件を満たすとき、$\mathcal{B}$ を S 上の **σ-加法族**(σ-field) であるという。また組 $(S, \mathcal{B})$ を**可測空間**(measurable space) という。

(i) $\emptyset \in \mathcal{B}$
(ii) $A \in \mathcal{B} \Rightarrow S \setminus A \in \mathcal{B}$ （全体集合が自明なら $S \setminus A$ でなく A^{c} と書く。）
(iii) $A_n \in \mathcal{B}\ \forall n \in \mathbb{N} \Rightarrow \bigcup_{n=1}^{\infty} A_n \in \mathcal{B}$

さらに次の条件を満たす関数 $\mu : \mathcal{B} \to \overline{\mathbb{R}}$ を S 上の**測度**(measure) と呼ぶ。

(i) $\mu(A) \geq 0\ \forall A \in \mathcal{B}$, $\mu(\emptyset) = 0$
(ii) $A_n \in \mathcal{B}\ \forall n \in \mathbb{N}$ 互いに素 $\Rightarrow \mu(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mu(A_n)$ σ **加法性**

三つ組 $(S, \mathcal{B}, \mu)$ を**測度空間**(measure space) という。

以後、S は一般的な集合を表す記号として使い特定のものを意識しない。とはいうものの S は標本空間 Ω を指すかあるいは $\mathbb{R}^d$ の部分集合を指すことが多い。但し $S=\emptyset$ だとそれを定義域とする写像はつまらないものしかないので、$S\neq\emptyset$ としておいた方がよいだろう。

2.2 定義. $S\neq\emptyset$ かつ $\mathcal{B}$ が S 上の σ-加法族であるとき可測空間 $(S,\mathcal{B})$ は**非自明**であるという。

この講義ノートでは必要な場合でもいちいち $S\neq\emptyset$ と断らないこともある。なお確率空間に関しては $P(\Omega)=1$ であるから必然的に $\Omega\neq\emptyset$ である。確率空間の最も重要な例としては区間 $(0,1]$, その上の Borel 集合族と Lebesgue 測度からなる三つ組がある（まだこの概念に不案内でも構わない）。Borel 集合族についてはこの節で正式に導入し、また第 3 節で Lebesgue 測度について一つのとらえ方を紹介する。

2.3 定義. **確率変数**(random variable) とは $\mathcal{F}$ 可測関数 $X:\Omega\to\overline{\mathbb{R}}$ をいう。とくに $X(\omega)\in\mathbb{R}$ $\forall\omega\in\Omega$ であるような確率変数を**実確率変数**(real random variable) と呼ぶことにする。

復習

$(S,\mathcal{B})$ を非自明可測空間とする。関数 $f:S\to\overline{\mathbb{R}}$ が $\mathcal{B}$ **可測**(measurable) であるとは

$$\{s\in S: f(s)>a\}\in\mathcal{B}\ \forall a\in\mathbb{R}$$

が成り立つことをいう。そのような f を $\mathcal{B}$ **可測関数**(measurable function) と呼ぶ。

次の命題を思い出しておこう。

2.4 補題. $(S,\mathcal{B})$ を非自明な可測空間とする。$f:S\to\overline{\mathbb{R}}$ を Image f が可算集合であるような関数とする。このときその $\mathcal{B}$ 可測性は $\{s\in S: f(s)=y\}\in\mathcal{B}\ \forall y\in\overline{\mathbb{R}}$ と同値である。

さて確率をはかる対象としての σ 加法族 $\mathcal{F}$ であるが、大きすぎると逆に確率そのものの内容が乏しいものになることが知られている。そこで必要最小限のものを取り扱うというスタンスに立つわけだが、その立脚根拠となるのが次の命題である。

2.5 補題. S 上の σ-加法族たち $\mathcal{B}_\alpha$ に対してそれらの共通部 $\bigcap_\alpha\mathcal{B}_\alpha$ も S 上の σ-加法族である。

証明. (i) 任意の α について $\emptyset\in\mathcal{B}_\alpha$ であるから、$\emptyset\in\bigcap_\alpha\mathcal{B}_\alpha$ である。

(ii) $A\in\bigcap_\alpha\mathcal{B}_\alpha$ とする。これは $A\in\mathcal{B}_\alpha\ \forall\alpha$ を意味する。各 $\mathcal{B}_\alpha$ は σ-加法族なので、$A^c\in\mathcal{B}_\alpha$ である。これが任意の α について成り立つので $A^c\in\bigcap_\alpha\mathcal{B}_\alpha$ が導かれる。

(iii) $A_n\in\bigcap_\alpha\mathcal{B}_\alpha\ \forall n\in\mathbb{N}$ とする。これは $A_n\in\mathcal{B}_\alpha\ \forall n\in\mathbb{N}\ \forall\alpha$ を意味する。各 $\mathcal{B}_\alpha$ は σ-加法族なので、$\bigcup_{n=1}^\infty A_n\in\mathcal{B}_\alpha$ である。α は任意なので $\bigcup_{n=1}^\infty A_n\in\bigcap_\alpha\mathcal{B}_\alpha$ が導かれる。 □

2.6 系. S の部分集合の族 $\mathcal{A}$ に対して次の条件を満たす集合族がただ一つ存在する。

(i) $\mathcal{B}$ は S 上の σ-加法族かつ $\mathcal{A}\subset\mathcal{B}$ である。
(ii) 条件 (i) を満たす任意の集合族 $\mathcal{B}'$ に対して $\mathcal{B}\subset\mathcal{B}'$ である。**最小性**

証明. 先ず $\mathcal{A}$ を内包する S 上の σ-加法族は存在する。実際、$\{A : A \subset S\}$ がそうである。そこで条件 (i) を満たす任意の集合族たちすべての共通部をとれば、それは補題 2.5 により σ-加法族である。しかもそれは条件 (ii) も満たす。一意性の確認は読者にゆだねる。 □

記号

集合 S とその部分集合の族 $\mathcal{A}$ に対し系 2.6 で規定される S 上の σ-加法族を $\sigma_S(\mathcal{A})$ と表記する。前後関係から分かるときは S を省いて $\sigma(\mathcal{A})$ と書くことが多い。

2.7 定義. (i) $\sigma_S(\mathcal{A})$ を $\mathcal{A}$ **で生成される** S **上の** σ-**加法族**(the σ-field generated by $\mathcal{A}$) と呼ぶ。
(ii) $a < b$ なる $a, b \in \mathbb{R}$ に対し区間 $(a, b]$ を**左半開区間**といい、それらすべてで生成される $\mathbb{R}$ 上の σ-加法族を *Borel* **集合族**(Borel σ-field) といい Borel($\mathbb{R}$) と表記する。その各元を *Borel* **集合**(Borel set) という。

2.8 注意. 上であげた Borel 集合族の定義は当座のものであるが 1 次元で利用するにはもっとも回り道の少ないものと考えている。正式のものは補題 2.10(ii) で取り扱うものであり、多次元空間に関わる第 5 節以後では後者を定義として採用する。

2.9 演習問題. この問題では開区間すべてで生成される $\mathbb{R}$ 上の σ-加法族を $\mathcal{B}$ とかく。
(i) $a < b$ なる $a, b \in \mathbb{R}$ に対し $(a, b) \in \mathrm{Borel}(\mathbb{R})$ かつ $(a, b] \in \mathcal{B}$ であることを示せ。
(ii) $\mathrm{Borel}(\mathbb{R}) = \mathcal{B}$ であることを示せ。
(iii) 任意の $\mathbb{R}$ の開部分集合は開区間の可算合併でかけることを示せ。

2.10 補題. (i) $\{(a, +\infty)\,; a \in \mathbb{R}\}$ で生成される $\mathbb{R}$ 上の σ-加法族は Borel($\mathbb{R}$) である。
(ii) $\mathbb{R}$ の開部分集合すべてで生成される $\mathbb{R}$ 上の σ-加法族は Borel($\mathbb{R}$) である。

証明. (i) まず $(a, +\infty) = \bigcup_{n=1}^{\infty}(a, a+n] \in \mathrm{Borel}(\mathbb{R})$ である。他方 $(a, b] = (a, +\infty) \cap (b, +\infty)^{\mathrm{c}}$ は $\{(a, +\infty)\,; a \in \mathbb{R}\}$ で生成される σ-加法族に属する。

(ii) $\mathbb{R}$ の開部分集合全体の族を $\mathcal{O}$ と書こう。また $\mathcal{B}$ は演習問題 2.9 と同じ集合族とする。開区間は開集合であるから、演習問題 2.9(ii) で得たことより $\mathrm{Borel}(\mathbb{R}) = \mathcal{B} \subset \sigma(\mathcal{O})$ が導かれる。次に演習問題 2.9(iii) で調べたことを使うと任意の開部分集合が $\mathcal{B} = \mathrm{Borel}(\mathbb{R})$ に属することがわかる。従って Borel($\mathbb{R}$) は開部分集合すべてを元とする σ 加法族になるが、そのようなものの最小が $\sigma(\mathcal{O})$ なので、$\sigma(\mathcal{O}) \subset \mathrm{Borel}(\mathbb{R})$ を得た。 □

Borel 集合族の重要性の一つは次に述べることの成立である。

2.11 演習問題. 連続関数 $f : \mathbb{R} \to \mathbb{R}$ は Borel($\mathbb{R}$) 可測であることを示せ。

記号

関数 $f : S \to \mathbb{R}$ および $B \subset \mathbb{R}$ に対し $f^{-1}(B) := \{s \in S : f(s) \in B\}$ と書く。

2.12 補題. 関数 $f : S \to \mathbb{R}$ が与えられたとき、次が成り立つ。

$$f^{-1}(\emptyset) = \emptyset, f^{-1}(\mathbb{R}) = S, f^{-1}(B \setminus A) = f^{-1}(B) \setminus f^{-1}(A), f^{-1}(\bigcup_{n=1}^{\infty} B_n) = \bigcup_{n=1}^{\infty} f^{-1}(B_n).$$

すなわち対応 $B \mapsto f^{-1}(B)$ は σ-加法族としての準同型写像である。

2.13 演習問題. 補題 2.12 を示せ。

次の補題では証明の全般にわたって補題 2.12 を適用する。その使いどころをみてほしい。

記号

関数 $f : S \to \mathbb{R}$ に対し $\sigma\{f\} := \{f^{-1}(B)\,;B \in \mathrm{Borel}(\mathbb{R})\}$ と書く。

2.14 補題. (i) $\sigma\{f\}$ は S 上の σ 加法族である。
(ii) $\mathcal{B}$ を S 上の σ 加法族とする。このとき $\{B \subset \mathbb{R} : f^{-1}(B) \in \mathcal{B}\}$ は $\mathbb{R}$ 上の σ 加法族である。
(iii) $\sigma\{f\} = \sigma_S(\{f^{-1}((a, +\infty))\,; a \in \mathbb{R}\})$.

証明. (i) まず $\emptyset \in \mathrm{Borel}(\mathbb{R})$ により $\emptyset = f^{-1}(\emptyset) \in \sigma\{f\}$ である。次に $A \in \sigma\{f\}$ とすると $\exists B \in \mathrm{Borel}(\mathbb{R})$ s.t. $A = f^{-1}(B)$ であるが、$\mathbb{R} \setminus B \in \mathrm{Borel}(\mathbb{R})$ により、

$$S \setminus A = S \setminus f^{-1}(B) = f^{-1}(\mathbb{R} \setminus B) \in \sigma\{f\}$$

が従う。最後に $A_n \in \sigma\{f\}$ $\forall n \in \mathbb{N}$ とすると $\exists B_n \in \mathrm{Borel}(\mathbb{R})$ s.t. $A_n = f^{-1}(B_n)$ であるが、$\bigcup_{n=1}^{\infty} B_n \in \mathrm{Borel}(\mathbb{R})$ により、

$$\bigcup_{n=1}^{\infty} A_n = \bigcup_{n=1}^{\infty} f^{-1}(B_n) = f^{-1}(\bigcup_{n=1}^{\infty} B_n) \in \sigma\{f\}$$

となる。(ii) を示すのは演習問題とする。

(iii) S 上の σ 加法族 $\sigma(\{f^{-1}((a, +\infty))\,; a \in \mathbb{R}\})$ に対して (ii) を適用すると

$$\{B \subset \mathbb{R} : f^{-1}(B) \in \sigma(\{f^{-1}((a, +\infty))\,; a \in \mathbb{R}\})\}$$

は $\mathbb{R}$ 上の σ 加法族であることがわかる。しかもその定義により $(a, +\infty)$ という形の区間はすべて属する。補題 2.10(i) によれば、そのようなものの最小が $\mathrm{Borel}(\mathbb{R})$ であるから、

$$\mathrm{Borel}(\mathbb{R}) \subset \{B \subset \mathbb{R} : f^{-1}(B) \in \sigma(\{f^{-1}((a, +\infty))\,; a \in \mathbb{R}\})\}$$

が従う。これは $f^{-1}(B) \in \sigma(\{f^{-1}((a, +\infty))\,; a \in \mathbb{R}\})$ $\forall B \in \mathrm{Borel}(\mathbb{R})$ を意味する。さらに $\sigma\{f\}$ の定義により次のように書き換えられる。

$$\sigma\{f\} \subset \sigma(\{f^{-1}((a, +\infty))\,; a \in \mathbb{R}\}).$$

逆向きの包含関係は $\sigma\{f\}$ が S 上の σ 加法族であることと $f^{-1}((a, +\infty)) \in \sigma\{f\}$ $\forall a \in \mathbb{R}$ であることから導かれる。 □

2.15 定義. $\sigma\{f\}$ は f によって**生成される σ 加法族**と呼ばれる。あるいは f を**可測にする最小の σ 加法族**と呼ぶこともある。

2.16 演習問題. 補題 2.14(ii) を示せ。

2.17 定理. 関数 $f : S \to \mathbb{R}$ と S 上の σ 加法族 $\mathcal{B}$ に対し、f が $\mathcal{B}$ 可測であるための必要十分条件は $f^{-1}(B) \in \mathcal{B}$ $\forall B \in \mathrm{Borel}(\mathbb{R})$ である。

証明. $\mathcal{B}$ 可測性は $\{f^{-1}((a,+\infty))\,;a\in\mathbb{R}\}\subset\mathcal{B}$ と表現でき、条件 $f^{-1}(B)\in\mathcal{B}\ \forall B\in\mathrm{Borel}(\mathbb{R})$ は $\sigma\{f\}\subset\mathcal{B}$ と表現できる。従って目標は次の同値性を示すこととなる。

$$\{f^{-1}((a,+\infty))\,;a\in\mathbb{R}\}\subset\mathcal{B}\Leftrightarrow\sigma\{f\}\subset\mathcal{B}$$

さて $\mathcal{B}$ は σ 加法族なので左の条件は $\sigma(\{f^{-1}((a,+\infty))\,;a\in\mathbb{R}\})\subset\mathcal{B}$ とよめる。ところが補題 2.14(iii) によれば、まさに $\sigma(\{f^{-1}((a,+\infty))\,;a\in\mathbb{R}\})=\sigma\{f\}$ である。 □

2.18 系. 関数 $X:\Omega\to\overline{\mathbb{R}}$ が確率空間 $(\Omega,\mathcal{F},P)$ 上の確率変数であるための必要十分条件は $X^{-1}(B)\in\mathcal{F}\ \forall B\in\mathrm{Borel}(\mathbb{R})$ かつ $X^{-1}(\{+\infty\})\in\mathcal{F}$ である。

証明. 定理 2.17 との違いは、関数が $+\infty$, $-\infty$ なる値を取りうるかどうかである。今の段階では、この点は些細なこととして気にしなくてもよい。略証をあげるにとどめよう。まず

$$\mathrm{Borel}(\mathbb{R})\cup\{\overline{\mathbb{R}}\}\cup\{A\cup\{+\infty\};A\in\mathrm{Borel}(\mathbb{R})\}\cup\{A\cup\{-\infty\};A\in\mathrm{Borel}(\mathbb{R})\}$$

は $\overline{\mathbb{R}}$ 上の σ 加法族である。これは集合族 $\{(a,+\infty)\cup\{+\infty\}\,;a\in\mathbb{R}\}$ で生成される $\overline{\mathbb{R}}$ 上の σ 加法族と一致する。あとは定理 2.17 の証明に倣えばよい。 □

2.19 例. $\alpha\in\mathbb{R}_{>0}$, $\beta\in\mathbb{R}$ とする。1 次関数 $f:\mathbb{R}\to\mathbb{R}$, $x\mapsto\alpha^{-1}(x-\beta)$ を考えよう。連続関数ゆえ f は $\mathrm{Borel}(\mathbb{R})$ 可測である（演習問題 2.11)。よって定理 2.17 を適用して

$$\alpha B+\beta=\{\alpha x+\beta;x\in B\}=f^{-1}(B)\in\mathrm{Borel}(\mathbb{R})\ \forall B\in\mathrm{Borel}(\mathbb{R})$$

が導かれる。当然ながら $\alpha<0$ であってもこれは成り立つ。

2.20 演習問題. $a,b\in\mathbb{R}$, $a<b$ に対して集合族 $\{B\subset(a,b]:B\in\mathrm{Borel}(\mathbb{R})\}$ は半開区間 $(a,b]$ 上の σ 加法族であることを示せ。

2.21 注意. $\{B\subset(a,b]:B\in\mathrm{Borel}(\mathbb{R})\}$ が半開区間 $(a,b]$ 上の σ 加法族であることを補題 2.14(i) の観点から確認してみるには包含写像 $\iota:(a,b]\to\mathbb{R}$ を導入するとよい。実際すぐ分かるように $\sigma\{\iota\}=\{B\subset(a,b]:B\in\mathrm{Borel}(\mathbb{R})\}$ である。

記号

$a,b\in\mathbb{R}$, $a<b$ に対して σ 加法族 $\{B\subset(a,b]:B\in\mathrm{Borel}(\mathbb{R})\}$ を $\mathrm{Borel}((a,b])$ と書く。

2.22 例. 1 節で述べた関数 $\xi_1:(0,1]\to\mathbb{R}$ は補題 2.4 により $\mathrm{Borel}((0,1])$ 可測である。

3 確率変数と分布–Lebesgue 積分論からの準備

確率空間 $(\Omega,\mathcal{F},P)$ が直に見えるわけではなくて、確率変数を介在して観測するという視点に立つとき、確率もまた観測にかかる集合上に実現されるというのが自然である。そのようなとき実現された確率を分布という。着目点は系 2.18 にある。即ち X が確率変数なら

$$X^{-1}(B)\in\mathcal{F}\ \forall B\in\mathrm{Borel}(\mathbb{R})$$

記号

確率変数 X と $B \in \mathrm{Borel}(\mathbb{R})$ に対し $\mathcal{L}(X, B) := P(X^{-1}(B))$ と書く。

確率論の慣用では $P(X^{-1}(B))$ を $P(X \in B)$ と表現することが多い。

3.1 演習問題. X を確率変数とする。このときつぎを示せ（補題 2.12 の役割が重要）。
(i) $\mathcal{L}(X, \emptyset) = 0$.
(ii) $B_n \in \mathrm{Borel}(\mathbb{R})\ \forall n \in \mathbb{N},\ B_n \cap B_m = \emptyset\ n \neq m \Rightarrow \mathcal{L}(X, \bigcup_{n=1}^{\infty} B_n) = \sum_{n=1}^{\infty} \mathcal{L}(X, B_n)$.
(iii) $\mathcal{L}(X, \mathbb{R}) = 1 \Leftrightarrow P(X = -\infty \text{ or } X = +\infty) = 0$.

3.2 補題. 関数 $\mathcal{L}(X, \cdot) : \mathrm{Borel}(\mathbb{R}) \to \overline{\mathbb{R}},\ B \mapsto P(X^{-1}(B))$ は測度である。これが確率測度であるための必要十分条件は $P(X = -\infty \text{ or } X = +\infty) = 0$ である。

証明. 演習問題 3.1 で確かめたとおりである。 □

通常は $P(X = -\infty \text{ or } X = +\infty) = 0$ を満たす場合を考えることが多い。

3.3 定義. (i) 確率変数 X に対し測度 $\mathcal{L}(X, \cdot)$ を X の**分布**(distribution) あるいは**法則**(law) という。$P(X = -\infty \text{ or } X = +\infty) = 0$ ならそれは $(\mathbb{R}, \mathrm{Borel}(\mathbb{R}))$ 上の確率測度であり、特に X の**確率分布**(probability distribution) あるいは**確率法則**(propability law) という。
(ii) $(\mathbb{R}, \mathrm{Borel}(\mathbb{R}))$ 上の確率測度 μ が与えられたとする。確率変数 X が**分布** μ **に従う**とは $\mathcal{L}(X, \cdot) = \mu$ が成り立つことをいう。

3.4 注意. 写像 $X : \Omega \to \mathbb{R}$ によって値域 $\mathbb{R}$ 上に測度 $\mathcal{L}(X, \cdot)$ が定義域 Ω 上の測度 P から誘導されたわけである。この観点から $\mathcal{L}(X, \cdot)$ のことを写像 X による P の**像測度**(image measure) あるいは**誘導測度**(induced measure) と呼ぶことも多い。

重要な確率測度の多くは積分を使って表現されるので復習しておこう。$(S, \mathcal{B})$ を非自明な可測空間とする。次の条件を満たす関数 $g : S \to \overline{\mathbb{R}}$ を $\mathcal{B}$-**単関数**(simple function) と呼ぶ。

$\mathcal{B}$-可測、$-\infty, +\infty$ の値はとらない、$\mathrm{Image}\, g := \{g(x)\,;\, x \in S\}$ は有限集合

記号

$S \neq \emptyset$ とする。$A \subset S$ に対し A の S に関する**定義関数**(indicator function) を $1_{A:S}$ と表わす。前後関係から分かるときは S を省いて 1_A と書くことが多い。

$$s \in A \ \Rightarrow\ 1_{A:S}(s) = 1,\, s \in S \setminus A \ \Rightarrow\ 1_{A:S}(s) = 0$$

$\mathcal{B}$-単関数 $g : S \to \overline{\mathbb{R}}$ は以下のように表現される。

$$g = \sum_{y \in \mathrm{Image}\, g} y 1_{g^{-1}(\{y\}):S}$$

測度論においては通常 $0\infty = 0$ という既約をもうける。以下ではこれを採用する。

復習

μ を $(S,\mathcal{B})$ 上の測度とする。非負値 $\mathcal{B}$-可測関数 $f : S \to \overline{\mathbb{R}}$ に対し量

$$\int_S f\,\mu := \sup\{\sum_{y\in \mathrm{Image}\,g} y\mu(g^{-1}(\{y\}))\,;g\text{ 非負値 }\mathcal{B}\text{-単関数}, g\leq f\}$$

を**積分**(integral) と言う。f 自身が単関数なら $\displaystyle\sum_{y\in \mathrm{Image}\,f} y\mu(f^{-1}(\{y\}))$ に一致する。

3.5 演習問題. $f : S \to \overline{\mathbb{R}}$ を $\mathcal{B}$-可測関数とするとき、$\max\{f,0\}$ ならびに $|f|$ も $\mathcal{B}$-可測であることを示せ。

復習の続き

一般に $f : S \to \overline{\mathbb{R}}$ を $\mathcal{B}$-可測関数とする。

$$\int_S \max\{f,0\}\,\mu < +\infty \text{ または } \int_S \max\{-f,0\}\,\mu < +\infty$$

であるとき f の μ についての**積分**を次で定義する。

$$\int_S f\,\mu := \int_S \max\{f,0\}\,\mu - \int_S \max\{-f,0\}\,\mu.$$

$\displaystyle\int_S |f|\,\mu < +\infty$ が成り立つとき f は μ-**可積分**(integrable) であるという。

まずいわゆる**離散型の分布**(distribution of discrete type) を紹介する。

3.6 補題. $(S,\mathcal{B})$ を非自明な可測空間、$s\in S$ とする。このとき関数 $\mu : \mathcal{B}\to\overline{\mathbb{R}}$, $A\mapsto 1_A(s)$ は測度であり、非負値 $\mathcal{B}$ 可測関数 $f : S\to\overline{\mathbb{R}}$ に対して $\displaystyle\int_S f\,\mu = f(s)$ である。

証明. $A_n\in\mathcal{B}$ $\forall n\in\mathbb{N}$, $A_n\cap A_m=\emptyset$ $n\neq m$ として $A=\bigcup_{n=1}^{\infty} A_n$ とおく。互いに交わらないので $s\in A_n$ となるような番号 n があったとしてもそれは高々一つである。よって

$$\sum_{n=1}^{\infty} 1_{A_n}(s) = 0 = 1_A(s) \qquad s\in A_n \text{ となるような番号 } n \text{ が存在しない}$$
$$\sum_{n=1}^{\infty} 1_{A_n}(s) = 1 = 1_A(s) \qquad s\in A_n \text{ となるような番号 } n \text{ が存在する}$$

というように場合分けされる。これで μ の σ 加法性が確かめられた。ほかの測度としての要件は明らかに満たされている。次に $\displaystyle\int_S f\,\mu = f(s)$ の検証に取りかかるがまず f が単関数である場合を考えよう。$\mathrm{Image}\,f$ は有限集合である。$y\in\mathrm{Image}\,f$ としよう。μ の定義より次が

成り立つ。

$$\mu(f^{-1}(\{y\})) = \begin{cases} 1 & s \in f^{-1}(\{y\}) \ (\Leftrightarrow y = f(s) \ ) \\ 0 & s \notin f^{-1}(\{y\}) \ (\Leftrightarrow y \neq f(s) \ ) \end{cases}$$

従って積分の線形性を適用して次を得る。

$$\int_S f\,\mu = \sum_{y \in \mathrm{Image}\, f} y\mu(f^{-1}(\{y\})) = \sum_{y \in \mathrm{Image}\, f : y = f(s)} y = f(s)$$

一般には非負値 $\mathcal{B}$-単関数 $S \to \mathbb{R}$ の列 f_n で $f_n \le f_{n+1}\ \forall n \in \mathbb{N}$, $\sup_{n\in\mathbb{N}} f_n(x) = f(x)\ \forall x \in S$ を満たすものが存在する。各 $n \in \mathbb{N}$ に対して

$$\int_S f_n\,\mu = f_n(s)$$

が成り立つ。単調収束定理によれば左辺は $\displaystyle\int_S f\,\mu$ に収束する。 $\square$

記号

$(S, \mathcal{B})$ を非自明な可測空間とする。各 $a \in S$ に対して測度 $\mathcal{B} \to \overline{\mathbb{R}}$, $A \mapsto 1_{A:S}(a)$ を a に質量 (mass) を持つ S 上の *Dirac* **測度**といい、δ_a^S と表記する。前後関係から分かるときは S を省いて δ_a と書くことが多い。

3.7 補題. $(S, \mathcal{B})$ を非自明な可測空間、$\mu_n\ n \in \mathbb{N}$ をそれ上の測度の列、また $\alpha_n\ n \in \mathbb{N}$ を非負実数列とする。このとき関数 $\mu : \mathcal{B} \to \overline{\mathbb{R}}$, $A \mapsto \sum_{n=1}^{\infty} \alpha_n \mu_n(A)$ は測度であり、非負値 $\mathcal{B}$ 可測関数 $f : S \to \overline{\mathbb{R}}$ に対して

$$\int_S f\,\mu = \sum_{n=1}^{\infty} \alpha_n \int_S f\,\mu_n$$

3.8 演習問題. 補題 3.7 を補題 3.6 の証明手順に従って示せ。

記号

$(S, \mathcal{B})$ を非自明な可測空間とする。測度の列 μ_n と非負実数列 α_n に対して測度 $\mathcal{B} \to \overline{\mathbb{R}}$, $A \mapsto \sum_{n=1}^{\infty} \alpha_n \mu_n(A)$ を $\sum_{n=1}^{\infty} \alpha_n \mu_n$ と書く。

3.9 系. $p_k\ k \in \mathbb{Z}$ を非負実数列であって $\sum_{k\in\mathbb{Z}} p_k = 1$ を満たすものとする。このとき

$$\mu = \sum_{k \in \mathbb{Z}} p_k \delta_k$$

は $(\mathbb{R}, \mathrm{Borel}(\mathbb{R}))$ 上の確率測度であり、非負値 $\mathrm{Borel}(\mathbb{R})$ 可測関数 f に対して

$$\int_{\mathbb{R}} f\,\mu = \sum_{k \in \mathbb{Z}} p_k f(k)$$

3.10 例. 与えられたパラメータ $n \in \mathbb{N}$, $p \in (0,1)$ に対して $(\mathbb{R}, \mathrm{Borel}(\mathbb{R}))$ 上の測度

$$\sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} \delta_k \text{ ここで } \binom{n}{k} := \frac{n!}{k!(n-k)!} \text{ 二項係数}$$

を**二項分布**(binomial distribution) という。これを μ と書くと系 3.9 により $z \in \mathbb{R}$ に対し

$$\int_{\mathbb{R}} e^{zx} \mu(dx) = \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} e^{zk} = (pe^z + 1 - p)^n$$

ここで $e^{zx} > 0$ であることに注意する。2 番目の等号は二項定理による。また

$$\int_{\mathbb{R}} x\mu(dx) = np, \int_{\mathbb{R}} x^2 \mu(dx) = np + n(n-1)p^2$$

3.11 例. 与えられたパラメータ $n \in \mathbb{N}$, $p \in (0,1)$ に対して $(\mathbb{R}, \mathrm{Borel}(\mathbb{R}))$ 上の測度

$$\sum_{k=0}^{\infty} \binom{n+k-1}{k} p^n (1-p)^k \delta_k$$

を**負の二項分布**(negative binomial distribution) という。これを μ と書くと $z \le 0$ に対し

$$\int_{\mathbb{R}} e^{zx} \mu(dx) = \sum_{k=0}^{\infty} \binom{n+k-1}{k} p^n (1-p)^k e^{zk} = \frac{p^n}{(1-(1-p)e^z)^n}$$

ここで $\binom{n+k-1}{k} = \binom{-n}{k}(-1)^k$ であることと二項定理を使っている。また

$$\int_{\mathbb{R}} x\mu(dx) = \frac{n(1-p)}{p}, \int_{\mathbb{R}} x^2 \mu(dx) = \frac{n(1-p)}{p} + \frac{n(n+1)(1-p)^2}{p^2}$$

3.12 例. 与えられたパラメータ $c \in \mathbb{R}_{>0}$ に対し $(\mathbb{R}, \mathrm{Borel}(\mathbb{R}))$ 上の測度

$$\sum_{k=0}^{\infty} \frac{e^{-c} c^k}{k!} \delta_k$$

を強度 (intensity) c の *Poisson* **分布**という。これを μ と書くと

$$\int_{\mathbb{R}} x\mu(dx) = c, \int_{\mathbb{R}} x^2 \mu(dx) = c + c^2, \int_{\mathbb{R}} e^{zx} \mu(dx) = \exp\{c(e^z - 1)\} \ z \in \mathbb{R}.$$

3.13 演習問題. 例 3.10, 例 3.11 および例 3.12 の計算をフォローせよ。

復習

次を満たす $(\mathbb{R}, \mathrm{Borel}(\mathbb{R}))$ 上の測度が一意に存在する。

$$\lambda((a,b]) = b - a \ \forall a, \forall b \in \mathbb{R},\ a < b$$

この測度 λ を 1 **次元** *Lebesgue* **測度**という。

ここで Lebesgue 測度が一意であるとは次が成り立つことをいう。実際の適用例としては補題 3.21 の証明を見よ。

重要

$(\mathbb{R}, \mathrm{Borel}(\mathbb{R}))$ 上の測度 μ が $\mu((a,b]) = b-a \ \forall a < \forall b$ を満たすなら $\mu(A) = \lambda(A) \ \forall A \in \mathrm{Borel}(\mathbb{R})$ である。

状況によっては、Lebesgue 可測な集合全体から構成される σ-加法族、*Lebesgue* **可測集合族**という、を考察の対象にする必要があるが、ここではそうしない。（Lebesgue 可測集合族は完備な σ-加法族となるので都合がよいことがある。他方それは Lebesgue 測度が 0 であるような集合全体と $\mathrm{Borel}(\mathbb{R})$ から生成される σ-加法族と一致することが示せる。）

約束

$\mathrm{Borel}(\mathbb{R})$ 可測関数 f に対してそれが Lebesgue 測度に関して可積分であることを単に *Lebesgue* **可積分**(Lebesgue integrable) といい、Lebesgue 可積分関数 f の Lebesgue 測度 λ に関する積分 $\int_{\mathbb{R}} f\,\lambda$ を *Lebesgue* **積分**と呼ぶ。

3.14 補題. μ を $(\mathbb{R}, \mathrm{Borel}(\mathbb{R}))$ 上の測度、$a, b \in \mathbb{R}$, $a < b$ とする。$\mu((a,b]) = 1$ であれば、関数 μ の定義域を $\mathrm{Borel}((a,b])$ に制限したものは $((a,b], \mathrm{Borel}((a,b]))$ 上の確率測度である。

証明. μ の σ 加法性は定義域の $\mathrm{Borel}((a,b])$ への制限によって壊れることはない。 □

記号

Lebesgue 測度の定義域を $\mathrm{Borel}((0,1])$ に制限したものもやはり λ で記述する。確率空間 $((0,1], \mathrm{Borel}((0,1]), \lambda)$ をここでは *Lebesgue* **モデル**という。

3.15 例. 第 1 節で登場した関数 $\xi_1 : (0,1] \to \mathbb{R}$ は例 2.22 により Lebesgue モデル上の確率変数である。その分布は $p\delta_1 + (1-p)\delta_{-1}$ である。

3.16 演習問題. Lebesgue モデルで考える。$n \in \mathbb{N}$, $p \in (0,1)$ とする。

$$X(\omega) := k \text{ if } \sum_{i=0}^{k-1} \binom{n}{i} p^i (1-p)^{n-i} < \omega \leq \sum_{i=0}^{k} \binom{n}{i} p^i (1-p)^{n-i} \ k = 0, 1, \ldots, n$$

このとき確率変数 X の分布はパラメータ n, p に対応する二項分布であることを示せ。

3.17 演習問題. $\rho : \mathbb{R} \to \mathbb{R}$ を連続関数とする。ρ は非負値かつある $a \in \mathbb{R}$ において $\rho(a) > 0$ なら $\int_{\mathbb{R}} \rho\,\lambda > 0$ であることを示せ。

3.18 演習問題. $(S, \mathcal{B})$ を可測空間、$f, g : S \to \overline{\mathbb{R}}$ を $\mathcal{B}$-可測関数とするとき、fg も $\mathcal{B}$-可測であることを示せ。既約 $0\infty = 0$ を採用していたことに注意。

3.19 定理. $\rho : \mathbb{R} \to \overline{\mathbb{R}}$ を非負値 Borel($\mathbb{R}$) 可測関数で $\int_{\mathbb{R}} \rho\,\lambda = 1$ を満たすものとする。
(i) 関数 $\mu : \mathrm{Borel}(\mathbb{R}) \to \overline{\mathbb{R}},\ A \mapsto \displaystyle\int_A \rho\,\lambda$ は $(\mathbb{R}, \mathrm{Borel}(\mathbb{R}))$ 上の確率測度である。
(ii) 任意の非負値 Borel($\mathbb{R}$) 可測関数 $f : \mathbb{R} \to \overline{\mathbb{R}}$ に対して次が成り立つ。

$$\int_{\mathbb{R}} f\,\mu = \int_{\mathbb{R}} f\rho\,\lambda$$

(iii) 任意の Borel($\mathbb{R}$) 可測関数 $f : \mathbb{R} \to \overline{\mathbb{R}}$ に対してその μ 可積分性は $f\rho$ の λ 可積分性と同値であり、可積分であるなら (ii) であげた関係が成り立つ。

証明. (i) $A_n \in \mathrm{Borel}(\mathbb{R})\ \forall n \in \mathbb{N},\ A_n \cap A_m = \emptyset\ n \neq m$ とする。このとき

$$1_A = \sum_{n=1}^{\infty} 1_{A_n} \text{ ただし } A := \bigcup_{n=1}^{\infty} A_n$$

である。従って非負関数項級数についての項別積分定理を適用して次が導ける。

$$\int_A \rho\,\lambda = \sum_{n=1}^{\infty} \int_{A_n} \rho\,\lambda$$

これは μ の σ 加法性を意味する。ほかの測度としての要件は明らかに満たされている。
(ii) まず f が単関数である場合を考える。このとき積分の線形性により

$$\int_{\mathbb{R}} f\,\mu = \sum_{y \in \mathrm{Image}\, f} y\mu(f^{-1}(\{y\})) = \sum_{y \in \mathrm{Image}\, f} y \int_{f^{-1}(\{y\})} \rho\,\lambda = \int_{\mathbb{R}} \sum_{y \in \mathrm{Image}\, f} y 1_{f^{-1}(\{y\})} \rho\,\lambda$$

となるが、右辺の被積分関数はちょうど $f\rho$ である。一般には非負値 Borel($\mathbb{R}$)-単関数 $\mathbb{R} \to \mathbb{R}$ の列 f_n で $f_n \leq f_{n+1}\ \forall n \in \mathbb{N},\ \sup_{n \in \mathbb{N}} f_n = f$ を満たすものが存在する。各 $n \in \mathbb{N}$ に対して

$$\int_{\mathbb{R}} f_n\,\mu = \int_{\mathbb{R}} f_n\rho\,\lambda$$

が成り立つ。単調収束定理によれば左辺は $\displaystyle\int_{\mathbb{R}} f\,\mu$ に右辺は $\displaystyle\int_{\mathbb{R}} f\rho\,\mu$ にそれぞれ収束する。
(iii) 非負値関数 $|f|$ に対して (ii) を適用して可積分性に関する同値性が示せる。次に (ii) を $\max\{f, 0\}$ と $\max\{-f, 0\}$ のそれぞれに適用すると求める関係式が得られる。 □

3.20 定義. 定理 3.19 で規定されるような確率測度は（測度 λ に関して）**絶対連続**(absolutely continuous) であるといい、そのような確率測度を**絶対連続分布**(absolutely continuous distribution) と呼ぶ。また関数 ρ を（λ に関する）**密度関数**(density function) という。

定理 3.19 の証明に用いた論法、すなわち単関数について証明し次に単調収束定理によって処理するという手順、は非常に有効なものであり、**スタンダードマシン**(standard machine) と呼ぶ研究者もいる。じつは補題 3.6 の証明において既に登場していたので確認しておこう。ポイントがつかめたかを試すには以下の演習問題 3.22 を解いてみるとよいだろう。

3.21 補題. $f : \mathbb{R} \to \overline{\mathbb{R}}$ を Borel($\mathbb{R}$) 可測関数、$\alpha, \beta \in \mathbb{R},\ \alpha \neq 0$ とする。このとき関数 $x \mapsto f(\alpha x + \beta)$ も Borel($\mathbb{R}$) 可測である。非負値であれば次が成り立つ。

$$\int_{\mathbb{R}} f(\alpha x + \beta)\,\lambda(dx) = \frac{1}{|\alpha|}\int_{\mathbb{R}} f(x)\,\lambda(dx)$$

また $x \mapsto f(\alpha x + \beta)$ の λ 可積分性は f のそれと同値であり、可積分なら上の等式が成り立つ。

証明. $g(x) := f(\alpha x + \beta)$ とおく。注目点は次である。

$$g^{-1}(B) = \alpha^{-1}(f^{-1}(B) - \beta)\ \forall B \subset \mathbb{R}$$

定理 2.17 と例 2.19 適用して、次が成り立つことが分かる。

$$B \in \mathrm{Borel}(\mathbb{R}) \Rightarrow f^{-1}(B) \in \mathrm{Borel}(\mathbb{R}) \Rightarrow \alpha^{-1}(f^{-1}(B) - \beta) \in \mathrm{Borel}(\mathbb{R})$$

従って定理 2.17 により関数 g の Borel($\mathbb{R}$) 可測性がわかる。f が $A \in \mathrm{Borel}(\mathbb{R})$ の定義関数であるとする。このとき g は集合 $\alpha^{-1}(A - \beta)$ の定義関数である。従って次が成り立つ。

$$(\star) \qquad \int_{\mathbb{R}} g\,\lambda = \lambda(\alpha^{-1}(A - \beta)).$$

さて関数 $A \mapsto \lambda(\alpha^{-1}(A - \beta))$ は写像 $x \mapsto \alpha x + \beta$ による測度 λ の誘導測度を表す。ここで

$$|\alpha|\lambda(\alpha^{-1}(a - \beta, b - \beta]) = \begin{cases} |\alpha|\lambda((a/\alpha - \beta/\alpha, b/\alpha - \beta/\alpha]) & \alpha > 0 \\ |\alpha|\lambda([b/\alpha - \beta/\alpha, a/\alpha - \beta/\alpha)) & \alpha < 0 \end{cases}$$

であるが、これは $b - a$ に等しい。従って Lebesgue 測度の一意性により

$$|\alpha|\lambda(\alpha^{-1}(A - \beta)) = \lambda(A)\ \forall A \in \mathrm{Borel}(\mathbb{R})$$

であり、($\star$) とあわせて f が Borel 集合の定義関数である場合に以下が導かれた。

$$\int_{\mathbb{R}} g\,\lambda = \frac{\lambda(A)}{|\alpha|} = \frac{1}{|\alpha|}\int_{\mathbb{R}} f\,\lambda$$

一般の非負関数および可積分関数の場合の議論は演習問題に委ねる。 □

3.22 演習問題. 補題 3.21 について残された部分を実行せよ。

ここまで複素数値関数を避けてきたが、いろいろ不便が生じるので対処しておこう。

記号

$\alpha \in \mathbb{C}$ についてその実部を $\mathrm{Re}\,\alpha$ 虚部を $\mathrm{Im}\,\alpha$ と表記する。即ち $\alpha = \mathrm{Re}\,\alpha + \sqrt{-1}\,\mathrm{Im}\,\alpha$

3.23 定義. $(S, \mathcal{B})$ を非自明な可測空間、μ をそれ上の測度とする。関数 $f : S \to \mathbb{C}$ の $\mathcal{B}$ 可測性あるいは μ 可積分性は $x \mapsto \mathrm{Re}\,f(x)$ と $x \mapsto \mathrm{Im}\,f(x)$ の両方がそれぞれ対応する性質を持つことをいう。$\mathcal{B}$ 可測かつ μ 可積分なとき f の μ についての積分を次で定義する。

$$\int_S f\,\mu := \int_S \mathrm{Re}\,f\,\mu + \sqrt{-1}\int_S \mathrm{Im}\,f\,\mu.$$

複素数に関わるときは面倒をさけるため ∞ を除外しておく。多くの場合、複素数値関数の積分に関する命題は実部と虚部に分けて実数値バージョンの命題を適用すれば導ける。

3.24 演習問題. $f, g : S \to \mathbb{C}$ を $\mathcal{B}$ 可測関数とする。
(i) f の μ 可積分性は条件 $\int_S |f|\,\mu < +\infty$ と同値であることを示せ。
(ii) $a, b \in \mathbb{C}$ とする。線形結合 $af + bg$ も $\mathcal{B}$ 可測であることを示せ。また f, g ともに μ 可積分なら $af + bg$ も μ 可積分であり次が成り立つことを示せ。

$$\int_S (af + bg)\,\mu = a\int_S f\,\mu + b\int_S g\,\mu \text{ 複素線形性}$$

次は演習問題に含めてもよかったのだが少し工夫が必要なので証明をつけておく。

3.25 補題. $f : S \to \mathbb{C}$ を $\mathcal{B}$ 可測関数する。μ 可積分なら次が成り立つ。

$$\Big|\int_S f\,\mu\Big| \le \int_S |f|\,\mu.$$

証明. 一般に $z \in \mathbb{C}$ に対して $|z| = \max\{\mathrm{Re}(\overline{a}z); a \in \mathbb{C}, |a| = 1\}$ である（$\overline{a}$ は a の複素共役)。そこで $a \in \mathbb{C}$, $|a| = 1$ とする。積分の複素線形性と単調性により

$$\mathrm{Re}\Big(\overline{a}\int_S f\,\mu\Big) = \int_S \mathrm{Re}(\overline{a}f)\,\mu \le \int_S |f|\,\mu.$$

左辺の a に関する最大値が $|\int_S f\,\mu|$ である。 □

3.26 補題. $a, b \in \overline{\mathbb{R}}$, $a < b$ かつ関数 $f : (a, b) \to \mathbb{C}$ は連続かつ Lebesgue 可積分とする。
(i) 関数 $(a, b) \to \mathbb{C}$, $x \mapsto \displaystyle\int_{(a,x]} f\,\lambda$ は f の原始関数である。**原始関数の存在**
(ii) 関数 f の原始関数の一つを $F : (a, b) \to \mathbb{C}$ とすると極限 $\lim_{x\to a} F(x)$, $\lim_{x\to b} F(x)$ が存在し $\displaystyle\int_{(a,b)} f\,\lambda = \lim_{x\to b} F(x) - \lim_{x\to a} F(x)$ が成り立つ。**微積分の基本定理**

証明. (i) $c \in (a, b)$ における微分可能性を議論しよう。任意に $\varepsilon > 0$ が与えられたとする。連続性により $\exists\delta > 0$ s.t. $|y - c| < \delta \Rightarrow |f(y) - f(c)| < \varepsilon$ となる。さて $c \le x < b$ のとき

$$\int_{(a,x]} f\,\lambda - \int_{(a,c]} f\,\lambda - f(c)(x - c) = \int_{(c,x]} (f - f(c))\,\lambda$$

より、補題 3.25 を適用して $c \le x < \min\{c + \delta, b\}$ なら

$$\Big|\int_{(a,x]} f\,\lambda - \int_{(a,c]} f\,\lambda - f(c)(x - c)\Big| \le \int_{(c,x]} |f - f(c)|\,\lambda \le \varepsilon|x - c|$$

が成り立つことを得る。$\varepsilon|x-c|$ で抑えるという評価は $\max\{c-\delta, a\} < x \le c$ であっても有効である。よって $x \mapsto \displaystyle\int_{(a,x]} f\,\lambda$ は c において微分可能であり、微分係数は $f(c)$ に等しい。 □

3.27 演習問題. 補題 3.26(ii) を示せ。Lebesgue 可積分という条件が重要である。

原始関数の正体がよく分かっているときは次の事実に基づいて**可積分性判定**を行う。その効力は優関数判定法と協調して発揮されることが多い。

3.28 系. $a, b \in \overline{\mathbb{R}}$ かつ $a < b$ とする。
(i) 連続関数 $f : (a, b) \to \mathbb{C}$ は原始関数をもつ。
(ii) 関数 $f : (a, b) \to \mathbb{R}$ は非負値かつ連続であるとする。このとき関数 f の原始関数の一つを $F : (a, b) \to \mathbb{R}$ とすると次が成り立つ。とくに、原始関数が有界なら f は可積分である。

$$\int_{(a,b)} f\,\lambda = \sup_{x \in (a,b)} F(x) - \inf_{x \in (a,b)} F(x)$$

3.29 演習問題. 系 3.28 を示せ。

3.30 例. 関数 $\mathbb{R} \to \mathbb{R},\ x \mapsto e^{-x^2/2}$ は Lebesgue 可積分である。

証明. そのものの原始関数がよく分からないときは、なじみのある優関数を見つけて優関数判定法を適用するというのが常套手段である。さて $2|x| \leq x^2 + 1\ \forall x$ であるから

$$0 < e^{-x^2/2} \leq e^{-|x|+1/2}\ \forall x \in \mathbb{R}$$

$\mathbb{R}$ を $(-\infty, 0)$, $\{0\}$ と $(0, +\infty)$ に分割し、変数変換公式（補題 3.21）$(-\infty, 0)$ 上の積分に適用することにより

$$\int_{\mathbb{R}} e^{-x^2/2} \lambda(dx) = 2 \int_{(0,+\infty)} e^{-x^2/2} \lambda(dx) \leq 2 \int_{(0,+\infty)} e^{-x+1/2} \lambda(dx) = 2e^{1/2} < +\infty$$

を得る。2 番目の等号は系 3.28(ii) による。 □

3.31 例. (i) 各 $t > 0$ に対して関数 $(0, +\infty) \to \mathbb{R},\ x \mapsto x^{t-1} e^{-x}$ は Lebesgue 可積分である。$(0, +\infty)$ 上で定義された次の関数を *gamma* **関数**(gamma function) という。

$$\Gamma : t \mapsto \int_{(0,+\infty)} x^{t-1} e^{-x}\ \lambda(dx)$$

(ii) 各 $s, t > 0$ に対して関数 $(0, 1) \to \mathbb{R},\ x \mapsto (1 - x)^{s-1} x^{t-1}$ は Lebesgue 可積分である。$(0, +\infty) \times (0, +\infty)$ 上で定義された次の関数を *beta* **関数**(beta function) という。

$$B : (s, t) \mapsto \int_{(0,1)} (1 - x)^{s-1} x^{t-1}\ \lambda(dx)$$

証明. (i) まず $0 < t \leq 1$ とする。$x^{t-1} e^{-x} \leq x^{t-1}\ \forall x \in (0, 1]$, $x^{t-1} e^{-x} \leq e^{-x}\ \forall x \in (1, +\infty)$ であるから、$(0, +\infty)$ を $(0, 1]$ と $(1, +\infty)$ に分割し系 3.28(ii) を適用することにより

$$\int_{(0,+\infty)} x^{t-1} e^{-x}\ \lambda(dx) \leq \int_{(0,1]} x^{t-1}\ \lambda(dx) + \int_{(1,+\infty)} e^{-x}\ \lambda(dx) = \frac{1}{t} + \frac{1}{e} < +\infty$$

と評価できる。次に $t \geq 1$ とする。各 $x > 0$ に対して $x/t \leq e^{x/t}$ であるから両辺を $t - 1$ 乗して比較することにより $x^{t-1} \leq t^{t-1} e^{(1-1/t)x}$ を得る。従って系 3.28(ii) より

$$\int_{(0,+\infty)} x^{t-1} e^{-x}\ \lambda(dx) \leq \int_{(0,+\infty)} t^{t-1} e^{(1-1/t)x} e^{-x}\ \lambda(dx) = t^t < +\infty$$

(ii) $(1-x)^{s-1}x^{t-1} \leq (1-x)^{s-1}/2 + x^{t-1}/2\ \forall x \in (0,1)$ であるから、

$$\int_{(0,1)} (1-x)^{s-1}x^{t-1}\ \lambda(dx) \leq \int_{(0,1)} \frac{(1-x)^{s-1}}{2} + \frac{x^{t-1}}{2}\ \lambda(dx) = \frac{1}{2s} + \frac{1}{2t} < +\infty$$

と評価できる。もちろん系 3.28(ii) を適用して等号を導いている。 □

3.32 演習問題. gamma 関数 Γ は微分可能であり、その導関数は

$$t \mapsto \int_{(0,+\infty)} x^{t-1}e^{-x}\log x\ \lambda(dx)$$

であることを示せ。

3.33 補題. $a, b \in \overline{\mathbb{R}},\ a < b$ かつ関数 $f : (a,b) \to \mathbb{R}$ は非負値かつ連続であるとする。また $v : I \to \mathbb{R}$ を開区間上の連続微分可能な関数で $a < v(x) < b\ \forall x \in I,\ v'(x) \geq 0\ \forall x \in I$, $\inf_{x\in I} v(x) = a$ かつ $\sup_{x\in I} v(x) = b$ を満たすものとする。このとき次が成り立つ。

$$\int_{(a,b)} f\,\lambda = \int_I f(v(x))v'(x)\ \lambda(dx).$$

証明. 関数 f の原始関数の一つを $F : (a,b) \to \mathbb{R}$ とすると系 3.28(ii) より

$$\int_{(a,b)} f\,\lambda = \sup_{y\in(a,b)} F(y) - \inf_{y\in(a,b)} F(y)$$

が成り立つ。さて関数 v は連続であり、$\inf_{x\in I} v(x) = a$ かつ $\sup_{x\in I} v(x) = b$ を満たす。従って中間値の定理より $(a,b) \subset \{v(x); x \in I\}$ である。逆向きの包含関係も仮定されているので

$$\sup_{y\in(a,b)} F(y) = \sup_{x\in I} F(v(x)),\quad \inf_{y\in(a,b)} F(y) = \inf_{x\in I} F(v(x))$$

となる。連鎖律により関数 $I \to \mathbb{R},\ x \mapsto F(v(x))$ は非負連続関数 $I \to \mathbb{R},\ x \mapsto f(v(x))v'(x)$ の原始関数である。再び系 3.28(ii) を適用して次の関係を得る。

$$\int_I f(v(x))v'(x)\ \lambda(dx) = \sup_{x\in I} F(v(x)) - \inf_{x\in I} F(v(x))$$

以上をまとめると結論にいたる。 □

3.34 注意. $v'(x) \leq 0\ \forall x \in I$, $\inf_{x\in I} v(x) = b$ かつ $\sup_{x\in I} v(x) = a$ であるなら次が成り立つ。

$$\int_{(a,b)} f\,\lambda = \int_I f(v(x))|v'(x)|\ \lambda(dx).$$

3.35 例. $t, c \in \mathbb{R}_{>0}$ とする。このとき各 $x > 0$ に対して

$$2\int_{(0,\sqrt{x}]} y^{2t-1}\exp\{-cy^2\}\lambda(dy) = \int_{(0,x]} y^{t-1}\exp\{-cy\}\lambda(dy).$$

3.36 例. 各 $s,t>0$ に対して次が成り立つ。

$$\begin{aligned}B(s,t) &= 2\int_{(0,\pi/2)}(\cos\theta)^{2s-1}(\sin\theta)^{2t-1}\lambda(d\theta)\\ &= \int_{(1,+\infty)}(x-1)^{s-1}x^{-t-s}\ \lambda(dx) = \int_{(0,+\infty)}x^{s-1}(x+1)^{-t-s}\ \lambda(dx)\\ &= 2\int_{(0,+\infty)}x^{2s-1}(x^2+1)^{-t-s}\ \lambda(dx) = \int_{\mathbb{R}}\frac{e^{sx}}{(e^x+1)^{t+s}}\ \lambda(dx)\end{aligned}$$

とくに $B(1/2,1/2)=\pi$ である。

3.37 演習問題. 例 3.35, 例 3.36 を示せ。

4　絶対連続な分布の例ならびに分布関数

　ここでは、第 3 節でのお膳立てのもと絶対連続な分布の例をいくつか紹介するとともに、それらの（累積）分布関数による特徴付けについて解説する。

記号

λ: 1 次元 Lebesgue 測度

4.1 例. $a,b\in\mathbb{R}$, $a<b$ とする。区間 (a,b) 上の**一様分布**(uniform distribution) とは関数

$$x\mapsto\frac{1}{b-a}1_{(a,b)}(x)$$

を密度とする絶対連続分布をいう。これを μ と書くと定理 3.19(ii) により $z\in\mathbb{R}$ に対し

$$\int_{\mathbb{R}}e^{zx}\mu(dx)=\int_{\mathbb{R}}e^{zx}\frac{1}{b-a}1_{(a,b)}(x)\lambda(dx)=\frac{1}{b-a}\int_{(a,b)}e^{zx}\lambda(dx)=\frac{e^{zb}-e^{za}}{z(b-a)}$$

ここで $e^{zx}>0$ であることに注意する。3 番目の等号は系 3.28(ii) による。また

$$\int_{\mathbb{R}}x\mu(dx)=\frac{a+b}{2},\int_{\mathbb{R}}x^2\mu(dx)=\frac{a^2+ab+b^2}{3}$$

4.2 例. $c\in\mathbb{R}_{>0}$ とする。パラメータ c の**指数分布**(exponential distribution) とは関数

$$x\mapsto c1_{(0,+\infty)}(x)e^{-cx}$$

を密度とする絶対連続分布をいう。これを μ と書くと $z<c$ に対し

$$\int_{\mathbb{R}}e^{zx}\mu(dx)=\int_{\mathbb{R}}e^{zx}c1_{(0,+\infty)}(x)e^{-cx}\lambda(dx)=c\int_{(0,+\infty)}e^{(z-c)x}\lambda(dx)=\frac{c}{c-z}.$$

可積分性が確かめられた後は上の計算は $\mathrm{Re}\,z<c$ なる $z\in\mathbb{C}$ に対しても有効である。また

$$\int_{\mathbb{R}}x\mu(dx)=\frac{1}{c},\int_{\mathbb{R}}x^2\mu(dx)=\frac{2}{c^2}.$$

4.3 例. $t, c \in \mathbb{R}_{>0}$ とする。パラメータ t, c の *gamma* **分布**とは関数

$$x \mapsto \frac{c^t}{\Gamma(t)} 1_{(0,+\infty)}(x) x^{t-1} e^{-cx} \text{ ただし } \Gamma \text{ は gamma 関数}$$

を密度とする絶対連続分布をいう。パラメータ t の役割を重要視するときは t を指数と呼ぶことにする。指数が 1 のときは指数分布である。また、$n \in \mathbb{N}$ に対してパラメータ $n/2$, $1/2$ の gamma 分布を自由度 n の χ^2 **分布**(chi-square distribution) という。

さてパラメータ t, c の gamma 分布 μ に対し $z < c$ のとき次が成り立つ。

$$\begin{aligned}\int_{\mathbb{R}} e^{zx} \mu(dx) &= \int_{\mathbb{R}} e^{zx} \frac{c^t}{\Gamma(t)} 1_{(0,+\infty)}(x) x^{t-1} e^{-cx} \lambda(dx) \\ &= \frac{c^t}{\Gamma(t)} \frac{1}{c-z} \int_{(0,+\infty)} \frac{x^{t-1}}{(c-z)^{t-1}} e^{-x} \lambda(dx) = \frac{c^t}{(c-z)^t}.\end{aligned}$$

２番目の等号は補題 3.21 による。z が虚数なら上の計算そのものはもはや無効である。しかし $\operatorname{Re} z < c$ なる $z \in \mathbb{C}$ に対しても可積分性は大丈夫であり、次の関係も成り立つ。

$$\int_{\mathbb{R}} e^{zx} \mu(dx) = \frac{c^t}{(c-z)^t} \quad z \in \mathbb{C}, \operatorname{Re} z < c$$

ただしこれを証明するには解析接続の概念が必要なので詳細は割愛する。また

$$\int_{\mathbb{R}} x^p \mu(dx) = \frac{\Gamma(t+p)}{c^p \Gamma(t)} \quad p > 0.$$

4.4 例. $a \in \mathbb{R}$, $c \in \mathbb{R}_{>0}$ とする。**両側指数分布**(two sided exponential distribution) とは関数

$$x \mapsto \frac{c}{2} e^{-c|x-a|}$$

を密度とする絶対連続分布をいう。これを μ と書くと

$$\int_{\mathbb{R}} x \mu(dx) = a, \int_{\mathbb{R}} x^2 \mu(dx) = \frac{2}{c^2} + a^2, \int_{\mathbb{R}} e^{zx} \mu(dx) = \frac{c^2 e^{za}}{c^2 - z^2} \quad z \in \mathbb{C}, |\operatorname{Re} z| < c.$$

4.5 演習問題. 例 4.2, 例 4.3 および例 4.4 の計算をフォローせよ。

4.6 例. $a \in \mathbb{R}$, $t \in \mathbb{R}_{>0}$ とする。中心 a 半値幅 t の *Cauchy* **分布**とは関数

$$x \mapsto \frac{t}{\pi} \frac{1}{t^2 + (x-a)^2}$$

を密度とする絶対連続分布をいう。これを μ と書くと定理 3.19(ii), 補題 3.21 と例 8.20 により

$$\int_{\mathbb{R}} e^{\sqrt{-1}zx} \mu(dx) = \frac{t}{\pi} \int_{\mathbb{R}} \frac{e^{\sqrt{-1}zx}}{t^2 + (x-a)^2} \lambda(dx) = \frac{t}{\pi} \int_{\mathbb{R}} \frac{e^{\sqrt{-1}z(x+a)}}{t^2 + x^2} \lambda(dx) = e^{\sqrt{-1}za - t|z|}$$

が $z \in \mathbb{R}$ に対して成り立つ。Cauchy 分布については $\int_{\mathbb{R}} |x| \mu(dx) = +\infty$ である。

4.7 補題. ここでは $t \in \mathbb{R}_{>0}$, $x \in \mathbb{R}$ に対し $p(t,x) := t/\{\pi(t^2+x^2)\}$ とおく。
(i) $\displaystyle\int_{\mathbb{R}} p(t,x-y)p(s,y)\lambda(dy) = p(t+s,x)$ $\forall t, \forall s \in \mathbb{R}_{>0}$, $\forall x \in \mathbb{R}$ が成り立つ。
(ii) 任意の有界 Borel($\mathbb{R}$) 可測関数 $f : \mathbb{R} \to \mathbb{C}$ と $a \in \mathbb{R}$ に対して f が a において連続なら $\displaystyle\int_{\mathbb{R}} p(t,x-a)f(x)\lambda(dx)$ は $t \to 0$ の極限で $f(a)$ に収束する。

証明. (i) $D = \{(t+s)^2+x^2\}\{(t-s)^2+x^2\}$ とおく。少々の計算により

$$\begin{aligned}\frac{ts}{(t^2+(y-x)^2)(s^2+y^2)} &= \frac{2ts}{D}\Big(\frac{xy+(t-s)s}{s^2+y^2} - \frac{x(y-x)+(t-s)t}{t^2+(y-x)^2}\Big)\\ &\quad + \frac{s}{(t+s)^2+x^2}\frac{t}{t^2+(y-x)^2} + \frac{t}{(t+s)^2+x^2}\frac{s}{s^2+y^2}\end{aligned}$$

さて連続関数 $y \mapsto 2xy/(s^2+y^2) - 2x(y-x)/\{t^2+(y-x)^2\}$ は λ 可積分であり、その原始関数は $y \mapsto x\log((s^2+y^2)/\{t^2+(y-x)^2\})$ で与えられる。従って補題 3.26(ii) を適用して

$$\frac{2ts}{D}\int_{\mathbb{R}}\Big(\frac{xy}{s^2+y^2} - \frac{x(y-x)}{t^2+(y-x)^2}\Big)\,\lambda(dy) = 0$$

を得る。あとは $\int_{\mathbb{R}} p(t,\cdot-x)\lambda = 1$, $\int_{\mathbb{R}} p(s,\cdot)\lambda = 1$ を使って変形すると

$$\frac{1}{\pi}\int_{\mathbb{R}}\frac{ts}{(t^2+(y-x)^2)(s^2+y^2)}\ \lambda(dy) = \frac{2ts}{D}\{(t-s)-(t-s)\} + \frac{s+t}{(t+s)^2+x^2}$$

言うまでもないが、以上の計算が有効なのは $D \neq 0$ 即ち $t \neq s$ あるいは $x \neq 0$ の時に限る。しかしながら Lebesgue の優収束定理が適用できるので

$$\int_{\mathbb{R}}\frac{t^2}{(t^2+y^2)^2}\ \lambda(dy) = \lim_{x\to 0}\int_{\mathbb{R}}\frac{t^2}{(t^2+(y-x)^2)(t^2+y^2)}\ \lambda(dy) = \lim_{x\to 0}\frac{\pi 2t}{(2t)^2+x^2} = \frac{\pi}{2t}$$

となり、$t = s$ かつ $x = 0$ の場合も示すべき関係が成り立つ。

(ii) 変数変換公式（補題 3.21）を適用して次を得る。

$$\frac{t}{\pi}\int_{\mathbb{R}}\frac{f(x)}{t^2+(x-a)^2}\,\lambda(dx) = \frac{1}{\pi}\int_{\mathbb{R}}\frac{f(tx+a)}{1+x^2}\,\lambda(dx)$$

右辺の被積分関数は $t \to 0$ の極限において関数 $x \mapsto f(a)/(1+x^2)$ に各点収束する。そこで $M := \sup_{x\in\mathbb{R}}|f(x)|$ とおく。λ 可積分関数 $x \mapsto M/(1+x^2)$ を優関数に選んで Lebesgue の収束定理を適用して結論に到達する。 □

4.8 例. $t \in \mathbb{R}_{>0}$, $a \in \mathbb{R}$ とする。平均 a 分散 t の**正規分布**(normal distribution) とは関数

$$x \mapsto \frac{1}{\sqrt{2\pi t}}\exp\{-\frac{(x-a)^2}{2t}\}$$

を密度とする絶対連続分布をいう。*Gauss* **分布**という別名もよく使われる。特に平均 0 分散 1 の正規分布を**標準正規分布**(standard normal distribution) という。基本となるのは

$$\int_{\mathbb{R}}\exp\{-x^2/2\}\lambda(dx) = \sqrt{2\pi}$$

である。その標準的な算出手段を例 8.19 と例 20.20 で紹介する。だが重要なのはむしろ

$$0 < \int_{\mathbb{R}} \exp\{-x^2/2\}\lambda(dx) < +\infty$$

を確認することであり、それぞれ演習問題 3.17 と例 3.30 で実行済みである。

4.9 補題. μ を平均 $a \in \mathbb{R}$ 分散 $t \in \mathbb{R}_{>0}$ の正規分布とする。
(i) $\int_{\mathbb{R}} |x|^p \mu(dx) < +\infty\ \forall p \in \mathbb{N},\ \int_{\mathbb{R}} x\mu(dx) = a,\ \int_{\mathbb{R}} x^2\mu(dx) = a^2 + t.$
(ii) 各 $z \in \mathbb{C}$ に対して $\int_{\mathbb{R}} |e^{zx}|\mu(dx) < +\infty$ かつ $\int_{\mathbb{R}} e^{zx}\mu(dx) = \exp\{za + tz^2/2\}$ である。

証明. (ii) を検討する。まず $z \in \mathbb{R}$ とする。このとき $e^{zx} \in \mathbb{R}_{\geq 0}\ \forall x \in \mathbb{R}$ である。さて

$$e^{zx} \exp\{-\frac{(x-a)^2}{2t}\} = \exp\{za + tz^2/2\} \exp\{-\frac{(x-a-tz)^2}{2t}\}$$

であるが、右辺の第 2 因子は平均 $a + tz$ 分散 t の正規分布の密度関数を構成する。よって定理 3.19(ii) により次が導かれ、よって $z \in \mathbb{R}$ の場合に (ii) が示された。

$$\int_{\mathbb{R}} e^{zx}\mu(dx) = \int_{\mathbb{R}} e^{zx} \frac{1}{\sqrt{2\pi t}} \exp\{-\frac{(x-a)^2}{2t}\}\lambda(dx) = \exp\{za + tz^2/2\}$$

以下 $a = 0$ として議論を続ける。$|e^{zx}| \leq e^{|zx|} \leq e^{|z|x} + e^{-|z|x}\ \forall x \in \mathbb{R}$ であるから

$$\int_{\mathbb{R}} |e^{zx}|\,\mu(dx) \leq \int_{\mathbb{R}} e^{|zx|}\,\mu(dx) \leq 2\exp\{t|z|^2/2\} < +\infty.$$

が従う。中辺の被積分関数を整級数展開する。各項は非負値であり項別積分が許される。

$$\sum_{n=0}^{\infty} \int_{\mathbb{R}} \frac{|zx|^n}{n!}\mu(dx) = \int_{\mathbb{R}} e^{|zx|}\mu(dx) < +\infty.$$

ゆえに級数 $\sum_{n=0}^{\infty} \int_{\mathbb{R}} x^n/n!\,\mu(dx)z^n$ はすべての $z \in \mathbb{C}$ に対して絶対収束し、

$$(\star) \qquad \sum_{n=0}^{\infty} \int_{\mathbb{R}} \frac{x^n}{n!}\mu(dx)z^n = \int_{\mathbb{R}} e^{zx}\mu(dx)\ \forall z \in \mathbb{C}$$

が成り立つ。$z \in \mathbb{R}$ のとき既に確かめたように右辺は $\exp\{tz^2/2\}$ に等しく、またそれは $\sum_{k=0}^{\infty} t^k z^{2k}/(2^k k!)$ と整級数展開される。 よって

$$\sum_{n=0}^{\infty} \int_{\mathbb{R}} \frac{x^n}{n!}\mu(dx)z^n = \sum_{k=0}^{\infty} \frac{t^k}{2^k k!} z^{2k}\ \forall z \in \mathbb{R}$$

整級数の一意性により上の関係はすべての $z \in \mathbb{C}$ に対して成立する。再び $(\star)$ に戻って

$$\int_{\mathbb{R}} e^{zx}\mu(dx) = \exp\{tz^2/2\}\ \forall z \in \mathbb{C}$$

に到達する。一般の a については補題 3.21 を使って $a = 0$ の場合に帰着できる。 □

4.10 演習問題. ここでは $t \in \mathbb{R}_{>0}$, $x \in \mathbb{R}$ に対し $p(t,x) := \exp\{-x^2/2t\}/\sqrt{2\pi t}$ とおく。
(i) $\dfrac{\partial}{\partial t}p(t,x) = \dfrac{1}{2}\dfrac{\partial^2}{\partial x^2}p(t,x)\ \forall t \in \mathbb{R}_{>0}, \forall x \in \mathbb{R}$ を示せ。
(ii) $\displaystyle\int_{\mathbb{R}} p(t,x-y)p(s,y)\lambda(dy) = p(t+s,x)\ \forall t, \forall s \in \mathbb{R}_{>0}, \forall x \in \mathbb{R}$ を示せ。
(iii) 任意の有界 $\mathrm{Borel}(\mathbb{R})$ 可測関数 $f : \mathbb{R} \to \mathbb{C}$ と $a \in \mathbb{R}$ に対して f が a において連続なら $\displaystyle\int_{\mathbb{R}} p(t,x-a)f(x)\lambda(dx)$ は $t \to 0$ の極限で $f(a)$ に収束することを示せ。

4.11 定義. (i) $(\mathbb{R}, \mathrm{Borel}(\mathbb{R}))$ 上の確率測度 μ に対して関数

$$\mathbb{R} \to \mathbb{R}, x \mapsto \mu((-\infty, x])$$

を μ の**分布関数**(distribution function) という。
(ii) 確率変数 X に対して $\mathbb{R} \to \mathbb{R}$, $x \mapsto \mathcal{L}(X, (-\infty, x])$ を X の分布関数という。

4.12 例. (i) パラメータ c の指数分布の分布関数は $x \mapsto \max\{1 - e^{-cx}, 0\}$ である。
(ii) 中心 a 半値幅 t の Cauchy 分布の分布関数は $x \mapsto \arctan\{(x-a)/t\}/\pi + 1/2$ である。
(iii) $a \in \mathbb{R}$ に質量を持つ Dirac 測度の分布関数は $1_{[a,+\infty)}$ である。

4.13 補題. $(\mathbb{R}, \mathrm{Borel}(\mathbb{R}))$ 上の確率測度 μ の分布関数は次の性質を持つ。

(i) $x \mapsto \mu((-\infty, x])$ は非減少である。
(ii) $x \mapsto \mu((-\infty, x])$ は右連続である。
(iii) $\lim_{x \to -\infty} \mu((-\infty, x]) = 0$, $\lim_{x \to +\infty} \mu((-\infty, x]) = 1$.

4.14 演習問題. 補題 4.13 を示せ。また $\lim_{x \uparrow a} \mu((-\infty, x]) = \mu((-\infty, a))$ であることを示せ。

復習

任意の右連続な非減少関数 $F : \mathbb{R} \to \mathbb{R}$ に対して次の条件を満たす $(\mathbb{R}, \mathrm{Borel}(\mathbb{R}))$ 上の測度がただ一つ存在する。

$$\mu((a,b]) = F(b) - F(a)\ \forall a, \forall b \in \mathbb{R},\ a < b$$

これを関数 F が誘導する *Lebesgue-Stieltjes* **測度**という。とくに $\sup_{x\in\mathbb{R}} F(x) = \inf_{x\in\mathbb{R}} F(x) + 1$ ならば、確率測度が対応する。

一意性に関しては後で再論する（定理 7.12)。一意性を適用すると直ちに次のことがわかる。

4.15 補題. $(\mathbb{R}, \mathrm{Borel}(\mathbb{R}))$ 上の確率測度 μ, ν に対して $\mu = \nu$ となるための必要十分条件はそれらの分布関数が一致することである。

4.16 定義. 確率変数 X, Y に対して $\mathcal{L}(X, \cdot) = \mathcal{L}(Y, \cdot)$ であるとき**同分布に従う**という。

補題 4.15 は次のような言い換えを持つ。

4.17 系. 確率変数 X, Y に対してそれらが同分布に従う必要十分条件はそれらの分布関数が一致することである。

4.18 例. 確率変数 X が標準正規分布に従う必要十分条件は

$$P(X \leq x) = \int_{(-\infty,x]} \frac{1}{\sqrt{2\pi}} \exp\{-y^2/2\}\lambda(dy)\ \forall x \in \mathbb{R}$$

である。右辺の関数を**誤差関数**(error function) という。

4.19 例. $\alpha, \beta \in \mathbb{R}$ かつ $\alpha \neq 0$ とする。確率変数 X が平均 $a \in \mathbb{R}$ 分散 $t \in \mathbb{R}_{>0}$ の正規分布に従うなら $\alpha X + \beta$ は平均 $\alpha a + \beta$ 分散 $\alpha^2 t$ の正規分布に従う。

証明. $\alpha < 0$ とする。確率変数 $\alpha X + \beta$ の分布関数を調べる。

$$\begin{aligned} P(\alpha X + \beta \leq x) = P(X \geq (x-\beta)/\alpha) &= \int_{[(x-\beta)/\alpha,+\infty)} \frac{1}{\sqrt{2\pi t}} \exp\{-\frac{(y-a)^2}{2t}\}\lambda(dy) \\ &= \frac{1}{|\alpha|}\int_{(-\infty,x]} \frac{1}{\sqrt{2\pi t}} \exp\{-\frac{((y-\beta)/\alpha - a)^2}{2t}\}\lambda(dy) \end{aligned}$$

ここでは変数変換公式（補題 3.21）を適用して 3 番目の等号を導いている。右辺を整理すれば分かるようにそれはまさに平均 $\alpha a + \beta$ 分散 $\alpha^2 t$ の正規分布の分布関数である。 □

4.20 注意. X を標準正規分布に従う確率変数とする。このとき例 4.19 によれば確率変数 $-X$ も標準正規分布に従う。しかしながら $P(X = -X) = P(X = 0) = 0$ である。すなわち分布が同じということと確率変数として同じということは異なる概念である。

4.21 演習問題. ここでは誤差関数を Φ と表記する。平均 $a \in \mathbb{R}$ 分散 $t \in \mathbb{R}_{>0}$ の正規分布についてその分布関数は $x \mapsto \Phi((x-a)/\sqrt{t})$ で与えられることを示せ。

4.22 例. 確率変数 X が標準正規分布に従うなら X^2 は自由度 1 の χ^2 分布に従う。

証明. 確率変数 X^2 の分布関数を調べる。まず $x < 0$ なら $\{X^2 \leq x\} = \emptyset$ であるからその確率は 0 となる。$x \geq 0$ なら $\{X^2 \leq x\} = \{-\sqrt{x} \leq X \leq \sqrt{x}\}$ であるから

$$P(X^2 \leq x) = \int_{[-\sqrt{x},\sqrt{x}]} \frac{1}{\sqrt{2\pi}} \exp\{-y^2/2\}\lambda(dy) = 2\int_{(0,\sqrt{x}]} \frac{1}{\sqrt{2\pi}} \exp\{-y^2/2\}\lambda(dy)$$

ここでは $[-\sqrt{x}, \sqrt{x}]$ を $[-\sqrt{x}, 0)$, $\{0\}$ と $(0, \sqrt{x}]$ に分割し、変数変換公式（補題 3.21）を $[-\sqrt{x}, 0)$ 上の積分に適用し 2 番目の等号を導いている。例 3.35 によれば

$$2\int_{(0,\sqrt{x}]} \frac{1}{\sqrt{2\pi}} \exp\{-y^2/2\}\lambda(dy) = \int_{(0,x]} \frac{1}{\sqrt{2\pi}} y^{-1/2} \exp\{-y/2\}\lambda(dy)\ x > 0$$

また $x = 0$ なら両辺ともに 0 なので自明に等号が成り立つ。以上より

$$P(X^2 \leq x) = \int_{(-\infty,x]} \frac{1}{\sqrt{2\pi}} 1_{(0,+\infty)}(y) y^{-1/2} \exp\{-y/2\}\lambda(dy)\ x \in \mathbb{R}$$

右辺はまさにパラメータ 1/2, 1/2 を持つ gamma 分布の分布関数である。 □

4.23 演習問題. モデル $((0,1], \mathrm{Borel}((0,1]), \lambda)$ で考える。
(i) $a \in \mathbb{R}$ とする。次の確率変数 X の分布は a に質量を持つ Dirac 測度であることを示せ。

$$X(\omega) := a \ \omega \in (0,1]$$

(ii) $c \in \mathbb{R}_{>0}$ とする。次の確率変数 X の分布はパラメータ c の指数分布であることを示せ。

$$X(\omega) := -\frac{\log \omega}{c} \ \omega \in (0,1]$$

演習問題 4.23(ii) のからくりは次の補題に述べるものに拡張される。これを確率変数の *Skorokhod* **表現**という。

4.24 補題. 右連続な非減少関数 $F : \mathbb{R} \to \mathbb{R}$ であって条件 $\inf_{x\in\mathbb{R}} F(x) = 0, \ \sup_{x\in\mathbb{R}} F(x) = 1$ を満たすものが与えられたとし、次の二つの関数 X, Y を導入する。

$$X : (0,1] \to \overline{\mathbb{R}}, \omega \mapsto \inf\{x \in \mathbb{R} : F(x) \geq \omega\}; \ Y : (0,1] \to \overline{\mathbb{R}}, \omega \mapsto \inf\{x \in \mathbb{R} : F(x) > \omega\}$$

(i) 各 $\omega \in (0,1)$ に対して $-\infty < X(\omega) \leq Y(\omega) < +\infty$ かつ $\omega \leq F(X(\omega))$ が成り立つ。
(ii) 各 $x \in \mathbb{R}$ に対し $(0, F(x)) \subset \{\omega \in (0,1) : Y(\omega) \leq x\} \subset \{\omega : X(\omega) \leq x\} = (0, F(x)]$
(iii) 各 $x \in \mathbb{R}$ に対し $\lambda(\{\omega \in (0,1) : X(\omega) \leq x\}) = F(x) = \lambda(\{\omega \in (0,1) : Y(\omega) \leq x\})$ であり、また $\lambda(\{\omega \in (0,1) : X(\omega) \neq Y(\omega)\}) = 0$ が成り立つ。

証明. $0 < \omega < 1$ とする。$\inf_{x\in\mathbb{R}} F(x) = 0$ かつ $\sup_{x\in\mathbb{R}} F(x) = 1$ により、次が成り立つ。

$$\emptyset \neq F^{-1}((-\infty, \omega)), \emptyset \neq F^{-1}((\omega, +\infty)) \subset F^{-1}([\omega, +\infty)).$$

F は非減少なので $y \in F^{-1}((-\infty, \omega))$ かつ $x \leq y$ なら $F(x) \leq F(y) < \omega$ が成り立つ。対偶を取ると $x \in F^{-1}([\omega, +\infty))$ かつ $y \in F^{-1}((-\infty, \omega))$ なら $y < x$ であることが分かる。よって

$$-\infty < \sup F^{-1}((-\infty, \omega)) \leq \inf F^{-1}([\omega, +\infty)) \leq \inf F^{-1}((\omega, +\infty)) < +\infty$$

となり、$-\infty < X(\omega) \leq Y(\omega) < +\infty$ を得る。他方 F の右連続性により

$$\omega \leq F(\inf F^{-1}([\omega, +\infty))) = F(X(\omega))$$

が成り立つ。ここで $x \geq X(\omega)$ なら右辺は $F(x)$ 以下であるから

$$[X(\omega), +\infty) \subset F^{-1}([\omega, +\infty))$$

を得るが、$X(\omega) = \inf F^{-1}([\omega, +\infty))$ なので実は等号が成り立つ。明らかに $x \in F^{-1}((\omega, +\infty))$ なら $x \geq \inf F^{-1}((\omega, +\infty)) = Y(\omega)$ である。従って

$$F^{-1}((\omega, +\infty)) \subset [Y(\omega), +\infty) \subset [X(\omega), +\infty) = F^{-1}([\omega, +\infty))$$

が成り立つ。$x \in \mathbb{R}$ を固定して上の関係を ω に関する条件で書き直すと (ii) が得られる。(iii) の前半は (ii) から直ちに従う。さて $X \leq Y$ であるから

$$\begin{aligned} &\{\omega \in (0,1) : X(\omega) \neq Y(\omega)\} = \{\omega \in (0,1) : X(\omega) < Y(\omega)\} \\ &= \bigcup_{r\in\mathbb{Q}} (\{\omega \in (0,1) : X(\omega) \leq r\} \setminus \{\omega \in (0,1) : Y(\omega) \leq r\}) \end{aligned}$$

各 $r \in \mathbb{Q}$ に対して $\lambda(\{\omega \in (0,1) : X(\omega) \leq r\}) = F(r) = \lambda(\{\omega \in (0,1) : Y(\omega) \leq r\})$ であるから、測度の可算劣加法性により $\lambda(\{\omega \in (0,1) : X(\omega) \neq Y(\omega)\}) = 0$ を得る。 $\square$

確率分布とは直接の関係はないが分布関数に類似の概念を導入しておこう。これは部分積分においてその役割を持つものである。

4.25 定義. $f:\mathbb{R}\to\mathbb{C}$ を Borel($\mathbb{R}$) 可測関数とする。各有界区間 (a,b) 上で λ 可積分となるとき f は**局所可積分**(locally integrable) であるという。

連続関数 $\mathbb{R}\to\mathbb{C}$ は局所可積分であることを注意しておこう。

4.26 演習問題. $f:\mathbb{R}\to\mathbb{C}$ を Borel($\mathbb{R}$) 可測関数で局所可積分なものとする。
(i) 関数 $F:\mathbb{R}\to\mathbb{C}$ であって次を満たすものが存在することを示せ。

$$F(b)-F(a)=\int_{(a,b]} f\,\lambda\ \forall a\in\mathbb{R},\forall b\in\mathbb{R}\ \text{但し}\ a<b.$$

このような関数 F を f の**不定積分**(indefinite integral) という。
(ii) f の不定積分 F は連続であることを示せ。

5 確率変数と多次元確率変数

実確率変数系 $X_1, X_2, \ldots, X_n$ が与えられたとき、それらを束ねて

$$\Omega\to\mathbb{R}^n, \omega\mapsto(X_1(\omega),X_2(\omega),\ldots,X_n(\omega))$$

という写像が構成できる。これを議論の対象として取り込もう。たとえば結合分布という概念を導入したいのだが、それには定理 2.17 に相当する命題が必要である。最終的な結論となる系 5.15 に至る途上で空間 $\mathbb{R}^n$ の直積構造と可分性の色濃い反映が判明するであろう。

5.1 定義. $\mathbb{R}^d$ の開集合すべてで生成される $\mathbb{R}^d$ 上の σ-加法族を d **次元** *Borel* **集合族**といい記号 Borel($\mathbb{R}^d$) で表す。各 $B\in$ Borel($\mathbb{R}^d$) を d **次元** *Borel* **集合**という。

ここでは別の題材との接続上、次元を表すのに n の代わりに d を使った。ところで $d=1$ の場合は、補題 2.10 により、以前の定義と同等なものになっていることを注意しよう。

5.2 演習問題. 連続関数 $f:\mathbb{R}^d\to\mathbb{R}$ は Borel($\mathbb{R}^d$) 可測であることを示せ。

約束

$\mathbb{R}^d$ の部分集合であって開区間の直積で表現できるものを d **次元開区間**と呼ぶ。

5.3 演習問題. $\mathcal{C}$ を両端が有理数であるような開区間の直積で表現できる d 次元開区間の全体の族とする。任意に与えられた $\mathbb{R}^d$ の開集合 A に対して次が成り立つことを示せ。

$$A=\bigcup_{I\in\mathcal{C}:I\subset A} I$$

5.4 補題. d 次元開区間すべてで生成される $\mathbb{R}^d$ 上の σ-加法族は Borel($\mathbb{R}^d$) である。

証明. d 次元開区間すべてで生成される σ-加法族を $\mathcal{B}$ とかく。d 次元開区間は開集合であるから、$\mathcal{B} \subset \mathrm{Borel}(\mathbb{R}^d)$ が導かれる。他方 $\mathcal{C}$ を演習問題 5.3 と同じ集合族とする。$\mathcal{C}$ は可算族であるから、演習問題 5.3 により、任意に与えられた開集合が開区間の可算合併でかけることになる。すると $\mathcal{B}$ は開部分集合すべてを含む σ 加法族となるが、そのようなものの最小が $\mathrm{Borel}(\mathbb{R}^d)$ なので、$\mathrm{Borel}(\mathbb{R}^d) \subset \mathcal{B}$ を得た。 $\square$

5.5 定義. $(S, \mathcal{B})$ を非自明な可測空間とする。写像 $f : S \to \mathbb{R}^d$ が $\mathcal{B}$ **可測**であるとは条件 $f^{-1}(B) \in \mathcal{B}\ \forall B \in \mathrm{Borel}(\mathbb{R}^d)$ を満たすことをいう。

記号

写像 $f : S \to \mathbb{R}^d$ に対し $\sigma\{f\} := \{f^{-1}(B)\,;B \in \mathrm{Borel}(\mathbb{R}^d)\}$ と書く。

この記号を用いると写像 $f : S \to \mathbb{R}^d$ に対しその $\mathcal{B}$ 可測性は条件 $\sigma\{f\} \subset \mathcal{B}$ と同値である。ここで、集合族 $\sigma\{f\}$ が S 上の σ 加法族であることに注目する。その証明は 1 次元のバージョンの補題 2.14 と同じく補題 2.12 の適用が基本線である。省略せず証明を与えておこう。

5.6 補題. f を写像 $S \to T$ とする
(i) T 上の σ 加法族 $\mathcal{M}$ に対し $\{f^{-1}(B)\,;B \in \mathcal{M}\}$ は S 上の σ 加法族である。
(ii) S 上の σ 加法族 $\mathcal{B}$ に対し $\{B \subset T : f^{-1}(B) \in \mathcal{B}\}$ は T 上の σ 加法族である。
(iii) T 上の集合族 $\mathcal{C}$ に対して $\{f^{-1}(B)\,;B \in \sigma_T(\mathcal{C})\} = \sigma_S(\{f^{-1}(B)\,;B \in \mathcal{C}\})$.
(iv) S 上の σ 加法族 $\mathcal{B}$ と T 上の集合族 $\mathcal{C}$ に対して次が成り立つ。

$$f^{-1}(B) \in \mathcal{B}\ \forall B \in \sigma_T(\mathcal{C}) \Leftrightarrow f^{-1}(C) \in \mathcal{B}\ \forall C \in \mathcal{C}.$$

証明. (i) $f^*\mathcal{M} := \{f^{-1}(B)\,;B \in \mathcal{M}\}$ とおく。まず $\emptyset \in \mathcal{M}$ より $\emptyset = f^{-1}(\emptyset) \in f^*\mathcal{M}$ である。次に $A \in f^*\mathcal{M}$ とすると $\exists B \in \mathcal{M}$ s.t. $A = f^{-1}(B)$ であるが、$T \setminus B \in \mathcal{M}$ により、次が従う。

$$S \setminus A = S \setminus f^{-1}(B) = f^{-1}(T \setminus B) \in f^*\mathcal{M}$$

他方 $A_n \in f^*\mathcal{M}\ \forall n \in \mathbb{N}$ とすると $\exists B_n \in \mathcal{M}$ s.t. $A_n = f^{-1}(B_n)$ である。そこで $\bigcup_{n=1}^{\infty} A_n \in \mathcal{M}$ を使って以下を得る。

$$\bigcup_{n=1}^{\infty} A_n = \bigcup_{n=1}^{\infty} f^{-1}(B_n) = f^{-1}(\bigcup_{n=1}^{\infty} B_n) \in f^*\mathcal{M}$$

(ii) $f_*\mathcal{B} := \{B \subset T : f^{-1}(B) \in \mathcal{B}\}$ とおく。まず $f^{-1}(\emptyset) = \emptyset \in \mathcal{B}$ により $\emptyset \in f_*\mathcal{B}$ である。$B \in f_*\mathcal{B}$ とする。このとき $B \subset T$ かつ $f^{-1}(B) \in \mathcal{B}$ であるから、次より $T \setminus B \in f_*\mathcal{B}$ が従う。

$$f^{-1}(T \setminus B) = S \setminus f^{-1}(B) \in \mathcal{B}$$

$B_n \in f_*\mathcal{B}\ \forall n \in \mathbb{N}$ とする。このとき $B_n \subset T$ かつ $f^{-1}(B_n) \in \mathcal{B}$ であるから、

$$f^{-1}(\bigcup_{n=1}^{\infty} B_n) = \bigcup_{n=1}^{\infty} f^{-1}(B_n) \in \mathcal{B}$$

より $\bigcup_{n=1}^{\infty} B_n \in f_*\mathcal{B}$ となる。

(iii) $f^*\mathcal{C} := \{f^{-1}(B)\,;B \in \mathcal{C}\}$ と書く。σ 加法族 $\sigma_S(f^*\mathcal{C})$ に対して (ii) を適用すると

$$\{B \subset T : f^{-1}(B) \in \sigma_S(f^*\mathcal{C})\}$$

は T 上の σ 加法族であることがわかる。しかもその定義により $\mathcal{C}$ を内包する。そのようなものの最小が $\sigma_T(\mathcal{C})$ であるから次を得る。

$$\sigma_T(\mathcal{C}) \subset \{B \subset T : f^{-1}(B) \in \sigma_S(f^*\mathcal{C})\}$$

これは $f^{-1}(B) \in \sigma_S(f^*\mathcal{C})\ \forall B \in \sigma_T(\mathcal{C})$ を意味する。さらに $f^*(\sigma_T(\mathcal{C}))$ の定義により

$$f^*(\sigma_T(\mathcal{C})) \subset \sigma_S(f^*\mathcal{C}).$$

というように書き換えられる。逆向きの包含関係は $f^*(\sigma_T(\mathcal{C}))$ が S 上の σ 加法族であることと $f^*\mathcal{C} \subset f^*(\sigma_T(\mathcal{C}))$ であることから導かれる。ゆえに $f^*(\sigma_T(\mathcal{C})) = \sigma_S(f^*\mathcal{C})$.

(iv) 示そうとしている同値性は $f^*(\sigma_T(\mathcal{C})) \subset \mathcal{B} \Leftrightarrow f^*\mathcal{C} \subset \mathcal{B}$ と表現できる。さて $\mathcal{B}$ は σ 加法族なので右の条件は $\sigma_S(f^*\mathcal{C}) \subset \mathcal{B}$ と書き換えられる。ところが (iii) によれば、まさに $\sigma_S(f^*\mathcal{C}) = f^*(\sigma_T(\mathcal{C}))$ であるから同値性が導かれた。 □

記号

写像 $S \to T$ と T 上の集合族 $\mathcal{C}$ に対し $f^*\mathcal{C} := \{f^{-1}(B)\,;B \in \mathcal{C}\}$ と書く。

5.7 注意. 写像 $f : S \to \mathbb{R}^d$ に対しては $\sigma\{f\} = f^*\mathrm{Borel}(\mathbb{R}^d)$ という関係である。

5.8 例. (i) $\{(a,b]\,;0 \le a < b \le 1\}$ で生成される $(0,1]$ 上の σ 加法族は $\mathrm{Borel}((0,1])$ に等しい。
(ii) 第 1 節の写像 $\varphi : (0,1] \to (0,1]$ は $\varphi^{-1}(B) \in \mathrm{Borel}((0,1])\ \forall B \in \mathrm{Borel}((0,1])$ を満たす。

証明. (i) $\iota : (0,1] \to \mathbb{R}$ を包含写像とする。左半開区間全体の族を $\mathcal{I}$ と書くと、注意 2.21 で述べたことと補題 5.6(iii) により、次がわかる。

$$\mathrm{Borel}((0,1]) = \sigma\{\iota\} = \iota^*\mathrm{Borel}(\mathbb{R}) = \iota^*\sigma_{\mathbb{R}}(\mathcal{I}) = \sigma_{(0,1]}(\iota^*\mathcal{I})$$

さて $\iota^{-1}(A) = A \cap (0,1]\ \forall A \subset \mathbb{R}$ である（これが実は注意 2.21 のポイント）から

$$\iota^*\mathcal{I} = \{(a,b]\,;0 \le a < b \le 1\}$$

となり結論に到達した。

(ii) (i) と補題 5.6(iv) により、$\varphi^{-1}((a,b]) \in \mathrm{Borel}((0,1])\ 0 \le \forall a < \forall b \le 1$ を示せばよいが、実際 $\varphi^{-1}((a,b]) = (pa, pb] \cup ((1-p)a + p, (1-p)b + p]$ であるから成り立つ。 □

記号

写像系 $f_i : S \to \mathbb{R}^{d(i)}\ i = 1, 2, \ldots, n$ （値域はそれぞれ違うものでもよい）に対して

$$\sigma\{f_1, f_2, \ldots, f_n\} := \sigma_S(\sigma\{f_1\} \cup \sigma\{f_2\} \cup \cdots \cup \sigma\{f_n\})$$

と表記し、これを系 $f_i\ i = 1, 2, \ldots, n$ によって**生成される σ 加法族**という。

5.9 定理. 写像 $f:S\to\mathbb{R}^d$ に対しその成分を $g_1,\,g_2,\,\ldots,\,g_d$ とする。このとき次が成り立つ。

$$\sigma\{f\}=\sigma\{g_1,g_2,\ldots,g_d\}.$$

証明. 開区間の全体の族を $\mathcal{I}$ とし、d 次元開区間の全体の族を $\mathcal{I}^d$ とかく。まず補題 5.4 によれば、$\mathrm{Borel}(\mathbb{R}^d)=\sigma(\mathcal{I}^d)$ である。そこで補題 5.6(iii) を適用して

$$\sigma\{f\}=f^*\mathrm{Borel}(\mathbb{R}^d)=f^*(\sigma(\mathcal{I}^d))=\sigma_S(f^*\mathcal{I}^d).$$

次に $\sigma\{g_1\}=\sigma_S(g_1^*\mathcal{I})$ も同様に成立することに注目する。

$$g_1^*\mathcal{I}=\{g_1^{-1}(I)\,;I\in\mathcal{I}\}=\{f^{-1}(I\times\mathbb{R}^{d-1})\,;I\in\mathcal{I}\}\subset f^*\mathrm{Borel}(\mathbb{R}^d)=\sigma\{f\}$$

より $\sigma\{f\}$ は $g_1^*\mathcal{I}$ を内包する S 上の σ 加法族である。$\sigma\{g_1\}=\sigma_S(g_1^*\mathcal{I})$ がそのようなものの最小であることから関係 $\sigma\{g_1\}\subset\sigma\{f\}$ が導かれる。ほかについても同様であるから

$$\sigma\{g_1\}\cup\sigma\{g_2\}\cup\cdots\cup\sigma\{g_n\}\subset\sigma\{f\}.$$

再び、最小の σ 加法族であることを使って次の関係を得る。

$$\sigma_S(\sigma\{g_1\}\cup\sigma\{g_2\}\cup\cdots\cup\sigma\{g_n\})\subset\sigma\{f\}.$$

以下、煩雑さをさけるため $d=2$ として議論を進める。さて

$$\begin{aligned}\{f^{-1}(I\times J)\,;I,J\in\mathcal{I}\}&=\{g_1^{-1}(I)\cap g_2^{-1}(J)\,;I,J\in\mathcal{I}\}\\ &\subset\{g_1^{-1}(A)\cap g_2^{-1}(B)\,;A,B\in\mathrm{Borel}(\mathbb{R})\}\subset\sigma_S(\sigma\{g_1\}\cup\sigma\{g_2\})\end{aligned}$$

である。従って $\sigma_S(\sigma\{g_1\}\cup\sigma\{g_2\})$ は $f^*\mathcal{I}^2$ を内包する S 上の σ 加法族である。そのようなものの最小であることから次の関係が導かれる。

$$\sigma\{f\}=\sigma_S(f^*\mathcal{I}^2)\subset\sigma_S(\sigma\{g_1\}\cup\sigma\{g_2\})$$

ここで左の等号は冒頭の結果を適用したものである。よって結論に到達した。 □

5.10 系. $(S,\mathcal{B})$ を非自明な可測空間とする。写像 $f:S\to\mathbb{R}^d$ が $\mathcal{B}$ 可測であることとその各成分が $\mathcal{B}$ 可測であることは同値である。

証明. f の成分を $g_1,\,g_2,\,\ldots,\,g_d$ とする。目標は次を示すことである。

$$\sigma\{f\}\subset\mathcal{B}\Leftrightarrow\sigma\{g_i\}\subset\mathcal{B}\ \forall i=1,2,\ldots,d$$

さて $\mathcal{B}$ は σ 加法族なので右条件は $\sigma(\sigma\{g_1\}\cup\sigma\{g_2\}\cup\cdots\cup\sigma\{g_d\})\subset\mathcal{B}$ と書き換えられる。ところが定理 5.9 によればまさに $\sigma\{f\}=\sigma\{g_1,g_2,\ldots,g_d\}$ なので同値性が導かれた。 □

5.11 演習問題. 連続写像 $f:\mathbb{R}^k\to\mathbb{R}^d$ は $\mathrm{Borel}(\mathbb{R}^k)$ 可測であることを示せ。

5.12 例. $\alpha\in GL(d)$, $\beta\in\mathbb{R}^d$ とする。1 次変換 $f:\mathbb{R}^d\to\mathbb{R}^d$, $x\mapsto\alpha^{-1}(x-\beta)$ を考えよう。連続ゆえ f は $\mathrm{Borel}(\mathbb{R}^d)$ 可測である。よって系 5.10 を適用して以下が導かれる。

$$\alpha B + \beta = f^{-1}(B) \in \mathrm{Borel}(\mathbb{R}^d)\ \forall B \in \mathrm{Borel}(\mathbb{R}^d)$$

5.13 例. $\mathbb{R}^2$ と $\mathbb{C}$ は $(x_1, x_2) \mapsto x_1 + \sqrt{-1}x_2$ により全単写対応している。2 次元空間とみなすとき $|b-a|$ は 2 点 $a, b \in \mathbb{C}$ の距離を表し、これが $\mathbb{C}$ に位相を導入する。よって $\mathbb{C}$ の開集合すべてで生成される $\mathbb{C}$ 上の σ-加法族という概念が有効である。$(S, \mathcal{B})$ を非自明な可測空間とする。系 5.10 により関数 $f : S \to \mathbb{C}$ の $\mathcal{B}$ 可測性は定義 3.23 におけるものと一致している。

5.14 定義. $\mathbb{R}^n$ **値確率変数**($\mathbb{R}^n$-valued random variable) あるいは n **次元確率変数**とは $\mathcal{F}$ 可測写像 $X : \Omega \to \mathbb{R}^n$ をいう。また**複素数値確率変数**とは $\mathcal{F}$ 可測関数 $X : \Omega \to \mathbb{C}$ をいう。

系 5.10 を確率空間上であてはめると次が得られる。

5.15 系. 写像 $X : \Omega \to \mathbb{R}^n$ が $\mathbb{R}^n$ 値確率変数であることと各成分が実確率変数であることは同値である。

5.16 演習問題. S 上の集合族 $\mathcal{A}_1, \mathcal{A}_2, \ldots, \mathcal{A}_n$ に対して次の関係が成り立つことを示せ。

$$\sigma(\sigma(\mathcal{A}_1) \cup \sigma(\mathcal{A}_2) \cup \cdots \cup \sigma(\mathcal{A}_n)) = \sigma(\mathcal{A}_1 \cup \mathcal{A}_2 \cup \cdots \cup \mathcal{A}_n).$$

6 確率変数と結合分布

実確率変数系 $X_1, X_2, \ldots, X_n$ を束ねるのは、いくつかの量を統合的に観測することである。束ねてできた $\mathbb{R}^n$ 値確率変数を別の写像と合成することは、特定できる複合要因を持ちしかも因果関係がきちんと記述できるような現象の観測を意味する。

さて X が $\mathbb{R}^n$ 値確率変数なら $B \in \mathrm{Borel}(\mathbb{R}^n) \Rightarrow X^{-1}(B) \in \mathcal{F}$ が成り立つので関数

$$\mathcal{L}(X, \cdot) : \mathrm{Borel}(\mathbb{R}^n) \to \overline{\mathbb{R}}, B \mapsto P(X^{-1}(B))$$

が定義される。それは確率測度である。$\mathcal{L}(X, \cdot)$ は観測量の平均などを記述する。

6.1 定義. (i) $\mathbb{R}^n$ 値確率変数 X に対し確率測度 $\mathcal{L}(X, \cdot)$ を X の**分布**(distribution) あるいは、その成分を $X_1, X_2, \ldots, X_n$ とするとき、それらの**結合分布**(joint distribution) という。他方、各成分の分布 $\mathcal{L}(X_1, \cdot), \mathcal{L}(X_2, \cdot), \ldots, \mathcal{L}(X_n, \cdot)$ を**周辺分布**(marginal distribution) という。
(ii) $(\mathbb{R}^n, \mathrm{Borel}(\mathbb{R}^n))$ 上の確率測度 μ が与えられたとする。確率変数 X が**分布** μ **に従う**とは $\mathcal{L}(X, \cdot) = \mu$ を満たすことをいう。

6.2 注意. 確率変数系 $X_1, \ldots, X_n$ と $Y_1, \ldots, Y_n$ が与えられたとする。$\mathcal{L}(X_i, \cdot) = \mathcal{L}(Y_i, \cdot)$ $\forall i = 1, 2, \ldots, n$ であっても $\mathcal{L}((X_1, X_2, \ldots, X_n), \cdot) = \mathcal{L}((Y_1, Y_2, \ldots, Y_n), \cdot)$ とは限らない。

6.3 演習問題. Lebesgue モデル $([0,1), \mathrm{Borel}([0,1)), \lambda)$ で考える。

$$X(\omega) = \begin{cases} 0 & 0 \le \omega < 1/2 \\ 1 & 1/2 \le \omega < 1 \end{cases}, \quad Y(\omega) = \begin{cases} 0 & \omega \in [0, 1/4) \cup [1/2, 3/4) \\ 1 & \omega \in [1/4, 1/2) \cup [3/4, 1) \end{cases}$$

二つの確率変数系として $X_1 = X, X_2 = X$ と $Y_1 = X, Y_2 = Y$ を与える。
(i) $\mathcal{L}(X_1, \cdot) = \mathcal{L}(Y_1, \cdot), \mathcal{L}(X_2, \cdot) = \mathcal{L}(Y_2, \cdot)$ を確かめよ。
(ii) $\mathcal{L}((X_1, X_2), \cdot) \ne \mathcal{L}((Y_1, Y_2), \cdot)$ を確かめよ。

6.4 定義. ある性質が P に関しほとんどいたるところ成立することを P-a.s. と略記する。

復習

$(S, \mathcal{B}, \mu)$ を測度空間とする。部分集合 N が μ **零集合**(null set) であるとは条件

$$\exists B \in \mathcal{B} \text{ s.t. } \mu(B) = 0, N \subset B$$

を満たすことをいう。また集合対 $A \subset B$ に対して $B \setminus A$ が μ 零集合であるとき A は B に関して μ-*a.e.* **集合**であるという。通常 A μ-a.e. on B と略記する。

A をある性質の成立する集合とする。A μ-a.e. on B であるときその性質は**測度 μ に関し B 上ほとんどいたるところ**(almost everywhere) 成立するという。

6.5 注意. $N \in \mathcal{F}$ とするとき N が P 零集合であるとは $P(N) = 0$ と同じことである。また $A \in \mathcal{F}$ とするとき A が（Ω に関して）P-a.s. 集合であるとは $P(A) = 1$ と同じことである。しかし P-a.s. 集合や P 零集合が $\mathcal{F}$ に属さないこともあり得る。そのような不都合を避けるときは完備な確率空間上で議論する。

6.6 演習問題. (i) A は P-a.s. 集合かつ $A \subset B$ なら B も P-a.s. 集合であることを示せ。
(ii) A_n を P-a.s. 集合の列とすると $\bigcap_{n=1}^{\infty} A_n$ も P-a.s. 集合であることを示せ。

6.7 補題. $\mathbb{R}^n$ 値確率変数 X, Y に対し $X = Y$ P-a.s. なら $\mathcal{L}(X, \cdot) = \mathcal{L}(Y, \cdot)$ である。

証明. $\Omega_0 := \{\omega : X(\omega) = Y(\omega)\} \in \mathcal{F}$ である。従って $X = Y$ P-a.s. とは $P(\Omega_0) = 1$ のことである。さて補集合 Ω_0^c の任意の部分集合は測度 0 となるので

$$P(A) = P(A \cap \Omega_0) + P(A \cap \Omega_0^c) = P(A \cap \Omega_0) \ \forall A \in \mathcal{F}$$

が成り立つ。従って $X^{-1}(B) \cap \Omega_0 = Y^{-1}(B) \cap \Omega_0$ $\forall B \in \mathrm{Borel}(\mathbb{R}^d)$ とあわせて結論を得る。 □

6.8 演習問題. 二つの実確率変数系 $X_1, X_2, \ldots, X_n$ と $Y_1, Y_2, \ldots, Y_n$ に対し $X_i = Y_i$ P-a.s. $\forall i = 1, 2, \ldots, n$ なら $\mathcal{L}((X_1, X_2, \ldots, X_n), \cdot) = \mathcal{L}((Y_1, Y_2, \ldots, Y_n), \cdot)$ であることを示せ。

6.9 系. $\mathbb{R}^n$ 値確率変数 X と $a \in \mathbb{R}^n$ に対し $X = a$ P-a.s. と $\mathcal{L}(X, \cdot) = \delta_a$ は同値である。

証明. $B \in \mathrm{Borel}(\mathbb{R}^n)$ とする。定数写像 $\Omega \to \mathbb{R}^n$, $\omega \mapsto a$ を Y とかくと、$a \in B$ のとき $Y^{-1}(B) = \Omega$ であり $a \notin B$ のとき $Y^{-1}(B) = \emptyset$ であるから、$\mathcal{L}(Y, \cdot) = \delta_a$ が成り立つ。従って補題 6.7 より、$X = a$ P-a.s. であるなら $\mathcal{L}(X, \cdot) = \delta_a$ が成り立つ。逆に $\mathcal{L}(X, \cdot) = \delta_a$ であるなら $P(X = a) = P(X^{-1}(\{a\})) = \delta_a(\{a\}) = 1$ が成り立つ。 □

定義 5.5 を一段抽象化した概念を導入する。

6.10 定義. $(S, \mathcal{B})$, $(T, \mathcal{M})$ を非自明な可測空間とする。写像 $\varphi : S \to T$ が**対 $\mathcal{B}$, $\mathcal{M}$ に関して可測**であるとは $\varphi^{-1}(A) \in \mathcal{B}$ $\forall A \in \mathcal{M}$ を満たすことをいう。

6.11 注意. 可測性に言及するとき対 $\mathcal{B}$, $\mathcal{M}$ のうち値域 T 上の σ 加法族 $\mathcal{M}$ のほうを暗黙の了解のもと省略することも多い。たとえば、$T = \mathbb{R}^d$, $\mathcal{M} = \mathrm{Borel}(\mathbb{R}^d)$ の場合がその典型で、実際、定義 5.5 ではそのような省略がなされている。

6.12 定義. 写像 $f : \mathbb{R}^k \to \mathbb{R}^d$ が *Borel* **可測**であるとは、f が対 $\mathrm{Borel}(\mathbb{R}^k)$, $\mathrm{Borel}(\mathbb{R}^d)$ に関して可測であることをいう。

6.13 定理. $(S, \mathcal{B})$, $(T, \mathcal{M})$ を非自明な可測空間とする。写像 $\varphi : S \to T$ が対 $\mathcal{B}$, $\mathcal{M}$ に関して可測かつ関数 $f : T \to \overline{\mathbb{R}}$ が $\mathcal{M}$ 可測なら合成関数 $f \circ \varphi : S \to \overline{\mathbb{R}}$ は $\mathcal{B}$ 可測である。

証明. $a \in \mathbb{R}$ とする。まず f の可測性により $\{t \in T : f(t) > a\} \in \mathcal{M}$ であり、さらに φ の可測性により $\{s \in S : f(\varphi(s)) > a\} = \varphi^{-1}(\{t \in T : f(t) > a\}) \in \mathcal{B}$ が従う。 □

6.14 系. $X_1, X_2, \ldots, X_n$ を実確率変数とし $X = (X_1, X_2, \ldots, X_n)$ とおく。関数 $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ が Borel 可測なら $f(X) : \Omega \to \overline{\mathbb{R}}$ は確率変数である。

証明. 系 5.15 と定理 6.13 による。 □

6.15 注意. 本来なら写像 $X : \Omega \to \mathbb{R}^n$ と関数 $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ の合成関数であるから $f \circ X$ と書くべきであるが、確率論の慣用では $f(X)$ と表現することが多い。

6.16 例. 第 1 節で述べた関数 $\xi_k : (0, 1] \to \mathbb{R}$, $k \in \mathbb{N}$ はすべて $\mathrm{Borel}((0, 1])$ 可測である。

証明. 関数 ξ_1 の $\mathrm{Borel}((0, 1])$ 可測性は例 2.22 で述べた。写像 $\varphi : (0, 1] \to (0, 1]$ をやはり第 1 節のものとすると例 5.8(ii) により対 $\mathrm{Borel}((0, 1])$, $\mathrm{Borel}((0, 1])$ に関して可測である。従って定理 6.13 の助けを借りて帰納法を適用して結論に至る。 □

次はある意味で系 6.14 の逆が成り立つことを主張している。スタンダードマシンの使い方の一例であるので証明をつけておこう。

6.17 定理. $X : \Omega \to \mathbb{R}^n$ と $Y : \Omega \to \overline{\mathbb{R}}$ について Y が $\sigma\{X\}$ 可測であるなら $\mathrm{Borel}(\mathbb{R}^n)$ 可測関数 $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ が存在して $Y = f \circ X$ が成り立つ。

証明. まず Y が $\sigma\{X\}$ 単関数である場合を検討しよう。その $\sigma\{X\}$ 可測性により

$$Y^{-1}\{y\} \in \sigma\{X\} \ \forall y \in \mathrm{Image}\, Y$$

である。従って σ 加法族 $\sigma\{X\} = X^*\mathrm{Borel}(\mathbb{R}^n)$ の定義により

$$\forall y \in \mathrm{Image}\, Y \ \exists B(y) \in \mathrm{Borel}(\mathbb{R}^n) \text{ such that } Y^{-1}\{y\} = X^{-1}(B(y))$$

このとき $1_{Y^{-1}\{y\}} = 1_{X^{-1}(B(y))} = 1_{B(y)} \circ X$ であるから次が得られる。

$$Y = \sum_{y \in \mathrm{Image}\, Y} y 1_{Y^{-1}\{y\}} = \sum_{y \in \mathrm{Image}\, Y} y 1_{B(y)} \circ X.$$

従って $\mathrm{Borel}(\mathbb{R}^n)$ 可測関数 $f := \sum_{y \in \mathrm{Image}\, Y} y 1_{B(y)}$ に対して $Y = f \circ X$ が成り立つ。

次に Y が非負値 $\sigma\{X\}$ 可測である場合の議論に移る。このとき $\sigma\{X\}$ 単関数の列 Y_n であって $0 \leq Y_n \leq Y_{n+1} \ \forall n$, $\sup_{n \in \mathbb{N}} Y_n = Y$ を満たすものが存在する。各 Y_n は単関数であるから、前段落で示したように

$$\forall n \in \mathbb{N} \ \exists f_n \ \mathrm{Borel}(\mathbb{R}^n) \text{ 可測関数 such that } Y_n = f_n \circ X$$

すると $f := \sup_{n\in n\mathbb{N}} f_n$ は Borel($\mathbb{R}^n$) 可測関数であり次が成り立つ。

$$Y = \sup_{n\in\mathbb{N}} Y_n = \sup_{n\in\mathbb{N}}(f_n \circ X) = (\sup_{n\in\mathbb{N}} f_n) \circ X = f \circ X.$$

一般には $\max\{Y, 0\}$ と $\max\{-Y, 0\}$ に上の結果を適用すればよい。 □

蛇足. 更に系 5.15 も考慮に入れると次の命題が得られる。

> $X : \Omega \to \mathbb{R}^n$ と $Y : \Omega \to \mathbb{R}^d$ について $\sigma\{Y\} \subset \sigma\{X\}$ 即ち Y が $\sigma\{X\}$ 可測であるなら Borel 可測写像 $\varphi : \mathbb{R}^n \to \mathbb{R}^d$ が存在して $Y = \varphi \circ X$ が成り立つ。

記号

$(S, \mathcal{B})$, $(T, \mathcal{M})$ を非自明な可測空間とする。可測写像 $\varphi : S \to T$ と $(S, \mathcal{B})$ 上の測度 μ に対して関数 $\mathcal{M} \to \overline{\mathbb{R}}$, $A \mapsto \mu(\varphi^{-1}(A))$ を $\varphi_*\mu$ と表記する。

6.18 定義. $(T, \mathcal{M})$ 上の測度 $\varphi_*\mu$ を φ による μ の**像測度**(image measure) という。

$$\begin{array}{ccc} (S, \mathcal{B}) & \xrightarrow{\varphi} & (T, \mathcal{M}) \\ \mu & & \varphi_*\mu \end{array}$$

6.19 例. (i) $\alpha, \beta \in \mathbb{R}$ かつ $\alpha \neq 0$ とする。写像 $\mathbb{R} \to \mathbb{R}$, $x \mapsto \alpha x + \beta$ による平均 $a \in \mathbb{R}$ 分散 $t \in \mathbb{R}_{>0}$ の正規分布の像測度は平均 $\alpha a + \beta$ 分散 $\alpha^2 t$ の正規分布である。
(ii) 写像 $\mathbb{R} \to \mathbb{R}$, $x \mapsto x^2$ による標準正規分布の像測度は自由度 1 の χ^2 分布である。
(iii) $c \in \mathbb{R}_{>0}$ とする。写像 $(0, 1] \to \mathbb{R}$, $x \mapsto -\log x/c$ による $(0, 1]$ 上の Lebesgue 測度の像測度はパラメータ c の指数分布である。
(iv) $a \in \mathbb{R}^d$ とする。定数写像 $(0, 1] \to \mathbb{R}$, $x \mapsto a$ による $(0, 1]$ 上の Lebesgue 測度の像測度は a に質量を持つ Dirac 測度である。

証明. (i), (ii) および (iii) はそれぞれ例 4.19, 例 4.22 と演習問題 4.23(ii) の言い換えである。 □

次においてもスタンダードマシンが有効に機能している。

6.20 定理. $(S, \mathcal{B})$, $(T, \mathcal{M})$ を非自明な可測空間とする。更に対 $\mathcal{B}$, $\mathcal{M}$ に関して可測な写像 $\varphi : S \to T$ と $(S, \mathcal{B})$ 上の測度 μ が与えられたとする。
(i) 非負値 $\mathcal{M}$ 可測関数 $f : T \to \overline{\mathbb{R}}$ に対して次が成り立つ。

$$\int_S f \circ \varphi \, \mu = \int_T f \, \varphi_*\mu$$

(ii) $\mathcal{M}$ 可測関数 $f : T \to \overline{\mathbb{R}}$ に対してその $\varphi_*\mu$ 可積分性は $f \circ \varphi$ の μ 可積分性と同値であり、可積分なら上の等式が成り立つ。

証明. (i) f が $A \in \mathcal{M}$ の定義関数であるとする。このとき $f \circ \varphi$ は集合 $\varphi^{-1}(A)$ の定義関数である。従って次が成り立つ。

$$\int_S f \circ \varphi \, \mu = \mu(\varphi^{-1}(A)) = \varphi_*\mu(A) = \int_T f \, \varphi_*\mu.$$

積分の線形性により上の関係は f が非負値 $\mathcal{M}$ 単関数である場合にも成り立つ。一般には非負値 $\mathcal{M}$ 単関数 $T \to \mathbb{R}$ の列 f_k で $f_k \le f_{k+1}\ \forall k \in \mathbb{N}$, $\sup_{k\in\mathbb{N}} f_k(t) = f(t)\ \forall t \in T$ を満たすものが存在する。各 $k \in \mathbb{N}$ に対して

$$\int_S f_k \circ \varphi\, \mu = \int_T f_k\, \varphi_* \mu$$

が成り立つので単調収束定理を適用して結論に至る。

(ii) 非負値関数 $|f|$ に対して (i) を適用して可積分性に関する同値性が示せる。次に (i) を $\max\{f, 0\}$ と $\max\{-f, 0\}$ のそれぞれに適用すると求める関係式が得られる。 □

6.21 例. (i) $[0, +\infty)$ 上の非負 Borel 可測関数 f に対して

$$\int_{\mathbb{R}} f(x^2) \frac{1}{\sqrt{2\pi}} \exp\{-x^2/2\} \lambda(dx) = \int_{(0,+\infty)} f(x) \frac{1}{\sqrt{2\pi}} \exp\{-x/2\} x^{-1/2} \lambda(dx)$$

(ii) $c \in \mathbb{R}_{>0}$ とする。$(0, +\infty)$ 上の非負 Borel 可測関数 f に対して

$$\int_{(0,1)} f(-\log x/c) \lambda(dx) = \int_{(0,+\infty)} f(x) c e^{-cx} \lambda(dx)$$

証明. 例 6.19(ii), (iii) でそれぞれ像測度が特定されている。 □

6.22 例. 第 1 節の写像 $\varphi : (0, 1] \to (0, 1]$ と非負値 Borel$((0, 1])$ 可測関数 $f : (0, 1] \to \overline{\mathbb{R}}$ に対して $\int_{(0,1]} f \circ \varphi\, \lambda = \int_{(0,1]} f\, \lambda$ が成り立つ。

6.23 例. $\mathcal{B}$, $\mathcal{M}$ を集合 $S \neq \emptyset$ 上の σ 加法族の対で $\mathcal{M} \subset \mathcal{B}$ を満たすものとする。$(S, \mathcal{B})$ 上の測度 μ に対してその定義域を $\mathcal{M}$ へ制限して得られる $(S, \mathcal{M})$ 上の測度を ν とする。即ち

$$\nu : \mathcal{M} \to \overline{\mathbb{R}}, A \mapsto \mu(A)$$

である。恒等写像 $S \to S$ を対 $\mathcal{B}$, $\mathcal{M}$ に関する可測写像と見なすときそれによる μ の像測度がまさに ν である。従って定理 6.20 によれば非負値 $\mathcal{M}$ 可測関数 $f : S \to \overline{\mathbb{R}}$ に対して

$$\int_S f\, \mu = \int_S f\ \nu$$

6.24 定理. $(S, \mathcal{B})$, $(T, \mathcal{M})$, $(U, \mathcal{W})$ を非自明な可測空間とする。写像 $\varphi : S \to T$ が対 $\mathcal{B}$, $\mathcal{M}$ に関して、写像 $\psi : T \to U$ が対 $\mathcal{M}$, $\mathcal{W}$ に関して可測なら

1° 合成写像 $\psi \circ \varphi : S \to U$ は対 $\mathcal{B}$, $\mathcal{W}$ に関して可測であり、
2° $(S, \mathcal{B})$ 上の測度 μ に対して $(\psi \circ \varphi)_* \mu = \psi_*(\varphi_* \mu)$ が成り立つ。

証明. 合成写像の可測性の証明は定理 6.13 のそれと同様である。また

$$(\psi \circ \varphi)_* \mu(A) = \mu(\varphi^{-1}(\psi^{-1}(A))) = (\varphi_* \mu)(\psi^{-1}(A)) = \psi_*(\varphi_* \mu)(A)\ \forall A \in \mathcal{W}.$$

であるから後半の主張も示せた。 □

6.25 例. (i) 第 1 節で述べた写像 $\varphi:(0,1]\to(0,1]$ は $\varphi_*\lambda=\lambda$ を満たす。
(ii) 第 1 節で述べた関数 $\xi_k:(0,1]\to\mathbb{R},\ k\in\mathbb{N}$ は Lebesgue モデル上の確率変数である。その分布はすべて $p\delta_1+(1-p)\delta_{-1}$ である。
(iii) $\lambda(\{\xi_1=1\}\cap\varphi^{-1}(A))=p\lambda(A)\ \forall A\in\mathrm{Borel}((0,1])$ が成り立つ。

証明. (i) Lebesgue 測度の一意性により $0\le a<b\le 1\Rightarrow\lambda(\varphi^{-1}((a,b]))=b-a$ を示せばよいが、実際 $\varphi^{-1}((a,b])=(pa,pb]\cup((1-p)a+p,(1-p)b+p]$ であるから成り立つ。
(ii) 定理 6.24 と (i) により $\xi_k=\xi_{k-1}\circ\varphi$ の分布は ξ_{k-1} の分布と等しい。また例 3.15 により ξ_1 の分布は $p\delta_1+(1-p)\delta_{-1}$ である。
(iii) (i) のときと同じく次を示せばよい。

$$0\le a<b\le 1\Rightarrow\lambda(\{\xi_1=1\}\cap\varphi^{-1}((a,b]))=p(b-a)$$

実際 $\{\xi_1=1\}=(0,p]$ より $\{\xi_1=1\}\cap\varphi^{-1}((a,b])=(pa,pb]$ であるから成り立つ。 □

6.26 定義. $(S,\mathcal{B})$ を非自明な可測空間、$\varphi:S\to S$ を対 $\mathcal{B},\mathcal{B}$ に関する可測写像、μ を $(S,\mathcal{B})$ 上の測度とする。$\varphi_*\mu=\mu$ であるとき写像 φ は**測度 μ を保存する**(μ-preserving) あるいは**測度 μ は φ 不変である**(φ-invariant) という。

6.27 例. Lebesgue モデルを確率空間とするとき第 1 節で述べた確率変数 $\xi_k:(0,1]\to\mathbb{R}$, $k\in\mathbb{N}$ の結合分布は $\mathbb{N}$ の有限部分集合 I に対して $\lambda(\xi_i=1\,\forall i\in I)=p^{\#I}$ を満たす。

証明. $p^0=1$ と見れば $I=\emptyset$ でも成り立つ。$\#I$ に関し帰納法を適用する。例 6.25(ii) より $\#I=1$ なら成り立つ。$\#I=n$ のとき成り立つと仮定して $\#I=n+1$ の場合を考察する。

$$m:=\min I,\ J:=\{i-m\,;i\in I,i>m\}$$

とおく。写像 $\varphi:(0,1]\to(0,1]$ を第 1 節のものとすると

$$\{\omega:\xi_i=1\,\forall i\in I\}=(\varphi^{m-1})^{-1}(\{\omega:\xi_1=1,\xi_{j+1}=1\,\forall j\in J\})$$

という関係が $\xi_i=\xi_{i-m+1}\circ\varphi^{m-1}$ から派生する。他方、定理 6.24 を適用して、写像 φ と同じく φ^{m-1} も測度 λ を保存するすることが分かる。従って

$$\lambda(\xi_i=1\,\forall i\in I)=\lambda(\xi_1=1,\xi_{j+1}=1\,\forall j\in J)=\lambda(\{\xi_1=1\}\cap\varphi^{-1}(\{\xi_j=1\,\forall j\in J\})).$$

ここで $\#J=\#I-1=n$ である。例 6.25(iii) および帰納法の仮定を適用して

$$\lambda(\{\xi_1=1\}\cap\varphi^{-1}(\{\xi_j=1\,\forall j\in J\}))=p\lambda(\xi_j=1\,\forall j\in J)=pp^{\#J}$$

ゆえに $\#I=n+1$ に対しても命題は成り立つ。 □

記号

非負値あるいは P 可積分な確率変数（可積分な複素数値確率変数）X に対して次を X の**期待値**(expectation) あるいは平均 (mean) という。

$$E[X]:=\int_\Omega X\,P,\ \ E[X\,;A]:=\int_A X\,P\ \text{ ただし } A\in\mathcal{F}$$

6.28 演習問題. $(S, \mathcal{B})$ を非自明な可測空間、$p > 0$ とする。$\mathcal{B}$ 可測関数 $f : S \to \overline{\mathbb{R}}$ について $|f|^p$ も $\mathcal{B}$ 可測であることを示せ。

6.29 定義. $(S, \mathcal{B}, \mu)$ を非自明な測度空間、$p > 0$ とする。$\mathcal{B}$ 可測関数 $f : S \to \overline{\mathbb{R}}$ が p **乗可積分**であるとは $|f|^p$ が μ 可積分であることをいう。(複素数値関数についても同様な概念あり)

次の命題は全測度が有限である空間でしか通用しない。

6.30 演習問題. $p > 1$ とする。p 乗可積分な確率変数は可積分であることを示せ。

次の命題を確認しておこう。

6.31 定理. $(S, \mathcal{B}, \mu)$ を非自明な測度空間、$f, g : S \to \overline{\mathbb{R}}$ を $\mathcal{B}$ 可測関数、$p > 0$ とする。
(i) f が p 乗可積分であれば $|f| < +\infty$ μ-a.e. である。
(ii) $\displaystyle\int_S |f|^p \, \mu = 0$ と $f = 0$ μ-a.e. は同値である。
(iii) f が可積分かつ $f = g$ μ-a.e. であれば g も可積分かつ $\displaystyle\int_S f \, \mu = \int_S g \, \mu$

記号

確率変数 X と $p > 0$ に対して $E[|X|^p]$ を p **次絶対モーメント**という。また $n \in \mathbb{N}$ とするとき n 乗可積分な確率変数 X に対して $E[X^n]$ を n **次モーメント**という。

6.32 補題. $(S, \mathcal{B}, \mu)$ を非自明な測度空間、$f, g : S \to \mathbb{R}$ を 2 乗可積分 $\mathcal{B}$ 可測関数とする。
(i) 積 fg は可積分、和 $f + g$ は 2 乗可積分である。
(ii) 不等式 $\displaystyle\int_S fg \, \mu \leq \Big(\int_S f^2 \, \mu \int_S g^2 \, \mu \Big)^{1/2}$ がなりたつ。これを *Schwarz* **の不等式**という。

証明. (i) 関係 $|fg| \leq (f^2 + g^2)/2$, $(f + g)^2 \leq 2f^2 + 2g^2$ と可積分関数の線形結合は再び可積分であることから従う。

(ii) $a, b \geq 0$ とすると $2ab \leq a^2 t + b^2/t \ \forall t > 0$ が成り立つ。よって

$$2 \int_S |fg| \, \mu \leq t \int_S f^2 \, \mu + \frac{1}{t} \int_S g^2 \, \mu \ \forall t > 0$$

他方 $a, b \geq 0$ に対して $\inf_{t>0}(at + b/t) = 2\sqrt{ab}$ であるから Schwarz の不等式が導ける。 □

記号

2 乗可積分な確率変数 X, Y に対して次を X, Y の**共分散**(covariance) という。

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

後者を X の**分散**(variance) とよぶ。

6.33 演習問題. X, Y を 2 乗可積分な確率変数とする。
(i) $\mathrm{Var}[X] = 0$ ならある $c \in \mathbb{R}$ が存在して $X = c$ P-a.s. が成り立つことを示せ。
(ii) $\mathrm{Cov}[X, Y] = E[XY] - E[X]E[Y]$, $\mathrm{Cov}[X, Y]^2 \leq \mathrm{Var}[X]\mathrm{Var}[Y]$ が成り立つことを示せ。

補題 6.34 の不等式を *Markov* **の不等式**、系 6.35 の不等式を *Chebyshev* **の不等式**という。

6.34 補題. 非負値確率変数 X と $\lambda \in \mathbb{R}_{>0}$ に対して $P(X \geq \lambda) \leq E[X]/\lambda$ が成り立つ。

証明. $\lambda P(X \geq \lambda) \leq E[X\,; X \geq \lambda] \leq E[X]$. □

6.35 系. 2 乗可積分確率変数 X と $\delta \in \mathbb{R}_{>0}$ に対し $P(|X - E[X]| \geq \delta) \leq \mathrm{Var}[X]/\delta^2$ である。

証明. 非負値確率変数 $|X - E[X]|^2$ と δ^2 に対して補題 6.34 を適用する。 □

補題 6.32 は $p > 1$ のときに p 乗可積分な関数についての命題へ拡張できる。

6.36 演習問題. $a, b \in \mathbb{R}_{\geq 0}$, $p, q > 1$ かつ $1/p + 1/q = 1$ とする。
(i) $(a+b)^p \leq t^{1-p}a^p + (1-t)^{1-p}b^p$ $0 < \forall t < 1$ であることを示せ。
(ii) $ab \leq \dfrac{1}{p}a^p t^{p-1} + \dfrac{1}{q}b^q \dfrac{1}{t}$ $\forall t > 0$ であることを示せ。
(iii) さらに $a > 0$, $b > 0$ とせよ。(i) において等号が成立するのは $t = a/(a+b)$ のときに限りまた (ii) において等号が成立するのは $t = b^{q-1}/a$ のときに限ることを示せ。

6.37 定理. $(S, \mathcal{B}, \mu)$ を非自明な測度空間、$p, q > 1$ かつ $1/p + 1/q = 1$ とする。$\mathcal{B}$ 可測関数 $f, g, h : S \to \mathbb{R}$ について f, g は p 乗可積分、h は q 乗可積分であるとする。
(i) 和 $f + g$ は p 乗可積分、積 fh は可積分である。
(ii) $\left(\displaystyle\int_S |f+g|^p \,\mu\right)^{1/p} \leq \left(\displaystyle\int_S |f|^p \,\mu\right)^{1/p} + \left(\displaystyle\int_S |g|^p \,\mu\right)^{1/p}$ が成り立つ。*Minkowski* **の不等式**
(iii) $\displaystyle\int_S fh\,\mu \leq \left(\int_S |f|^p \,\mu\right)^{1/p}\left(\int_S |h|^q \,\mu\right)^{1/q}$ が成り立つ。*Hölder* **の不等式**

証明. (i) 演習問題 6.36(i) で得た不等式において $t = 1/2$ としたものと演習問題 6.36(ii) で得た不等式において $t = 1$ としたものから次が成り立つことが分かる。

$$|f+g|^p \leq (|f| + |g|)^p \leq 2^{p-1}(|f|^p + |g|^p), |fh| \leq |f|^p/p + |h|^q/q$$

(ii) 演習問題 6.36(i) で得た不等式を適用して

$$\int_S (|f| + |g|)^p \,\mu \leq t^{1-p} \int_S |f|^p \,\mu + (1-t)^{1-p} \int_S |g|^p \,\mu \;\; 0 < \forall t < 1.$$

他方 $a, b \geq 0$ に対して $\inf_{0<t<1}\{t^{1-p}a + (1-t)^{1-p}b\} = (a^{1/p} + b^{1/p})^p$ である。
(iii) (ii) と同様であるが今度は演習問題 6.36(ii) で得た不等式を適用する。 □

6.38 定理. $X_1, X_2, \ldots, X_n$ を実確率変数とする。
(i) 非負値 $\mathrm{Borel}(\mathbb{R}^n)$ 可測関数 $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ に対して次が成り立つ。

$$E[f(X_1, X_2, \ldots, X_n)] = \int_{\mathbb{R}^n} f\, \mathcal{L}((X_1, X_2, \ldots, X_n), \cdot)$$

(ii) $\mathrm{Borel}(\mathbb{R}^n)$ 可測関数 $f : T \to \overline{\mathbb{R}}$ に対してその $\mathcal{L}((X_1, X_2, \ldots, X_n), \cdot)$ 可積分性は確率変数 $f(X_1, X_2, \ldots, X_n)$ の P 可積分性と同値であり、可積分なら上の等式が成り立つ。

証明. X_1, X_2, ..., X_n を束ねて得られる写像 $\Omega \to \mathbb{R}^n$ は系 5.15 により $\mathbb{R}^n$ 値確率変数である。従って定理 6.20 を適用することにより結論を得る。 □

6.39 系. X, Y を実確率変数とする。
(i) $X \geq 0$ P-a.s. あるいは X が P 可積分であるなら $E[X] = \int_{\mathbb{R}} x\,\mathcal{L}(X, dx)$
(ii) X, Y ともに 2 乗可積分なら $\mathrm{Cov}[X, Y] = \int_{\mathbb{R}^2} (x - E[X])(y - E[Y])\,\mathcal{L}((X, Y), dxdy)$
(iii) $p > 0$ に対して $E[|X|^p] = \int_{\mathbb{R}} |x|^p\,\mathcal{L}(X, dx)$

まず離散分布の例について期待値と分散を述べる。

6.40 例. (i) X をパラメータ $n \in \mathbb{N}$, $p \in (0, 1)$ を持つ二項分布に従う確率変数とすると

$$E[X] = np, \mathrm{Var}[X] = np(1 - p).$$

(ii) X をパラメータ $n \in \mathbb{N}$, $p \in (0, 1)$ を持つ負の二項分布に従う確率変数とすると

$$E[X] = \frac{n(1-p)}{p}, \mathrm{Var}[X] = \frac{n(1-p)}{p^2}.$$

(iii) X を強度 c の Poisson 分布に従う確率変数とすると

$$E[X] = c, \mathrm{Var}[X] = c.$$

証明. (i), (ii), (iii) はそれぞれ例 3.10, 例 3.11, 例 3.12 から従う。 □

次に絶対連続分布の場合にふれる。

6.41 例. (i) X をパラメータ t, c の gamma 分布に従う確率変数とすると

$$E[X] = \frac{t}{c}, \mathrm{Var}[X] = \frac{t}{c^2}.$$

(ii) X をパラメータ $a \in \mathbb{R}$, $c \in \mathbb{R}_{>0}$ を持つ両側指数分布に従う確率変数とすると

$$E[X] = a, \mathrm{Var}[X] = \frac{2}{c^2}.$$

(iii) X を平均 a 分散 t の正規分布に従う確率変数とすると

$$E[X] = a, \mathrm{Var}[X] = t.$$

これがパラメータ a, t の名前の由来である。

証明. (i), (ii), (iii) はそれぞれ例 4.3, 例 4.4, 補題 4.9(i) から従う。 □

6.42 定義. n 次元確率変数 X の成分を $X_1, X_2, \ldots, X_n$ とする。各 X_i が可積分である場合にそれらの平均値からなる $\mathbb{R}^n$ の元 $(E[X_1], E[X_2], \ldots, E[X_n])$ を $E[X]$ と略記し、これを X の期待値あるいは平均ベクトルという。また各 X_i が 2 乗可積分である場合に $\mathrm{Cov}[X_i, X_j]$ を成分とする行列を $\mathrm{Var}[X]$ と表記し、これを X の**共分散行列**(covariance matrix) という。

Chebyshev の不等式の典型的な応用例として *Weierstrass* **の多項式近似定理**（例 6.46）の証明を紹介しておこう。X をパラメータ $n \in \mathbb{N}$, $p \in (0,1)$ を持つ二項分布に従う確率変数とする。このとき有界 Borel 可測関数 $f : \mathbb{R} \to \mathbb{C}$ に対して次が成り立つ。

$$E[f(X/n)] = \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} f(k/n)$$

右辺を p の多項式と見るときこれを関数 f に付随する *Bernstein* **多項式**という。$E[X/n] = p$, $\mathrm{Var}[X/n] = p(1-p)/n \le 1/4n$ であることに注意しておく。

6.43 補題. n 次元確率変数 X, Y と有界かつ一様連続な関数 $f : \mathbb{R}^n \to \mathbb{C}$ に対して

$$|E[f(X)] - E[f(Y)]| \le \sup_{x,y:\|x-y\|<\delta} |f(x) - f(y)| + 2 \sup_{x\in\mathbb{R}^n} |f(x)| P(\|X-Y\| \ge \delta)\ \forall\delta > 0$$

が成り立つ。ここで $\|\ \ \|$ は n 次元ユークリッドノルムを表す。

6.44 演習問題. 補題 6.43 を示せ。ヒント　$\{\|X-Y\| < \delta\}$ と $\{\|X-Y\| \ge \delta\}$ に分割する。

6.45 系. 2 乗可積分確率変数 X と有界かつ一様連続な関数 $f : \mathbb{R} \to \mathbb{C}$ に対して

$$|E[f(X)] - f(E[X])| \le \sup_{x,y:|x-y|<\delta} |f(x) - f(y)| + 2 \sup_{x\in\mathbb{R}} |f(x)| \mathrm{Var}[X]/\delta^2\ \forall\delta > 0$$

証明. 補題 6.43 と Chebyshev の不等式（系 6.35）を適用する。 □

6.46 例. 連続関数 $f : [0,1] \to \mathbb{C}$ と $\delta > 0$ に対して $p \in [0,1]$ に関する一様評価

$$|\sum_{k=0}^{n} \binom{n}{k} f(k/n) p^k (1-p)^{n-k} - f(p)| \le \sup_{x,y:|x-y|<\delta} |f(x) - f(y)| + \frac{\sup_{x\in[0,1]} |f(x)|}{2n\delta^2}$$

が成り立つ。従って関数 f に付随する Bernstein 多項式は $n \to \infty$ の極限で f に一様収束する。

7　Dynkin 族定理と測度の一意性

与えられた二つの確率変数が同じ分布に従うかどうかは重要な問題である。多くの場合その判定は測度の定義域すべてで行うのではなく、コアとなる部分だけでのチェックで十分である。補題 4.15 はその典型例であったわけである。4 節ではそれを Lebesgue-Stieltjes 測度の存在・一意性定理に帰着して証明していたが、一意性は存在とは切り離して議論ができる。確率論でよく利用される Dynkin 族という概念を紹介して話を進めていこう。

7.1 定理. $(\mathbb{R}^d, \mathrm{Borel}(\mathbb{R}^d))$ 上の確率測度 μ, ν が一致するための必要十分条件は

$$\mu((a_1, +\infty) \times (a_2, +\infty) \times \cdots \times (a_d, +\infty)) = \nu((a_1, +\infty) \times (a_2, +\infty) \times \cdots \times (a_d, +\infty))$$

が任意の $a_1, a_2, \ldots, a_d$ に対して成り立つことである。

証明. 集合族 $\{(a_1,+\infty)\times(a_2,+\infty)\times\cdots\times(a_d,+\infty)\}$ で生成される $\mathbb{R}^d$ 上の σ 加法族は $\mathrm{Borel}(\mathbb{R}^d)$ であることおよび定理 7.12 から導かれる。 □

上にあげた定理は補題 4.15 を多次元に一般化したものがであるが、ここではさらに抽象化した設定で議論を展開する。

7.2 定義. 集合族 $\mathcal{D}$ で次の条件を満たすものを *Dynkin* **族**(Dynkin system) という。

(i) $\emptyset \in \mathcal{D}$.
(ii) $A \in \mathcal{D}$, $B \in \mathcal{D}$, $A \supset B \Rightarrow A \setminus B \in \mathcal{D}$.
(iii) $A_n \in \mathcal{D}\ \forall n \in \mathbb{N}$, $A_n \subset A_{n+1}\ \forall n \in \mathbb{N} \Rightarrow \bigcup_{n=1}^{\infty} A_n \in \mathcal{D}$.

単調収束定理との関連を次の演習問題 7.3 を解くことによって確認しよう。

7.3 演習問題. 非自明な可測空間 $(S,\mathcal{B})$ 上にふたつの確率測度 μ, ν が与えられたとする。このとき集合族 $\{A \in \mathcal{B} : \mu(A)=\nu(A)\}$ は Dynkin 族であることを示せ。

さて集合 S を一つ固定するとき任意の S 上の σ-加法族は Dynkin 族である。他方、演習問題 7.3 で述べた Dynkin 族が S 上の σ-加法族であれば、$\mu=\nu$ である可能性が高くなる。どのような付加条件があれば Dynkin 族が S 上の σ-加法族になるのだろうか？

記号

$\mathrm{Sbset}(S)$ 集合 S の部分集合全体の族

7.4 補題. $\mathcal{D}$ を S の部分集合からなる Dynkin 族とする.
(i) $S \in \mathcal{D}$ かつ $A \cap B \in \mathcal{D}\ \forall A \in \mathcal{D}\ \forall B \in \mathcal{D}$ なら $\mathcal{D}$ は S 上の σ-加法族である。
(ii) 任意の $B \in \mathrm{Sbset}(S)$ に対し集合族 $\{A \in \mathrm{Sbset}(S) : A \cap B \in \mathcal{D}\}$ は Dynkin 族である。

証明. (i) $S \in \mathcal{D}$ と Dynkin 族の条件 (ii) より

$$A \in \mathcal{D} \Rightarrow A^{\mathrm{c}} = S \setminus A \in \mathcal{D}.$$

次に上のことと仮定 $A \cap B \in \mathcal{D}\ \forall A \in \mathcal{D}\ \forall B \in \mathcal{D}$ より

$$A \in \mathcal{D},\ B \in \mathcal{D} \Rightarrow A \cup B = (A^{\mathrm{c}} \cap B^{\mathrm{c}})^{\mathrm{c}} \in \mathcal{D}.$$

従って $A_1, A_2, \ldots, A_n \in \mathcal{D} \Rightarrow \bigcup_{k=1}^{n} A_k \in \mathcal{D}$ である。他方 $\bigcup_{k=1}^{n} A_k \subset \bigcup_{k=1}^{n+1} A_k$ が成り立つので、Dynkin 族の条件 (iii) より

$$A_n \in \mathcal{D}\ \forall n \in \mathbb{N} \Rightarrow \bigcup_{n=1}^{\infty} A_n = \bigcup_{n=1}^{\infty}\Bigl(\bigcup_{k=1}^{n} A_k\Bigr) \in \mathcal{D}.$$

(ii) まず $\emptyset \cap B = \emptyset \in \mathcal{D}$ である。次に $A_1 \cap B \in \mathcal{D}$, $A_2 \cap B \in \mathcal{D}$, $A_1 \supset A_2$ とすると

$$(A_1 \setminus A_2) \cap B = (A_1 \cap B) \setminus (A_2 \cap B) \in \mathcal{D}$$

である。最後に $A_n \cap B \in \mathcal{D}\ \forall n \in \mathbb{N}$, $A_n \subset A_{n+1}\ \forall n \in \mathbb{N}$ とすると

$$\Bigl(\bigcup_{n=1}^{\infty} A_n\Bigr) \cap B = \bigcup_{n=1}^{\infty}(A_n \cap B) \in \mathcal{D}$$

が導かれる。よって Dynkin 族の条件がすべて確認できた。 □

補題 2.5 でのべたように S 上の σ-加法族たちの共通部はやはり S 上の σ-加法族である。これと同じことが Dynkin 族についてもいえる。

7.5 補題. Dynkin 族たち $\mathcal{D}_\alpha$ に対しそれらの共通部 $\bigcap_\alpha \mathcal{D}_\alpha$ も Dynkin 族である。

7.6 演習問題. 補題 7.5 を示せ。

7.7 定義. 集合の族 $\mathcal{A}$ に対し次の集合族を $\mathcal{A}$ **で生成される** *Dynkin* **族**と呼ぶ。

$$\bigcap \{\mathcal{D};\ \mathcal{A} \subset \mathcal{D} \subset \mathrm{Sbset}(S)\ \text{Dynkin 族}\ \}$$

ここで S は集合族 $\mathcal{A}$ に属する集合全体の合併集合である。

定理 7.1 で登場した集合族は $\{(a_1, +\infty) \times (a_2, +\infty) \times \cdots \times (a_d, +\infty)\}$ であるが、それを簡単のため $\mathcal{C}$ と表記すると条件 $A \cap B \in \mathcal{C}\ \forall A \in \mathcal{C}\ \forall B \in \mathcal{C}$ が成り立っている。

7.8 定義. 集合族 $\mathcal{C}$ で条件 $A \cap B \in \mathcal{C}\ \forall A \in \mathcal{C}\ \forall B \in \mathcal{C}$ を満たすものを π **システム**という。

7.9 例. (i) 有界な d 次元開区間全体に空集合 $\emptyset$ を付加した集合族は π システムである。
(ii) $\mathbb{R}^d$ のコンパクト部分集合のなす集合族は π システムである。

7.10 補題. π システム $\mathcal{C}$ で生成される Dynkin 族 $\mathcal{D}$ は π システムである。

証明. 集合族 $\mathcal{C}$ は π システムであることと関係 $\mathcal{C} \subset \mathcal{D}$ により

$$A \in \mathcal{C} \Rightarrow A \cap B \in \mathcal{D}\ \forall B \in \mathcal{C}.$$

さて $\mathcal{C}$ に属する集合全体の合併を S とする。次の集合族は $\mathcal{C} \subset \mathcal{D}_1$ を満たす。

$$\mathcal{D}_1 := \{A \in \mathrm{Sbset}(S) : A \cap B \in \mathcal{D}\ \forall B \in \mathcal{C}\} = \bigcap_{B \in \mathcal{C}} \{A \in \mathrm{Sbset}(S) : A \cap B \in \mathcal{D}\}$$

$\mathcal{D}$ は S の部分集合からなる Dynkin 族なので補題 7.4(ii) と補題 7.5 を適用して $\mathcal{D}_1$ が Dynkin 族であることがわかる。したがって $\mathcal{C} \subset \mathcal{D}_1 \subset \mathrm{Sbset}(S)$ であることと $\mathcal{D}$ の定義により $\mathcal{D} \subset \mathcal{D}_1$ すなわち $A \in \mathcal{D} \Rightarrow A \cap B \in \mathcal{D}\ \forall B \in \mathcal{C}$ を得る。これは次と同値である。

$$A \in \mathcal{D}, B \in \mathcal{C} \Rightarrow A \cap B \in \mathcal{D}.$$

A, B の役割を入れ替えてみよう。

$$A \in \mathcal{C} \Rightarrow A \cap B \in \mathcal{D}\ \forall B \in \mathcal{D}.$$

よって次の集合族は $\mathcal{C} \subset \mathcal{D}_2$ を満たす。

$$\mathcal{D}_2 := \{A \in \mathrm{Sbset}(S) : A \cap B \in \mathcal{D}\ \forall B \in \mathcal{D}\}$$

$\mathcal{D}_1$ に対するのと同じ論法を繰り返して $A \in \mathcal{D} \Rightarrow A \cap B \in \mathcal{D}\ \forall B \in \mathcal{D}$ を得る。 □

次は *Dynkin* **族定理**と呼ばれる。

7.11 定理. $\mathcal{C}$ を S の部分集合からなる π システムとする。
(i) 集合族 $\mathcal{C} \cup \{S\}$ で生成される Dynkin 族は σ-加法族 $\sigma_S(\mathcal{C})$ に等しい。
(ii) $\mathcal{C}$ の元からなる列 C_n であって $C_n \subset C_{n+1}\ \forall n \in \mathbb{N}$ かつ $\bigcup_{n=1}^{\infty} C_n = S$ を満たすものが存在するなら集合族 $\mathcal{C}$ で生成される Dynkin 族は σ-加法族 $\sigma_S(\mathcal{C})$ に等しい。

証明. (i) 集合族 $\tilde{\mathcal{C}} := \mathcal{C} \cup \{S\}$ も π システムである。従って補題 7.10 により $\tilde{\mathcal{C}}$ で生成される Dynkin 族 $\mathcal{D}$ も π システムになる。しかも $\mathcal{D}$ は S の部分集合族であってかつ $S \in \mathcal{D}$ なので補題 7.4(i) により $\mathcal{D}$ は S 上の σ-加法族である。またそれは $\mathcal{C}$ を含む。従って $\sigma_S(\mathcal{C})$ の定義により $\sigma_S(\mathcal{C}) \subset \mathcal{D}$ を得る。

他方 σ-加法族は Dynkin 族である。よって $\sigma_S(\mathcal{C})$ は $\mathcal{C} \cup \{S\}$ を含みかつ S の部分集合からなる Dynkin 族となる。$\mathcal{D}$ の定義により $\mathcal{D} \subset \sigma_S(\mathcal{C})$ である。

(ii) この場合、集合族 $\mathcal{C}$ で生成される Dynkin 族は S を要素とするので (i) に帰着する。 □

一般的な**測度の一意性定理**は次のようになる。

7.12 定理. $\mathcal{C}$ を S の部分集合からなる π システム、μ, ν を $(S, \sigma_S(\mathcal{C}))$ 上の測度とする。
(i) $\mu(S) = \nu(S) < +\infty$ かつ $\mu(C) = \nu(C)\ \forall C \in \mathcal{C}$ であるなら測度 μ, ν は一致する。
(ii) $\mathcal{C}$ の元からなる列 C_n であって条件 $\bigcup_{n=1}^{\infty} C_n = S,\ \mu(C_n) < +\infty\ \forall n \in \mathbb{N}$ を満たすものが存在しかつ $\mu(C) = \nu(C)\ \forall C \in \mathcal{C}$ であるなら測度 μ, ν は一致する。

証明. (i) 演習問題 7.3 で確認したように集合族 $\{A \in \sigma_S(\mathcal{C}) : \mu(A) = \nu(A)\}$ は Dynkin 族である。仮定によりそれは $\mathcal{C} \cup \{S\}$ を包含する。定理 7.11 によれば $\mathcal{C} \cup \{S\}$ で生成される Dynkin 族は $\sigma(\mathcal{C})$ に一致するので

$$\sigma_S(\mathcal{C}) \subset \{A \in \sigma_S(\mathcal{C}) : \mu(A) = \nu(A)\}$$

すなわち $A \in \sigma_S(\mathcal{C}) \Rightarrow \mu(A) = \nu(A)$ が成り立つ。

(ii) (i) と違うところは $\{A \in \sigma_S(\mathcal{C}) : \mu(A \cap C_n) = \nu(A \cap C_n)\ \forall n \in \mathbb{N}\}$ という集合族を導入する点にある。後は同様の議論により次が得られる。

$$A \in \sigma_S(\mathcal{C}) \Rightarrow \mu(A \cap C_n) = \nu(A \cap C_n)\ \forall n \in \mathbb{N}.$$

さて $A = \bigcup_{n=1}^{\infty} (A \setminus \bigcup_{k=1}^{n-1} C_k) \cap C_n$ であり、右辺は disjoint union である。よって

$$\mu(A) = \sum_{n=1}^{\infty} \mu((A \setminus \bigcup_{k=1}^{n-1} C_k) \cap C_n) = \sum_{n=1}^{\infty} \nu((A \setminus \bigcup_{k=1}^{n-1} C_k) \cap C_n) = \nu(A)$$

となり示すべき等式が導かれた。 □

7.13 演習問題. $(S, \mathcal{B}, \mu)$ を測度空間、$f, g : S \to \overline{\mathbb{R}}$ を μ 可積分な $\mathcal{B}$ 可測関数とする。
(i) 集合族 $\{A \in \mathcal{B} : \int_A f\mu = \int_A g\mu\}$ は Dynkin 族であることを示せ。
(ii) $\int_A f\mu \le \int_A g\mu\ \forall A \in \mathcal{B}$ であるなら $f \le g$ μ-a.e. であることを示せ。

上の結果を利用すると定理 7.12(i) と同様にして次の定理を示すことができる。

7.14 定理. $(S, \mathcal{B}, \mu)$ を測度空間、$f, g : S \to \overline{\mathbb{R}}$ を μ 可積分な $\mathcal{B}$ 可測関数とする。このとき $\int_S f\mu = \int_S g\mu$ であり、S の部分集合からなる π システム $\mathcal{C}$ が存在して $\sigma(\mathcal{C}) = \mathcal{B}$ かつ $\int_A f\mu = \int_A g\mu\ \forall A \in \mathcal{C}$ であるなら $f = g$ μ-a.e. である。

7.15 演習問題. $f, g : \mathbb{R} \to \mathbb{C}$ を Borel($\mathbb{R}$) 可測関数で局所可積分なもの、F, G をそれぞれ f, g の不定積分とする。このとき

$$F(b) - F(a) = G(b) - G(a)\ \forall a \in \mathbb{R}, \forall b \in \mathbb{R}$$

であるなら $f = g$ λ-a.e. であることを示せ。ヒント 定理 7.12(ii) を参照せよ。

定理 7.12(ii) を適用できる設定として好まれる条件は次のものである。

7.16 定義. $\mathbb{R}^d$ 上の *Radon* **測度**とは $(\mathbb{R}^d, \mathrm{Borel}(\mathbb{R}^d))$ 上の測度 μ で任意の有界な d 次元区間 J に対して $\mu(J) < +\infty$ を満たすものをいう。

7.17 補題. $\mathbb{R}^d$ 上の Radon 測度 μ, ν が与えられたとする。任意の有界な d 次元開区間 U に対して $\mu(U) = \nu(U)$ が成り立つなら測度 μ, ν は一致する。

証明. 有界な d 次元開区間全体に空集合 $\emptyset$ を付加した集合族は π システムである。またそれが生成する $\mathbb{R}^d$ 上の σ 加法族が Borel($\mathbb{R}^d$) と等しいことは補題 5.4 で述べられている。また

$$\bigcup_{n=1}^{\infty} (-n, n) \times \cdots \times (-n, n) = \mathbb{R}^d$$

が成り立つ。よって定理 7.12(ii) を適用して $\mu = \nu$ が得られる。 □

7.18 定義. 連続関数 $f : \mathbb{R}^d \to \mathbb{C}$ の**台**(support) とは $\{x \in \mathbb{R}^d : f(x) \neq 0\}$ の閉包をいう。

7.19 演習問題. 台が有界な連続関数 $f : \mathbb{R}^d \to \mathbb{C}$ が与えられたとする。関数 f は有界かつ一様連続であることを示せ。また任意の Radon 測度に対して f は可積分であることを示せ。

記号

$\operatorname{supp} f :=$ 関数 f の台

7.20 演習問題. (i) $0 < C := \int_{(0,1)} \exp\{-1/x(1-x)\}\,\lambda(dx) < +\infty$ であることを示せ。
(ii) 次で定義される関数 $\mathbb{R} \to \mathbb{R}$ は C^∞ 級であることを示せ。

$$x \mapsto \begin{cases} \exp\{-1/x(1-x)\}/C & 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

(iii) 関数 $\psi : \mathbb{R} \to \mathbb{R}$ を (ii) で導入した関数の原始関数であって $\psi(0) = 0$ を満たすものとする。与えられた $a < b$ に対して関数列 $\phi_n(x) := \psi(n(x-a))\,\psi(n(b-x))$ は

$$\phi_n \text{は}\ C^\infty\ \text{級}, \operatorname{supp}\phi_n = [a, b]\,\forall n \in \mathbb{N},\ a + 1/n \le x \le b - 1/n \Rightarrow \phi_n(x) = 1$$
$$0 \le \phi_n(x) \le \phi_{n+1}(x) \le 1\ \forall x \in \mathbb{R}\,\forall n \in \mathbb{N}, \sup_{n \in \mathbb{N}} \phi_n(x) = 1_{(a,b)}(x)\ \forall x \in \mathbb{R}$$

を満たすことを確認せよ。

記号

$C_0(\mathbb{R}^d)$: 台が有界な連続関数 $\mathbb{R}^d \to \mathbb{R}$ 全体の集合; $C_0^r(\mathbb{R}^d) := \{f \in C_0(\mathbb{R}^d) :\ C^r$ 級$\}$

このノートでは $C_0(\mathbb{R}^d)$ の各元は実数値であることに注意。

記号

$f_1, f_2, \ldots, f_d : \mathbb{R} \to \mathbb{C}$ に対して次の関数を $f_1 \otimes f_2 \otimes \cdots \otimes f_d$ と表記する。

$$\mathbb{R}^d \to \mathbb{C}, (x_1, x_2, \ldots, x_d) \mapsto f_1(x_1) f_2(x_2) \ldots f_d(x_d)$$

7.21 補題. U を有界な d 次元開区間とする。このとき $C_0^\infty(\mathbb{R}^d)$ の列 f_n であって

$$0 \leq f_n(x) \leq 1_U(x)\ \forall n \in \mathbb{N}\, \forall x \in \mathbb{R}^d$$

を満たしかつ 1_U に各点収束するものが存在する。

証明. 煩雑化を防ぐため $d = 2$ として議論を進める。$a_1 < b_1,\ a_2 < b_2$ なる $a_1, b_1, a_2, b_2 \in \mathbb{R}$ によって U は $(a_1, b_1) \times (a_2, b_2)$ と表現される。関数列 $\phi_n^1,\ \phi_n^2$ を次で定義する。

$$\phi_n^1(x) := \psi(n(x - a_1))\, \psi(n(b_1 - x)),\ \ \phi_n^2(x) := \psi(n(x - a_2))\, \psi(n(b_2 - x)).$$

ここで ψ は演習問題 7.20(iii) で登場した関数である。そこで確認したように $\phi_n^1, \phi_n^2 \in C_0^\infty(\mathbb{R})$ であり、従って $\phi_n^1 \otimes \phi_n^2 \in C_0^\infty(\mathbb{R}^2)$ である。関数列 $\phi_n^1 \otimes \phi_n^2$ が求めるものであることは演習問題 7.20(iii) で調べたところから直ちに分かる。 □

測度の一意性定理には次のようなタイプのものもある。このような場合 $C_0^\infty(\mathbb{R})$ を**試験関数**(test function) の空間というが、問題に応じて適切なものを選ぶ必要がある。

7.22 定理. $\mathbb{R}^d$ 上の Radon 測度 μ, ν が一致するための必要十分条件は

$$\int_{\mathbb{R}^d} f_1 \otimes f_2 \otimes \cdots \otimes f_d\, \mu = \int_{\mathbb{R}^d} f_1 \otimes f_2 \otimes \cdots \otimes f_d\, \nu\ \forall f_1, f_2, \ldots, f_d \in C_0^\infty(\mathbb{R})$$

証明. 十分性を示せばよい。次を仮定する。

$$\int_{\mathbb{R}^d} f_1 \otimes f_2 \otimes \cdots \otimes f_d\, \mu = \int_{\mathbb{R}^d} f_1 \otimes f_2 \otimes \cdots \otimes f_d\, \nu\ \ \forall f_1, f_2, \ldots, f_d \in C_0^\infty(\mathbb{R})$$

U を有界な d 次元開区間とする。補題 7.21 により $C_0^\infty(\mathbb{R}^d)$ の列 f_n であって $0 \leq f_n \leq 1_U$ を満たしかつ 1_U に各点収束するものが存在する。しかも補題 7.21 の証明においては各 f_n は $C_0^\infty(\mathbb{R})$ の要素 $\phi^1, \phi^2, \ldots, \phi^d$ によって $\phi^1 \otimes \phi^2 \otimes \cdots \otimes \phi^d$ という表現を持っている。従って

$$\int_{\mathbb{R}^d} f_n\, \mu = \int_{\mathbb{R}^d} f_n\, \nu\ \forall n \in \mathbb{N}$$

が成り立つものとできる。他方、可積分関数 1_U を優関数として Lebesgue の収束定理を適用すれば、左辺は $\int_{\mathbb{R}^d} 1_U\, \mu$ に右辺は $\int_{\mathbb{R}^d} 1_U\, \nu$ に収束することが分かる。よって

$$\text{任意の有界な } d \text{ 次元開区間 } U \text{ に対して } \mu(U)=\nu(U) \text{ が成り立つ。}$$

補題 7.17 を適用して $\mu=\nu$ が得られる。 □

7.23 注意. 上の証明から明らかなように次も $\mu=\nu$ であるための必要十分条件である。

$$\int_{\mathbb{R}^d} f\,\mu = \int_{\mathbb{R}^d} f\,\nu \ \ \forall f \in C_0^\infty(\mathbb{R}^d)$$

次の定理の証明は読者に委ねる。

7.24 定理. μ を $(\mathbb{R}^d, \mathrm{Borel}(\mathbb{R}^d))$ 上の測度、$f, g : \mathbb{R}^d \to \overline{\mathbb{R}}$ を μ 可積分な $\mathrm{Borel}(\mathbb{R}^d)$ 可測関数とする。このとき $\int_{\mathbb{R}^d} f\phi\mu = \int_{\mathbb{R}^d} g\phi\mu \ \forall \phi \in C_0^\infty(\mathbb{R}^d)$ であるなら $f=g$ μ-a.e. である。

定理 7.22 の応用例をあげておく。

7.25 例. $a, b \in \overline{\mathbb{R}}$ かつ $a<b$ とする。また $v : I \to \mathbb{R}$ を開区間上の連続微分可能な関数で $a < v(x) < b \ \forall x \in I$ であり次のいずれかを満たすものとする。

- $v'(x) \geq 0 \ \forall x \in I$, $\inf_{x\in I} v(x) = a$ かつ $\sup_{x\in I} v(x) = b$
- $v'(x) \leq 0 \ \forall x \in I$, $\inf_{x\in I} v(x) = b$ かつ $\sup_{x\in I} v(x) = a$

このとき非負 Borel 可測関数 $f : (a,b) \to \overline{\mathbb{R}}$ に対して次が成り立つ。

$$\int_{(a,b)} f\,\lambda = \int_I f(v(x))|v'(x)|\ \lambda(dx).$$

証明. $(\mathbb{R}, \mathrm{Borel}(\mathbb{R}))$ 上の集合関数 μ と $(I, \mathrm{Borel}(I))$ 上の集合関数 ν を以下で導入する。

$$\mu(A) = \lambda(A \cap (a,b)) \ A \in \mathrm{Borel}(\mathbb{R}),\ \ \nu(B) = \int_B |v'(x)|\ \lambda(dx)\ B \in \mathrm{Borel}(I).$$

$\nu(I) < +\infty$ であるかはともかく、それ以外については定理 3.19 の主張が ν に当てはまることに注意する。特に ν は測度である。また示すべきは

$$(\star) \qquad \int_{\mathbb{R}} f\,\mu = \int_I f(v(x))\ \nu(dx) = \int_{\mathbb{R}} f\ v_*\nu$$

が非負 Borel 可測関数 $f : \mathbb{R} \to \overline{\mathbb{R}}$ に対して成り立つことと同値である。ここで 2 番目の等号は定理 6.20 による。従って $(\mathbb{R}, \mathrm{Borel}(\mathbb{R}))$ 上の測度である μ と $v_*\nu$ が一致すればよい。補題 3.33 とその直後の注意によれば f が非負値連続関数なら $(\star)$ は成り立つ。$n \in \mathbb{N}$ に対して

$$0 \leq h(x) \leq 1 \ \forall x \in \mathbb{R}, h(x) = 1 \ \forall x \in (-n, n) \text{ かつ } h(x) = 0 \ \forall x \notin (-n-1, n+1)$$

を満たす非負値連続関数 h が存在する。従って

$$v_*\nu((-n,n)) \leq \int_{\mathbb{R}} h\ v_*\nu = \int_{\mathbb{R}} h\ \mu \leq 2(n+1) < +\infty$$

上のことから $v_*\nu$ は Radon 測度である。さらに $f = \max\{f,0\} - \max\{-f,0\}$ と分解して、非負値連続関数に対して成り立つ $(\star)$ を適用することにより

$$\begin{aligned}\int_{\mathbb{R}} f\ \mu &= \int_{\mathbb{R}} \max\{f,0\}\ \mu - \int_{\mathbb{R}} \max\{-f,0\}\ \mu \\ &= \int_{\mathbb{R}} \max\{f,0\}\ v_*\nu - \int_{\mathbb{R}} \max\{-f,0\}\ v_*\nu = \int_{\mathbb{R}} f\ v_*\nu\end{aligned}$$

であることが分かる。定理 7.22 より結論 $\mu = v_*\nu$ を得る。 □

次はいわゆる**変数変換公式**(change of variable formula) である。

7.26 系. $v : I \to \mathbb{R}$ を開区間上の連続微分可能かつ単調な関数であり、その導関数 v' は

$$\lambda(\{x \in \mathbb{R} : v'(x) = 0\}) = 0$$

を満たすとする。このとき v は狭義単調でありかつ $J := v(I)$ は開区間である。任意の非負値 Borel 可測関数 $f : I \to \overline{\mathbb{R}}$ と $g : J \to \overline{\mathbb{R}}$ に対して次が成り立つ。

$$\int_J f(v^{-1}(y))g(y)\ \lambda(dy) = \int_I f(x)g(v(x))\ |v'(x)|\ \lambda(dx)$$

証明. 非減少であるとして証明する。与えられた条件より v' は連続かつ非負値である。さて $a, b \in I$, $a < b$ とする。このとき $v' > 0$ λ-a.e. であるから系 3.28(ii) と定理 6.31(ii) より

$$v(b) - v(a) = \int_{(a,b)} v'\lambda > 0$$

よって v は狭義増加でありかつ v は I から $J := v(I)$ への位相同型である。例 7.25 における f を $J \to \overline{\mathbb{R}}$, $y \mapsto f(v^{-1}(y))g(y)$ に取り替えて結論を得る。 □

7.27 例. $\rho : \mathbb{R} \to \overline{\mathbb{R}}$ を非負値 Borel 可測関数とする。
(i) $p > 0$ と任意の非負値 Borel 可測関数 $f : [0, +\infty) \to \overline{\mathbb{R}}$ に対して次が成り立つ。

$$\int_{\mathbb{R}} f(|y|^p)\rho(y)\ \lambda(dy) = \int_{(0,+\infty)} f(x)\frac{\rho(-\sqrt[p]{x}) + \rho(\sqrt[p]{x})}{p}x^{1/p-1}\ \lambda(dx)$$

(ii) 任意の非負値 Borel 可測関数 $f : \mathbb{R} \to \overline{\mathbb{R}}$ に対して次が成り立つ。

$$\int_{\mathbb{R}\setminus\{0\}} f(\log y^2)\rho(y)\ \lambda(dy) = \int_{\mathbb{R}} f(x)\frac{\rho(-e^{x/2}) + \rho(e^{x/2})}{2e^{-x/2}}\ \lambda(dx)$$

(iii) 任意の非負値 Borel 可測関数 $f : (0, 1] \to \overline{\mathbb{R}}$ に対して次が成り立つ。

$$\int_{\mathbb{R}} f(1/(1+y^2))\rho(y)\ \lambda(dy) = \int_{(0,1)} f(x)\frac{\rho(-\sqrt{(1-x)/x}) + \rho(\sqrt{(1-x)/x})}{2x\sqrt{(1-x)x}}\ \lambda(dx)$$

証明. (i) $\mathbb{R}$ を $(-\infty,0)$, $\{0\}$ と $(0,+\infty)$ に分割することにより

$$\int_{\mathbb{R}} f(|y|^p)\rho(y)\ \lambda(dy) = \int_{(-\infty,0)} f(|y|^p)\rho(y)\ \lambda(dy) + \int_{(0,+\infty)} f(|y|^p)\rho(y)\ \lambda(dy)$$

ここで狭義減少な C^1 級関数 $v:(0,+\infty)\to\mathbb{R}$, $x\mapsto -\sqrt[p]{x}$ を導入する。このとき

$$v((0,+\infty)) = (-\infty,0), v^{-1}(y) = |y|^p, |v'(x)| = \frac{\sqrt[p]{x}}{px}$$

である。よって系 7.26 を適用して次を得る。

$$\int_{(-\infty,0)} f(|y|^p)\rho(y)\ \lambda(dy) = \int_{(0,+\infty)} f(x)\frac{\rho(-\sqrt[p]{x})\sqrt[p]{x}}{px}\ \lambda(dx)$$

$\int_{(0,+\infty)} f(|y|^p)\rho(y)\ \lambda(dy)$ に関する議論も同様である。 □

例 7.27(i) は例 6.21(i) の一般化である。実際、以下が成り立つ。

$$\rho(y) = \frac{1}{\sqrt{2\pi}}\exp\{-y^2/2\} \text{ のとき } \frac{\rho(-\sqrt{x})+\rho(\sqrt{x})}{2\sqrt{x}} = \frac{1}{\sqrt{2\pi}}x^{-1/2}\exp\{-x/2\}$$

ついでながら ρ が中心 0 半値幅 1 の Cauchy 分布の密度関数であるときは

$$\sum_{\varepsilon=\pm1}\frac{\rho(\varepsilon e^{x/2})}{2e^{-x/2}} = \frac{1}{2\pi}\frac{1}{\cosh(x/2)},\ \sum_{\varepsilon=\pm1}\frac{\rho(\varepsilon\sqrt{(1-x)/x})}{2x\sqrt{(1-x)x}} = \frac{1}{\pi}\frac{1}{\sqrt{(1-x)x}}$$

8　測度の直積と確率変数の独立性

n 次元確率変数が与えられたとしよう。一般にはその周辺分布だけでは自身の分布を知ることはできない。しかしながら空間 $\mathbb{R}^n$ の直積構造が結合分布にも反映する特別な場合があり、それが確率変数系の独立性として定式化される。ここでは、まず測度の直積についてしっかりと確認してから独立性を論じていくことにする。

記号

集合族 $\mathcal{A}_1$ と $\mathcal{A}_2$ に対して $\mathcal{A}_1\times\mathcal{A}_2 := \{A_1\times A_2\,;A_1\in\mathcal{A}_1, A_2\in\mathcal{A}_2\}$ と書く。

8.1 定義. S_1 上の σ-加法族 $\mathcal{B}_1$ と S_2 上の σ-加法族 $\mathcal{B}_2$ に対して $S_1\times S_2$ 上の σ-加法族 $\sigma(\mathcal{B}_1\times\mathcal{B}_2)$ を**直積 σ-加法族**(product σ-field) と呼び記号 $\mathcal{B}_1\otimes\mathcal{B}_2$ で表す。

復習

$(S,\mathcal{B})$ 上の測度 μ が σ-**有限**(σ-finite) であるとは列 $B_n\in\mathcal{B}$ $n\in\mathbb{N}$ で $\bigcup_{n=1}^\infty B_n = S$ かつ $\mu(B_n) < +\infty$ $\forall n\in\mathbb{N}$ であるものが存在することをいう。

8.2 演習問題. $\mathbb{R}^d$ 上の Radon 測度は σ 有限であることを示せ。

直積測度に関する前提

σ 有限な測度空間 $(S_1, \mathcal{B}_1, \mu_1)$, $(S_2, \mathcal{B}_2, \mu_2)$

復習

次を満たす $(S_1 \times S_2, \mathcal{B}_1 \otimes \mathcal{B}_2)$ 上の測度 $\mu_1 \otimes \mu_2$ が唯一存在する。

$$(\mu_1 \otimes \mu_2)(A \times B) = \mu_1(A)\mu_2(B) \ \forall A \in \mathcal{B}_1, \forall B \in \mathcal{B}_2$$

これを**直積測度**(product measure) と呼ぶ。三つ組み $(S_1 \times S_2, \mathcal{B}_1 \otimes \mathcal{B}_2, \mu_1 \otimes \mu_2)$ を**直積測度空間**(product measure space) という。

8.3 定義. 実確率変数系 X_1, X_2, ..., X_n が**独立**(independent) であるとはそれらの結合分布が各分布の直積に等しいことをいう。すなわち次が成り立つことである。

$$\mathcal{L}((X_1, X_2, \ldots, X_n), \cdot) = \mathcal{L}(X_1, \cdot) \otimes \mathcal{L}(X_2, \cdot) \otimes \cdots \otimes \mathcal{L}(X_n, \cdot)$$

8.4 定理. 実確率変数系 X_1, X_2, ..., X_n が独立であるのは次の各々と同値である。

(i) $$P(X_1 \in A_1, X_2 \in A_2, \ldots, X_n \in A_n) = \prod_{k=1}^{n} P(X_k \in A_k) \ \forall A_1, A_2, \ldots, A_n \in \mathrm{Borel}(\mathbb{R}),$$

(ii) $$P(X_1 > a_1, X_2 > a_2, \ldots, X_n > a_n) = \prod_{k=1}^{n} P(X_k > a_k) \ \forall a_1, a_2, \ldots, a_n \in \mathbb{R}.$$

証明. 独立であれば、条件 (i) が成り立つ。他方条件 (i) が成り立てば、条件 (ii) が成り立つ。したがって条件 (ii) から独立性を導けばよい。スペースの節約のため X_1, X_2, ..., X_n の結合分布を μ また各分布の直積を ν と書くと条件 (ii) は次を意味する。

任意の $A = (a_1, +\infty) \times (a_2, +\infty) \times \cdots \times (a_n, +\infty)$ に対して $\mu(A) = \nu(A)$ が成り立つ。

よって定理 7.1 を適用して結論を得る。 □

8.5 系. 実確率変数系 X_1, X_2, ..., X_n は独立であるとする。
(i) $m \le n$ に対して実確率変数系 X_1, X_2, ..., X_m も独立である。
(ii) Borel 可測関数 $f_1, f_2, \ldots, f_n : \mathbb{R} \to \mathbb{R}$ に対して実確率変数系 $f_1(X_1)$, $f_2(X_2)$, ..., $f_n(X_n)$ も独立である。

8.6 演習問題. 系 8.5 を示せ。

8.7 例. Lebesgue モデルを確率空間とする。このとき第 1 節で述べた実確率変数系 ξ_k, $k \in \mathbb{N}$ は独立である。これは例 6.27 と演習問題 8.8 で述べることより示せる。

8.8 演習問題. $0 < p < 1$ とする。実確率変数系 X_1, X_2, ..., X_n が

$$P(X_i = 1 \, \forall i \in I) = p^{\# I} \ \forall I \subset \{1, 2, \ldots, n\}, \ P(X_i = -1) = 1 - p \ \forall i = 1, 2, \ldots, n.$$

を満たすなら独立であることを示せ。ヒント $a_1 < 1$, $a_2 < 1$, ..., $a_n < 1$ とする。このとき $J := \{i \in \mathbb{N} : 1 \le i \le n, a_i \ge -1\}$ とおくと次が成り立つことを示せ。

$$P(X_1 > a_1, X_2 > a_2, \ldots, X_n > a_n) = P(X_i = 1\,\forall i \in J)$$

直積測度の構造をその断面から観察するのが *Fubini* **の定理**であり非常に重要な役割を果たす。これを復習しておこう。設定は次の通りである。

Fubini の定理での前提

σ 有限な測度空間 $(S_1, \mathcal{B}_1, \mu_1)$, $(S_2, \mathcal{B}_2, \mu_2)$

8.9 補題. 関数 $f : S_1 \times S_2 \to \overline{\mathbb{R}}$ は $\mathcal{B}_1 \otimes \mathcal{B}_2$-可測とする。
(i) 各 $x \in S_1$ に対して関数 $S_2 \to \overline{\mathbb{R}}$, $y \mapsto f(x, y)$ は $\mathcal{B}_2$-可測である。
(ii) 各 $y \in S_2$ に対して関数 $S_1 \to \overline{\mathbb{R}}$, $x \mapsto f(x, y)$ は $\mathcal{B}_1$-可測である。

非負値可測関数については *Fubini* **の定理**、むしろ *Tonelli* **の定理**あるいは *Fubini-Tonelli* **の定理**と呼ぶべきであろう、は非常に明解である。その証明の核心は定理 8.10(i) に述べる部分にあり単調族あるいは Dynkin 族の概念が不可欠である。

8.10 定理. 関数 $f : S_1 \times S_2 \to \overline{\mathbb{R}}$ は非負値かつ $\mathcal{B}_1 \otimes \mathcal{B}_2$-可測とする。
(i) 関数 $S_1 \to \overline{\mathbb{R}}$, $x \mapsto \displaystyle\int_{S_2} f(x, y)\,\mu_2(dy)$ は $\mathcal{B}_1$-可測である。
(ii) $\displaystyle\int_{S_1} \Big(\int_{S_2} f(x, y)\,\mu_2(dy) \Big)\,\mu_1(dx) = \int_{S_1 \times S_2} f \;\mu_1 \otimes \mu_2.$

8.11 演習問題. $\mathcal{B}_1 \otimes \mathcal{B}_2$ 可測関数 $f : S_1 \times S_2 \to \overline{\mathbb{R}}$ が $\int_{S_2} |f(x, y)|\,\mu_2(dy) < +\infty\ \forall x \in S_1$ を満たすなら関数 $S_1 \to \overline{\mathbb{R}}$, $x \mapsto \int_{S_2} f(x, y)\,\mu_2(dy)$ は $\mathcal{B}_1$-可測であることを示せ。

次が通常 *Fubini* **の定理**と呼んで引用されるものである。Fubini の定理を適用するには可積分性が重要な前提条件である。それを確かめるのに定理 8.10 の協力を得ることが多い。

8.12 定理. 関数 $f : S_1 \times S_2 \to \overline{\mathbb{R}}$ は $\mathcal{B}_1 \otimes \mathcal{B}_2$-可測かつ $\mu_1 \otimes \mu_2$-可積分とする。
(i) $A := \{x \in S_1 : \int_{S_2} |f(x, y)|\,\mu_2(dy) < +\infty\} \in \mathcal{B}_1$ は μ_1-a.e. 集合であり、かつ $\mathcal{B}_1$ 可測関数 $g : S_1 \to \overline{\mathbb{R}}$ が存在して $g(x) = \int_{S_2} f(x, y)\,\mu_2(dy)$ μ_1-a.e. $x \in A$.
(ii) $\mathcal{B}_1$-可測関数 $g : S_1 \to \overline{\mathbb{R}}$ が $g(x) = \int_{S_2} f(x, y)\,\mu_2(dy)$ μ_1-a.e. $x \in A$ を満たせば、それは μ_1-可積分であり $\displaystyle\int_{S_1 \times S_2} f \;\mu_1 \otimes \mu_2 = \int_{S_1} g\,\mu_1$ が成り立つ。

8.13 注意. 実際の応用例では $\int_{S_2} |f(x, y)|\,\mu_2(dy) < +\infty\ \forall x \in S_1$ も成り立つことが多い。そのようなときは、演習問題 8.11 の結論により、定理 8.12 における補助的な関数 g として関数 $x \mapsto \int_{S_2} f(x, y)\,\mu_2(dy)$ を選ぶことができる。

8.14 定理. 関数 $f : S_1 \to \overline{\mathbb{R}}$ は $\mathcal{B}_1$-可測、関数 $g : S_2 \to \overline{\mathbb{R}}$ は $\mathcal{B}_2$-可測とする。
(i) 関数 $f \otimes g : S_1 \times S_2 \to \overline{\mathbb{R}}$, $(x, y) \mapsto f(x)g(y)$ は $\mathcal{B}_1 \otimes \mathcal{B}_2$-可測である。
(ii) f, g ともに非負値であるなら

$$\int_{S_1 \times S_2} f \otimes g \;\mu_1 \otimes \mu_2 = \int_{S_1} f\,\mu_1 \int_{S_2} g\,\mu_2$$

(iii) f が μ_1 可積分 g が μ_2 可積分なら $f \otimes g$ は $\mu_1 \otimes \mu_2$ 可積分であり上の等式が成立する。

証明. (i) まず関数 $S_1 \times S_2 \to \overline{\mathbb{R}}$, $(x,y) \mapsto f(x)$ の $\mathcal{B}_1 \otimes \mathcal{B}_2$-可測性を確かめる。関数 f の $\mathcal{B}_1$-可測性により $\{x \in S_1 : f(x) < a\} \in \mathcal{B}_1\ \forall a \in \mathbb{R}$ である。よって $\mathcal{B}_1 \otimes \mathcal{B}_2$ の定義により

$$\{(x,y) \in S_1 \times S_2 : f(x) < a\} = \{x \in S_1 : f(x) < a\} \times S_2 \in \mathcal{B}_1 \times \mathcal{B}_2 \subset \mathcal{B}_1 \otimes \mathcal{B}_2\ \forall a \in \mathbb{R}.$$

同様にして関数 $S_1 \times S_2 \to \overline{\mathbb{R}}$, $(x,y) \mapsto g(y)$ の $\mathcal{B}_1 \otimes \mathcal{B}_2$-可測性もわかる。従ってそれらの積関数 $(x,y) \mapsto f(x)g(y)$ は $\mathcal{B}_1 \otimes \mathcal{B}_2$-可測である。

(ii) (i) と定理 8.10 から従う。(iii) (ii) と定理 8.12 から従う。 □

複素数値関数についても定理 8.12 に対応するものがあるが、しかるべき読み替えを行えばよいのでわざわざ書くのは差し控える。

さて $\mathbb{R}^d$ 上で Fubini の定理を適用するには、Borel 集合族が直積 σ 加法族として表現できることを確認しておく必要がある。

8.15 補題. S_1 の部分集合族 $\mathcal{C}_1$ と S_2 の部分集合族 $\mathcal{C}_2$ に対し、S_1 の可算 $\mathcal{C}_1$ 被覆と S_2 の可算 $\mathcal{C}_2$ 被覆が存在するなら $\sigma_{S_1 \times S_2}(\mathcal{C}_1 \times \mathcal{C}_2) = \sigma_{S_1}(\mathcal{C}_1) \otimes \sigma_{S_2}(\mathcal{C}_2)$ が成り立つ。

証明. まず $\mathcal{C}_1 \times \mathcal{C}_2 \subset \sigma(\mathcal{C}_1) \times \sigma(\mathcal{C}_2) \subset \sigma(\mathcal{C}_1) \otimes \sigma(\mathcal{C}_2)$ である。右辺は $\mathcal{C}_1 \times \mathcal{C}_2$ を含む σ-加法族であるがそのようなものの最小が $\sigma(\mathcal{C}_1 \times \mathcal{C}_2)$ なので $\sigma(\mathcal{C}_1 \times \mathcal{C}_2) \subset \sigma(\mathcal{C}_1) \otimes \sigma(\mathcal{C}_2)$ を得る。

次に標準写像 $f : S_1 \times S_2 \to S_1$, $(x,y) \mapsto x$ を導入する。補題 5.6(iii) により集合族

$$\mathcal{B} := \{A \in \mathrm{Sbset}(S_1) : A \times S_2 \in \sigma(\mathcal{C}_1 \times \mathcal{C}_2)\} = \{A \in \mathrm{Sbset}(S_1) : f^{-1}(A) \in \sigma(\mathcal{C}_1 \times \mathcal{C}_2)\}$$

は S_1 上の σ-加法族である。S_2 は $\mathcal{C}_2$ に属する集合の可算合併であるから次が成り立つ。

$$A \times S_2 \in \sigma(\mathcal{C}_1 \times \mathcal{C}_2)\ \forall A \in \mathcal{C}_1$$

従って $\mathcal{B}$ は $\mathcal{C}_1$ を含む S_1 上の σ-加法族である。$\sigma(\mathcal{C}_1)$ の定義により $\sigma(\mathcal{C}_1) \subset \mathcal{B}$ すなわち

$$A \in \sigma(\mathcal{C}_1) \Rightarrow A \times S_2 \in \sigma(\mathcal{C}_1 \times \mathcal{C}_2)$$

を得る。同様に $B \in \sigma(\mathcal{C}_2) \Rightarrow S_1 \times B \in \sigma(\mathcal{C}_1 \times \mathcal{C}_2)$ も成り立つ。以上より

$$A \in \sigma(\mathcal{C}_1), B \in \sigma(\mathcal{C}_2) \Rightarrow A \times B = (A \times S_2) \cap (S_1 \times B) \in \sigma(\mathcal{C}_1 \times \mathcal{C}_2).$$

即ち $\sigma(\mathcal{C}_1 \times \mathcal{C}_2)$ は $\sigma(\mathcal{C}_1) \times \sigma(\mathcal{C}_2)$ を含む $S_1 \times S_2$ 上の σ-加法族である。従って $\sigma(\mathcal{C}_1) \otimes \sigma(\mathcal{C}_2)$ の定義により $\sigma(\mathcal{C}_1) \otimes \sigma(\mathcal{C}_2) \subset \sigma(\mathcal{C}_1 \times \mathcal{C}_2)$ を得る。 □

復習

次の条件を満たす $(\mathbb{R}^d, \mathrm{Borel}(\mathbb{R}^d))$ 上の測度 $\lambda^{(d)}$ がただ一つ存在する。

$$\lambda^{(d)}((a_1,b_1] \times (a_2,b_2] \times \cdots \times (a_d,b_d]) = \textstyle\prod_{i=1}^d (b_i - a_i)\ \forall a, \forall b \in \mathbb{R}^d,\ a_i < b_i$$

これを d **次元** *Lebesgue* **測度**という。1 次元 Lebesgue 測度は λ とかく。

8.16 系. $\mathrm{Borel}(\mathbb{R}^{d_1+d_2}) = \mathrm{Borel}(\mathbb{R}^{d_1}) \otimes \mathrm{Borel}(\mathbb{R}^{d_2})$ かつ $\lambda^{(d_1+d_2)} = \lambda^{(d_1)} \otimes \lambda^{(d_2)}$ である。

証明. 補題 5.4 によれば、d 次元 Borel 集合族は d 次元開区間すべてから生成されているので、補題 8.15 を適用して結論を得る。 □

8.17 例. $p > 0$ と $(\mathbb{R}, \mathrm{Borel}(\mathbb{R}))$ 上の σ 有限な測度 μ に対して次が成り立つ。

$$\int_{(0,\infty)} x^p \,\mu(dx) = \int_{(0,\infty)} px^{p-1}\mu((x,+\infty))\,\lambda(dx)$$

特に μ を指数分布にとると $\int_{(0,\infty)} x^p \,\mu(dx) = \Gamma(p+1)$, $\mu((x,+\infty)) = e^{-x}$ であるから

$$\Gamma(p+1) = p\Gamma(p) \ \textit{gamma}\,\textbf{関数の関数等式}$$

証明. $\mathbb{R}^2$ の部分集合 $A := \{(x,y) : 0 < x < y\}$ は開集合である。非負値 Borel 可測関数 $f : (x,y) \mapsto px^{p-1}1_A(x,y)$ を直積測度 $\lambda \otimes \mu$ で積分する。Fubini-Tonelli の定理（系 8.16 と定理 8.10）を適用して

$$\int_{\mathbb{R}} \Big(\int_{\mathbb{R}} px^{p-1}1_A(x,y)\,\lambda(dx) \Big) \mu(dy) = \int_{\mathbb{R}^2} f \ \lambda \otimes \mu = \int_{\mathbb{R}} \Big(\int_{\mathbb{R}} px^{p-1}1_A(x,y)\,\mu(dy) \Big) \lambda(dx)$$

を得る。左辺および右辺において累次積分は次のように評価される。

$$\int_{\mathbb{R}} px^{p-1}1_A(x,y)\,\lambda(dx) = \begin{cases} 0 & y \le 0 \\ y^p & y > 0 \end{cases}, \quad \int_{\mathbb{R}} 1_A(x,y)\,\mu(dy) = \begin{cases} 0 & x \le 0 \\ \mu((x,+\infty)) & x > 0 \end{cases}$$

もちろん系 3.28(ii) を適用して左の等号を導いている。 □

8.18 演習問題. $p > 0$ と $(\mathbb{R}, \mathrm{Borel}(\mathbb{R}))$ 上の σ 有限な測度 μ に対して次が成り立つことを示せ。

$$\int_{(0,\infty)} \frac{1}{x^p}\,\mu(dx) = \int_{(0,\infty)} \frac{p\mu((0,x))}{x^{p+1}}\,\lambda(dx).$$

8.19 例. $\Big(\int_{\mathbb{R}} e^{-x^2}\,\lambda(dx) \Big)^2 = \int_{\mathbb{R}^2} e^{-x^2-y^2}\,\lambda^{(2)}(dxdy) = \int_{\mathbb{R}} \frac{1}{1+x^2}\,\lambda(dx).$

証明. 左の等号は Fubini-Tonelli の定理（系 8.16 と定理 8.14(ii)）から従う。変数変換公式（補題 3.21）によれば、$y \neq 0$ に対して

$$\int_{\mathbb{R}} e^{-x^2y^2-y^2}\,\lambda(dx) = \frac{1}{|y|}\int_{\mathbb{R}} e^{-x^2-y^2}\,\lambda(dx)$$

が成り立つ。$\lambda(\{0\}) = 0$ であるから、定理 6.31(iii) により

$$\int_{\mathbb{R}} \Big(\int_{\mathbb{R}} e^{-x^2-y^2}\,\lambda(dx) \Big)\,\lambda(dy) = \int_{\mathbb{R}} \Big(\int_{\mathbb{R}} |y|e^{-x^2y^2-y^2}\,\lambda(dx) \Big)\,\lambda(dy)$$

Fubini-Tonelli の定理（定理 8.10）を適用する。右辺は次に等しい。

$$\int_{\mathbb{R}^2} |y|e^{-x^2y^2-y^2}\,\lambda^{(2)}(dxdy) = \int_{\mathbb{R}} \Big(\int_{\mathbb{R}} |y|e^{-x^2y^2-y^2}\,\lambda(dy) \Big)\,\lambda(dx)$$

$\mathbb{R}$ を $(-\infty,0)$, $\{0\}$ と $(0,+\infty)$ に分割し、変数変換公式（補題 3.21）を $(-\infty,0)$ 上の積分に適用することにより次のように評価される。

$$\int_{\mathbb{R}} |y| e^{-x^2y^2-y^2}\,\lambda(dy) = 2\int_{(0,+\infty)} y e^{-(1+x^2)y^2}\lambda(dy) = \left.\frac{-e^{-(1+x^2)y^2}}{(1+x^2)}\right|_0^{+\infty} = \frac{1}{1+x^2}$$

2 番目の等号は系 3.28(ii) による。従って Fubini-Tonelli の定理（定理 8.10）により

$$\int_{\mathbb{R}^2} e^{-x^2-y^2}\,\lambda^{(2)}(dxdy) = \int_{\mathbb{R}}\Big(\int_{\mathbb{R}} e^{-x^2-y^2}\,\lambda(dx)\Big)\,\lambda(dy) = \int_{\mathbb{R}} \frac{1}{1+x^2}\,\lambda(dx)$$

を得る。右辺は Cauchy 分布の密度関数と関わっている。その背景を知るには例 8.32 を参照するとよいだろう。ついでながら $\displaystyle\int_{\mathbb{R}} e^{-x^2/2}\,\lambda(dx) = \sqrt{2}\int_{\mathbb{R}} e^{-x^2}\,\lambda(dx) = \sqrt{2\pi}$ である。 □

8.20 例. $t\in\mathbb{R}_{>0}$, $\xi\in\mathbb{R}$ に対して次が成り立つ。

$$\frac{t}{\pi}\int_{\mathbb{R}} \frac{e^{\sqrt{-1}\xi x}}{t^2+x^2}\,\lambda(dx) = e^{-t|\xi|}$$

Cauchy **核の** *Fourier* **変換**(Fourier transform of Cauchy kernel)

証明. ここでは複素解析を援用しない方法を紹介する。$\varepsilon\in\mathbb{R}_{>0}$ とする。関数

$$(x,y)\mapsto |e^{-t|y+\xi|}e^{-\sqrt{-1}yx}e^{-\varepsilon|x|}| = e^{-t|y+\xi|}e^{-\varepsilon|x|}$$

は Fubini-Tonelli の定理（系 8.16 と定理 8.14）により 2 次元 Lebesgue 測度 $\lambda^{(2)}$ に関して可積分である。また例 4.4 で述べた最後の公式より

$$\int_{\mathbb{R}} e^{-\sqrt{-1}yx}e^{-\varepsilon|x|}\,\lambda(dx) = \frac{2\varepsilon}{\varepsilon^2+y^2}\ \forall y\in\mathbb{R}$$

である。従って Fubini の定理（定理 8.12(ii)）を適用して

$$\begin{aligned}&\int_{\mathbb{R}}\Big(\int_{\mathbb{R}} e^{-t|y+\xi|}e^{-\sqrt{-1}yx}\,\lambda(dy)\Big)e^{-\varepsilon|x|}\lambda(dx) = \int_{\mathbb{R}^2} e^{-t|y+\xi|}e^{-\sqrt{-1}yx}e^{-\varepsilon|x|}\,\lambda^{(2)}(dxdy)\\ =&\int_{\mathbb{R}} e^{-t|y+\xi|}\Big(\int_{\mathbb{R}} e^{-\sqrt{-1}yx}e^{-\varepsilon|x|}\,\lambda(dx)\Big)\lambda(dy) = 2\varepsilon\int_{\mathbb{R}} e^{-t|y+\xi|}\frac{1}{\varepsilon^2+y^2}\,\lambda(dy)\end{aligned}$$

再び例 4.4 で述べた最後の公式より

$$\frac{t}{\pi}\frac{e^{\sqrt{-1}\xi x}}{t^2+x^2} = \frac{1}{2\pi}\int_{\mathbb{R}} e^{-t|y+\xi|}e^{-\sqrt{-1}yx}\,\lambda(dy)\ \forall x\in\mathbb{R}.$$

以上より次の公式が得られた。

$$\frac{t}{\pi}\int_{\mathbb{R}} \frac{e^{\sqrt{-1}\xi x}}{t^2+x^2}e^{-\varepsilon|x|}\,\lambda(dx) = \frac{\varepsilon}{\pi}\int_{\mathbb{R}} \frac{e^{-t|y+\xi|}}{\varepsilon^2+y^2}\,\lambda(dy)\ \forall\varepsilon\in\mathbb{R}_{>0}.$$

補題 4.7(ii) によれば $\varepsilon \to 0$ のとき右辺は有界連続関数 $y \mapsto e^{-t|y+\xi|}$ の 0 における値 $e^{-t|\xi|}$ に収束する。他方、左辺の被積分関数は次のように評価される。

$$\Big|\frac{e^{\sqrt{-1}\xi x}}{t^2+x^2}e^{-\varepsilon|x|}\Big| \le \frac{1}{t^2+x^2}\ \forall x\in\mathbb{R}\ \forall\varepsilon\in\mathbb{R}_{>0}.$$

Lebesgue の収束定理により $\varepsilon \to 0$ のとき左辺は目標とする積分に収束する。人為的に導入した因子 $e^{-\varepsilon|x|}$ は最終結果のどこにもその痕跡がない。このようなものを**収束因子**といい、解析学においては有用なトリックの一つである。 □

8.21 例. $\alpha > 0$, $x \in \mathbb{R}$ に対して次が成り立つ。

$$\int_{(0,+\infty)} \frac{1}{\sqrt{\pi u}}\exp\{-\frac{x^2}{4u}-\alpha u\}\ \lambda(du) = \frac{\exp\{-\sqrt{\alpha}|x|\}}{\sqrt{\alpha}}.$$

熱核の *Laplace* 変換(Laplace transform of heat kernel)

証明. $u > 0$ とする。平均 0 分散 $1/2u$ の正規分布に補題 4.9(ii) を適用して次を得る。

$$\exp\{-\frac{x^2}{4u}\} = \int_{\mathbb{R}} e^{\sqrt{-1}xy}\frac{\sqrt{u}}{\sqrt{\pi}}e^{-uy^2}\lambda(dy)$$

他方、関数 $(u,y)\mapsto e^{\sqrt{-1}xy}1_{(0,+\infty)}(u)e^{-uy^2-\alpha u}$ は $\lambda^{(2)}$ 可積分である。実際、その絶対値の積分は Fubini-Tonelli の定理（定理 8.10）より

$$\int_{\mathbb{R}^2} 1_{(0,+\infty)}(u)e^{-uy^2-\alpha u}\lambda^{(2)}(du,dy) = \int_{\mathbb{R}}\Big(\int_{(0,+\infty)} e^{-(y^2+\alpha)u}\lambda(du)\Big)\lambda(dy) = \int_{\mathbb{R}}\frac{1}{\alpha+y^2}\lambda(dy)$$

となり有限である。従って Fubini の定理（定理 8.12(ii)）を適用して

$$\begin{aligned}&\int_{(0,+\infty)} \frac{1}{\sqrt{\pi u}}\exp\{-\frac{x^2}{4u}-\alpha u\}\ \lambda(du) = \int_{(0,+\infty)}\frac{1}{\pi}\Big(\int_{\mathbb{R}} e^{\sqrt{-1}xy}e^{-uy^2-\alpha u}\ \lambda(dy)\Big)\lambda(du)\\ &= \frac{1}{\pi}\int_{\mathbb{R}^2} 1_{(0,+\infty)}(u)e^{\sqrt{-1}xy}e^{-uy^2-\alpha u}\lambda^{(2)}(du,dy) = \frac{1}{\pi}\int_{\mathbb{R}} e^{\sqrt{-1}xy}\Big(\int_{(0,+\infty)} e^{-(y^2+\alpha)u}\lambda(du)\Big)\lambda(dy)\end{aligned}$$

が成り立つことが分かる。右辺は $y\mapsto e^{\sqrt{-1}xy}/(y^2+\alpha)$ の積分を π で割ったものである。例 8.20 を適用して積分を評価すると $e^{-\sqrt{\alpha}|x|}/\sqrt{\alpha}$ が得られる。 □

8.22 定理. $a,b\in\mathbb{R}$, $a<b$, $f,g:\mathbb{R}\to\mathbb{C}$ を Borel($\mathbb{R}$) 可測かつ局所可積分な関数とする。f, g それぞれの不定積分を F, G とするとき（演習問題 4.26(i) 参照）次が成り立つ。

$$\int_{(a,b)} Fg\,\lambda = F(b)G(b) - F(a)G(a) - \int_{(a,b)} fG\,\lambda$$

これを**部分積分公式**(integration by parts formula) という。

証明. 集合 $A := \{(x,y) \in \mathbb{R}^2 : a < x < y < b\} \in \mathcal{B}(\mathbb{R}^2)$ を導入する。 $A \subset (a,b) \times (a,b)$ より

$$|1_A(x,y)f(x)g(y)| \le 1_{(a,b)}(x)|f(x)|\ 1_{(a,b)}(y)|g(y)|\ \forall (x,y) \in \mathbb{R}^2$$

が成り立つ。また Fubini-Tonelli の定理（系 8.16 と定理 8.14）によれば右辺は $\lambda^{(2)}$ 可積分である。従って関数 $(x,y) \mapsto 1_A(x,y)f(x)g(y)$ の $\lambda^{(2)}$ 可積分性が確認され Fubini の定理（定理 8.12）が適用できる。各 $x \in \mathbb{R}$ に対して

$$1_A(x,y) = \begin{cases} 1_{(x,b)}(y) & a < x < b \\ 0 & \text{otherwise} \end{cases}$$

であるから次が得られる。

$$(\star) \qquad \int_A f(x)g(y)\,\lambda^{(2)}(dx,dy) = \int_{(a,b)} \Big(\int_{(x,b)} f(x)g(y)\,\lambda(dy) \Big) \lambda(dx)$$

さて演習問題 4.26(i) の条件及び $\lambda(\{b\}) = 0$ という事実により

$$\int_{(x,b)} g(y)\,\lambda(dy) = G(b) - G(x), \int_{(a,b)} f(x)\,\lambda(dx) = F(b) - F(a)$$

ゆえに $(\star)$ の右辺は次に等しい。

$$\int_{(a,b)} f(x)(G(b) - G(x))\,\lambda(dx) = (F(b) - F(a))G(b) - \int_{(a,b)} f(x)G(x)\,\lambda(dx).$$

x, y の役割を入れ替えて同様の議論をすると次のように変形できる。

$$\int_A f(x)g(y)\,\lambda^{(2)}(dx,dy) = \int_{(a,b)} F(y)g(y)\,\lambda(dy) - F(a)(G(b) - G(a)).$$

得られた二つの表式を比較して部分積分公式に到達する。 □

8.23 演習問題. 台が有界な C^2 級関数 $f : \mathbb{R} \to \mathbb{C}$ が与えられたとする。
(i) 各 $x \in \mathbb{R}$ に対し $\displaystyle\int_{\mathbb{R}} e^{-\sqrt{-1}xy} f''(y)\,\lambda(dy) = -x^2 \int_{\mathbb{R}} e^{-\sqrt{-1}xy} f(y)\,\lambda(dy)$ であることを示せ。
(ii) 関数 $x \mapsto \displaystyle\int_{\mathbb{R}} e^{-\sqrt{-1}xy} f(y)\,\lambda(dy)$ は連続かつ λ 可積分であることを示せ。

8.24 例. $f : \mathbb{R} \to \mathbb{C}$ を有界連続関数でかつ λ 可積分なものとする。

$$g : \mathbb{R} \to \mathbb{C}, x \mapsto \int_{\mathbb{R}} e^{-\sqrt{-1}xy} f(y)\,\lambda(dy)$$

で定義される関数 g が λ 可積分であるなら

$$\int_{\mathbb{R}} e^{\sqrt{-1}xa} g(x)\,\lambda(dx) = 2\pi f(a)\ \forall a \in \mathbb{R}.$$

これを *Fourier* **逆変換公式**という。関連する事項を補題 10.13 でも取り上げる。

証明．関数 $(x,y)\mapsto e^{\sqrt{-1}xa}e^{-\sqrt{-1}xy}f(y)$ は $\lambda^{(2)}$ 可積分ではないので、直接 Fubini の定理を適用するわけにはいかない。そこで一工夫が必要なのだが、例 8.20 で登場した収束因子の導入というテクニックを再び使う。$\varepsilon\in\mathbb{R}_{>0}$ とする。関数

$$(x,y)\mapsto e^{\sqrt{-1}xa}e^{-\varepsilon|x|}e^{-\sqrt{-1}xy}f(y)$$

は $\lambda^{(2)}$ 可積分である。Fubini の定理（定理 8.12(ii)）を適用して積分の順序交換を行う。

$$\begin{aligned}\int_{\mathbb{R}} e^{\sqrt{-1}xa}e^{-\varepsilon|x|}g(x)\,\lambda(dx) &= \int_{\mathbb{R}^2} e^{\sqrt{-1}xa}e^{-\varepsilon|x|}e^{-\sqrt{-1}xy}f(y)\,\lambda^{(2)}(dxdy)\\ &= \int_{\mathbb{R}}\Big(\int_{\mathbb{R}} e^{\sqrt{-1}xa}e^{-\varepsilon|x|}e^{-\sqrt{-1}xy}\,\lambda(dx)\Big)f(y)\,\lambda(dy).\end{aligned}$$

例 4.4 で述べたように次が成り立つ。

$$\int_{\mathbb{R}} e^{-\varepsilon|x|}e^{-\sqrt{-1}x(y-a)}\,\lambda(dx)=\frac{2\varepsilon}{\varepsilon^2+(y-a)^2}\ \forall y\in\mathbb{R}.$$

従って

$$\frac{1}{2\pi}\int_{\mathbb{R}} e^{\sqrt{-1}xa}e^{-\varepsilon|x|}g(x)\,\lambda(dx)=\frac{\varepsilon}{\pi}\int_{\mathbb{R}}\frac{f(y)}{\varepsilon^2+(y-a)^2}\,\lambda(dy).$$

ここまでの展開では f の λ 可積分性だけが必要であった。f は有界連続であるから、補題 4.7(ii) より、$\varepsilon\to 0$ のとき右辺は $f(a)$ に収束する。他方、左辺の被積分関数は次のように可積分関数 g で評価される。

$$|e^{\sqrt{-1}xa}e^{-\varepsilon|x|}g(x)|\le|g(x)|\ \forall x\in\mathbb{R}\ \forall\varepsilon\in\mathbb{R}_{>0}.$$

Lebesgue の収束定理により $\varepsilon\to 0$ のとき左辺は目標とする積分に収束する。 □

8.25 演習問題. (i) $n\in\mathbb{N}$, $n\ge 2$, $x\in\mathbb{R}$ とする。$\displaystyle\int_{(0,+\infty)} e^{-\sqrt{-1}xy}y^{n-1}e^{-y}\,\lambda(dy)$ を評価せよ。
(ii) $n\in\mathbb{N}$, $n\ge 2$, $a\in\mathbb{R}$ とする。次を示せ。

$$\frac{1}{2\pi}\int_{\mathbb{R}}\frac{e^{\sqrt{-1}xa}}{(1+\sqrt{-1}x)^n}\,\lambda(dx)=\begin{cases}a^{n-1}e^{-a}/(n-1)! & a\ge 0\\ 0 & a\le 0\end{cases}$$

つぎは大筋において定理 3.19 で述べたものであるから、その証明は読者に委ねる。違いは関数 ρ の可積分性を要求するか否かである。定理 3.19 においては確率分布に言及していた。

8.26 定理. $\rho:\mathbb{R}^d\to\overline{\mathbb{R}}$ を非負値 Borel$(\mathbb{R}^d)$ 可測関数とする。
(i) 関数 $\mu:\mathrm{Borel}(\mathbb{R}^d)\to\overline{\mathbb{R}}$, $A\mapsto\displaystyle\int_A\rho\lambda$ は $(\mathbb{R}^d,\mathrm{Borel}(\mathbb{R}^d))$ 上の測度である。
(ii) 任意の非負値 Borel$(\mathbb{R}^d)$ 可測関数 $f:\mathbb{R}^d\to\overline{\mathbb{R}}$ に対して次が成り立つ。

$$\int_{\mathbb{R}^d} f\,\mu=\int_{\mathbb{R}^d} f\rho\,\lambda^{(d)}$$

(iii) 任意の Borel$(\mathbb{R}^d)$ 可測関数 $f:\mathbb{R}^d\to\overline{\mathbb{R}}$ に対してその μ 可積分性は $f\rho$ の $\lambda^{(d)}$ 可積分性と同値であり、可積分であるなら (ii) であげた関係が成り立つ。
(iv) 非負値 Borel$(\mathbb{R}^d)$ 可測関数 $g:\mathbb{R}^d\to\overline{\mathbb{R}}$ が $\int_A g\,\lambda=\mu(A)\ \forall A\in\mathrm{Borel}(\mathbb{R}^d)$ （測度 μ は (i) で述べたもの）を満たすなら $g=\rho\ \lambda^{(d)}$-a.e. である。

8.27 演習問題. 定理 8.26 を示せ。ヒント (i), (ii), (iii) については定理 3.19 の証明を参照せよ。(iv) については演習問題 7.13(ii) に帰着すればよいがそれには一工夫必要である。

8.28 定義. 定理 8.26 で規定されるような測度を d 次元**絶対連続測度**(absolutely continuous measure) という。また関数 ρ を（$\lambda^{(d)}$ に関する）**密度関数**(density function) という。

定理 8.26(iv) より絶対連続測度の密度関数は $\lambda^{(d)}$-a.e. の曖昧さを許す限り一意に定まる。

8.29 演習問題. 非負値 Borel($\mathbb{R}^d$) 可測関数 $\rho : \mathbb{R}^d \to \overline{\mathbb{R}}$ を密度関数に持つ d 次元絶対連続測度 μ が σ 有限であるための必要十分条件は $\rho < +\infty$ $\lambda^{(d)}$-a.e. であることを示せ。

8.30 補題. μ_1 を σ 有限な d_1 次元絶対連続測度で ρ_1 を密度関数とするもの、μ_2 を σ 有限な d_2 次元絶対連続測度で ρ_2 を密度関数とするものとする。このとき直積測度 $\mu_1 \otimes \mu_2$ は絶対連続で $\rho_1 \otimes \rho_2$ を密度関数に持つ。

証明. $A_1 \in \mathrm{Borel}(\mathbb{R}^{d_1})$, $A_2 \in \mathrm{Borel}(\mathbb{R}^{d_2})$ とする。Fubini-Tonelli の定理（定理 8.14(ii)）を適用して

$$\int_{\mathbb{R}^d} 1_{A_1 \times A_2} \rho_1 \otimes \rho_2 \, \lambda^{(d)} = \int_{\mathbb{R}^{d_1}} 1_{A_1} \rho_1 \, \lambda^{(d_1)} \int_{\mathbb{R}^{d_2}} 1_{A_2} \rho_2 \, \lambda^{(d_2)} = \mu_1(A_1)\mu_2(A_2)$$

従って絶対連続な測度 $A \mapsto \int_A \rho_1 \otimes \rho_2 \, \lambda^{(d)}$ は直積測度 $\mu_1 \otimes \mu_2$ に一致する。 □

8.31 例. 1 次元標準正規分布の n 重直積測度を n **次元標準正規分布**という。あるいは関数

$$\mathbb{R}^n \to \mathbb{R}, x \mapsto \frac{1}{\sqrt{(2\pi)^n}} \exp\{-\|x\|^2/2\}$$

を密度とする絶対連続測度といっても同じである。ここで $\|\cdot\|$ はユークリッドノルムである。

8.32 例. $t > 0$ とする。実確率変数系 X, Y は独立であり、X は平均 0 分散 t の正規分布に従い、Y はパラメータ $t/2$, $1/2$ の gamma 分布に従うとする。$P(Y > 0) = 1$ であることに注意する。このとき確率変数 $X1_{Y>0}/\sqrt{Y}$ の分布は関数

$$\mathbb{R} \to \overline{\mathbb{R}}, x \mapsto \frac{\Gamma((t+1)/2)}{\sqrt{\pi}\Gamma(t/2)\sqrt{t}} \frac{1}{(1 + x^2/t)^{(t+1)/2}}$$

を密度とする絶対連続分布である。これを自由度 t の t **分布**という。特に $t = 1$ のとき X, Y, $X1_{Y>0}/\sqrt{Y}$ の分布はそれぞれ標準正規分布、自由度 1 の χ^2 分布、Cauchy 分布である。

証明. 補題 8.30 により 2 次元確率変数 (X, Y) は関数

$$\mathbb{R}^2 \to \overline{\mathbb{R}}, (x, y) \mapsto \frac{1}{\sqrt{2\pi t}} \exp\{-x^2/2t\} \frac{1}{\Gamma(t/2) 2^{t/2}} 1_{(0,+\infty)}(y) y^{t/2-1} e^{-y/2}$$

を密度とする絶対連続分布に従う。f を非負 Borel 可測関数 $\mathbb{R} \to \overline{\mathbb{R}}$ とする。定理 6.38(i) と定理 8.26(ii) により $E[f(X1_{Y>0}/\sqrt{Y})]$ は

$$\frac{1}{\sqrt{\pi}\Gamma(t/2)\sqrt{t} 2^{(t+1)/2}} \int_{\mathbb{R} \times (0,+\infty)} f(x/\sqrt{y}) \exp\{-x^2/2t\} y^{t/2-1} e^{-y/2} \, \lambda^{(2)}(dx, dy)$$

で与えられる。Fubini-Tonelli の定理（定理 8.10）より上の積分は次に等しい。

$$\begin{aligned}&\int_{(0,+\infty)}\Big(\int_{\mathbb{R}} f(x/\sqrt{y})\exp\{-x^2/2t\}\lambda(dx)\Big)y^{t/2-1}e^{-y/2}\,\lambda(dy)\\ =&\int_{(0,+\infty)}\sqrt{y}\Big(\int_{\mathbb{R}} f(x)\exp\{-(\sqrt{y}x)^2/2t\}\lambda(dx)\Big)y^{t/2-1}e^{-y/2}\,\lambda(dy)\end{aligned}$$

ここで変数変換公式（補題 3.21）を適用している。積分の順序交換を行うと

$$\int_{\mathbb{R}} f(x)\Big(\int_{(0,+\infty)} y^{(t+1)/2-1}\exp\{-(1/2+x^2/2t)y\}\lambda(dy)\Big)\lambda(dx)$$

が得られる。y についての積分を実行すると

$$E[f(X1_{Y>0}/\sqrt{Y})] = \frac{1}{\sqrt{\pi}\Gamma(t/2)\sqrt{t}2^{(t+1)/2}}\int_{\mathbb{R}} f(x)\frac{\Gamma((t+1)/2)}{(1/2+x^2/2t)^{(t+1)/2}}\lambda(dx)$$

が成り立つことが分かる。以上を例 8.19 と比較すると面白い。 □

つぎは**期待値の乗法定理**と称して引用される。

8.33 定理. 実確率変数系 $X_1,\ X_2,\ \dots,\ X_n$ は独立であるとする。
(i) 非負値 Borel 可測関数 $f_1, f_2, \dots, f_n : \mathbb{R} \to \overline{\mathbb{R}}$ に対して次が成り立つ。

$$E[f_1(X_1)f_2(X_2)\dots f_n(X_n)] = E[f_1(X_1)]E[f_2(X_2)]\dots E[f_n(X_n)].$$

(ii) Borel 可測関数 $f_1, f_2, \dots, f_n : \mathbb{R} \to \overline{\mathbb{R}}$ に対して $E[|f_i(X_i)|] < +\infty\ \forall i = 1, 2, \dots, n$ なら積 $f_1(X_1)f_2(X_2)\dots f_n(X_n)$ は可積分であって (i) で述べた等式が成り立つ。

証明. (i) 煩雑化を避けるため $n = 2$ として話を進める。定理 6.38 を適用して

$$E[f_1(X_1)f_2(X_2)] = \int_{\mathbb{R}^2} f_1(x_1)f_2(x_1)\,\mathcal{L}((X_1, X_2), dx)$$

を得る。$\mathcal{L}((X_1, X_2), \cdot) = \mathcal{L}(X_1, \cdot)\otimes\mathcal{L}(X_2, \cdot)$ であるから Fubini-Tonelli の定理（定理 8.14(ii)）により右辺は

$$\int_{\mathbb{R}} f_1(x_1)\,\mathcal{L}(X_1, dx_1)\int_{\mathbb{R}^2} f_2(x_2)\,\mathcal{L}(X_2, dx_2)$$

に等しい。再び定理 6.38 を適用して結論が導けた。
(ii) (i) と同様であるがこんどは定理 8.14(iii) も必要である。 □

8.34 系. 実確率変数系 $X_1,\ X_2,\ \dots,\ X_n$ は独立であるとする。このとき次が成り立つ。

$$E[\exp\{\sum_{i=1}^n \sqrt{-1}z_iX_i\}] = \prod_{i=1}^n E[\exp\{\sqrt{-1}z_iX_i\}]\ \forall(z_1, z_2, \dots, z_n) \in \mathbb{R}^n.$$

8.35 例. n 次元標準正規分布 μ に対し次が成り立つ。

$$\int_{\mathbb{R}^n} e^{\sqrt{-1}(z,x)}\mu(dx) = e^{-\|z\|^2/2}\ \forall z \in \mathbb{R}^n$$

ここで $(\cdot,\cdot)$ はユークリッド内積 $\|\cdot\|$ はユークリッドノルムである。

8.36 定理. 実確率変数系 $X_1, X_2, \ldots, X_n$ が独立であるのは次と同値である。

$$E[f_1(X_1)f_2(X_2)\ldots f_n(X_n)] = E[f_1(X_1)]E[f_2(X_2)]\ldots E[f_n(X_n)] \ \forall f_1, f_2, \ldots, f_n \in C_0^\infty(\mathbb{R}).$$

証明. $X_1, X_2, \ldots, X_n$ の結合分布を μ とし、それらの分布の直積測度を ν とする。すなわち

$$\mu := \mathcal{L}((X_1, X_2, \ldots, X_n), \cdot),\ \nu := \mathcal{L}(X_1, \cdot) \otimes \mathcal{L}(X_2, \cdot) \otimes \cdots \otimes \mathcal{L}(X_n, \cdot).$$

定理 6.38 と定理 8.33 により、与えられた条件の左辺および右辺は次のようにかける。

$$\text{左辺} = \int_{\mathbb{R}^n} f_1 \otimes f_2 \otimes \cdots \otimes f_n \, \mu,\ \text{右辺} = \int_{\mathbb{R}^n} f_1 \otimes f_2 \otimes \cdots \otimes f_n \, \nu.$$

従って定理 7.22 を適用して $\mu = \nu$ を得る。 □

9 可逆アファイン写像と絶対連続測度

$(\mathbb{R}^d, \mathrm{Borel}(\mathbb{R}^d))$ 上では絶対連続な測度が重要である。

記号

λ: 1 次元 Lebesgue 測度; $\lambda^{(d)}$: d 次元 Lebesgue 測度

9.1 例. C を n 次正定値対称行列、$a \in \mathbb{R}^n$ とする。対称行列 C^{-1} が誘導する 2 次形式を $C^{-1}(\cdot, \cdot)$ と表す。平均 a 共分散 C の n **次元正規分布**とは関数

$$\mathbb{R}^n \to \mathbb{R}, x \mapsto \frac{1}{\sqrt{(2\pi)^n \det C}} \exp\{-C^{-1}(x-a, x-a)/2\}$$

を密度に持つ n 次元絶対連続分布をいう。ただしこれは暫定定義であり、最終決着は定義 10.6 で行う。平均が 0 で共分散が単位行列の場合は n 次元標準正規分布（例 8.31）である。

$\alpha \in GL(d)$, $\beta \in \mathbb{R}^d$ とする。絶対連続な測度のアファイン写像 $\varphi : \mathbb{R}^d \to \mathbb{R}^d$, $x \mapsto \alpha x + \beta$ による像測度について考察したい。それには d 次元 Lebesgue 測度 $\lambda^{(d)}$ の像測度 $\varphi_* \lambda^{(d)}$ を確定することが求められる。まず、アファイン写像が直積構造を持つ場合を検討する。

記号

与えられた写像 $\varphi_1 : S_1 \to T_1$, $\varphi_2 : S_2 \to T_2$ に対して写像 $S_1 \times S_2 \to T_1 \times T_2$, $(x_1, x_2) \mapsto (\varphi_1(x_1), \varphi_2(x_2))$ を $\varphi_1 \times \varphi_2$ と表記する。

9.2 補題. $(S_1, \mathcal{B}_1)$, $(S_2, \mathcal{B}_2)$, $(T_1, \mathcal{M}_1)$, $(T_2, \mathcal{M}_2)$ を非自明な可測空間、$\varphi_1 : S_1 \to T_1$ を対 $\mathcal{B}_1$, $\mathcal{M}_1$ に関し可測な写像、$\varphi_2 : S_2 \to T_2$ を $\mathcal{B}_2$, $\mathcal{M}_2$ に関し可測な写像、μ_1 を $(S_1, \mathcal{B}_1)$ 上の測度、μ_2 を $(S_2, \mathcal{B}_2)$ 上の測度とする。このとき

1° $\varphi_1 \times \varphi_2$ は対 $\mathcal{B}_1 \otimes \mathcal{B}_2$, $\mathcal{M}_1 \otimes \mathcal{M}_2$ に関して可測であり、
2° $\varphi_{1*}\mu_1$, $\varphi_{2*}\mu_2$ が σ 有限なら $(\varphi_1 \times \varphi_2)_*(\mu_1 \otimes \mu_2) = (\varphi_{1*}\mu_1) \otimes (\varphi_{2*}\mu_2)$ が成立。

証明. 可測性の判定に補題 5.6(iv) を適用する。$\sigma(\mathcal{M}_1 \times \mathcal{M}_2) = \mathcal{M}_1 \otimes \mathcal{M}_2$ であるから条件

$$(\varphi_1 \times \varphi_2)^{-1}(A_1 \times A_2) \in \mathcal{B}_1 \otimes \mathcal{B}_2 \ \forall A_1 \in \mathcal{M}_1, \forall A_2 \in \mathcal{M}_2$$

を確かめればよいが、それは関係 $(\varphi_1 \times \varphi_2)^{-1}(A_1 \times A_2) = \varphi_1^{-1}(A_1) \times \varphi_2^{-1}(A_2)$ より従う。他方、測度については直積測度の特徴付けにより

$$(\varphi_1 \times \varphi_2)_*(\mu_1 \otimes \mu_2)(A_1 \times A_2) = \mu_1(\varphi_1^{-1}(A_1))\mu_2(\varphi_2^{-1}(A_2)) \ \forall A_1 \in \mathcal{M}_1, \forall A_2 \in \mathcal{M}_2$$

を確かめればよいが、それも上で述べた関係より従う。 □

9.3 系. $\alpha \in GL(d)$, $\beta \in \mathbb{R}^d$ とする。α が対角行列であればアファイン写像 $\varphi : \mathbb{R}^d \to \mathbb{R}^d$, $x \mapsto \alpha x + \beta$ に対して $\varphi_*\lambda^{(d)} = \lambda^{(d)}/|\det \alpha|$ が成り立ち、また

$$\int_{\mathbb{R}^d} f(\alpha x + \beta)\, \lambda^{(d)}(dx) = \frac{1}{|\det \alpha|} \int_{\mathbb{R}^d} f\, \lambda^{(d)}$$

が非負値 $\mathrm{Borel}(\mathbb{R}^d)$ 可測関数 $f : \mathbb{R}^d \to \overline{\mathbb{R}}$ に対して成り立つ。

証明. $d = 2$ として議論を進める（一般次元の証明には帰納法を使う）。α の対角成分を α_1, α_2 また β の成分を β_1, β_2 とする。写像 $\varphi_i : \mathbb{R} \to \mathbb{R}$, $x \mapsto \alpha_i x + \beta_i$ を導入する。補題 3.21 によれば、$\varphi_{i*}\lambda = \lambda/|\alpha_i|$ である。従って系 8.16 と補題 9.2 より次が成り立つ。

$$(\varphi_1 \times \varphi_2)_*\lambda^{(2)} = \frac{1}{|\alpha_1\alpha_2|}\lambda^{(2)} = \frac{1}{|\det \alpha|}\lambda^{(2)}$$

写像 $\varphi_1 \times \varphi_2$ は $\mathbb{R}^2 \to \mathbb{R}^2$, $x \mapsto \alpha x + \beta$ を表す。定理 6.20 を適用して結論を得る。 □

系 9.3 から特に d **次元** *Lebesgue* **測度** $\lambda^{(d)}$ **の平行移動不変性**がわかる。

9.4 定理. μ_1, μ_2 を σ 有限な d 次元絶対連続測度であって、それぞれ関数 ρ_1, ρ_2 を密度に持つとする。このとき Borel 可測写像 $\mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$, $(x, y) \mapsto x + y$ による直積測度 $\mu_1 \otimes \mu_2$ の像測度は絶対連続で $x \mapsto \int \rho_1(x - y)\rho_2(y)\, \lambda^{(d)}(dy)$ を密度関数に持つ。

証明. 補題 8.30 によれば $\mu_1 \otimes \mu_2$ は関数 $\rho_1 \otimes \rho_2$ を密度とする絶対連続測度である。ν を求める像測度、$A \in \mathrm{Borel}(\mathbb{R}^d)$ とする。定理 6.20(i) と定理 8.26(ii) により

$$\int_{\mathbb{R}^d} f\nu = \int_{\mathbb{R}^d \times \mathbb{R}^d} f(x + y)(\mu_1 \otimes \mu_2)(dx, dy) = \int_{\mathbb{R}^d \times \mathbb{R}^d} f(x + y)\rho_1(x)\rho_2(y)\, \lambda^{(d+d)}(dx, dy)$$

右辺は Fubini-Tonelli の定理（定理 8.10）と平行移動不変性（系 9.3）により次に等しい。

$$\begin{aligned} &\int_{\mathbb{R}^d} \Big(\int_{\mathbb{R}^d} f(x + y)\rho_1(x)\rho_2(y)\lambda^{(d)}(dx)\Big)\lambda^{(d)}(dy) \\ &= \int_{\mathbb{R}^d} \Big(\int_{\mathbb{R}^d} f(x)\rho_1(x - y)\rho_2(y)\lambda^{(d)}(dx)\Big)\lambda^{(d)}(dy) \\ &= \int_{\mathbb{R}^d} f(x)\Big(\int_{\mathbb{R}^d} \rho_1(x - y)\rho_2(y)\lambda^{(d)}(dy)\Big)\lambda^{(d)}(dx). \end{aligned}$$

従って ν は $x \mapsto \int \rho_1(x - y)\rho_2(y)\, \lambda^{(d)}(dy)$ を密度とする絶対連続測度である。 □

9.5 定義. 非負値 Borel 可測関数 $\rho_1, \rho_2 : \mathbb{R}^d \to \overline{\mathbb{R}}$ に対して関数 $x \mapsto \int_{\mathbb{R}^d} \rho_1(x-y)\rho_2(y)\,\lambda^{(d)}(dy)$ を ρ_1 と ρ_2 の**たたみ込み**(convolution) あるいは**合成積**という。

9.6 系. X_1, X_2 を互いに独立な d 次元確率変数とする。それぞれが関数 ρ_1, ρ_2 を密度とする絶対連続分布に従うなら $X_1 + X_2$ は ρ_1 と ρ_2 のたたみ込みを密度とする絶対連続分布に従う。

9.7 例. X, Y を互いに独立な確率変数とする。また $t, s > 0$ とする。
(i) X が中心 0 半値幅 t の Cauchy 分布に従い、Y が中心 0 半値幅 s の Cauchy 分布に従うなら $X+Y$ は中心 0 半値幅 $t+s$ の Cauchy 分布に従う。
(ii) X が平均 0 分散 t の正規分布に従い、Y が平均 0 分散 s の正規分布に従うなら $X+Y$ は平均 0 分散 $t+s$ の正規分布に従う。
(iii) X が指数 t の gamma 分布に従い Y が指数 s の gamma 分布に従うなら $X+Y$ は指数 $t+s$ の gamma 分布に従う。

証明. (i) は補題 4.7(i) を、また (ii) は演習問題 4.10(i) で確かめたことをそれぞれ系 9.6 の観点から焼き直したものである。(iii) を示すために次を計算する。

$$\begin{aligned}&\frac{1}{\Gamma(t)\Gamma(s)}\int_{\mathbb{R}} 1_{(0,+\infty)}(x-y)\ (x-y)^{t-1}e^{-(x-y)}1_{(0,+\infty)}(y)\ y^{s-1}e^{-y}\,\lambda(dy)\\ &= \frac{1}{\Gamma(t)\Gamma(s)}\int_{(0,x)}(x-y)^{t-1}y^{s-1}\,\lambda(dy)\ e^{-x}\ \text{ただし}\ x>0\end{aligned}$$

$x \le 0$ なら左辺は 0 である。ここで $x > 0$ として変数変換公式（補題 3.21）を適用する。

$$\int_{(0,x)}(x-y)^{t-1}y^{s-1}\,\lambda(dy) = \int_{(0,1)}(1-y)^{t-1}y^{s-1}\,\lambda(dy)\ x^{t+s-1}.$$

よって系 9.6 より $X+Y$ の分布は絶対連続でありその密度関数は次で与えられる。

$$x \mapsto \frac{B(t,s)}{\Gamma(t)\Gamma(s)}1_{(0,+\infty)}(x)x^{t+s-1}e^{-x}\ \text{ただし}\ B\ \text{は beta 関数}$$

これは確率測度の密度であるから、よく知られた公式

$$\frac{\Gamma(t)\Gamma(s)}{\Gamma(t+s)} = B(t,s)\ \textit{gamma}\ \textbf{関数と}\ \textit{beta}\ \textbf{関数の関係}$$

が得られる。また $X+Y$ は指数 $t+s$ の gamma 分布に従う。 □

上の例にあげたように分布に付随するパラメータがたたみ込みに関して加法性を持つときその分布族は**再生性**(reproducing property) を持つという。

9.8 演習問題. $a_1 < b_1$, $a_2 < b_2$ かつ $b_1 - a_1 \le b_2 - a_2$ とする。X, Y が互いに独立な確率変数であり、それぞれ区間 (a_1, b_1), (a_2, b_2) 上の一様分布に従うなら $X+Y$ は関数

$$x \mapsto \begin{cases} (x-a_1-a_2)/(b_1-a_1)(b_2-a_2) & a_1+a_2 < x < b_1+a_2 \\ 1/(b_2-a_2) & b_1+a_2 \le x \le a_1+b_2 \\ (b_1+b_2-x)/(b_1-a_1)(b_2-a_2) & a_1+b_2 < x < b_1+b_2 \\ 0 & \text{otherwise} \end{cases}$$

を密度とする絶対連続分布に従うことを示せ。

9.9 演習問題. $f,\, g : \mathbb{R}^d \to \overline{\mathbb{R}}$ を $\lambda^{(d)}$ 可積分な Borel 可測関数とする。
(i) $\{x \in \mathbb{R}^d :$ 積分 $\int_{\mathbb{R}^d} f(x-y)f(y)\,\lambda^{(d)}(dy)$ が存在 $\}$ は $\lambda^{(d)}$-a.e. 集合であることを示せ。
(ii) $x \mapsto \int_{\mathbb{R}^d} f(x-y)f(y)\,\lambda^{(d)}(dy)$ は $\lambda^{(d)}$ 可積分関数であることを示せ。

9.10 補題. $p > 0$ とする。また $\|\ \|$ を $\mathbb{R}^d$ のユークリッドノルムとする。可測写像 $\varphi : \mathbb{R}^d \to \mathbb{R}$, $x \mapsto \|x\|^p$ による $\lambda^{(d)}$ の像測度は

$$\mathbb{R} \to \overline{\mathbb{R}}, r \mapsto 1_{(0,+\infty)}(r)\lambda^{(d)}(\{x \in \mathbb{R}^d : \|x\| \le 1\})\frac{d}{p}r^{d/p-1}$$

を密度関数に持つ絶対連続測度である。

証明. $a < b$ とする。このとき $\varphi^{-1}((a,b]) = \{x \in \mathbb{R}^d : \|x\|^p \le b\} \setminus \{x \in \mathbb{R}^d : \|x\|^p \le a\}$ であり、$\{x \in \mathbb{R}^d : \|x\|^p \le b\}$ が空でないのは $b \ge 0$ である時に限る。$b \ge 0$ なら系 9.3 より

$$\lambda^{(d)}(\{x \in \mathbb{R}^d : \|x\|^p \le b\}) = (b^{1/p})^d \lambda^{(d)}(\{x \in \mathbb{R}^d : \|x\| \le 1\}) < +\infty$$

が成り立つ。従って $\varphi_*\lambda^{(d)}((a,b]) < +\infty$ でありかつ

$$\begin{aligned}
&\varphi_*\lambda^{(d)}((a,b]) = \lambda^{(d)}(\{x \in \mathbb{R}^d : \|x\|^p \le b\}) - \lambda^{(d)}(\{x \in \mathbb{R}^d : \|x\|^p \le a\}) \\
&= (\max\{b,0\}^{d/p} - \max\{a,0\}^{d/p})\lambda^{(d)}(\{x \in \mathbb{R}^d : \|x\| \le 1\}) \\
&= \lambda^{(d)}(\{x \in \mathbb{R}^d : \|x\| \le 1\})\frac{d}{p}\int_{(a,b]} 1_{(0,+\infty)}(r)r^{d/p-1}\lambda(dr).
\end{aligned}$$

左半開区間全体に空集合 $\emptyset$ を付加した集合族 $\mathcal{C}$ は π システムをなし、かつ $\mathcal{C}$ は Borel($\mathbb{R}$) を生成する。よって定理 7.12(ii) を適用して $\varphi_*\lambda^{(d)}$ が所定の関数を密度に持つ絶対連続測度に一致することが結論づけられる。 □

9.11 系. (i) $\lambda^{(d)}(\{x \in \mathbb{R}^d : \|x\| \le 1\}) = \dfrac{2\sqrt{\pi^d}}{d\Gamma(d/2)}$.
(ii) $p > 0$ とする。非負値 Borel($\mathbb{R}$) 可測関数 $f : \mathbb{R} \to \overline{\mathbb{R}}$ に対して次が成り立つ。

$$\int_{\mathbb{R}^d} f(\|x\|^p)\,\lambda^{(d)}(dx) = \frac{2\sqrt{\pi^d}}{p\Gamma(d/2)}\int_{(0,+\infty)} f(r)r^{d/p-1}\,\lambda(dr).$$

証明. 補題 9.10 で像測度が特定されているので、定理 6.20(i) と定理 8.26(ii) により

$$\int_{\mathbb{R}^d} f(\|x\|^p)\,\lambda^{(d)}(dx) = \lambda^{(d)}(\{x \in \mathbb{R}^d : \|x\| \le 1\})\frac{d}{p}\int_{(0,+\infty)} f(r)r^{d/p-1}\,\lambda(dr)$$

が成り立つ。以下、p と関数 f を $p = 2$, $f(r) = e^{-r}$ と選ぶ。右辺を評価する。

$$\frac{d}{2}\int_{(0,+\infty)} e^{-r}r^{d/2-1}\,\lambda(dr) = \frac{d\Gamma(d/2)}{2}$$

他方、右辺は Fubini-Tonelli の定理を適用して例 8.19 に帰着させる。

$$\int_{\mathbb{R}^d} e^{-\|x\|^2}\,\lambda^{(d)}(dx) = \Big(\int_{\mathbb{R}} e^{-|x|^2}\,\lambda(dx)\Big)^d = \sqrt{\pi^d}$$

以上より $\lambda^{(d)}(\{x \in \mathbb{R}^d : \|x\| \le 1\})$ が評価できた。 □

9.12 例. (i) 可測写像 $\mathbb{R}^d \to \mathbb{R}$, $x \mapsto \|x\|^2$ による d 次元標準正規分布の像測度は

$$\mathbb{R} \to \overline{\mathbb{R}}, u \mapsto 1_{(0,+\infty)}(u)\frac{1}{\sqrt{2^d}\Gamma(d/2)}e^{-u/2}u^{d/2-1}$$

を密度関数に持つ絶対連続測度（即ち自由度 d の χ^2 分布）である。
(ii) 各 $t > 0$ に対して次が成り立つ。

$$\int_{\mathbb{R}^d} \frac{t}{(t^2 + \|x\|^2)^{(d+1)/2}} \lambda^{(d)}(dx) = \frac{\sqrt{\pi^{d+1}}}{\Gamma((d+1)/2)}.$$

9.13 演習問題. (i) 例 9.12(ii) を示せ。ヒント　例 3.36
(ii) 確率変数系 X, Y は独立であり、X は d 次元標準正規分布に従い、Y は自由度 1 の χ^2 分布に従うとする（$P(Y > 0) = 1$）。このとき d 次元確率変数 $X1_{Y>0}/\sqrt{Y}$ の分布は関数

$$\mathbb{R}^d \to \overline{\mathbb{R}}, x \mapsto \frac{\Gamma((d+1)/2)}{\sqrt{\pi^{d+1}}}\frac{1}{(1+\|x\|^2)^{(d+1)/2}}$$

を密度とする d 次元絶対連続分布であることを示せ。ヒント　例 8.32

平行移動不変性が Lebesgue 測度の特質であることを使って可逆アファイン写像による d 次元 Lebesgue 測度 $\lambda^{(d)}$ の像測度を確定する。

9.14 定理. μ を $\mathbb{R}^d$ 上の Radon 測度とする。それが平行移動不変 、すなわち $\mu(A+t) = \mu(A)$ $\forall A \in \mathrm{Borel}(\mathbb{R}^d)$ $\forall t \in \mathbb{R}^d$ を満たす、なら $\mu = \mu((0,1]^d)\lambda^{(d)}$ である。

証明. 複雑化を避けるために $d = 2$ として議論を進める（一般次元の証明には帰納法を使う)。有界な $A \in \mathrm{Borel}(\mathbb{R})$ を固定する。各 $t, s \in \mathbb{R}_{>0}$ に対して次が成り立つ。

$$\mu((0,t+s] \times A) = \mu((0,t] \times A) + \mu((t,s+t] \times A)$$

μ は Radon 測度なので各項は有限の値である。よって関数 $f : \mathbb{R}_{>0} \to \mathbb{R}$, $t \mapsto \mu((0,t] \times A)$ は右連続であり、さらに平行移動不変性により次を満たすことになる。

$$f(t+s) = f(t) + f(s)\ \forall t, \forall s$$

そのような関数は $f : t \mapsto f(1)t$ に限る。従って $\mu((0,t] \times A) = \mu((0,1] \times A)t$ $\forall t \in \mathbb{R}_{>0}$ となるが、ふたたび平行移動不変性により次を得る。

$$\mu((a,b] \times A) = \mu((0,1] \times A)(b-a)\ \forall a, b \in \mathbb{R} \text{ with } a < b$$

同様に $\mu(A \times (a,b]) = \mu(A \times (0,1])(b-a)$ $\forall a, b \in \mathbb{R}$, $a < b$ も導くことができる。以上より

$$\mu((a_1,b_1] \times (a_2,b_2]) = \mu((0,1]^2)(b_1-a_1)(b_2-a_2) = \mu((0,1]^2)\lambda^{(2)}((a_1,b_1] \times (a_2,b_2])$$

である。ゆえに Lebesgue 測度の特徴付けを適用して結論 $\mu = \mu((0,1]^2)\lambda^{(2)}$ に至った。　□

9.15 系. d 次元 Lebesgue 測度 $\lambda^{(d)}$ は回転不変である。すなわち $\alpha \in O(d)$ であればアファイン写像 $\varphi : \mathbb{R}^d \to \mathbb{R}^d$, $x \mapsto \alpha x$ に対して $\varphi_*\lambda^{(d)} = \lambda^{(d)}$ が成り立つ。

証明. 写像 φ による有界集合の逆像は再び有界なので、像測度 $\varphi_*\lambda^{(d)}$ は Radon 測度である。他方 $\varphi(x)+t=\varphi(x+\alpha^{-1}t)\ \forall x,t\in\mathbb{R}^d$ であるから Lebesgue 測度の平行移動不変性と定理 6.24 を適用して測度 $\varphi_*\lambda^{(d)}$ の平行移動不変性を得る。従って定理 9.14 の前提条件が満たされる。よって

$$\exists c\in\mathbb{R}_{\geq 0}\ \text{s.t.}\ \varphi_*\lambda^{(d)}=c\lambda^{(d)}.$$

特に、$\|\cdot\|$ をユークリッドノルムとするとき、開集合 $A=\{x\in\mathbb{R}^d:\|x\|<1\}$ については

$$\varphi^{-1}(A)=A, 0<\lambda^{(d)}(A)<+\infty$$

なので $c=1$ を得る。 □

9.16 補題. $\alpha\in GL(d)$ に対して $K,L\in O(d)$ と対角行列 H が存在して $\alpha=KHL$ を満たす。

証明. ${}^{\mathrm{t}}\alpha\alpha$ は正定値対称行列である。よって ${}^{\mathrm{t}}\alpha\alpha=L^{-1}H^2L$ を満たすような対角行列 H と $L\in O(d)$ が存在する。このとき $\alpha L^{-1}H^{-1}$ は直交行列である。 □

アファイン写像に対する**変数変換公式**(change of variable formula) を証明しておく。特に行列式が 1 または -1 であるなら *Lebesgue* **積分は保存される**。

9.17 定理. $\alpha\in GL(d)$, $\beta\in\mathbb{R}^d$ とする。
(i) 非負値 Borel($\mathbb{R}^d$) 可測関数 $f:\mathbb{R}^d\to\overline{\mathbb{R}}$ に対して次が成り立つ。

$$\int_{\mathbb{R}^d}f(\alpha x+\beta)\,\lambda^{(d)}(dx)=\frac{1}{|\det\alpha|}\int_{\mathbb{R}^d}f\,\lambda^{(d)}$$

(ii) Borel($\mathbb{R}^d$) 可測関数 $f:\mathbb{R}^d\to\overline{\mathbb{R}}$ に対してその $\lambda^{(d)}$ 可積分性は $x\mapsto f(\alpha x+\beta)$ のそれと同値であり、可積分なら上の等式が成り立つ。

証明. アファイン写像 $\mathbb{R}^d\to\mathbb{R}^d$, $x\mapsto\alpha x+\beta$ を φ と書こう。補題 9.16 によりこの写像は系 9.3 あるいは系 9.15 が適用可能な写像の合成として表現できる。また補題 9.16 における表現では $|\det\alpha|=|\det H|$ が成り立っている。従って定理 6.24 により $\varphi_*\lambda^{(d)}=\lambda^{(d)}/|\det\alpha|$ が導かれる。さらに定理 6.20 を適用して結論に到達する。 □

9.18 系. $\alpha\in GL(d)$, $\beta\in\mathbb{R}^d$, μ を d 次元絶対連続測度、関数 ρ を μ の密度とする。このときアファイン写像 $\varphi:\mathbb{R}^d\to\mathbb{R}^d$, $x\mapsto\alpha x+\beta$ による μ の像測度 $\varphi_*\mu$ は関数

$$\mathbb{R}^d\to\overline{\mathbb{R}}, x\mapsto\rho(\alpha^{-1}(x-\beta))/|\det\alpha|$$

を密度とする絶対連続測度である。

9.19 例. $\alpha\in GL(n)$, $\beta\in\mathbb{R}^n$ とし、X を n 次元標準正規分布に従う確率変数とする。このとき $\mathbb{R}^n$ 値確率変数 $\alpha X+\beta$ は平均 β 共分散 $\alpha\,{}^{\mathrm{t}}\alpha$ の n 次元正規分布に従う。

9.20 演習問題. 系 9.18, 例 9.19 を示せ。

9.21 補題. 平均 a 共分散 C の n 次元正規分布 μ に対し次が成り立つ。

$$\int_{\mathbb{R}^n} |e^{(z,x)}|\mu(dx) < +\infty, \int_{\mathbb{R}^n} e^{(z,x)}\mu(dx) = \exp\{(z,a) + C(z,z)/2\} \ \forall z \in \mathbb{C}^n.$$

ここで $(\cdot,\cdot)$ はユークリッド内積であり $C(\cdot,\cdot)$ は対称行列 C が誘導する 2 次形式を表す。

証明. C は n 次正定値対称行列であるから $C = \alpha\,{}^{t}\alpha$ を満たす $\alpha \in GL(n)$ が存在する。従って例 9.19 により例 8.35 に帰着する。 □

9.22 例. ここでは $\|\ \|$ は $\mathbb{R}^d$ のユークリッドノルムを表す。
(i) **熱核の** *Fourier* **積分表示**　$t > 0$, $x \in \mathbb{R}^d$ に対して次が成り立つ。

$$\frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} \exp\{\sqrt{-1}(\xi,x) - t\|\xi\|^2\}\ \lambda^{(d)}(d\xi) = \frac{1}{\sqrt{(4\pi t)^d}}\exp\{-\frac{\|x\|^2}{4t}\}.$$

(ii) **多次元** *Cauchy* **核の** *Fourier* **積分表示**　$t > 0$, $x \in \mathbb{R}^d$ に対して次が成り立つ。

$$\frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} \exp\{\sqrt{-1}(\xi,x) - t\|\xi\|\}\ \lambda^{(d)}(d\xi) = \frac{\Gamma((d+1)/2)}{\sqrt{\pi^{d+1}}}\frac{t}{(t^2+\|x\|^2)^{(d+1)/2}}.$$

9.23 定義. 次の関数を d **次元熱核** (heat kernel) という。

$$\mathbb{R}_{>0}\times\mathbb{R}^d \to \mathbb{R}, (t,x) \mapsto \exp\{-\|x\|^2/4t\}/\sqrt{(4\pi t)^d}$$

9.24 演習問題. 例 9.22(ii) を示せ。ヒント 例 8.21 より次が成り立つ。

$$\int_{(0,+\infty)} \frac{1}{\sqrt{\pi u}}\exp\{-\frac{t^2\|\xi\|^2}{4u} - u\}\ \lambda(du) = \exp\{-t\|\xi\|\}$$

Fubini の定理が適用できることを確かめた上で、積分の順序交換を行え。累次積分の評価には、熱核の Forier 積分表示（例 9.22(i)）と gamma 分布の密度関数（例 4.3）を参照せよ。

10　特性関数と正規分布

　第 4 節において調べたように実数値確率変数についてはその分布と分布関数が全単写対応をしていた。多次元の場合も分布関数を定義できるが、$\mathbb{R}^d$ には自然な順序構造が存在しないので、必ずしも使いやすいものではない。そこで、線形空間の構造を反映する特性関数を活用していくことにする。これにより正規分布と線形構造との結びつきが明らかになる。

10.1 定義. (i) $(\mathbb{R}^d, \mathrm{Borel}(\mathbb{R}^d))$ 上の確率測度 μ に対して関数

$$\mathrm{char}(\mu,\cdot) : \mathbb{R}^d \to \mathbb{C}, \xi \mapsto \int_{\mathbb{R}^d} e^{\sqrt{-1}(\xi,x)}\,\mu(dx)$$

を μ の**特性関数**(characteristic function) という。ここで $(\cdot,\cdot)$ はユークリッド内積を表す。
(ii) d 次元確率変数 X に対してその分布 $\mathcal{L}(X,\cdot)$ の特性関数を X の特性関数といい $\mathrm{char}(X,\cdot)$ と表記する。なお定理 6.38 により $\mathrm{char}(X,\xi) = E[e^{\sqrt{-1}(\xi,X)}]$ である。

10.2 補題. (i) $(\mathbb{R}^d, \mathrm{Borel}(\mathbb{R}^d))$ 上の確率測度 μ に対して $\mathrm{char}(\mu,\cdot)$ は有界かつ連続である。
(ii) $n \in \mathbb{N}$ に対して $\int_{\mathbb{R}^d} \|x\|^n \,\mu(dx) < +\infty$ であれば、$\mathrm{char}(\mu,\cdot)$ は C^n 級である。とくに $\int_{\mathbb{R}^d} \|x\|^2 \,\mu(dx) < +\infty$ であれば、原点 0 における偏微分係数は次を満たす。

$$\frac{\partial \mathrm{char}(\mu,\cdot)}{\partial \xi_i}(0) = \sqrt{-1}\int_{\mathbb{R}^d} x_i \,\mu(dx) \quad \frac{\partial^2 \mathrm{char}(\mu,\cdot)}{\partial \xi_i \partial \xi_j}(0) = -\int_{\mathbb{R}^d} x_i x_j \,\mu(dx)$$

証明. (i) 補題 3.25 により次が成り立つ。

$$\Big|\int_{\mathbb{R}^d} e^{\sqrt{-1}(\xi,x)}\,\mu(dx)\Big| \le \int_{\mathbb{R}^d} |e^{\sqrt{-1}(\xi,x)}|\,\mu(dx) = 1 \ \forall \xi \in \mathbb{R}^d$$

次に $\xi \in \mathbb{R}^d$ を固定し、ξ_k を ξ に収束する $\mathbb{R}^d$ の点列とする。このとき各 $x \in \mathbb{R}^d$ に対して $\exp\{\sqrt{-1}(\xi_k, x)\}$ は $\exp\{\sqrt{-1}(\xi, x)\}$ に収束する。他方、

$$|\exp\{\sqrt{-1}(\xi_k, x)\}| \le 1 \ \forall x \in \mathbb{R}^d, \forall k \in \mathbb{N}$$

が成り立つ。定数関数 1 は μ 可積分であるから Lebesgue の収束定理を適用して収束

$$\lim_{n\to\infty}\int_{\mathbb{R}^d} \exp\{\sqrt{-1}(\xi_k, x)\}\,\mu(dx) = \int_{\mathbb{R}^d} e^{\sqrt{-1}(\xi,x)}\,\mu(dx)$$

がわかる。以上は ξ に収束する任意の点列にあてはまるので連続性が得られる。

(ii) $0 < p < n$ とする。このとき $\|x\|^p \le 1 + \|x\|^n \ \forall x \in \mathbb{R}^d$ であるから

$$\int_{\mathbb{R}^d} \|x\|^p \,\mu(dx) \le 1 + \int_{\mathbb{R}^d} \|x\|^n \,\mu(dx) < +\infty$$

が成り立つことに注意する。以下、記号が煩雑になるのを回避するため、$d = 1$ の場合に議論する（$d > 1$ であっても基本線は同じであるが、偏微分係数をシステマティックに表現するのに少なからず工夫が必要となる）。$\xi \in \mathbb{R}$ を固定し、ξ_k を ξ に収束する点列とする。さて

$$e^z = 1 + \int_{(0,1)} e^{zt}\,\lambda(dt)\, z \ \forall z \in \mathbb{C}$$

を利用すると、（もちろん $\xi_k \ne \xi$ であると仮定して）各 $x \in \mathbb{R}$ に対して差分商は

$$\frac{\exp\{\sqrt{-1}\xi_k x\} - \exp\{\sqrt{-1}\xi x\}}{\xi_k - \xi} = \sqrt{-1}x e^{\sqrt{-1}\xi x}\int_{(0,1)} e^{\sqrt{-1}(\xi_k - \xi)xt}\,\lambda(dt)$$

という表現を持つ。これは $\sqrt{-1}xe^{\sqrt{-1}\xi x}$ に収束する。またその絶対値は $|x|$ で抑えられる。仮定より $\int_{\mathbb{R}} |x|\,\mu(dx) < +\infty$ であるから、Lebesgue の収束定理が適用でき

$$\lim_{k\to\infty}\frac{\int_{\mathbb{R}} \exp\{\sqrt{-1}\xi_k x\}\,\mu(dx) - \int_{\mathbb{R}} \exp\{\sqrt{-1}\xi x\}\,\mu(dx)}{\xi_k - \xi} = \sqrt{-1}\int_{\mathbb{R}} x e^{\sqrt{-1}\xi x}\,\mu(dx)$$

と収束することが分かる。以上は ξ に収束する任意の点列にあてはまるので $\mathrm{char}(\mu,\cdot)$ は ξ において微分可能であって、右辺が微分係数を与える。以下帰納法により、$\mathrm{char}(\mu,\cdot)$ は n 回微分可能であって、ξ における n 階微分係数は

$$(\sqrt{-1})^n \int_{\mathbb{R}} x^n e^{\sqrt{-1}\xi x}\,\mu(dx)$$

であることが導ける。また n 階導関数は連続である。 □

補題10.2(ii)の逆の主張もある意味で成り立つが、その前提条件には注意を払う必要がある。

10.3 補題. μ を $(\mathbb{R}, \mathrm{Borel}(\mathbb{R}))$ 上の確率測度、$n \in \mathbb{N}$ を偶数とする。 関数 $\xi \mapsto \mathrm{Re}\,\mathrm{char}(\mu, \xi)$ が0のある近傍で C^n 級なら $\int_{\mathbb{R}} |x|^n \mu(dx) < +\infty$ が成り立つ。

証明. 一般に f を原点0の δ 近傍 $(-\delta, \delta)$ で定義された C^n 級関数とする。Taylorの定理より

$$\lim_{t \to 0} \frac{1}{t^n} \sum_{i=0}^{n} \binom{n}{i} (-1)^k f((n/2 - i)t) = f^{(n)}(0)$$

が成り立つ。ここで ξ_k を0に収束する点列とする。n は偶数であるから

$$\sum_{i=0}^{n} \binom{n}{i} (-1)^i \exp\{\sqrt{-1}(n/2 - i)\xi_k x\} = \{2\sqrt{-1}\sin(\xi_k x/2)\}^n$$

は実数値である。Fatouの補題により $\int_{\mathbb{R}} x^n \mu(dx)$ は次で抑えられる。

$$\liminf_{k \to \infty} \frac{(-1)^{n/2}}{(\xi_k)^n} \sum_{i=0}^{n} \binom{n}{i} (-1)^i \mathrm{Re}\,\mathrm{char}(\mu, (n/2 - i)\xi_k).$$

さて n 回連続微分可能性により上の極限は関数 $\xi \mapsto \mathrm{Re}\,\mathrm{char}(\mu, \xi)$ の0における n 階微分係数に等しい。従って $\int_{\mathbb{R}} x^n \mu(dx) < +\infty$ が成り立つ。 □

次の補題では、依然として共分散 C は正定値対称行列である。

10.4 補題. α を $d \times n$ 行列, $\beta \in \mathbb{R}^d$ とし、X を平均 a 共分散 C の n 次元正規分布に従う確率変数とする。このとき各 $z \in \mathbb{C}^d$ に対して $\exp\{(z, \alpha X + \beta)\}$ は可積分であり次が成り立つ。

$$E[\exp\{(z, \alpha X + \beta)\}| = \exp\{(z, \alpha a + \beta) + (\alpha C {}^{\mathrm{t}}\alpha)(z, z)/2\}.$$

ここで $(\alpha C {}^{\mathrm{t}}\alpha)(\cdot, \cdot)$ は行列 $\alpha C {}^{\mathrm{t}}\alpha$ が誘導する2次形式を表す。

証明. $\exp\{(z, \alpha X + \beta)\} = \exp\{({}^{\mathrm{t}}\alpha z, X)\} \exp\{(z, \beta)\}$ であるから補題9.21に帰着する。 □

d 次元確率変数 $\alpha X + \beta$ の特性関数は形式的に平均 $\alpha a + \beta$ 共分散 $\alpha C {}^{\mathrm{t}}\alpha$ の正規分布のそれである。ここで形式的にといったのは行列 $\alpha C {}^{\mathrm{t}}\alpha$ は対称であるが必ずしも正定値ではない、従って例9.1で述べた定義の守備範囲ではないからである。二つの問題が明らかになった。

- 一般に非負定値な n 次対称行列 C と $a \in \mathbb{R}^n$ が与えられたとき、特性関数が $\xi \mapsto \exp\{\sqrt{-1}(\xi, a) - C(\xi, \xi)/2\}$ であるような確率測度は存在するか？
- 存在するとしてそのような確率測度は特性関数だけで特徴づけられるか？

１番目の方については今すぐ解決することができる。

10.5 定理. 非負定値な n 次対称行列 C と $a \in \mathbb{R}^n$ が与えられたとき、

$$\mathrm{char}(\mu, \xi) = \exp\{\sqrt{-1}(\xi, a) - C(\xi, \xi)/2\} \ \forall \xi \in \mathbb{R}^n$$

を満たす $(\mathbb{R}^n, \mathrm{Borel}(\mathbb{R}^n))$ 上の確率測度 μ が存在する。

証明. アイデァは補題 9.21 と類似である。即ち n 次行列 α で $C = \alpha\,{}^{t}\alpha$ を満たすものが存在する。ただし α は逆行列を持つとは限らない。ν を n 次元標準正規分布とする。補題 10.4 を適用して写像 $x \mapsto \alpha x + a$ による ν の像測度が求めるものであることがわかる。 □

そこで正規分布を改めて定義し直すことにする。

10.6 定義. 非負定値な n 次対称行列 C と $a \in \mathbb{R}^n$ が与えられたとする。平均 a 共分散 C の **n 次元正規分布**あるいは **n 次元 *Gauss* 分布**とは

$$\xi \mapsto \exp\{\sqrt{-1}(\xi, a) - C(\xi, \xi)/2\}$$

を特性関数とする $(\mathbb{R}^n, \mathrm{Borel}(\mathbb{R}^n))$ 上の確率測度をいう。

このように定義すると補題 10.4 は次のような言い換えを持つこととなる。

10.7 定理. α を $d \times n$ 行列, $\beta \in \mathbb{R}^d$ とする。アファイン写像 $x \mapsto \alpha x + \beta$ による平均 a 共分散 C の n 次元正規分布の像測度は平均 $\alpha a + \beta$ 共分散 $\alpha C\,{}^{t}\alpha$ の d 次元正規分布である。

10.8 定理. $X = (X_1, X_2, \ldots, X_n)$ をその結合分布が平均 a 共分散 C の n 次元正規分布に従う n 次元確率変数とする。このとき各 X_i はすべての次数のモーメントをもち、

$$E[X_i] = a \text{ の第 } i \text{ 成分、} \mathrm{Cov}[X_i, X_j] = C \text{ の第 } ij \text{ 成分}$$

が成り立つ（即ち X の平均ベクトルは a で共分散行列は C である)。

証明. 正規分布の特性関数は明らかに無限回微分可能である。従って補題 10.3 により各 X_i はすべての次数のモーメントを持つことになる。あとは原点 0 における偏微分係数を計算して補題 10.2(ii) を適用すればよい。 □

だがこれだけでは単なるその場しのぎである。根本的な解決は特性関数が確率測度を一意に決定することを示してはかられる。以下 Fourier 変換の理論を少々展開する。

記号

$\mathcal{L}^1(\mathbb{R}^d)$: $\mathrm{Borel}(\mathbb{R}^d)$ 可測かつ $\lambda^{(d)}$ 可積分な関数 $\mathbb{R}^d \to \mathbb{C}$ の全体

10.9 定義. $f \in \mathcal{L}^1(\mathbb{R}^d)$ に対して次の関数を f の ***Fourier* 変換**(Fourier transform) という。

$$\widehat{f} : \mathbb{R}^d \to \mathbb{C},\ \xi \mapsto \int_{\mathbb{R}^d} f(x) e^{-\sqrt{-1}(\xi, x)} \lambda^{(d)}(dx).$$

10.10 補題. (i) $f \in \mathcal{L}^1(\mathbb{R}^d)$ の Fourier 変換 $\widehat{f}$ は有界かつ連続である。
(ii) $f \in \mathcal{L}^1(\mathbb{R}^d)$ の Fourier 変換 $\widehat{f}$ は $\lim_{\xi \to \infty} \widehat{f}(\xi) = 0$ を満たす（***Riemann-Lebesgue* の定理**)。

10.11 演習問題. 補題 10.10(i) を示せ。

10.12 注意. 補題 10.10(ii) を示すには工夫が必要である。

10.13 補題. p を d 次元熱核とする。$t>0$ と $f\in\mathcal{L}^1(\mathbb{R}^d)$ に対して以下が成り立つ。

(i) $$\frac{1}{(2\pi)^d}\int_{\mathbb{R}^d}\widehat{f}(\xi)\exp\{\sqrt{-1}(\xi,y)-t\|\xi\|^2\}\,\lambda^{(d)}(d\xi)=\int_{\mathbb{R}^d}p(t,x-y)f(x)\,\lambda^{(d)}(dx)\ \forall y\in\mathbb{R}^d.$$

(ii) 関数 $\mathbb{R}^d\to\mathbb{C}, y\mapsto\displaystyle\int_{\mathbb{R}^d}p(t,x-y)f(x)\,\lambda^{(d)}(dx)$ は有界かつ一様連続である。

(iii) $(\mathbb{R}^d,\mathrm{Borel}(\mathbb{R}^d))$ 上の確率測度 μ に対して次が成り立つ。

$$\frac{1}{(2\pi)^d}\int_{\mathbb{R}^d}\widehat{f}(\xi)\mathrm{char}(\mu,\xi)\exp\{-t\|\xi\|^2\}\,\lambda^{(d)}(d\xi)=\int_{\mathbb{R}^d}\Big(\int_{\mathbb{R}^d}p(t,x-y)f(x)\,\lambda^{(d)}(dx)\Big)\mu(dy).$$

証明. (i) 関数 $\mathbb{R}^d\times\mathbb{R}^d\to\mathbb{C}$, $(x,\xi)\mapsto f(x)\exp\{\sqrt{-1}(\xi,y-x)-t\|\xi\|^2\}$ は $\lambda^{(d)}\otimes\lambda^{(d)}$ 可積分である。よって Fubini の定理と系 9.22(i) を適用して結論を得る。

(iii) 補題 10.10(i) により関数 $\mathbb{R}^d\times\mathbb{R}^d\to\mathbb{C}$, $(\xi,y)\mapsto\widehat{f}(\xi)\exp\{\sqrt{-1}(\xi,y)-t\|\xi\|^2\}$ は連続かつ $\lambda^{(d)}\otimes\mu$ 可積分である。従って Fubini の定理を適用して次の関係を得る。

$$\begin{aligned}&\int_{\mathbb{R}^d}\widehat{f}(\xi)\mathrm{char}(\mu,\xi)\exp\{-t\|\xi\|^2\}\,\lambda^{(d)}(d\xi)\\&=\int_{\mathbb{R}^d}\Big(\int_{\mathbb{R}^d}\widehat{f}(\xi)\exp\{\sqrt{-1}(\xi,y)-t\|\xi\|^2\}\,\lambda^{(d)}(d\xi)\Big)\mu(dy).\end{aligned}$$

ゆえに (i) により求める結論を得る。 □

10.14 演習問題. 補題 10.13(ii) を示せ。

当面の目的のためには演習問題 4.10(iii) で述べた結果と Lebesgue の優収束定理だけで間に合うという意味で次の命題は必要ないが応用上重要なのでふれておく。

10.15 補題. p を d 次元熱核とする。有界かつ一様連続な関数 $f:\mathbb{R}^d\to\mathbb{C}$ に対して $t>0$ をパラメータとする関数族 $y\mapsto\int_{\mathbb{R}^d}p(t,x-y)f(x)\,\lambda^{(d)}(dx)$ は $t\downarrow 0$ のとき f に一様収束する。

証明. $t>0$ と $y\in\mathbb{R}^d$ を固定する。平均が y で共分散が d 次単位行列の $2t$ 倍の d 次元正規分布に従う確率変数 X に対して次が成り立つ。

$$\int_{\mathbb{R}^d}p(t,x-y)f(x)\,\lambda^{(d)}(dx)=E[f(X)],\ \int_{\|x\|\geq\delta}p(t,x)\,\lambda^{(d)}(dx)=P(\|X-y\|\geq\delta)$$

従って補題 6.43 より各 $\delta>0$ に対して評価

$$\begin{aligned}&\Big|\int_{\mathbb{R}^d}p(t,x-y)f(x)\,\lambda^{(d)}(dx)-f(y)\Big|\\&\leq\sup_{x,z:\|x-z\|<\delta}|f(x)-f(z)|+2\sup_{x\in\mathbb{R}^d}|f(x)|\int_{\|x\|\geq\delta}p(t,x)\,\lambda^{(d)}(dx).\end{aligned}$$

を得る。右辺は y に依存しない。$\delta>0$ を固定するとき、右辺第 2 項は $t\downarrow 0$ の極限で 0 に収束するがこれを示すのは演習問題 10.16 に委ねる。よって

$$\limsup_{t\downarrow 0}\sup_{y\in\mathbb{R}^d}\Big|\int_{\mathbb{R}^d}p(t,x-y)f(x)\,\lambda^{(d)}(dx)-f(y)\Big|\leq\sup_{x,z:\|x-z\|<\delta}|f(x)-f(z)|.$$

$\delta\downarrow 0$ のとき右辺は 0 に収束するので結論に到達した。 □

10.16 演習問題. 各 $\delta > 0$ に対し $\lim_{t\downarrow 0}\int_{\|x\|\geq\delta} p(t,x)\,\lambda^{(d)}(dx) = 0$ が成り立つことを示せ。

特性関数に基づく**測度の一意性定理**を以下にあげる。

10.17 定理. $(\mathbb{R}^d, \mathrm{Borel}(\mathbb{R}^d))$ 上の確率測度 μ, ν に対して $\mu = \nu$ となるための必要十分条件はそれらの特性関数が一致することである。

証明. $\mathrm{char}(\mu,\xi) = \mathrm{char}(\nu,\xi)\ \forall\xi\in\mathbb{R}^d$ と仮定する。$f \in C_0(\mathbb{R}^d)$ を固定する。関数 f は $\lambda^{(d)}$ 可積分なので補題 10.13(iii) により各 $t > 0$ に対して次が成り立つ。

$$\int_{\mathbb{R}^d}\Big(\int_{\mathbb{R}^d} p(t,x-y)f(x)\,\lambda^{(d)}(dx)\Big)\mu(dy) = \int_{\mathbb{R}^d}\Big(\int_{\mathbb{R}^d} p(t,x-y)f(x)\,\lambda^{(d)}(dx)\Big)\nu(dy).$$

ここで p は d 次元熱核である。関数 f は一様連続であるから、補題 10.15 を適用して、$t\downarrow 0$ の極限で左辺は $\int_{\mathbb{R}^d} f\,\mu$ に右辺は $\int_{\mathbb{R}^d} f\,\nu$ にそれぞれ収束することがわかる。従って

$$\int_{\mathbb{R}^d} f\,\mu = \int_{\mathbb{R}^d} f\,\nu\ \forall f\in C_0(\mathbb{R}^d).$$

定理 7.22 によればこれは $\mu = \nu$ を意味している。 □

10.18 例. $n \in \mathbb{N}$, $0 < p < 1$ とする。実確率変数系 X_1, X_2, ..., X_n が独立であり、それぞれ分布 $(1-p)\delta_0 + p\delta_1$ に従うなら $X_1 + X_2 + \cdots + X_n$ はパラメータ n, p の二項分布に従う。

証明. 分布 $(1-p)\delta_0 + p\delta_1$ の特性関数は $\xi \mapsto 1 - p + pe^{\sqrt{-1}\xi}$ である。系 8.34 により

$$\mathrm{char}(X_1 + X_2 + \cdots + X_n, \xi) = \prod_{i=1}^{n}\mathrm{char}(X_i,\xi) = (1 - p + pe^{\sqrt{-1}\xi})^n$$

右辺はパラメータ n, p の二項分布の特性関数を表すので、定理 10.17 より結論を得る。 □

10.19 定理. 実確率変数系 X_1, X_2, ..., X_n が独立であるのは次と同値である。

$$\mathrm{char}((X_1, X_2, \ldots, X_n), (\xi_1, \xi_2, \ldots, \xi_n)) = \prod_{i=1}^{n}\mathrm{char}(X_i,\xi_i)\ \forall(\xi_1,\xi_2,\ldots,\xi_n)\in\mathbb{R}^n.$$

証明. 系 8.34 と定理 10.17 を適用する点を除けば定理 8.36 と同じ議論である。 □

10.20 例. $X_1, X_2, \ldots, X_n$ をその結合分布が n 次元正規分布に従う確率変数系とする。このとき $X_1, X_2, \ldots, X_n$ が独立であるための必要十分条件は $\mathrm{Cov}[X_i, X_j] = 0\ i \neq j$ である。

10.21 演習問題. 例 10.20 を示せ。

例 8.24 で述べた Fourier 逆変換公式の観点から定理 10.17 を証明することができる。

10.22 補題. $f \in C_0^2(\mathbb{R})$ の Fourier 変換 $\widehat{f}$ は λ 可積分であり *Fourier* **逆変換公式**が成り立つ。

$$\frac{1}{2\pi}\int_{\mathbb{R}} e^{\sqrt{-1}\xi x}\widehat{f}(\xi)\,\lambda(d\xi) = f(x)\ \forall x\in\mathbb{R}.$$

証明. 演習問題 8.23(ii) で調べたように関数 $\xi \mapsto \displaystyle\int_{\mathbb{R}} e^{-\sqrt{-1}\xi x} f(y)\,\lambda(dy)$ は連続かつ λ 可積分である。よって例 8.24 を適用して Fourier 逆変換公式をえる。 □

10.23 演習問題. 補題 10.22 を使って定理 10.17 の別証明を与えよ。

11　ランダムウォークと中心極限定理

ランダム要因の積み重なりの時間経過を表現するのがランダムウォークであり、実際の観測ではミクロ的には長時間の振る舞いをマクロ時間にスケールし直したものが誤差として現れるわけである。この節では 2 乗可積分な増分を持つランダムウォークについての中心極限定理を議論する。その際、特性関数の利用が見通しのよい展開を提供する。

11.1 定義. d 次元確率変数の無限系 $X_k;\ k \in \mathbb{N}\cup\{0\}$ を d 次元（離散時間）**確率過程**(stochastic process) という。$X_0,\ X_1 - X_0,\ \ldots,\ X_n - X_{n-1},\ \ldots$ が独立かつ $X_k - X_{k-1}$ $k \in \mathbb{N}$ が同じ分布に従うときそれを d 次元**ランダムウォーク**(random walk) という。

演習問題 6.30 で確かめたように 2 乗可積分な確率変数は可積分でもある。

前提

X_n は $\mathbb{R}$ 上のランダムウォーク, $X_0 = 0$, $E[|X_1|^2] < +\infty$, $m := E[X_1]$, $v := \mathrm{Var}[X_1] > 0$.

11.2 補題. (i) $E[X_n] = mn$, $\mathrm{Var}[X_n] = vn$ $\forall n \in \mathbb{N}$.
(ii) $\mathrm{char}((X_n - mn)/\sqrt{vn}, \xi) = \big(\mathrm{char}(X_1 - m, \xi/\sqrt{vn})\big)^n$ $\forall n \in \mathbb{N}$, $\forall \xi \in \mathbb{R}$.

11.3 演習問題. 補題 11.2 を示せ。

特に各 $n \in \mathbb{N}$ に対して $(X_n - mn)/\sqrt{vn}$ は平均 0 分散 1 である。

11.4 補題. 複素数列 a_n に対して na_n が α に収束するとき $(1 + a_n)^n$ は e^α に収束する。

証明. $z \in \mathbb{C}$, $n \in \mathbb{N}$ とする。このとき次の評価が成り立つ。

$$|(1+z)^n - 1| \le \sum_{k=1}^{n} \binom{n}{k} |z|^k = (1+|z|)^n - 1 \le e^{n|z|} - 1.$$

これを $z = (1 + a_n)e^{-\alpha/n} - 1$ に対して適用して

$$|(1 + a_n)^n e^{-\alpha} - 1| \le \exp\{n|(1 + a_n)e^{-\alpha/n} - 1|\} - 1\ \forall n \in \mathbb{N}$$

が導ける。ここで $n|(1 + a_n)e^{-\alpha/n} - 1|$ を上から次のものでおさえる。

$$|na_n e^{-\alpha/n} - na_n| + |na_n - \alpha| + |\alpha + ne^{-\alpha/n} - n|$$

各項はいずれも $n \to \infty$ の極限で 0 に収束する。よって $(1 + a_n)^n e^{-\alpha}$ は 1 に収束する。 □

11.5 例. X_1 は分布 $\frac{1}{2}\delta_1 + \frac{1}{2}\delta_{-1}$ に従うとする。このとき $E[X_1] = 0$, $\mathrm{Var}[X_1] = 1$ であり、$\mathrm{char}(X_1, \xi) = \cos\xi$ である。従って

$$\mathrm{char}((X_n - mn)/\sqrt{vn}, \xi) = (\mathrm{char}(X_1, \xi/\sqrt{n})))^n = (\cos(\xi/\sqrt{n}))^n$$

である。さて数列 $n(\cos(\xi/\sqrt{n}) - 1)$ は $-\xi^2/2$ に収束する。よって補題 11.4 を適用して上は関数 $\xi \mapsto \exp\{-\xi^2/2\}$ に各点収束することが分かる。

関数 $\xi \mapsto \exp\{-\xi^2/2\}$ は標準正規分布の特性関数である。定理 10.17 によれば特性関数は分布を一意に特徴づける。従って上の例では何らかの意味で $(Y_n - mn)/\sqrt{vn}$ の分布が標準正規分布に収束していることになる。まず特性関数の収束を一般的に取り扱う。

11.6 補題. 各 $\xi \in \mathbb{R}$ に対して $\mathrm{char}((X_n - mn)/\sqrt{vn}, \xi)$ は $\exp\{-\xi^2/2\}$ に収束する。

証明. 補題 11.2(ii) により、$\bigl(\mathrm{char}(X_1 - m, \xi/\sqrt{vn})\bigr)^n$ が $e^{-\xi^2/2}$ に収束することを示せばよい。一般性を失うことなしに $E[X_1] = 0$, $\mathrm{Var}[X_1] = 1$ であるとしても差し支えない。

$$e^z = 1 + z + \int_{(0,1)} (1-t) e^{tz}\, \lambda(dt)\, z^2 \ \forall z \in \mathbb{C}$$

を利用して次を得る。

$$\mathrm{char}(X_1, \xi) - 1 = E[\sqrt{-1}\xi X_1 - \int_{(0,1)} (1-t) \exp\{t\sqrt{-1}\xi X_1\}\, \lambda(dt)\, \xi^2 {X_1}^2].$$

さて $E[X_1] = 0$ である。また $|(1-t)e^{t\sqrt{-1}\xi X_1} {X_1}^2| \le |X_1|^2$ であるから $\lambda \otimes P$ に関する積分に Fubini の定理が適用できる。したがって ξ に $\xi/\sqrt{n}$ を代入して次を得る。

$$\begin{aligned} &n\bigl(\mathrm{char}(X_1, \xi/\sqrt{n}) - 1\bigr) \\ &= -\int_{(0,1)\times\Omega} (1-t) \exp\{t\sqrt{-1}\xi X_1(\omega)/\sqrt{n}\} X_1(\omega)^2\, (\lambda \otimes P)(dt, d\omega)\, \xi^2 \end{aligned}$$

ここで $n(\xi/\sqrt{n})^2 = \xi^2$ という関係が単純だが重要である。さて $(0,1) \times \Omega$ 上の関数列

$$(t, \omega) \mapsto (1-t) \exp\{t\sqrt{-1}\xi X_1(\omega)/\sqrt{n}\} X_1(\omega)^2$$

は関数 $(t, \omega) \mapsto (1-t) X_1(\omega)^2$ に各点収束している。他方 $|X_1|^2$ は $\lambda \otimes P$ について可積分な優関数である。よって Lebesgue の収束定理が適用できるので各 $\xi \in \mathbb{R}$ に対して

$n\bigl(\mathrm{char}(X_1, \xi/\sqrt{n}) - 1\bigr)$ は $-\int_{(0,1)\times\Omega} (1-t) X_1(\omega)^2\, (\lambda \otimes P)(dt, d\omega) \xi^2 = -\xi^2/2$ に収束する。

故に補題 11.4 により $\bigl(\mathrm{char}(X_1, \xi/\sqrt{n})\bigr)^n$ は $e^{-\xi^2/2}$ に収束する。 □

11.7 例. $(\mathbb{R}, \mathrm{Borel}(\mathbb{R}))$ 上の確率測度 μ が $(\mu \otimes \mu)(\{(x,y) \in \mathbb{R}^2 : (x+y)/\sqrt{2} \in A\}) = \mu(A)$ $\forall A \in \mathrm{Borel}(\mathbb{R})$ かつ $\int_{\mathbb{R}} x^2\, \mu(dx) = 1$ を満たすなら μ は標準正規分布である。

証明. まず $\int_{\mathbb{R}} x^2\, \mu(dx) = 1$ から $\int_{\mathbb{R}} |x|\, \mu(dx) < +\infty$ が導ける。定理 6.20 を適用して

$$\int_{\mathbb{R}^2} \frac{x+y}{\sqrt{2}}\, (\mu \otimes \mu)(dx, dy) = \int_{\mathbb{R}} x\, \mu(dx)$$

が最初の条件から導ける。Fubini の定理（定理 8.14）により左辺は $\sqrt{2} \int_{\mathbb{R}} x\, \mu(dx)$ に等しいので

$$\int_{\mathbb{R}} x\, \mu(dx) = 0$$

である。よって補題 11.6 の証明から次がわかる。

各 $\xi \in \mathbb{R}$ に対して $\mathrm{char}(\mu, \xi/\sqrt{n})^n$ は $e^{-\xi^2/2}$ に収束する。

同様に最初の条件により $\mathrm{char}(\mu \otimes \mu, (\xi/\sqrt{2}, \xi/\sqrt{2})) = \mathrm{char}(\mu, \xi)$ である。左辺は定理 10.19 により $\mathrm{char}(\mu, \xi/\sqrt{2})^2$ に等しい。得られた関係式を n 回繰り返し適用することにより

$$\mathrm{char}(\mu, \xi/\sqrt{2^n})^{2^n} = \mathrm{char}(\mu, \xi) \ \forall n \in \mathbb{N}\, \forall \xi \in \mathbb{R}$$

を得る。左辺は $e^{-\xi^2/2}$ に収束していたので $\mathrm{char}(\mu, \xi) = e^{-\xi^2/2}$ が導けた。 □

特性関数の収束をもう少し直接的に表現しよう。

確率測度の収束に関する前提

μ_n を $(\mathbb{R}^d, \mathrm{Borel}(\mathbb{R}^d))$ 上の確率測度の列、μ を $(\mathbb{R}^d, \mathrm{Borel}(\mathbb{R}^d))$ 上の確率測度とする。

11.8 補題. $\lim_{n\to\infty} \mathrm{char}(\mu_n, \xi) = \mathrm{char}(\mu, \xi) \ \forall \xi \in \mathbb{R}^d$ なら $\lim_{n\to\infty} \int_{\mathbb{R}^d} f\, \mu_n = \int_{\mathbb{R}^d} f\, \mu \ \forall f \in C_0(\mathbb{R}^d)$

証明. 関数 p を d 次元熱核とし、$f \in C_0(\mathbb{R}^d)$ を固定して関数

$$u : (0, +\infty) \times \mathbb{R}^d \to \mathbb{R}, (t, y) \mapsto \int_{\mathbb{R}^d} p(t, x-y) f(x)\, \lambda^{(d)}(dx)$$

を導入する。補題 10.13(iii) より $t > 0$ と $(\mathbb{R}^d, \mathrm{Borel}(\mathbb{R}^d))$ 上の確率測度 ν に対して

$$\frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \widehat{f}(\xi) \mathrm{char}(\nu, \xi) \exp\{-t\|\xi\|^2\}\, \lambda^{(d)}(d\xi) = \int_{\mathbb{R}^d} u(t, y) \nu(dy).$$

が満たされる。補題 10.10(i) によりある $M < +\infty$ が存在して $|\widehat{f}(\xi)| \le M \ \forall \xi \in \mathbb{R}^d$ であり、また補題 10.2(i) により $|\mathrm{char}(\nu, \xi)| \le 1$ である。さて各 $\xi \in \mathbb{R}^d$ に対して

$$\lim_{n\to\infty} \widehat{f}(\xi) \mathrm{char}(\mu_n, \xi) \exp\{-t\|\xi\|^2\} = \widehat{f}(\xi) \mathrm{char}(\mu, \xi) \exp\{-t\|\xi\|^2\}$$

よって $\lambda^{(d)}$ 可積分関数 $\xi \mapsto M \exp\{-t\|\xi\|^2\}$ を優関数に選んで Lebesgue の収束定理を適用すると

$$\lim_{n\to\infty} \int_{\mathbb{R}^d} u(t, y) \mu_n(dy) = \int_{\mathbb{R}^d} u(t, y) \mu(dy) \ \forall t > 0$$

となることがわかる。次に $\int_{\mathbb{R}^d} f \mu_n$ を以下のように評価する。

$$\int_{\mathbb{R}^d} u(t, y) \mu_n(dy) + \int_{\mathbb{R}^d} (f(y) - u(t, y)) \mu_n(dy) \le \int_{\mathbb{R}^d} u(t, y) \mu_n(dy) + \sup_{y \in \mathbb{R}^d} |f(y) - u(t, y)|.$$

右辺第 2 項は n に依存しない。従って

$$\limsup_{n\to\infty} \int_{\mathbb{R}^d} f \mu_n \le \int_{\mathbb{R}^d} u(t, y) \mu(dy) + \sup_{y \in \mathbb{R}^d} |f(y) - u(t, y)|.$$

補題 10.15 によれば、$t \downarrow 0$ の極限で $u(t, \cdot)$ は f に一様収束するので

$$\limsup_{n\to\infty} \int_{\mathbb{R}^d} f \mu_n \le \int_{\mathbb{R}^d} f \mu.$$

$-f$ にも上の議論は適用されるので $\int_{\mathbb{R}^d} f \mu \le \liminf_{n\to\infty} \int_{\mathbb{R}^d} f \mu_n$ も成り立つ。 □

11.9 定義. $\mathbb{R}^d$ 上の Radon 測度の列 μ_n が $\mathbb{R}^d$ 上の Radon 測度 μ に**漠収束**(vague convergence) するとは任意の $f \in C_0(\mathbb{R}^d)$ に対して $\int_{\mathbb{R}^d} f\,\mu_n$ が $\int_{\mathbb{R}^d} f\,\mu$ に収束することをいう。

11.10 補題. 列 μ_n が μ に漠収束するなら有界な $f \in C(\mathbb{R}^d)$ に対し $\displaystyle\lim_{n\to\infty} \int_{\mathbb{R}^d} f\,\mu_n = \int_{\mathbb{R}^d} f\,\mu$

証明. $g \in C_0(\mathbb{R}^d)$, $g(x) \le 1\ \forall x \in \mathbb{R}^d$ とする。まず $M := \sup_{y\in\mathbb{R}^d} |f(y)| < +\infty$ に注意して

$$\int_{\mathbb{R}^d} f\,\mu_n = \int_{\mathbb{R}^d} fg\,\mu_n + \int_{\mathbb{R}^d} f(1-g)\,\mu_n \le \int_{\mathbb{R}^d} fg\,\mu_n + M\int_{\mathbb{R}^d} (1-g)\,\mu_n.$$

と評価する。さて $fg \in C_0(\mathbb{R}^d)$ であるから、右辺第 1 項は $\int_{\mathbb{R}^d} fg\,\mu$ に収束する。他方

$$\int_{\mathbb{R}^d}(1-g)\,\mu_n = 1 - \int_{\mathbb{R}^d} g\,\mu_n, \int_{\mathbb{R}^d}(1-g)\,\mu = 1 - \int_{\mathbb{R}^d} g\,\mu,$$

であるから $\int_{\mathbb{R}^d}(1-g)\,\mu_n$ は $\int_{\mathbb{R}^d}(1-g)\,\mu$ に収束する。従って

$$\limsup_{n\to\infty} \int_{\mathbb{R}^d} f\,\mu_n \le \int_{\mathbb{R}^d} fg\,\mu + M\int_{\mathbb{R}^d}(1-g)\,\mu.$$

補題 11.11 によれば、$C_0(\mathbb{R}^d)$ の列 g_k であって次を満たすものを選ぶことができる。

$$g_k \le g_{k+1}\ \forall k \in \mathbb{N} \text{ かつ } \sup_{k\in\mathbb{N}} g_k(x) = 1\ \forall x \in \mathbb{R}^d$$

Lebesgue の収束定理より $\int_{\mathbb{R}^d} fg_k\,\mu$ は $\int_{\mathbb{R}^d} f\,\mu$ に収束し、また $\int_{\mathbb{R}^d}(1-g_k)\,\mu$ は 0 に収束する。従って

$$\text{有界な } f \in C(\mathbb{R}^d) \text{ に対し } \limsup_{n\to\infty} \int_{\mathbb{R}^d} f\,\mu_n \le \int_{\mathbb{R}^d} f\,\mu.$$

$-f$ にも上の議論は適用されるので $\int_{\mathbb{R}^d} f\mu \le \liminf_{n\to\infty} \int_{\mathbb{R}^d} f\mu_n$ も成り立つ。 □

11.11 補題. $\mathbb{R}^d$ の開集合 A に対して $C_0(\mathbb{R}^d)$ の列 g_k であって次を満たすものが存在する。

$$0 \le g_k \le g_{k+1}\ \forall k \in \mathbb{N} \text{ かつ } \sup_{k\in\mathbb{N}} g_k(x) = 1\ \forall x \in \mathbb{R}^d$$

11.12 演習問題. 補題 11.11 を示せ。

11.13 定義. μ を $(\mathbb{R}^d, \mathrm{Borel}(\mathbb{R}^d))$ 上の確率測度とする。
(i) $(\mathbb{R}^d, \mathrm{Borel}(\mathbb{R}^d))$ 上の確率測度の列 μ_n が μ に**弱収束**(weak convergence) するとは任意の有界な $f \in C(\mathbb{R}^d)$ に対して $\int_{\mathbb{R}^d} f\,\mu_n$ が $\int_{\mathbb{R}^d} f\,\mu$ に収束することをいう。
(ii) $\mathbb{R}^d$ 値確率変数の列 X_n が μ に**分布収束**する、あるいは**法則収束**(convergence in law) するとは確率分布の列 $\mathcal{L}(X_n, \cdot)$ が μ に弱収束することをいう。

このノートでは $C(\mathbb{R}^d)$ の各元は実数値であったことに注意。

11.14 定理. 以下は互いに同値である。

- 1° 特性関数の列 $\mathrm{char}(\mu_n, \cdot)$ が $\mathrm{char}(\mu, \cdot)$ に各点収束する。
- 2° μ_n が μ に漠収束する。
- 3° μ_n が μ に弱収束する。

証明. 補題 11.8 と補題 11.10 により $1^\circ \Rightarrow 2^\circ \Rightarrow 3^\circ$ という関係であることがわかっている。さて $x \mapsto \cos(\xi, x)$, $x \mapsto \sin(\xi, x)$ はともに有界連続関数である。従って 3° を仮定すると $\mathrm{char}(\mu_n, \cdot)$ の実部、虚部が $\mathrm{char}(\mu, \cdot)$ の実部、虚部にそれぞれ各点収束することになる。従って $3^\circ \Rightarrow 1^\circ$ という関係も成り立つ。 □

11.15 注意. 漠収束と弱収束の違いについては補題 12.1 以降の記述が参考になるであろう。

以下では弱収束あるいは法則収束をわかりやすく表現する。

11.16 定理. μ_n が μ に弱収束することと以下は同値である。

$$\text{任意の } \mathbb{R}^d \text{ の開集合 } A \text{ に対して } \liminf_{n\to\infty} \mu_n(A) \geq \mu(A)$$

証明. μ_n が μ に弱収束するとしよう。$\mathbb{R}^d$ の開集合 A に対して、補題 11.11 より $g_k \leq g_{k+1}$ $\forall k \in \mathbb{N}$ かつ $\sup_{k\in\mathbb{N}} g_k(x) = 1_A(x)$ $\forall x \in \mathbb{R}^d$ を満たす $C_0(\mathbb{R}^d)$ の列 g_k が存在する。このとき

$$\liminf_{n\to\infty} \mu_n(A) \geq \liminf_{n\to\infty} \int_{\mathbb{R}^d} g_k\, \mu_n = \int_{\mathbb{R}^d} g_k\, \mu \ \forall k \in \mathbb{N}.$$

単調収束定理により右辺は $\mu(A)$ に収束する。

逆に $\liminf_{n\to\infty} \mu_n(A) \geq \mu(A)$ $\forall A$ open としよう。まず非負値の $f \in C(\mathbb{R}^d)$ について検討する。各 $\alpha \in \mathbb{R}$ に対して $A(\alpha) := \{x \in \mathbb{R}^d : f(x) > \alpha\}$ は開集合であるから仮定より直ちに $\liminf_{n\to\infty} \mu_n(A(\alpha)) \geq \mu(A(\alpha))$ が従う。ここで非負値 Borel 関数列

$$g_k(x) := \frac{\#\{i \in \mathbb{N} : i < 2^k f(x)\}}{2^k} = \frac{1}{2^k} \sum_{i=1}^{\infty} 1_{A(i/2^k)}(x)$$

を導入する。級数和に Fatou の補題を適用して

$$\liminf_{n\to\infty} \int_{\mathbb{R}^d} g_k\, \mu_n \geq \frac{1}{2^k} \sum_{i=1}^{\infty} \liminf_{n\to\infty} \mu_n(A(i/2^k)) \geq \frac{1}{2^k} \sum_{i=1}^{\infty} \mu(A(i/2^k)) = \int_{\mathbb{R}^d} g_k\, \mu$$

を得る。さて $0 \leq g_k \leq g_{k+1}$ $\forall k \in \mathbb{N}$ かつ $\sup_{k\in\mathbb{N}} g_k(x) = f(x)$ $\forall x \in \mathbb{R}^d$ であるから、左辺は $\liminf_{n\to\infty} \int_{\mathbb{R}^d} f\, \mu_n$ で抑えられ、単調収束定理により右辺は $\int_{\mathbb{R}^d} f\, \mu$ に収束する。従って

$$\liminf_{n\to\infty} \int_{\mathbb{R}^d} f\, \mu_n \geq \int_{\mathbb{R}^d} f\, \mu.$$

非負値でなくてもある実数 M に対して $f + M$ が非負となるような $f \in C(\mathbb{R}^d)$ については

$$\liminf_{n\to\infty} \int_{\mathbb{R}^d} f\, \mu_n = \liminf_{n\to\infty} \int_{\mathbb{R}^d} (f+M)\, \mu_n - M \geq \int_{\mathbb{R}^d} (f+M)\, \mu - M = \int_{\mathbb{R}^d} f\, \mu$$

という議論ができる。従って有界な $f \in C(\mathbb{R}^d)$ については $-f$ にも同じ論法があてはまり

$$\limsup_{n\to\infty} \int_{\mathbb{R}^d} f\, \mu_n \leq \int_{\mathbb{R}^d} f\, \mu$$

も導くことができる。よって μ_n は μ に弱収束する。 □

11.17 系. μ_n が μ に弱収束することと以下は同値である。

$$\text{任意の } \mathbb{R}^d \text{ の閉集合 } A \text{ に対して } \limsup_{n\to\infty} \mu_n(A) \leq \mu(A)$$

証明. 閉集合 A に対してその補集合 A^c は開集合であり $\mu_n(A) = 1 - \mu_n(A^c)$ であることに注目すれば、定理 11.16 から直ちに導くことができる。 □

以上でこの節で取り扱う範囲に限ってではあるが中心極限定理を証明する準備が整った。

記号

Φ: 標準正規分布の分布関数

11.18 定理. 確率変数列 $(X_n - mn)/\sqrt{vn}$ は標準正規分布に法則収束する。とくに $a, b \in \mathbb{R}$ ただし $a < b$ に対して確率 $P(a < (X_n - mn)/\sqrt{vn} < b)$ は $\Phi(b) - \Phi(a)$ に収束する。

証明. 法則収束は補題 11.6 と定理 11.14 で述べられている。次に定理 11.16 と系 11.17 より

$$\int_{(a,b)} \frac{e^{-x^2/2}}{\sqrt{2\pi}}\,\lambda(dx) \leq \liminf_{n\to\infty} P(a < (X_n - mn)/\sqrt{vn} < b)$$
$$\leq \limsup_{n\to\infty} P(a \leq (X_n - mn)/\sqrt{vn} \leq b) \leq \int_{[a,b]} \frac{e^{-x^2/2}}{\sqrt{2\pi}}\,\lambda(dx)$$

最左辺と最右辺はともに $\Phi(b) - \Phi(a)$ に等しい。したがって 2 番目の結論も得られた。 □

定理 11.18 は広く**中心極限定理**(central limit theorem) と呼ばれるものの基本形である。

11.19 例. X_1 がパラメータ 1 の指数分布に従うとする。このとき

$$E[X_1] = 1, \mathrm{Var}[X_1] = 1$$

パラメータ 1 の指数分布とはパラメータ 1, 1 の gamma 分布に他ならない。以後、2 番目のパラメータは重要でない。例 9.7(iii) の助けを借りて帰納法を適用すると、X_n は指数 n の gamma 分布に従うことがわかる。従って次が成り立つ。

$$P(0 < (X_n - n)/\sqrt{n} < 1) = P(n < X_n < \sqrt{n} + n) = \frac{1}{\Gamma(n)} \int_{(n, \sqrt{n}+n)} y^{n-1} e^{-y}\,\lambda(dy).$$

ここで Γ は gamma 関数である。右辺の積分に変数変換公式（補題 3.21）を適用すると

$$\int_{(0,1)} (\sqrt{n}y + n)^{n-1} e^{-(\sqrt{n}y+n)} \sqrt{n}\,\lambda(dy) = n^{n-1} e^{-n} \sqrt{n} \int_{(0,1)} (1 + y/\sqrt{n})^{n-1} e^{-\sqrt{n}y}\,\lambda(dy).$$

従って各 $n \in \mathbb{N}$ に対して

$$\frac{\Gamma(n)}{n^{n-1} e^{-n} \sqrt{n}} = \frac{\int_{(0,1)} (1 + y/\sqrt{n})^{n-1} e^{-\sqrt{n}y}\,\lambda(dy)}{P(0 < (X_n - n)/\sqrt{n} < 1)}.$$

中心極限定理（定理 11.18）によれば右辺の分母は $\Phi(1) - \Phi(0)$ に収束する。他方、

$$\lim_{n\to\infty} (1 + y/\sqrt{n})^{n-1} e^{-\sqrt{n}y} = e^{-y^2/2}\ \forall y \in (0,1).$$

Lebesgue の収束定理を適用するために $0 < 1 + x \le e^x \ \forall x > 0$ から直ちに導かれる評価

$$0 < (1 + y/\sqrt{n})^{n-1} e^{-\sqrt{n}y} \le e^{(n-1)y/\sqrt{n}} e^{-\sqrt{n}y} \le 1 \ \forall y \in (0,1) \ \forall n \in \mathbb{N}$$

を使う。従って Lebesgue の収束定理により

$$\lim_{n\to\infty} \frac{\Gamma(n)}{n^{n-1}e^{-n}\sqrt{n}} = \frac{\int_{(0,1)} e^{-y^2/2}\,\lambda(dy)}{\Phi(1) - \Phi(0)} = \sqrt{2\pi}.\ \textit{Stirling}\,\textbf{の公式}$$

定理 11.18 を多次元 $\mathbb{R}^d$ でも対応できる形に少し拡張しておこう。$\|\ \ \|$ をユークリッドノルムとし $(\cdot,\cdot)$ をユークリッド内積とする。

前提

X_n は $\mathbb{R}^d$ 上のランダムウォーク, $X_0 = 0$, $E[\|X_1\|^2] < +\infty$, $E[X_1] = 0$.

記号

X を $\mathbb{R}^d$ 値確率変数とする。$\xi, \eta \in \mathbb{R}^d$ に対し $\mathrm{Var}[X](\xi,\eta) := \mathrm{Cov}[(\xi, X), (\eta, X)]$.

このとき $\mathrm{Var}[X] : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ は非負定値な対称 2 次形式である。次の定理においては極限として現れる正規分布は必ずしも密度関数を持たない。

11.20 定理. 確率変数列 $X_n/\sqrt{n}$ は平均 0 共分散 $\mathrm{Var}[X_1]$ の d 次元正規分布に法則収束する。

証明. 各 $\xi \in \mathbb{R}^d$ に対して $\mathrm{char}(X_n/\sqrt{n}, \xi)$ が $\exp\{-\mathrm{Var}[X_1](\xi,\xi)/2\}$ に収束することを示せばよいが、補題 11.6 の証明を今の状況に焼き直すだけでよいので省略する。 □

11.21 注意. 非負定値な対称 2 次形式 C が与えられたとき

$$\ker C := \{\xi \in \mathbb{R}^d : C(\xi,\xi) = 0\}$$

は $\mathbb{R}^d$ の線形部分空間である。平均 m 共分散 C の d 次元正規分布はアファイン部分空間 $\{x \in \mathbb{R}^d : (\xi, x - m) = 0\,\forall \xi \in \ker C\} = (\ker C)^\perp + m$ に集中している。条件 $\ker C = \{0\}$ が満たされるとき C を非退化というが、そのときは対応する正規分布は絶対連続である。

12　分布関数と弱収束

第 11 節であつかった状況では極限となる分布は正規分布という既に馴染みのあるものであり、補題 11.6 によってあらかじめ特定できていた。これが特性関数という道具がよく機能する原因である。しかしながら一般には極限となる分布の特定は難しく、それどころか収束するかどうかも自明でない場合が多い。そのようなときでも集積点の存在が分かれば、それをもとにして議論を展開する可能性が出てくる。たとえば集積点が唯一であるなら収束する。集積点の存在を主張する命題で代表的なものが $\mathbb{R}^d$ の有界集合に関する Bolzano-Weierstrass の定理であった。次は収束する点列は有界であるという事実の類似物である。

12.1 補題. $(\mathbb{R}^d, \mathrm{Borel}(\mathbb{R}^d))$ 上の確率測度の列 μ_n が $(\mathbb{R}^d, \mathrm{Borel}(\mathbb{R}^d))$ 上の確率測度 μ に弱収束するなら任意の $\varepsilon \in \mathbb{R}_{>0}$ に対して次を満たすような $K \in \mathbb{R}_{>0}$ が存在する。

$$\mu_n(\{x \in \mathbb{R}^d : \|x\| \leq K\}) > 1 - \varepsilon \ \forall n \in \mathbb{N}.$$

証明. まず $\bigcup_{k=1}^{\infty}\{x \in \mathbb{R}^d : \|x\| < k\} = \mathbb{R}^d$ であるから単調収束定理により

$$k \to \infty \text{ の極限で } \mu(\{x \in \mathbb{R}^d : \|x\| < k\}) \nearrow 1.$$

従って $\varepsilon \in \mathbb{R}_{>0}$ に対して $\mu(\{x \in \mathbb{R}^d : \|x\| < k\}) > 1 - \varepsilon$ を満たすような $k \in \mathbb{R}_{>0}$ が存在する。他方、$\{x \in \mathbb{R}^d : \|x\| < k\}$ は開部分集合である。定理 11.16 によれば

$$\liminf_{n\to\infty} \mu_n(\{x \in \mathbb{R}^d : \|x\| < k\}) \geq \mu(\{x \in \mathbb{R}^d : \|x\| < k\}) > 1 - \varepsilon.$$

よって $m \in \mathbb{N}$ が存在して $n \geq m$ なる $n \in \mathbb{N}$ に対しては $\mu_n(\{x \in \mathbb{R}^d : \|x\| < k\}) > 1 - \varepsilon$ が成り立つ。他方 $n < m$ なる各 $n \in \mathbb{N}$ に対して $\mu_n(\{x \in \mathbb{R}^d : \|x\| < k(n)\}) > 1 - \varepsilon$ を満たすように $k(n) \in \mathbb{R}_{>0}$ を選ぶことができる。そこで $K := \max\{k(1), \ldots, k(m-1), k\}$ とおくと

$$\mu_n(\{x \in \mathbb{R}^d : \|x\| \leq K\}) \geq \mu_n(\{x \in \mathbb{R}^d : \|x\| < k(n)\}) > 1 - \varepsilon \ \forall n \in \mathbb{N}$$

が成り立つ。ただし $n \geq m$ のときは $k(n) = k$ と読んでいる。 □

$\mathbb{R}^d$ における有界性と同じ役割を果たす概念を導入しておこう。

12.2 定義. $(\mathbb{R}^d, \mathrm{Borel}(\mathbb{R}^d))$ 上の確率測度の列 μ_n が**タイト**(tight) であるとは任意の $\varepsilon \in \mathbb{R}_{>0}$ に対して次を満たすような $K \in \mathbb{R}_{>0}$ が存在することをいう。

$$\mu_n(\{x \in \mathbb{R}^d : \|x\| \leq K\}) > 1 - \varepsilon \ \forall n \in \mathbb{N}.$$

タイト性については次にあげる *Prokhorov* **の定理**が基本的である。

12.3 定理. タイトな $(\mathbb{R}^d, \mathrm{Borel}(\mathbb{R}^d))$ 上の確率測度の列 μ_n は弱収束する部分列を持つ。

$d = 1$ の場合、以下のような証明手順が標準である。即ち、まず対角線論法を用いて分布関数の収束部分列を選びだす（*Helly* **の選出定理**という、補題 12.7 を参照)。次に極限関数の誘導する Lebesgue-Stieltjes 測度が確率測度であることおよび実際それに収束していることを確かめる（系 12.10 と系 12.13 を参照)。タイト性がないと漠収束の意味では極限があっても極限は確率測度にはならない。以上の議論は少し手間がかかるので、その前に極限が確率測度であることをタイト性が保証するというのを例を挙げて説明する。

12.4 例. 各 $n \in \mathbb{N}$ に対して μ_n を区間 $(n, n+1)$ 上の一様分布とする。その分布関数は

$$F_n(x) = \begin{cases} 0 & x \leq n \\ x - n & n < x < n+1 \\ 1 & x \geq n+1 \end{cases}$$

である。$K \in \mathbb{R}_{>0}$ をいかように選んでも n が十分大きいなら $\mu_n(\{x \in \mathbb{R} : |x| \leq K\}) = 0$ となってしまうのでタイトではない。また $n \to \infty$ における極限関数は恒等的に値 0 を取る。これは確率測度に対応しないが漠収束の意味では 0 測度に収束している。直感的には分布が $+\infty$ の彼方に逃げたとでもいえよう。

12.5 演習問題. 各 $n \in \mathbb{N}$ に対して μ_n を平均 0 分散 n の正規分布とする。このとき確率測度の列 μ_n はタイトでないことおよび分布関数の極限は確率測度に対応しないことを確かめよ。

12.6 演習問題. 与えられた $\mathbb{R}^d$ の点列 a_n に対して μ_n を a_n に質量を持つ Dirac 測度とする。このとき確率測度列 μ_n のタイト性は点列 a_n の有界性と同値であることを確かめよ。

Prokhorov の定理は更に進んで確率論を深めていくと必ず遭遇するものである。記憶にとどめておくとよいだろう。それでは $d=1$ の議論を紹介する。

12.7 補題. F_n を非減少関数 $\mathbb{R} \to \mathbb{R}$ の列であってある $M \in \mathbb{R}_{>0}$ に対して

$$|F_n(x)| \leq M \ \forall x \in \mathbb{R} \ \forall n \in \mathbb{N}$$

を満たすものとする。このとき $\alpha(n) < \alpha(n+1) \ \forall n \in \mathbb{N}$ を満たす自然数列 α と右連続な非減少関数 $F : \mathbb{R} \to \mathbb{R}$ が存在して関数列 $F_{\alpha(n)}$ は F に F の連続点において各点収束する。

証明. まず集合 $\mathbb{Q}$ の番号付け $q : \mathbb{N} \to \mathbb{Q}$ を一つ固定する。有界数列 $F_n(q(1))$ に対して Bolzano-Weierstrass の定理を適用する。

$$\text{ある狭義に増加な } \phi(1,\cdot) : \mathbb{N} \to \mathbb{N} \text{ が存在して数列 } F_{\phi(1,n)}(q(1)) \text{ は収束する。}$$

数列 $F_{\phi(1,n)}(q(2))$ も有界であることに注意して Bolzano-Weierstrass の定理を適用しよう。

$$\text{ある狭義に増加な } \psi : \mathbb{N} \to \mathbb{N} \text{ が存在して数列 } F_{\phi(1,\psi(n))}(q(2)) \text{ は収束する。}$$

合成写像 $\phi(1,\psi(\cdot))$ を $\phi(2,\cdot)$ と表すとそれは狭義に増加である。以後、ψ は表に出さずに単に $\phi(2,\cdot)$ は $\phi(1,\cdot)$ の部分列であるという表現をする。帰納的に推論すること（および選択公理）により各 $k \in \mathbb{N}$ に対して狭義に増加な $\phi(k,\cdot) : \mathbb{N} \to \mathbb{N}$ であって

$$\phi(k,\cdot) \text{ は } \phi(k-1,\cdot) \text{ の部分列であり数列 } F_{\phi(k,n)}(q(k)) \ n \in \mathbb{N} \text{ は収束する}$$

ものが存在することが示せる。各 $k \in \mathbb{N}$ に対して数列 $F_{\phi(n,n)}(q(k))$ は収束数列 $F_{\phi(k,n)}(q(k))$ の部分列であるから収束する（これを対角線論法という）。$\alpha(n) := \phi(n,n)$ とおくと

$$\alpha(\cdot) \text{ は狭義に増加であり各 } r \in \mathbb{Q} \text{ に対して数列 } F_{\alpha(n)}(r) \ n \in \mathbb{N} \text{ は収束する。}$$

各 $r \in \mathbb{Q}$ に対して数列 $F_{\alpha(n)}(r) \ n \in \mathbb{N}$ の極限値を $\widetilde{F}(r)$ とおく。このとき $|\widetilde{F}(r)| \leq M$ が成り立つ。関数 F を

$$F : \mathbb{R} \to \mathbb{R}, x \mapsto \inf_{r \in \mathbb{Q} : r > x} \widetilde{F}(r)$$

によって導入する。定義より F は非減少である。$a \in \mathbb{R}$ とする。数列 $F(a+1/k) \ k \in \mathbb{N}$ は単調減少かつ $F(a)$ を下界に持つので極限を持つ。実は

$$\inf_{k \in \mathbb{N}} F(a+1/k) = \inf_{k \in \mathbb{N}} \inf_{r \in \mathbb{Q} : r > a+1/k} \widetilde{F}(r) = \inf_{r \in \mathbb{Q} : r > a} \widetilde{F}(r) = F(a)$$

であるから数列 $F(a+1/k) \ k \in \mathbb{N}$ は $F(a)$ に収束する。F の非減少性によりこれは F が a において右連続であることを意味する。従って F は右連続な非減少関数である。$a \in \mathbb{R}$ とする。各 F_n が非減少関数であったことから次が成り立つ。

$$\limsup_{n \to \infty} F_{\alpha(n)}(a) \leq \limsup_{n \to \infty} F_{\alpha(n)}(r) = \widetilde{F}(r) \ \forall r \in \mathbb{Q} \text{ with } r > a$$

従って関数 F の定義より $\limsup_{n\to\infty} F_{\alpha(n)}(a) \le F(a)$ である。更に a は F の連続点であると仮定する。与えられた $\varepsilon \in \mathbb{R}_{>0}$ に対して $b < a$ が存在して $F(b) > F(a) - \varepsilon$ が成り立つ。ここで $q \in \mathbb{Q} \cap (b, a)$ を一つ選ぶ。各 F_n が非減少関数であったことから次が成り立つ。

$$\liminf_{n\to\infty} F_{\alpha(n)}(a) \ge \liminf_{n\to\infty} F_{\alpha(n)}(q) = \widetilde{F}(q) \ge \inf_{r\in\mathbb{Q}:r>b} \widetilde{F}(r) = F(b) > F(a) - \varepsilon$$

$\varepsilon \in \mathbb{R}_{>0}$ は任意であったので $\liminf_{n\to\infty} F_{\alpha(n)}(a) \ge F(a)$ である。よって F の連続点 a において数列 $F_{\alpha(n)}(a)$ は $F(a)$ に収束する。 □

12.8 補題. (i) $(\mathbb{R}, \mathrm{Borel}(\mathbb{R}))$ 上の σ 有限な測度 μ に対して集合 $\{x \in \mathbb{R} : \mu(\{x\}) > 0\}$ は高々可算である。
(ii) 右連続な非減少関数 $F : \mathbb{R} \to \mathbb{R}$ に対して F の不連続点は高々可算である。

証明. (i) 集合 $D := \{x \in \mathbb{R} : \mu(\{x\}) > 0\}$ は有限でないとしてよい。σ 有限性より列 $B_n \in \mathrm{Borel}(\mathbb{R})$ で $\bigcup_{n=1}^{\infty} B_n = S$ かつ $\mu(B_n) < +\infty$ $\forall n \in \mathbb{N}$ であるものが存在する。着眼点は

$$D = \bigcup_{n=1}^{\infty} D_n \text{ ただし } D_n := \{x \in \bigcup_{k=1}^{n} B_k : \mu(\{x\}) \ge 1/n\}$$

である。各 D_n の有限部分集合 A に対して次が成り立つ。

$$\#A = \sum_{x\in A} 1 \le \sum_{x\in A} n\mu(\{x\}) = n\mu(A) \le n\mu(\bigcup_{k=1}^{n} B_k)$$

また $n\mu(\bigcup_{k=1}^{n} B_k) \le n \sum_{k=1}^{n} \mu(B_k) < +\infty$ である。よって D_n 自身が有限集合であり、

$$\#D_n \le n\mu(\bigcup_{k=1}^{n} B_k)$$

であることが分かる。D は有限集合の可算合併であるから高々可算であるといえばよいが、ここはもう少し掘り下げてみよう。便宜上 $D_0 = \emptyset$ とする。定義より $D_{n-1} \subset D_n$ $\forall n \in \mathbb{N}$ である。さて各 $x \in D$ に対して $n \in \mathbb{N}$ であって $x \in D_n \setminus D_{n-1}$ を満たすものがただ一つ存在する。その様な n に対して $\phi(x) = \#D_{n-1} + \#\{y \in D_n \setminus D_{n-1} : y \le x\}$ とおく。このようにして定義される写像 $\phi : D \to \mathbb{N}$ は全単射である。 □

12.9 定義. $(\mathbb{R}, \mathrm{Borel}(\mathbb{R}))$ 上の確率測度 μ に対して $x \in \mathbb{R}$ が**測度 μ の連続点**であるとは $\mu(\{x\}) = 0$ を満たすことをいう。

12.10 系. μ_n をタイトな $(\mathbb{R}, \mathrm{Borel}(\mathbb{R}))$ 上の確率測度の列とする。このとき $\alpha(n) < \alpha(n+1)$ $\forall n \in \mathbb{N}$ を満たす自然数列 α と $(\mathbb{R}, \mathrm{Borel}(\mathbb{R}))$ 上の確率測度 μ が存在して μ の各連続点 x において数列 $\mu_{\alpha(n)}((-\infty, x])$ は $\mu((-\infty, x])$ に収束する。

証明. 各 $n \in \mathbb{N}$ に対して F_n を確率測度 μ_n の分布関数とする。明らかに F_n は補題 12.7 の前提を満たす。従って $\alpha(n) < \alpha(n+1)$ $\forall n \in \mathbb{N}$ を満たす自然数列 α と右連続な非減少関数

$F:\mathbb{R}\to\mathbb{R}$ が存在して関数列 $F_{\alpha(n)}$ は F に F の連続点において各点収束する。さて $\varepsilon\in\mathbb{R}_{>0}$ とする。確率測度の列 μ_n はタイトであったので次を満たすような $K\in\mathbb{R}_{>0}$ が存在する。

$$\mu_n(\{x\in\mathbb{R}:-K<x\leq K\})>1-\varepsilon\ \forall n\in\mathbb{N}.$$

補題 12.8(ii) により F の不連続点は高々可算に過ぎないので $a\leq -K,\ K\leq b$ を満たすような連続点 a,b が存在する。その様な a,b に対して評価

$$\begin{aligned}0\leq F_{\alpha(n)}(a)=\mu_{\alpha(n)}((-\infty,a])\leq\mu_{\alpha(n)}((-\infty,-K])+\mu_{\alpha(n)}((K,+\infty))<\varepsilon,\\ 1\geq F_{\alpha(n)}(b)=\mu_{\alpha(n)}((-\infty,b])\geq\mu_{\alpha(n)}(\{x\in\mathbb{R}:-K<x\leq K\})>1-\varepsilon\end{aligned}$$

が成り立つ。従って $0\leq F(a)\leq\varepsilon,\ 1\geq F(b)\geq 1-\varepsilon$ を得る。F の非減少性によりこれは

$F(x)$ が極限 $x\to-\infty$ において 0 に収束し極限 $x\to+\infty$ において 1 に収束する

ことを意味する。従って極限関数 F の誘導する Lebesgue-Stieltjes 測度を μ とすると、それは確率測度であって次が成り立つ。

$$\mu((-\infty,x])=F(x)\ \forall x\in\mathbb{R}$$

F の連続点全体の集合は $\{x\in\mathbb{R}:\mu(\{x\})=0\}$ に一致するので、$x\in\mathbb{R}$ かつ $\mu(\{x\})=0$ ならば $\mu_n((-\infty,x])$ は $\mu((-\infty,x])$ に収束する。 □

12.11 補題. $a,b\in\mathbb{R}$ with $a<b,\ f\in C^1(\mathbb{R})$ と $\mathbb{R}$ 上の Radon 測度 μ に対して

$$\int_{(a,b]}f\mu=f(a)\mu((a,b])+\int_{(a,b)}f'(x)\mu((x,b])\ \lambda(dx)$$

証明. $A:=\{(x,y)\in\mathbb{R}^2:a<x<y\leq b\}$ とおくと $A\subset(a,b)\times(a,b]$ である。$\lambda\otimes\mu$ 可積分関数 $(x,y)\mapsto f'(x)1_A(x,y)$ に Fubini の定理を適用する（定理 8.22 参照)。 □

12.12 補題. μ_n を $(\mathbb{R},\mathrm{Borel}(\mathbb{R}))$ 上の確率測度の列とする。$(\mathbb{R},\mathrm{Borel}(\mathbb{R}))$ 上の確率測度 μ と $\mathbb{R}$ の可算部分集合 D が存在して各 $x\in\mathbb{R}\setminus D$ に対して数列 $\mu_n((-\infty,x])$ が $\mu((-\infty,x])$ に収束するなら測度列 μ_n は μ に弱収束する。

証明. $a,b\in\mathbb{R}\setminus D,\ a<b,\ f\in C^1(\mathbb{R})$ とする。集合 D は可算なので $D\in\mathrm{Borel}(\mathbb{R})$ かつ $\lambda(D)=0$ である。また $|f'(x)\mu_n((x,b])|\leq\sup_{y\in[a,b]}|f'(y)|\ \forall x\in(a,b)\ \forall n\in\mathbb{N}$ かつ $\lim_{n\to\infty}\mu_n((x,b])=\mu((x,b])\ \forall x\in(a,b)\setminus D$ であるから、Lebesgue の優収束定理より

$$\int_{(a,b)}f'(x)\mu_n((x,b])\ \lambda(dx)\ \text{は}\ \int_{(a,b)}f'(x)\mu((x,b])\ \lambda(dx)\ \text{に収束する。}$$

他方、$a,b\notin D$ より差 $\mu_n((-\infty,b])-\mu_n((-\infty,a])$ は差 $\mu((-\infty,b])-\mu((-\infty,a])$ に収束する。即ち $\mu_n((a,b])$ は $\mu((a,b])$ に収束する。従って補題 12.11 を適用して次が導かれる。

$$\int_{(a,b]}f\mu_n\ \text{は}\ \int_{(a,b]}f\mu\ \text{に収束する。}$$

ここで $\xi \in \mathbb{R}$ と $\varepsilon \in \mathbb{R}_{>0}$ を任意に固定する。$K \in \mathbb{R}_{>0}$ を $\mu((-K, K]) > 1 - \varepsilon/2$ が成り立つように十分大きくとる。D は可算集合に過ぎないので $a \le -K$, $K \le b$ を満たすような $a, b \in \mathbb{R} \setminus D$ が存在する。その様な a, b と C^1 級関数 $f : x \mapsto e^{\sqrt{-1}\xi x}$ に対して評価

$$\begin{aligned}|\mathrm{char}(\mu_n, \xi) - \mathrm{char}(\mu, \xi)| &\le |\int_{\mathbb{R}\setminus(a,b]} f \ \mu_n| + |\int_{(a,b]} f \ \mu_n - \int_{(a,b]} f \ \mu| + |\int_{\mathbb{R}\setminus(a,b]} f \ \mu| \\ &\le \mu_n(\mathbb{R} \setminus (a, b]) + |\int_{(a,b]} f \ \mu_n - \int_{(a,b]} f \ \mu| + \mu(\mathbb{R} \setminus (a, b])\end{aligned}$$

が成り立つ。$\mu_n(\mathbb{R} \setminus (a, b]) = 1 - \mu_n((a, b])$ は $\mu(\mathbb{R} \setminus (a, b])$ に収束するので

$$\limsup_{n\to\infty} |\mathrm{char}(\mu_n, \xi) - \mathrm{char}(\mu, \xi)| \le 2\mu(\mathbb{R} \setminus (a, b]) \le 2\mu(\mathbb{R} \setminus (-K, K]) \le \varepsilon$$

を得る。$\varepsilon \in \mathbb{R}_{>0}$ は任意であったので特性関数の列 $\mathrm{char}(\mu_n, \cdot)$ は $\mathrm{char}(\mu, \cdot)$ に各点収束する。よって定理 11.14 を適用して測度列 μ_n は μ に弱収束することが導かれた。 □

分布関数列の収束と確率測度列の弱収束との関係を述べておく。

12.13 系. μ_n を $(\mathbb{R}, \mathrm{Borel}(\mathbb{R}))$ 上の確率測度の列、μ を $(\mathbb{R}, \mathrm{Borel}(\mathbb{R}))$ 上の確率測度とする。このとき以下は同値である。

1° 測度列 μ_n は μ に弱収束する。
2° μ の各連続点 x において $\mu_n((-\infty, x])$ は $\mu((-\infty, x])$ に収束する。

証明. 測度列 μ_n は μ に弱収束し、x を μ の連続点とする。定理 11.16 と系 11.17 より

$$\begin{aligned}\mu((-\infty, x]) = \mu((-\infty, x)) &\le \liminf_{n\to\infty} \mu_n((-\infty, x)) \\ &\le \liminf_{n\to\infty} \mu_n((-\infty, x]) \le \limsup_{n\to\infty} \mu_n((-\infty, x]) \le \mu((-\infty, x])\end{aligned}$$

したがって $1^\circ \Rightarrow 2^\circ$ が成り立つ。逆は補題 12.8(i) と補題 12.12 から導かれる。 □

確率測度列の弱収束に関しては次にあげる *Skorokhod* **の定理**も基本的である。

12.14 定理. $(\mathbb{R}, \mathrm{Borel}(\mathbb{R}))$ 上の確率測度 μ に弱収束する確率測度の列 μ_n が与えられたとし、

$$\begin{aligned}X &: (0, 1) \to \overline{\mathbb{R}}, \omega \mapsto \inf\{x \in \mathbb{R} : \mu((-\infty, x]) \ge \omega\}; \\ X_n &: (0, 1) \to \overline{\mathbb{R}}, \omega \mapsto \inf\{x \in \mathbb{R} : \mu_n((-\infty, x]) \ge \omega\}\end{aligned}$$

で定義される関数 X と関数列 X_n を導入する。このとき $X_*\lambda = \mu$ かつ $(X_n)_*\lambda = \mu_n$ $\forall n \in \mathbb{N}$ であり、さらに次を満たすような $A \in \mathrm{Borel}((0, 1))$ が存在する。

$$\lambda(A) = 1, \forall \omega \in A \lim_{n\to\infty} X_n(\omega) = X(\omega)$$

証明. 補題 4.24(iii) より $\lambda(\{\omega : X(\omega) \leq x\}) = \mu((-\infty, x])\ \forall x \in \mathbb{R}$ であるが、これは $X_*\lambda = \mu$ を意味する。しばらく $\omega \in (0,1)$ を固定する。μ の分布関数を F と書くと、補題 4.24 の証明で登場したように $[X(\omega), +\infty) = F^{-1}([\omega, +\infty))$ が成り立つ。従って次の同値性を得る。

$$(\star) \qquad x < X(\omega) \Leftrightarrow \mu((-\infty, x]) < \omega$$

ここで $x < X(\omega)$ とする。系 11.17 を適用すると

$$\limsup_{n\to\infty} \mu_n((-\infty, x]) \leq \mu((-\infty, x]) < \omega$$

従ってある番号から先の n に対して $\mu_n((-\infty, x]) < \omega$ が成り立つ。同値性 $(\star)$ によりこれは

$$x \leq \liminf_{n\to\infty} X_n(\omega)$$

を意味する。$x \in \mathbb{R}$ は $x < X(\omega)$ なら任意なので次を得る。

$$X(\omega) \leq \liminf_{n\to\infty} X_n(\omega).$$

ここで補助的に次の関数 Z を導入する。

$$Z : (0,1) \to \overline{\mathbb{R}}, \omega \mapsto \inf\{x \in \mathbb{R} : \mu((-\infty, x)) > \omega\};$$

包含関係 $\{x \in \mathbb{R} : \mu((-\infty, x)) > \omega\} \subset \{x \in \mathbb{R} : \mu((-\infty, x]) \geq \omega\}$ より $X(\omega) \leq Z(\omega)$ が従う。今度は $\omega \in (0,1)$ と $x \in \mathbb{R}$ に対して次の関係が成り立つがその証明は読者に委ねる。

$$(\star\star) \qquad Z(\omega) < x \Leftrightarrow \mu((-\infty, x)) > \omega$$

よって $y \in \mathbb{R}$ を固定すると $\{\omega \in (0,1) : Z(\omega) < y\} = (0, \mu((-\infty, y)))$ が成り立つので

$$\inf_{y>x} \lambda(\{\omega \in (0,1) : Z(\omega) < y\}) = \inf_{y>x} \mu((-\infty, y))\ \forall x \in \mathbb{R}$$

即ち $\lambda(\{\omega \in (0,1) : Z(\omega) \leq x\}) = \mu((-\infty, x]) = \lambda(\{\omega \in (0,1) : X(\omega) \leq x\})\ \forall x \in \mathbb{R}$ となる。補題 4.24 の証明によれば、これは $X \leq Z$ とあわせて

$$\lambda(\{\omega \in (0,1) : X(\omega) = Z(\omega)\}) = 1$$

を意味する。更に補助的な関数列 Z_n を導入する。

$$Z_n : (0,1] \to \overline{\mathbb{R}}, \omega \mapsto \inf\{x \in \mathbb{R} : \mu_n((-\infty, x)) > \omega\}$$

各 $n \in \mathbb{N}$ に対して $X_n \leq Z_n$ が成り立つ。他方、$(\star\star)$ をもとにして X_n に関するのと同様な議論展開を行うと次が導かれる。

$$\limsup_{n\to\infty} Z_n(\omega) \leq Z(\omega).$$

以上より各 $\omega \in (0,1)$ に対して $\liminf_{n\to\infty} X_n(\omega)$ と $\limsup_{n\to\infty} X_n(\omega)$ は $X(\omega)$ で下から、$Z(\omega)$ で上から評価される。既に確かめたように $A := \{\omega \in (0,1) : X(\omega) = Z(\omega)\}$ は $\lambda(A) = 1$ を満たすのでこれが求める集合である。 $\square$

12.15 定義. 確率変数の列 X_n が確率変数 Y に**概収束**(almost sure convergence) するとは $X_n(\omega)$ が $Y(\omega)$ に収束する $\omega \in \Omega$ 全体が P-a.s. 集合であることをいう。即ち $A \in \mathcal{F}$ であって次を満たすものが存在することである。

$$P(A) = 1, \forall \omega \in A\ \lim_{n\to\infty} X_n(\omega) = Y(\omega)$$

13　大数の弱法則と強法則

多数のランダム要因が積み重なると個々のサンプルの特性は失われて集団としての特性量が現れるという原理が確率論では知られている。ランダム要因の時間的集積を表現するのがランダムウォークであると第 11 節でいったが、その長時間平均が期待値を表すというのが大数の法則で、上で述べた原理の典型である。この節では大数の法則に関連して、確率変数列の収束について諸概念を導入し互いの関係について述べる。

まず Chebyshev の不等式を使って大数の弱法則を導く標準的な手続きを述べておく。

記号

集合 S に対して $\mathrm{Pair}(S) := \{I \subset S : \#I = 2\}$ とおく。

13.1 補題. 2 乗可積分確率変数 $X_1, X_2, \ldots, X_n$ と $C : \mathrm{Pair}(\mathbb{N}_{\leq n}) \to \mathbb{R}_{\geq 0}$ に対して

$$E[\prod_{i \in I} X_i] \leq C(I) \prod_{i \in I} \sqrt{E[{X_i}^2]} \ \forall I \in \mathrm{Pair}(\mathbb{N}_{\leq n})$$

が成り立つとする。このとき $\sum_{I \in \mathrm{Pair}(\mathbb{N}_{\leq n}) : i \in I} C(I) \leq R \ \forall i \in \mathbb{N}_{\leq n}$ なら

$$E[|\sum_{k=1}^{n} X_k|^2] \leq (1 + R) \sum_{k=1}^{n} E[{X_k}^2].$$

証明. 各 $I \in \mathrm{Pair}(\mathbb{N}_{\leq n})$ に対して次が成り立つ。

$$2E[\prod_{i \in I} X_i] \leq 2C(I) \prod_{i \in I} \sqrt{E[{X_i}^2]} \leq C(I) \sum_{i \in I} E[{X_i}^2].$$

2 番目の不等号は $2ab \leq a^2 + b^2$ による。I について $\mathrm{Pair}(\mathbb{N}_{\leq n})$ 全域にわたって和をとると

$$E[\sum_{I \in \mathrm{Pair}(\mathbb{N}_{\leq n})} 2 \prod_{i \in I} X_i] \leq \sum_{i=1}^{n} \sum_{I \in \mathrm{Pair}(\mathbb{N}_{\leq n}) : i \in I} C(I) E[|X_i|^2] \leq R \sum_{i=1}^{n} E[|X_i|^2].$$

左辺に $\sum_{k=1}^{n} E[|X_k|^2]$ を加えたものが $E[|\sum_{k=1}^{n} X_k|^2]$ である。 □

13.2 定理. 2 乗可積分確率変数の列 $Z_1, Z_2, \ldots, Z_n, \ldots$ と $m \in \mathbb{R}$, $v \in \mathbb{R}_{\geq 0}$, $N \in \mathbb{N}$ に対して

$$E[Z_n] = m \ \forall n \in \mathbb{N}, \ \mathrm{Var}[Z_n] \leq v \ \forall n \in \mathbb{N}, \ \mathrm{Cov}[Z_i, Z_j] = 0 \text{ if } |i - j| \geq N$$

が成り立つなら

$$P(|\frac{1}{n} \sum_{k=1}^{n} Z_k - m| \geq \varepsilon) \leq \frac{(2N-1)v}{\varepsilon^2 n} \ \forall n \in \mathbb{N}, \forall \varepsilon \in \mathbb{R}_{>0}.$$

証明. 確率変数 $\sum_{k=1}^{n} Z_k / n$ に対して Chebyshev の不等式（系 6.35）を適用すると

$$P(|\frac{1}{n} \sum_{k=1}^{n} Z_k - m| \geq \varepsilon) \leq \frac{1}{\varepsilon^2} \mathrm{Var}[\frac{1}{n} \sum_{k=1}^{n} Z_k] = \frac{1}{\varepsilon^2 n^2} E[|\sum_{k=1}^{n} (Z_k - m)|^2].$$

ここで記号の簡単化のため $X_i := Z_i - m$ と書くと Schwarz の不等式より

$$E[X_iX_j] \leq \sqrt{E[X_i^2]}\sqrt{E[X_j^2]}$$

であり、また $|i-j| \geq N$ なら $E[X_iX_j] = \mathrm{Cov}[Z_i, Z_j] = 0$ である。したがって $|i-j| < N$ のとき $C(\{i,j\}) = 1$ それ以外の時 $C(\{i,j\}) = 0$ とかくと

$$E[\prod_{i\in I} X_i] \leq C(I)\prod_{i\in I}\sqrt{E[X_i^2]}\ \forall I \in \mathrm{Pair}(\mathbb{N}_{\leq n}) \text{ かつ } \sum_{I\in\mathrm{Pair}(\mathbb{N}_{\leq n}):i\in I} C(I) \leq 2N-2\ \forall i \in \mathbb{N}_{\leq n}$$

が満たされる。よって補題 13.1 により

$$E[|\sum_{k=1}^n (Z_k - m)|^2] \leq (2N-1)\sum_{k=1}^n \mathrm{Var}[Z_k] \leq (2N-1)nv$$

を得るので結論が導ける。 □

13.3 例. 2 乗可積分な確率変数の列 $Z_1, Z_2, \ldots, Z_n, \ldots$ は**対ごとに独立**である、すなわち $i \neq j$ なら組 Z_i, Z_j は独立とする。このとき $\mathrm{Cov}[Z_i, Z_j] = 0$ if $i \neq j$ である。

13.4 演習問題. 例 13.3 を示せ。

13.5 定義. 確率変数の列 X_n が確率変数 Y に**確率収束**(convergence in probability) するとは各 $\varepsilon \in \mathbb{R}_{>0}$ に対して $P(|X_n - Y| \geq \varepsilon)$ が 0 に収束することをいう。

13.6 補題. 確率変数列 X_n が確率変数 Y に確率収束するなら、確率測度列 $\mathcal{L}(X_n, \cdot)$ は確率測度 $\mathcal{L}(Y, \cdot)$ に弱収束する（確率変数列 X_n は確率測度 $\mathcal{L}(Y, \cdot)$ に法則収束する）。

証明. 定理 11.14 によれば、測度の列 $\mathcal{L}(X_n, \cdot)$ が測度 $\mathcal{L}(Y, \cdot)$ に漠収束することを示せばよい。そこで $f \in C_0^\infty(\mathbb{R})$ とする。f は有界かつ一様連続である。定理 6.38(i) より

$$\int_{\mathbb{R}} f(x)\ \mathcal{L}(X_n, dx) - \int_{\mathbb{R}} f(x)\ \mathcal{L}(Y, dx) = E[f(X_n)] - E[f(Y)]$$

$\delta > 0$ とする。補題 6.43 によれば上の絶対値は次で抑えられる。

$$\sup_{x,y:|x-y|<\delta} |f(x) - f(y)| + 2\sup_{x\in\mathbb{R}} |f(x)| P(|X_n - Y| \geq \delta)$$

確率収束していたことから

$$\limsup_{n\to\infty} \Big| \int_{\mathbb{R}} f(x)\mathcal{L}(X_n, dx) - \int_{\mathbb{R}} f(x)\mathcal{L}(Y, dx) \Big| \leq \sup_{x,y:|x-y|<\delta} |f(x) - f(y)|$$

右辺は $\delta \downarrow 0$ のとき 0 に収束する。 □

定理 13.2 は広く**大数の弱法則**(weak law of large numbers) と呼ばれるものの基本形である。さて上では分散が 0 に収束することを通じて確率収束することを導いていたのである。

13.7 定義. $p \in \mathbb{R}_{>0}$ とする。確率変数の列 X_n が確率変数 Y に p **次平均収束**(convergence in L^p-sense) するとは $E[|X_n - Y|^p]$ が 0 に収束することをいう。

13.8 補題. $p \in \mathbb{R}_{>0}$ とする。確率変数の列 X_n が確率変数 Y に p 次平均収束するなら、列 X_n は Y に確率収束する。

13.9 演習問題. 補題 13.8 を示せ。

13.10 例. 2 乗可積分な確率変数の列 Z_n が $\mathrm{Cov}[Z_i, Z_j] = 0$ if $i \neq j$ を満たしかつ同じ平均と分散を持つとする。このとき $\sum_{k=1}^{n} Z_k/n$ は $E[Z_1]$ に 2 次平均収束する。従って確率変数の列 Z_n は大数の弱法則を満たす。

一般には確率収束から p 次平均収束は出てこないが、付帯条件があればうまくいく。

13.11 演習問題. 確率変数の列 X_n が 0 に確率収束し $\sup_{n\in\mathbb{N}} E[|X_n|^2] < +\infty$ であるなら、列 X_n は 0 に 1 次平均収束することを示せ。

13.12 補題. X, Y を実確率変数とする。
(i) $p > 0$, $K > 0$, $\lambda > 0$, $\varepsilon > 0$ に対して

$$E[\min\{|Y|^p, K\}] \leq (\lambda+\varepsilon)^p + (1+\varepsilon/\lambda)^p E[|X|^p; |X| \geq \lambda] + KP(|Y-X| \geq \varepsilon)$$

(ii) $p \geq 1$, $\lambda > 0$, $\varepsilon > 0$ に対して

$$E[|X-Y|^p] \leq \varepsilon^p + 2^{p-1}E[|X|^p\,; |X| \geq \lambda] + 2^{p-1}E[|Y|^p\,; |Y| \geq \lambda] + 2^p\lambda^p P(|X-Y| \geq \varepsilon)$$

(iii) $0 < p < 1$, $\lambda > 0$, $\varepsilon > 0$ に対して

$$E[|X-Y|^p] \leq \varepsilon^p + E[|X|^p\,; |X| \geq \lambda] + E[|Y|^p\,; |Y| \geq \lambda] + 2\lambda^p P(|X-Y| \geq \varepsilon)$$

証明. (i) 積分範囲を分割する。$\{|Y-X| \geq \varepsilon\}$ 上で $\min\{|Y|^p, K\} \leq K$ と評価し $\{|Y-X| < \varepsilon\}$ 上では $\min\{|Y|^p, K\} \leq |Y|^p \leq (|X| + |Y-X|)^p \leq (|X|+\varepsilon)^p$ と評価することにより

$$E[\min\{|Y|^p, K\}] \leq E[(|X|+\varepsilon)^p] + KP(|Y-X| \geq \varepsilon)$$

を得る。更に $E[(|X|+\varepsilon)^p]$ について積分範囲を $\{|X| < \lambda\}$ と $\{|X| \geq \lambda\}$ に分割する。前者上で $(|X|+\varepsilon)^p \leq (\lambda+\varepsilon)^p$ 後者上では $(|X|+\varepsilon)^p \leq (1+\varepsilon/\lambda)^p|X|^p$ と評価することにより

$$E[(|X|+\varepsilon)^p] \leq (\lambda+\varepsilon)^p + (1+\varepsilon/\lambda)^p E[|X|^p; |X| \geq \lambda].$$

(ii) 積分範囲を分割する。$\{|X-Y| < \varepsilon\}$ 上では $|X-Y|^p \leq \varepsilon^p$ と評価し $\{|X-Y| \geq \varepsilon\}$ 上では $|X-Y|^p \leq (|X|+|Y|)^p \leq 2^{p-1}(|X|^p + |Y|^p)$ と評価することにより

$$E[|X-Y|^p] \leq \varepsilon^p + 2^{p-1}E[|X|^p\,; |X-Y| \geq \varepsilon] + 2^{p-1}E[|Y|^p\,; |X-Y| \geq \varepsilon]$$

を得る。第 2 項について積分範囲を $\{|X-Y| \geq \varepsilon, |X| \geq \lambda\}$, $\{|X-Y| \geq \varepsilon, |X| < \lambda\}$ と分割する。後者上では $|X|^p \leq \lambda^p$ と評価することにより

$$E[|X|^p\,; |X-Y| \geq \varepsilon] \leq E[|X|^p\,; |X| \geq \lambda] + \lambda^p P(|X-Y| \geq \varepsilon)$$

を得る。第 3 項についても同様に評価できる。

(iii) (ii) と違うのは $(|X|+|Y|)^p \leq |X|^p + |Y|^p$ と評価する点だけである。 □

13.13 演習問題. $0 < p < 1$, $x, y \geq 0$ に対して $(x+y)^p \leq x^p + y^p$ が成り立つことを示せ。

13.14 定理. $p > 0$ とする。確率変数の列 X_n が確率変数 Y に確率収束し

$$\inf_{\lambda>0} \sup_{n\in\mathbb{N}} E[|X_n|^p\,;|X_n| \geq \lambda] = 0$$

であるなら、Y は p 乗可積分かつ列 X_n は Y に p 次平均収束する。

証明. $K > 0$, $\varepsilon > 0$, $n \in \mathbb{N}$ とする。ここで $\sup_{n\in\mathbb{N}} E[|X_n|^p\,;|X_n| \geq \lambda] \leq 1$ となるように λ をえらぶ。このとき補題 13.12(i) を適用して次を得る。

$$E[\min\{|Y|^p, K\}] \leq (\lambda+\varepsilon)^p + (1+\varepsilon/\lambda)^p + KP(|Y - X_n| \geq \varepsilon)$$

まず $n \to \infty$ の極限を取り、次に $\varepsilon \downarrow 0$ の極限を取る。確率収束していたことから

$$E[\min\{|Y|^p, K\}] \leq \lambda^p + 1 \ \forall K \in \mathbb{R}_{>0}.$$

最後に $K \to +\infty$ の極限を取る。単調収束定理により

$$E[|Y|^p] \leq \lambda^p + 1 < +\infty.$$

簡単化のため $Z_n := X_n - Y$ と書く。$p \geq 1$ のときについて p 次平均収束を議論する。$n \in \mathbb{N}$, $\varepsilon > 0$, $\lambda > 0$ とする。補題 13.12(ii) より

$$E[|Z_n|^p] \leq \varepsilon^p + 2^{p-1} \sup_{k\in\mathbb{N}} E[|X_k|^p\,;|X_k| \geq \lambda] + 2^{p-1} E[|Y|^p\,;|Y| \geq \lambda] + 2^p\lambda^p P(|Z_n| \geq \varepsilon)$$

が成り立つ。Z_n は 0 に確率収束するから

$$\limsup_{n\to\infty} E[|Z_n|^p] \leq \varepsilon^p + 2^{p-1} \sup_{k\in\mathbb{N}} E[|X_k|^p\,;|X_k| \geq \lambda] + 2^{p-1} E[|Y|^p\,;|Y| \geq \lambda].$$

仮定より右辺第 2 項は $\lambda \to +\infty$ の極限で 0 に収束する。第 3 項は $E[|Y|^p] < +\infty$ であるから単調収束定理により $\lambda \to +\infty$ の極限で 0 に収束する。よって $E[|Z_n|^p]$ は 0 に収束する。$0 < p < 1$ のときについて p 次平均収束を議論するには補題 13.12(iii) を適用すればよい。 □

13.15 定義. 実確率変数の列 X_n について

$$\inf_{\lambda>0} \sup_{n\in\mathbb{N}} E[|X_n|\,;|X_n| \geq \lambda] = 0$$

が成り立つとき**一様可積分**(uniformly integrable) であるという。

13.16 演習問題. $p \in \mathbb{R}_{>1}$ とする。確率変数の列 X_n が $\sup_{n\in\mathbb{N}} E[|X_n|^p] < +\infty$ を満たすなら、それは一様可積分であることを示せ。

13.17 補題. 2 乗可積分確率変数 Z_n と $C : \mathrm{Pair}(\mathbb{N}) \to \mathbb{R}_{\geq 0}$ に対して

$$E[\prod_{i\in I} Z_i] \leq C(I) \prod_{i\in I} \sqrt{E[{Z_i}^2]} \ \forall I \in \mathrm{Pair}(\mathbb{N}), \sup_{n\in\mathbb{N}} E[Z_n^2] < +\infty, \sup_{i\in\mathbb{N}} \sum_{I\in\mathrm{Pair}(\mathbb{N}):i\in I} C(I) < +\infty$$

が成り立つとする。このとき $\sum_{k=1}^n Z_k/n$ は 0 に概収束する。

証明. $m, n \in \mathbb{N}$, $m \leq n$ とする。ここで

$$R := \sup_{i \in \mathbb{N}} \sum_{I \in \mathrm{Pair}(\mathbb{N}): i \in I} C(I), M := \sup_{n \in \mathbb{N}} E[|Z_n|^2]$$

として補題 13.1 を適用すると

$$(\star) \qquad E[|\sum_{k=m}^{n} Z_k|^2] \leq (1+R) \sum_{k=m}^{n} E[|Z_k|^2] \leq (1+R)M(n-m+1)$$

が得られる。従って記号簡略化のため $S_n := \sum_{k=1}^{n} Z_k$ と書くと

$$E[\sum_{n=1}^{\infty} \left|\frac{S_{n^2}}{n^2}\right|^2] = \sum_{n=1}^{\infty} \frac{1}{n^4} E[|\sum_{k=1}^{n^2} Z_k|^2] \leq \sum_{n=1}^{\infty} \frac{(1+R)Mn^2}{n^4} < +\infty.$$

これにより

$$B_1 := \{\omega \in \Omega : \sum_{n=1}^{\infty} |S_{n^2}(\omega)/n^2|^2 < +\infty\} \in \mathcal{F} \text{ は } P(B_1) = 1 \text{ を満たす。}$$

また $\omega \in B_1$ なら $S_{n^2}(\omega)/n^2$ は 0 に収束している。さて次のように変形してみる。

$$\frac{S_n(\omega)}{n} = \frac{[\sqrt{n}]^2}{n} \frac{S_{[\sqrt{n}]^2}(\omega)}{[\sqrt{n}]^2} + \frac{S_n(\omega) - S_{[\sqrt{n}]^2}(\omega)}{n}.$$

ただし $[x] := \max\{n \in \mathbb{Z} : n \leq x\}$ である。$\omega \in B_1$ のとき $(\sqrt{n}-1)^2/n \leq [\sqrt{n}]^2/n \leq 1$ であるから右辺第 1 項は 0 に収束する。他方 $(\star)$ により

$$E[\sum_{n=1}^{\infty} \left|\frac{S_n - S_{[\sqrt{n}]^2}}{n}\right|^2] = \sum_{n=1}^{\infty} \frac{1}{n^2} E[|\sum_{k=[\sqrt{n}]^2+1}^{n} Z_k|^2] \leq \sum_{n=1}^{\infty} \frac{(1+R)M(n-[\sqrt{n}]^2)}{n^2}$$

と評価されるが、$n - [\sqrt{n}]^2 = (\sqrt{n} + [\sqrt{n}])(\sqrt{n} - [\sqrt{n}]) \leq 2\sqrt{n}$ なので右辺の級数は収束する。従って

$$B_2 := \{\omega \in \Omega : \sum_{n=1}^{\infty} \left|\frac{S_n(\omega) - S_{[\sqrt{n}]^2}(\omega)}{n}\right|^2 < +\infty\} \in \mathcal{F} \text{ は } P(B_2) = 1 \text{ を満たす。}$$

こんどは $\omega \in B_2$ なら $(S_n(\omega) - S_{[\sqrt{n}]^2}(\omega))/n$ が 0 に収束する。ゆえに $\omega \in B_1 \cap B_2$ なら $S_n(\omega)/n$ は 0 に収束しかつ $P(B_1 \cap B_2) = 1$ である。 □

次は補題 13.17 で議論済みのものを適用しやすい条件で設定してみたものである。

13.18 定理. 2 乗可積分確率変数の列 Z_n と $m \in \mathbb{R}$, $N \in \mathbb{N}$ に対して

$$E[Z_n] = m \ \forall n \in \mathbb{N}, \ \sup_{n \in \mathbb{N}} \mathrm{Var}[Z_n] < +\infty, \ \mathrm{Cov}[Z_i, Z_j] = 0 \text{ if } |i-j| \geq N$$

が成り立つなら $\sum_{k=1}^{n} Z_k/n$ は m に概収束する。

証明. $X_i := Z_i - m$ と書き、$|i-j| < N$ のとき $C(\{i,j\}) = 1$ それ以外の時 $C(\{i,j\}) = 0$ とかくと、定理 13.2 の証明で論じたように

$$E[\prod_{i\in I} X_i] \le C(I) \prod_{i\in I} \sqrt{E[X_i^2]}\ \forall I \in \mathrm{Pair}(\mathbb{N})\ \text{かつ} \sum_{I\in \mathrm{Pair}(\mathbb{N}):i\in I} C(I) \le 2N-2\ \forall i \in \mathbb{N}$$

が満たされる。また $E[X_n^2] = \mathrm{Var}[Z_n]$ である。よって補題 13.17 を適用して $\sum_{k=1}^n Z_k/n$ が 0 に概収束することを得る。 □

上は広く**大数の強法則**(strong law of large numbers) と呼ばれるものの基本形である。

13.19 注意. 2 乗可積分な確率変数の列 Z_n は対ごとに独立であるとする。このとき例 13.3 で述べたように $\mathrm{Cov}[Z_i, Z_j] = 0$ if $i \ne j$ である。したがって X_n がすべて同じ平均と分散を持つなら大数の強法則が成り立つ。

13.20 例. $\mathbb{R}$ 上のランダムウォーク X_n について $0 < p < 1$, $P(X_1 = 1) = p$, $P(X_1 = -1) = 1 - p$ とする。このとき大数の強法則より $\lim_{n\to\infty} X_n/n = 2p - 1$ P-a.s. が成り立つ。

13.21 注意. 証明は割愛するが、独立な確率変数の列 Z_n が同じ分布を持ちかつ可積分なら大数の強法則が成り立つ。

13.22 補題. 確率変数の列 X_n が確率変数 Y に概収束するなら、列 X_n は Y に確率収束する。

証明. $A \in \mathcal{F}$ で $P(A) = 1$ を満たしかつすべての $\omega \in A$ に対して $X_n(\omega)$ は $Y(\omega)$ に収束するようなものが存在する。$A(n) := \{\omega \in \Omega : |X_n(\omega) - Y(\omega)| < \varepsilon\}$ とおいて非負値関数列 $1_{A(n)}$ に Fatou の補題を適用する。$\omega \in A$ ならある番号より大きい n については $1_{A(n)}(\omega) = 1$ であるから

$$\liminf_{n\to\infty} P(|X_n - Y| < \varepsilon) \ge E[\liminf_{n\to\infty} 1_{A(n)}] \ge P(A) = 1\ \forall \varepsilon \in \mathbb{R}_{>0}.$$

したがって各 $\varepsilon \in \mathbb{R}_{>0}$ に対して $P(|X_n - Y| \ge \varepsilon)$ は 0 に収束する。 □

補題 13.22 によると、大数の強法則が成り立てば大数の弱法則は必ず成り立つ。しかしながら、定理 13.2 における $P(|(Z_1 + \cdots + Z_n)/n - m| \ge \varepsilon)$ の収束オーダーを評価するという点まで定理 13.18 の主張がカバーしているわけではない。収束オーダーの評価に関しては第 15 節で続けて検討する。

14 モーメント母関数とキュムラント母関数

第 15 節において大数の法則について収束のオーダーを評価するがその準備をここで行う。

14.1 定義. (i) $(\mathbb{R}, \mathrm{Borel}(\mathbb{R}))$ 上の確率測度 μ が次を満たすとき**指数可積分的**であるという。

$$\int_{\mathbb{R}} e^{\xi x}\,\mu(dx) < +\infty\ \forall \xi \in \mathbb{R}$$

このとき $\xi \mapsto \int_{\mathbb{R}} e^{\xi x}\,\mu(dx)$ を μ の**モーメント母関数**(moment generating function) という。
(ii) 実確率変数 X に対してその分布のモーメント母関数を X のモーメント母関数という。なお定理 6.38 によりそれは $\xi \mapsto E[e^{\xi X}]$ で与えられる。

14.2 補題. 指数可積分的な $(\mathbb{R}, \mathrm{Borel}(\mathbb{R}))$ 上の確率測度 μ のモーメント母関数は $\mathbb{C}$ 全体へ正則関数として拡張できる。その $c \in \mathbb{C}$ における n 次微分係数は $\int_{\mathbb{R}} x^n e^{cx}\,\mu(dx)$ で与えられる。

証明. $z \in \mathbb{C}$ とする。$a := |z-c| + |c|$ とおくと $|e^{zx}| \le e^{a|x|} \le e^{ax} + e^{-ax}\ \forall x \in \mathbb{R}$ であるから

$$\int_{\mathbb{R}} |e^{zx}|\,\mu(dx) \le \int_{\mathbb{R}} e^{(|z-c|+|c|)|x|}\,\mu(dx) \le \int_{\mathbb{R}} e^{ax}\mu(dx) + \int_{\mathbb{R}} e^{-ax}\mu(dx) < +\infty$$

が従う。中辺の被積分関数を整級数展開する。各項は非負値であり項別積分が許される。

$$\sum_{n=0}^{\infty} \int_{\mathbb{R}} \frac{|(z-c)x|^n}{n!} e^{|cx|}\mu(dx) = \int_{\mathbb{R}} e^{(|z-c|+|c|)|x|}\mu(dx) < +\infty.$$

ゆえに級数 $\sum_{n=0}^{\infty} \int_{\mathbb{R}} x^n e^{cx}/n!\,\mu(dx)(z-c)^n$ はすべての $z \in \mathbb{C}$ に対して絶対収束し、

$$\sum_{n=0}^{\infty} \frac{1}{n!} \int_{\mathbb{R}} x^n e^{cx}\mu(dx)(z-c)^n = \int_{\mathbb{R}} e^{zx}\mu(dx)\ \forall z \in \mathbb{C}$$

が成り立つ。左辺は c を中心とする Taylor 展開を与える。 □

特性関数とモーメント母関数はたがいに複素解析接続の関係にある。後者は $\mathbb{R}$ 上で正実数値であるから、その対数を偏角の曖昧さなしに定義できる。

14.3 定義. (I) 指数可積分的な $(\mathbb{R}, \mathrm{Borel}(\mathbb{R}))$ 上の確率測度 μ に対し $\mathbb{R} \to \mathbb{R}, \xi \mapsto \log \int_{\mathbb{R}} e^{\xi x}\,\mu(dx)$ を μ の**キュムラント母関数**(cumulant generating function) という。
(ii) 実確率変数 X に対してその分布のキュムラント母関数を X のキュムラント母関数という。記号としては $\mathrm{cml}(X, \cdot)$ を用いる。

記号

H を実確率変数であって $E[e^H] < +\infty$ を満たすものとする。このとき

$$\mathcal{F} \to \mathbb{R}, A \mapsto E[e^H; A]/E[e^H]$$

で定義される確率測度 P^H に関する期待値を E^H 分散を Var^H と表記する。即ち

$$E^H[X] = E[Xe^H]/E[e^H],\ \mathrm{Var}^H[X] = E^H[(X - E^H[X])^2].$$

14.4 補題. X を指数可積分的な分布に従う実確率変数とする。
(i) $\theta := \mathrm{cml}(X, \cdot)$ は解析的であり $\xi \in \mathbb{R}$ に対し $\theta'(\xi) = E^{\xi X}[X],\ \theta''(\xi) = \mathrm{Var}^{\xi X}[X]$ である。
(ii) X の分布は Dirac 測度ではないとする。このとき $\theta''(\xi) > 0\ \forall \xi \in \mathbb{R}$ が成り立つ。
(iii) $P(X \le x) = 1$ のとき $\mathbb{R}$ 上の関数 $\xi \mapsto \mathrm{cml}(X, \xi) - \xi x$ は単調非増加であり

$$\inf_{\xi \in \mathbb{R}} (\mathrm{cml}(X, \xi) - \xi x) = \log P(X = x) \text{ ここで } \log 0 = -\infty$$

$P(X \ge x) = 1$ なら $\xi \mapsto \mathrm{cml}(X, \xi) - \xi x$ は単調非減少であり下限は $\log P(X = x)$ である。
(iv) $P(X > x) > 0$ のとき $\lim_{\xi \to +\infty}(\mathrm{cml}(X, \xi) - \xi x) = +\infty$ が成り立つ。他方 $P(X < x) > 0$ なら $\lim_{\xi \to -\infty}(\mathrm{cml}(X, \xi) - \xi x) = +\infty$ である。

証明. (i) X のモーメント母関数を ϕ とかこう。補題 14.2 によれば ϕ は実解析的で

$$\phi'(\xi) = E[Xe^{\xi X}] = \phi(\xi)E^{\xi X}[X],\ \phi''(\xi) = E[X^2e^{\xi X}] = \phi(\xi)E^{\xi X}[X^2]\ \forall\xi \in \mathbb{R}$$

が成り立つ。実解析性は関数の合成により保存されるので $\theta = \log \circ \phi$ は実解析的である。また微分に関する連鎖律と演習問題 6.33(ii) で確認した関係により以下が得られる。

$$\theta'(\xi) = \frac{\phi'(\xi)}{\phi(\xi)} = E^{\xi X}[X],\ \theta''(\xi) = \frac{\phi''(\xi)}{\phi(\xi)} - \frac{\phi'(\xi)^2}{\phi(\xi)^2} = E^{\xi X}[X^2] - E^{\xi X}[X]^2 = \mathrm{Var}^{\xi X}[X].$$

(ii) X の分布は Dirac 測度ではないことより $P(X \neq x) > 0\ \forall x \in \mathbb{R}$ が成り立つ。他方 $e^{\xi X(\omega)} > 0\ \forall\omega \in \Omega$ であるから定理 6.31(ii) を適用して次が導ける。

$$P^{\xi X}(X \neq x) = E[e^{\xi X}; X \neq x]/E[e^{\xi X}] > 0\ \forall x \in \mathbb{R}.$$

従って演習問題 6.33(i) で確認したことより $\theta''(\xi) = \mathrm{Var}^{\xi X}[X] > 0$ が成り立つ。

(iii), (iv) $\mathrm{cml}(X,\xi) - \xi x = \log E[e^{\xi(X-x)}]$ を使えばすぐに分かる。 □

14.5 演習問題. X を指数可積分的な分布に従う実確率変数とする。$\theta := \mathrm{cml}(X,\cdot)$ の微分係数と X のモーメントには次の関係が成り立つことを示せ。

$$E[X^n] = \theta^{(n)}(0) + \sum_{k=1}^{n-1} \binom{n-1}{k} E[X^k]\theta^{(n-k)}(0)\ n \in \mathbb{N}.$$

14.6 定義. X を指数可積分的な分布に従う実確率変数とする。そのキュムラント母関数の原点における n 次微分係数を X の n **次キュムラント**(cumulant) という。

記号

指数可積分的な分布に従う実確率変数 X に対し関数 $\mathbb{R} \to \overline{\mathbb{R}}$ を次で導入する。

$$\beta(X,\cdot) : x \mapsto \begin{cases} \inf\{\xi \in \mathbb{R} : E^{\xi X}[X] > x\} & x < E[X] \\ 0 & x = E[X] \\ \sup\{\xi \in \mathbb{R} : E^{\xi X}[X] < x\} & x > E[X] \end{cases}$$

$$\mathrm{cml}^*(X,\cdot) : x \mapsto \inf_{\xi\in\mathbb{R}}(\mathrm{cml}(X,\xi) - \xi x)$$

一般に関数 $\theta : \mathbb{R} \to \mathbb{R}$ が与えられたとき関数 $x \mapsto \inf_{\xi\in\mathbb{R}}(\theta(\xi) - \xi x)$ を θ の *Legendre* **変換**(Legendre transform) という。

14.7 例. X を Dirac 測度を分布に持つ実確率変数とする。即ち $P(X = a) = 1$ を満たす $a \in \mathbb{R}$ が存在する。キュムラント母関数 $\theta := \mathrm{cml}(X,\cdot)$ に対して次が成り立つ。

$\theta(\xi) = \xi a.\ \theta'(\xi) = a.\ \theta''(0) = 0.$

	$x < a$	$x = a$	$x > a$
$\beta(X,\cdot)$	$-\infty$	0	$+\infty$
$\mathrm{cml}^*(X,\cdot)$	$-\infty$	0	$-\infty$

14.8 例. X を実確率変数であって $P(X=1)=p$, $P(X=-1)=1-p$ を満たすものとする。ただし $0<p<1$ である。キュムラント母関数 $\theta := \mathrm{cml}(X,\cdot)$ に対して次が成り立つ。

$$\theta(\xi)=\log(pe^{\xi}+(1-p)e^{-\xi}).\ \theta'(\xi)=(pe^{\xi}+(p-1)e^{-\xi})/(pe^{\xi}+(1-p)e^{-\xi}).$$
$$-1<\theta'(\xi)<1.\ \theta''(\xi)=1-\theta'(\xi)^2>0.\ \theta'(0)=2p-1.\ \theta''(0)=4p(1-p).$$

	$x<-1$	$x=-1$	$-1<x<1$	$x=1$	$x>1$
$\beta(X,\cdot)$	$-\infty$	$-\infty$	$\frac{1}{2}\log\frac{(1-p)(1+x)}{p(1-x)}$	$+\infty$	$+\infty$
$\mathrm{cml}^*(X,\cdot)$	$-\infty$	$\log(1-p)$	$\frac{1-x}{2}\log\frac{2(1-p)}{1-x}+\frac{1+x}{2}\log\frac{2p}{1+x}$	$\log p$	$-\infty$

14.9 例. X を平均 $a\in\mathbb{R}$ 分散 $t\in\mathbb{R}_{>0}$ の正規分布に従う実確率変数とする。キュムラント母関数 $\theta := \mathrm{cml}(X,\cdot)$ に対して次が成り立つ。

$$\theta(\xi)=a\xi+t\xi^2/2.\ \theta'(\xi)=a+t\xi.\ \theta''(\xi)=t>0.$$

$$\beta(X,x)=(x-a)/t.\ \mathrm{cml}^*(X,x)=-(x-a)^2/(2t)$$

14.10 演習問題. 例 14.7、例 14.8 と例 14.9 を確認せよ。

Dirac 測度ではない指数可積分的な分布に従う実確率変数のキュムラント母関数に対してその Legendre 変換の持つ性質を調べておこう。

14.11 補題. X を Dirac 測度ではない指数可積分的な分布に従う実確率変数とする。
(i) 各 $x\in\mathbb{R}$ に対して次の分類表が成り立つ。

$\beta(X,x)$	$E[X]$	$P(X<x)$	$P(X>x)$
$-\infty$	$(x,+\infty)$	0	>0
$(-\infty,0)$	$(x,+\infty)$	>0	>0
0	x	>0	>0
$(0,+\infty)$	$(-\infty,x)$	>0	>0
$+\infty$	$(-\infty,x)$	>0	0

(ii) $P(X<x)P(X>x)=0$ のとき次の増減表が成り立つ。

$P(X<x)=0$ のとき

	$\beta(X,x)=-\infty$	$\mathbb{R}$	$+\infty$
$\xi\mapsto E^{\xi X}[X]-x$	$*$	>0	$*$
$\xi\mapsto \mathrm{cml}(X,\xi)-\xi x$	$\log P(X=x)$	増加	$+\infty$

$P(X>x)=0$ のとき

	$-\infty$	$\mathbb{R}$	$\beta(X,x)=+\infty$
$\xi\mapsto E^{\xi X}[X]-x$	$*$	<0	$*$
$\xi\mapsto \mathrm{cml}(X,\xi)-\xi x$	$+\infty$	減少	$\log P(X=x)$

(iii) $P(X<x)P(X>x)>0$ のとき $-\infty<\beta(X,x)<+\infty$ であり次の増減表が成り立つ。

	$-\infty$	$(-\infty,\beta(X,x))$	$\beta(X,x)$	$(\beta(X,x),+\infty)$	$+\infty$
$\xi\mapsto E^{\xi X}[X]-x$	$*$	<0	0	>0	$*$
$\xi\mapsto \mathrm{cml}(X,\xi)-\xi x$	$+\infty$	減少	最小値	増加	$+\infty$

証明. (ii) まず $P(X < x) = 0$ すなわち $P(X \geq x) = 1$ とする。補題 14.4(iii) によれば $\xi \mapsto \mathrm{cml}(X,\xi) - \xi x$ は非減少でありそれは $\xi \to -\infty$ の極限で $\log P(X = x)$ に収束する（$P(X = x) = 0$ のときは $-\infty$ に定発散するとよむ）。また補題 14.4(i), (ii) よりその導関数は $\xi \mapsto E^{\xi X}[X] - x$ であり狭義に増加する。従ってもとの関数の非減少性を考慮に入れて $E^{\xi X}[X] - x > 0\ \forall\xi \in \mathbb{R}$ が得られる。これから直ちに $\beta(X, x) = -\infty$ が分かる。他方 $P(X = x) < 1$ であるから $P(X > x) > 0$ となり、補題 14.4(iv) を適用して $\xi \to +\infty$ の極限で $+\infty$ に定発散することが分かる。$P(X > x) = 0$ の場合についても同様の議論である。

(iii) $P(X < x) > 0$ かつ $P(X > x) > 0$ であるから補題 14.4(iv) を適用して $\xi \to -\infty$, $\xi \to +\infty$ のどちらの極限でも関数 $\xi \mapsto \mathrm{cml}(X,\xi) - \xi x$ は $+\infty$ に定発散することが分かる。従ってこの関数は最小値を持つ。補題 14.4(i), (ii) よりその導関数は $\xi \mapsto E^{\xi X}[X] - x$ であり狭義に増加する。従って極小点は高々一つなので関数の最小値を実現する点はただ一つ存在することになる。ここで $E^{\xi X}[X]|_{\xi=0} = E[X]$ に注意しつつ $\beta(X, x)$ の定義に従うとそれが有限値であることならびに増減表の内容が分かる。

(i) $E^{\xi X}[X]|_{\xi=0} = E[X]$ および (ii), (iii) から分類表が得られる。 □

記号

実確率変数 X に対し $\mathrm{cnvx}(X) := \{x \in \mathbb{R} : P(X < x)P(X > x) > 0\}$ とおく。

14.12 補題. X を Dirac 測度ではない指数可積分的な分布に従う実確率変数とする。
(i) $\mathrm{cnvx}(X)$ は $E[X]$ を要素とする開区間である。
(ii) $\xi \mapsto E^{\xi X}[X]$ は狭義に増加し次が成り立つ。

$$\lim_{\xi \to -\infty} E^{\xi X}[X] = \inf \mathrm{cnvx}(X), \ \lim_{\xi \to +\infty} E^{\xi X}[X] = \sup \mathrm{cnvx}(X).$$

(iii) $\xi \mapsto E^{\xi X}[X]$ の逆写像は $\beta(X, \cdot)$ を $\mathrm{cnvx}(X)$ に制限したものである。

証明. (i) X の分布は Dirac 測度ではないので $P(X < E[X]) > 0$ または $P(X > E[X]) > 0$ である。他方、次が成り立つので $P(X < E[X]) > 0$ と $P(X > E[X]) > 0$ は同値である。

$$0 \leq E[E[X] - X; X < E[X]] = E[X - E[X]; X > E[X]].$$

よって $P(X < E[X]) > 0$ かつ $P(X > E[X]) > 0$ が導けた。すなわち $E[X] \in \mathrm{cnvx}(X)$ である。関数 $x \mapsto P(X < x)$ は非減少かつ左連続であり、関数 $x \mapsto P(X > x)$ は非増加かつ右連続なので $\mathrm{cnvx}(X)$ は開区間である。

(ii) $\xi < 0$, $x \in \mathrm{cnvx}(X)$ とする。まず次のように上から評価しておく。

$$\begin{aligned} E[Xe^{\xi X}] &\leq xE[e^{\xi X}; X < x] + E[Xe^{\xi X}; X \geq x] \\ &= xE[e^{\xi X}] + E[(X - x)e^{\xi X}; X \geq x] \leq xE[e^{\xi X}] + e^{\xi x}E[(X - x); X \geq x] \end{aligned}$$

ここで $\xi < 0$ であることに注意せよ。補題 6.34 により $\delta > 0$ に対して

$$P(X < x - \delta) = P(e^{\xi X} > e^{\xi(x-\delta)}) \leq e^{-\xi(x-\delta)}E[e^{\xi X}]$$

(i) により $x-\delta \in \mathrm{cnvx}(X)$ を満たす $\delta > 0$ が存在する。$P(X < x-\delta) > 0$ であるから

$$0 \le \frac{e^{\xi x} E[(X-x); X \ge x]}{E[e^{\xi X}]} \le \frac{e^{\xi\delta} E[(X-x); X \ge x]}{P(X < x-\delta)}$$

が成り立つ。$\xi \to -\infty$ の極限で右辺は 0 に収束する。従って

$$\limsup_{\xi \to -\infty} \frac{E[X e^{\xi X}]}{E[e^{\xi X}]} \le x \ \forall x \in \mathrm{cnvx}(X).$$

他方、明らかに $\inf \mathrm{cnvx}(X) \le E^{\xi X}[X] \ \forall \xi \in \mathbb{R}$ が成り立つ。以上より

$$\lim_{\xi \to -\infty} E^{\xi X}[X] = \inf \mathrm{cnvx}(X)$$

が導けた。$\lim_{\xi \to +\infty} E^{\xi X}[X] = \sup \mathrm{cnvx}(X)$ についても同様の議論である。補題 14.4(i), (ii) より $\xi \mapsto E^{\xi X}[X]$ は微分可能でありその導関数は正実数値をとる。よって狭義に増加し、従って単写である。その像は $\mathrm{cnvx}(X)$ に一致する。

(iii) さて補題 14.11(iii) によれば $x \in \mathrm{cnvx}(X)$ のとき $\beta(X,x)$ は有限値でありまた微分可能関数 $\xi \mapsto \mathrm{cml}(X,\xi) - \xi x$ の最小点であった。従ってその導関数 $\xi \mapsto E^{\xi X}[X] - x$ が値 0 を取る点であるから $\xi \mapsto E^{\xi X}[X]$ の逆写像は $\beta(X,\cdot)$ を $\mathrm{cnvx}(X)$ に制限したもので与えられる。 □

14.13 定理. X を Dirac 測度ではない指数可積分的な分布に従う実確率変数とする。
(i) 関数 $\beta(X,\cdot)$, $\mathrm{cml}^*(X,\cdot)$ は $\mathrm{cnvx}(X)$ 上で実解析的である。
(ii) $\mathrm{cml}^*(X,\cdot)$ の導関数は $-\beta(X,\cdot)$ に等しくまた 2 階導関数は負実数値をとる。

	$(\inf \mathrm{cnvx}(X), E[X])$	$E[X]$	$(E[X], \sup \mathrm{cnvx}(X))$
$-\beta(X,\cdot)$	> 0（狭義減少）	0	< 0（狭義減少）
$\mathrm{cml}^*(X,\cdot)$	< 0（狭義増加）	0（最大値）	< 0（狭義減少）

(iii) $\mathrm{cml}^*(X,x) \le 0 \ \forall x \in \mathbb{R}$. $\mathrm{cml}^*(X,x) = -\infty \Leftrightarrow P(X \le x) = 0$ あるいは $P(X \ge x) = 0$.
$I := \{x \in \mathbb{R} : \mathrm{cml}^*(X,x) > -\infty\}$ は区間でありかつ $P(X \in I) = 1$ が成り立つ。
(iv) $a := \inf \mathrm{cnvx}(X) > -\infty$ なら $\lim_{x \downarrow\downarrow a} \mathrm{cml}^*(X,x) = \log P(X=a) = \mathrm{cml}^*(X,a)$ であり $b := \sup \mathrm{cnvx}(X) < +\infty$ なら $\lim_{x \uparrow\uparrow b} \mathrm{cml}^*(X,x) = \log P(X=b) = \mathrm{cml}^*(X,b)$ である。
(v) $\inf \mathrm{cnvx}(X) = -\infty$ なら $\lim_{x \to -\infty} \mathrm{cml}^*(X,x) = -\infty$ であり $\sup \mathrm{cnvx}(X) = +\infty$ なら $\lim_{x \to +\infty} \mathrm{cml}^*(X,x) = -\infty$ である。

証明. (i), (ii), (iii) 補題 14.12(iii) によれば関数 $\beta(X,\cdot)$ を $\mathrm{cnvx}(X)$ 上に制限したものは実解析的かつ導関数が正実数値をとるものの逆関数である。よって $\beta(X,\cdot)$ は実解析的であり、その導関数は正実数値をとる。補題 14.11(iii) によれば $\beta(X,x)$ は微分可能関数 $\xi \mapsto \mathrm{cml}(X,\xi) - \xi x$ の最小点であり、最小値は $\mathrm{cml}^*(X,x)$ ということであったので

$$(\star) \qquad \mathrm{cml}^*(X,x) = \mathrm{cml}(X, \beta(X,x)) - \beta(X,x)x \ \forall x \in \mathrm{cnvx}(X)$$

が成り立つ。実解析的関数の合成ゆえ $\mathrm{cml}^*(X,\cdot)$ もまたそうである。$\beta(X,x)$ は微分可能関数 $\xi \mapsto \mathrm{cml}(X,\xi) - \xi x$ の導関数が値 0 を取る点であることを使って $\mathrm{cml}^*(X,\cdot)$ の導関数が

$-\beta(X,\cdot)$ に等しいことを得る。さて補題 14.12(i) より $E[X] \in \mathrm{cnvx}(X)$ である。そこで $\beta(X,\cdot)$ は値 0 をとり、また $(\star)$ により次が成り立つ。

$$\mathrm{cml}^*(X, E[X]) = (\mathrm{cml}(X,\xi) - \xi E[X])|_{\xi=0} = 0.$$

以上より増減表が完成する。さて $x \notin \mathrm{cnvx}(X)$ なら補題 14.11(ii) より

$$\mathrm{cml}^*(X,x) = \log P(X=x) < 0$$

である。よって $\mathrm{cml}^*(X,x) = -\infty$ は $P(X=x)=0$ かつ $x \notin \mathrm{cnvx}(X)$ と同値である。

(iv) $\inf \mathrm{cnvx}(X) > -\infty$ とする。補題 14.12(i) を示す際に注意したように $P(X<a)=0$ である。従って補題 14.11(ii) により $\mathrm{cml}^*(X,a) = \log P(X=a)$ である。(ii) によれば $\mathrm{cml}^*(X,\cdot)$ は $(a, E[X])$ 上で非減少である。そこで $(\star)$ と補題 14.12(iii) を考慮に入れて

$$\begin{aligned}\lim_{x\to a: x\in \mathrm{cnvx}(X)} \mathrm{cml}^*(X,x) &= \lim_{x\to a: x\in \mathrm{cnvx}(X)} (\mathrm{cml}(X,\beta(X,x)) - \beta(X,x)x) \\ &= \lim_{\xi\to-\infty} (\mathrm{cml}(X,\xi) - \xi E^{\xi X}[X])\end{aligned}$$

を得る。$\xi < 0$ とする。$a \le E^{\xi X}[X]$ であるから次が成り立つ。

$$\mathrm{cml}(X,\xi) - \xi a \le \mathrm{cml}(X,\xi) - \xi a + \xi(a - E^{\xi X}[X]) = \mathrm{cml}(X,\xi) - \xi E^{\xi X}[X].$$

補題 14.4(iii) により左辺は $\xi \to -\infty$ の極限で $\log P(X=a)$ に収束する。よって

$$\mathrm{cml}^*(X,a) = \log P(X=a) \le \lim_{x\to a: x\in \mathrm{cnvx}(X)} \mathrm{cml}^*(X,x)$$

である。上からの評価を得るため $y > \mathrm{cml}^*(X,a)$ とする。次が成り立つ。

$$\begin{aligned}\{x\in\mathbb{R} : \mathrm{cml}^*(X,x) < y\} &= \{x\in\mathbb{R} : \inf_{\xi\in\mathbb{R}}(\mathrm{cml}(X,\xi) - \xi x) < y\} \\ &= \bigcup_{\xi\in\mathbb{R}} \{x\in\mathbb{R} : \mathrm{cml}(X,\xi) - \xi x < y\}.\end{aligned}$$

右辺は開集合の合併であるからやはり開集合である。また左辺は a を要素とする。従って a のある近傍は $\{x\in\mathbb{R} : \mathrm{cml}^*(X,x) < y\}$ に含まれる。即ち

$$|x-a| < \delta \Rightarrow \mathrm{cml}^*(X,x) < y.$$

を満たすような $\delta > 0$ が存在する。よって

$$\lim_{x\to a: x\in \mathrm{cnvx}(X)} \mathrm{cml}^*(X,x) \le y \ \forall y > \mathrm{cml}^*(X,a).$$

下からの評価とあわせて $\lim_{x\to a: x\in \mathrm{cnvx}(X)} \mathrm{cml}^*(X,x) = \mathrm{cml}^*(X,a)$ を得る。

(v) $c := \lim_{x\upuparrows \sup \mathrm{cnvx}(X)} \mathrm{cml}^*(X,x) > -\infty$ と仮定する。$x \in \mathrm{cnvx}(X)$, $x > E[X]$ なら (ii) より $c \le \mathrm{cml}^*(X,x) \le 0$ が成り立つ。よって

$$c \le \mathrm{cml}(X,\xi) - \xi x \ \forall \xi\in\mathbb{R} \ \forall x \in (E[X], \sup \mathrm{cnvx}(X)).$$

特に $\xi = 1$ と選ぶと $x \le \mathrm{cml}(X,1) - c \ \forall x \in (E[X], \sup \mathrm{cnvx}(X))$ であることが分かる。よって $\sup \mathrm{cnvx}(X) < +\infty$ である。対偶を取って結論を得る。 □

15 大偏差原理

定理 13.2 における $P(|(X_1+\cdots+X_n)/n-m|\geq\varepsilon)$ の収束オーダーを第 14 節で導入された概念を使って定量的に評価する。

前提

μ を Dirac 測度ではない指数可積分的な $(\mathbb{R},\mathrm{Borel}(\mathbb{R}))$ 上の確率測度とする。
$X_1,X_2,\ldots,X_n,\ldots$ をすべて分布 μ に従う独立な確率変数の列とする。

次の命題は Cramér によって得られたものの一部分である。

15.1 補題. $x\in\mathrm{cnvx}(X_1)$ と仮定し確率変数列

$$Y(n):=\beta(X_1,x)(X_1+\cdots+X_n)\ n\in\mathbb{N}$$

を導入する。このとき各 $y>x$ に対して $P^{Y(n)}(nx<X_1+\cdots+X_n<ny)$ は $1/2$ に収束する。

証明. 補題 14.11(iii) より $-\infty<\beta(X_1,x)<+\infty$ であることを確認しておこう。$\xi_1,\ldots,\xi_n\in\mathbb{R}$ とする。$X_1,X_2,\ldots,X_n,\ldots$ は独立なので定理 8.33 により次が成り立つ。

$$\begin{aligned}&\frac{E[\exp\{(\sqrt{-1}\xi_1+\beta(X_1,x))X_1+\cdots+(\sqrt{-1}\xi_n+\beta(X_1,x))X_n\}]}{E[\exp\{\beta(X_1,x)X_1+\cdots+\beta(X_1,x)X_n\}]}\\&=\prod_{k=1}^{n}\frac{E[\exp\{(\sqrt{-1}\xi_k+\beta(X_1,x))X_k\}]}{E[\exp\{\beta(X_1,x)X_k\}]}=\prod_{k=1}^{n}E^{Y(1)}[\exp\{\sqrt{-1}\xi_kX_1\}].\end{aligned}$$

2 番目の等号はすべて同じ分布 μ を持つことから従う。ここで $\xi\in\mathbb{R}$ を固定して $\xi_1=\cdots=\xi_n=\xi/\sqrt{vn}$ とおき両辺に $\exp\{\sqrt{-1}\xi x\sqrt{n/v}\}$ をかけると

$$(\star)\quad E^{Y(n)}[\exp\{\sqrt{-1}\xi(X_1+\cdots+X_n-nx)/\sqrt{vn}\}]=E^{Y(1)}[\exp\{\sqrt{-1}\xi(X_1-x)/\sqrt{vn}\}]^n$$

さて $Y(1)=\beta(X_1,x)X_1$ である。よって以下において補題 14.12(iii) により最初の関係が、また補題 14.4(i), (ii) により 2 番目の関係が導かれる。

$$E^{Y(1)}[X_1]=x,\ v:=\mathrm{Var}^{Y(1)}[X_1]>0.$$

補題 11.6 によれば $(\star)$ 右辺は $e^{-\xi^2/2}$ に収束する。定理 11.18 の証明で使った論法を適用して

$$\lim_{n\to\infty}P^{Y(n)}(0<(X_1+\cdots+X_n-nx)/\sqrt{vn}<b)=\Phi(b)-\Phi(0)\ \forall b\in\mathbb{R}_{>0}$$

を得る。ここで誤差関数 Φ は $\Phi(0)=1/2$ を満たしている。$y>x$, $b>0$ とする。$n\in\mathbb{N}$ が十分大きい限り $nx+b\sqrt{vn}\leq ny$ が成り立ち、従って

$$\begin{aligned}&P^{Y(n)}(0<(X_1+\cdots+X_n-nx)/\sqrt{vn}<b)\\&\leq P^{Y(n)}(nx<X_1+\cdots+X_n<ny)\leq P^{Y(n)}(0<(X_1+\cdots+X_n-nx)/\sqrt{vn})\end{aligned}$$

である。以上より $n \to \infty$ の極限において次が成り立つ。

$$\begin{aligned}\Phi(b) - 1/2 &\leq \liminf_{n\to\infty} P^{Y(n)}(nx < X_1 + \cdots + X_n < ny) \\ &\leq \limsup_{n\to\infty} P^{Y(n)}(nx < X_1 + \cdots + X_n < ny) \leq 1/2 \ \forall b \in \mathbb{R}_{>0}.\end{aligned}$$

$\lim_{b\to+\infty} \Phi(b) = 1$ であるから $P^{Y(n)}(nx < X_1 + \cdots + X_n < ny)$ は $1/2$ に収束する。 □

補題 15.1 では基礎となる確率測度を $P^{Y(n)}$ に取り替えている。確率空間 $(\Omega, \mathcal{F}, P^{Y(n)})$ では確率変数 X_1 の平均値は着目している値 x にちょうど等しい。また確率変数のシステム $X_1, X_2, \ldots, X_n$ は独立であり同分布に従うという構造が保たれていることも重要である。測度 P から測度 $P^{Y(n)}$ への取り替えを *Cramér* **変換**(Cramér transformation) という。

15.2 補題. I を閉区間（有界でなくてもよい）であって $\inf I < \sup I$ を満たすものとする。
(i) $a := \inf I \geq E[X_1]$ なら $n^{-1} \log P((X_1 + \cdots + X_n)/n \in I)$ は $\mathrm{cml}^*(X_1, a)$ に収束する。
(ii) $b := \sup I \leq E[X_1]$ なら $n^{-1} \log P((X_1 + \cdots + X_n)/n \in I)$ は $\mathrm{cml}^*(X_1, b)$ に収束する。
(iii) $\inf I < E[X_1] < \sup I$ なら $n^{-1} \log P((X_1 + \cdots + X_n)/n \in I)$ は 0 に収束する。

証明. (i) 補題 14.11(i) より条件 $a \geq E[X_1]$ は $\beta(X_1, a) < 0$ を排除するので次が成り立つ。

$$\inf_{\xi\geq 0}(\mathrm{cml}(X_1, \xi) - \xi a) = \mathrm{cml}^*(X_1, a).$$

$a \in I \subset [a, +\infty)$ に注意する。補題 6.34 により $P(X_1 + \cdots + X_n \geq na)$ は

$$P(\exp\{\xi(X_1 + \cdots + X_n)\} \geq e^{\xi na}) \leq e^{-\xi na} E[\exp\{\xi(X_1 + \cdots + X_n)\}] \ \forall \xi \in \mathbb{R}_{\geq 0}$$

と評価される。定理 8.33 によれば右辺は $\exp\{n(\mathrm{cml}(X_1, \xi) - \xi a)\}$ に等しい。よって

$$P(X_1 + \cdots + X_n \geq na) \leq \exp\{n \,\mathrm{cml}^*(X_1, a)\} \ \forall n \in \mathbb{N}$$

を得る。場合分けを行う。$P(X_1 \geq a) = 0$ のとき次が成り立つ。

$$P(X_1 + \cdots + X_n \geq na) = 0 \ \forall n \in \mathbb{N} \text{ かつ } \mathrm{cml}^*(X_1, a) = -\infty.$$

後者は定理 14.13(iii) を適用して分かる。よって主張は明らかである。$P(X_1 > a) = 0$ かつ $P(X_1 = a) > 0$ のとき $P(X_1 + \cdots + X_n \in nI) = P(X_1 = a, \ldots, X_n = a)$ であるから

$$P(X_1 + \cdots + X_n \in nI) = P(X = a)^n \ \forall n \in \mathbb{N}$$

（$P(X_1 + \cdots + X_n > na) = 0 \neq P(X_1 = a)^n \ \forall n \in \mathbb{N}$ に注意せよ。）また補題 14.11(ii) より

$$\mathrm{cml}^*(X, a) = \log P(X_1 = a) > -\infty.$$

よって主張は明らかである。残るは $P(X_1 > a) > 0$ かつ $P(X_1 < a) > 0$ の場合である。このとき補題 14.11(i) より $0 \leq \beta(X_1, a) < +\infty$ である。確率変数列

$$Y(n) := \beta(X_1, a)(X_1 + \cdots + X_n) \ n \in \mathbb{N}$$

を導入する。$y > a$ とする。$P(na < X_1 + \cdots + X_n < ny)$ は各 $\xi \in \mathbb{R}_{\geq 0}$ に対し

$$e^{-\xi ny} E[\exp\{\xi(X_1 + \cdots + X_n)\}; na < X_1 + \cdots + X_n < ny]$$

で下から評価される。これを $E[\exp\{\xi(X_1 + \cdots + X_n)\}] = \exp\{n\,\mathrm{cml}(X_1, \xi)\}$ を使って

$$\exp\{n(\mathrm{cml}(X_1, \xi) - \xi y)\} P^{\xi(X_1 + \cdots + X_n)}(na < X_1 + \cdots + X_n < ny)$$

と書き換える。ここで補題 14.11(iii) より

$$(\mathrm{cml}(X_1, \xi) - \xi y)|_{\xi = \beta(X_1, a)} = \mathrm{cml}^*(X_1, a) - \beta(X_1, a)(y - a)$$

である。従って $\xi = \beta(X_1, a)$ と選ぶことにより

$$\begin{aligned}&\exp\{n\,\mathrm{cml}^*(X, a) - n\beta(X_1, a)(y - a)\} P^{Y(n)}(na < X_1 + \cdots + X_n < ny) \\ &\leq P(na < X_1 + \cdots + X_n < ny) \leq P(X_1 + \cdots + X_n \in nI^{\mathrm{int}}).\end{aligned}$$

ここで y は $a < y < \sup I$ となるように選んでいる。補題 15.1 を適用しよう。

$$\mathrm{cml}^*(X, a) - \beta(X_1, a)(y - a) \leq \liminf_{n \to \infty} \frac{1}{n} \log P(X_1 + \cdots + X_n \in nI^{\mathrm{int}}).$$

この段階で y は $a < y < \sup I$ である限り任意である。上からの評価とあわせて

$$\begin{aligned}\mathrm{cml}^*(X, a) &\leq \liminf_{n \to \infty} \frac{1}{n} \log P(X_1 + \cdots + X_n \in nI^{\mathrm{int}}) \\ &\leq \limsup_{n \to \infty} \frac{1}{n} \log P(X_1 + \cdots + X_n \in nI) \leq \mathrm{cml}^*(X, a)\end{aligned}$$

が導かれた。(ii) についても同様の議論である。

(iii) この場合は大数の弱法則により $P((X_1 + \cdots + X_n)/n \in I)$ は 1 に収束する。 □

15.3 系. (i) $x > E[X_1]$ のとき $\sum_{n=1}^{\infty} P(X_1 + \cdots + X_n \geq nx) < +\infty$ である。
(ii) $x < E[X_1]$ のとき $\sum_{n=1}^{\infty} P(X_1 + \cdots + X_n \leq nx) < +\infty$ である。

証明. 定理 14.13(iii) によれば $x \neq E[X_1]$ のとき $\mathrm{cml}^*(X, x) < 0$ である。 □

Cramér **の定理**（補題 15.2）の内容を雑に表現すると次のようになる。

$a := \inf I \geq E[X_1]$ なら $P((X_1 + \cdots + X_n)/n \in I)$ は漸近的に $\exp\{\mathrm{cml}^*(X_1, a)n\}$ である。

$a = \inf I > E[X_1]$ のとき $\mathrm{cml}^*(X_1, a) < 0$ であるから確率 $P((X_1 + \cdots + X_n)/n \in I)$ は 0 に収束する。関数 $\mathrm{cml}^*(X_1, \cdot)$ は収束の指数的レートを決めているという意味で**レート関数**(rate function) と呼ばれる。補題 15.2 ではそれを X_1 のキュムラント母関数から Legendre 変換によって与えていたのである。

$a := \inf I \geq E[X_1]$ として話を続ける。関数 $\mathrm{cml}^*(X_1, \cdot)$ は収束の指数的レートと見なせるのだが、定理 14.13 によればそれは区間 $(E[X_1], +\infty)$ では非増加であり、従って $\mathrm{cml}^*(X_1, a)$ は区間 I 上における最大値である。他方、$\inf I < E[X_1] < \sup I$ ならレート値は 0 であるが、それもやはり区間 I 上における $\mathrm{cml}^*(X_1, \cdot)$ の最大値である。即ち $(X_1 + \cdots + X_n)/n \in I$ という事象の確率は区間 I の中でもレート関数の大きい部分によって支配されている。これを一般に**大偏差原理**(large deviation principle) が成立するという。

15.4 系. I を空でない開区間、J をその閉包とする。
(i) $n^{-1}\log P((X_1+\cdots+X_n)/n\in I)$ は $\sup_{x\in I}\mathrm{cml}^*(X_1,x)$ に収束する。
(ii) $n^{-1}\log P((X_1+\cdots+X_n)/n\in J)$ は $\max_{x\in J}\mathrm{cml}^*(X_1,x)$ に収束する。

証明. (ii) については既に上に述べたとおりである。

(i) $a:=\inf I\geq E[X_1]$ かつ $\inf I\in\mathrm{cnvx}(X_1)$ とする。このとき、補題 15.2 の証明によれば、$n^{-1}\log P((X_1+\cdots+X_n)/n\in I)$ は $\mathrm{cml}^*(X_1,a)$ に収束する。他方、定理 14.13 より

$$\sup_{x\in I}\mathrm{cml}^*(X_1,x)=\max_{x\in J}\mathrm{cml}^*(X_1,x)=\mathrm{cml}^*(X_1,a)$$

である。$\sup I\leq E[X_1]$ かつ $\sup I\in\mathrm{cnvx}(X_1)$ の場合も同様の議論ができる。また $E[X_1]\in I$ なら大数の弱法則により $0=\max_{x\in I}\mathrm{cml}^*(X_1,x)$ に収束する。その他に検討されていないのは $\sup I=\inf\mathrm{cnvx}(X_1)>-\infty$ または $\inf I=\sup\mathrm{cnvx}(X_1)<+\infty$ の場合である。このときは $\sup_{x\in I}\mathrm{cml}^*(X_1,x)=-\infty$ かつ $P((X_1+\cdots+X_n)/n\in I)=0$ である。 □

確率変数 X_1 の分布を特定すると詳しい漸近挙動が分かる。次の問題は一般論に頼ることなしに具体的に計算できる例である。

15.5 演習問題. (i) $x>0$ に対して次が成り立つことを示せ。

$$\int_{[x,+\infty)}e^{-y^2/2}\,\lambda(dy)=\frac{1}{x}e^{-x^2/2}-\int_{[x,+\infty)}\frac{1}{y^2}e^{-y^2/2}\,\lambda(dy).$$

(ii) $x>0$ に対して次が成り立つことを示せ。

$$\frac{x}{x^2+1}e^{-x^2/2}\leq\int_{[x,+\infty)}e^{-y^2/2}\,\lambda(dy)\leq\frac{1}{x}e^{-x^2/2}.$$

(iii) $X_1,X_2,\ldots,X_n,\ldots$ を平均 $a\in\mathbb{R}$ 分散 $t\in\mathbb{R}_{>0}$ の正規分布に従う独立な確率変数の列とする。$x>a$ のとき次が成り立つことを示せ。更に補題 15.2(i) と比較検討せよ。

$$\lim_{n\to\infty}\sqrt{\frac{2\pi n(x-a)}{t}}\exp\left\{\frac{n(x-a)^2}{2t}\right\}P((X_1+\cdots+X_n)/n\geq x)=1.$$

系 15.3 から確率変数列 X_n について大数の強法則を導くことができる。命題自身は定理 13.18 でカバーされているが *Borel-Cantelli* **の補題**（補題 15.6）を運用する典型を紹介するために証明を述べておくことにしよう。

15.6 補題. $A_n\in\mathcal{F}\ \forall n\in\mathbb{N},\ \sum_{n=1}^{\infty}P(A_n)<+\infty\Rightarrow P(\bigcap_{n=1}^{\infty}\bigcup_{k\geq n}A_k)=0.$

証明. 単調収束定理により $P(\bigcap_{n=1}^{\infty}\bigcup_{k\geq n}A_k)=\inf_{n\in\mathbb{N}}P(\bigcup_{k\geq n}A_k)$ である。劣加法性を使うと右辺は $\inf_{n\in\mathbb{N}}\sum_{k\geq n}P(A_k)=0$ で抑えられる。 □

15.7 系. 確率変数列 $(X_1+\cdots+X_n)/n$ は $E[X_1]$ に概収束する。

証明. $x > E[X_1]$ とする。系 15.3 を考慮に入れて、事象列 $A_n := \{(X_1 + \cdots + X_n)/n \geq x\}$ に対して補題 15.6 を適用する。以下の集合は P 零集合であることが分かる。

$$N_1(x) := \bigcap_{n=1}^{\infty} \bigcup_{k \geq n} \{(X_1 + \cdots + X_k)/k \geq x\}$$

さて $\omega \in \Omega \setminus N_1(x)$ ならある $n \in \mathbb{N}$ が存在して $(X_1(\omega) + \cdots + X_k(\omega))/k < x \ \forall k \geq n$ が成り立つ。従って以下が得られた。

$$\limsup_{n\to\infty} \frac{X_1(\omega) + \cdots + X_k(\omega)}{n} \leq x \ \forall \omega \in \Omega \setminus N_1(x).$$

各 $x > E[X_1]$ に対して $\Omega \setminus N_1(x)$ は P-a.s. 集合であるがそのようなものの可算共通集合も P-a.s. 集合である。そこで次の集合を導入する。

$$\Omega_1 := \bigcap_{l \in \mathbb{N}} (\Omega \setminus N_1(E[X_1] + 1/l)).$$

このとき P-a.s. 集合 Ω_1 上で $\limsup_{n\to\infty}(X_1 + \cdots + X_n)/n \leq E[X_1]$ が成り立つ。後者は

$$\limsup_{n\to\infty} \frac{X_1(\omega) + \cdots + X_k(\omega)}{n} \leq E[X_1] + 1/l \ \forall l \in \mathbb{N} \, \forall \omega \in \Omega_1$$

であることから分かる。下極限について考察するには

$$N_2(x) := \bigcap_{n=1}^{\infty} \bigcup_{k \geq n} \{(X_1 + \cdots + X_k)/k \leq x\}, \ \Omega_2 := \bigcap_{l \in \mathbb{N}} (\Omega \setminus N_2(E[X_1] - 1/l)).$$

を導入すればよい。P-a.s. 集合 Ω_2 上で $\liminf_{n\to\infty}(X_1 + \cdots + X_n)/n \geq E[X_1]$ が成り立つ。従って P-a.s. 集合 $\Omega_1 \cap \Omega_2$ 上で $(X_1 + \cdots + X_n)/n$ は $E[X_1]$ に収束する。 □

通常、大偏差原理の成立とは補題 15.2 ないしは系 15.4 を定理 15.11 にあるように拡張したものをいう。

15.8 補題. (i) 各 $L \in \mathbb{R}$ に対して $\{x \in \mathbb{R} : \mathrm{cml}^*(X_1, x) \geq L\}$ は $\emptyset$ または有界閉区間である。(ii) A を空でない $\mathbb{R}$ の閉集合とするとき関数 $\mathrm{cml}^*(X_1, \cdot)$ は A 上で最大値を持つ。

証明. (i) $L > 0$ なら $\{x \in \mathbb{R} : \mathrm{cml}^*(X_1, x) \geq L\} = \emptyset$ である。定理 14.13 より $L \leq 0$ なら

$$a := \inf\{x \in \mathbb{R} : \mathrm{cml}^*(X_1, x) \geq L\} > -\infty, \ b := \sup\{x \in \mathbb{R} : \mathrm{cml}^*(X_1, x) \geq L\} < -\infty$$

であって $\{x \in \mathbb{R} : \mathrm{cml}^*(X_1, x) \geq L\} = [a, b]$ である。

(ii) $\sup_{x\in A} \mathrm{cml}^*(X_1, x) = -\infty$ なら自明なので $\gamma := \sup_{x\in A} \mathrm{cml}^*(X_1, x) > -\infty$ と仮定する。空でない有界閉集合上の連続関数は最大値を持つという命題を適用する。まず $I := \{x \in \mathbb{R} : \mathrm{cml}^*(X_1, x) \geq \gamma - 1\}$ は $\mathrm{cnvx}(X_1)$ の閉包にふくまれ、$A \cap I \neq \emptyset$ でありかつ $\gamma = \sup_{x \in A\cap I} \mathrm{cml}^*(X_1, x)$ が成り立つ。定理 14.13 より $\mathrm{cnvx}(X_1)$ の閉包上で関数 $\mathrm{cml}^*(X_1, \cdot)$ は連続である。他方、(i) より $A \cap I$ は有界閉集合である。 □

15.9 補題. ある $A>0$ が存在して $0<a_n^i\le A\ \forall i=1,2\dots,k\ \forall n\in\mathbb{N}$ が満たされるなら

$$\limsup_{n\to\infty}\frac{1}{n}\log(a_n^1+\cdots+a_n^k)=\max_{i=1,2\dots,k}\limsup_{n\to\infty}\frac{1}{n}\log a_n^i$$

が成り立つ。

15.10 演習問題. 補題 15.9 を示せ。

一般には $n^{-1}\log P((X_1+\cdots+X_n)/n\in A)$ が収束するとは限らない。

15.11 定理. $\mathrm{cml}^*(X_1,\cdot)$ を**レート関数**とする**大偏差原理**が成立する。すなわち G を開集合、F を閉集合とするとき次が成り立つ。(ただし $\sup_{x\in\emptyset}\mathrm{cml}^*(X_1,x)=-\infty$ とよむ。)

$$\liminf_{n\to\infty}\frac{1}{n}\log P((X_1+\cdots+X_n)/n\in G)\ge\sup_{x\in G}\mathrm{cml}^*(X_1,x),$$
$$\limsup_{n\to\infty}\frac{1}{n}\log P((X_1+\cdots+X_n)/n\in F)\le\sup_{x\in F}\mathrm{cml}^*(X_1,x)$$

証明. $G=\emptyset$ なら自明なので $G\ne\emptyset$ と仮定する。$x\in G$ としよう。G は開集合なので $\delta>0$ を十分小さく選ぶと $I:=(x-\delta,x+\delta)\subset G$ が成り立つ。系 15.4(i) を適用して

$$\lim_{n\to\infty}\frac{1}{n}\log P((X_1+\cdots+X_n)/n\in I)=\sup_{y\in I}\mathrm{cml}^*(X_1,y)\ge\mathrm{cml}^*(X_1,x)$$

を得る。もちろん $P((X_1+\cdots+X_n)/n\in I)\le P((X_1+\cdots+X_n)/n\in G)$ であるから

$$\liminf_{n\to\infty}\frac{1}{n}\log P((X_1+\cdots+X_n)/n\in G)\ge\mathrm{cml}^*(X_1,x)\ \forall x\in G$$

が成り立つ。よって開集合 G に関する主張が得られた。

次に閉集合 F に関する考察に移る。($F=\emptyset$ もこめて)$\sup_{x\in F}\mathrm{cml}^*(X_1,x)=-\infty$ なら $F\cap\{x:\mathrm{cml}^*(X,x)>-\infty\}=\emptyset$ なので定理 14.13(iii) より $P((X_1+\cdots+X_n)/n\in F)=0$ となり不等式は自明である。また $\sup_{x\in F}\mathrm{cml}^*(X_1,x)=0$ の場合もそうである。従って

$F\ne\emptyset$ かつ $-\infty<\gamma:=\sup_{x\in F}\mathrm{cml}^*(X_1,x)<0$ と仮定する。

補題 15.8 によれば、$K:=\{x\in\mathbb{R}:\mathrm{cml}^*(X,x)\ge\gamma\}$ は内点を持つ有界閉区間であり、$\gamma=\mathrm{cml}^*(X_1,a)$ を満たすような $a\in F$ が存在する。 $a\in K$ であるが、定理 14.13(ii), (iv) より K の内部では $\mathrm{cml}^*(X,\cdot)>\gamma$ であるから a は K の端点である。そこで a は左端であるとして話を進める。K の右端を b とする。このとき開区間 (a,b) は上では $\mathrm{cml}^*(X,\cdot)>\gamma$ でありかつ $\mathrm{cml}^*(X,b)\ge\gamma$ である。従って次が成り立つ。

$$F\subset\{x\in\mathbb{R}:\mathrm{cml}^*(X,x)\le\gamma\}=\begin{cases}(-\infty,a]\cup[b,+\infty) & \mathrm{cml}^*(X,b)=\gamma\\(-\infty,a]\cup(b,+\infty) & \mathrm{cml}^*(X,b)>\gamma\end{cases}$$

ここで $\mathrm{cml}^*(X,b)>\gamma$ となり得るがそれは $b=\sup\mathrm{cnvx}(X_1)$ のときに限る。この場合を検討しよう。定理 14.13(iii) により $P((X_1+\cdots+X_n)/n>b)=0$ である。よって

$$P((X_1+\cdots+X_n)/n\in F)\le P((X_1+\cdots+X_n)/n\le a)$$

が成り立つ。補題 15.2 を適用する。$a < E[X_1]$ であるから次を得る。

$$\begin{aligned}&\limsup_{n\to\infty}\frac{1}{n}\log P((X_1+\cdots+X_n)/n \in F)\\ \le& \lim_{n\to\infty}\frac{1}{n}\log P((X_1+\cdots+X_n)/n \le a) = \mathrm{cml}^*(X_1,a)=\gamma\end{aligned}$$

$\mathrm{cml}^*(X,b)=\gamma$ のときを検討しよう。$E[X_1]<b$ であるから補題 15.2 により

$$\lim_{n\to\infty}\frac{1}{n}\log P((X_1+\cdots+X_n)/n \ge b) = \mathrm{cml}^*(X_1,b)=\gamma$$

を得る。ここで補題 15.9 を適用すると

$$\lim_{n\to\infty}\frac{1}{n}\log P((X_1+\cdots+X_n)/n \in (-\infty,a]\cup[b,+\infty)) = \gamma$$

が導かれる。よって閉集合 F に関する主張も得られた。 □

15.12 演習問題. 確率変数 X_1 の分布が Dirac 測度である場合も大偏差原理が成立することを直接検証せよ。

15.13 補題. 関数 $f:\mathbb{R}\to\mathbb{R}$ を $\mathrm{cnvx}(X_1)$ の閉包 I 上で上に有界なものとする。このとき次が成り立つとともにそれらの値は有限である。

$$\sup_{x\in\mathbb{R}}(f(x)+\mathrm{cml}^*(X_1,x)) = \sup\{L+\mathrm{cml}^*(X_1,x);(L,x)\in\mathbb{R}^2 : L<f(x)\}.$$

証明. まず $x \notin I$ なら $f(x)+\mathrm{cml}^*(X_1,x) = -\infty$ である。次に関数 f は I 上で上に有界、また関数 $\mathrm{cml}^*(X_1,\cdot)$ は非正値を取るので $f+\mathrm{cml}^*(X_1,\cdot)$ もまた上に有界である。他方 $\mathrm{cml}^*(X_1,E[X_1])=0$ であるから $a := \sup(f+\mathrm{cml}^*(X_1,\cdot))$ は有限な値である。さて $L<f(x)$ を満たす $(L,x)\in\mathbb{R}^2$ に対して

$$L+\mathrm{cml}^*(X_1,x) \le f(x)+\mathrm{cml}^*(X_1,x)\le a$$

が成り立つ（$\mathrm{cml}^*(X_1,x)=-\infty$ となる得るので左の不等号は = 付きである）。よって

$$\sup\{L+\mathrm{cml}^*(X_1,x);(L,x)\in\mathbb{R}^2 : L<f(x)\}\le a$$

さて $\varepsilon>0$ に対して $a-\varepsilon/2 < f(x)+\mathrm{cml}^*(X_1,x)$ を満たすような $x\in\mathbb{R}$ が存在する。そのような x に対して L を $f(x)-\varepsilon/2<L<f(x)$ を満たすように選ぶ。すると

$$a-\varepsilon < f(x)+\mathrm{cml}^*(X_1,x)-\varepsilon/2 < L+\mathrm{cml}^*(X_1,x)$$

が成り立つ。したがって $a\le \sup\{L+\mathrm{cml}^*(X_1,x);(L,x)\in\mathbb{R}^2 : L<f(x)\}$ である。 □

15.14 定理. 関数 $f:\mathbb{R}\to\mathbb{R}$ を $\mathrm{cnvx}(X_1)$ の閉包 I 上で上に有界かつ連続なものとする。このとき次が成り立つ。

$$\lim_{n\to\infty}\frac{1}{n}\log E[\exp\{nf((X_1+\cdots+X_n)/n)\}] = \sup_{x\in\mathbb{R}}(f(x)+\mathrm{cml}^*(X_1,x)).$$

証明. スペースを省くために $Y_n := (X_1 + \cdots + X_n)/n$ とおく。$L < M$ とする。

$$E[\exp\{nf(Y_n)\}; L \le f(Y_n) \le M] \le e^{nM} P(L \le f(Y_n) \le M) = e^{nM} P(Y_n \in f^{-1}([L, M]) \cap I)$$

ここで右の等号は $Y_n \in I$ P-a.s. に由来する。f の連続性より集合 $f^{-1}([L, M]) \cap I$ は閉集合である。定理 15.11 で述べた大偏差原理を適用する。

$$\limsup_{n\to\infty} \frac{1}{n} \log E[\exp\{nf(Y_n)\}; L \le f(Y_n) \le M] \le M + \sup_{x \in f^{-1}([L,M]) \cap I} \mathrm{cml}^*(X_1, x).$$

$x \in f^{-1}([L, M])$ なら $\mathrm{cml}^*(X_1, x) \le -L + f(x) + \mathrm{cml}^*(X_1, x)$ であるから

$$M + \sup_{x \in f^{-1}([L,M]) \cap I} \mathrm{cml}^*(X_1, x) \le M - L + \sup_{x \in f^{-1}([L,M]) \cap I} (f(x) + \mathrm{cml}^*(X_1, x))$$

を得る。よって $L < M$ である限り次が成り立つ。

$$\limsup_{n\to\infty} \frac{1}{n} \log E[\exp\{nf(Y_n)\}; L \le f(Y_n) \le M] \le M - L + \sup_{x \in \mathbb{R}} (f(x) + \mathrm{cml}^*(X_1, x)).$$

$C := \sup_{x \in I} f(x)$ とする。$\varepsilon > 0$ に対して次が満たされるように $k \in \mathbb{N}$ を十分大きく選ぶ。

$$C - k\varepsilon < \sup_{x \in \mathbb{R}} (f(x) + \mathrm{cml}^*(X_1, x)).$$

$A_i := f^{-1}([C - i\varepsilon, C - (i-1)\varepsilon])$ とおき各 $i = 1, \ldots, k$ に対して上の評価を適応する。

$$\limsup_{n\to\infty} \frac{1}{n} \log E[\exp\{nf(Y_n)\}; Y_n \in A_i] \le \varepsilon + \sup_{x \in \mathbb{R}} (f(x) + \mathrm{cml}^*(X_1, x)).$$

また $A_{k+1} := f^{-1}((-\infty, C - k\varepsilon])$ とおくと

$$\log E[\exp\{nf(Y_n)\}; Y_n \in A_{k+1}] \le n(C - k\varepsilon) \le n \sup_{x \in \mathbb{R}} (f(x) + \mathrm{cml}^*(X_1, x))$$

と評価できる。$I = \bigcup_{i=1}^{k+1} A_i$ かつ $Y_n \in I$ P-a.s. に注意して補題 15.9 を適用する。

$$\begin{aligned}
&\limsup_{n\to\infty} \frac{1}{n} \log E[\exp\{nf(Y_n)\}] \le \limsup_{n\to\infty} \frac{1}{n} \log \sum_{i=1}^{k+1} E[\exp\{nf(Y_n)\}; Y_n \in A_i] \\
&= \max_{i=1,\ldots,k+1} \limsup_{n\to\infty} \frac{1}{n} \log E[\exp\{nf(Y_n)\}; Y_n \in A_i] \le \varepsilon + \sup_{x \in \mathbb{R}} (f(x) + \mathrm{cml}^*(X_1, x)).
\end{aligned}$$

この段階で $\varepsilon > 0$ は任意である。下からの評価を行う。$L \in \mathbb{R}$ とする。f の連続性より集合 $f^{-1}((L, +\infty)) \cap I$ は I の開部分集合である。すなわち

$$f^{-1}((L, +\infty)) \cap I = A \cap I$$

を満たすような $\mathbb{R}$ の開集合 A が存在する。$Y_n \in I$ P-a.s. であるから

$$P(Y_n \in A) = P(Y_n \in A \cap I) = P(Y_n \in f^{-1}((L, +\infty)) \cap I) = P(f(Y_n) > L)$$

が成り立つ。従って補題 6.34 により次を得る。

$$P(Y_n \in A) = P(f(Y_n) > L) = P(\exp\{nf(Y_n)\} > e^{nL}) \le e^{-nL}E[\exp\{nf(Y_n)\}]$$

定理 15.11 で述べた大偏差原理を適用する。

$$\begin{aligned} L + \sup_{y\in A}\mathrm{cml}^*(X_1, y) &\le L + \liminf_{n\to\infty}\frac{1}{n}\log P(Y_n \in A) \\ &\le \liminf_{n\to\infty}\frac{1}{n}\log E[\exp\{nf(Y_n)\}]. \end{aligned}$$

$y \notin I$ なら $\mathrm{cml}^*(X_1, y) = -\infty$ であるから $f^{-1}((L, +\infty)) \cap I = A \cap I$ という関係より

$$\sup_{y\in f^{-1}((L,+\infty))}\mathrm{cml}^*(X_1, y) = \sup_{y\in A}\mathrm{cml}^*(X_1, y)$$

$L < f(x)$ を満たす $(L, x) \in \mathbb{R}^2$ に対して $\mathrm{cml}^*(X_1, x)$ は左辺で上からおさえられる。よって

$$L + \mathrm{cml}^*(X_1, x) \le \liminf_{n\to\infty}\frac{1}{n}\log E[\exp\{nf(Y_n)\}]$$

が $L < f(x)$ である限り成り立つ。ここで補題 15.13 を適用して

$$\sup_{x\in\mathbb{R}}(f(x) + \mathrm{cml}^*(X_1, x)) \le \liminf_{n\to\infty}\frac{1}{n}\log E[\exp\{nf(Y_n)\}]$$

を導くことができた。 □

16　無限次元確率変数とその分布

ここまでは同時に取り扱える確率変数はたかだか有限個であった。しかし応用上は不十分である。例えば第 1 節で紹介した $\omega \mapsto \tau(x, y, \omega)$ は無限系 $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ すべてが与えられて初めて決まる量である。そこで第 5 節、第 6 節、第 8 節での議論を拡張しておこう。即ち確率変数の無限系 $X_1, X_2, \ldots, X_n, \ldots$ が与えられたときそれらを束ねて

$$\Omega \to \mathbb{R}^{\mathbb{N}}, \omega \mapsto (X_1(\omega), X_2(\omega), \ldots, X_n(\omega), \ldots)$$

という写像を構成し、それを構造がわかりやすい写像と合成するというわけである。ここでも空間 $\mathbb{R}^{\mathbb{N}}$ の直積構造と構造の可算性が重要である。

他方、多数のランダムな要因を記述する数学的モデルが存在するのであろうか。それに答えるのが定理 16.18 であり、それによれば任意の $(\mathbb{R}^d, \mathrm{Borel}(\mathbb{R}^d))$ 上の確率測度 μ に対して独立な $\mathbb{R}^d$ 値確率変数の無限系 $X_1, X_2, \ldots, X_n, \ldots$ であって $\mathcal{L}(X_i, \cdot) = \mu\ \forall i \in \mathbb{N}$ を満たすものが存在する。これにより例えばランダムウォークというものが確固とした数学的土台にのることになる。

約束

$\mathbb{R}^{\mathbb{N}}$ の元 x に対してその第 i 座標を $x(i)$ と表記する。関数

$$\mathbb{R}^{\mathbb{N}} \times \mathbb{R}^{\mathbb{N}} \to \mathbb{R}, (x, y) \mapsto \sum_{i=1}^{\infty} \min\{|x(i) - y(i)|, 1\}/2^i$$

を dist と表記する。また $a \in \mathbb{R}^{\mathbb{N}}$, $\delta \in \mathbb{R}_{>0}$ に対して

$$\mathrm{Nbd}(a, \delta) := \{x \in \mathbb{R}^{\mathbb{N}} : \mathrm{dist}(x, a) < \delta\}.$$

16.1 演習問題. (i) 関数 dist は $\mathbb{R}^{\mathbb{N}}$ 上の距離であることを示せ。
(ii) $\mathbb{R}^{\mathbb{N}}$ の点列 a_n と $\mathbb{R}^{\mathbb{N}}$ の元 a について $\mathrm{dist}(a_n, a)$ が 0 に収束することと各 $i \in \mathbb{N}$ に対して数列 $a_n(i)$ が $a(i)$ に収束することは同値であることを示せ。
(iii) $\mathbb{R}^{\mathbb{N}}$ 上の距離 dist は完備であることを示せ。
(iv) $\{x \in \mathbb{R}^{\mathbb{N}} : x(i) \in \mathbb{Q}\, \forall i, \exists n \text{ s.t. } x(i) = 0\, \forall i \geq n\}$ は可算集合であることを示せ。
(v) $\mathbb{R}^{\mathbb{N}}$ 上の距離 dist は可分であることを示せ。

16.2 定義. 距離関数 dist によって誘導される $\mathbb{R}^{\mathbb{N}}$ 上の位相を標準位相という。また開集合すべてで生成される $\mathbb{R}^{\mathbb{N}}$ 上の σ-加法族を記号 $\mathrm{Borel}(\mathbb{R}^{\mathbb{N}})$ で表す。

以下、$\mathbb{R}^{\mathbb{N}}$ 上には標準位相および σ-加法族 $\mathrm{Borel}(\mathbb{R}^{\mathbb{N}})$ を導入しておく。

約束

$\mathbb{R}^{\mathbb{N}}$ の部分集合であって有限個の開区間と残りは $\mathbb{R}$ の直積で表現できるものを $\mathbb{R}^{\mathbb{N}}$ の開区間と呼ぶことにする。また $\mathbb{R}^{\mathbb{N}}$ の開区間の全体を $\mathcal{I}^{\mathbb{N}}$ で表記する。

16.3 補題. (i) $\forall a \in \mathbb{R}^{\mathbb{N}}, \forall \delta > 0\ \exists J \in \mathcal{I}^{\mathbb{N}}$ s.t. $a \in J \subset \mathrm{Nbd}(a, \delta)$.
(ii) $J \in \mathcal{I}^{\mathbb{N}}$ とする。$\forall a \in J\ \exists \delta > 0$ s.t. $\mathrm{Nbd}(a, \delta) \subset J$.

証明. (i) $\sum_{i=n+1}^{\infty} 1/2^i < \delta/2$ なる $n \in \mathbb{N}$ を選ぶとき次が成り立つ。

$$x \in \mathbb{R}^{\mathbb{N}},\ |x(i) - a(i)| < \delta/2\ \forall i \leq n \Rightarrow \mathrm{dist}(x, a) \leq \textstyle\sum_{i=1}^{n} \delta/2^{i+1} + \sum_{i=n+1}^{\infty} 1/2^i < \delta.$$

従って次の関係を得る。真ん中は $\mathbb{R}^{\mathbb{N}}$ の開区間である。

$$a \in (a(1) - \delta/2, a(1) + \delta/2) \times \cdots \times (a(n) - \delta/2, a(n) + \delta/2) \times \mathbb{R}^{\mathbb{N}} \subset \mathrm{Nbd}(a, \delta)$$

(ii) a を含む $\mathbb{R}^{\mathbb{N}}$ の開区間 J に対し有限個の開区間 $I_1, \ldots, I_n$ を $J = I_1 \times \cdots \times I_n \times \mathbb{R}^{\mathbb{N}}$ となるように選ぶ。$a \in J$ より $a(i) \in I_i\ \forall i \leq n$ が従う。各 I_i は開区間なので $\delta(i) > 0$ が存在して $(a(i) - \delta(i), a(i) + \delta(i)) \subset I_i$ が成り立つ。このとき $\delta := \min\{\delta(i)/2^i ; i \leq n\} > 0$ であり

$$x \in \mathbb{R}^{\mathbb{N}},\ \mathrm{dist}(x, a) < \delta \Rightarrow |x(i) - a(i)|/2^i < \delta \leq \delta(i)/2^i\ \forall i \leq n \Rightarrow a(i) \in I_i\ \forall i \leq n$$

という関係が成り立つ。即ち $\mathrm{Nbd}(a, \delta) \subset J$ である。 □

16.4 系. $A \subset \mathbb{R}^{\mathbb{N}}$ に対して A 開集合 $\Leftrightarrow \forall a \in A\ \exists J \in \mathcal{I}^{\mathbb{N}}$ s.t. $a \in J \subset A$ である。

16.5 補題. $\mathbb{R}^{\mathbb{N}}$ の開区間すべてで生成される $\mathbb{R}^{\mathbb{N}}$ 上の σ-加法族は $\mathrm{Borel}(\mathbb{R}^{\mathbb{N}})$ である。

証明. 証明の方針は補題 5.4 におけるものと同じである。即ち以下の 2 点を確かめればよい。両端が有理数であるような開区間を有限個と残りは $\mathbb{R}$ の直積で表現できる $\mathbb{R}^{\mathbb{N}}$ の部分集合全体 $\mathcal{C}$ は可算族である。また $\mathbb{R}^{\mathbb{N}}$ の開集合 A に対して $\mathcal{C}$ に属する集合で A の部分集合となるもののすべての合併は A である。後者は系 16.4 を利用して示すことができる。 □

記号

S 上の σ 加法族たち $\mathcal{B}_\alpha$ に対して S 上の σ 加法族 $\sigma(\bigcup_\alpha \mathcal{B}_\alpha)$ を $\bigvee_\alpha \mathcal{B}_\alpha$ と表記する。

以下は定理 5.9 から系 2.17 を経由して系 5.15 に至るまでと同じ筋道である。

16.6 定理. 写像 $f : S \to \mathbb{R}^{\mathbb{N}}$ に対しその第 i 成分を g_i とすると $\sigma\{f\} = \bigvee_{i=1}^{\infty} \sigma\{g_i\}$ である。

16.7 系. 写像 $f : S \to \mathbb{R}^{\mathbb{N}}$ と S 上の σ 加法族 $\mathcal{B}$ に対し、f が $\mathcal{B}$ 可測であることとその各成分が $\mathcal{B}$ 可測であることは同値である。

16.8 演習問題. 定理 16.6 と系 16.7 を示せ。

確率変数およびその分布の定義は以前と形式上同じであるが念のため明示しておく。

16.9 定義. 確率空間 $(\Omega, \mathcal{F}, P)$ 上の $\mathbb{R}^{\mathbb{N}}$ 値確率変数あるいは**無限次元確率変数**とは $\mathcal{F}$ 可測写像 $X : \Omega \to \mathbb{R}^{\mathbb{N}}$ をいう。$\mathbb{R}^{\mathbb{N}}$ 値確率変数 X による P の**像測度** $\mathcal{L}(X, \cdot)$ を X の**分布**という。その成分 $X_1, X_2, \ldots, X_n, \ldots$ からみたときは**結合分布**と呼ばれる。逆に $(\mathbb{R}^{\mathbb{N}}, \mathrm{Borel}(\mathbb{R}^{\mathbb{N}}))$ 上の確率測度 μ が与えられたとき、$\mathcal{L}(X, \cdot) = \mu$ を満たす確率変数 X は**分布** μ **に従う**という。

16.10 注意. 系 16.7 によれば写像 $X : \Omega \to \mathbb{R}^{\mathbb{N}}$ について確率空間 $(\Omega, \mathcal{F}, P)$ 上の $\mathbb{R}^{\mathbb{N}}$ 値確率変数であることと各成分が実確率変数であることは同値である。またこのとき非負値 $\mathrm{Borel}(\mathbb{R}^{\mathbb{N}})$ 可測関数 $f : \mathbb{R}^{\mathbb{N}} \to \overline{\mathbb{R}}$ に対して $E[f(X)] = \displaystyle\int_{\mathbb{R}^{\mathbb{N}}} f \, \mathcal{L}(X, \cdot)$ が成り立つ。

さて直積測度の概念をパーツとなる測度が有限個でない場合にも拡張しておこう。

記号

各 $i \in \mathbb{N}$ に対して集合 S_i とその部分集合の族 $\mathcal{A}_i$ が与えられたとして

$$\prod_{\mathbb{N}} S_{\cdot} := S_i \text{ たちの直積集合}, \ \mathrm{proj}_i : \prod_{\mathbb{N}} S_{\cdot} \to S_i \text{ 第 } i \text{ 座標への射影}.$$

$S_i = S \ \forall i \in \mathbb{N}$ が満たされる特殊な場合は $\prod_{\mathbb{N}} S_{\cdot}$ を $S^{\mathbb{N}}$ と表記する。

$$\mathrm{Cyl}_{\mathbb{N}} \mathcal{A}_{\cdot} := \{\textstyle\bigcap_{i \in I} \mathrm{proj}_i^{-1}(A_i) \, ; I \in \{\ \mathbb{N} \text{ の有限部分集合}\}, A_i \in \mathcal{A}_i \text{ for } i \in I\}.$$

16.11 定義. $\mathrm{Cyl}_{\mathbb{N}} \mathcal{A}_{\cdot}$ の元を $\prod_{\mathbb{N}} S_{\cdot}$ の $\mathcal{A}_{\cdot}$-*cylinder* **集合**という。

16.12 例. $\mathcal{I}$ を開区間全体、$\mathcal{I}^{\mathbb{N}}$ を $\mathbb{R}^{\mathbb{N}}$ の開区間全体とするとき $\mathrm{Cyl}_{\mathbb{N}} \mathcal{I} = \mathcal{I}^{\mathbb{N}}$ である。

16.13 補題. $\prod_{\mathbb{N}} S_{\cdot}$ 上の σ 加法族として $\sigma(\mathrm{Cyl}_{\mathbb{N}} \mathcal{A}_{\cdot}) = \sigma(\mathrm{Cyl}_{\mathbb{N}} \sigma(\mathcal{A}_{\cdot}))$ である。

証明. まず、各 $i \in \mathbb{N}$ に対して $\mathcal{A}_i \subset \sigma(\mathcal{A}_i)$ なので $\mathrm{Cyl}_{\mathbb{N}}\mathcal{A}. \subset \mathrm{Cyl}_{\mathbb{N}}\sigma(\mathcal{A}.)$ であり、従って $\sigma(\mathrm{Cyl}_{\mathbb{N}}\mathcal{A}.) \subset \sigma(\mathrm{Cyl}_{\mathbb{N}}\sigma(\mathcal{A}.))$ である。他方、補題 5.6(iii) によれば、各 $i \in \mathbb{N}$ に対して $\mathrm{proj}_i^*\sigma(\mathcal{A}_i) = \sigma(\mathrm{proj}_i^*\mathcal{A}_i)$ となるが、右辺は $\sigma(\mathrm{Cyl}_{\mathbb{N}}\mathcal{A}.)$ に包含される。従って I を $\mathbb{N}$ の有限部分集合、$A_i \in \sigma(\mathcal{A}_i)$ for $i \in I$ とするとき

$$\bigcap\nolimits_{i\in I} \mathrm{proj}_i^{-1}(A_i) \in \sigma(\mathrm{Cyl}_{\mathbb{N}}\mathcal{A}.)$$

が成り立つ。すなわち $\mathrm{Cyl}_{\mathbb{N}}\sigma(\mathcal{A}.) \subset \sigma(\mathrm{Cyl}_{\mathbb{N}}\mathcal{A}.)$ となる。右辺は σ 加法族なので包含関係 $\sigma(\mathrm{Cyl}_{\mathbb{N}}\sigma(\mathcal{A}.)) \subset \sigma(\mathrm{Cyl}_{\mathbb{N}}\mathcal{A}.)$ も導かれた。 □

前提

各 $i \in \mathbb{N}$ に対し S_i 上の σ-加法族 $\mathcal{B}_i$ と $(S_i, \mathcal{B}_i)$ 上の 確率測度 μ_i が与えられている。

16.14 定義. $\prod_{\mathbb{N}} S.$ 上の σ-加法族 $\sigma(\mathrm{Cyl}_{\mathbb{N}}\mathcal{B}.)$ を**直積** σ-**加法族**と呼び $\bigotimes_{\mathbb{N}} \mathcal{B}.$ と表す。あるいは $\bigotimes_{\mathbb{N}} \mathcal{B}. = \bigvee_{i\in\mathbb{N}} \mathrm{proj}_i^*\mathcal{B}_i$ と定義しても同じである。

16.15 例. $\bigotimes_{\mathbb{N}} \mathrm{Borel}(\mathbb{R}) = \mathrm{Borel}(\mathbb{R}^{\mathbb{N}})$.

証明. $\mathrm{Borel}(\mathbb{R}) = \sigma(\mathcal{I})$ であり、$\mathcal{I}^{\mathbb{N}} = \mathrm{Cyl}_{\mathbb{N}}\mathcal{I}$ である。補題 16.5 と補題 16.13 を適用して $\mathrm{Borel}(\mathbb{R}^{\mathbb{N}}) = \sigma(\mathrm{Cyl}_{\mathbb{N}}\mathcal{I}) = \sigma(\mathrm{Cyl}_{\mathbb{N}}\sigma(\mathcal{I})) = \bigotimes_{\mathbb{N}} \mathrm{Borel}(\mathbb{R})$ を得る。 □

16.16 補題. μ, ν を $(\prod_{\mathbb{N}} S., \bigotimes_{\mathbb{N}} \mathcal{B}.)$ 上の 確率測度 とするとき次が成り立つ

$$\mu = \nu \Leftrightarrow \mu(A) = \nu(A)\ \forall A \in \mathrm{Cyl}_{\mathbb{N}}\mathcal{B}.$$

証明. 集合族 $\mathrm{Cyl}_{\mathbb{N}}\mathcal{B}.$ は $\prod_{\mathbb{N}} S.$ を含む π システムであるから定理 7.12(i) が適用。 □

16.17 定義. $(\prod_{\mathbb{N}} S., \bigotimes_{\mathbb{N}} \mathcal{B}.)$ 上の測度 μ で次を満たすものを**直積測度**と呼ぶ。

$$I \in \{\ \mathbb{N}\text{の有限部分集合}\ \}, A_i \in \mathcal{A}_i \text{ for } i \in I \Rightarrow \mu(\bigcap_{i\in I} \mathrm{proj}_i^{-1}(A_i)) = \prod_{i\in I} \mu_i(A_i)$$

16.18 定理. 直積測度 $\bigotimes_{\mathbb{N}} \mu.$ が存在しかつ一意である。

証明. 一意性は補題 16.16 で述べられている。存在は系 17.9 と系 17.12 から導かれる。 □

16.19 定義. 実確率変数の無限系 $X_1, X_2, \ldots, X_n, \ldots$ が**独立**であるとはそれらの結合分布が各分布の直積に等しいことをいう。すなわち次が成り立つことである。

$$\mathcal{L}((X_1, X_2, \ldots, X_n, \ldots), \cdot) = \bigotimes\nolimits_{\mathbb{N}} \mathcal{L}(X., \cdot).$$

定理 16.18 のうちで存在に関わる部分の応用例として次をあげておく。

16.20 系. 各 $i \in \mathbb{N}$ に対し $(\mathbb{R}, \mathrm{Borel}(\mathbb{R}))$ 上の確率測度 μ_i が与えられたとき独立な実確率変数の無限系 $X_1, X_2, \ldots, X_n, \ldots$ であって $\mathcal{L}(X_i, \cdot) = \mu_i\ \forall i \in \mathbb{N}$ を満たすものが存在する。

証明. 確率空間 $(\mathbb{R}^{\mathbb{N}}, \mathrm{Borel}(\mathbb{R}^{\mathbb{N}}), \bigotimes_{\mathbb{N}} \mu.)$ 上の関数列 $\mathrm{proj}_i\ i \in \mathbb{N}$ が求めるものである。 □

16.21 例. 独立な実確率変数の無限系 $X_1, X_2, \ldots, X_n, \ldots$ であってその各々が標準正規分布に従うものが存在する。

次は直積測度の特徴付けからわかるが、その確認の意味も込めて演習問題とする。

16.22 補題. 実確率変数の無限系 $X_1, X_2, \ldots, X_n, \ldots$ が独立であるのはその任意の有限部分系が独立であるのと同値である。

16.23 演習問題. 補題 16.22 を示せ。

16.24 注意. 補題 16.22 によれば、無限系の独立性を議論する場合でも、結局は有限部分系がしかるべき条件を満たすかチェックすることに帰着するので、測度の無限直積に表面上はタッチしなくてすむのである。このような事情があるので、多くの場合「任意の有限部分系が独立であること」をもって無限系の独立性の定義としている。

16.25 例. 第 1 節で述べた関数の列 $\xi_k : (0,1] \to \mathbb{R}$, $k \in \mathbb{N}$ で与えられる Lebesgue モデル上の実確率変数系は例 8.7 でも述べたように無限系として独立である。

17 無限直積測度の構成

この節は定理 16.18 の証明にあてられる。記号など第 16 節のものを引き継ぐ。いつものように有限加法的測度を構成しその σ 加法性をチェックするという手順を踏む。ここで念のため有限加法的測度とその測度への拡張について復習しておこう。

定義の確認

S の部分集合族 $\mathcal{C}$ と関数 $m : \mathcal{C} \to \overline{\mathbb{R}}$ が次の条件

(i) $\emptyset \in \mathcal{C}$. $A \in \mathcal{C}, B \in \mathcal{C} \Rightarrow A \cap B \in \mathcal{C}$.
$A \in \mathcal{C}, B \in \mathcal{C}, B \subset A, A \neq B \Rightarrow A \setminus B$ の有限な $\mathcal{C}$-分割が存在する
(ii) $m(A) \geq 0 \ \forall A \in \mathcal{C}$, $m(\emptyset) = 0$
(iii) $A \in \mathcal{C}$ とその有限な $\mathcal{C}$-分割 Δ に対して $m(A) = \sum_{J \in \Delta} m(J)$.

を満たすとき、**有限加法的測度**(finitely additive measure) という。

定義の確認の続き

任意の可算無限な $\mathcal{C}$-分割についても条件 (iii) が成り立つとき m は σ-**加法的**(σ-additive) であるという。また S の可算 $\mathcal{C}$-被覆 C_n $n \in \mathbb{N}$ で $m(C_n) < +\infty \ \forall n \in \mathbb{N}$ をみたすものが存在するとき m は σ-**有限**(σ-finite) であるという。

17.1 例. $\mathbb{R}^d$ の部分集合であって左半開区間の直積で表現できるものすべてと $\emptyset$ からなる集合族 $\mathcal{I}$ は上の性質 (i) を満たす。また $m((a_1, b_1] \times \cdots \times (a_d, b_d]) = \prod_{i=1}^{d} (b_i - a_i) \ \forall a, \forall b \in \mathbb{R}^d$, $a_i < b_i$ で定義される $\mathcal{I}$ 上の関数 m は σ-加法的かつ σ-有限な有限加法的測度である。

復習

Hopf **の拡張定理**: 有限加法的測度が測度に拡張されるための必要十分条件はそれが σ-加法的なことである。また σ-有限なときは拡張は一意的である。

σ-有限性のもとでの一意性は定理 7.12(ii) を適用して導くことができる。
有限測度の場合 *Hopf* **の拡張定理**の前提となる条件を次の形で判定することが多い。

17.2 定義. S の部分集合の族 $\mathcal{C}$ が**有限加法族**であるとは

$$\emptyset \in \mathcal{C};\ A \in \mathcal{C} \Rightarrow S \setminus A \in \mathcal{C};\ A, B \in \mathcal{C} \Rightarrow A \cup B \in \mathcal{C}.$$

17.3 補題. $(\mathcal{C}, m)$ を S 上の有限加法的測度で定義域 $\mathcal{C}$ が有限加法族でありかつ $m(S) < +\infty$ をみたすものとする。このとき m の σ 加法性は次と同値である。

$$A_n \in \mathcal{C}\ \forall n \in \mathbb{N},\ A_n \supset A_{n+1}\ \forall n \in \mathbb{N},\ \inf_{n\in\mathbb{N}} m(A_n) > 0 \Rightarrow \textstyle\bigcap_{n=1}^{\infty} A_n \neq \emptyset.$$

証明. 与えられた条件から σ 加法性を導こう。そこで

$$B_n \in \mathcal{C}\ \forall n \in \mathbb{N},\ B_n \cap B_k = \emptyset\ n \neq k,\ A := \textstyle\bigcup_{n=1}^{\infty} B_n \in \mathcal{C}$$

であるとする。$\mathcal{C}$ は有限加法族であるから

$$A_n := A \setminus \bigcup_{k=1}^{n} B_k \in \mathcal{C}$$

またその定義により $A_n \supset A_{n+1}\ \forall n \in \mathbb{N}$ かつ $\bigcap_{n=1}^{\infty} A_n = A \setminus \bigcup_{k=1}^{\infty} B_k = \emptyset$ である。従って

$$\inf_{n\in\mathbb{N}} m(A_n) = 0$$

が与えられた条件から得られる。他方、有限加法性と $m(S) < +\infty$ により

$$\sum_{k=1}^{n} m(B_k) = m(A) - m(A_n)\ \forall n \in \mathbb{N}$$

である。よって $\sup_{n\in\mathbb{N}} \sum_{k=1}^{n} m(B_k) = m(A)$ が導けた。 □

17.4 演習問題. 補題 17.3 で与えられた条件が必要であることを示せ。

記号

$k, n \in \mathbb{N},\ n < k$ とする。

$\mathrm{proj}_{\leq k} : \prod_{\mathbb{N}} S. \to \prod_{\leq k} S.$ 第 k 座標までへの射影,
$\mathrm{proj}_{\leq n, \leq k} : \prod_{\leq k} S. \to \prod_{\leq n} S.$ 第 n 座標までへの射影,

17.5 補題. $\mathrm{proj}_{\leq k} = \mathrm{proj}_{\leq k, \leq k+1} \circ \mathrm{proj}_{\leq k+1},\ \mathrm{proj}_{\leq k+1}(\mathrm{proj}_{\leq k+1}^{-1}(A)) = A\ \forall A \subset \prod_{\leq k+1} S..$

記号

$k, n \in \mathbb{N}$, $n < k$ とする。

$\bigotimes_{(n,k]} \mathcal{B}_\cdot := \bigvee_{i \in \mathbb{N}: n < i \le k} \mathrm{proj}_i^* \mathcal{B}_i$ $\prod_{(n,k]} S_\cdot$ 上の σ 加法族,
$\bigotimes_{(n,k]} \mu_\cdot := \mu_{n+1}, \ldots, \mu_k$ の直積測度.

17.6 注意. $\bigotimes_{(n,k]} \mu_\cdot$ は次を満たす唯一の $\bigotimes_{(n,k]} \mathcal{B}_\cdot$ 上の測度 ν である。

$$\nu(\textstyle\bigcap_{i \in \mathbb{N}: n < i \le k} \mathrm{proj}_i^{-1}(A_i)) = \prod_{i \in \mathbb{N}: n < i \le k} \mu_i(A_i) \ \forall A_i \in \mathcal{B}_i.$$

17.7 補題. (i) $k, n \in \mathbb{N}$, $n < k$ とする。$\prod_{\le k} S_\cdot$ 上で次の関係が成り立つ。

$$\textstyle\bigotimes_{\le n} \mathcal{B}_\cdot \otimes \bigotimes_{(n,k]} \mathcal{B}_\cdot = \bigotimes_{\le k} \mathcal{B}_\cdot, \ \ \bigotimes_{\le n} \mu_\cdot \otimes \bigotimes_{(n,k]} \mu_\cdot = \bigotimes_{\le k} \mu_\cdot.$$

(ii) $n \in \mathbb{N}$ とし、ν を $\bigotimes_{\le n} \mathcal{B}_\cdot$ 上の確率測度とするとき次が成り立つ。

$$(\mathrm{proj}_{\le k, \le k+1})_*(\nu \otimes \textstyle\bigotimes_{(n,k+1]} \mu_\cdot) = \nu \otimes \bigotimes_{(n,k]} \mu_\cdot \ \ \forall k \in \mathbb{N}_{\ge n}.$$

ここで $k = n$ のときは右辺は ν を表すものとする。

17.8 演習問題. 補題 17.5 と補題 17.7 を示せ。

17.9 系. $\bigotimes_{\mathbb{N}}^\circ \mathcal{B}_\cdot := \bigcup_{n=1}^\infty \mathrm{proj}_{\le n}^* \bigotimes_{\le n} \mathcal{B}_\cdot$ は $\prod_{\mathbb{N}} S_\cdot$ 上の有限加法族である。またその上に

$$m(\mathrm{proj}_{\le n}^{-1}(A)) = (\textstyle\bigotimes_{\le n} \mu_\cdot)(A) \ \ \forall n \in \mathbb{N} \ \forall A \in \bigotimes_{\le n} \mathcal{B}_\cdot.$$

を満たす有限加法的測度 m が唯一存在する。これを $\bigotimes_{\mathbb{N}}^\circ \mu_\cdot$ と表記する。

σ 加法性の判定がなかなか厄介である。

17.10 補題. $\alpha \in \mathbb{R}$, $n \in \mathbb{N}$ とし、ν を $\bigotimes_{\le n} \mathcal{B}_\cdot$ 上の確率測度とする。

$$\mathrm{proj}_{\le k}^{-1}(A_k) \supset \mathrm{proj}_{\le k+1}^{-1}(A_{k+1}) \ \forall k \in \mathbb{N}_{\ge n}, \ \inf_{k \in \mathbb{N}: k \ge n} (\nu \otimes \textstyle\bigotimes_{(n,k]} \mu_\cdot)(A_k) > \alpha$$

を満たすように各 $k \in \mathbb{N}_{\ge n}$ に対して $A_k \in \bigotimes_{\le k} \mathcal{B}_\cdot$ が与えられているなら

$$\nu(\{x \in \textstyle\prod_{\le n} S_\cdot : (\delta_x \otimes \bigotimes_{(n,k]} \mu_\cdot)(A_k) \ge \alpha \ \forall k \in \mathbb{N}_{\ge n}\}) > 0.$$

証明. 軽装化のため $\mu_{n,k} := \bigotimes_{(n,k]} \mu_\cdot$ とかき、$\{x \in \prod_{\le n} S_\cdot : (\delta_x \otimes \mu_{n,k})(A_k) \ge \alpha\}$ を B_k とおく。定理 8.10(i) によれば $B_k \in \bigotimes_{\le n} \mathcal{B}_\cdot$ である。一方

$$1_{A_k}(x, y) \le 1_{B_k}(x) + 1_{B_k^c}(x) 1_{A_k}(x, y) \ \forall x \in \prod_{\le n} S_\cdot, \forall y \in \prod_{(n,k]} S_\cdot$$

が満たされる。したがって直積測度 $\nu \otimes \mu_{n,k}$ についての積分に Fubini の定理を適用して

$$(\nu \otimes \mu_{n,k})(A_k) \le \nu(B_k) + \int_{B_k^c} (\delta_x \otimes \mu_{n,k})(A_k) \, \nu(dx) \le \nu(B_k) + \alpha \nu(B_k^c).$$

右辺は $(1-\alpha)\nu(B_k)+\alpha$ に等しい。従って仮定 $\inf_{k\in\mathbb{N}:k\geq n}(\nu\otimes\mu_{n,k})(A_k)>\alpha$ から

$$\alpha<\inf_{k\in\mathbb{N}:k\geq n}(1-\alpha)\nu(B_k)+\alpha$$

および $\alpha<1$ であることが導かれる。よって

$$\inf_{k\in\mathbb{N}:k\geq n}\nu(B_k)>0.$$

つぎに仮定 $\mathrm{proj}_{\leq k}^{-1}(A_k)\supset\mathrm{proj}_{\leq k+1}^{-1}(A_{k+1})\ \forall k\in\mathbb{N}_{\geq n}$ を使う。補題 17.5 を適用して

$$\begin{aligned}\mathrm{proj}_{\leq k,\leq k+1}^{-1}(A_k)&=\mathrm{proj}_{\leq k+1}(\mathrm{proj}_{\leq k+1}^{-1}(\mathrm{proj}_{\leq k,\leq k+1}^{-1}(A_k)))\\&=\mathrm{proj}_{\leq k+1}(\mathrm{proj}_{\leq k}^{-1}(A_k))\\&\supset\mathrm{proj}_{\leq k+1}(\mathrm{proj}_{\leq k+1}^{-1}(A_{k+1}))=A_{k+1}.\end{aligned}$$

従って補題 17.7(ii) を適用して次を得る。

$$\begin{aligned}(\delta_x\otimes\mu_{n,k})(A_k)&=(\delta_x\otimes\mu_{n,k+1})(\mathrm{proj}_{\leq k,\leq k+1}^{-1}(A_k))\\&\geq(\delta_x\otimes\mu_{n,k+1})(A_{k+1})\ \forall x\in\prod\nolimits_{\leq n}S_\cdot.\end{aligned}$$

これは包含関係 $B_k\supset B_{k+1}\ \forall k\in\mathbb{N}_{\geq n}$ を意味する。ν は確率測度なので単調収束定理により

$$\nu(\bigcap_{k\in\mathbb{N}:k\geq n}B_k)=\inf_{k\in\mathbb{N}:k\geq n}\nu(B_k).$$

既に示されたように右辺は正である。 □

17.11 補題. $\delta>0$ とする。各 $k\in\mathbb{N}$ に対して $A_k\in\bigotimes_{\leq k}\mathcal{B}_\cdot$ が与えられ

$$\mathrm{proj}_{\leq k}^{-1}(A_k)\supset\mathrm{proj}_{\leq k+1}^{-1}(A_{k+1})\ \forall k\in\mathbb{N},\ \inf_{k\in\mathbb{N}}(\bigotimes\nolimits_{\leq k}\mu_\cdot)(A_k)>\delta$$

が成り立つなら

$$\exists x\in\prod\nolimits_{\mathbb{N}}S_\cdot\ \text{s.t.}\ (\delta_{\mathrm{proj}_{\leq n}(x)}\otimes\bigotimes\nolimits_{(n,k]}\mu_\cdot)(A_k)\geq\delta/n\ \forall n\in\mathbb{N},\forall k\in\mathbb{N}_{\geq n}.$$

証明. 軽装化のため $\mu_{n,k}:=\bigotimes_{(n,k]}\mu_\cdot$ とおく。各 $m\in\mathbb{N}$ に対して

$$(\star)\qquad\exists x\in\prod\nolimits_{\leq m}S_\cdot\ \text{s.t.}\ (\delta_{\mathrm{proj}_{\leq n}(x)}\otimes\mu_{n,k})(A_k)\geq\delta/n\ \forall n\in\mathbb{N}_{\leq m}\forall k\in\mathbb{N}_{\geq n}$$

を数学的帰納法を適用して示す。仮定から $\inf_{k\in\mathbb{N}}(\mu_1\otimes\mu_{1,k})(A_k)>\delta$ である。補題 17.10 より $\{x\in S_1:(\delta_x\otimes\mu_{1,k})(A_k)\geq\delta\forall k\in\mathbb{N}\}$ の μ_1 測度は正であるからそれは空集合でない。従って

$$\exists x(1)\in S_1\ \text{s.t.}\ (\delta_{x(1)}\otimes\mu_{1,k})(A_k)\geq\delta\ \forall k\in\mathbb{N}$$

が得られ $m=1$ の場合 $(\star)$ が示せた。ここで $\delta/m>\delta/(m+1)$ であるから帰納法のステップを次へ進めるには $x\in\prod_{\leq m}S_\cdot$ に対して

$$\inf_{k\in\mathbb{N}:k\geq m+1}((\delta_x\otimes\mu_{m+1})\otimes\mu_{m+1,k})(A_k)>\delta/(m+1)$$

という設定のもと補題 17.10 を適用すればよい。結論を得るためには $(\star)$ だけでは不十分で更に Zorn の補題を使う必要がある。即ち $x\in\prod_{\leq m}S_\cdot$ と $x'\in\prod_{\leq m'}S_\cdot$ に対して $x\prec x'\Leftrightarrow m\leq m',\ x=\mathrm{proj}_{\leq m,\leq m'}(x')$ という順序を導入して極大元を選び出せばよい。 □

17.12 系. 系 17.9 で述べた有限加法的測度 $\bigotimes_{\mathbb{N}}^{\circ} \mu.$ は σ 加法的である。

証明. 補題 17.3 で述べた判定法を適用する。そこで次を仮定する。

$$B_k \in \bigotimes_{\mathbb{N}}^{\circ} \mathcal{B}.\ \forall k \in \mathbb{N},\ B_k \supset B_{k+1}\ \forall k \in \mathbb{N},\ \inf_{k\in\mathbb{N}} \bigotimes_{\mathbb{N}}^{\circ} \mu.(B_k) > 0.$$

各 $k \in \mathbb{N}$ に対して $B_k = \mathrm{proj}_{\leq k}^{-1}(A_k)$ を満たす $A_k \in \bigotimes_{\leq k} \mathcal{B}.$ が存在するとしても一般性を失わない（ことを確認するのは読者に委ねる）。このとき次が成り立つ。

$$\mathrm{proj}_{\leq k}^{-1}(A_k) \supset \mathrm{proj}_{\leq k+1}^{-1}(A_{k+1})\ \forall k \in \mathbb{N},\ \inf_{k\in\mathbb{N}}(\bigotimes_{\leq k} \mu.)(A_k) > 0.$$

$\delta = \inf_{k\in\mathbb{N}}(\bigotimes_{\leq k} \mu.)(A_k)/2 > 0$ に対し $x \in \prod_{\mathbb{N}} S.$ を補題 17.11 で述べられたものとする。ここで各 $n \in \mathbb{N}$ について $k = n$ と選択することにより $\delta_{\mathrm{proj}_{\leq n}(x)}(A_n) \geq \delta/n$ を得る。従って $\mathrm{proj}_{\leq n}(x) \in A_n\ \forall n \in \mathbb{N}$ である。すなわち次が成り立つ。

$$x \in \bigcap_{k=1}^{\infty} \mathrm{proj}_{\leq k}^{-1}(A_k) = \bigcap_{k=1}^{\infty} B_k.$$

以上より $\bigcap_{k=1}^{\infty} B_k \neq \emptyset$ が導けた。故に $\bigotimes_{\mathbb{N}}^{\circ} \mu.$ は σ 加法的である。 □

18　独立性の σ 加法族による定式化

ここでは独立性を視点を変えて考察してみたい。さて、第 2 節で事象は確率変数を介して観測されるという立場を宣言したので、第 8 節でも確率変数に対して独立性を定義したわけだが、定理 8.4 をみると直接関わっているのは事象の確率そのものであることがわかる。よって事象に対して独立性を定義するのがより深く切り込むのに向いていると考えられる。

記号

$$\mathcal{F}^{\otimes n} := \overbrace{\mathcal{F} \otimes \cdots \otimes \mathcal{F}}^{n},\ P^{\otimes n} := \overbrace{P \otimes \cdots \otimes P}^{n},\ \mathrm{diag}_n(\omega) := (\overbrace{\omega, \ldots, \omega}^{n}).$$

18.1 補題. 各 $n \in \mathbb{N}$ と $A \in \mathcal{F}^{\otimes n}$ に対して $\{\omega \in \Omega : \mathrm{diag}_n(\omega) \in A\} \in \mathcal{F}$ である。

証明. 示すのは写像 $\mathrm{diag}_n : \Omega \to \Omega^n$ の対 $\mathcal{F}$, $\mathcal{F}^{\otimes n}$ についての可測性である。

$$A_i \in \mathcal{F}\ \forall i = 1, 2, \ldots, n \Rightarrow \mathrm{diag}_n^{-1}(A_1 \times \cdots \times A_n) = \textstyle\bigcap_{i=1}^{n} A_i \in \mathcal{F}$$

が成り立つので補題 5.6(iv) により、$\mathrm{diag}_n^{-1}(A) \in \mathcal{F}\ \forall A \in \mathcal{F}^{\otimes n}$ が従う。 □

以上の準備のもと独立性を直積測度と結びつけて導入する。

18.2 定義. Ω 上の σ 加法族 $\mathcal{G}$ で $\mathcal{G} \subset \mathcal{F}$ を満たすものを**部分 σ 加法族**(sub σ-field) という。部分 σ 加法族の系 $\mathcal{G}_1, \mathcal{G}_2, \ldots, \mathcal{G}_n$ が**独立**であるとは次が成り立つことをいう。

$$P(\{\omega \in \Omega : \mathrm{diag}_n(\omega) \in A\}) = P^{\otimes n}(A)\ \forall A \in \bigotimes_{\leq n} \mathcal{G}..$$

無限系の独立性についてもその形式的定義は同じである。すなわち部分 σ 加法族の系 $\mathcal{G}_1, \mathcal{G}_2, \ldots, \mathcal{G}_n, \ldots$ が独立であるとは次が成り立つことをいう。

$$P(\{\omega \in \Omega : \mathrm{diag}_{\mathbb{N}}(\omega) \in A\}) = P^{\otimes \mathbb{N}}(A)\ \forall A \in \bigotimes_{\mathbb{N}} \mathcal{G}..$$

次の定理は無限系で定式化されているが、有限系の独立性についても同様の命題が成り立つことはいうまでもない。

18.3 定理. $\mathcal{C}_1, \ldots, \mathcal{C}_n, \ldots$ を $\mathcal{F}$ 集合のなす π システムとする。このとき部分 σ 加法族の系 $\sigma(\mathcal{C}_1), \ldots, \sigma(\mathcal{C}_n), \ldots$ の独立性は次の条件と同値である。

$$I \subset \mathbb{N} \text{ 有限部分集合}, A_i \in \mathcal{C}_i\ \forall i \in I \Rightarrow P(\bigcap_{i \in I} A_i) = \prod_{i \in I} P(A_i).$$

証明. 与えられた条件は次と同じである。

$$(\mathrm{diag}_{\mathbb{N}})_* P(A) = P^{\otimes \mathbb{N}}(A)\ \forall A \in \mathrm{Cyl}_{\mathbb{N}} \mathcal{C}..$$

ところが各 $\mathcal{C}_i$ は π システムであるから集合族 $\mathrm{Cyl}_{\mathbb{N}} \mathcal{C}.$ も π システムである。すなわち

$$A, B \in \mathrm{Cyl}_{\mathbb{N}} \mathcal{C}. \Rightarrow A \cap B \in \mathrm{Cyl}_{\mathbb{N}} \mathcal{C}.$$

したがって定理 7.12(i) を適用して

$$(\mathrm{diag}_{\mathbb{N}})_* P(A) = P^{\otimes \mathbb{N}}(A)\ \forall A \in \sigma(\mathrm{Cyl}_{\mathbb{N}} \mathcal{C}.)$$

が導ける。補題 16.13 によれば $\sigma(\mathrm{Cyl}_{\mathbb{N}} \mathcal{C}.) = \bigotimes_{\mathbb{N}} \sigma(\mathcal{C}.)$ なので結論を得た。 □

18.4 系. 部分 σ 加法族の系 $\mathcal{G}_1, \mathcal{G}_2, \ldots, \mathcal{G}_n, \ldots$ が独立であるのは任意の有限部分系が独立であるのと同値である。

18.5 演習問題. 系 18.4 を示せ。

18.6 系. 確率変数系 $X_1, \ldots, X_n, \ldots$ の独立性は部分 σ 加法族系 $\sigma\{X_1\}, \ldots, \sigma\{X_n\}, \ldots$ の独立性と同値である。

証明. 定理 8.4 と定理 18.3 から従う。 □

18.7 例. モデル $((0,1], \mathrm{Borel}((0,1]), \lambda)$ で考える。次の集合を順に A_1, A_2, A_3 とする。

$$(0,1/2],\ (0,1/4] \cup (1/2,3/4],\ (0,1/8] \cup (1/4,3/8] \cup (1/2,5/8] \cup (3/4,7/8].$$

容易にわかるように $\lambda(A_i) = 1/2\ i = 1,2,3$ である。また

$$A_1 \cap A_2 = (0,1/4], A_1 \cap A_3 = (0,1/8] \cup (1/4,3/8], A_2 \cap A_3 = (0,1/8] \cup (1/2,5/8]$$

であり、λ 測度はすべて $(1/2)^2$ である。さらに

$$A_1 \cap A_2 \cap A_3 = (0,1/8],\ \lambda(A_1 \cap A_2 \cap A_3) = (1/2)^3.$$

従って定理 18.3 を適用して系 $\sigma(\{A_1\}), \sigma(\{A_2\}), \sigma(\{A_3\})$ が独立であることがわかる。

18.8 定義. 部分 σ 加法族 $\mathcal{G}, \mathcal{H}$ について $\mathcal{G} \subset \mathcal{H} \bmod P$ とは次の条件が成り立つことをいう。

$$\forall A \in \mathcal{G}\ \exists B \in \mathcal{H} \text{ s.t. } P(A \cap B^c) = 0,\ P(B \cap A^c) = 0.$$

$\mathcal{G} \subset \mathcal{H} \bmod P$ かつ $\mathcal{H} \subset \mathcal{G} \bmod P$ であることを $\mathcal{G} = \mathcal{H} \bmod P$ と書く。

18.9 定理. 部分 σ 加法族の系 $\mathcal{G}_1, \mathcal{G}_2, \dots, \mathcal{G}_n, \dots$ および系 $\mathcal{H}_1, \mathcal{H}_2, \dots, \mathcal{H}_n, \dots$ について $\mathcal{H}_i \subset \mathcal{G}_i \bmod P\ \forall i \in \mathbb{N}$ であるとする。このとき系 $\mathcal{G}.$ が独立であるなら系 $\mathcal{H}.$ も独立である。

18.10 演習問題. 定理 18.3 を使って定理 18.9 を示せ。

18.11 系. 確率変数系 $X_1, \dots, X_n, \dots$ および系 $Y_1, \dots, Y_n, \dots$ について $X_i = Y_i$ P-a.s. $\forall i \in \mathbb{N}$ であるとする。このとき系 $X.$ の独立性と系 $Y.$ の独立性は同値である。

18.12 演習問題. 系 18.11 を示せ。

18.13 定理. 部分 σ 加法族の系 $\mathcal{G}_1, \mathcal{G}_2, \dots, \mathcal{G}_n, \dots$ は独立であるとする。このとき共通部を持たない $\mathbb{N}$ の部分集合 I, J に対して部分 σ 加法族の組 $\bigvee_{i\in I} \mathcal{G}_i, \bigvee_{j\in J} \mathcal{G}_j$ は独立である。

証明. 定理 18.3 により K が $I \cup J$ の有限部分集合でかつ $A_i \in \mathcal{G}_i\ \forall i \in K$ ならば

$$\begin{aligned} P(\bigcap_{i\in K\cap I} A_i \cap \bigcap_{i\in K\cap J} A_i) &= P(\bigcap_{i\in K} A_i) = \prod_{i\in K} P(A_i) \\ &= \prod_{i\in K\cap I} P(A_i) \prod_{i\in K\cap J} P(A_i) = P(\bigcap_{i\in K\cap I} A_i)P(\bigcap_{i\in K\cap J} A_i). \end{aligned}$$

ここで、$\bigcup_{i\in I} \mathcal{G}_i$ に属する集合の有限共通部で表現できる集合全体の族を $\mathcal{C}_1$ とし、$\bigcup_{j\in J} \mathcal{G}_j$ に属する集合の有限共通部で表現できる集合全体の族を $\mathcal{C}_2$ とする。上の関係は

$$P(A \cap B) = P(A)P(B)\ \forall A \in \mathcal{C}_1, \forall B \in \mathcal{C}_2$$

を意味する。$\mathcal{C}_1, \mathcal{C}_2$ ともに π システムなので、定理 18.3 を適用して組 $\sigma(\mathcal{C}_1), \sigma(\mathcal{C}_2)$ の独立性が導ける。それぞれは $\bigvee_{i\in I} \mathcal{G}_i, \bigvee_{j\in J} \mathcal{G}_j$ に一致する。 □

18.14 系. $X_1, X_2, \dots, X_n, \dots$ を独立な確率変数系とする。共通部を持たない $\mathbb{N}$ の部分集合 I, J に対して部分 σ 加法族の組 $\sigma\{X_i\,; i \in I\}, \sigma\{X_j\,; j \in J\}$ は独立である。

以下、第 1 節で述べたモデルについて独立性に主眼をおいて考察する。

前提

基礎となる確率空間は Lebesgue モデル $((0, 1], \mathrm{Borel}((0, 1]), \lambda)$ である。関数の列 $\xi_k : (0, 1] \to \mathbb{R}$, $k \in \mathbb{N}$ および写像 $\varphi : (0, 1] \to (0, 1]$ は 1 節で述べたものとする。

18.15 補題. 部分 σ 加法族の組 $\sigma\{\xi_1\}, \varphi^*\mathrm{Borel}((0, 1])$ は独立である。

証明. 写像 φ が測度 λ を保存することを考慮に入れて例 6.25(iii) を書き直すことにより

$$\lambda((0, p] \cap B) = \lambda((0, p])\lambda(B)\ \forall B \in \varphi^*\mathrm{Borel}((0, 1])$$

を得る。さて A を $\mathbb{R}$ の部分集合とするとき

$$\xi_1^{-1}(A) = \begin{cases} (0, 1] & \text{if } 1 \in A, -1 \in A \\ (0, p] & \text{if } 1 \in A, -1 \notin A \\ (p, 1] & \text{if } 1 \notin A, -1 \in A \\ \emptyset & \text{if } 1 \notin A, -1 \notin A \end{cases}$$

である。従って $\sigma\{\xi_1\} = \{\emptyset, (0,p], (p,1], (0,1]\}$ であるから区間 $(0,p]$ は σ 加法族 $\sigma\{\xi_1\}$ を生成する。以上より定理 18.3 を適用して結論を導くことができる。 □

記号

$\eta_0(x,\omega) := x,\ \eta_k(x,\omega) := \eta_{k-1}(x,\omega) + \xi_k(\omega)$ for $k \in \mathbb{N}$,
$\tau(x,y,\omega) := \min\{k \in \mathbb{N} \cup \{0\} : \eta_k(x,\omega) = y\}$.

ここで $\min\emptyset = +\infty$ と約束している。

18.16 補題. 各 $x, y \in \mathbb{Z}$ に対して関数 $\tau(x,y,\cdot) : (0,1] \to \overline{\mathbb{R}}$ は $\mathrm{Borel}((0,1])$ 可測である。

証明. 定義により $\mathrm{Image}\,\tau(x,y,\cdot) \subset \mathbb{N} \cup \{0, +\infty\}$ であるから可測性の判定について補題 2.4 が適用できる。$n \in \mathbb{N} \cup \{0\}$ としよう。各 $\eta_k(x,\cdot)$ の $\mathrm{Borel}((0,1])$ 可測性により

$$\{\omega \in (0,1] : \tau(x,y,\omega) = n\} = \bigcap_{k=0}^{n-1} \{\omega \in (0,1] : \eta_k(x,\omega) \neq y\} \cap \{\omega \in (0,1] : \eta_n(x,\omega) = y\}$$

は $\mathrm{Borel}((0,1])$ に属する。 □

18.17 補題. $x \neq y$ のとき $\tau(x,y,\omega) = \tau(\eta_1(x,\omega), y, \varphi(\omega)) + 1\ \forall \omega \in (0,1]$ が成り立つ。

証明. まず $x \neq y$ より $\eta_0(x,\omega) \neq y$ である。次に各 $k \in \mathbb{N} \cup \{0\}$ に対して

$$\eta_{k+1}(x,\omega) = x + \xi_1(\omega) + \sum_{i=1}^{k} \xi_{i+1}(\omega) = \eta_1(x,\omega) + \sum_{i=1}^{k} \xi_i(\varphi(\omega)) = \eta_k(\eta_1(x,\omega), \varphi(\omega))$$

であることに注意する。従って $n \in \mathbb{N}$ について次の関係が成り立つ。

$$\begin{aligned} \tau(x,y,\omega) = n &\Leftrightarrow \eta_{k+1}(x,\omega) \neq y\ \forall k < n-1, \eta_n(x,\omega) = y \\ &\Leftrightarrow \eta_k(\eta_1(x,\omega), \varphi(\omega)) \neq y\ \forall k < n-1, \eta_{n-1}(\eta_1(x,\omega), \varphi(\omega)) = y. \end{aligned}$$

右辺は $\tau(\eta_1(x,\omega), y, \varphi(\omega)) = n-1$ ということである。同様にして

$$\tau(x,y,\omega) = +\infty \Leftrightarrow \tau(\eta_1(x,\omega), y, \varphi(\omega)) = +\infty$$

も示せる。よって $\tau(x,y,\omega) > 0$ を考慮に入れて結論が導ける。 □

18.18 補題. $a, b \in \mathbb{Z}$, $a \neq b$ とする。このとき次が成り立つ。

$$\begin{aligned} \lambda(\tau(x,a,\cdot) < \tau(x,b,\cdot)) &= p\lambda(\tau(x+1,a,\cdot) < \tau(x+1,b,\cdot)) \\ &\quad + (1-p)\lambda(\tau(x-1,a,\cdot) < \tau(x-1,b,\cdot))\ \forall x \in \mathbb{Z}, x \neq a, x \neq b. \end{aligned}$$

証明. 各 $x \in \mathbb{Z}$ に対して $A(x) := \{\omega \in (0,1] : \tau(x,a,\omega) < \tau(x,b,\omega)\}$ とおく。補題 18.16 により $A(x) \in \mathrm{Borel}((0,1])$ である。補題 18.17 によれば $x \neq a$, $x \neq b$ のとき

$$\tau(x,a,\omega) = \tau(\eta_1(x,\omega), a, \varphi(\omega)) + 1, \tau(x,b,\omega) = \tau(\eta_1(x,\omega), b, \varphi(\omega)) + 1$$

であるから次の関係が得られる。

$$\begin{aligned}A(x)\cap\{\xi_1=1\}&=\{\omega\in(0,1]:\tau(x+1,a,\varphi(\omega))<\tau(x+1,b,\varphi(\omega))\}\cap\{\xi_1=1\}\\&=\varphi^{-1}(A(x+1))\cap\{\xi_1=1\}.\end{aligned}$$

したがって補題 18.15 を適用すると

$$\lambda(A(x)\cap\{\xi_1=1\})=\lambda(\varphi^{-1}(A(x+1)))\lambda(\{\xi_1=1\}).$$

例 6.25 によれば右辺は $p\lambda(A(x+1))$ に等しい。同様にして

$$\lambda(A(x)\cap\{\xi_1=-1\})=(1-p)\lambda(A(x-1))$$

も導ける。$\{\xi_1=1\}\cup\{\xi_1=-1\}=(0,1]$ なので目標とする関係式が得られた。 □

18.19 定理. $a,b,x\in\mathbb{Z}$, $a\neq b$, $\min\{a,b\}\leq x\leq\max\{a,b\}$ とするとき

$$\lambda(\tau(x,a,\cdot)<\tau(x,b,\cdot))=\begin{cases}\dfrac{(1/p-1)^{x-a}-(1/p-1)^{b-a}}{1-(1/p-1)^{b-a}} & p\neq 1/2\\ \dfrac{b-x}{b-a} & p=1/2\end{cases}.$$

18.20 演習問題. 補題 18.18 を使って定理 18.19 を導け。

18.21 定理. $a,x\in\mathbb{Z}$, $x\neq a$ とするとき

$$\lambda(\tau(x,a,\cdot)<+\infty)=\begin{cases}(1/p-1)^{x-a} & p>1/2, x>a\\ (1/(1-p)-1)^{a-x} & p<1/2, x<a\\ 1 & \text{otherwise}\end{cases}.$$

従って $\lambda(\tau(x,a,\cdot)<+\infty)=1\Leftrightarrow p=1/2$ or "$p>1/2,\ x<a$" or "$p<1/2,\ x>a$" である。

証明. $\tau(x,x+n,\omega)+1\leq\tau(x,x+n+1,\omega)$ であるから、各 $\omega\in(0,1]$, $x\in\mathbb{Z}$ に対して $\tau(x,x+n,\omega)$, $n\in\mathbb{N}$ は単調非減少かつその極限は $+\infty$ である。従って

$$\{\omega\in(0,1]:\tau(x,a,\omega)<+\infty\}=\bigcup_{n=1}^{\infty}\{\omega\in(0,1]:\tau(x,a,\omega)<\tau(x,x+n,\omega)\}.$$

右辺の集合列は n について非減少であるから

$$\lambda(\tau(x,a,\cdot)<+\infty)=\sup_{n\in\mathbb{N}}\lambda(\{\tau(x,a,\cdot)<\tau(x,x+n,\cdot)\}).$$

$x>a$ の場合、右辺のそれぞれに定理 18.19 が適用できるので、結論が得られる。$x<a$ の場合は参照する事象の取り方を適宜変えて議論すればよい。 □

18.22 補題. $a,b\in\mathbb{Z}$, $a\neq b$ とする。このとき次が成り立つ。

$$\begin{aligned}&\lambda(\exists n\in\mathbb{N}\text{ s.t. }\eta_n(a,\cdot)=a, n<\tau(a,b,\cdot))\\&=p\lambda(\tau(a+1,a,\cdot)<\tau(a+1,b,\cdot))+(1-p)\lambda(\tau(a-1,a,\cdot)<\tau(a-1,b,\cdot)).\end{aligned}$$

18.23 演習問題. 補題 18.22 を示せ

18.24 定理. 各 $x\in\mathbb{R}$ に対して $\lambda(\exists n\in\mathbb{N}\text{ s.t. }\eta_n(x,\cdot)=x)=2\min\{p,1-p\}$.

18.25 演習問題. 補題 18.22 と定理 18.19 を使って定理 18.24 を導け。

19　ランダムウォークの再帰性と非再帰性

Lebesgue モデル $((0,1],\mathrm{Borel}((0,1]),\lambda)$ 上で第 1 節で述べた関数の列 $\xi_k:(0,1]\to\mathbb{R}$, $k\in\mathbb{N}$ は独立で同じ分布を持つ確率変数系である。それから派生する関数の列 $\eta_k(x,\cdot):(0,1]\to\mathbb{R}$, $k\in\mathbb{N}\cup\{0\}$ は $\mathbb{R}$ 上のランダムウォークとなるが、定理 18.24 によるとそれが出発点 x に戻る確率が 1 であるのは $p=1/2$ のときに限られる。ここでは、一般のランダムウォークについて考察を展開する。

前提

$Y.$ は $Y_0=0$ をみたす $\mathbb{R}^d$ 上のランダムウォーク。$X_k:=Y_k-Y_{k-1}$ for $k\in\mathbb{N}$.

19.1 補題. $n\in\mathbb{N}$ とする。部分 σ 加法族の組 $\sigma\{X_1,\dots,X_n\}$, $\sigma\{X_i\,;i\in\mathbb{N},i>n\}$ は独立である。また確率変数の系 $X_1,\ \dots,\ X_k,\ \dots$ と系 $X_{n+1},\ \dots,\ X_{n+k},\ \dots$ は同じ結合分布を持つ。

証明. 前半の主張は系 18.14 から従う。また後半についてはどちらも $\mathcal{L}(X_1,\cdot)$ の無限直積に等しいことからわかる。 □

記号

$Z_n:\Omega\to\prod_{\mathbb{N}}\mathbb{R}^d$, $\omega\mapsto(X_n(\omega),X_{n+1}(\omega),\dots,X_{n+k-1}(\omega),\dots)$ for $n\in\mathbb{N}$.
$\psi:\prod_{\mathbb{N}}\mathbb{R}^d\to\prod_{\mathbb{N}\cup\{0\}}\mathbb{R}^d$, $x\mapsto(0,x(1),x(1)+x(2),\dots,\sum_{i=1}^k x(i),\dots)$.

19.2 補題. (i) 写像 ψ は対 $\bigotimes_{\mathbb{N}}\mathrm{Borel}(\mathbb{R}^d)$, $\bigotimes_{\mathbb{N}\cup\{0\}}\mathrm{Borel}(\mathbb{R}^d)$ について可測である。
(ii) $\Lambda:=\{y\in\prod_{\mathbb{N}\cup\{0\}}\mathbb{R}^d : y(i)\neq 0\ \forall i>0\}\in\bigotimes_{\mathbb{N}\cup\{0\}}\mathrm{Borel}(\mathbb{R}^d)$ でありかつ

$$\{Y_i-Y_n\neq 0\ \forall i>n\}=\{\psi\circ Z_{n+1}\in\Lambda\}\ \forall n\in\mathbb{N}\cup\{0\}.$$

証明. (i) は系 16.7 を適用して得られる。また Λ は $\bigcap_{i=1}^{\infty}\mathrm{proj}_i^{-1}(\{y\in\mathbb{R}^d : y\neq 0\})$ と表されることから (ii) の前半がわかる。後半は $Y_i-Y_n=\sum_{k=1}^{i-n}X_{n+k}$ から従う。 □

19.3 補題. $P(Y_n=0,Y_i\neq 0\ \forall i>n)=P(Y_n=0)P(Y_i\neq 0\ \forall i>0)\ \forall n\in\mathbb{N}\cup\{0\}$.

証明. 左辺内の事象は $\{Y_n=0,Y_i-Y_n\neq 0\ \forall i>n\}$ に等しい。更に、補題 19.2(ii) によると、これは $\{Y_n=0,Z_{n+1}\in\psi^{-1}(\Lambda)\}$ と一致する。よって

$$P(Y_n=0,Y_i\neq 0\ \forall i>n)=P(Y_n=0,Z_{n+1}\in\psi^{-1}(\Lambda)).$$

さて $\{Y_n=0\}$ は $\sigma\{X_1,\dots,X_n\}$ に属する事象であるが、この部分 σ 加法族は補題 19.1 によれば $\sigma\{X_i\,;i\in\mathbb{N},i>n\}$ と独立である。他方、定理 16.6 により $\sigma\{Z_{n+1}\}=\sigma\{X_i\,;i\in\mathbb{N},i>n\}$ が成り立っている。従って

$$P(Y_n=0,Z_{n+1}\in\psi^{-1}(\Lambda))=P(Y_n=0)P(Z_{n+1}\in\psi^{-1}(\Lambda)).$$

ここでふたたび補題 19.1 と補題 19.2(ii) を適用すると

$$P(Z_{n+1}\in\psi^{-1}(\Lambda))=P(Z_1\in\psi^{-1}(\Lambda))=P(Y_i\neq 0\ \forall i>0)$$

が導け、ゆえに主張が示せた。 □

19.4 定理. (i) $P(Y_i \neq 0\ \forall i > 0) = 1/\sum_{n=0}^{\infty} P(Y_n = 0)$ ただし $0 = 1/\infty$ と読む。
(ii) $P(\exists n \in \mathbb{N} \text{ s.t. } Y_n = 0) = 1 \Leftrightarrow \sum_{n=0}^{\infty} P(Y_n = 0) = +\infty$.
(iii) $P(\exists n \in \mathbb{N} \text{ s.t. } Y_n = 0) < 1 \Rightarrow P(\#\{n \in \mathbb{N} : Y_n = 0\} < +\infty) = 1$.
(iv) $P(\exists n \in \mathbb{N} \text{ s.t. } Y_n = 0) = 1 \Rightarrow P(\#\{n \in \mathbb{N} : Y_n = 0\} = +\infty) = 1$.

証明. 事象 $\{\#\{n \in \mathbb{N} : Y_n = 0\} < +\infty\}$ は排反事象列 $\{Y_n = 0, Y_i \neq 0\ \forall i > n\}$ $n \in \mathbb{N} \cup \{0\}$ の合併である。従って補題 19.3 により

$$(\star 1) \qquad P(\#\{n \in \mathbb{N} : Y_n = 0\} < +\infty) = P(Y_i \neq 0\ \forall i > 0) \sum_{n=0}^{\infty} P(Y_n = 0).$$

ただし右辺においては $0\infty = 0$ と読む。左辺は有限量であるから、

$$\sum_{n=0}^{\infty} P(Y_n = 0) = +\infty \Rightarrow P(Y_i \neq 0\ \forall i > 0) = 0.$$

従って $\sum_{n=0}^{\infty} P(Y_n = 0) = +\infty$ のとき $0 = 1/\infty$ という形で (i) が成り立つことになる。さて

$$\{\#\{n \in \mathbb{N} : Y_n = 0\} < +\infty\} = \bigcup_{n=1}^{\infty} \bigcap_{k \geq n} \{Y_k \neq 0\}$$

である。$\sum_{n=0}^{\infty} P(Y_n = 0) < +\infty$ なら右辺の余事象の確率は補題 15.6 により 0 である。よって

$$(\star 2) \qquad \sum_{n=0}^{\infty} P(Y_n = 0) < +\infty \Rightarrow P(\#\{n \in \mathbb{N} : Y_n = 0\} < +\infty) = 1.$$

また $(\star 1)$ と組み合わせて $\sum_{n=0}^{\infty} P(Y_n = 0) < +\infty$ のときも (i) が成り立つことが分かる。同値性 (ii) は関係式 (i) から直ちに得られる。更に同値性 (ii) によって $(\star 2)$ を読み替えると (iii) が得られる。他方

$$P(\exists i > 0 \text{ s.t. } Y_i = 0) = 1 \Leftrightarrow P(Y_i \neq 0\ \forall i > 0) = 0$$

であるから $(\star 1)$ より (iv) が導かれる。 □

19.5 定義. 条件 $P(\exists n \in \mathbb{N} \text{ s.t. } Y_n = 0) = 1$ が成り立つときランダムウォーク $Y_{\cdot}$ は**再帰的**(recurrent) であるという。

19.6 例. $\mathbb{R}$ 上のランダムウォーク $Y_{\cdot}$ について $0 < p < 1$, $P(X_1 = 1) = p$, $P(X_1 = -1) = 1-p$ とする。このとき $Y_{\cdot}$ が再帰的であるための必要十分条件は $p = 1/2$ である。

証明. Y_n の分布を求めるためその特性関数を計算する。定理 10.19 により

$$\begin{aligned}\mathrm{char}(Y_n, \xi) &= \mathrm{char}(X_1, \xi)^n = (pe^{\sqrt{-1}\xi} + (1-p)e^{-\sqrt{-1}\xi})^n \\ &= \sum_{j=0}^{n} \binom{n}{j} p^j (1-p)^{n-j} e^{\sqrt{-1}\xi(2j-n)}\ \forall \xi \in \mathbb{R}.\end{aligned}$$

特性関数の一意性を適用して次を得る。

$$P(Y_n = x) = \begin{cases} 0 & x+n \text{ odd} \\ \binom{n}{(x+n)/2} p^{(x+n)/2}(1-p)^{(n-x)/2} & x+n \text{ even} \end{cases}.$$

さて $\binom{2m}{m} = \binom{-1/2}{m}(-4)^m$ であるから $p \neq 1/2$ のとき二項定理より次を得る。

$$\sum_{m=0}^{\infty} \binom{-1/2}{m}(-4)^m p^m (1-p)^m = \frac{1}{\sqrt{1-4p(1-p)}} = \frac{1}{|1-2p|}.$$

$p = 1/2$ のとき上の級数は発散する。従って定理 19.4(i) より次が成立する。

$$P(Y_i \neq 0\ \forall i > 0) = \frac{1}{\sum_{n=0}^{\infty} P(Y_n = 0)} = |1-2p| \text{ 但し } 1/\infty = 0$$

$1 - |1-2p| = 2\min\{p, 1-p\}$ であるから定理 18.24 で述べたことと符合している。 $\square$

19.7 注意. $\mathbb{R}$ 上のランダムウォーク $Y.$ について $0 < p < 1$, $P(X_1 = 1) = p$, $P(X_1 = -1) = 1-p$ とする。このとき大数の強法則より

$$\exists A \in \mathcal{F} \text{ s.t. } P(A) = 1, \lim_{n\to\infty} Y_n(\omega)/n = 2p-1\ \forall \omega \in A.$$

さて $p \neq 1/2$ としよう。$\omega \in A$ ならある番号より大きい n について $|Y_n(\omega)/n| > |2p-1|/2$ である。従って $P(\#\{n \in \mathbb{N} : Y_n = 0\} < +\infty) = 1$ が成り立ち非再帰性と符合している。

19.8 演習問題. $\mathbb{R}^2$ 上のランダムウォーク $Y.$ について

$$P(X_1 = (1,0)) = P(X_1 = (-1,0)) = P(X_1 = (0,1)) = P(X_1 = (0,-1)) = 1/4$$

とする。このとき $Y.$ は再帰的であることを示せ。

19.9 演習問題. $\mathbb{R}^3$ 上のランダムウォーク $Y.$ について

$$P(X_1 = e_i) = 1/6,\ P(X_1 = -e_i) = 1/6\ \forall i = 1,2,3$$

とする。ただし e_1, e_2, e_3 は $\mathbb{R}^3$ の標準基底である。このとき $Y.$ は再帰的でないことを示せ。

20 可微分同相写像と Lebesgue 測度

第 9 節では可逆アファイン写像についてその行列式の絶対値が測度の比を表していることを見た。さて $\mathbb{R}^d$ 上の写像については、微分可能性は局所的に線形化ができることを意味する。このとき局所的な測度の比は Jacobi 行列式の絶対値で与えられると類推するのが自然で、それを表すのがいわゆる変数変換公式である。

前提

D を $\mathbb{R}^d$ の空でない開集合、$\varphi : D \to \mathbb{R}^d$ を C^1 級写像で単射かつ $\det \varphi'(x) \neq 0\ \forall x \in D$ を満たすものとする。ここで $\varphi'(x)$ は x における微分を表す d 次正方行列である。

20.1 注意. 逆写像定理により $\varphi(D)$ は開集合かつ逆写像 $\varphi^{-1} : \varphi(D) \to D$ も C^1 級である。

以下で考察する行列の成分はすべて実数値である。

記号

$\|\cdot\|$ $\mathbb{R}^d$ 上のユークリッドノルム　$\mathbf{1} :=$ d 次単位行列　$\mathbf{0} :=$ d 次零行列

20.2 定義. d 次正方行列 A に対して $\|A\| := \sup\{\|Au\|\,;u \in \mathbb{R}^d, \|u\| = 1\}$

20.3 補題. d 次正方行列 A, B に対して以下が成り立つ。
(i) $0 \leq \|A\| < +\infty$. $\|Au\| \leq \|A\|\|u\|\ \forall u \in \mathbb{R}^d$. $\|AB\| \leq \|A\|\|B\|$.
(ii) $\|A + B\| \leq \|A\| + \|B\|$. $\|cA\| = |c|\|A\|\ \forall c \in \mathbb{R}$. $\|A\| = 0 \Leftrightarrow A = \mathbf{0}$.
(iii) $\|A\|^2$ は対称行列 ${}^{\mathrm{t}}AA$ の最大固有値に等しい。$\|A\|^2 \leq \mathrm{trace}({}^{\mathrm{t}}AA)$.

証明. (i), (ii) については演習問題とする。

(iii) ユークリッド内積を $(\cdot,\cdot)$ と表記する。その定義により $\|A\|^2$ は次で与えられる。

$$\|A\|^2 := \sup\{({}^{\mathrm{t}}AAu, u)\,; u \in \mathbb{R}^d, \|u\| = 1\}.$$

${}^{\mathrm{t}}AA$ の固有値を大きいものから順に並べたものを $c_1, c_2, \ldots, c_d$ また対応する固有ベクトルからなる正規直交基底を $e_1, e_2, \ldots, e_d$ とする。このとき

$$({}^{\mathrm{t}}AAu, u) = \sum_{i=1}^{d} c_i(u, e_i)^2 \leq c_1 \sum_{i=1}^{d}(u, e_i)^2 = c_1\|u\|^2\ \forall u \in \mathbb{R}^d$$

より $\|A\|^2 = c_1$ が従う。また $c_1 \leq \sum_{i=1}^{d} c_i = \mathrm{trace}({}^{\mathrm{t}}AA)$ である。 □

20.4 演習問題. 補題 20.3(i), (ii) を示せ。

20.5 系. d 次正方行列 A に対して $\|A - \mathbf{1}\| < 1$ が成り立つとする。
(i) $\det A \geq (1 - \|A - \mathbf{1}\|)^d$. A は逆行列を持つ。
(ii) $\|A^{-1}B - \mathbf{1}\| \leq (\|A - \mathbf{1}\| + \|B - \mathbf{1}\|)/(1 - \|A - \mathbf{1}\|)$.

証明. (i) c を非負定値対称行列 ${}^{\mathrm{t}}AA$ の最小固有値とし、対応する正規化された固有ベクトルを u とする。ユークリッド内積を $(\cdot,\cdot)$ と表記する。三角不等式と補題 20.3(i) により

$$\sqrt{c} = \sqrt{({}^{\mathrm{t}}AAu, u)} = \|Au\| \geq \|u\| - \|(A - \mathbf{1})u\| \geq \|u\| - \|(A - \mathbf{1})\|\|u\| = 1 - \|(A - \mathbf{1})\|$$

となる。右辺は正値であることに注意する。従って

$$(\det A)^2 = \det({}^{\mathrm{t}}AA) \geq c^d \geq (1 - \|(A - \mathbf{1})\|)^{2d}$$

である。とくに $\det A \neq 0$ がわかるので、A は逆行列を持つ。また $0 \leq t \leq 1$ なら、補題 20.3(ii) により $\mathbf{1} + t(A - \mathbf{1})$ も同じ条件を満たすので $\det(\mathbf{1} + t(A - \mathbf{1})) \neq 0$ である。次に関数

$$f : [0, 1] \to \mathbb{R},\ t \mapsto \det(\mathbf{1} + t(A - \mathbf{1}))$$

を考察する。この関数 f は連続で $f(0)=1>0$ かつ $f(t)\neq 0\ \forall t$ である。故に中間値の定理を適用して $\det A=f(1)>0$ が導かれた。

(ii) $u\in\mathbb{R}^d$ に対して $v=(A^{-1}B-\mathbf{1})u$ とおく。$v=(B-\mathbf{1})u-(A-\mathbf{1})(u+v)$ であるから三角不等式と補題 20.3(i) により

$$\|v\|\le\|B-\mathbf{1}\|\|u\|+\|A-\mathbf{1}\|(\|u\|+\|v\|)$$

が従う。$\|A-\mathbf{1}\|\|v\|$ を左辺に移項の後 $1-\|A-\mathbf{1}\|>0$ で両辺を割ることにより

$$\|(A^{-1}B-1)u\|=\|v\|\le\frac{\|B-\mathbf{1}\|+\|A-\mathbf{1}\|}{1-\|A-\mathbf{1}\|}\|u\|$$

を得る。$u\in\mathbb{R}^d$ は任意であるので結論に到達した。 $\square$

記号

$\mathrm{Nbd}(a,\delta):=\{x\in\mathbb{R}^d:\|x-a\|<\delta\}$ ただし $a\in\mathbb{R}^d$, $\delta>0$

20.6 補題. $t\in D$ と $\varepsilon>0$ に対して次の条件を満たす $\delta>0$ が存在する。

$$\mathrm{Nbd}(t,3\delta)\subset D,\ \|\varphi'(t)^{-1}\varphi'(x)-\mathbf{1}\|<\varepsilon/(2+\varepsilon)\ \forall x\in\mathrm{Nbd}(t,2\delta).$$

証明. 行列 A に対して $\mathrm{trace}({}^{\mathrm{t}}AA)$ は行列成分の 2 乗和に等しい。また $x\mapsto\varphi'(t)^{-1}\varphi'(x)$ は連続であるから、補題 20.3(ii), (iii) を適用して関数 $x\mapsto\|\varphi'(t)^{-1}\varphi'(x)-1\|$ の連続性が導かれる。D は開集合であることを考慮に入れて結論に到達する。 $\square$

上の結果を有界閉集合上での一様評価に拡張しておく。

20.7 補題. K を $K\subset D$ なる有界閉集合とする。任意の $\varepsilon>0$ に対し $\delta>0$ が存在して

$$\begin{aligned}&\mathrm{Nbd}(t,2\delta)\subset D\ \forall t\in K;\ \|\varphi'(t)^{-1}\varphi'(x)-\mathbf{1}\|<\varepsilon\ \forall t\in K,\forall x\in\mathrm{Nbd}(t,\delta);\\&\|\varphi'(t)^{-1}(\varphi(x)-\varphi(y))-(x-y)\|<\varepsilon\|x-y\|\ \forall t\in K\ \forall x\forall y\in\mathrm{Nbd}(t,\delta)\ x\neq y\end{aligned}$$

が成り立つ。

証明. 補題 20.6 により各 $z\in K$ に対して次のような $\delta(z)>0$ が存在する。

$$\mathrm{Nbd}(z,3\delta(z))\subset D,\ \|\varphi'(z)^{-1}\varphi'(x)-\mathbf{1}\|<\varepsilon/(2+\varepsilon)\ \forall x\in\mathrm{Nbd}(z,2\delta(z)).$$

K は有界閉集合であるから Heine-Borel の被覆定理によりある有限集合 F が存在して

$$F\subset K,\ K\subset\bigcup_{z\in F}\mathrm{Nbd}(z,\delta(z)).$$

このとき $\delta:=\min\{\delta(z)\,;z\in F\}>0$ である。また任意の $t\in K$ に対して $z\in F$ が存在して $\|t-z\|<\delta(z)$ が成り立つ。そのような t,z について

$$\|x-t\|<2\delta\Rightarrow\|x-z\|\le\|x-t\|+\|t-z\|<2\delta+\delta(z)\le3\delta(z)$$

である。もともと $\mathrm{Nbd}(z,3\delta(z))\subset D$ であったので $\mathrm{Nbd}(t,2\delta)\subset D$ が得られる。次に

$$\|x-t\| < \delta \Rightarrow \|x-z\| \le \|x-t\| + \|t-z\| < \delta + \delta(z) \le 2\delta(z)$$
$$\Rightarrow \|\varphi'(z)^{-1}\varphi'(x) - \mathbf{1}\| < \varepsilon/(2+\varepsilon).$$

他方、$\|\varphi'(z)^{-1}\varphi'(t) - \mathbf{1}\| < \varepsilon/(2+\varepsilon)$ も成り立っている。ここで

$$\frac{2\varepsilon/(2+\varepsilon)}{1-\varepsilon/(2+\varepsilon)} = \varepsilon$$

かつ $\varphi'(t)^{-1}\varphi'(x) = (\varphi'(z)^{-1}\varphi'(t))^{-1}\varphi'(z)^{-1}\varphi'(x)$ なので系 20.5(ii) を適用して $x \in \mathrm{Nbd}(t,\delta)$ なら $\|\varphi'(t)^{-1}\varphi'(x) - \mathbf{1}\| < \varepsilon$ であることが導かれた。3 番目の条件も成り立つことを示すのは演習問題とする。 □

20.8 演習問題. 補題 20.7 の証明を完成させよ。ヒント 次の関係を使え。

$$\varphi(x) - \varphi(y) = \int_0^1 \varphi'(y + s(x-y))(x-y)\,ds.$$

記号

$E \in \mathrm{Borel}(\mathbb{R}^d)$ に対し集合族 $\{B \subset E : B \in \mathrm{Borel}(\mathbb{R}^d)\}$ を $\mathrm{Borel}(E)$ と書く。

20.9 補題. (i) $E \in \mathrm{Borel}(\mathbb{R}^d)$ に対し $\mathrm{Borel}(E)$ は集合 E 上の σ 加法族である。
(ii) $\mathcal{C}$ を $\mathbb{R}^d$ の部分集合族で $\sigma(\mathcal{C}) = \mathrm{Borel}(\mathbb{R}^d)$ をみたすもの、$E \in \mathrm{Borel}(\mathbb{R}^d)$ とする。このとき $\mathrm{Borel}(E)$ は集合族 $\{E \cap C\,; C \in \mathcal{C}\}$ で生成される。
(iii) D, E をそれぞれ $\mathbb{R}^d$, $\mathbb{R}^n$ の空でない開集合とする。このとき連続写像 $f : D \to E$ は対 $\mathrm{Borel}(D)$, $\mathrm{Borel}(E)$ に関して可測である。

証明. (ii) 包含写像 $\iota : E \to \mathbb{R}^d$ を使うと $\mathrm{Borel}(E)$ は $\iota^*\mathrm{Borel}(\mathbb{R}^d)$ と $\{E \cap C\,; C \in \mathcal{C}\}$ は $\iota^*\mathcal{C}$ と一致する。したがって補題 5.6(iii) を適用して結論に至る。 □

20.10 演習問題. 補題 20.9(iii) を示せ。

逆写像 φ^{-1} も連続なので $\varphi(A) \in \mathrm{Borel}(\varphi(D))\ \forall A \in \mathrm{Borel}(D)$ であることに注意しよう。

記号

$\mathcal{I} := \{J \subset \mathbb{R}^d :$ 有界な左半開区間の直積で表現できる $\}$ ここでは $\emptyset \notin \mathcal{I}$ とする。
$A \subset \mathbb{R}^d$ に対してそのユークリッド位相に関する閉包を $\overline{A}$ と表記する。

20.11 補題. $\mathrm{Borel}(D)$ は集合族 $\{J \in \mathcal{I} : \overline{J} \subset D\}$ で生成される。

証明. 両端が有理数であるような左半開区間の直積で表現できる集合 J であって $\overline{J} \subset D$ を満たすもの全体を $\mathcal{I}_0$ とする。演習問題 5.3 と同様な考察により $\bigcup_{J\in\mathcal{I}_0} J = D$ であることが導ける。集合族 $\mathcal{I}_0$ は可算であるから

$$D \cap K = \bigcup_{J\in\mathcal{I}_0} (J \cap K) \in \sigma(\{J \in \mathcal{I} : \overline{J} \subset D\})\ \forall K \in \mathcal{I}.$$

従って $\sigma(\{D \cap K\,; K \in \mathcal{I}\}) \subset \sigma(\{J \in \mathcal{I} : \overline{J} \subset D\}) \subset \mathrm{Borel}(D)$ である。他方、補題 20.9(ii) を適用して $\sigma(\{D \cap K\,; K \in \mathcal{I}\}) = \mathrm{Borel}(D)$ が分かるので証明できた。 □

記号

$J \subset \mathbb{R}^d$ に対し $\operatorname{diam} J := \sup\{\|x-y\| \,; x, y \in J\}$, $\operatorname{proj}_i : \mathbb{R}^d \to \mathbb{R}$ 第 i 座標への射影

以下しばらく $\overline{I} \subset D$ なる $I \in \mathcal{I}$ と次を満たす $r > 0$ を固定する。

$$\operatorname{diam} I \le r \min_{1 \le i \le d} \lambda(\operatorname{proj}_i I)$$

ここで λ は 1 次元 Lebesgue 測度である。

20.12 補題. $K = \overline{I}$ と $0 < \varepsilon < 1$ なる ε に対し $\delta > 0$ を補題 20.7 で述べたものとする。このとき $J \in \mathcal{I}$ が条件 $J \subset I$, $\operatorname{diam} J < 2\delta$, $\operatorname{diam} J \le r \min_{1 \le i \le d} \lambda(\operatorname{proj}_i J)$ を満たせば

$$\lambda^{(d)}(\varphi(J)) \le \frac{(1+r\varepsilon)^d}{(1-\varepsilon)^d} \int_J |\det \varphi'| \lambda^{(d)}.$$

証明. $x \in J$ とする。各 $i = 1, 2, \ldots, d$ に対して $\operatorname{proj}_i u \le \|u\|$ であるから

$$\operatorname{proj}_i\{\varphi'(t)^{-1}(\varphi(x) - \varphi(t))\} \le \operatorname{proj}_i(x-t) + \|\varphi'(t)^{-1}(\varphi(x) - \varphi(t)) - (x-t)\|.$$

ここで t は J の中心である。右辺第 2 項に補題 20.7 を適用して

$$\operatorname{proj}_i\{\varphi'(t)^{-1}(\varphi(x) - \varphi(t))\} < \operatorname{proj}_i(x-t) + \varepsilon\|x-t\| \quad \text{但し } x \ne t.$$

右辺第 1 項は $\lambda(\operatorname{proj}_i J)/2$ で、第 2 項は $\varepsilon \operatorname{diam} J/2$ で押さえられる。よって

$$\operatorname{proj}_i\{\varphi'(t)^{-1}(\varphi(x) - \varphi(t))\} < (1+r\varepsilon)\lambda(\operatorname{proj}_i J)/2$$

が得られる。ここで $\operatorname{diam} J \le r\lambda(\operatorname{proj}_i J)$ という仮定を使った。同様にして

$$-(1+r\varepsilon)\lambda(\operatorname{proj}_i J)/2 < \operatorname{proj}_i\{\varphi'(t)^{-1}(\varphi(x) - \varphi(t))\}$$

もわかるので

$$\operatorname{proj}_i\{\varphi'(t)^{-1}(\varphi(x) - \varphi(t))\} \in (1+r\varepsilon)\operatorname{proj}_i(J-t) \ \forall i = 1, 2, \ldots, d, \ \forall x \in J$$

が導かれる。すなわち次の包含関係が成り立つ。

$$\varphi(J) \subset (1+r\varepsilon)\varphi'(t)(J-t) + \varphi(t).$$

右辺の集合の測度について定理 9.17 を適用すると

$$\lambda^{(d)}(\varphi(J)) \le (1+r\varepsilon)^d |\det \varphi'(t)| \lambda^{(d)}(J).$$

他方、補題 20.7 と系 20.5(i) から $(1-\varepsilon)^d |\det \varphi'(t)| \le |\det \varphi'(x)| \ \forall x \in J$ なので

$$|\det \varphi'(t)| \lambda^{(d)}(J) \le \frac{1}{(1-\varepsilon)^d} \int_J |\det \varphi'| \lambda^{(d)}$$

が従う。以上より結論に到達した。 □

20.13 系. $I \in \mathcal{I}$ とする。$\overline{I} \subset D$ であれば

$$\lambda^{(d)}(\varphi(I)) \leq \int_I |\det \varphi'|\lambda^{(d)}.$$

証明. $K = \overline{I}$ と $0 < \varepsilon < 1$ なる ε に対し $\delta > 0$ を補題 20.7 で述べたものとする。$n \in \mathbb{N}$ を $\operatorname{diam} I/2^n < 2\delta$ となるように選ぶ。I の各辺を 2^n 等分割して得られる I の $\mathcal{I}$ 分割を Δ とする。各 $J \in \Delta$ について補題 20.12 が適用でき、また $A \mapsto \lambda^{(d)}(\varphi(A))$ は $(D, \mathrm{Borel}(D))$ 上の測度であるので次が成り立つ。

$$\lambda^{(d)}(\varphi(I)) \leq \frac{(1+r\varepsilon)^d}{(1-\varepsilon)^d}\int_I |\det \varphi'|\lambda^{(d)}.$$

ε は $0 < \varepsilon < 1$ である限り任意なので結論が導かれた。 □

20.14 補題. $\mathcal{C}$ を S の部分集合の族で有限加法的測度の条件 (i) を満たすもの、μ, ν を $(S, \sigma(\mathcal{C}))$ 上の測度とする。ν が $\mathcal{C}$ 上 σ 有限でありかつ $\mu(A) \leq \nu(A)\ \forall A \in \mathcal{C}$ が成り立てば、$\mu(A) \leq \nu(A)$ $\forall A \in \sigma(\mathcal{C})$ である。

証明. σ 有限性により ν を $\mathcal{C}$ 上に制限してできる有限加法的測度から作られる Carathéodory の外測度は $\sigma(\mathcal{C})$ 上でもとの ν と一致する。すなわち

$$\nu(A) = \inf\{\sum_{J\in\Delta}\nu(J)\,;\Delta\ A\text{ の可算 }\mathcal{C}\text{ 被覆 }\}\ \forall A \in \sigma(\mathcal{C}).$$

よって $\mu(A) \leq \nu(A)\ \forall A \in \mathcal{C} \Rightarrow \mu(A) \leq \nu(A)\ \forall A \in \sigma(\mathcal{C})$ が成り立つ。 □

20.15 系. (i) $\lambda^{(d)}(\varphi(A)) \leq \int_A |\det \varphi'|\lambda^{(d)}\ \forall A \in \mathrm{Borel}(D)$.
(ii) 非負 Borel 可測関数 $f : D \to \overline{\mathbb{R}}$ に対して $\int_{\varphi(D)} f \circ \varphi^{-1}\,\lambda \leq \int_D f|\det \varphi'|\,\lambda^{(d)}$

証明. (i) 集合族 $\mathcal{C} = \{I \in \mathcal{I} : \overline{I} \subset D\} \cup \{\emptyset\}$ は有限加法的測度の条件 (i) を満たす。また補題 20.11 によれば $\mathrm{Borel}(D)$ は $\{I \in \mathcal{I} : \overline{I} \subset D\}$ で生成される。他方、$|\det \varphi'|$ の連続性より測度 $A \mapsto \int_A |\det \varphi'|\lambda^{(d)}$ の σ 有限性が導かれる（これを実行するのは演習問題に委ねる）。従って系 20.13 と補題 20.14 から (i) が得られる。

(ii) $A \in \mathrm{Borel}(D)$ とするとき (i) により $f = 1_A$ については成り立つ。一般の場合はスタンダードマシンが機能する。 □

20.16 演習問題. (i) 非負値連続関数 $\rho : D \to \mathbb{R}$ を密度関数に持つ絶対連続な $(D, \mathrm{Borel}(D))$ 上の測度は集合族 $\{I \in \mathcal{I} : \overline{I} \subset D\} \cup \{\emptyset\}$ について σ 有限であることを示せ。
(ii) 系 20.15(ii) について一般の場合の証明をつけよ。

C^1 級可微分同相写像に関して**変数変換公式**は以下のように述べられる。

20.17 定理. (i) 任意の非負値 $\mathrm{Borel}(\varphi(D))$ 可測関数 $f : \varphi(D) \to \overline{\mathbb{R}}$ に対して次が成り立つ。

$$\int_D f \circ \varphi\,\lambda^{(d)} = \int_{\varphi(D)} f\frac{1}{|\det \varphi'| \circ \varphi^{-1}}\,\lambda^{(d)}.$$

(ii) $\mathrm{Borel}(\varphi(D))$ 可測関数 $f : \varphi(D) \to \overline{\mathbb{R}}$ に対して $f \circ \varphi$ の $\lambda^{(d)}$ 可積分性は $f/|\det \varphi'| \circ \varphi^{-1}$ のそれと同値であり、可積分なら上の等式が成り立つ。

証明. $B \in \mathrm{Borel}(\varphi(D))$ とする。$\varphi^{-1}: \varphi(D) \to D$ に系 20.15(i) を適用して

$$\lambda^{(d)}(\varphi^{-1}(B)) \le \int_B \frac{1}{|\det \varphi'| \circ \varphi^{-1}} \lambda^{(d)}.$$

他方 $f: D \to \overline{\mathbb{R}}$, $x \mapsto 1_B(\varphi(x))/|\det \varphi'(x)|$ に系 20.15(ii) を適用して

$$\int_{\varphi(D)} 1_B \frac{1}{|\det \varphi'| \circ \varphi^{-1}} \lambda^{(d)} \le \int_D 1_B \circ \varphi \lambda^{(d)} = \lambda^{(d)}(\varphi^{-1}(B)).$$

二つを組み合わせて $f = 1_B$ について成り立つことがわかった。あとはスタンダードマシンを適用すればよい。 $\square$

絶対連続な測度の可微分同相写像による**像測度の絶対連続性**については次がいえる。

20.18 系. 非負値 Borel 可測関数 $\rho: D \to \overline{\mathbb{R}}$ を密度関数に持つ絶対連続な $(D, \mathrm{Borel}(D))$ 上の測度を μ とする。このとき像測度 $\varphi_* \mu$ は絶対連続でありその密度関数は

$$\mathbb{R}^d \to \overline{\mathbb{R}}, y \mapsto \begin{cases} \rho(\varphi^{-1}(y))/|\det \varphi'(\varphi^{-1}(y))| & y \in \varphi(D) \\ 0 & y \notin \varphi(D) \end{cases}$$

で与えられる。

20.19 演習問題. 系 20.18 を示せ。

定理 20.17 の典型的な応用例として**極座標変換**を取り上げよう。

20.20 例. $D := \mathbb{R}^2 \setminus \{x \in \mathbb{R}^2 : x_1 \le 0, x_2 = 0\}$ とし、極座標変換

$$\varphi: D \to \mathbb{R}^2,\ x \mapsto (\|x\|, 2 \arctan \frac{x_2}{\|x\| + x_1})$$

に定理 20.17 を適用しよう。$\varphi(D) = (0, +\infty) \times (-\pi, \pi)$ であり

$$\varphi'(x) = \begin{pmatrix} x_1/\|x\| & x_2/\|x\| \\ -x_2/\|x\|^2 & x_1/\|x\|^2 \end{pmatrix},\ \det \varphi'(x) = \frac{1}{\|x\|}\ \forall x \in D$$

である。したがって非負値 Borel 可測関数 $f: (0, +\infty) \times (-\pi, \pi) \to \overline{\mathbb{R}}$ に対して

$$\int_D f(\|x\|, 2 \arctan \frac{x_2}{\|x\| + x_1}) \lambda^{(2)}(dx) = \int_{(0,+\infty) \times (-\pi, \pi)} f(r, \theta)\, r\, \lambda^{(2)}(d(r, \theta))$$

という変数変換公式を得る。特に

$$f(r, \theta) = e^{-r^2/2}$$

の場合、右辺に定理 8.14 と補題 3.26(ii) を適用して

$$\int_D e^{-\|x\|^2/2} \lambda^{(2)}(dx) = 2\pi.$$

$1_D = 1$ $\lambda^{(2)}$-a.e. であるから定理 8.14 により左辺は $\int_{\mathbb{R}} e^{-x^2/2} \lambda(dx)$ の 2 乗である。従って

$$\int_{\mathbb{R}} e^{-x^2/2} \lambda(dx) = \sqrt{2\pi}.$$

20.21 例. $D := (0, +\infty) \times (0, +\infty)$ とし、座標変換

$$\varphi : D \to \mathbb{R}^2,\, x \mapsto (x_1 + x_2, \frac{x_2}{x_1 + x_2})$$

に系 20.18 を適用しよう。$\varphi(D) = (0, +\infty) \times (0, 1)$ であり

$$\varphi'(x) = \begin{pmatrix} 1 & 1 \\ -x_2/(x_1+x_2)^2 & x_1/(x_1+x_2)^2 \end{pmatrix},\ \det \varphi'(x) = \frac{1}{x_1 + x_2}\ \forall x \in D$$

である。したがって非負値 Borel 可測関数 $\rho : D \to \overline{\mathbb{R}}$ と $A \in \mathrm{Borel}(\varphi(D))$ に対して

$$\int_{\varphi^{-1}(A)} \rho\, \lambda^{(2)} = \int_A \rho(y_1(1-y_2), y_1y_2)\, y_1\, \lambda^{(2)}(dy)$$

という変数変換公式を得る。特に $a, b \in \mathbb{R}_{>0}$ として $\rho(x) = e^{-x_1-x_2}{x_1}^{a-1}{x_2}^{b-1}$ の場合には

$$(\star)\qquad \int_{\varphi^{-1}(A)} e^{-x_1}{x_1}^{a-1}e^{-x_2}{x_2}^{b-1}\, \lambda^{(2)}(dx) = \int_A e^{-y_1}{y_1}^{a+b-1}(1-y_2)^{a-1}{y_2}^{b-1}\, \lambda^{(2)}(dy)$$

となる。従って $A = (0, +\infty) \times (0, 1)$ として両辺に定理 8.14 を適用すると次が得られる。

$$\Gamma(a)\Gamma(b) = \Gamma(a+b)B(a,b)\ \forall a, b \in \mathbb{R}_{>0}$$

ここで Γ は gamma 関数、B は beta 関数である。さて指数 a, b の *beta* **分布**とは関数

$$y \mapsto \frac{1}{B(a,b)} 1_{(0,1)}(y)(1-y)^{a-1}y^{b-1}$$

を密度関数とする絶対連続分布をいう。これを $\beta_{a,b}$ と書く。また $t > 0$ に対して指数 $t, 1$ の gamma 分布を γ_t と書くことにしよう。$(\star)$ は以下のように読むことができる。

$$\varphi_*(\gamma_a \otimes \gamma_b) = \gamma_{a+b} \otimes \beta_{a,b}.$$

gamma **分布と** *beta* **分布の関係**は応用上重要である。$A = B \times (0, 1)$ に対して適用すると

$$\gamma_a \otimes \gamma_b(\{(x_1, x_2) : x_1 + x_2 \in B\}) = \gamma_{a+b}(B)\ \forall B \in \mathrm{Borel}(\mathbb{R})$$

が得られるが、これはまさに例 9.7(iii) で述べたことである。

21 Sard の定理と面積公式

写像 φ の単写性が崩れたり、また Jaccobi 行列式が 0 となるような点が存在するときは、定理 20.17 には修正が必要である。記号などは第 20 節におけるものを引き続いて使用する。

前提

D を $\mathbb{R}^d$ の空でない開集合、$\varphi : D \to \mathbb{R}^d$ を C^1 級写像とする。

21.1 定義. φ の**正則点**(regular point) とは $\operatorname{rank}\varphi'(x) = d$ を満たす $x \in D$ のことをいう。正則点でない D の要素を φ の**臨界点**(critical point) という。また臨界点の像を**臨界値**(critical value) という。

21.2 補題. (i) K を $K \subset D$ なる有界閉集合とする。このとき $M := \sup_{t\in K} \|\varphi'(t)\| < +\infty$ かつ K に含まれる凸集合 I に対して $\|\varphi(x) - \varphi(y)\| \le M\|x - y\|\ \forall x\forall y \in I$ が成り立つ。
(ii) $K \subset D$ なる有界閉集合 K と $\varepsilon > 0$ に対し $\delta > 0$ が存在して $\mathrm{Nbd}(t, 2\delta) \subset D\ \forall t \in K$ かつ $\|\varphi(x) - \varphi(t) - \varphi'(t)(x - t)\| \le \varepsilon\|x - t\|\ \forall t \in K\ \forall x \in \mathrm{Nbd}(t, \delta)$ が成り立つ。

21.3 演習問題. 補題 21.2 を示せ。

記号

$\mathrm{crp}(\varphi)$ を φ の臨界点集合、即ち $\mathrm{crp}(\varphi) := \{x \in D : \operatorname{rank}\varphi'(x) < d\}$ とする。

21.4 補題. その閉包が D に含まれる $I \in \mathcal{I}$ に対して $\varphi(\mathrm{crp}(\varphi) \cap I)$ は $\lambda^{(d)}$ 零集合である。

証明. 簡単のため I の閉包を K とかく。また $C := \mathrm{crp}(\varphi)$, $M := \sup_{x\in K} \|\varphi'(x)\|$ とおく。K と $\varepsilon > 0$ に対し $\delta > 0$ を補題 21.2(ii) で述べたものとする。$n \in \mathbb{N}$ を $\operatorname{diam} I/2^n < \delta$ となるように選ぶ。I の各辺を 2^n 等分割して得られる I の $\mathcal{I}$ 分割を Δ とする。このとき

$$(\star)\qquad \lambda^{(d)\star}(\varphi(C \cap I)) \le \lambda^{(d)\star}\Big(\bigcup_{J\in\Delta:J\cap C\neq\emptyset} \varphi(J)\Big) \le \sum_{J\in\Delta:J\cap C\neq\emptyset} \lambda^{(d)\star}(\varphi(J)).$$

ここで $\lambda^{(d)\star}$ は d 次元 Lebesgue 外測度である。ひとまず $J \cap C \neq \emptyset$ を満たす $J \in \Delta$ を固定する。すると $\operatorname{rank}\varphi'(c) < d$ を満たす $c \in J$ が存在する。そこで $\mathrm{proj}_d L\varphi'(c) = 0$ となるように直交行列 L を選ぶ。$x \in J$ とする。J は凸集合であるから補題 21.2(i) より

$$M < +\infty \text{ かつ } \|\varphi(x) - \varphi(c)\| \le M\|x - c\| \le M \operatorname{diam} J.$$

よって各 $x \in J$ と $i = 1, \dots, d - 1$ に対して

$$-M \operatorname{diam} J \le \mathrm{proj}_i\{L(\varphi(x) - \varphi(c))\} \le M \operatorname{diam} J.$$

ここで $|\mathrm{proj}_i Lu| \le \|Lu\| = \|u\|\ \forall i = 1, \dots, d\ \forall u \in \mathbb{R}^d$ を使った。d 座標については

$$\mathrm{proj}_d\{L(\varphi(x) - \varphi(c))\} \le \mathrm{proj}_d L\varphi'(c)(x - c) + \|\varphi(x) - \varphi(c) - \varphi'(c)(x - c)\|$$

と評価する。右辺第 1 項は 0 である。右辺第 2 項に補題 21.2(ii) を適用して

$$-\varepsilon \operatorname{diam} J \le \mathrm{proj}_d\{L(\varphi(x) - \varphi(c))\} \le \varepsilon \operatorname{diam} J\ \forall x \in J$$

が得られる。以上より $\varphi(J)$ は $\lambda^{(d)}$ 測度が $(2M \operatorname{diam} J)^{d-1} 2\varepsilon \operatorname{diam} J$ であるような d 次元閉区間をアファイン写像 $x \mapsto L^{-1}x + \varphi(c)$ で写した像に含まれることになる。よって

$$\lambda^{(d)\star}(\varphi(J)) \le (2M \operatorname{diam} J)^{d-1} 2\varepsilon \operatorname{diam} J = \varepsilon M^{d-1}(2 \operatorname{diam} I/2^n)^d$$

と押さえられる。$\#\Delta = (2^n)^d$ であることを考慮して $(\star)$ を適用する。

$$\lambda^{(d)\star}(\varphi(C \cap I)) \le \varepsilon M^{d-1}(2 \operatorname{diam} I/2^n)^d \#\Delta = \varepsilon M^{d-1}(2 \operatorname{diam} I)^d$$

以上より結論に到達した。 □

21.5 演習問題. (i) $A \subset D,\ K \subset D$ とする。$D \setminus A$ が $\mathbb{R}^d$ の開集合であり K が $\mathbb{R}^d$ の閉集合であるなら $A \cap K$ は $\mathbb{R}^d$ の閉集合であることを示せ。
(ii) $\mathbb{R}^d$ の有界閉集合 B 上の連続写像 $\psi : B \to \mathbb{R}^k$ に対してその像 $\psi(B)$ は $\mathbb{R}^k$ の有界閉集合であることを示せ。ヒント：Bolzano-Weierstrass の定理

21.6 補題. A を D の部分集合、また ψ を連続写像 $D \to \mathbb{R}^k$ とする。$D \setminus A$ が $\mathbb{R}^d$ の開集合であるなら A の像 $\psi(A)$ は $\mathbb{R}^k$ の Borel 集合である。

証明. 両端が有理数であるような閉区間の直積で表現できる集合 K であって $K \subset D$ を満たすもの全体を $\mathcal{K}$ とする。演習問題 5.3 と同様な考察により $\bigcup_{K \in \mathcal{K}} K = D$ が導けるので

$$A = \bigcup_{K \in \mathcal{K}} (A \cap K)$$

が成り立つ。集合族 $\mathcal{K}$ は可算である。演習問題 21.5(i) で確認したことから各 $K \in \mathcal{K}$ に対して $A \cap K$ は $\mathbb{R}^d$ の有界閉集合であり、演習問題 21.5(ii) で確認したことから $\psi(A \cap K)$ は $\mathbb{R}^k$ の有界閉集合である。閉集合は Borel 集合であるから可算合併 $\bigcup_{K \in \mathcal{K}} \psi(A \cap K) = \psi(A)$ は $\mathbb{R}^k$ の Borel 集合である。 □

21.7 系. (i) φ の正則点集合 $D \setminus \mathrm{crp}(\varphi)$ は $\mathbb{R}^d$ の開集合である。
(ii) φ の像 $\varphi(D)$ は $\mathbb{R}^d$ の Borel 集合である。また φ の臨界値集合 $\varphi(\mathrm{crp}(\varphi))$ もそうである。

証明. (i) $\mathrm{rank}\, \varphi'(x) = d \Leftrightarrow \det \varphi'(x) \neq 0$ であるから $D \setminus \mathrm{crp}(\varphi) = \{x \in D : \det \varphi'(x) \neq 0\}$ と表される。関数 $x \mapsto \det \varphi'(x)$ の連続性により $D \setminus \mathrm{crp}(\varphi)$ は開集合である。
(ii) $A = D$ として補題 21.6 が適用できる。また (i) により $A = \mathrm{crp}(\varphi)$ として補題 21.6 が適用できる。 □

21.8 定理. φ の臨界値集合 $\varphi(\mathrm{crp}(\varphi))$ は $\lambda^{(d)}$ 測度 0 の Borel 集合である。

証明. 系 21.7(ii) で述べたように $\varphi(\mathrm{crp}(\varphi))$ は Borel 集合である。両端が有理数であるような左半開区間の直積で表現できる集合 I であって $\overline{I} \subset D$ を満たすもの全体を $\mathcal{I}_0$ とする。既に何回か利用した関係 $\bigcup_{I \in \mathcal{I}_0} I = D$ を再び使って次が成り立つことが分かる。

$$\bigcup_{I \in \mathcal{I}_0} \varphi(\mathrm{crp}(\varphi) \cap I) = \varphi(\mathrm{crp}(\varphi))$$

補題 21.4 によれば各 $I \in \mathcal{I}_0$ に対して $\varphi(\mathrm{crp}(\varphi) \cap I)$ は $\lambda^{(d)}$ 零集合であり、零集合の可算合併は再び零集合であるから $\varphi(\mathrm{crp}(\varphi))$ は $\lambda^{(d)}$ 零集合である。 □

定理 21.8 は *Sard* **の定理**と呼ばれる。臨界値集合は Lebesgue 測度に関しては無視されるような集合である。しかしながら臨界点集合 $\mathrm{crp}(\varphi)$ が無視できるとは限らない。

21.9 例. 定数値写像では定義域全体が臨界点集合である。

ひとまず正則点集合 $D \setminus \mathrm{crp}(\varphi)$ 上で考察を進めよう。

記号

$\mathrm{rgp}(\varphi)$ を φ の正則点集合、即ち $\mathrm{rgp}(\varphi) := \{x \in D : \mathrm{rank}\, \varphi'(x) = d\}$ とする。

21.10 注意. 補題 21.7(i) で述べたように正則点集合 $\mathrm{rgp}(\varphi)$ は $\mathbb{R}^d$ の開集合であり、また局所逆写像定理によれば各 $x \in \mathrm{rgp}(\varphi)$ に対して次が成り立つような近傍 $U \subset \mathrm{rgp}(\varphi)$ が存在する。

$\varphi(U)$ は $\mathbb{R}^d$ の開集合、φ の U への制限は U から $\varphi(U)$ への C^1 級可微分同相写像。

なお上の例でも述べたように $\mathrm{rgp}(\varphi) = \emptyset$ が起こりうる。

局所可微分同相写像に関して**変数変換公式**は以下のように述べられる。

21.11 定理. (i) 任意の開集合 U に対して $U \subset \mathrm{rgp}(\varphi)$ なら $\varphi(U)$ も開集合である。
(ii) 各 $y \in \mathbb{R}^d$ に対して $\varphi^{-1}(\{y\}) \cap \mathrm{rgp}(\varphi)$ は可算集合である。
(iii) 任意の非負値 Borel 可測関数 $\rho : \mathrm{rgp}(\varphi) \to \overline{\mathbb{R}}$ に対して関数

$$\varphi_*\rho : \mathbb{R}^d \to \overline{\mathbb{R}},\ y \mapsto \sum_{x \in \varphi^{-1}(\{y\}) \cap \mathrm{rgp}(\varphi)} \frac{\rho(x)}{|\det \varphi'(x)|}$$

（$y \notin \varphi(\mathrm{rgp}(\varphi))$ なら定義より $\varphi_*\rho(y) = 0$ である）は Borel 可測でありかつ次が成り立つ。

$$\int_{\varphi^{-1}(A) \cap \mathrm{rgp}(\varphi)} \rho\, \lambda^{(d)} = \int_A \varphi_*\rho\, \lambda^{(d)}\ \forall A \in \mathrm{Borel}(\mathbb{R}^d).$$

(iv) $\rho : \mathrm{rgp}(\varphi) \to \overline{\mathbb{R}}$ を密度関数に持つ絶対連続な $(\mathrm{rgp}(\varphi), \mathrm{Borel}(\mathrm{rgp}(\varphi)))$ 上の測度を μ とする。その像測度 $\varphi_*\mu$ は絶対連続で $\varphi_*\rho$ を密度関数に持つ。

証明. 記述を簡略化するために φ の定義域は $\mathrm{rgp}(\varphi)$ に制限されているとする。従って本来 $\mathrm{rgp}(\varphi)$ であるべきところを D と書くことにする。$\mathcal{C}$ を両端が有理数であるような開区間の直積で表現できる d 次元開区間の全体の族とする。次のような部分族 $\mathcal{O}$ を導入する

$$\mathcal{O} := \{I \in \mathcal{C} : \text{注意 21.10 の条件を満たす}\}.$$

まず $\mathcal{C}$ は可算族であるから $\mathcal{O}$ もそうである。$\mathcal{O}$ の番号付けを $\{I_n\,; n \in \mathbb{N}\}$ としよう。任意の $x \in D$ に対して注意 21.10 の条件を満たす近傍 V が存在する。さて練習問題 5.3 で見たように、$I \in \mathcal{C}$ を $x \in I \subset V$ となるように選べる。そのような I は注意 21.10 の条件を満たす、すなわち $I \in \mathcal{O}$ であるので次が導けた。

$$D = \bigcup_{n=1}^{\infty} I_n.$$

以上の準備のもとで (i) から順に示していこう。次の関係が成り立つ。

$$\varphi(U) = \varphi(\bigcup_{n=1}^{\infty}(U \cap I_n)) = \bigcup_{n=1}^{\infty} \varphi(U \cap I_n)$$

注意 21.10 の条件より各 $\varphi(U \cap I_n)$ は開集合である。従って $\varphi(U)$ も開集合である。つぎに (ii) を検討するのだが $\{I_n\,; n \in \mathbb{N}\}$ は非交差とは限らないので以下のようにおく。

$$E_n := I_n \cap \Big(\bigcup_{k=1}^{n-1} I_k\Big)^{\mathrm{c}}.$$

各 E_n は $E_n \subset I_n$ なる Borel 集合でありかつ族 $\{E_n\,; n \in \mathbb{N}\}$ は D の可算分割である。$y \in \varphi(D)$ を固定すると注意 21.10 の条件より各 $n \in \mathbb{N}$ に対して $\varphi^{-1}(\{y\}) \cap E_n$ はたかだか１点からなるので $\varphi^{-1}(\{y\})$ は可算である。いよいよ (iii) に取りかかろう。

$$\int_{\varphi^{-1}(A)} \rho\,\lambda^{(d)} = \sum_{n=1}^{\infty} \int_{\varphi^{-1}(A)\cap E_n} \rho\,\lambda^{(d)} = \sum_{n=1}^{\infty} \int_{\varphi_n^{-1}(A\cap\varphi(E_n))} \rho\,\lambda^{(d)}.$$

ここで φ_n は φ の I_n への制限である。φ_n に対しては系 20.18 が適用できるので右辺は

$$\sum_{n=1}^{\infty} \int_{A\cap\varphi(E_n)} \frac{\rho(\varphi_n^{-1}(y))}{|\det\varphi'(\varphi_n^{-1}(y))|}\,\lambda^{(d)}(dy) = \int_A \sum_{n=1}^{\infty} \frac{1_{\varphi(E_n)}(y)\rho(\varphi_n^{-1}(y))}{|\det\varphi'(\varphi_n^{-1}(y))|}\,\lambda^{(d)}(dy)$$

と変形できる。右辺の被積分関数を取り出そう。これは次のようにかける。

$$\sum_{n=1}^{\infty} \frac{1_{\varphi(E_n)}(y)\rho(\varphi_n^{-1}(y))}{|\det\varphi'(\varphi_n^{-1}(y))|} = \sum_{n=1}^{\infty} \sum_{x\in\varphi^{-1}(\{y\})\cap E_n} \frac{\rho(x)}{|\det\varphi'(x)|} = \sum_{x\in\varphi^{-1}(\{y\})} \frac{\rho(x)}{|\det\varphi'(x)|}.$$

２番目の等号が成り立つのは族 $\{E_n\,; n \in \mathbb{N}\}$ が D の可算分割を成すからである。さて左辺は非負値 Borel 可測関数の級数和であるから変数 y について Borel 可測である。従って関数

$$\varphi_*\rho : y \mapsto \sum_{x\in\varphi^{-1}(\{y\})} \frac{\rho(x)}{|\det\varphi'(x)|}$$

も Borel 可測でありかつ (iii) にあげた等式が成り立つ。(iv) は (iii) の言い換えである。 □

定理 21.11 と次の定理との関係は定理 8.10 と定理 8.12 の間のそれと同じである。

21.12 定理. 関数 $f : \mathrm{rgp}(\varphi) \to \overline{\mathbb{R}}$ は Borel 可測かつ $\lambda^{(d)}$ 可積分であるとする。このとき

$$A := \{y \in \mathbb{R}^d : \varphi_*|f|(y) < +\infty\} \in \mathrm{Borel}(\mathbb{R}^d)$$

は $\lambda^{(d)}$-a.e. 集合であり、かつ次を満たすような Borel 可測関数 $g : \mathbb{R}^d \to \overline{\mathbb{R}}$ が存在する。

$$g(y) = \sum_{x\in\varphi^{-1}(\{y\})\cap\mathrm{rgp}(\varphi)} \frac{f(x)}{|\det\varphi'(x)|} \quad \lambda^{(d)}\text{-a.e. } y \in A.$$

そのような Borel 可測関数 g は $\lambda^{(d)}$ 可積分でありかつ次が成り立つ。

$$\int_{\mathrm{rgp}(\varphi)} f\,\lambda^{(d)} = \int_{\mathbb{R}^d} g\,\lambda^{(d)}.$$

21.13 定義. φ の**正則値**(regular value) とは φ の臨界値でない $\mathbb{R}^d$ の要素をいう。

記号

rgv(φ) を φ の正則値集合、即ち $\mathrm{rgv}(\varphi) := \mathbb{R}^d \setminus \varphi(\mathrm{crp}(\varphi))$ とする。

21.14 注意. 誤解しやすいが $\mathrm{rgv}(\varphi)=\varphi(\mathrm{rgp}(\varphi))$ とは限らない。$\varphi(\mathrm{rgp}(\varphi))\cap\varphi(\mathrm{crp}(\varphi))\neq\emptyset$ かもしれない。また $\varphi(D)$ に属さない点も正則値であるから

$$\mathrm{rgv}(\varphi)=\{\varphi(\mathrm{rgp}(\varphi))\setminus\varphi(\mathrm{crp}(\varphi))\}\cup\{\mathbb{R}^d\setminus\varphi(D)\}$$

という関係である。系 21.7(ii) で述べたように $\varphi(\mathrm{crp}(\varphi))$ は $\mathbb{R}^d$ の Borel 集合であるから $\mathrm{rgv}(\varphi)$ もそうである。さて $y\in\mathrm{rgv}(\varphi)$ なら $\varphi^{-1}(\{y\})\subset\mathrm{rgp}(\varphi)$ であり、従って定理 21.11(ii) より $\varphi^{-1}(\{y\})$ は可算集合である。

次は**面積公式**(area formula) と呼ばれる。

21.15 定理. 任意の非負値 Borel 可測関数 $f:D\to\overline{\mathbb{R}}$ に対して関数

$$\varphi_! f:\mathbb{R}^d\to\overline{\mathbb{R}},\ y\mapsto\begin{cases}\displaystyle\sum_{x\in\varphi^{-1}(\{y\})}f(x) & y\in\mathrm{rgv}(\varphi)\\ 0 & y\in\varphi(\mathrm{crp}(\varphi))\end{cases}$$

は Borel 可測でありかつ次が成り立つ。

$$\int_D f|\det\varphi'|\,\lambda^{(d)}=\int_{\mathbb{R}^d}\varphi_! f\,\lambda^{(d)}.$$

証明. 定理 21.11(iii) を次の関数に対して適用する。

$$\rho:\mathrm{rgp}(\varphi)\to\overline{\mathbb{R}},\ x\mapsto f(x)|\det\varphi'(x)|.$$

定理 21.8 によれば $\mathrm{rgv}(\varphi)$ は $\lambda^{(d)}$-a.e. 集合であり、また $\mathrm{rgv}(\varphi)$ 上では $\varphi_*\rho=\varphi_! f$ であるから

$$\int_{\mathrm{rgp}(\varphi)}f|\det\varphi'|\,\lambda^{(d)}=\int_{\mathrm{rgp}(\varphi)}\rho\,\lambda^{(d)}=\int_{\mathbb{R}^d}\varphi_*\rho\,\lambda^{(d)}=\int_{\mathbb{R}^d}\varphi_! f\,\lambda^{(d)}$$

が成り立つ。$D\setminus\mathrm{rgp}(\varphi)$ 上では $\det\varphi'=0$ であるから求める公式に到達した。 □

21.16 定理. 関数 $f:D\to\overline{\mathbb{R}}$ は Borel 可測かつ $f\det\varphi'$ は $\lambda^{(d)}$ 可積分であるとする。このとき

$$A:=\{y\in\mathbb{R}^d:\varphi_!|f|(y)<+\infty, y\in\mathrm{rgv}(\varphi)\}\in\mathrm{Borel}(\mathbb{R}^d)$$

は $\lambda^{(d)}$-a.e. 集合であり、かつ次を満たすような Borel 可測関数 $g:\mathbb{R}^d\to\overline{\mathbb{R}}$ が存在する。

$$g(y)=\sum_{x\in\varphi^{-1}(\{y\})}f(x)\frac{\det\varphi'(x)}{|\det\varphi'(x)|}\ \ \lambda^{(d)}\text{-a.e. }y\in A.$$

そのような Borel 可測関数 g は $\lambda^{(d)}$ 可積分でありかつ次が成り立つ。

$$\int_D f\det\varphi'\,\lambda^{(d)}=\int_{\mathbb{R}^d}g\,\lambda^{(d)}.$$

記号

次の関数 $\mathbb{R}^d \to \mathbb{R}$ を $\deg(\varphi,\cdot)$ と表記する。

$$y \mapsto \begin{cases} \displaystyle\sum_{x\in\varphi^{-1}(\{y\})} \frac{\det\varphi'(x)}{|\det\varphi'(x)|} & y \in \mathrm{rgv}(\varphi), \#\varphi^{-1}(\{y\}) < +\infty \\ 0 & \text{otherwise} \end{cases}$$

21.17 系. (i) 関数 $\deg(\varphi,\cdot)$ は Borel 可測である。
(ii) Borel 可測関数 $\rho : \mathbb{R}^d \to \overline{\mathbb{R}}$ は $\rho > 0$ $\lambda^{(d)}$-a.e. をみたし関数 $\rho\circ\varphi\ \det\varphi'$ は $\lambda^{(d)}$ 可積分であるとする。このとき次は $\lambda^{(d)}$-a.e. 集合である。

$$\{y \in \mathbb{R}^d : \#\varphi^{-1}(\{y\}) < +\infty, y \in \mathrm{rgv}(\varphi)\} \in \mathrm{Borel}(\mathbb{R}^d).$$

更に関数 $\rho\deg(\varphi,\cdot)$ は $\lambda^{(d)}$ 可積分であり次が成り立つ。

$$\int_D \rho\circ\varphi\ \det\varphi'\,\lambda^{(d)} = \int_{\mathbb{R}^d} \rho\deg(\varphi,\cdot)\,\lambda^{(d)}.$$

21.18 例. $K = [0,1]\times[0,1]$ 上の C^2 級写像 $\phi : K \to \mathbb{R}^2$ が次の条件を満たすとする。

$$\phi(1,t) - \phi(0,t) \in \mathbb{Z}^2\ \forall t \in [0,1],\ \ \phi(s,1) - \phi(s,0) \in \mathbb{Z}^2\ \forall s \in [0,1]$$

これを $D = (0,1)\times(0,1)$ に制限した写像 φ に対して関数 $\deg(\varphi,\cdot)$ は $\lambda^{(2)}$ 可積分であり

$$\int_{\mathbb{R}^2} \deg(\varphi,\cdot)\,\lambda^{(2)} = \det\begin{pmatrix}\phi(1,0)-\phi(0,0) & \phi(0,1)-\phi(0,0)\end{pmatrix}$$

が成り立つ。条件より右辺は整数値であるがこれは注目すべきである。

証明. 系 21.17 によれば、積分 $\int_D \det\varphi'\,\lambda^{(2)}$ を評価すればよい。関係

$$\det\varphi' = \partial_1(\phi_1\partial_2\phi_2) - \partial_2(\phi_1\partial_1\phi_2)$$

に着目する。ここで ∂_1, ∂_2 はそれぞれ第 1 座標、第 2 座標についての偏微分である。Fubini の定理を適用することにより

$$\begin{aligned}\int_D \det\varphi'\,\lambda^{(2)} &= \int_{(0,1)}\Big(\int_{(0,1)} \partial_1(\phi_1\partial_2\phi_2)(x,y)\lambda(dx)\Big)\lambda(dy) \\ &\qquad - \int_{(0,1)}\Big(\int_{(0,1)} \partial_2(\phi_1\partial_1\phi_2)(x,y)\lambda(dy)\Big)\lambda(dx) \\ &= \int_{(0,1)} (\phi_1(1,y)\partial_2\phi_2(1,y) - \phi_1(0,y)\partial_2\phi_2(0,y))\lambda(dy) \\ &\qquad - \int_{(0,1)} (\phi_1(x,1)\partial_1\phi_2(x,1) - \phi_1(x,1)\partial_1\phi_2(x,1))\lambda(dx) \\ &= I - J.\end{aligned}$$

さて右辺第 1 項を次のように変形する。

$$I = \int_{(0,1)} \{\phi_1(1,y)\big(\partial_2\phi_2(1,y) - \partial_2\phi_2(0,y)\big) + \big(\phi_1(1,y) - \phi_1(0,y)\big)\partial_2\phi_2(0,y)\}\lambda(dy).$$

ここでまだ手つかずの条件

$$\phi(1,t) - \phi(0,t) \in \mathbb{Z}^2 \ \forall t \in [0,1]$$

を活用する。区間 $[0,1]$ 上の連続関数が値として整数しか許さないということは中間値の定理から定数以外にはあり得ない。よって

$$\phi_1(1,t) - \phi_1(0,t) = \phi_1(1,0) - \phi_1(0,0) \text{ かつ } \partial_2\phi_2(1,t) - \partial_2\phi_2(0,t) = 0 \ \forall t \in [0,1].$$

従って

$$I = \int_{(0,1)} \big(\phi_1(1,0) - \phi_1(0,0)\big)\partial_2\phi_2(0,y)\lambda(dy) = (\phi_1(1,0) - \phi_1(0,0))(\phi_2(0,1) - \phi_2(0,0)).$$

同様に条件 $\phi(s,1) - \phi(s,0) \in \mathbb{Z}^2 \ \forall s \in [0,1]$ を活用して

$$J = (\phi_1(0,1) - \phi_1(0,0))(\phi_2(1,0) - \phi_2(0,0)).$$

以上より $I - J = \det\begin{pmatrix}\phi(1,0) - \phi(0,0) & \phi(0,1) - \phi(0,0)\end{pmatrix}$ が導かれた。 $\square$

索 引