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

2004 年 10 月 9 日

目 次

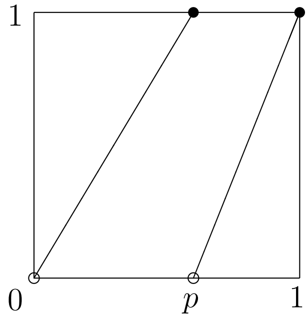
1　導入–あるモデル

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。

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

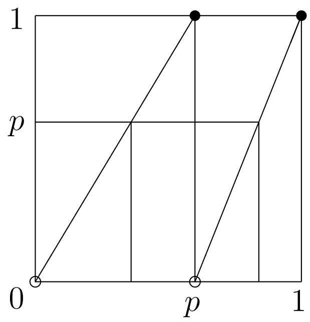


写像 φ のグラフ

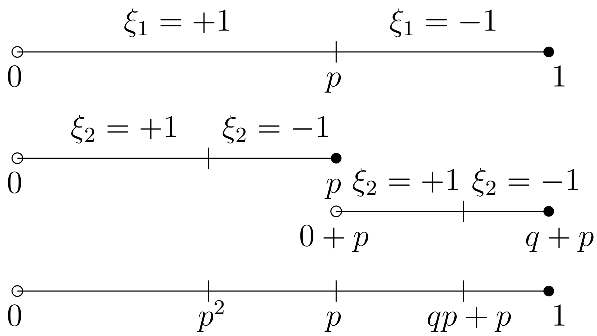
次に関数 $(0,1] \to \mathbb{R}$ の列 ξ_k $k \in \mathbb{N}$ を下のように定義する。

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

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

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

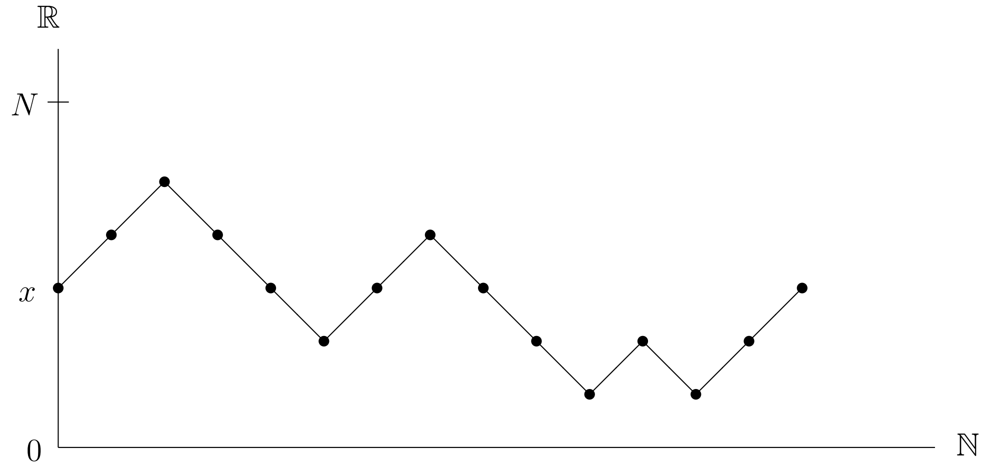
以上から各 $k \in \mathbb{N}$ に対して ξ_k が値 1 を返すような分割区分のそれぞれの長さの合計は p であり ξ_k が値 -1 を返すような分割区分のそれぞれの長さの合計は $1-p$ であることが予想できるであろう。実際、次が成り立つが、その厳密な証明は後で行う。

> $\varepsilon(1), \varepsilon(2), \ldots, \varepsilon(n)$ を 1 または -1 からなる有限数列とする。このとき $\xi_k(\omega) = \varepsilon(k)$ $\forall k \in \{1, 2, \ldots, n\}$ 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。

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

$x \in \mathbb{R}$. $\omega \in (0,1]$ 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. 三つ組み $(\Omega,\mathcal{F},P)$ 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、$(\Omega,\mathcal{F},P)$ は確率空間を表す記号として固定される。

復習

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

(i) $\emptyset\in\mathcal{B}$
(ii) $A\in\mathcal{B}\Rightarrow A^{c}\in\mathcal{B}$ （A^{c} は S における A の補集合を表す。）
(iii) $A_n\in\mathcal{B}\ \forall n\in\mathbb{N}\Rightarrow\bigcup_{n=1}^{\infty}A_n\in\mathcal{B}$

さらに関数 $\mu:\mathcal{B}\to\overline{\mathbb{R}}$ が次の条件を満たすとき、**測度**(measure) であるという。

(i) $\mu(A)\geq 0\ \forall A\in\mathcal{B}$, $\mu(\emptyset)=0$
(ii) $A_n\in\mathcal{B}\ \forall n\in\mathbb{N}$, $A_n\cap A_m=\emptyset\ n\neq m\Rightarrow\mu(\bigcup_{n=1}^{\infty}A_n)=\sum_{n=1}^{\infty}\mu(A_n)$

三つ組み $(S,\mathcal{B},\mu)$ 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が、多くの場合 Ω あるいは $\mathbb{R}^d$ の部分集合を指している。確率空間の最も重要な例としては $((0,1],\mathrm{Borel}((0,1]),\lambda)$ がある、ただし $\mathrm{Borel}((0,1])$ は $(0,1]$ 上の Borel 集合族、λ は Lebesgue 測度である。

2.2 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

関数 $f:S\to\overline{\mathbb{R}}$ が $\mathcal{B}$ **可測**(measurable) とは $\{s\in S:f(s)>a\}\in\mathcal{B}\ \forall a\in\mathbb{R}$ であることをいう。そのような関数を $\mathcal{B}$ **可測関数**(measurable function) と呼ぶ。

次の命題を思い出しておこう。

2.3 補題. $f:S\to\overline{\mathbb{R}}$ を $\mathrm{Image}\,f$ 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。

2.4 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、$A^{\mathrm{c}}\in\mathcal{B}_\alpha$ である。これが任意の α について成り立つので $A^{\mathrm{c}}\in\bigcap_\alpha\mathcal{B}_\alpha$ 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。 □

記号

Sbset(S) 集合 S の部分集合全体の族

2.5 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、Sbset(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 2.4 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

S の部分集合の族 $\mathcal{A}$ に対し系 2.5 で規定される S 上の σ-加法族を $\sigma(\mathcal{A})$ と表記する。

2.6 定義. (i) S の部分集合の族 $\mathcal{A}$ に対し $\sigma(\mathcal{A})$ 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。

2.7 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。

2.8 補題. (i) $\{(a,+\infty)\,;a\in\mathbb{R}\}$ で生成される $\mathbb{R}$ 上の σ-加法族は $\mathrm{Borel}(\mathbb{R})$ 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 $\mathrm{Borel}(\mathbb{R})$ 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 $\mathcal{B}$ は上の演習問題と同じ記号とする。開区間は開集合であるから、演習問題 2.7(ii) を適用して $\mathrm{Borel}(\mathbb{R})=\mathcal{B}\subset\sigma(\mathcal{O})$ が導かれる。次に演習問題 2.7(iii) 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 $\mathrm{Borel}(\mathbb{R})$ 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。 □

2.9 演習問題. 連続関数 $f:\mathbb{R}\to\mathbb{R}$ は $\mathrm{Borel}(\mathbb{R})$ 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。

2.10 補題. 関数 $f:S\to\mathbb{R}$ が与えられたとき、対応 $\mathrm{Sbset}(\mathbb{R})\to\mathrm{Sbset}(S)$, $B\mapsto f^{-1}(B)$ は σ-加法族としての準同型写像である。すなわち次が成り立つ。

$$f^{-1}(\emptyset)=\emptyset,\, f^{-1}(B^{\mathrm{c}})=(f^{-1}(B))^{\mathrm{c}},\, f^{-1}(\bigcup_{n=1}^{\infty}B_n)=\bigcup_{n=1}^{\infty}f^{-1}(B_n).$$

2.11 演習問題. 補題 2.10 を示せ。

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。

2.12 補題. (i) 関数 $f:S\to\mathbb{R}$ に対し $\sigma\{f\}$ は S 上の σ 加法族である。
(ii) S 上の σ 加法族 $\mathcal{B}$ に対し $\{B\in\mathrm{Sbset}(\mathbb{R}):f^{-1}(B)\in\mathcal{B}\}$ は $\mathbb{R}$ 上の σ 加法族である。
(iii) $\sigma\{f\}=\sigma(\{f^{-1}((a,+\infty))\,;a\in\mathbb{R}\})$.

証明. (i) 証明の全般にわたって補題 2.10 を適用するのでその使いどころをみてほしい。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、$B^{\mathrm{c}}\in\mathrm{Borel}(\mathbb{R})$ により、

$$A^{\mathrm{c}}=(f^{-1}(B))^{\mathrm{c}}=f^{-1}(B^{\mathrm{c}})\in\sigma\{f\}$$

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

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

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 2.8(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、

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

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。 □

2.13 定義. 関数 $f:S\to\mathbb{R}$ に対し $\sigma\{f\}$ を f によって**生成される** σ **加法族**あるいは f を**可測にする最小の** σ **加法族**という。

2.14 演習問題. 補題 2.12(ii) を示せ。

2.15 演習問題. 包含写像 $\iota:(a,b]\to\mathbb{R}$ に対し $\sigma\{\iota\}=\{B\subset(a,b]:B\in\mathrm{Borel}(\mathbb{R})\}$ を示せ。

記号

$a,b\in\mathbb{R},\ a<b$ に対して集合族 $\{B\subset(a,b]:B\in\mathrm{Borel}(\mathbb{R})\}$ を $\mathrm{Borel}((a,b])$ と書く。これは補題 2.12(i) と演習問題 2.15 の結論により半開区間 $(a,b]$ 上の σ 加法族である。

2.16 例. 1 節で述べた関数 $\xi_1:(0,1]\to\mathbb{R}$ は補題 2.3 により $\mathrm{Borel}((0,1])$ 可測である。

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. 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

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

を示すこととなる。さて $\mathcal{B}$ は σ 加法族であるから左の条件は $\sigma(\{f^{-1}((a,+\infty))\,;a\in\mathbb{R}\})\subset\mathcal{B}$ と同値である。ところが補題 2.12(iii) によれば、まさに $\sigma(\{f^{-1}((a,+\infty))\,;a\in\mathbb{R}\})=\sigma\{f\}$ であるから同値性が導かれた。 □

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。 □

2.19 例. $\alpha\in\mathbb{R}_{>0}$, $\beta\in\mathbb{R}$ に対して 1 次関数 $f:\mathbb{R}\to\mathbb{R},\ x\mapsto\alpha x+\beta$ 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 2.9）。よって定理 2.17 を適用して

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

が導かれる。当然ながら $\alpha<0$ であってもこれは成り立つ。

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。実現されたものを分布という。系 2.18 により確率変数 X について $X^{-1}(B) \in \mathcal{F}\ \forall B \in \mathrm{Borel}(\mathbb{R})$ であることに注目しよう。

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。
右辺を $P(X \in B)$ と表現することも多い。

3.1 演習問題. X を確率変数とする。このときつぎを示せ。
(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. 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 $(\mathbb{R}, \mathrm{Borel}(\mathbb{R}))$ 上の確率測度であるが、このときは特に X の**確率分布**(probability distribution) あるいは**確率法則**(propability law) という。逆に $(\mathbb{R}, \mathrm{Borel}(\mathbb{R}))$ 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。

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 Lebesgue 積分を使って表現される。積分を復習しておこう。S 上の σ 加法族 $\mathcal{B}$ に対し次の条件を満たす関数 $f : S \to \overline{\mathbb{R}}$ を $\mathcal{B}$-**単関数**(simple function) と呼ぶ。

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

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 $\int_S f\,\mu = \sum_{y\in \mathrm{Image}\, f} y\mu(f^{-1}(\{y\}))$ である。

復習の続き

一般に $f: S \to \overline{\mathbb{R}}$ を $\mathcal{B}$-可測関数とする。f が μ-**可積分**(integrable) であるとは

$$\int_S |f|\,\mu < +\infty.$$

が成り立つことをいう。このとき f の μ についての**積分**を次で定義する。

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

記号

S の部分集合 A に対しその**定義関数**(indicator function) を 1_A と表記する。

$$s\in A \Rightarrow 1_A(s) = 1,\ s\notin A \Rightarrow 1_A(s) = 0$$

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

3.5 補題. $\mathcal{B}$ を S 上の σ 加法族とする。ここでは $(S,\mathcal{B})$ 上の測度を単に測度という。
(i) 各 $s\in S$ に対して関数 $\mathcal{B}\to\overline{\mathbb{R}},\ A\mapsto 1_A(s)$ は測度である。この測度を μ と書くと非負値 $\mathcal{B}$ 可測関数 $f: S\to\overline{\mathbb{R}}$ に対して $\int_S f\,\mu = f(s)$ である。
(ii) 測度の列 μ_n $n\in\mathbb{N}$ と非負実数列 α_n $n\in\mathbb{N}$ に対して関数 $\mathcal{B}\to\overline{\mathbb{R}},\ A\mapsto\sum_{n=1}^{\infty}\alpha_n\mu_n(A)$ は測度である。この測度を μ と書くと非負値 $\mathcal{B}$ 可測関数 $f: S\to\overline{\mathbb{R}}$ に対して

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

3.6 演習問題. 補題 3.5 を示せ。

記号

$\mathcal{B}$ を S 上の σ 加法族とする。各 $s\in S$ に対して測度 $\mathcal{B}\to\overline{\mathbb{R}},\ A\mapsto 1_A(s)$ を s に質量 (mass) を持つ S 上の *Dirac* **測度**といい、δ_s と表記する。

3.7 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 $\sum_{k=0}^{\infty}(e^{-c}c^k/k!)\delta_k$ を強度 (intensity)c の *Poisson* **分布**という。これを μ と書くと補題 3.5 により

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

3.8 例. 与えられたパラメータ $n\in\mathbb{N},\ p\in(0,1)$ に対して測度 $\sum_{k=0}^{n}\binom{n}{k}p^k(1-p)^{n-k}\delta_k$ を**二項分布**(binomial distribution) という。これを μ と書くと補題 3.5 により

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

復習

次を満たす $(\mathbb{R}, \mathrm{Borel}(\mathbb{R}))$ 上の測度が唯一存在する。

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

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

状況によっては、完備化された σ-加法族（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）を考察の対象にする必要があるが、ここではそうしない。さて $\mathrm{Borel}(\mathbb{R})$ 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。

3.9 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。

3.10 例. 1 節の関数 $\xi_1 : (0,1] \to \mathbb{R}$ は例 2.16 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。

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

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

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。

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

証明. (i) $A_n \in \mathrm{Borel}(\mathbb{R})\ \forall n \in \mathbb{N}$, $A_n \cap A_m = \emptyset\ n \neq m$ なら

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

が成り立つことが Lebesgue 積分の主要な性質の一つである。

(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

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

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 $f_n \le f_{n+1}\ \forall n\in\mathbb{N}$, $\sup_{n\in\mathbb{N}} f_n(x) = f(x)\ \forall x\in\mathbb{R}$ を満たすものが存在する。各 $n\in\mathbb{N}$ に対して $\int_{\mathbb{R}} f_n\,\mu = \int_{\mathbb{R}} f_n\rho\,\lambda$ が成り立つので単調収束定理を適用して結論に至る。 □

3.13 定義. 定理 3.12 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。

3.14 補題. $f:\mathbb{R}\to\overline{\mathbb{R}}$ を Borel($\mathbb{R}$) 可測関数、$\alpha,\beta\in\mathbb{R}$, $\alpha\ne 0$ 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。

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

証明. $g(x) := f(\alpha x+\beta)$ とおくと $g^{-1}(A) = \alpha^{-1}(f^{-1}(A)-\beta)\ \forall A\subset\mathbb{R}$ であることが注目点である。定理 2.17 と例 2.19 によれば

$$g^{-1}(A) = \alpha^{-1}(f^{-1}(A)-\beta) \in \mathrm{Borel}(\mathbb{R})\,\forall A\in\mathrm{Borel}(\mathbb{R})$$

である。再度定理 2.17 を適用して関数 g の Borel($\mathbb{R}$) 可測性を得た。f が非負値単関数であるなら g もそうであって、$\mathrm{Image}\,g = \mathrm{Image}\,f$ および次が成り立つ。

$$(\star)\qquad \int_{\mathbb{R}} g\,\lambda = \sum_{y\in\mathrm{Image}\,g} y\lambda(g^{-1}(\{y\})) = \sum_{y\in\mathrm{Image}\,f} y\lambda(\alpha^{-1}(f^{-1}(\{y\})-\beta)).$$

ここで関数 $\mu:\mathrm{Borel}(\mathbb{R})\to\overline{\mathbb{R}},\ A\mapsto|\alpha|\lambda(\alpha^{-1}(A-\beta))$ は測度である（演習問題 3.1）。他方、

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

であるが、これは $b-a=\lambda((a,b])$ 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

$$\int_{\mathbb{R}} g\,\lambda = \sum_{y\in\mathrm{Image}\,f} y\lambda(f^{-1}(\{y\}))/|\alpha| = \frac{1}{|\alpha|}\int_{\mathbb{R}} f\,\lambda$$

が導かれた。一般の場合の議論は演習問題とする。 □

3.15 演習問題. 一般の場合に補題 3.14 を示せ。

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 $\alpha = \mathrm{Re}\,\alpha + \sqrt{-1}\,\mathrm{Im}\,\alpha$

3.16 定義. $\mathcal{B}$ を S 上の σ 加法族、μ を $(\mathcal{B}, S)$ 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 $x \mapsto \mathrm{Re}\, f(x)$, $x \mapsto \mathrm{Im}\, f(x)$ 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とりあえず ∞ を除外しておくと面倒が少ない。

3.17 注意. 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対応する実数値関数の積分に関する命題を適用すれば導ける。

3.18 演習問題. $f, g : S \to \mathbb{C}$ を $\mathcal{B}$ 可測関数とし、$a, b \in \mathbb{C}$ とする。このとき以下を示せ。
(i) 線形結合 $af + bg$ も $\mathcal{B}$ 可測である。さらに f, g ともに μ 可積分なら $af + bg$ も μ 可積分であり $\int_S (af + bg)\,\mu = a\int_S f\,\mu + b\int_S g\,\mu$ が成り立つ。
(ii) f μ 可積分 $\Leftrightarrow$ $\int_S |f|\,\mu < +\infty$. μ 可積分なら $|\int_S f\,\mu| \leq \int_S |f|\,\mu$

次は 1 次元区間上の連続関数には必ず原始関数が存在することおよび原始関数と Lebesgue 積分との関係、いわゆる**微積分の基本定理**、を述べており解析学において基幹の位置をしめるものである。

3.19 補題. $a, b \in \mathbb{R}$, $a < b$ 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 $(a, b) \to \mathbb{C}$, $x \mapsto \displaystyle\int_{(a,x]} f\,\lambda$ は微分可能であり、その導関数は f に等しい。
(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。

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 $c \leq x < b$ 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、積分における三角不等式を適用して $c \leq x < \min\{c + \delta, b\}$ なら

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

が成り立つことを得る。$\varepsilon|x-c|$ で抑えるという評価は $\max\{c-\delta, a\} < x \leq c$ 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。 $\square$

3.20 演習問題. 補題 3.19(ii) 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、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。

4.1 例. 与えられたパラメータ $a, b \in \mathbb{R}$, $a < b$ に対して測度

$$\mathrm{Borel}(\mathbb{R}) \to \overline{\mathbb{R}}, A \mapsto \lambda(A \cap (a,b))/(b-a)$$

を区間 (a,b) 上の**一様分布**(uniform distribution) という。これを μ と書くと補題 3.19(ii) により

$$\int_{\mathbb{R}} x\mu(dx) = \frac{a+b}{2}, \int_{\mathbb{R}} x^2\mu(dx) = \frac{a^2+ab+b^2}{3}, \int_{\mathbb{R}} e^{zx}\mu(dx) = \frac{e^{zb}-e^{za}}{z(b-a)}\ z \in \mathbb{C}.$$

4.2 例. 与えられた $c \in \mathbb{R}_{>0}$ に対して測度

$$\mathrm{Borel}(\mathbb{R}) \to \overline{\mathbb{R}}, A \mapsto \int_{A\cap(0,+\infty)} ce^{-cx}\,\lambda(dx)$$

をパラメータ c の**指数分布**(exponential distribution) という。これを μ と書くと定理 3.12(ii) と補題 3.19(ii) により

$$\int_{\mathbb{R}} x\mu(dx) = \frac{1}{c}, \int_{\mathbb{R}} x^2\mu(dx) = \frac{2}{c^2}, \int_{\mathbb{R}} e^{zx}\mu(dx) = \frac{c}{c-z}\ z \in \mathbb{C}, \mathrm{Re}\, z < c.$$

4.3 例. 与えられた $a \in \mathbb{R}$, $c \in \mathbb{R}_{>0}$ に対して測度

$$\mathrm{Borel}(\mathbb{R}) \to \overline{\mathbb{R}}, A \mapsto \int_A \frac{c}{2}e^{-c|x-a|}\,\lambda(dx)$$

を中心 a パラメータ c の**両側指数分布**(two sided exponential distribution) という。これを μ と書くと定理 3.12(ii) と補題 3.19(ii) により

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

4.4 例. 与えられた $a \in \mathbb{R}$, $t \in \mathbb{R}_{>0}$ に対して測度

$$\mathrm{Borel}(\mathbb{R}) \to \overline{\mathbb{R}}, A \mapsto \frac{t}{\pi}\int_A \frac{1}{t^2+(x-a)^2}\,\lambda(dx)$$

を中心 a 半値幅 t の *Cauchy* **分布**という。これを μ と書くと定理 3.12(ii) により

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

その標準的な算出技術は、複素線積分に持ち込んで留数定理を適用するものである。のちに例 8.17 で Fubini の定理を応用した計算方法を紹介する。なお Cauchy 分布については $\int_{\mathbb{R}} |x|\,\mu(dx) = +\infty$ であることに注意せよ。

4.5 演習問題. 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) $\displaystyle\int_{\mathbb{R}} p(t,x-y)p(s,y)\lambda(dy) = p(t+s,x)\ \forall t, \forall s \in \mathbb{R}_{>0}, \forall x \in \mathbb{R}$ を示せ。
(ii) 任意の有界連続関数 $f : \mathbb{R} \to \mathbb{C}$ と $a \in \mathbb{R}$ に対して $\displaystyle\int_{\mathbb{R}} p(t,x-a)f(x)\lambda(dx)$ は $t \to 0$ の極限で $f(a)$ に収束することを示せ。

4.6 例. 与えられたパラメータ $t \in \mathbb{R}_{>0}$, $a \in \mathbb{R}$ に対して次の測度を平均 a 分散 t の**正規分布**(normal distribution) あるいは *Gauss* **分布**という。

$$\mathrm{Borel}(\mathbb{R}) \to \overline{\mathbb{R}}, A \mapsto \int_A \frac{1}{\sqrt{2\pi t}} \exp\{-\frac{(x-a)^2}{2t}\}\lambda(dx)$$

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。補題 3.14 が適用できるので確率測度であるのを示すのに基本となるのは

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

である。その標準的な算出手段は、重積分に持ち込んで極座標変換を適用するものであり例 10.20 で紹介する。だが重要なのは値の特定ではなくてむしろ次を確認することである。

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

可積分性をみるにはたとえば $-x^2/2 \le -|x| + 1/2\ \forall x \in \mathbb{R}$ という類の評価式が必要になる。他方、被積分関数は a.e. で正であるから積分値は正であることが従う。

4.7 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 $e^{zx} \in \mathbb{R}_{\ge 0}\ \forall x \in \mathbb{R}$ 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 3.12(ii) により

$$\int_{\mathbb{R}} e^{zx}\mu(dx) = \exp\{za + tz^2/2\}$$

となり、$z \in \mathbb{R}$ の場合に (ii) が示された。以下、煩雑性を回避するために $a = 0$ として議論を続ける。$|e^{zx}| \le e^{|zx|} \le e^{|z|x} + e^{-|z|x}\ \forall x \in \mathbb{R}$ であるから

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

が従う。中辺の被積分関数を整級数展開すると各項は非負値であるから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、従ってその収束半径は ∞ でありさらに

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

が成り立つ。$z\in\mathbb{R}$ のとき右辺は $\exp\{tz^2/2\}$ に等しく、またそれは $\sum_{k=0}^{\infty}t^kz^{2k}/(2^kk!)$ と整級数展開される。 よって

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

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 3.14 を使って $a=0$ の場合に帰着できる。 □

4.8 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) 任意の有界連続関数 $f:\mathbb{R}\to\mathbb{C}$ と $a\in\mathbb{R}$ に対して $\displaystyle\int_{\mathbb{R}}p(t,x-a)f(x)\lambda(dx)$ は $t\to 0$ の極限で $f(a)$ に収束することを示せ。

4.9 定義. $(\mathbb{R},\mathrm{Borel}(\mathbb{R}))$ 上の確率測度 μ に対して $\mathbb{R}\to\mathbb{R},\ x\mapsto\mu((-\infty,x])$ を μ の**分布関数**(distribution function) という。また確率変数 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。

4.10 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。

4.11 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) $\lim_{x\to-\infty}\mu((-\infty,x])=0,\ \lim_{x\to+\infty}\mu((-\infty,x])=1.$

4.12 演習問題. 補題 4.11 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。一意性を適用すると直ちに次のことがわかる。

4.13 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。

4.14 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 4.13 は次のような言い換えを持つ。

4.15 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。

4.16 例. 確率変数 X が標準正規分布に従う必要十分条件は

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

である。右辺の関数をここでは**誤差関数**(error function) という（$x \mapsto \int_0^x e^{-y^2} dy$ を誤差関数と呼んでいるケースが多いので注意せよ）。

4.17 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。

4.18 演習問題. モデル $((0,1], \mathrm{Borel}((0,1]), \lambda)$ で考える。$c \in \mathbb{R}_{>0}$ とする。

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

このとき確率変数 X の分布はパラメータ c の指数分布であることを示せ。

5 確率変数と多次元確率変数

実確率変数系 X_1, X_2, ..., 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 $\mathrm{Borel}(\mathbb{R}^d)$ で表す。各 $B \in \mathrm{Borel}(\mathbb{R}^d)$ 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 2.8 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ておく。

5.2 演習問題. 連続関数 $f : \mathbb{R}^d \to \mathbb{R}$ は $\mathrm{Borel}(\mathbb{R}^d)$ 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. この問題では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を $\mathcal{C}$ とかく。A を任意に与えられた $\mathbb{R}^d$ の開集合とする。このとき $\mathcal{C}$ に属する区間で A の部分集合となるものすべての合併は A であることを示せ。

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 $\mathrm{Borel}(\mathbb{R}^d)$ 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、$\mathcal{C}$ を上の演習問題と同じ記号とすると、$\mathcal{C}$ は可算族である。したがって演習問題 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. $\mathcal{B}$ を S 上の σ 加法族とする。写像 $f : S \to \mathbb{R}^d$ は、条件 $f^{-1}(B) \in \mathcal{B}\ \forall B \in \mathrm{Borel}(\mathbb{R}^d)$ を満たすとき、$\mathcal{B}$ **可測**であるという。

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。1 次元の時と同じく集合族 $\sigma\{f\}$ は S 上の σ 加法族であることを確かめよう。

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 $\{B \in \mathrm{Sbset}(T) : f^{-1}(B) \in \mathcal{B}\}$ は T 上の σ 加法族である。

(iii) T 上の集合族 $\mathcal{C}$ に対して $\{f^{-1}(B)\,;B \in \sigma(\mathcal{C})\} = \sigma(\{f^{-1}(B)\,;B \in \mathcal{C}\})$.

(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。

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

証明. 補題 2.12 の証明と同じく補題 2.10 の適用が基本線であるが、抽象化が進んでいるので念のため省略せず証明を与えておこう。

(i) $f^*\mathcal{M} := \{f^{-1}(B)\,;B \in \mathcal{M}\}$ と書く。まず $\emptyset \in \mathcal{M}$ により $\emptyset = f^{-1}(\emptyset) \in f^*\mathcal{M}$ 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、$B^{\mathrm{c}} \in \mathcal{M}$ により、$A^{\mathrm{c}} = (f^{-1}(B))^{\mathrm{c}} = f^{-1}(B^{\mathrm{c}}) \in f^*\mathcal{M}$ が従う。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. $A_n = f^{-1}(B_n)$ であるが、$\bigcup_{n=1}^{\infty} A_n \in \mathcal{M}$ により、$\bigcup_{n=1}^{\infty} A_n = \bigcup_{n=1}^{\infty} f^{-1}(B_n) = f^{-1}(\bigcup_{n=1}^{\infty} B_n) \in f^*\mathcal{M}$ となる。

(ii) $f_*\mathcal{B} := \{B \in \mathrm{Sbset}(T) : f^{-1}(B) \in \mathcal{B}\}$ と書く。まず $f^{-1}(\emptyset) = \emptyset \in \mathcal{B}$ により $\emptyset \in f_*\mathcal{B}$ である。次に $B \in f_*\mathcal{B}$ とすると $B \subset T$ かつ $f^{-1}(B) \in \mathcal{B}$ であるが、$f^{-1}(B^{\mathrm{c}}) = (f^{-1}(B))^{\mathrm{c}} \in \mathcal{B}$ により、$B^{\mathrm{c}} \in f_*\mathcal{B}$ が従う。他方 $B_n \in f_*\mathcal{B}\ \forall n \in \mathbb{N}$ とすると $B_n \subset T$ かつ $f^{-1}(B_n) \in \mathcal{B}$ であるが、$f^{-1}(\bigcup_{n=1}^{\infty} B_n) = \bigcup_{n=1}^{\infty} f^{-1}(B_n) \in \mathcal{B}$ により、$\bigcup_{n=1}^{\infty} B_n \in f_*\mathcal{B}$ となる。

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

$$\{B\in\mathrm{Sbset}(T) : f^{-1}(B)\in\sigma(f^*\mathcal{C})\}$$

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 $\sigma(\mathcal{C})$ であるから、

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

が従う。これは $f^{-1}(B)\in\sigma(f^*\mathcal{C})\ \forall B\in\sigma(\mathcal{C})$ を意味する。さらに $f^*(\sigma(\mathcal{C}))$ の定義により次のように書き換えられる。

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

逆向きの包含関係は $f^*(\sigma(\mathcal{C}))$ が S 上の σ 加法族であることと $f^*\mathcal{C}\subset f^*(\sigma(\mathcal{C}))$ であることから導かれる。ゆえに $f^*(\sigma(\mathcal{C}))=\sigma(f^*\mathcal{C})$.

(iv) 同値性は $f^*(\sigma(\mathcal{C}))\subset\mathcal{B}\Leftrightarrow f^*\mathcal{C}\subset\mathcal{B}$ と表現できる。さて $\mathcal{B}$ は σ 加法族なので右の条件は $\sigma(f^*\mathcal{C})\subset\mathcal{B}$ と同値である。ところが (iii) によれば、まさに $\sigma(f^*\mathcal{C})=f^*(\sigma(\mathcal{C}))$ 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 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) 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、演習問題 2.15 の結果と補題 5.6(iii) により、$\mathrm{Borel}((0,1])=\sigma(\iota^*\mathcal{I})$ がわかる。さて $\iota^{-1}(A)=A\cap(0,1]$ $\forall A\subset\mathbb{R}$ であるから $\iota^*\mathcal{I}=\{(a,b]\,;0\le a<b\le 1\}$ 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。 □

5.8 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。

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

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 S 上の σ 加法族 $\sigma(\sigma\{f_1\}\cup\sigma\{f_2\}\cup\cdots\cup\sigma\{f_n\})$ を $\sigma\{f_1,f_2,\ldots,f_n\}$ 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**あるいは系 f_i $i=1,2,\ldots,n$ を**可測にする最小の** 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

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

などにより、$\sigma\{f\}$ は $g_1^*\mathcal{I} \cup g_2^*\mathcal{I} \cup \cdots \cup g_d^*\mathcal{I}$ 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。

$$\sigma(g_1^*\mathcal{I} \cup g_2^*\mathcal{I} \cup \cdots \cup g_d^*\mathcal{I}) \subset \sigma\{f\}$$

他方、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

$$\sigma\{f\} = f^*(\sigma(\mathcal{I}^d)) = \sigma(f^*\mathcal{I}^d).$$

以下、煩雑さをさけるため $d = 2$ として議論を進める。さて

$$f^*\mathcal{I}^2 = \{f^{-1}(I \times J)\,;I,J \in \mathcal{I}\} = \{g_1^{-1}(I) \cap g_2^{-1}(J)\,;I,J \in \mathcal{I}\} \subset \sigma(g_1^*\mathcal{I} \cup g_2^*\mathcal{I})$$

である。すなわち $\sigma(g_1^*\mathcal{I} \cup g_2^*\mathcal{I})$ 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が $\sigma(f^*\mathcal{I}^2)$ であるから、$\sigma(f^*\mathcal{I}^2) \subset \sigma(g_1^*\mathcal{I} \cup g_2^*\mathcal{I})$ がわかる。従って

$$\sigma\{f\} = \sigma(g_1^*\mathcal{I} \cup g_2^*\mathcal{I} \cup \cdots \cup g_d^*\mathcal{I}) = \sigma(\sigma(g_1^*\mathcal{I}) \cup \sigma(g_2^*\mathcal{I}) \cup \cdots \cup \sigma(g_d^*\mathcal{I}))$$

ここで 2 番目の等号は演習問題 5.8 の結果を適用したものである。再び補題 5.6(iii) と補題 5.4 を適用して $\sigma(g_1^*\mathcal{I}) = g_1^*(\sigma(\mathcal{I})) = \sigma\{g_1\}$ などがわかるので結論に到達した。 □

5.10 系. 写像 $f : S \to \mathbb{R}^d$ 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。

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 加法族なので右の条件は $\sigma(\sigma\{g_1\} \cup \sigma\{g_2\} \cup \cdots \cup \sigma\{g_d\}) \subset \mathcal{B}$ と同値である。ところが定理 5.9 によれば、まさに $\sigma\{f\} = \sigma\{g_1, g_2, \ldots, g_d\}$ であるから同値性が導かれた。 □

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. $\alpha \in GL(d)$, $\beta \in \mathbb{R}^d$ に対して 1 次変換 $f : \mathbb{R}^d \to \mathbb{R}^d$, $x \mapsto \alpha x + \beta$ 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

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

が導かれる。

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。系 5.10 により関数 $f : S \to \mathbb{C}$ と S 上の σ 加法族 $\mathcal{B}$ に対し、f の $\mathcal{B}$ 可測性は定義 3.16 におけるものと一致している。

5.14 定義. 確率空間 $(\Omega, \mathcal{F}, P)$ 上の $\mathbb{R}^n$ **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。

5.15 系. 写像 $X : \Omega \to \mathbb{R}^n$ について確率空間 $(\Omega, \mathcal{F}, P)$ 上の $\mathbb{R}^n$ 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。

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 $\mathbb{R}^n$ 値確率変数 X について $X^{-1}(B) \in \mathcal{F}\ \forall B \in \mathrm{Borel}(\mathbb{R}^n)$ であるから、関数 $\mathcal{L}(X, \cdot) : \mathrm{Borel}(\mathbb{R}^n) \to \overline{\mathbb{R}},\ B \mapsto P(X^{-1}(B))$ が矛盾なく定義される。それは確率測度であり、また観測量の平均などを記述する。

6.1 定義. $\mathbb{R}^n$ 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。逆に $(\mathbb{R}^n, \mathrm{Borel}(\mathbb{R}^n))$ 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**というのは以前と同じである。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。

6.2 注意. 二つの確率変数系 $X_1, X_2, \ldots, X_n$ と $Y_1, Y_2, \ldots, Y_n$ が与えられたとする。仮に $\mathcal{L}(X_i, \cdot) = \mathcal{L}(Y_i, \cdot)\ \forall i = 1, 2, \ldots, n$ 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。誤解しやすいから注意せよ。

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) $\mathcal{L}((X_1, X_2), \cdot) \neq \mathcal{L}((Y_1, Y_2), \cdot)$ を確かめよ。

6.4 補題. $\mathbb{R}^n$ 値確率変数 X, Y に対し $P(X = Y) = 1$ なら $\mathcal{L}(X, \cdot) = \mathcal{L}(Y, \cdot)$ である。

証明. $\Omega_0 := \{\omega : X(\omega) = Y(\omega)\}$ とおくと $\Omega_0 \in \mathcal{F}$ かつ $P(\Omega_0) = 1$ である。さてその補集合 Ω_0^c の任意の部分集合は測度 0 であるから

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

が成り立つ。従って $X^{-1}(A) \cap \Omega_0 = Y^{-1}(A) \cap \Omega_0$ とあわせて結論を得る。 □

6.5 演習問題. 二つの確率変数系 $X_1, X_2, \ldots, X_n$ と $Y_1, Y_2, \ldots, Y_n$ に対し $P(X_i = Y_i) = 1$ $\forall i = 1, 2, \ldots, n$ 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。

6.6 定義. $\mathcal{B}$ を S 上の σ 加法族、$\mathcal{M}$ を T 上の σ 加法族とする。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。

6.7 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。また写像 $f : \mathbb{R}^k \to \mathbb{R}^d$ については、対 $\mathrm{Borel}(\mathbb{R}^k), \mathrm{Borel}(\mathbb{R}^d)$ に関して可測であることを単に *Borel* **可測**といってしまうことが多い。

6.8 定理. $\mathcal{B}$ を S 上の σ 加法族、$\mathcal{M}$ を T 上の σ 加法族とする。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。 □

6.9 系. $X_1, X_2, \ldots, X_n$ を実確率変数とする。
(i) 関数 $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ が $\mathrm{Borel}(\mathbb{R}^n)$ 可測なら $f(X_1, X_2, \ldots, X_n) : \Omega \to \overline{\mathbb{R}}$ は確率変数である。
(ii) 写像 $\varphi : \mathbb{R}^n \to \mathbb{R}^d$ が Borel 可測、すなわち対 $\mathrm{Borel}(\mathbb{R}^n)$, $\mathrm{Borel}(\mathbb{R}^d)$ に関して可測なら合成写像 $\varphi(X_1, X_2, \ldots, X_n) : \Omega \to \mathbb{R}^d$ は $\mathbb{R}^d$ 値確率変数である。

証明. 系 5.15 と定理 6.8 による。 □

蛇足. $X : \Omega \to \mathbb{R}^n$ と $Y : \Omega \to \mathbb{R}^d$ について $\sigma\{Y\} \subset \sigma\{X\}$ 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。これを証明するのはかなり難しい。

6.10 例. 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 $\xi_1 : (0, 1] \to \mathbb{R}$ は例 2.16 により $\mathrm{Borel}((0, 1])$ 可測であり、写像 $\varphi : (0, 1] \to (0, 1]$ は例 5.7(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 6.8 を適用して結論に至る。 □

6.11 定義. $\mathcal{B}$ を S 上の σ 加法族、$\mathcal{M}$ を T 上の σ 加法族とする。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 に対して関数 $\mathcal{M} \to \overline{\mathbb{R}}$, $A \mapsto \mu(\varphi^{-1}(A))$ を φ による μ の**像測度**(image measure) といい $\varphi_*\mu$ と表記する。

6.12 定理. $\mathcal{B}$, $\mathcal{M}$, $\mathcal{W}$ をそれぞれ S, T, U 上の σ 加法族とする。写像 $\varphi : S \to T$ が対 $\mathcal{B}$, $\mathcal{M}$ に関して、写像 $\psi : T \to U$ が対 $\mathcal{M}$, $\mathcal{W}$ に関して可測なら合成写像 $\psi \circ \varphi : S \to U$ は対 $\mathcal{B}$, $\mathcal{W}$ に関して可測である。また $(S, \mathcal{B})$ 上の測度 μ に対して $(\psi \circ \varphi)_*\mu = \psi_*(\varphi_*\mu)$ である。

証明. 合成写像の可測性の証明は定理 6.8 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。 □

6.13 例. (i) 1 節で述べた写像 $\varphi : (0, 1] \to (0, 1]$ は $\varphi_*\lambda = \lambda$ を満たす。
(ii) 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。

証明. (i) Lebesgue-Stieltjes 測度の一意性により $0 \leq a < b \leq 1 \Rightarrow \lambda(\varphi^{-1}((a, b])) = b - a$ 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 6.12 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 3.10 により ξ_1 の分布は $p\delta_1 + (1 - p)\delta_{-1}$ である。 □

6.14 定義. $\mathcal{B}$ を S 上の σ 加法族、μ を $(S, \mathcal{B})$ 上の測度、$\varphi : S \to S$ を対 $\mathcal{B}, \mathcal{B}$ に関する可測写像とする。$\varphi_* \mu = \mu$ 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。

6.15 例. Lebesgue モデル上で 1 節で述べた関数 $\xi_k : (0,1] \to \mathbb{R}$, $k \in \mathbb{N}$ の結合分布は $\mathbb{N}$ の空でない有限部分集合 I に対して $\lambda(\xi_i = 1 \, \forall i \in I) = p^{\# I}$ を満たす。

証明. $\# I$ に関する帰納法により証明する。場合 $\# I = 1$ の検証は例 6.13 で済んでいる。$\# I = n$ のとき成り立つと仮定して場合 $\# I = n+1$ を考察する。$\min I = m$ とする。目の付け所は $\xi_i = \xi_{i-m+1} \circ \varphi^{m-1}$ から派生する次の関係である。

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

写像 φ は測度 λ を保存するので、定理 6.12 を適用して

$$\lambda(\xi_i = 1 \, \forall i \in I) = \lambda(\xi_{i-m+1} = 1 \, \forall i \in I).$$

従って $\min I = 1$ と見なしてよいので $I = \{1\} \cup \{i+1 \,; i \in J\}$ としよう。一般に

$$\lambda(\{\xi_1 = 1\} \cap \varphi^{-1}(A)) = p\lambda(A) \ \forall A \in \mathrm{Borel}((0,1]) \tag{6.16}$$

が成り立つ。これは例 6.13 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 $\{\xi_1 = 1\} = (0,p]$, $\varphi^{-1}((a,b]) = (pa, pb] \cup ((1-p)a + p, (1-p)b + p]$ であるから成り立つ。従って (6.16) および帰納法の仮定を適用して

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

ゆえに $\# I = n+1$ に対しても命題は成り立つ。 □

6.17 定理. 定義 6.11 での設定のもと非負値 $\mathcal{M}$ 可測関数 $f : T \to \overline{\mathbb{R}}$ に対して次が成り立つ。

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

証明. f が非負値 $\mathcal{M}$ 単関数である場合をまず考察する。このとき $f \circ \varphi$ は非負値 $\mathcal{B}$ 単関数であり、さらに $\mathrm{Image}\, f \circ \varphi \subset \mathrm{Image}\, f$ であるから次が成り立つ。

$$\int_S f \circ \varphi \, \mu = \sum_{y \in \mathrm{Image}\, f} y \mu(\varphi^{-1}(f^{-1}(\{y\}))) = \sum_{y \in \mathrm{Image}\, f} y \varphi_* \mu(f^{-1}(\{y\})) = \int_{\mathbb{R}^n} f \, \varphi_* \mu.$$

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。 □

6.18 例. 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。

6.19 定義. μ を $(S, \mathcal{B})$ 上の測度、また $p > 0$ とする。$\mathcal{B}$ 可測関数 $f : S \to \overline{\mathbb{R}}$ について $|f|^p$ が μ 可積分であるときそれを p **乗可積分**という。（複素数値関数についても同様な概念がある。）

記号

非負値あるいは P 可積分な確率変数 X に対して $E[X] := \int_\Omega X \, P$ と表記し、これを X の**期待値**(expectation) あるいは平均 (mean) という。（可積分な複素数値確率変数についても同様）また $A \in \mathcal{F}$ に対して $E[X\,;A] := \int_A X \, P$ という表記も用いる。$p > 0$ 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。

6.20 補題. μ を $(S, \mathcal{B})$ 上の測度、$f, g : S \to \mathbb{R}$ を $\mathcal{B}$ 可測関数で 2 乗可積分なものとする。
(i) 積 fg は可積分、和 $f + g$ は 2 乗可積分である。
(ii) 不等式 $\Big(\displaystyle\int_S fg \, \mu \Big)^2 \le \int_S f^2 \, \mu \int_S g^2 \, \mu$ がなりたつ。これを *Schwarz* **の不等式**という。

証明. (i) 関係 $|fg| \le (f^2 + g^2)/2$, $(f+g)^2 \le 2f^2 + 2g^2$ 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) $a, b \ge 0$ とすると $2ab \le a^2 t + b^2/t \; \forall t > 0$ が成り立つ。よって

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

他方 $a, b \ge 0$ 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。 □

6.21 注意. $p, q > 1$ 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 乗可積分であるなら和 $f + g$ は p 乗可積分、積 fh は可積分であり、

$$\Big(\int_S |f+g|^p \, \mu \Big)^{1/p} \le \Big(\int_S |f|^p \, \mu \Big)^{1/p} + \Big(\int_S |g|^p \, \mu \Big)^{1/p}, \quad \int_S fh \, \mu \le \Big(\int_S |f|^p \, \mu \Big)^{1/p} \Big(\int_S |h|^q \, \mu \Big)^{1/q}$$

が成り立つ。前者を *Minkowski* **の不等式**、後者を *Hölder* **の不等式**という。

次の命題は全測度が有限である空間でしか通用しない。

6.22 演習問題. 2 乗可積分な確率変数は可積分であることを示せ。

記号

2 乗可積分な確率変数 X, Y に対して $\mathrm{Cov}[X, Y] := E[(X - E[X])(Y - E[Y])]$ と表記し、これを X, Y の**共分散**(covariance) という。また $\mathrm{Var}[X] := \mathrm{Cov}[X, X]$ と書き、X の**分散**(variance) とよぶ。

6.23 演習問題. 2 乗可積分な確率変数 X, Y に対し $\mathrm{Cov}[X, Y]^2 \le \mathrm{Var}[X]\mathrm{Var}[Y]$ を示せ。

6.24 定理. $X_1,\ X_2,\ \ldots,\ X_n$ を実確率変数とする。このとき非負値 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。

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

証明. $X_1,\ X_2,\ \ldots,\ X_n$ 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 6.17 を適用することにより結論を得る。 □

6.25 系. (i) 確率変数 X が非負値あるいは P 可積分であるとき $E[X]=\int_{\mathbb{R}} x\,\mathcal{L}(X,dx)$
(ii) 2 乗可積分な確率変数 X,Y に対し $\mathrm{Cov}[X,Y]=\int_{\mathbb{R}^2}(x-E[X])(y-E[Y])\,\mathcal{L}((X,Y),dxdy)$

6.26 例. 平均 a 分散 t の正規分布に従う確率変数 X に対して

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

であるというのが補題 4.7(i) の主張であった。

6.27 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。

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 4.13 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 $\mu,\ \nu$ 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 Borel$(\mathbb{R}^d)$ であることおよび定理 7.11 から導かれる。 □

上にあげた定理は補題 4.13 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) $A_n\in\mathcal{D}\ \forall n\in\mathbb{N},\ A_n\subset A_{n+1}\ \forall n\in\mathbb{N}\Rightarrow\bigcup_{n=1}^{\infty}A_n\in\mathcal{D}$.

上で性質 (iii) は単調収束定理と呼応することが次の例 7.3 で判明する。

7.3 例. 集合 S 上の σ-加法族 $\mathcal{B}$ および $(S,\mathcal{B})$ 上の確率測度 $\mu,\ \nu$ が与えられたとき集合族 $\{A \in \mathcal{B} : \mu(A) = \nu(A)\}$ は Dynkin 族である。

7.4 演習問題. 例 7.3 を確認せよ。

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、例 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？

7.5 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

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

(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

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

が導かれる。よって Dynkin 族の条件がすべて確認できた。 □

補題 2.4 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。

7.6 補題. (i) 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。
(ii) 集合族 $\mathcal{A}$ に対し $\mathcal{A}$ を包含する Dynkin 族たちに最小のものが存在する。

証明. (i) 証明の実行は補題 2.4 のそれと同じすじなので演習問題とする。

(ii) 系 2.5 の証明と違うところを指摘しておこう。今度は $\mathcal{A}$ を含む Dynkin 族が存在するか明らかでない。そこで集合族 $\mathcal{A}$ に属する集合全体の合併集合 S を考える。このとき $\mathrm{Sbset}(S)$ は Dynkin 族をなし、かつ集合族 $\mathcal{A}$ を包含する。あとは系 2.5 の証明に準じればよい。 □

7.7 演習問題. 補題 7.6(i) を示せ。

7.8 定義. 集合族 $\mathcal{A}$ を含む Dynkin 族で最小のものを $\mathcal{A}$ **で生成される** *Dynkin* **族**(Dynkin system generated by $\mathcal{A}$) と呼ぶ。

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 $A \cap B \in \mathcal{C}\ \forall A \in \mathcal{C}\ \forall B \in \mathcal{C}$ が成り立っている。

7.9 補題. 集合族 $\mathcal{C}$ は次の条件を満たすとする。

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

このとき $\mathcal{C}$ で生成される Dynkin 族 $\mathcal{D}$ も $A \cap B \in \mathcal{D}\ \forall A \in \mathcal{D}\ \forall B \in \mathcal{D}$ を満たす。

証明. ここでも補題 7.6(ii) の証明と同じトリックが必要である。

集合族 $\mathcal{C}$ に属する集合全体の合併集合を S とする。

$\mathcal{D}$ は S の部分集合からなる Dynkin 族なので補題 7.5(ii) および補題 7.6(i) により

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

は Dynkin 族である。$\mathcal{C}$ に対する条件と関係 $\mathcal{C} \subset \mathcal{D}$ により

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

従って集合族 $\mathcal{D}_1$ は $\mathcal{C}$ を含む Dynkin 族である。$\mathcal{D}$ はそのようなものの中で最小であるから $\mathcal{D} \subset \mathcal{D}_1$ 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 $\mathcal{C}$ を含む。

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

しかも $\mathcal{D}_1$ に対するのと同じ根拠により集合族 $\mathcal{D}_2$ は Dynkin 族である。$\mathcal{D}$ は $\mathcal{C}$ を含む Dynkin 族の中で最小であるから $\mathcal{D} \subset \mathcal{D}_2$ すなわち $A \in \mathcal{D} \Rightarrow A \cap B \in \mathcal{D}\ \forall B \in \mathcal{D}$ を得る。 $\square$

次は *Dynkin* **族定理**と呼ばれる。

7.10 定理. S の部分集合族 $\mathcal{C}$ が条件 $A \cap B \in \mathcal{C}\ \forall A \in \mathcal{C}\ \forall B \in \mathcal{C}$ を満たすとする。
(i) 集合族 $\mathcal{C} \cup \{S\}$ で生成される Dynkin 族は S 上の σ-加法族 $\sigma(\mathcal{C})$ 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 S 上の σ-加法族 $\sigma(\mathcal{C})$ に等しい。

証明. (i) 集合族 $\tilde{\mathcal{C}} := \mathcal{C} \cup \{S\}$ も条件 $A \cap B \in \tilde{\mathcal{C}}\ \forall A \in \tilde{\mathcal{C}}\ \forall B \in \tilde{\mathcal{C}}$ を満たす。従って補題 7.9 により $\tilde{\mathcal{C}}$ で生成される Dynkin 族 $\mathcal{D}$ は条件

$$A \cap B \in \mathcal{D}\ \forall A \in \mathcal{D}\ \forall B \in \mathcal{D}$$

を満たすことになる。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 7.5(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。$\mathcal{C}$ を包含する S 上の σ-加法族で最小のものが $\sigma(\mathcal{C})$ であるから $\sigma(\mathcal{C}) \subset \mathcal{D}$ を得る。

他方、S 上の σ-加法族は Dynkin 族である。よって $\sigma(\mathcal{C})$ は $\mathcal{C} \cup \{S\}$ を含む Dynkin 族となる。そのようなもので最小が $\mathcal{D}$ であるから $\mathcal{D} \subset \sigma(\mathcal{C})$ を得る。

(ii) この場合、集合族 $\mathcal{C}$ で生成される Dynkin 族は S を含むので (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。

7.11 定理. $\mathcal{C}$ を S 上の集合族であって条件 $A \cap B \in \mathcal{C}\ \forall A \in \mathcal{C}\ \forall B \in \mathcal{C}$ を満たすもの、また μ, ν を $(S, \sigma(\mathcal{C}))$ 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 $C_n \subset C_{n+1}\ \forall n \in \mathbb{N}$, $\bigcup_{n=1}^{\infty} C_n = S$, $\mu(C_n) < +\infty\ \forall n \in \mathbb{N}$ 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. (ii) 集合族 $\{A \in \sigma(\mathcal{C}) : \mu(A \cap C_n) = \nu(A \cap C_n)\, \forall n \in \mathbb{N}\}$ は Dynkin 族であり $\mathcal{C} \cup \{S\}$ を包含する。定理 7.10 によればそのようなもので最小が $\sigma(\mathcal{C})$ であるから

$$\sigma(\mathcal{C}) \subset \{A \in \sigma(\mathcal{C}) : \mu(A \cap C_n) = \nu(A \cap C_n)\, \forall n \in \mathbb{N}\}$$

すなわち $A \in \sigma(\mathcal{C}) \Rightarrow \mu(A \cap C_n) = \nu(A \cap C_n)\ \forall n \in \mathbb{N}$ が成り立つ。一方、単調収束定理から $\sup_{n \in \mathbb{N}} \mu(A \cap C_n) = \mu(\bigcup_{n=1}^{\infty}(A \cap C_n)) = \mu(A)$ であり、同様に $\nu(A) = \sup_{n \in \mathbb{N}} \nu(A \cap C_n)$ である。従って $\mu(A) = \nu(A)\ \forall A \in \sigma(\mathcal{C})$ が導かれた。 □

7.12 演習問題. 定理 7.11(i) を示せ。

記号

$C(\mathbb{R}^d) :=$ $\mathbb{R}^d$ 上の実数値連続関数の全体、$\mathrm{supp}\, f := \{x \in \mathbb{R}^d : f(x) \neq 0\}$ の閉包
$C_0(\mathbb{R}^d) := \{f \in C(\mathbb{R}^d) : \mathrm{supp}\, f \text{ is bounded}\}$

7.13 補題. I, J を d 次元開区間であって I は有界かつその閉包は J に含まれているとする。このとき $f \in C_0(\mathbb{R}^d)$ であって $0 \leq f(x) \leq 1\ \forall x \in \mathbb{R}^d$, $f(x) = 1\ \forall x \in I$, $\mathrm{supp}\, f \subset J$ を満たすものが存在する。

7.14 演習問題. 補題 7.13 を示せ。（ヒント まず $d = 1$ として具体的に関数を構成せよ。）

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

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 $C_0(\mathbb{R})$ を**試験関数**(test function) の空間というが、試験関数の空間の選び方は問題に応じて適切なものをとる必要がある。

7.15 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

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

証明. 煩雑化を防ぐため $d = 1$ として議論を進める。定理 7.1 によれば $\mu = \nu$ は

$$\mu((a, +\infty)) = \nu((a, +\infty)) \ \forall a \in \mathbb{R}$$

と同値である。従って $\int_{\mathbb{R}} f \, \mu = \int_{\mathbb{R}} f \, \nu \ \forall f \in C_0(\mathbb{R})$ と仮定して上の条件を導けばよい。補題 7.13 によれば任意の $n \in \mathbb{N}$ に対して次のような $f \in C_0(\mathbb{R})$ が存在する。

$$0 \leq f(x) \leq 1 \ \forall x \in \mathbb{R}, \ f(x) = 1 \ \forall x \in (a + 1/n, a + n), \ \mathrm{supp}\, f \subset (a, +\infty).$$

積分の単調性により次が成り立つ。

$$\mu((a + 1/n, a + n)) \leq \int_{\mathbb{R}} f \, \mu = \int_{\mathbb{R}} f \, \nu \leq \nu((a, +\infty))$$

さて $(a + 1/n, a + n) \subset (a + 1/(n + 1), a + n + 1)$ かつ $\bigcup_{n=1}^{\infty} (a + 1/n, a + n) = (a, +\infty)$ であるから、$\mu((a + 1/n, a + n))$ は $n \to \infty$ の極限で $\mu((a, +\infty))$ に収束する。よって

$$\mu((a, +\infty)) \leq \nu((a, +\infty)) \ \forall a \in \mathbb{R}.$$

μ と ν の役割を入れ替えることができるので結論が得られた。 □

7.16 注意. $f_1, f_2, \ldots, f_d \in C_0(\mathbb{R})$ なら $f_1 \otimes f_2 \otimes \cdots \otimes f_d \in C_0(\mathbb{R}^d)$ なので次も μ, ν が一致するための必要十分条件である。

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

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 $\mathcal{A}_1 * \mathcal{A}_2 := \{A_1 \times A_2 \, ; A_1 \in \mathcal{A}_1, A_2 \in \mathcal{A}_2\}$ と書く。

復習

$(\mathcal{C}_1, m_1)$ を S_1 上の、$(\mathcal{C}_2, m_2)$ を S_2 上の有限加法的測度でともに σ-加法的かつ σ-有限なものとする。このとき $(S_1 \times S_2, \sigma(\mathcal{C}_1 * \mathcal{C}_2))$ 上の測度 μ であって、条件

$$\mu(A \times B) = m_1(A) m_2(B) \ \forall A \in \mathcal{C}_1, \forall B \in \mathcal{C}_2$$

を満たすものが唯一存在する。

8.1 定義. (i) 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 $\sigma(\mathcal{B}_1 * \mathcal{B}_2)$ 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。
(ii) $(S_1, \mathcal{B}_1)$ 上の σ-有限な測度 μ_1 と $(S_2, \mathcal{B}_2)$ 上の σ-有限な測度 μ_2 に対し

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

で規定される $(S_1 \times S_2, \mathcal{B}_1 \otimes \mathcal{B}_2)$ 上の測度 $\mu_1 \otimes \mu_2$ を**直積測度**(product measure) と呼ぶ。また三つ組み $(S_1 \times S_2, \mathcal{B}_1 \otimes \mathcal{B}_2, \mu_1 \otimes \mu_2)$ を**直積測度空間**(product measure space) という。

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

有限加法的測度の条件 (iii) が任意の可算無限な $\mathcal{C}$-分割についても成り立つとき 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。

8.2 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 7.11 を適用して導くことができる。

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。

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

これを d **次元** *Lebesgue* **測度**という。

以前にも述べたように $\mathrm{Borel}(\mathbb{R}^d)$ を Lebesgue 測度により完備化した σ-加法族上で考察するのが自然な場合があるが、ここではそうしない。

8.3 定義. 実確率変数系 $X_1,\ X_2,\ \ldots,\ X_n$ 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 $X_1,\ X_2,\ \ldots,\ X_n$ 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 $X_1,\ X_2,\ \ldots,\ X_n$ 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

$$\mu(A) = \nu(A) \text{ が } A = (a_1, +\infty) \times (a_2, +\infty) \times \cdots \times (a_n, +\infty) \text{ に対して成り立つ。}$$

よって定理 7.1 を適用して結論を得る。 □

8.5 演習問題. 実確率変数系 $X_1,\ X_2,\ \ldots,\ X_n$ は独立であるとする。
(i) $m \leq n$ に対して実確率変数系 $X_1,\ X_2,\ \ldots,\ X_m$ も独立であることを示せ。
(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 $f_1(X_1),\ f_2(X_2),\ \ldots,\ f_n(X_n)$ も独立であることを示せ。

8.6 演習問題. $0 < p < 1$ について次が成り立つなら実確率変数系 $X_1,\ X_2,\ \ldots,\ X_n$ は独立であることを示せ。

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

ヒント $a_1 < 1,\ a_2 < 1,\ \ldots,\ a_n < 1$ とする。このとき $J := \{i \in \mathbb{N} : 1 \leq i \leq n, a_i \geq -1\}$ とおくと $P(X_1 > a_1, X_2 > a_2, \ldots, X_n > a_n) = P(X_i = 1\, \forall i \in J)$ であることを示せ。

8.7 例. 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 6.15 と演習問題 8.6 により独立である。

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

S_1, S_2 上にそれぞれ σ-加法族 $\mathcal{B}_1$ と $\mathcal{B}_2$ が与えられ、さらに $(S_1, \mathcal{B}_1)$, $(S_2, \mathcal{B}_2)$ 上にそれぞれ σ 有限な測度 μ_1 と μ_2 が与えられている。

8.8 演習問題. 非負値 Borel 可測関数 $\rho : \mathbb{R}^d \to \overline{\mathbb{R}}$ を密度関数に持つ絶対連続な $(\mathbb{R}^d, \mathrm{Borel}(\mathbb{R}^d))$ 上の測度 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。

8.9 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。

約束

S 上に測度 $(\mathcal{B}, \mu)$ が与えられているとする。S の部分集合 A に対し $\exists B \in \mathcal{B}$ s.t. $\mu(B^{\mathrm{c}}) = 0$, $B \subset A$ が成り立つときそれを μ-a.e. 集合という。

次が通常 *Fubini* **の定理**と呼んで引用されるものである。非負値とは限らない関数に Fubini の定理を適用する場合にその前提条件となる可積分性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) $A := \{x \in S_1 : \int_{S_2} |f(x, y)|\, \mu_2(dy) < +\infty\}$ 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。

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。

8.14 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 $\sigma(\mathcal{C}_1 * \mathcal{C}_2) = \sigma(\mathcal{C}_1) \otimes \sigma(\mathcal{C}_2)$ が成り立つ。

証明. まず $\mathcal{C}_1 * \mathcal{C}_2 \subset \sigma(\mathcal{C}_1) * \sigma(\mathcal{C}_2) \subset \sigma(\mathcal{C}_1) \otimes \sigma(\mathcal{C}_2)$ である。右辺は $\mathcal{C}_1 * \mathcal{C}_2$ を含む σ-加法族であるがそのようなものの最小が $\sigma(\mathcal{C}_1 * \mathcal{C}_2)$ なので $\sigma(\mathcal{C}_1 * \mathcal{C}_2) \subset \sigma(\mathcal{C}_1) \otimes \sigma(\mathcal{C}_2)$ 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) により次の集合族

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

は S_1 上の σ-加法族で、S_2 は $\mathcal{C}_2$ に属する集合の可算合併であるから、$\mathcal{B}$ は $\mathcal{C}_1$ を含む。そのようなものの最小が $\sigma(\mathcal{C}_1)$ なので $\sigma(\mathcal{C}_1) \subset \mathcal{B}$ すなわち

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

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

$$A \times B = (A \times S_2) \cap (S_1 \times B)$$

だから $\sigma(\mathcal{C}_1) * \sigma(\mathcal{C}_2) \subset \sigma(\mathcal{C}_1 * \mathcal{C}_2)$ が導かれた。右辺は $\sigma(\mathcal{C}_1) * \sigma(\mathcal{C}_2)$ を含む σ-加法族であるがそのようなものの最小が $\sigma(\mathcal{C}_1) \otimes \sigma(\mathcal{C}_2)$ なので $\sigma(\mathcal{C}_1) \otimes \sigma(\mathcal{C}_2) \subset \sigma(\mathcal{C}_1 * \mathcal{C}_2)$ を得る。 □

8.15 系. $d = d_1 + d_2$ とすると $\mathcal{B}(\mathbb{R}^d) = \mathcal{B}(\mathbb{R}^{d_1}) \otimes \mathcal{B}(\mathbb{R}^{d_2})$, $\lambda^{(d)} = \lambda^{(d_1)} \otimes \lambda^{(d_2)}$.

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 8.14 を適用して結論を得る。 □

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

$$\int_{(0,\infty)} x\, \mu(dx) = \int_{(0,\infty)} \mu((x, +\infty))\, \lambda(dx)$$

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 1_A を直積測度 $\lambda \otimes \mu$ で積分する。系 8.15 と定理 8.10 を適用して

$$\int_{\mathbb{R}} \Big(\int_{\mathbb{R}} 1_A(x, y)\, \lambda(dx) \Big) \mu(dy) = (\lambda \otimes \mu)(A) = \int_{\mathbb{R}} \Big(\int_{\mathbb{R}} 1_A(x, y)\, \mu(dy) \Big) \lambda(dx)$$

を得る。左辺および右辺において

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

であるから結論に達する。 □

8.17 例. $t \in \mathbb{R}_{>0}$, $a, \xi \in \mathbb{R}$ に対して $\displaystyle\frac{t}{\pi}\int_{\mathbb{R}} \frac{e^{\sqrt{-1}\xi x}}{t^2+(x-a)^2}\,\lambda(dx) = \exp\{\sqrt{-1}\xi a - t|\xi|\}$

証明. ここでは複素解析を援用しない方法を紹介する。まず例 4.3 の最後の公式において適当な変数の置き換えを行うことにより次の関係を得る。

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

$\varepsilon \in \mathbb{R}_{>0}$ とする。次の関数は 2 次元 Lebesgue 測度 $\lambda^{(2)}$ に関して可積分である。

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

定理 8.12(ii) を適用して

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

さて例 4.3 の最後の公式を再び使うと

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

定理 8.12(ii) を適用して

$$\frac{e^{\sqrt{-1}\xi a}}{2\pi}\int_{\mathbb{R}^2} e^{-t|y+\xi|}e^{\sqrt{-1}y(a-x)}e^{-\varepsilon|x|}\,\lambda^{(2)}(dxdy) = \frac{e^{\sqrt{-1}\xi a}}{\pi}\int_{\mathbb{R}} e^{-t|y+\xi|}e^{\sqrt{-1}ya}\frac{\varepsilon}{\varepsilon^2+y^2}\,\lambda(dy)$$

以上を組み合わせて次の公式を得る。

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

演習問題 4.5(ii) で述べたように、$\varepsilon \to 0$ のとき右辺は有界連続関数 $y \mapsto e^{-t|y+\xi|}e^{\sqrt{-1}(y+\xi)a}$ の 0 における値 $e^{-t|\xi|}e^{\sqrt{-1}\xi a}$ 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。

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

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。 □

8.18 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 11.11 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 8.17 で登場したテクニックを使う。$\varepsilon\in\mathbb{R}_{>0}$ 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。定理 8.12(ii) を適用して積分の順序交換を行う。

$$\int_{\mathbb{R}} e^{\sqrt{-1}xa}e^{-\varepsilon|x|}g(x)\,\lambda(dx)=\int_{\mathbb{R}}\Big(\int_{\mathbb{R}} e^{\sqrt{-1}xa}e^{-\varepsilon|x|}e^{-\sqrt{-1}xy}\,\lambda(dx)\Big)f(y)\,\lambda(dy).$$

例 8.17 でも述べたように

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

従って

$$\int_{\mathbb{R}} e^{\sqrt{-1}xa}e^{-\varepsilon|x|}g(x)\,\lambda(dx)=2\pi\frac{\varepsilon}{\pi}\int_{\mathbb{R}}\frac{f(y)}{\varepsilon^2+(y-a)^2}\,\lambda(dy).$$

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、演習問題 4.5(ii) で述べたように、$\varepsilon\to 0$ のとき右辺は $2\pi f(a)$ 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。 □

8.19 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。

$$\frac{1}{2\pi}\int_{\mathbb{R}}\frac{e^{\sqrt{-1}xa}}{(1+\sqrt{-1}x)^n}\,\lambda(dx)=\begin{cases}a^{n-1}e^{-a}/(n-1)! & a\ge 0\\ 0 & a\le 0\end{cases}$$

8.20 補題. $d=d_1+d_2$ とする。μ_1, μ_2 をそれぞれ非負値 Borel 可測関数 $\rho_i:\mathbb{R}^{d_i}\to\overline{\mathbb{R}}$ を密度関数に持つ絶対連続かつ σ 有限な $(\mathbb{R}^{d_i},\mathrm{Borel}(\mathbb{R}^{d_i}))$ 上の測度とする。このとき直積測度 $\mu_1\otimes\mu_2$ は絶対連続で $\rho_1\otimes\rho_2:\mathbb{R}^d\to\overline{\mathbb{R}}$, $(x,y)\mapsto\rho_1(x)\rho_2(y)$ 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。系 8.15 および定理 8.10 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。 □

8.21 例. 1 次元標準正規分布の n 次直積測度を n **次元標準正規分布**という。あるいは測度

$$\mathrm{Borel}(\mathbb{R}^n)\to\overline{\mathbb{R}},A\mapsto\int_A\frac{1}{\sqrt{(2\pi)^n}}\exp\{-\|x\|^2/2\}\lambda^{(n)}(dx)$$

のこと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。

つぎは**期待値の乗法定理**と称して引用される。

8.22 定理. 実確率変数系 $X_1,\ X_2,\ \ldots,\ X_n$ は独立であるとする。
(i) 非負値 Borel 可測関数 $f_1, f_2, \ldots, f_n : \mathbb{R} \to \overline{\mathbb{R}}$ に対して次が成り立つ。

$$E[f_1(X_1)f_2(X_2)\ldots f_n(X_n)] = E[f_1(X_1)]E[f_2(X_2)]\ldots E[f_n(X_n)].$$

(ii) Borel 可測関数 $f_1, f_2, \ldots, f_n : \mathbb{R} \to \overline{\mathbb{R}}$ に対して $E[|f_i(X_i)|] < +\infty\ \forall i = 1, 2, \ldots, n$ なら積 $f_1(X_1)f_2(X_2)\ldots f_n(X_n)$ 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 6.24 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。$\mathcal{L}((X_1, X_2), \cdot) = \mathcal{L}(X_1, \cdot) \otimes \mathcal{L}(X_2, \cdot)$ であるから定理 8.10 により右辺は次に等しい

$$\int_{\mathbb{R}} f_1(x_1)\,\mathcal{L}(X_1, dx_1) \int_{\mathbb{R}^2} f_2(x_2)\,\mathcal{L}(X_2, dx_2)$$

再び定理 6.24 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 8.12 も必要である。 □

8.23 系. 実確率変数系 $X_1,\ X_2,\ \ldots,\ X_n$ は独立であるとする。このとき次が成り立つ。

$$E[\exp\{\sum_{i=1}^{n} \sqrt{-1}z_iX_i\}] = \prod_{i=1}^{n} E[\exp\{\sqrt{-1}z_iX_i\}]\ \forall(z_1, z_2, \ldots, z_n) \in \mathbb{R}^n.$$

8.24 例. n 次元標準正規分布 μ に対し $\int_{\mathbb{R}^n} e^{\sqrt{-1}(z,x)}\mu(dx) = e^{-\|z\|^2/2}\ \forall z \in \mathbb{R}^n$ が成り立つ。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。

8.25 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。

$$E[f_1(X_1)f_2(X_2)\ldots f_n(X_n)] = E[f_1(X_1)]E[f_2(X_2)]\ldots E[f_n(X_n)]\ \forall f_1, f_2, \ldots, f_n \in C_0(\mathbb{R}).$$

証明. $X_1,\ X_2,\ \ldots,\ X_n$ の結合分布を μ とし、またそれらの分布の直積測度を ν とすると定理 6.24 と定理 8.22 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 7.15 を適用して $\mu = \nu$ を得る。 □

9 可逆アファイン写像とルベーグ測度

$(\mathbb{R}^d, \mathrm{Borel}(\mathbb{R}^d))$ 上では Lebesgue 測度について絶対連続な測度が重要である。

9.1 例. 与えられた n 次正定値対称行列 C と $a \in \mathbb{R}^n$ に対して測度

$$\mathrm{Borel}(\mathbb{R}^n) \to \overline{\mathbb{R}}, A \mapsto \int_A \frac{1}{\sqrt{(2\pi)^n \det C}} \exp\{-C^{-1}(x-a, x-a)/2\} \lambda^{(n)}(dx)$$

を平均 a 共分散 C の n **次元正規分布**という。ここで $C^{-1}(\cdot,\cdot)$ は C^{-1} が誘導する 2 次形式を表す。平均が 0 で共分散が単位行列の場合は例 8.21 で述べた n 次元標準正規分布である。

$\alpha \in GL(d)$, $\beta \in \mathbb{R}^d$ とする。このときアファイン写像 $\varphi : \mathbb{R}^d \to \mathbb{R}^d$, $x \mapsto \alpha x + \beta$ は Borel 可測である。絶対連続な測度の写像 φ による像測度について考察したいが、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. $\mathcal{B}_1, \mathcal{B}_2, \mathcal{M}_1, \mathcal{M}_2$ をそれぞれ S_1, S_2, T_1, T_2 上の σ 加法族とする。写像 $\varphi_1 : S_1 \to T_1$ が対 $\mathcal{B}_1, \mathcal{M}_1$ に関し、$\varphi_2 : S_2 \to T_2$ が $\mathcal{B}_2, \mathcal{M}_2$ に関しそれぞれ可測なら $\varphi_1 \times \varphi_2$ は対 $\mathcal{B}_1 \otimes \mathcal{B}_2$, $\mathcal{M}_1 \otimes \mathcal{M}_2$ に関して可測である。またそれぞれ $(S_1, \mathcal{B}_1)$, $(S_2, \mathcal{B}_2)$ 上の測度 μ_1, μ_2 に対して像測度 $\varphi_{1*}\mu_1$, $\varphi_{2*}\mu_2$ が σ 有限なら $(\varphi_1 \times \varphi_2)_*(\mu_1 \otimes \mu_2) = (\varphi_{1*}\mu_1) \otimes (\varphi_{2*}\mu_2)$ である。

証明. 可測性の判定に補題 5.6 を適用する。$\sigma(\mathcal{M}_1 * \mathcal{M}_2) = \mathcal{M}_1 \otimes \mathcal{M}_2$ 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 $\varphi_* \lambda^{(d)} = \lambda^{(d)}/|\det \alpha|$ が成り立つ。

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。) 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。補題 3.14 によれば、

$$\lambda(\alpha_i^{-1}(A - \beta_i)) = \lambda(A)/|\alpha_i| \ \forall A \in \mathrm{Borel}(\mathbb{R}), i = 1, 2$$

である。ただし λ は 1 次元 Lebesgue 測度である。これは写像 $\varphi_i : \mathbb{R} \to \mathbb{R}$, $x \mapsto \alpha_i x + \beta_i$ を導入すると $\varphi_{i*}\lambda = \lambda/|\alpha_i|$ と表現できる。系 8.15 と補題 9.2 を適用して

$$(\varphi_1 \times \varphi_2)_* \lambda^{(2)} = \frac{1}{|\alpha_1 \alpha_2|} \lambda^{(2)} = \frac{1}{|\det \alpha|} \lambda^{(2)}$$

が導かれる。写像 $\varphi_1 \times \varphi_2$ は $\mathbb{R}^2 \to \mathbb{R}^2$, $x \mapsto \alpha x + \beta$ に一致している。 □

系 9.3 から特に d **次元** *Lebesgue* **測度** $\lambda^{(d)}$ **の平行移動不変性**がわかるが、これがまさに Lebesgue 測度の特質なのである。

9.4 定義. $(\mathbb{R}^d, \mathrm{Borel}(\mathbb{R}^d))$ 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 Radon 測度という。

9.5 例. 非負値連続関数 $\rho : \mathbb{R}^d \to \mathbb{R}$ を密度関数にもつ絶対連続な $(\mathbb{R}^d, \mathrm{Borel}(\mathbb{R}^d))$ 上の測度は Radon 測度である。

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

9.7 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 $\mu((a,b] \times A) = \mu((0,1] \times A)(b-a)\ \forall a, b \in \mathbb{R},\ a < b$ を得る。同様にして $\mu(A \times (a,b]) = \mu(A \times (0,1])(b-a)\ \forall a, b \in \mathbb{R},\ a < b$ 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。ゆえに定理 7.11 を適用して $\mu = \mu((0,1]^2)\lambda^{(2)}$ という結論に至った。 □

9.8 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 $\mathbb{R}^d$ 上の測度 $\varphi_*\lambda^{(d)}$ 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 6.12 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 9.7 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。 □

9.9 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**。

9.10 定理. $\alpha \in GL(d)$, $\beta \in \mathbb{R}^d$ とする。任意の非負値 $\mathrm{Borel}(\mathbb{R}^d)$ 可測関数 $f : \mathbb{R}^d \to \overline{\mathbb{R}}$ に対して次が成り立つ。

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

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 9.9 によりこの写像は系 9.3 あるいは系 9.8 が適用可能な写像の合成として表現できる。また補題 9.9 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 6.12 により $\varphi_*\lambda^{(d)} = \lambda^{(d)}/|\det \alpha|$ が導かれる。さらに定理 6.17 を適用して結論に到達する。 □

9.11 系. 非負値 Borel 可測関数 $\rho : \mathbb{R}^d \to \overline{\mathbb{R}}$ を密度関数に持つ絶対連続な $(\mathbb{R}^d, \mathrm{Borel}(\mathbb{R}^d))$ 上の測度を μ とする。$\alpha \in GL(d)$, $\beta \in \mathbb{R}^d$ に対しアファイン写像 $\varphi : \mathbb{R}^d \to \mathbb{R}^d$, $x \mapsto \alpha x + \beta$ による μ の像測度 $\varphi_*\mu$ は絶対連続で $\mathbb{R}^d \to \overline{\mathbb{R}}$, $x \mapsto \rho(\alpha^{-1}(x - \beta))/|\det \alpha|$ を密度関数に持つ。

9.12 演習問題. 系 9.11 を示せ。

9.13 演習問題. $\alpha \in GL(n)$, $\beta \in \mathbb{R}^n$ 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ことを示せ。

9.14 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 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 $C = \alpha\,{}^{\mathrm{t}}\alpha$ を満たす $\alpha \in GL(n)$ が存在する。従って演習問題 9.13 の結論により例 8.24 に帰着する。 □

9.15 演習問題. $X_1, X_2, \ldots, X_n$ 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確率変数系とする。このとき次を示せ。

$$E[X_i] = a \text{ の第 } i \text{ 成分、} \mathrm{Cov}[X_i, X_j] = C \text{ の第 } ij \text{ 成分}$$

即ち、n 次元確率変数 $(X_1, X_2, \ldots, X_n)$ の平均ベクトルは a で共分散行列は C である。

10　可微分同相写像とルベーグ測度

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。

10.1 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

10.2 定義. d 次正方行列 A に対して $\|A\| := \sup\{\|Au\|\,; u \in \mathbb{R}^d, \|u\| = 1\}$

10.3 補題. d 次正方行列 A, B に対して以下が成り立つ。
(i) $\|Au\| \leq \|A\|\|u\|\ \forall u \in \mathbb{R}^d$. $\|AB\| \leq \|A\|\|B\|$.
(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。 □

10.4 演習問題. 補題 10.3(i), (ii) を示せ。

10.5 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。三角不等式と補題 10.3(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 10.3(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 10.3(i) により

$$\|v\| \leq \|B - \mathbf{1}\|\|u\| + \|A - \mathbf{1}\|(\|u\| + \|v\|)$$

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

$$\|(A^{-1}B - 1)u\| = \|v\| \leq \frac{\|B - \mathbf{1}\| + \|A - \mathbf{1}\|}{1 - \|A - \mathbf{1}\|}\|u\|$$

を得る。$u \in \mathbb{R}^d$ は任意であるので結論に到達した。 □

記号

$\mathrm{Nbd}(a, \delta) := \{x \in \mathbb{R}^d : \|x - a\| < \delta\}$ ただし $a \in \mathbb{R}^d$, $\delta > 0$

10.6 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 10.3(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。 □

上の結果を有界閉集合上での一様評価に拡張しておく。

10.7 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

$$\begin{aligned}
&\mathrm{Nbd}(t, 2\delta) \subset D\ \forall t \in K;\ \|\varphi'(t)^{-1}\varphi'(x) - \mathbf{1}\| < \varepsilon\ \forall t \in K, \forall x \in \mathrm{Nbd}(t, \delta);\\
&\|\varphi'(t)^{-1}(\varphi(x) - \varphi(y)) - (x - y)\| \leq \varepsilon\|x - y\|\ \forall t \in K, \forall x, \forall y \in \mathrm{Nbd}(t, \delta)
\end{aligned}$$

が成り立つ。

証明. 補題 10.6 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

$$\|x - t\| < 2\delta \Rightarrow \|x - z\| \leq \|x - t\| + \|t - z\| < 2\delta + \delta(z) \leq 3\delta(z)$$

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

$$\begin{aligned}\|x - t\| < \delta \Rightarrow \|x - z\| &\leq \|x - t\| + \|t - z\| < \delta + \delta(z) \leq 2\delta(z) \\ &\Rightarrow \|\varphi'(z)^{-1}\varphi'(x) - \mathbf{1}\| < \varepsilon/(2+\varepsilon).\end{aligned}$$

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 10.5(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。 □

10.8 演習問題. 補題 10.7 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。

10.9 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。 □

10.10 演習問題. 補題 10.9(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

10.11 補題. 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 10.9(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 $\mathrm{diam}\, J := \sup\{\|x - y\|\,; x, y \in J\}$, $\mathrm{proj}_i : \mathbb{R}^d \to \mathbb{R}$ 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。

$$\mathrm{diam}\, I \le r \min_{1 \le i \le d} \lambda(\mathrm{proj}_i I)$$

ここで λ は 1 次元 Lebesgue 測度である。

10.12 補題. $K = \overline{I}$ と $0 < \varepsilon < 1$ なる ε に対し $\delta > 0$ を補題 10.7 で述べたものとする。このとき $J \in \mathcal{I}$ が条件 $J \subset I$, $\mathrm{diam}\, J < 2\delta$, $\mathrm{diam}\, J \le r \min_{1 \le i \le d} \lambda(\mathrm{proj}_i J)$ 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 $\mathrm{proj}_i u \le \|u\|$ であるから

$$\mathrm{proj}_i\{\varphi'(t)^{-1}(\varphi(x) - \varphi(t))\} \le \mathrm{proj}_i(x - t) + \|\varphi'(t)^{-1}(\varphi(x) - \varphi(t)) - (x - t)\|.$$

ここで t は J の中心である。右辺第 2 項に補題 10.7 を適用して

$$\mathrm{proj}_i\{\varphi'(t)^{-1}(\varphi(x) - \varphi(t))\} < \mathrm{proj}_i(x - t) + \varepsilon\|x - t\|.$$

右辺第 1 項は $\lambda(\mathrm{proj}_i J)/2$ で、第 2 項は $\mathrm{diam}\, J/2$ で押さえられる。よって

$$\mathrm{proj}_i\{\varphi'(t)^{-1}(\varphi(x) - \varphi(t))\} < (1 + r\varepsilon)\lambda(\mathrm{proj}_i J)/2$$

が得られる。ここで $\mathrm{diam}\, J \le r\lambda(\mathrm{proj}_i J)$ という仮定を使った。同様にして

$$-(1 + r\varepsilon)\lambda(\mathrm{proj}_i J)/2 < \mathrm{proj}_i\{\varphi'(t)^{-1}(\varphi(x) - \varphi(t))\}$$

もわかるので

$$\mathrm{proj}_i\{\varphi'(t)^{-1}(\varphi(x) - \varphi(t))\} \in (1 + r\varepsilon)\mathrm{proj}_i(J - t) \ \forall i = 1, 2, \ldots, d,\ \forall x \in J$$

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 9.10 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 10.7 と系 10.5(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。 □

10.13 系. $I\in\mathcal{I}$ とする。$\overline{I}\subset D$ であれば

$$\lambda^{(d)}(\varphi(I)) \le \int_I |\det\varphi'|\lambda^{(d)}.$$

証明. $K=\overline{I}$ と $0<\varepsilon<1$ なる ε に対し $\delta>0$ を補題 10.7 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 10.12 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。

$$\lambda^{(d)}(\varphi(I)) \le \frac{(1+r\varepsilon)^d}{(1-\varepsilon)^d}\int_I |\det\varphi'|\lambda^{(d)}.$$

ε は $0<\varepsilon<1$ である限り任意なので結論が導かれた。 □

10.14 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 $\mu(A)\le\nu(A)\ \forall A\in\mathcal{C}$ が成り立てば、$\mu(A)\le\nu(A)$ $\forall A\in\sigma(\mathcal{C})$ 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 $\mu(A)\le\nu(A)\ \forall A\in\mathcal{C} \Rightarrow \mu(A)\le\nu(A)\ \forall A\in\sigma(\mathcal{C})$ が成り立つ。 □

10.15 系. (i) $\lambda^{(d)}(\varphi(A)) \le \int_A |\det\varphi'|\lambda^{(d)}\ \forall A\in\mathrm{Borel}(D)$.

(ii) 非負 Borel 可測関数 $f: D\to\overline{\mathbb{R}}$ に対して $\int_{\varphi(D)} f\circ\varphi^{-1}\,\lambda \le \int_D f|\det\varphi'|\,\lambda^{(d)}$

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 10.11 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（これは演習問題とする）。従って系 10.13 と補題 10.14 から (i) が得られる。

(ii) $A\in\mathrm{Borel}(D)$ とするとき (i) により $f=1_A$ については成り立っている。 □

10.16 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 $\{I \in \mathcal{I} : \bar{I} \subset D\} \cup \{\emptyset\}$ について σ 有限であることを示せ。
(ii) 系 10.15(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。

10.17 定理. 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。

$$\int_D f \circ \varphi \, \lambda^{(d)} = \int_{\varphi(D)} f \frac{1}{|\det \varphi'| \circ \varphi^{-1}} \, \lambda^{(d)}.$$

証明. $B \in \mathrm{Borel}(\varphi(D))$ とする。$\varphi^{-1} : \varphi(D) \to D$ に系 10.15(i) を適用して

$$\lambda^{(d)}(\varphi^{-1}(B)) \leq \int_B \frac{1}{|\det \varphi'| \circ \varphi^{-1}} \lambda^{(d)}.$$

他方 $f : D \to \overline{\mathbb{R}},\, x \mapsto 1_B(\varphi(x))/|\det \varphi'(x)|$ に系 10.15(ii) を適用して

$$\int_{\varphi(D)} 1_B \frac{1}{|\det \varphi'| \circ \varphi^{-1}} \, \lambda^{(d)} \leq \int_D 1_B \circ \varphi \lambda^{(d)} = \lambda^{(d)}(\varphi^{-1}(B)).$$

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。 $\square$

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。

10.18 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

$$\varphi(D) \to \overline{\mathbb{R}}, y \mapsto \rho(\varphi^{-1}(y))/|\det \varphi'(\varphi^{-1}(y))|$$

を密度関数に持つ。

10.19 演習問題. 系 10.18 を示せ。

定理 10.17 の典型的な応用例として**極座標変換**を取り上げよう。

10.20 例. $D := \mathbb{R}^2 \setminus \{x \in \mathbb{R}^2 : x_1 \leq 0, x_2 = 0\}$ 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 10.17 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 8.10 と補題 3.19(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 8.10 により左辺は $\int_{\mathbb{R}} e^{-x^2/2}\lambda(dx)$ の 2 乗である。従って

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

10.21 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 10.18 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

$$\int_{\varphi^{-1}(A)}e^{-x_1}{x_1}^{a-1}e^{-x_2}{x_2}^{b-1}\,\lambda^{(2)}(dx)=\int_A e^{-y_1}{y_1}^{a+b-1}(1-y_2)^{a-1}{y_2}^{b-1}\,\lambda^{(2)}(dy)$$

となる。従って $A=(0,+\infty)\times(0,1)$ として両辺に Fubini の定理を適用すると

$$\Gamma(a)\Gamma(b)=\Gamma(a+b)B(a,b)\ \forall a,b\in\mathbb{R}_{>0}$$

が得られる。ここで Γ, B はそれぞれ gamma 関数、beta 関数である。また

$\gamma_a:=x\mapsto e^{-x}x^{a-1}/\Gamma(a)$ を密度に持つ $(0,+\infty)$ 上の絶対連続測度
$\beta_{a,b}:=y\mapsto(1-y)^{a-1}y^{b-1}/B(a,b)$ を密度に持つ $(0,1)$ 上の絶対連続測度

と定義すると次が成り立つ。

$$\varphi_*(\gamma_a\otimes\gamma_b)=\gamma_{a+b}\otimes\beta_{a,b}.$$

γ_a を指数 a の *gamma* **分布**といい、$\beta_{a,b}$ を指数 a, b の *beta* **分布**という。$B\in\mathrm{Borel}((0,+\infty))$ とする。$A=B\times(0,1)$ に対して適用すると

$$\gamma_a\otimes\gamma_b(\{(x_1,x_2):x_1+x_2\in B\})=\gamma_{a+b}(B).$$

さて X を指数 a の gamma 分布に従う確率変数、Y を指数 b の gamma 分布に従う確率変数とする。上の公式により

X と Y が互いに独立であるなら $X+Y$ は指数 $a+b$ の gamma 分布に従う

ことが分かる。これを gamma 分布の再生性という。

写像 φ の単射性が崩れるときは、定理 10.17 には修正が必要である。

前提

D は $\mathbb{R}^d$ の開集合、$\varphi : D \to \mathbb{R}^d$ は C^1 級写像で $\det \varphi'(x) \neq 0\ \forall x \in D$ を満たすもの。

10.22 注意. 局所逆写像定理により各 $x \in D$ に対して次が成り立つような近傍 U が存在する。

$U \subset D$, $\varphi(U)$ は開集合、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。

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。

10.23 定理. (i) 任意の開集合 U に対して $U \subset D$ なら $\varphi(U)$ も開集合である。
(ii) 各 $y \in \varphi(D)$ に対して $\varphi^{-1}(\{y\})$ は可算集合である。
(iii) 任意の非負値 Borel 可測関数 $\rho : D \to \overline{\mathbb{R}}$ に対して関数

$$\varphi(D) \to \overline{\mathbb{R}},\ y \mapsto \textstyle\sum_{x\in\varphi^{-1}(\{y\})} \rho(x)/|\det \varphi'(x)|$$

は Borel 可測でかつ次が成り立つ。

$$\int_{\varphi^{-1}(A)} \rho\,\lambda^{(d)} = \int_A \sum_{x\in\varphi^{-1}(\{y\})} \frac{\rho(x)}{|\det \varphi'(x)|}\,\lambda^{(d)}(dy)\ \forall A \in \mathrm{Borel}(\varphi(D)).$$

(iv) ρ を密度関数に持つ絶対連続な $(D, \mathrm{Borel}(D))$ 上の測度 μ の像測度 $\varphi_*\mu$ は絶対連続で

$$y \mapsto \sum_{x\in\varphi^{-1}(\{y\})} \rho(x)/|\det \varphi'(x)|$$

を密度関数に持つ。

証明. $\mathcal{C}$ 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。

$$\mathcal{O} := \{I \in \mathcal{C} : \text{注意 10.22 の条件を満たす}\}$$

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 10.22 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 10.22 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 10.22 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 $n\in\mathbb{N}$ に対して E_n は $E_n\subset I_n$ なるボレル集合でありかつ D の可算分割を構成する。$y\in\varphi(D)$ を固定すると注意 10.22 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 10.18 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 番目の等号では族 $\{E_n\,;n\in\mathbb{N}\}$ が D の可算分割を構成することが利いている。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 $y\mapsto\sum_{x\in\varphi^{-1}(\{y\})}\rho(x)/|\det\varphi'(x)|$ も Borel 可測でありかつ (iii) に挙げた等式が成り立つ。 □

10.24 例. $D=\mathbb{R}\setminus\{0\}$ とする。

(i) $\varphi: D\to\mathbb{R},\ x\mapsto x^2/2$ に定理 10.23 を適用しよう。$\varphi(D)=\mathbb{R}_{>0}$ であり $\varphi'(x)=x\ \forall x\in D$ である。したがって非負値 Borel 可測関数 $\rho: D\to\overline{\mathbb{R}}$ と $A\in\mathrm{Borel}(\mathbb{R}_{>0})$ に対して

$$\int_{\varphi^{-1}(A)}\rho\,\lambda = \int_A\frac{\rho(\sqrt{2y})}{\sqrt{2y}}.$$

特に標準正規分布の像測度は密度 $y\mapsto e^{-y}/\sqrt{\pi y}$ を持つ絶対連続測度である。これはパラメータ 1/2 の gamma 分布である。

(ii) $\varphi: D\to\mathbb{R},\ x\mapsto\log x^2$ に定理 10.23 を適用しよう。$\varphi(D)=\mathbb{R}$ であり $\varphi'(x)=2/x\ \forall x\in D$ である。したがって非負値 Borel 可測関数 $\rho: D\to\overline{\mathbb{R}}$ と $A\in\mathrm{Borel}(\mathbb{R})$ に対して

$$\int_{\varphi^{-1}(A)}\rho\,\lambda = \int_A\frac{\rho(e^{y/2})}{2/e^{y/2}}+\frac{\rho(-e^{y/2})}{2/e^{y/2}}\,\lambda(dy) = \int_A\frac{\rho(e^{y/2})+\rho(-e^{y/2})}{2}e^{-y/2}\,\lambda(dy).$$

特に Cauchy 分布の像測度は密度 $y\mapsto 1/2\pi\cosh(y/2)$ を持つ絶対連続測度である。

(iii) $\varphi: D\to\mathbb{R},\ x\mapsto 1/(1+x^2)$ に定理 10.23 を適用しよう。$\varphi(D)=\mathbb{R}_{(0,1)}$ であり $\varphi'(x)=-2x/(1+x^2)^2\ \forall x\in D$ である。したがって非負値 Borel 可測関数 $\rho: D\to\overline{\mathbb{R}}$ と $A\in\mathrm{Borel}(\mathbb{R}_{(0,1)})$ に対して

$$\int_{\varphi^{-1}(A)}\rho\,\lambda = \int_A\sum_{\varepsilon=+1,-1}\frac{\rho(\varepsilon\sqrt{(1-y)/y})}{2\sqrt{(1-y)/y}y^2}\,\lambda(dy) = \int_A\sum_{\varepsilon=+1,-1}\frac{\rho(\varepsilon\sqrt{(1-y)/y})}{2y\sqrt{y(1-y)}}\,\lambda(dy).$$

特に Cauchy 分布の像測度は密度 $y\mapsto 1/\pi\sqrt{y(1-y)}$ を持つ絶対連続測度である。

11　特性関数と正規分布

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。

11.1 定義. $(\mathbb{R}^d, \mathrm{Borel}(\mathbb{R}^d))$ 上の確率測度 μ に対して $\mathbb{R}^d \to \mathbb{C}$, $\xi \mapsto \int_{\mathbb{R}^d} e^{\sqrt{-1}(\xi,x)}\,\mu(dx)$ を μ の**特性関数**(characteristic function) といい $\mathrm{char}(\mu,\cdot)$ と表記する。また確率変数 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。

11.2 補題. $(\mathbb{R}^d, \mathrm{Borel}(\mathbb{R}^d))$ 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。

証明. 三角不等式により $|\int_{\mathbb{R}^d} e^{\sqrt{-1}(\xi,x)}\,\mu(dx)| \le \int_{\mathbb{R}^d} |e^{\sqrt{-1}(\xi,x)}|\,\mu(dx) = 1$ である。次に ξ_n を ξ に収束する $\mathbb{R}^d$ の点列とする。関数列 $x \mapsto \exp\{\sqrt{-1}(\xi_n, x)\}$ は関数 $x \mapsto \exp\{\sqrt{-1}(\xi, x)\}$ に各点収束する。他方、定数 1 が μ について可積分な優関数の役割を果たす。従って Lebesgue の収束定理を適用して $\int_{\mathbb{R}^d} \exp\{\sqrt{-1}(\xi_n,x)\}\,\mu(dx)$ が $\int_{\mathbb{R}^d} e^{\sqrt{-1}(\xi,x)}\,\mu(dx)$ に収束することがわかる。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。 □

11.3 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 $(\cdot,\cdot)$ はユークリッド内積であり $(\alpha C^{\mathrm{t}}\alpha)(\cdot,\cdot)$ は $\alpha C^{\mathrm{t}}\alpha$ 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 9.14 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。

(i) 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？
(ii) 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？

(i) については今すぐ解決することができる。

11.4 系. 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、$(\mathbb{R}^n, \mathrm{Borel}(\mathbb{R}^n))$ 上の確率測度 μ で $\mathrm{char}(\mu,\xi) = \exp\{\sqrt{-1}(\xi,a) - C(\xi,\xi)/2\}$ $\forall \xi \in \mathbb{R}^n$ を満たすものが存在する。

証明. アイデァは補題 9.14 と類似である。即ち n 次行列 α で $C = \alpha\,{}^{\mathrm{t}}\alpha$ 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 11.3 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。

11.5 定義. 非負定値な n 次対称行列 C と $a \in \mathbb{R}^n$ が与えられたとき、$(\mathbb{R}^n, \mathrm{Borel}(\mathbb{R}^n))$ 上の確率測度で特性関数が $\xi \mapsto \exp\{\sqrt{-1}(\xi, a) - C(\xi, \xi)/2\}$ であるものを平均 a 共分散 C の n **次元正規分布**という。

このように定義すると補題 11.3 は次のような言い換えを持つこととなる。

11.6 系. 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 $\alpha C {}^{\mathrm{t}}\alpha$ の d 次元正規分布である。

だがこれ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。

11.7 定義. $\mathrm{Borel}(\mathbb{R}^d)$ 可測かつ $\lambda^{(d)}$ 可積分な関数 $f : \mathbb{R}^d \to \mathbb{C}$ の全体を $\mathcal{L}^1(\mathbb{R}^d)$ と表記する。$f \in \mathcal{L}^1(\mathbb{R}^d)$ に対して次の関数を f の *Fourier* **変換**(Fourier transform) という。

$$\widehat{f} : \mathbb{R}^d \to \mathbb{C},\ \xi \mapsto \int_{\mathbb{R}^d} f(x) e^{-\sqrt{-1}(\xi, x)} \lambda^{(d)}(dx).$$

11.8 補題. $f \in \mathcal{L}^1(\mathbb{R}^d)$ の Fourier 変換 $\widehat{f}$ は有界かつ連続である。

11.9 演習問題. 補題 11.8 を示せ。

11.10 注意. 実は $\lim_{\xi \to \infty} \widehat{f}(\xi) = 0$ が成り立つがこれ（*Riemann-Lebesgue* **の定理**）を示すには工夫が必要である。

記号

$t > 0$ と $x \in \mathbb{R}^d$ に対して $p(t, x) := \exp\{-\|x\|^2/2t\}/\sqrt{(2\pi t)^d}$ とかく。

11.11 補題. $t > 0$, $y \in \mathbb{R}^d$ と $f \in \mathcal{L}^1(\mathbb{R}^d)$ に対して

$$\frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \widehat{f}(\xi) \exp\{\sqrt{-1}(\xi, y) - t\|\xi\|^2/2\}\, \lambda^{(d)}(d\xi) = \int_{\mathbb{R}^d} p(t, x-y) f(x)\, \lambda^{(d)}(dx).$$

証明. 補題 9.14 より $\int_{\mathbb{R}^d} \exp\{\sqrt{-1}(\xi, y-x) - t\|\xi\|^2/2\}\, \lambda^{(d)}(d\xi) = (2\pi)^d p(t, x-y)$ が成り立つ。また関数 $\mathbb{R}^d \times \mathbb{R}^d \to \mathbb{C}$, $(x, \xi) \mapsto f(x) \exp\{\sqrt{-1}(\xi, y-x) - t\|\xi\|^2/2\}$ は $\lambda^{(d)} \otimes \lambda^{(d)}$ 可積分である。従って Fubini の定理を適用して結論を得る。 $\square$

11.12 演習問題. $t > 0$ と $f \in \mathcal{L}^1(\mathbb{R}^d)$ に対して関数 $\mathbb{R}^d \to \mathbb{C}$, $y \mapsto \int_{\mathbb{R}^d} p(t, x-y) f(x)\, \lambda^{(d)}(dx)$ は有界かつ一様連続であることを示せ。

11.13 補題. $t > 0$, $f \in \mathcal{L}^1(\mathbb{R}^d)$ と $(\mathbb{R}^d, \mathrm{Borel}(\mathbb{R}^d))$ 上の確率測度 μ に対して

$$\frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \widehat{f}(\xi) \mathrm{char}(\mu, \xi) \exp\{-t\|\xi\|^2/2\}\, \lambda^{(d)}(d\xi) = \int_{\mathbb{R}^d} \Big(\int_{\mathbb{R}^d} p(t, x-y) f(x)\, \lambda^{(d)}(dx) \Big) \mu(dy).$$

証明. 補題 11.8 により関数 $\mathbb{R}^d \times \mathbb{R}^d \to \mathbb{C}$, $(\xi, y) \mapsto \widehat{f}(\xi)\exp\{\sqrt{-1}(\xi, y) - t\|\xi\|^2/2\}$ 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。

$$\begin{aligned}&\int_{\mathbb{R}^d} \widehat{f}(\xi)\mathrm{char}(\mu, \xi)\exp\{-t\|\xi\|^2/2\}\,\lambda^{(d)}(d\xi)\\ &= \int_{\mathbb{R}^d}\Big(\int_{\mathbb{R}^d} \widehat{f}(\xi)\exp\{\sqrt{-1}(\xi, y) - t\|\xi\|^2/2\}\,\lambda^{(d)}(d\xi)\Big)\mu(dy).\end{aligned}$$

ゆえに補題 11.11 により求める結論を得る。 □

当面の目的のためには演習問題 4.8(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。

11.14 補題. 有界かつ一様連続な関数 $f : \mathbb{R}^d \to \mathbb{C}$ に対して $y \mapsto \int_{\mathbb{R}^d} p(t, x-y)f(x)\,\lambda^{(d)}(dx)$ は $t \downarrow 0$ のとき $\mathbb{R}^d$ 上で f に一様収束する。

証明. 各 $\delta > 0$ に対して $\mathrm{amp}(\delta) := \sup_{x,z:\|x-z\|<\delta} |f(x) - f(z)|$ とおく。一様連続性により $\mathrm{amp}(\delta) < +\infty$ かつ $\delta \downarrow 0$ のときそれは 0 に収束する。一方 $M := \sup_{x\in\mathbb{R}^d} |f(x)|$ とおくと $M < +\infty$ である。各 $\delta > 0$, $t > 0$ と $y \in \mathbb{R}^d$ に対して

$$\begin{aligned}&\int_{\mathrm{Nbd}(y,\delta)} p(t, x-y)|f(x) - f(y)|\,\lambda^{(d)}(dx) \le \mathrm{amp}(\delta)\int_{\mathrm{Nbd}(y,\delta)} p(t, x-y)\,\lambda^{(d)}(dx) \le \mathrm{amp}(\delta),\\ &\int_{\mathrm{Nbd}(y,\delta)^c} \cdots\,\lambda^{(d)}(dx) \le 2M\int_{\mathrm{Nbd}(y,\delta)^c} p(t, x-y)\,\lambda^{(d)}(dx) = 2M\int_{\mathrm{Nbd}(0,\delta)^c} p(t, x)\,\lambda^{(d)}(dx).\end{aligned}$$

従って積分の加法性、三角不等式などにより

$$\Big|\int_{\mathbb{R}^d} p(t, x-y)f(x)\,\lambda^{(d)}(dx) - f(y)\Big| \le \mathrm{amp}(\delta) + 2M\int_{\mathrm{Nbd}(0,\delta)^c} p(t, x)\,\lambda^{(d)}(dx).$$

右辺は y に依存しないことに注意する。$\delta > 0$ 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とする。よって

$$\limsup_{t\downarrow 0}\sup_{y\in\mathbb{R}^d}\Big|\int_{\mathbb{R}^d} p(t, x-y)f(x)\,\lambda^{(d)}(dx) - f(y)\Big| \le \mathrm{amp}(\delta)\ \forall\delta > 0.$$

$\delta \downarrow 0$ のとき右辺は 0 に収束していたので結論に到達した。 □

11.15 演習問題. 各 $\delta > 0$ に対し $t \downarrow 0$ の極限で $\int_{\mathrm{Nbd}(0,\delta)^c} p(t, x)\,\lambda^{(d)}(dx)$ は 0 に収束することを示せ。

特性関数に基づく**測度の一意性定理**を以下にあげる。

11.16 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。各 $t>0$ と $f\in C_0(\mathbb{R}^d)$ に対して f は $\lambda^{(d)}$ 可積分なので補題 11.13 により次が成り立つ。

$$\int_{\mathbb{R}^d}\Big(\int_{\mathbb{R}^d}p(t,x-y)f(x)\,\lambda^{(d)}(dx)\Big)\mu(dy)=\int_{\mathbb{R}^d}\Big(\int_{\mathbb{R}^d}p(t,x-y)f(x)\,\lambda^{(d)}(dx)\Big)\nu(dy).$$

f は一様連続であるがこれと可積分性を示すのは演習問題とする。補題 11.14 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 7.15 によればこれは $\mu=\nu$ を意味している。 □

11.17 演習問題. $f\in C_0(\mathbb{R}^d)$ は $\lambda^{(d)}$ 可積分であることおよび一様連続であることを示せ。

11.18 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。

$$E[\exp\{\sum_{i=1}^{n}\sqrt{-1}\xi_iX_i\}]=\prod_{i=1}^{n}E[\exp\{\sqrt{-1}\xi_iX_i\}]\ \forall(\xi_1,\xi_2,\ldots,\xi_n)\in\mathbb{R}^n.$$

証明. $X_1,\,X_2,\,\ldots,\,X_n$ の結合分布を μ とし、またそれらの分布の直積測度を ν とすると定理 6.24 と系 8.23 により、与えられた条件は μ の特性関数と ν の特性関数が一致することを意味する。従って定理 11.16 を適用して $\mu=\nu$ を得る。 □

11.19 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。

11.20 演習問題. 例 11.19 を示せ。

12　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 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 5 節、6 節、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。

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。

12.1 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。

12.2 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。

12.3 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 $\mathbb{R}^{\mathbb{N}}$ の開区間 $(a(1) - \delta/2, a(1) + \delta/2) \times \cdots \times (a(n) - \delta/2, a(n) + \delta/2) \times \mathbb{R}^{\mathbb{N}}$ は a を含みさらに a の δ 近傍 $\mathrm{Nbd}(a, \delta)$ に包含されることになる。

(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。 □

12.4 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。

12.5 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 12.4 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。

以下の議論は定理 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ので、証明は演習問題とする。

12.6 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。

12.7 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。

12.8 演習問題. 定理 12.6 と系 12.7 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。

12.9 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。

12.10 注意. 系 12.7 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。

さてここからが問題である。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なければならないが、それは易しくはない。

12.11 定義. 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

$$\prod(\mathbb{N}, S_.) := S_i \text{ たちの直積集合}, \ \mathrm{proj}_i := \text{第 } i \text{ 座標への射影} \prod(\mathbb{N}, S_.) \to S_i.$$

$S_i = S \ \forall i \in \mathbb{N}$ が満たされる特殊な場合は $\prod(\mathbb{N}, S_.)$ を $S^{\mathbb{N}}$ と表記する。

$$\mathrm{Cyl}(\mathbb{N}, \mathcal{A}_.) := \{\textstyle\bigcap_{i \in I} \mathrm{proj}_i^{-1}(A_i)\,;\, I\ \mathbb{N} \text{ の有限部分集合}, A_i \in \mathcal{A}_i \text{ for } i \in I\}.$$

$\mathrm{Cyl}(\mathbb{N}, \mathcal{A}_.)$ の元を $\prod(\mathbb{N}, S_.)$ の $\mathcal{A}_.$-*cylinder* **集合**という。

12.12 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。

12.13 補題. $\prod(\mathbb{N}, S_.)$ 上の σ 加法族として $\sigma(\mathrm{Cyl}(\mathbb{N}, \mathcal{A}_.)) = \sigma(\mathrm{Cyl}(\mathbb{N}, \sigma(\mathcal{A}_.)))$ 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

$$\textstyle\bigcap_{i \in I} \mathrm{proj}_i^{-1}(A_i) \in \sigma(\mathrm{Cyl}(\mathbb{N}, \mathcal{A}_.))$$

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。 $\square$

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。

12.14 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。

12.15 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 12.5 と補題 12.13 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。 □

集合族 $\mathcal{C} = \mathrm{Cyl}(\mathbb{N}, \mathcal{B}.)$ は全体集合 $\prod(\mathbb{N}, S.)$ を含みかつ条件 $A \cap B \in \mathcal{C}\ \forall A \in \mathcal{C}\ \forall B \in \mathcal{C}$ を満たしている。従って μ, ν を $(\prod(\mathbb{N}, S.), \bigotimes(\mathbb{N}, \mathcal{B}.))$ 上の 確率測度 とするとき

$$\mu = \nu \Leftrightarrow \mu(A) = \nu(A)\ \forall A \in \mathrm{Cyl}(\mathbb{N}, \mathcal{B}.)$$

が成り立つことが定理 7.11(i) より導ける。

12.16 定義. $\bigcap_{i\in I} \mathrm{proj}_i^{-1}(A_i)$ という表現をもつ $\prod(\mathbb{N}, S.)$ の $\mathcal{B}.$-cylinder 集合に対して

$$\bigotimes(\mathbb{N}, \mu.)(\textstyle\bigcap_{i\in I} \mathrm{proj}_i^{-1}(A_i)) = \prod_{i\in I} \mu_i(A_i)$$

を満たす $(\prod(\mathbb{N}, S.), \bigotimes(\mathbb{N}, \mathcal{B}.))$ 上の測度 $\bigotimes(\mathbb{N}, \mu.)$ を**直積測度**と呼ぶ。

12.17 定理. 直積測度 $\bigotimes(\mathbb{N}, \mu.)$ が存在しかつ一意である。

証明のうちで一意性については説明済みである。

12.18 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。

$$\mathcal{L}((X_1, X_2, \ldots, X_n, \ldots), \cdot) = \bigotimes(\mathbb{N}, \mathcal{L}(X., \cdot)).$$

定理 12.17 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。

12.19 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。 □

12.20 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。

12.21 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。

12.22 演習問題. 補題 12.21 を示せ。

12.23 注意. 補題 12.21 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。

12.24 例. 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。

これから無限直積測度の存在について述べるが、記号がかなり複雑なうえ選択公理を適用しなければならないので、補題 12.26 までを確認してその後は割愛しても差し支えない。

12.25 定義. S の部分集合の族 $\mathcal{C}$ が**有限加法族**であるとは

$$\emptyset \in \mathcal{C};\ A \in \mathcal{C} \Rightarrow A^{\mathrm{c}} \in \mathcal{C};\ A, B \in \mathcal{C} \Rightarrow A \cup B \in \mathcal{C}.$$

有限測度の場合 *Hopf* **の拡張定理**は次の形で適用されることが多い。

12.26 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 が σ 加法的であることは次と同値である。

$$A_n \in \mathcal{C}\ \forall n \in \mathbb{N},\ A_n \supset A_{n+1}\ \forall n \in \mathbb{N},\ \inf_{n\in\mathbb{N}} m(A_n) > 0 \Rightarrow \bigcap_{n=1}^{\infty} A_n \neq \emptyset.$$

12.27 注意. 補題 12.26 の条件が成り立てば、Hopf の拡張定理と定理 7.11 により m は σ 加法族 $\sigma(\mathcal{C})$ 上の測度へ一意的に拡張される。

証明. 与えられた条件から σ 加法性を導こう。そこで

$$B_n \in \mathcal{C}\ \forall n \in \mathbb{N},\ B_n \cap B_k = \emptyset\ n \neq k,\ A := \bigcup_{n=1}^{\infty} B_n \in \mathcal{C}$$

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。 □

12.28 演習問題. 補題 12.26 で与えられた条件が必要であることを示せ。

12.29 定義. $k \in \mathbb{N}$ と $n \leq k$ なる $n \in \mathbb{N}$ に対して

$$\mathrm{proj}_{\leq k} := \text{第 } k \text{ 座標までへの射影 } \prod(\mathbb{N}, S.) \to \prod(\mathbb{N}_{\leq k}, S.),$$
$$\mathrm{proj}_{\leq n, \leq k} := \text{第 } n \text{ 座標までへの射影 } \prod(\mathbb{N}_{\leq k}, S.) \to \prod(\mathbb{N}_{\leq n}, S.).$$

また $\prod(\mathbb{N}_{(n,k]}, S.)$ 上の σ 加法族 $\bigvee_{i\in\mathbb{N}:n<i\le k}\mathrm{proj}_i^*\mathcal{B}_i$ を $\bigotimes(\mathbb{N}_{(n,k]}, \mathcal{B}.)$ と書き、その上の測度であって次を満たす唯一のものを $\bigotimes(\mathbb{N}_{(n,k]}, \mu.)$ と表記する。

$$\bigotimes(\mathbb{N}_{(n,k]}, \mu.)(\bigcap_{i\in\mathbb{N}:n<i\le k}\mathrm{proj}_i^{-1}(A_i)) = \prod_{i\in\mathbb{N}:n<i\le k}\mu_i(A_i)\ \forall A_i \in \mathcal{B}_i.$$

このとき $\prod(\mathbb{N}_{\le k}, S.)$ 上の σ 加法族として $\bigotimes(\mathbb{N}_{\le n}, \mathcal{B}.) \otimes \bigotimes(\mathbb{N}_{(n,k]}, \mathcal{B}.)$ は $\bigotimes(\mathbb{N}_{\le k}, \mathcal{B}.)$ に一致する、また測度として $\bigotimes(\mathbb{N}_{\le n}, \mu.) \otimes \bigotimes(\mathbb{N}_{(n,k]}, \mu.)$ は $\bigotimes(\mathbb{N}_{\le k}, \mu.)$ に一致する。

12.30 補題. (i) $n \in \mathbb{N}$ とし、ν を $\bigotimes(\mathbb{N}_{\le n}, \mathcal{B}.)$ 上の確率測度とするとき

$$(\mathrm{proj}_{\le k,\le k+1})_*(\nu \otimes \bigotimes(\mathbb{N}_{(n,k+1]}, \mu.)) = \nu \otimes \bigotimes(\mathbb{N}_{(n,k]}, \mu.)\ \forall k \in \mathbb{N}_{\ge n}.$$

ここで $k = n$ のときは右辺は ν を表すものとする。
(ii) $\mathrm{proj}_{\le k} = \mathrm{proj}_{\le k,\le k+1} \circ \mathrm{proj}_{\le k+1}$, $\mathrm{proj}_{\le k+1}(\mathrm{proj}_{\le k+1}^{-1}(A)) = A\ \forall A \subset \prod(\mathbb{N}_{\le k+1}, S.)$.

12.31 演習問題. 補題 12.30 を示せ。

12.32 系. $\bigcup_{n=1}^\infty \mathrm{proj}_{\le n}^* \bigotimes(\mathbb{N}_{\le n}, \mathcal{B}.)$ は $\prod(\mathbb{N}, S.)$ 上の有限加法族である。またその上の有限加法的測度 m であって次を満たすものが唯一存在する。

$$m(\mathrm{proj}_{\le n}^{-1}(A)) = \bigotimes(\mathbb{N}_{\le n}, \mu.)(A)\ \forall n \in \mathbb{N}, \forall A \in \bigotimes(\mathbb{N}_{\le n}, \mathcal{B}.).$$

あとは系 12.32 で述べた有限加法的測度の σ 加法性を補題 12.26 を適用して証明すればよい。系 12.35 が目標であるがここからがなかなか厄介である。

12.33 補題. $\alpha \in \mathbb{R}$, $n \in \mathbb{N}$ とし、ν を $\bigotimes(\mathbb{N}_{\le n}, \mathcal{B}.)$ 上の確率測度とする。各 $k \in \mathbb{N}_{\ge n}$ に対して $A_k \in \bigotimes(\mathbb{N}_{\le k}, \mathcal{B}.)$ が与えられ

$$\mathrm{proj}_{\le k}^{-1}(A_k) \supset \mathrm{proj}_{\le k+1}^{-1}(A_{k+1})\ \forall k \in \mathbb{N}_{\ge n},$$
$$\inf_{k\in\mathbb{N}:k\ge n}(\nu \otimes \bigotimes(\mathbb{N}_{(n,k]}, \mu.))(A_k) > \alpha$$

が成り立つなら

$$\nu(\{x \in \prod(\mathbb{N}_{\le n}, S.) : (\delta_x \otimes \bigotimes(\mathbb{N}_{(n,k]}, \mu.))(A_k) \ge \alpha\ \forall k \in \mathbb{N}_{\ge n}\}) > 0.$$

証明. 軽装化のため $\mu_{n,k} := \bigotimes(\mathbb{N}_{(n,k]}, \mu.)$ とかき、$\{x \in \prod(\mathbb{N}_{\le n}, S.) : (\delta_x \otimes \mu_{n,k})(A_k) \ge \alpha\}$ を B_k とおく。定理 8.10(i) によれば $B_k \in \bigotimes(\mathbb{N}_{\le n}, \mathcal{B}.)$ である。一方

$$1_{A_k}(x,y) \le 1_{B_k}(x) + 1_{B_k^c}(x)1_{A_k}(x,y)\ \forall x \in \prod(\mathbb{N}_{\le n}, S.), \forall y \in \prod(\mathbb{N}_{(n,k]}, S.)$$

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 $\inf_{k\in\mathbb{N}:k\ge n}(\nu \otimes \mu_{n,k})(A_k) > \alpha$ から

$$\alpha < \inf_{k\in\mathbb{N}:k\ge n}(1-\alpha)\nu(B_k) + \alpha$$

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 $\mathrm{proj}^{-1}_{\leq k}(A_k) \supset \mathrm{proj}^{-1}_{\leq k+1}(A_{k+1})\ \forall k \in \mathbb{N}_{\geq n}$ を使う。補題 12.30(ii) を適用して

$$\begin{aligned}\mathrm{proj}^{-1}_{\leq k,\leq k+1}(A_k) &= \mathrm{proj}_{\leq k+1}(\mathrm{proj}^{-1}_{\leq k+1}(\mathrm{proj}^{-1}_{\leq k,\leq k+1}(A_k)))\\ &= \mathrm{proj}_{\leq k+1}(\mathrm{proj}^{-1}_{\leq k}(A_k))\\ &\supset \mathrm{proj}_{\leq k+1}(\mathrm{proj}^{-1}_{\leq k+1}(A_{k+1})) = A_{k+1}.\end{aligned}$$

従って補題 12.30(i) を適用して次を得る。

$$\begin{aligned}(\delta_x \otimes \mu_{n,k})(A_k) &= (\delta_x \otimes \mu_{n,k+1})(\mathrm{proj}^{-1}_{\leq k,\leq k+1}(A_k))\\ &\geq (\delta_x \otimes \mu_{n,k+1})(A_{k+1})\ \forall x \in \prod(\mathbb{N}_{\leq n}, S.).\end{aligned}$$

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。 □

12.34 補題. $\delta > 0$ とする。各 $k \in \mathbb{N}$ に対して $A_k \in \bigotimes(\mathbb{N}_{\leq k}, \mathcal{B}.)$ が与えられ

$$\mathrm{proj}^{-1}_{\leq k}(A_k) \supset \mathrm{proj}^{-1}_{\leq k+1}(A_{k+1})\ \forall k \in \mathbb{N},\ \inf_{k\in\mathbb{N}} \bigotimes(\mathbb{N}_{\leq k}, \mu.)(A_k) > \delta$$

が成り立つなら

$$\exists x \in \prod(\mathbb{N}, S.)\ \text{s.t.}\ (\delta_{\mathrm{proj}_{\leq n}(x)} \otimes \bigotimes(\mathbb{N}_{(n,k]}, \mu.))(A_k) \geq \delta/n\ \forall n \in \mathbb{N},\ \forall k \in \mathbb{N}_{\geq n}.$$

証明. 軽装化のため $\mu_{n,k} := \bigotimes(\mathbb{N}_{(n,k]}, \mu.)$ とおく。各 $m \in \mathbb{N}$ に対して

$$\exists x \in \prod(\mathbb{N}_{\leq m}, S.)\ \text{s.t.}\ \delta_{\mathrm{proj}_{\leq n}(x)} \otimes \mu_{n,k}(A_k) \geq \delta/n\ \forall n \in \mathbb{N}_{\leq m},\ \forall k \in \mathbb{N}_{\geq n}$$

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 12.33 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 0 でない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 $x \in \prod(\mathbb{N}_{\leq m}, S.)$ 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 12.33 を適用すればよい。結論を得るためには更に選択公理を使わねばならない。 □

12.35 系. 各 $k \in \mathbb{N}$ に対して $A_k \in \bigotimes(\mathbb{N}_{\leq k}, \mathcal{B}.)$ が与えられ

$$\mathrm{proj}^{-1}_{\leq k}(A_k) \supset \mathrm{proj}^{-1}_{\leq k+1}(A_{k+1})\ \forall k \in \mathbb{N},\ \inf_{k\in\mathbb{N}} \bigotimes(\mathbb{N}_{\leq k}, \mu.)(A_k) > 0$$

が成り立つなら $\bigcap_{k=1}^{\infty} \mathrm{proj}^{-1}_{\leq k}(A_k) \neq \emptyset$ である。

証明. $\delta = \inf_{k\in\mathbb{N}} \bigotimes(\mathbb{N}_{\leq k}, \mu.)(A_k)/2 > 0$ に対して $x \in \prod(\mathbb{N}, S.)$ を補題 12.34 で述べられたものとする。このとき $\delta_{\mathrm{proj}_{\leq n}(x)}(A_n) \geq \delta/n\ \forall n \in \mathbb{N}$ である。従って $\mathrm{proj}_{\leq n}(x) \in A_n\ \forall n \in \mathbb{N}$ すなわち $x \in \mathrm{proj}^{-1}_{\leq n}(A_n)\ \forall n \in \mathbb{N}$ である。 □

13 ランダムウォークと中心極限定理

定理 12.17 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。これを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と解釈するのが通常である。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。

13.1 定義. $\mathbb{R}^d$ 値の確率変数の無限系 $Y_k\ k \in \mathbb{N}\cup\{0\}$ を $\mathbb{R}^d$ 値の（離散時間）**確率過程**(stochastic process) という。$\mathbb{R}^d$ 値確率過程 $Y_\cdot$ について $Y_0, Y_1 - Y_0, \ldots, Y_n - Y_{n-1}, \ldots$ が独立かつ $Y_k - Y_{k-1}$ $k \in \mathbb{N}$ が同じ分布に従うときそれを $\mathbb{R}^d$ 上の**ランダムウォーク**(random walk) という。

前提

$Y_\cdot$ は $\mathbb{R}$ 上のランダムウォーク, $Y_0 = 0$, $E[|Y_1|^2] < +\infty$, $m := E[Y_1]$, $v := \mathrm{Var}[Y_1] > 0$.

13.2 補題. $E[Y_n] = mn$, $\mathrm{Var}[Y_n] = vn\ \forall n \in \mathbb{N}$.

13.3 演習問題. 補題 13.2 を示せ。

即ち各 $n \in \mathbb{N}$ に対して $(Y_n - mn)/\sqrt{vn}$ は平均 0 分散 1 である。一方、定理 11.18 によればその特性関数は次で与えられる。

$$\mathrm{char}((Y_n - mn)/\sqrt{vn}, \xi) = \bigl(\mathrm{char}(Y_1, \xi/\sqrt{vn}) \exp\{-\sqrt{-1}\xi m/\sqrt{vn}\}\bigr)^n. \tag{13.4}$$

13.5 例. $P(Y_1 = 1) = 1/2$, $P(Y_1 = -1) = 1/2$ の時は $m := E[Y_1] = 0$, $v := \mathrm{Var}[Y_1] = 1$, $\mathrm{char}(Y_1, \xi) = \cos\xi$ である。従って $\mathrm{char}(Y_n/\sqrt{n}, \xi) = (\cos(\xi/\sqrt{n}))^n$ である。これは関数 $\xi \mapsto \exp\{-\xi^2/2\}$ に各点収束する。

13.6 演習問題. 数列 $(\cos(\xi/\sqrt{n}))^n$ は $\exp\{-\xi^2/2\}$ に収束すること示せ。ヒント 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 に収束することを示せ。

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 11.16 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 $Y_n/\sqrt{n}$ 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が一般的に成り立つことを示そう。

13.7 補題. 各 $\xi \in \mathbb{R}$ に対して $\mathrm{char}((Y_n - mn)/\sqrt{vn}, \xi)$ は $\exp\{-\xi^2/2\}$ に収束する。

証明. 一般性を失うことなしに $E[Y_1] = 0$, $\mathrm{Var}[Y_1] = 1$ 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。

$$\mathrm{char}(Y_1, \xi) - 1 = E[\sqrt{-1}\xi Y_1 - \int_{(0,1)} (1-t)\exp\{t\sqrt{-1}\xi Y_1\}\,\lambda(dt)\,\xi^2 {Y_1}^2].$$

さて $E[Y_1]=0$ である。また $|(1-t)e^{t\sqrt{-1}\xi Y_1}{Y_1}^2| \le |Y_1|^2$ 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

$$n\bigl(\mathrm{char}(Y_1,\xi/\sqrt{n})-1\bigr) = -\int_{(0,1)\times\Omega}(1-t)\exp\{t\sqrt{-1}\xi Y_1(\omega)/\sqrt{n}\}Y_1(\omega)^2\,(\lambda\otimes P)(dt,d\omega)\,\xi^2.$$

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ことを記憶にとどめてほしい。関数列 $(0,1)\times\Omega\to\mathbb{C}$, $(t,\omega)\mapsto(1-t)\exp\{t\sqrt{-1}\xi Y_1(\omega)/\sqrt{n}\}Y_1(\omega)^2$ は関数 $(t,\omega)\mapsto(1-t)Y_1(\omega)^2$ に各点収束している。他方 $|Y_1|^2$ が $\lambda\otimes P$ について可積分な優関数の役割を果たす。よって Lebesgue の収束定理を適用すると各 $\xi\in\mathbb{R}$ に対して

$n\bigl(\mathrm{char}(Y_1,\xi/\sqrt{n})-1\bigr)$ は $-\int_{(0,1)\times\Omega}(1-t)Y_1(\omega)^2\,(\lambda\otimes P)(dt,d\omega)\xi^2=-\xi^2/2$ に収束する

ことが導ける。例 13.5 に引き続く演習問題でも述べたように $na_n\to\alpha$ が成り立てば $(1+a_n)^n$ は e^α に収束する。故に $\bigl(\mathrm{char}(Y_1,\xi/\sqrt{n})\bigr)^n$ は $e^{-\xi^2/2}$ に収束する。(13.4) により結論を得る。 □

13.8 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 6.17 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 13.7 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 11.18 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。

13.9 補題. 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。特性関数の列 $\mathrm{char}(\mu_n,\cdot)$ が $\mathrm{char}(\mu,\cdot)$ に各点収束するなら任意の $f\in C_0(\mathbb{R}^d)$ に対して $\int_{\mathbb{R}^d}f\,\mu_n$ は $\int_{\mathbb{R}^d}f\,\mu$ に収束する。

証明. 関数 p を補題 11.11 で現れたものとし、$f \in C_0(\mathbb{R}^d)$ を固定して関数 $(0,+\infty)\times\mathbb{R}^d \to \mathbb{R}$, $(t,y) \mapsto \int_{\mathbb{R}^d} p(t,x-y)f(x)\,\lambda^{(d)}(dx)$ を u と書く。補題 11.13 より $t>0$ と $(\mathbb{R}^d,\mathrm{Borel}(\mathbb{R}^d))$ 上の確率測度 ν に対して次が満たされる。

$$\frac{1}{(2\pi)^d}\int_{\mathbb{R}^d}\widehat{f}(\xi)\mathrm{char}(\nu,\xi)\exp\{-t\|\xi\|^2/2\}\,\lambda^{(d)}(d\xi)=\int_{\mathbb{R}^d}u(t,y)\nu(dy).$$

補題 11.8 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 11.2 により $|\mathrm{char}(\nu,\xi)|\le 1$ である。さて関数列 $\mathbb{R}^d\to\mathbb{C}$, $\xi\mapsto\widehat{f}(\xi)\mathrm{char}(\mu_n,\xi)\exp\{-t\|\xi\|^2/2\}$ は $\xi\mapsto\widehat{f}(\xi)\mathrm{char}(\mu,\xi)\exp\{-t\|\xi\|^2/2\}$ に収束する。他方 $\xi\mapsto M\exp\{-t\|\xi\|^2/2\}$ が $\lambda^{(d)}$ について可積分な優関数の役割を果たす。よって Lebesgue の収束定理を適用すると

$$\text{各 } t>0 \text{ に対して } \int_{\mathbb{R}^d}u(t,y)\mu_n(dy) \text{ は } \int_{\mathbb{R}^d}u(t,y)\mu(dy) \text{ に収束する}$$

ことがわかる。次に $\int_{\mathbb{R}^d} f\mu_n$ を次のように評価する。

$$\int_{\mathbb{R}^d}u(t,y)\mu_n(dy)+\int_{\mathbb{R}^d}(f(y)-u(t,y))\mu_n(dy)\le\int_{\mathbb{R}^d}u(t,y)\mu_n(dy)+\sup_{y\in\mathbb{R}^d}|f(y)-u(t,y)|.$$

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 11.14 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。 □

13.10 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。

13.11 補題. $(\mathbb{R}^d,\mathrm{Borel}(\mathbb{R}^d))$ 上の確率測度の列 μ_n $(\mathbb{R}^d,\mathrm{Borel}(\mathbb{R}^d))$ 上の確率測度 μ に漠収束するなら任意の有界な $f\in C(\mathbb{R}^d)$ に対して $\int_{\mathbb{R}^d}f\,\mu_n$ は $\int_{\mathbb{R}^d}f\,\mu$ に収束する。

証明. $g\in C_0(\mathbb{R}^d)$, $g(x)\le 1\ \forall x\in\mathbb{R}^d$ とする。

$$\int_{\mathbb{R}^d}f\,\mu_n=\int_{\mathbb{R}^d}fg\,\mu_n+\int_{\mathbb{R}^d}f(1-g)\,\mu_n\le\int_{\mathbb{R}^d}fg\,\mu_n+M\int_{\mathbb{R}^d}(1-g)\,\mu_n.$$

ここで $M:=\sup_{y\in\mathbb{R}^d}|f(y)|<+\infty$ である。さて $fg\in C_0(\mathbb{R}^d)$ かつ $\int_{\mathbb{R}^d}(1-g)\,\mu_n=1-\int_{\mathbb{R}^d}g\,\mu_n$ であるから

$$\limsup_{n\to\infty}\int_{\mathbb{R}^d}f\,\mu_n\le\int_{\mathbb{R}^d}fg\,\mu+M\int_{\mathbb{R}^d}(1-g)\,\mu.$$

他方 $g_k\le g_{k+1}\ \forall k\in\mathbb{N}$ かつ $\sup_{k\in\mathbb{N}}g_k(x)=1\ \forall x\in\mathbb{R}^d$ を満たす $C_0(\mathbb{R}^d)$ の列 g_k が存在する（演習問題 13.12）。このとき Lebesgue の収束定理を適用して $\int_{\mathbb{R}^d}fg_k\,\mu$ が $\int_{\mathbb{R}^d}f\,\mu$ に収束し、また単調収束定理を適用して $\int_{\mathbb{R}^d}(1-g_k)\,\mu$ が 0 に収束することがわかる。従って

$$\limsup_{n\to\infty}\int_{\mathbb{R}^d}f\,\mu_n\le\int_{\mathbb{R}^d}f\,\mu.$$

$-f$ 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。 □

13.12 演習問題. $\mathbb{R}^d$ の開集合 A に対して $0 \le g_k \le g_{k+1}$ $\forall k \in \mathbb{N}$, $\sup_{k\in\mathbb{N}} g_k(x) = 1_A(x)$ $\forall x \in \mathbb{R}^d$ を満たしかつ $C_0(\mathbb{R}^d)$ に属する関数の列 g_k が存在することを示せ。

13.13 定義. $(\mathbb{R}^d, \mathrm{Borel}(\mathbb{R}^d))$ 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**(weak convergence) するとは任意の有界な $f \in C(\mathbb{R}^d)$ に対して $\int_{\mathbb{R}^d} f\,\mu_n$ は $\int_{\mathbb{R}^d} f\,\mu$ に収束することをいう。また $\mathbb{R}^d$ 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。

13.14 定理. 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。このとき (i) 特性関数の列 $\mathrm{char}(\mu_n, \cdot)$ が $\mathrm{char}(\mu, \cdot)$ に各点収束する、(ii) μ_n が μ に漠収束する、(iii) μ_n が μ に弱収束する、は互いに同値である。

証明. 補題 13.9 と補題 13.11 により (i) $\Rightarrow$ (ii) $\Rightarrow$ (iii) 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 (iii) 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 (iii) $\Rightarrow$ (i) という関係も成り立つ。 □

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。

13.15 定理. $(\mathbb{R}^d, \mathrm{Borel}(\mathbb{R}^d))$ 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 に弱収束することと任意の $\mathbb{R}^d$ の開集合 A に対して $\liminf_{n\to\infty} \mu_n(A) \ge \mu(A)$ であるのは同値である。

証明. μ_n が μ に弱収束するとしよう。$\mathbb{R}^d$ の開集合 A と $g_k \le g_{k+1}$ $\forall k \in \mathbb{N}$ 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 に対して

$$\liminf_{n\to\infty} \mu_n(A) \ge \liminf_{n\to\infty} \int_{\mathbb{R}^d} g_k\,\mu_n = \int_{\mathbb{R}^d} g_k\,\mu \ \forall k \in \mathbb{N}.$$

単調収束定理により右辺は $\mu(A)$ に収束する。

逆に $\liminf_{n\to\infty} \mu_n(A) \ge \mu(A)$ $\forall A$ 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 $\liminf_{n\to\infty} \mu_n(A(\alpha)) \ge \mu(A(\alpha))$ が従う。ここで

$$g_k(x) := \frac{\#\{i \in \mathbb{N} : i < 2^k f(x)\}}{2^k} = \frac{1}{2^k}\sum_{i=1}^{\infty} 1_{A(i/2^k)}(x)$$

と非負値 Borel 関数列 g_k を定義する。級数和に Fatou の補題を適用して

$$\liminf_{n\to\infty} \int_{\mathbb{R}^d} g_k\,\mu_n \ge \frac{1}{2^k}\sum_{i=1}^{\infty} \liminf_{n\to\infty} \mu_n(A(i/2^k)) \ge \frac{1}{2^k}\sum_{i=1}^{\infty} \mu(A(i/2^k)) = \int_{\mathbb{R}^d} g_k\,\mu$$

を得る。さて $g_k \le g_{k+1}$ $\forall k \in \mathbb{N}$ 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

$$\liminf_{n\to\infty} \int_{\mathbb{R}^d} f\,\mu_n \ge \int_{\mathbb{R}^d} f\,\mu.$$

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

$$\limsup_{n\to\infty}\int_{\mathbb{R}^d} f\,\mu_n \leq \int_{\mathbb{R}^d} f\,\mu$$

を導くことができる。 □

13.16 系. $(\mathbb{R}^d, \mathrm{Borel}(\mathbb{R}^d))$ 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 に弱収束することと任意の $\mathbb{R}^d$ の閉集合 A に対して $\limsup_{n\to\infty}\mu_n(A) \leq \mu(A)$ であるのは同値である。

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 13.15 から直ちに導くことができる。 □

標準正規分布の分布関数を Φ と表記していたことを思い出そう。

13.17 定理. 確率変数列 $(Y_n - mn)/\sqrt{vn}$ は標準正規分布に法則収束する。とくに $a, b \in \mathbb{R}$ $a < b$ に対して確率 $P(a < (Y_n - mn)/\sqrt{vn} < b)$ は $\Phi(b) - \Phi(a)$ に収束する。

証明. 法則収束は補題 13.7 で述べられている。つぎに定理 13.15 と系 13.16 により

$$\begin{aligned}\Phi(b) - \Phi(a) &\leq \liminf_{n\to\infty} P(a < (Y_n - mn)/\sqrt{vn} < b) \\ &\leq \limsup_{n\to\infty} P(a \leq (Y_n - mn)/\sqrt{vn} \leq b) \leq \Phi(b) - \Phi(a)\end{aligned}$$

したがって 2 番目の結論も得られた。 □

13.18 例. Y_1 がパラメータ 1 の指数分布に従うとする。このとき

$$E[Y_1] = 1, E[|Y_1|^2] = 1$$

指数 1 の gamma 分布とはパラメータ 1 の指数分布に他ならない。よって例 10.21 を適用すると、帰納法により Y_n は指数 n の gamma 分布であることがわかる。

$$P(0 < (Y_n - n)/\sqrt{n} < 1) = P(n < Y_n < \sqrt{n} + n) = \frac{1}{\Gamma(n)}\int_{(n,\sqrt{n}+n)} y^{n-1}e^{-y}\,\lambda(dy).$$

ここで Γ は gamma 関数である。右辺の積分に補題 3.14 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

$$\frac{\Gamma(n)}{n^{n-1}e^{-n}\sqrt{n}} = \frac{\int_{(0,1)}(1 + y/\sqrt{n})^{n-1}e^{-\sqrt{n}y}\,\lambda(dy)}{P(0 < (Y_n - n)/\sqrt{n} < 1)}.$$

定理 13.17 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。従って定理 13.17 と Lebesgue の収束定理により

$$\lim_{n\to\infty}\frac{\Gamma(n)}{n^{n-1}e^{-n}\sqrt{n}}=\frac{\int_{(0,1)}e^{-y^2/2}\,\lambda(dy)}{\Phi(1)-\Phi(0)}=\sqrt{2\pi}.$$

これはまさに *Stirling* **の公式**である。

定理 13.17 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。多次元でも対応できる形に少し拡張しておこう。

前提

$Y.$ は $\mathbb{R}^d$ 上のランダムウォーク, $Y_0=0$, $E[\|Y_1\|^2]<+\infty$, $E[Y_1]=0$.

記号

Y を $\mathbb{R}^d$ 値確率変数とする。$\xi,\eta\in\mathbb{R}^d$ に対し $\mathrm{Var}[Y](\xi,\eta):=\mathrm{Cov}[(\xi,Y),(\eta,Y)]$.

このとき $\mathrm{Var}[Y]:\mathbb{R}^d\times\mathbb{R}^d\to\mathbb{R}$ 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。

13.19 定理. 確率変数列 $Y_n/\sqrt{n}$ は平均 0 共分散 $\mathrm{Var}[Y_1]$ 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 $\mathrm{char}(Y_n/\sqrt{n},\xi)$ が $\exp\{-\mathrm{Var}[Y_1](\xi,\xi)/2\}$ に収束することを示せばよいが、補題 13.7 の証明を今の状況に焼き直すだけでよいので省略する。 □

13.20 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。

14 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、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、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

$$\mathcal{F}^{\otimes n} := \overbrace{\mathcal{F} \otimes \cdots \otimes \mathcal{F}}^{n},\ P^{\otimes n} := \overbrace{P \otimes \cdots \otimes P}^{n},\ \mathrm{diag}_n(\omega) := (\overbrace{\omega, \ldots, \omega}^{n}).$$

14.1 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。

14.2 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。

$$P(\{\omega \in \Omega : \mathrm{diag}_n(\omega) \in A\}) = P^{\otimes n}(A)\ \forall A \in \bigotimes(\mathbb{N}_{\leq n}, \mathcal{G}.).$$

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。

14.3 定理. Ω の部分集合の族 $\mathcal{C}_1, \ldots, \mathcal{C}_n, \ldots$ はそれぞれ次を満たすとする。

$$\mathcal{C}_i \subset \mathcal{F};\ A, B \in \mathcal{C}_i \Rightarrow A \cap B \in \mathcal{C}_i.$$

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 $\mathcal{C}_i$ についての条件から集合族 $\mathrm{Cyl}(\mathbb{N}, \mathcal{C}.)$ は条件

$$A, B \in \mathrm{Cyl}(\mathbb{N}, \mathcal{C}.) \Rightarrow A \cap B \in \mathrm{Cyl}(\mathbb{N}, \mathcal{C}.)$$

を満たす。したがって定理 7.11(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 12.13 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。 □

14.4 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。

14.5 演習問題. 系 14.4 を示せ。

14.6 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 14.3 から従う。 □

14.7 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 14.3 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。

14.8 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。

$$\forall A\in\mathcal{G}\ \exists B\in\mathcal{H}\ \text{s.t.}\ P(A\cap B^{\mathrm{c}})=0,\ P(B\cap A^{\mathrm{c}})=0.$$

$\mathcal{G}\subset\mathcal{H} \bmod P$ かつ $\mathcal{H}\subset\mathcal{G} \bmod P$ であることを $\mathcal{G}=\mathcal{H} \bmod P$ と書く。

14.9 定理. 部分 σ 加法族の系 $\mathcal{G}_1,\ \mathcal{G}_2,\ \ldots,\ \mathcal{G}_n,\ \ldots$ および系 $\mathcal{H}_1,\ \mathcal{H}_2,\ \ldots,\ \mathcal{H}_n,\ \ldots$ 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。

14.10 演習問題. 定理 14.3 を使って定理 14.9 を示せ。

14.11 系. 確率変数系 $X_1,\ \ldots,\ X_n,\ \ldots$ および系 $Y_1,\ \ldots,\ Y_n,\ \ldots$ について $P(X_i=Y_i)=1$ $\forall i\in\mathbb{N}$ 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。

14.12 演習問題. 系 14.11 を示せ。

14.13 定理. 部分 σ 加法族の系 $\mathcal{G}_1,\ \mathcal{G}_2,\ \ldots,\ \mathcal{G}_n,\ \ldots$ 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 14.3 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。$\mathcal{C}_1$, $\mathcal{C}_2$ ともに条件 $A,B\in\mathcal{C}_i \Rightarrow A\cap B\in\mathcal{C}_i$ を満たすので、定理 14.3 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。 □

14.14 系. $X_1, X_2, \ldots, X_n, \ldots$ 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 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。

14.15 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 (6.16) 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 14.3 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。

14.16 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 2.3 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。 □

14.17 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。 □

14.18 補題. $a, b \in \mathbb{Z}$, $a \neq b$ とする。このとき次が成り立つ。

$$\begin{aligned}\lambda(\tau(x,a,\cdot) < \tau(x,b,\cdot)) = p\lambda(\tau(x+1,a,\cdot) < \tau(x+1,b,\cdot)) \\ + (1-p)\lambda(\tau(x-1,a,\cdot) < \tau(x-1,b,\cdot))\ \forall x \in \mathbb{Z}, x \neq a, x \neq b.\end{aligned}$$

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 14.16 により $A(x) \in \mathrm{Borel}((0,1])$ である。補題 14.17 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 14.15 を適用すると

$$\lambda(A(x) \cap \{\xi_1 = 1\}) = \lambda(\varphi^{-1}(A(x+1)))\lambda(\{\xi_1 = 1\}).$$

例 6.13 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。 □

14.19 補題. $a, b \in \mathbb{Z}$, $a \neq b$ とする。このとき次が成り立つ。

$$\begin{aligned}&\lambda(\exists n \in \mathbb{N} \text{ s.t. } \eta_n(a,\cdot) = a, n < \tau(a,b,\cdot)) \\ &= p\lambda(\tau(a+1,a,\cdot) < \tau(a+1,b,\cdot)) + (1-p)\lambda(\tau(a-1,a,\cdot) < \tau(a-1,b,\cdot)).\end{aligned}$$

14.20 演習問題. 補題 14.19 を示せ

14.21 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

$$\lambda(\tau(x,a,\cdot) < \tau(x,b,\cdot)) = \begin{cases} \dfrac{(1/p-1)^{x-a} - (1/p-1)^{b-a}}{1-(1/p-1)^{b-a}} & p \neq 1/2 \\ \dfrac{b-x}{b-a} & p = 1/2 \end{cases}.$$

14.22 演習問題. 補題 14.18 を使って定理 14.21 を導け。

14.23 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 14.21 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。 □

14.24 定理. 各 $x \in \mathbb{R}$ に対して $\lambda(\exists n \in \mathbb{N} \text{ s.t. } \eta_n(x,\cdot) = x) = 2\min\{p, 1-p\}$.

14.25 演習問題. 補題 14.19 と定理 14.21 を使って定理 14.24 を導け。

15 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 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 14.24 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 $k \in \mathbb{N}$.

15.1 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 14.14 から従う。また後半についてはどちらも X_1 の分布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 $n \in \mathbb{N}$.
$\psi : \prod(\mathbb{N}, \mathbb{R}^d) \to \prod(\mathbb{N} \cup \{0\}, \mathbb{R}^d)$, $x \mapsto (0, x(1), x(1)+x(2), \dots, \sum_{i=1}^{k} x(i), \dots)$.

15.2 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 $\{Y_i - Y_n \neq 0\ \forall i > n\} = \{\psi \circ Z_{n+1} \in \Lambda\}\ \forall n \in \mathbb{N} \cup \{0\}$ が成り立つ。

証明. (i) は系 12.7 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。 □

15.3 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 15.2(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 15.1 によれば $\sigma\{X_i\,; i \in \mathbb{N}, i > n\}$ と独立である。他方、定理 12.6 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 15.1 と補題 15.2(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。 □

次は *Borel-Cantelli* **の補題**と呼ばれよく利用される。

15.4 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。 □

15.5 定理. (i) $P(\exists n \in \mathbb{N} \text{ s.t. } Y_n = 0) = 1 \Leftrightarrow \sum_{n=0}^{\infty} P(Y_n = 0) = +\infty$.
(ii) $P(\exists n \in \mathbb{N} \text{ s.t. } Y_n = 0) = 1 \Rightarrow P(\#\{n \in \mathbb{N} : Y_n = 0\} = +\infty) = 1$.
(iii) $P(\exists n \in \mathbb{N} \text{ s.t. } Y_n = 0) < 1 \Rightarrow P(\#\{n \in \mathbb{N} : Y_n = 0\} < +\infty) = 1$.

証明. $P(\exists i>0 \text{ s.t. } Y_i=0)<1$ 即ち $p:=P(Y_i\neq 0\ \forall i>0)>0$ なら補題 15.3 により

$$\sum_{n=0}^{\infty}P(Y_n=0)=\frac{1}{p}\sum_{n=0}^{\infty}P(Y_n=0,Y_i\neq 0\ \forall i>n).$$

右辺内の事象たちは互いに排反であり、それらの合併は $\{\#\{n\in\mathbb{N}:Y_n=0\}<+\infty\}$ である。従って右辺は $1/p$ を超えないので

$$(\star)\qquad P(\exists i>0 \text{ s.t. } Y_i=0)<1\Rightarrow\sum_{n=0}^{\infty}P(Y_n=0)<+\infty.$$

次に $\sum_{n=0}^{\infty}P(Y_n=0)<+\infty$ とする。$\{\#\{n\in\mathbb{N}:Y_n=0\}<+\infty\}=\bigcup_{n=1}^{\infty}\bigcap_{k\geq n}\{Y_k\neq 0\}$ であるが、その余事象の確率は補題 15.4 により 0 である。よって

$$(\star\star)\qquad \sum_{n=0}^{\infty}P(Y_n=0)<+\infty\Rightarrow P(\#\{n\in\mathbb{N}:Y_n=0\}<+\infty)=1.$$

ここで $(\star)$ と $(\star\star)$ を組み合わせて (iii) を得る。

$P(Y_i\neq 0\ \forall i>0)=0$ とする。補題 15.3 により $P(Y_n=0,Y_i\neq 0\ \forall i>n)=0\ \forall n\in\mathbb{N}\cup\{0\}$ となる。すでに述べたようにこれら確率の和は $P(\#\{n\in\mathbb{N}:Y_n=0\}<+\infty)$ を表すので

$$P(\exists i>0 \text{ s.t. } Y_i=0)=1\Rightarrow P(\#\{n\in\mathbb{N}:Y_n=0\}<+\infty)=0.$$

これで (ii) が得られた。(ii) と $(\star\star)$ の対偶と組み合わせて

$$P(\exists i>0 \text{ s.t. } Y_i=0)=1\Rightarrow \textstyle\sum_{n=0}^{\infty}P(Y_n=0)=+\infty$$

が得られる。これと $(\star)$ をまとめたのが (i) である。 □

15.6 注意. いずれの場合も $P(Y_i\neq 0\ \forall i>0)=1/\sum_{n=0}^{\infty}P(Y_n=0)$ である。

15.7 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。

次の例の結論はもちろん定理 14.24 で述べたことと符合している。

15.8 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 11.18 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 $\binom{2m}{m} = \binom{-1/2}{m}(-4)^m$ for $m \in \mathbb{N} \cup \{0\}$ であるから $p \neq 1/2$ のとき

$$\sum_{n=0}^{\infty} P(Y_n = 0) = \sum_{m=0}^{\infty} \binom{-1/2}{m}(-4)^m p^m (1-p)^m = \frac{1}{\sqrt{1-4p(1-p)}} = \frac{1}{|1-2p|} < +\infty.$$

$p = 1/2$ のとき上の級数は発散する。ついでながら $1 - |1-2p| = 2\min\{p, 1-p\}$ である。 □

15.9 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。

15.10 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。

16 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 13 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。

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

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

証明. 非負値確率変数 $|X - E[X]|^2$ と ε^2 に対して補題 16.1 を適用する。 □

補題 16.1 の不等式を *Markov* **の不等式**、系 16.2 の不等式を *Chebyshev* **の不等式**という。ときには Markov の不等式も Chebyshev の不等式と呼ばれることがある。

記号

集合 S に対して $\mathrm{Pair}(S) := \{I \in \mathrm{Sbset}(S) : \#I = 2\}$ とおく。

16.3 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

$$E[\textstyle\prod_{i \in I} X_i] \leq C(I) \prod_{i \in I} \sqrt{E[{X_i}^2]}\ \forall I \in \mathrm{Pair}(\mathbb{N}_{\leq n})$$

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 $I \in \mathrm{Pair}(\mathbb{N}_{\le n})$ に対して次が成り立つ。

$$2E[\prod_{i\in I} X_i] \le 2C(I)\prod_{i\in I}\sqrt{E[{X_i}^2]} \le C(I)\sum_{i\in I} E[{X_i}^2].$$

2 番目の不等号は $2ab \le a^2+b^2$ による。I について $\mathrm{Pair}(\mathbb{N}_{\le n})$ 全域にわたって和をとると

$$E[\sum_{I\in\mathrm{Pair}(\mathbb{N}_{\le n})} 2\prod_{i\in I} X_i] \le \sum_{i=1}^{n}\sum_{I\in\mathrm{Pair}(\mathbb{N}_{\le n}):i\in I} C(I)E[|X_i|^2] \le R\sum_{i=1}^{n} E[|X_i|^2].$$

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。 □

16.4 定理. 2 乗可積分確率変数の列 $X_1, X_2, \dots, X_n, \dots$ と $m \in \mathbb{R}$, $v \in \mathbb{R}_{\ge 0}$, $N \in \mathbb{N}$ に対して

$$E[X_n] = m\ \forall n \in \mathbb{N},\ \mathrm{Var}[X_n] \le v\ \forall n \in \mathbb{N},\ \mathrm{Cov}[X_i, X_j] = 0 \text{ if } |i-j| \ge N$$

が成り立つなら

$$P(|\frac{1}{n}\sum_{k=1}^{n} X_k - m| \ge \varepsilon) \le \frac{(2N-1)v}{\varepsilon^2 n}\ \forall n \in \mathbb{N}, \forall \varepsilon \in \mathbb{R}_{>0}.$$

証明. 確率変数 $\sum_{k=1}^{n} X_k/n$ に対して系 16.2 を適用すると

$$P(|\frac{1}{n}\sum_{k=1}^{n} X_k - m| \ge \varepsilon) \le \frac{1}{\varepsilon^2}\mathrm{Var}[\frac{1}{n}\sum_{k=1}^{n} X_k] = \frac{1}{\varepsilon^2 n^2}E[|\sum_{k=1}^{n}(X_k - m)|^2].$$

ここで記号の簡単化のため $Y_i := X_i - m$ と書くと Schwarz の不等式より

$$E[Y_iY_j] \le \sqrt{E[Y_i^2]}\sqrt{E[Y_j^2]}$$

であり、また $|i-j| \ge N$ なら $E[Y_iY_j] = \mathrm{Cov}[X_i, X_j] = 0$ 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

$$E[\prod_{i\in I} Y_i] \le C(I)\prod_{i\in I}\sqrt{E[Y_i^2]}\ \forall I \in \mathrm{Pair}(\mathbb{N}_{\le n}) \text{ かつ } \sum_{I\in\mathrm{Pair}(\mathbb{N}_{\le n}):i\in I} C(I) \le 2N-2\ \forall i \in \mathbb{N}_{\le n}$$

が満たされる。よって補題 16.3 により

$$E[|\sum_{k=1}^{n}(X_k - m)|^2] \le (2N-1)\sum_{k=1}^{n}\mathrm{Var}[X_k] \le (2N-1)nv$$

を得るので結論が導ける。 □

16.5 定義. 確率変数の列 $X_1, X_2, \dots, X_n, \dots$ 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 $P(|X_n - Y| \ge \varepsilon)$ が 0 に収束することをいう。

16.6 演習問題. 2 乗可積分な確率変数の列 $X_1, X_2, \dots, X_n, \dots$ は**対ごとに独立**である、すなわち $i \ne j$ なら組 X_i, X_j は独立とする。このとき $\mathrm{Cov}[X_i, X_j] = 0$ if $i \ne j$ であることを示せ。

16.7 例. 2 乗可積分な確率変数の列 $X_1, X_2, \ldots, X_n, \ldots$ は対ごとに独立であるとする。このとき演習問題 16.6 で述べたように $\mathrm{Cov}[X_i, X_j] = 0$ if $i \neq j$ である。したがって系 $X_.$ が同じ分布に従うなら定理 16.4 により $\sum_{k=1}^{n} X_k/n$ は $E[X_1]$ に確率収束する。

上は広く**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。さて定理 16.4 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。

16.8 定義. $p \in \mathbb{R}_{>0}$ とする。確率変数の列 $X_1, X_2, \ldots, X_n, \ldots$ 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。

16.9 例. 2 乗可積分な確率変数の列 $X_1, X_2, \ldots, X_n, \ldots$ が $\mathrm{Cov}[X_i, X_j] = 0$ if $i \neq j$ を満たしかつ同じ平均と分散を持つなら $\sum_{k=1}^{n} X_k/n$ は $E[X_1]$ に 2 次平均収束する。

16.10 補題. $p \in \mathbb{R}_{>0}$ とする。確率変数の列 $X_1, X_2, \ldots, X_n, \ldots$ が確率変数 Y に p 次平均収束するなら、列 $X_.$ は Y に確率収束する。

16.11 演習問題. 補題 16.10 を示せ。

16.12 補題. $p \in \mathbb{R}_{\geq 1}$ とする。確率変数の列 $X_1, X_2, \ldots, X_n, \ldots$ 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 $X_.$ は Y に p 次平均収束する。

証明. 仮定よりある $K \in \mathbb{R}_{>0}$ に対して $\sup_{n\in\mathbb{N}} E[|X_n|^p\,; |X_n| \geq K] \leq 1$ が成り立つ。$n \in \mathbb{N}$ とする。$E[|X_n|^p]$ について積分範囲を $\{|X_n| < K\}$ と $\{|X_n| \geq K\}$ とに分割する。前者上で $|X_n|^p \leq K^p$ と評価することにより

$$E[|X_n|^p] = E[|X_n|^p\,; |X_n| \geq K] + K^p P(|X_n| < K) \leq 1 + K^p$$

を得る。$\lambda > 0, \varepsilon > 0, n \in \mathbb{N}$ とする。$E[\min\{|Y|^p, \lambda\}]$ について積分範囲を $\{|Y| < |X_n| + \varepsilon\}$ と $\{|Y| \geq |X_n| + \varepsilon\}$ とに分割する。前者上では $\min\{|Y|^p, \lambda\} \leq (|X_n| + \varepsilon)^p$ 後者上では $\min\{|Y|^p, \lambda\} \leq \lambda$ と評価することにより

$$E[\min\{|Y|^p, \lambda\}] \leq E[(|X_n| + \varepsilon)^p\,; |Y| < |X_n| + \varepsilon] + \lambda P(|Y| \geq |X_n| + \varepsilon)$$

を得る。更に $\{|Y| \geq |X_n| + \varepsilon\} \subset \{|X_n - Y| \geq \varepsilon\}$ であるから右辺第 2 項は $\lambda P(|X_n - Y| \geq \varepsilon)$ で抑えられる。他方右辺第 1 項は

$$E[(|X_n| + \varepsilon)^p] \leq (E[|X_n|^p]^{1/p} + \varepsilon)^p \leq ((1 + K^p)^{1/p} + \varepsilon)^p$$

で抑えられる。ここで L^p ノルムに関する Minkowski の不等式を適用した。従って

$$E[\min\{|Y|^p, \lambda\}] \leq ((1 + K^p)^{1/p} + \varepsilon)^p + \lambda P(|X_n - Y| \geq \varepsilon)$$

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

$$E[\min\{|Y|^p, \lambda\}] \leq 1 + K^p \ \forall \lambda \in \mathbb{R}_{>0}.$$

最後に $\lambda \to +\infty$ の極限を取る。単調収束定理により

$$E[|Y|^p] \leq 1 + K^p < +\infty.$$

簡単化のため $Z_n := X_n - Y$ と書く。$E[|Z_n|^p]$ について積分範囲を $\{|Z_n| < \varepsilon\}$ と $\{|Z_n| \geq \varepsilon\}$ とに分割する。前者上では $|Z_n|^p \leq \varepsilon^p$ 後者上では $|Z_n|^p \leq (|X_n| + |Y|)^p \leq 2^{p-1}(|X_n|^p + |Y|^p)$ と評価することにより

$$E[|Z_n|^p] \leq \varepsilon^p P(|Z_n| < \varepsilon) + 2^{p-1} E[|X_n|^p ; |Z_n| \geq \varepsilon] + 2^{p-1} E[|Y|^p ; |Z_n| \geq \varepsilon]$$

を得る。右辺第 1 項は ε^p で抑えられる。以下第 2 項、第 3 項に議論を絞る。$E[|X_n|^p ; |Z_n| \geq \varepsilon]$ について積分範囲を $\{|Z_n| \geq \varepsilon, |X_n| \geq \lambda\}$, $\{|Z_n| \geq \varepsilon, |X_n| < \lambda\}$ と分割する。前者上では積分範囲を $\{|X_n| \geq \lambda\}$ と広げ後者上では $|X_n|^p \leq \lambda^p$ と評価することにより

$$E[|X_n|^p ; |Z_n| \geq \varepsilon] \leq E[|X_n|^p ; |X_n| \geq \lambda] + \lambda^p P(|Z_n| \geq \varepsilon, |X_n| < \lambda)$$

を得る。右辺第 2 項は $\lambda^p P(|Z_n| \geq \varepsilon)$ で抑えられる。$E[|Y|^p ; |Z_n| \geq \varepsilon]$ についても同様にして

$$E[|Y|^p ; |Z_n| \geq \varepsilon] \leq E[|Y|^p ; |Y| \geq \lambda] + \lambda^p P(|Z_n| \geq \varepsilon)$$

を得る。従って

$$E[|Z_n|^p] \leq \varepsilon^p + 2^{p-1} \sup_{k \in \mathbb{N}} E[|X_k|^p ; |X_k| \geq \lambda] + 2^{p-1} E[|Y|^p ; |Y| \geq \lambda] + 2^p \lambda^p P(|Z_n| \geq \varepsilon)$$

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。 □

16.13 定義. 実確率変数の列 $X_1, X_2, \ldots, X_n, \ldots$ 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。

16.14 演習問題. $p \in \mathbb{R}_{>1}$ とする。確率変数の列 $X_1, X_2, \ldots, X_n, \ldots$ 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。

16.15 演習問題. 確率変数の列 $X_1, X_2, \ldots, X_n, \ldots$ が 0 に確率収束し $\sup_{n \in \mathbb{N}} E[|X_n|^2] < +\infty$ であるなら、列 $X_.$ は 0 に 1 次平均収束することを示せ。

16.16 補題. 2 乗可積分確率変数 $X_1, X_2, \ldots, X_n, \ldots$ と $C : \mathrm{Pair}(\mathbb{N}) \to \mathbb{R}_{\geq 0}$ に対して

$$E[\prod_{i \in I} X_i] \leq C(I) \prod_{i \in I} \sqrt{E[{X_i}^2]} \ \forall I \in \mathrm{Pair}(\mathbb{N}), \sup_{n \in \mathbb{N}} E[X_n^2] < +\infty, \sup_{i \in \mathbb{N}} \sum_{I \in \mathrm{Pair}(\mathbb{N}) : i \in I} C(I) < +\infty$$

が成り立つとする。このとき $A \in \mathcal{F}$ で $P(A) = 1$ を満たしかつすべての $\omega \in A$ に対して $\sum_{k=1}^n X_k(\omega)/n$ は 0 に収束するようなものが存在する。

証明. $m,n\in\mathbb{N}$, $m\le n$とする。ここで

$$R:=\sup_{i\in\mathbb{N}}\sum_{I\in\mathrm{Pair}(\mathbb{N}):i\in I}C(I),\ M:=\sup_{n\in\mathbb{N}}E[|X_n|^2]$$

として補題 16.3 を適用すると

$$(\star)\qquad E[|\sum_{k=m}^{n}X_k|^2]\le(1+R)\sum_{k=m}^{n}E[|X_k|^2]\le(1+R)M(n-m+1)$$

が得られる。従って記号簡略化のため $S_n:=\sum_{k=1}^{n}X_k$ と書くと

$$E[\sum_{n=1}^{\infty}\Big|\frac{S_{n^2}}{n^2}\Big|^2]=\sum_{n=1}^{\infty}\frac{1}{n^4}E[|\sum_{k=1}^{n^2}X_k|^2]\le\sum_{n=1}^{\infty}\frac{(1+R)Mn^2}{n^4}<+\infty.$$

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 $[x]:=\max\{n\in\mathbb{Z}:n\le x\}$ である。$\omega\in B_1$ のとき $(\sqrt{n}-1)^2/n\le[\sqrt{n}]^2/n\le 1$ であるから右辺第 1 項は 0 に収束する。他方 $(\star)$ により

$$E[\sum_{n=1}^{\infty}\Big|\frac{S_n-S_{[\sqrt{n}]^2}}{n}\Big|^2]=\sum_{n=1}^{\infty}\frac{1}{n^2}E[|\sum_{k=[\sqrt{n}]^2+1}^{n}X_k|^2]\le\sum_{n=1}^{\infty}\frac{(1+R)M(n-[\sqrt{n}]^2)}{n^2}$$

と評価されるが、$n-[\sqrt{n}]^2=(\sqrt{n}+[\sqrt{n}])(\sqrt{n}-[\sqrt{n}])\le 2\sqrt{n}$ なので右辺の級数は収束する。従って

$$B_2:=\{\omega\in\Omega:\sum_{n=1}^{\infty}\Big|\frac{S_n(\omega)-S_{[\sqrt{n}]^2}(\omega)}{n}\Big|^2<+\infty\}\in\mathcal{F}\text{ は }P(B_2)=1\text{ を満たす。}$$

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 $P(B_1\cap B_2)=1$ であるから $A:=B_1\cap B_2$ が求める集合である。 □

16.17 定義. 確率変数の列 $X_1,X_2,\ldots,X_n,\ldots$ が確率変数 Y に**概収束**(almost sure convergence) するとは $A\in\mathcal{F}$ 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ことをいう。

つぎは、その内容は補題 16.16 において既に議論済みのものであるが、適用しやすい形に整理してみたものである。

16.18 定理. 2 乗可積分確率変数の列 $X_1, X_2, \ldots, X_n, \ldots$ と $m \in \mathbb{R}$, $N \in \mathbb{N}$ に対して

$$E[X_n] = m \ \forall n \in \mathbb{N},\ \sup_{n\in\mathbb{N}} \mathrm{Var}[X_n] < +\infty,\ \mathrm{Cov}[X_i, X_j] = 0 \text{ if } |i-j| \geq N$$

が成り立つなら $\sum_{k=1}^n X_k/n$ は m に概収束する。

証明. $Y_i := X_i - m$ 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 16.4 の証明で論じたように

$$E[\prod_{i\in I} Y_i] \leq C(I) \prod_{i\in I} \sqrt{E[Y_i^2]} \ \forall I \in \mathrm{Pair}(\mathbb{N}) \text{ かつ } \sum_{I\in\mathrm{Pair}(\mathbb{N}): i\in I} C(I) \leq 2N-2 \ \forall i \in \mathbb{N}$$

が満たされる。また $E[Y_n^2] = \mathrm{Var}[X_n]$ である。よって補題 16.16 を適用して $\sum_{k=1}^n Y_k/n$ が 0 に概収束することを得る。 □

16.19 注意. 2 乗可積分な確率変数の列 $X_1, X_2, \ldots, X_n, \ldots$ は対ごとに独立であるとする。このとき演習問題 16.6 で述べたように $\mathrm{Cov}[X_i, X_j] = 0$ if $i \neq j$ である。したがって系 $X.$ が同じ分布に従うなら定理 16.18 の前提は満たされる。

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。

16.20 例. $\mathbb{R}$ 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。

16.21 補題. 確率変数の列 $X_1, X_2, \ldots, X_n, \ldots$ が確率変数 Y に概収束するなら、列 $X.$ は確率変数 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

$$\liminf_{n\to\infty} P(|X_n - Y| < \varepsilon) \geq E[\liminf_{n\to\infty} 1_{A(n)}] \geq P(A) = 1 \ \forall \varepsilon \in \mathbb{R}_{>0}.$$

したがって各 $\varepsilon \in \mathbb{R}_{>0}$ に対して $P(|X_n - Y| \geq \varepsilon)$ は 0 に収束する。 □

補題 16.21 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 16.4 における $P(|X_n - Y| < \varepsilon)$ の収束オーダーを評価するという点まで定理 16.18 がカバーしているわけではない。

16.22 注意. 証明は割愛するが、独立な確率変数の列 $X_1, X_2, \ldots, X_n, \ldots$ 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。

索 引