確率論

箱

2024 年 5 月 22 日

概要

確率論の基本事項をまとめた．後半では，大数の強法則と中心極限定理を証明する．

目次

記号と用語

- 自然数，実数，複素数全体の集合を，それぞれ $\mathbb{N}$, $\mathbb{R}$, $\mathbb{C}$ と書く．0 は自然数に含める．また，$\mathbb{R}_{\geq 0} = [0,\infty)$，$\overline{\mathbb{R}}_{\geq 0} = [0,\infty]$ と書く．
- 位相空間は，常に Borel 集合族によって可測空間とみなす．位相空間上の測度とは，常に Borel 測度のことをいう．
- 積分値 $\int f\,d\mu$ を，$\mu(f)$ とも書く．
- 集合 Ω を考えているとき，その部分集合 A の特性関数を χ_A と書く．
- 位相空間の部分集合 A に対して，その閉包を $\overline{A}$，内部を A°，境界を ∂A と書く．
- 距離 d に関する中心 x，半径 r の閉球を $B_d(x;r)$，開球を $B_d^\circ(x;r)$ と書く．
- V を有限次元実線型空間とするとき，V^* と V の間，および $V^* \otimes V^*$ と $V \otimes V$ の間の自然なペアリングを，ともに $\langle -,- \rangle$ と書く．

1　確率変数

1.1　確率空間と確率変数

定義 1.1（確率空間）　可測空間 $(\Omega,\mathfrak{A})$ 上の全測度 1 の（正値）測度 P を，$(\Omega,\mathfrak{A})$ 上の**確率測度**（probability）という．このとき，測度空間 $(\Omega,\mathfrak{A},P)$ を，**確率空間**（probability space）といい，可測集合 $A \in \mathfrak{A}$ を，この確率空間における**事象**（event）という．

確率空間 $(\Omega,\mathfrak{A},P)$ における事象 A に対して，$P(A)$ を，「A の確率」や「A が起こる確率」という．また，確率空間を考えているときには，「ほとんど至るところで……が成り立つ」という代わりに，「ほとんど確実に（almost surely, a.s.）……が成り立つ」ということが多い．

定義 1.2（確率変数）　確率空間 $(\Omega,\mathfrak{A},P)$ から可測空間 $(E,\mathfrak{B})$ への可測写像 $X\colon \Omega \to E$ を，$(\Omega,\mathfrak{A},P)$ 上の E に値をとる（あるいは，単に E 値）**確率変数**（random variable）という．このとき，P の X による像測度 X_*P を，確率変数 X の**分布**（distribution）という．

$\mathbb{R}$ に値をとる確率変数を実確率変数，$\mathbb{C}$ に値をとる確率変数を複素確率変数という．

X を確率空間 $(\Omega,\mathfrak{A},P)$ 上の確率変数とするとき，しばしば，$\{\omega \in \Omega \mid X(\omega) \in B\}$ を $\{X \in B\}$ と略記したり，$P(\{\omega \in \Omega \mid X(\omega) \in B\})$ を $P(X \in B)$ と略記したりする[*1]．

確率変数 X の分布が μ であることを，「X は分布 μ に従う」ともいう[*2]．同じ可測空間に値をとる確率変数の族 $(X_i)_{i\in I}$ の分布が等しいことを，$(X_i)_{i\in I}$ は**同分布**（identically distributed）であるという．確率変数が列挙されている場合には，「X と Y は同分布である」，「X_0, X_1, ... は同分布である」などといういい方もする．

命題 1.3　$(\Omega,\mathfrak{A},P)$ を確率空間とする．

[*1] このような記法の利点は，確率空間 $(\Omega,\mathfrak{A},P)$ を意識せず，確率変数 X に注目して議論ができるようになることである．
[*2] 本稿では用いないが，X が分布 μ に従うことを，$X \sim \mu$ と書くことが多い．

(1) $(A_n)_{n\in\mathbb{N}}$ を $(\Omega, \mathfrak{A}, P)$ における事象の増加列とすると，$P(\bigcup_{n\in\mathbb{N}} A_n) = \lim_{n\to\infty} P(A_n)$ である．
(2) $(A_n)_{n\in\mathbb{N}}$ を $(\Omega, \mathfrak{A}, P)$ における事象の減少列とすると，$P(\bigcap_{n\in\mathbb{N}} A_n) = \lim_{n\to\infty} P(A_n)$ である．

証明　測度空間の一般論である． □

1.2　独立性

定義 1.4（独立性）　$(\Omega, \mathfrak{A}, P)$ を確率空間とする．

(1) $(\Omega, \mathfrak{A}, P)$ における事象の族 $(A_i)_{i\in I}$ が**独立**（independent）であるとは，任意の有限部分集合 $F \subseteq I$ に対して，
$$P\left(\bigcup_{i\in F} A_i\right) = \prod_{i\in F} P(A_i)$$
が成り立つことをいう．
(2) $\mathfrak{A}$ の部分 σ-代数の族 $(\mathfrak{A}_i)_{i\in I}$ が**独立**であるとは，任意の有限部分集合 $F \subseteq I$ と各 $i \in F$ に対する任意の $A_i \in \mathfrak{A}_i$ に対して，事象の族 $(A_i)_{i\in F}$ が独立であることをいう．
(3) $(\Omega, \mathfrak{A}, P)$ 上の確率変数の族 $(X_i)_{i\in I}$（X_i の終域である可測空間は，i ごとに異なってもよい）が**独立**であるとは，$\mathfrak{A}$ の部分 σ-代数の族 $(\sigma(X_i))_{i\in I}$ が独立であることをいう．

事象・部分 σ-代数・確率変数が列挙されている場合には，「X と Y は独立である」，「X_0, X_1, $\ldots$ は独立である」などといういい方もする．

同じ可測空間に値をとる確率変数の族 $(X_i)_{i\in I}$ が独立かつ同分布であるとき，$(X_i)_{i\in I}$ は**独立同分布**（independent and identically distributed, i.i.d.）であるという（確率変数が列挙されている場合のいい方も同様である）．$\mathbb{R}$ や有限次元実線型空間に値をとる確率変数の独立同分布列 $(X_i)_{i\in\mathbb{N}_{>0}}$ の極限における性質を，3 節と 4 節で述べる．

注意 1.5　確率空間 $(\Omega, \mathfrak{A}, P)$ における事象の族 $(A_i)_{i\in I}$ が独立であるとする．このとき，I の異なる元からなる任意の列 $(i_n)_{n\in\mathbb{N}}$ に対して，命題 1.3 (2) より，
$$P\left(\bigcup_{k\in\mathbb{N}} A_{i_k}\right) = \lim_{n\to\infty} P\left(\bigcup_{k=0}^{n-1} A_{i_k}\right) = \lim_{n\to\infty} \prod_{k=0}^{n-1} P(A_{i_k}) = \prod_{k\in\mathbb{N}} P(A_{i_k})$$
である．よって，任意の可算部分集合 $J \subseteq I$ に対しても，
$$P\left(\bigcup_{i\in J} A_i\right) = \prod_{i\in J} P(A_i)$$
が成り立つ．

注意 1.6　$(\Omega, \mathfrak{A}, P)$ を確率空間とし，$(\mathfrak{A}_i)_{i\in I}$ を $\mathfrak{A}$ の部分 σ-代数の有限族とする．もし

($*$) 各 $i \in I$ に対する任意の $A_i \in \mathfrak{A}_i$ に対して，
$$P\left(\bigcup_{i\in I} A_i\right) = \prod_{i\in I} P(A_i)$$
が成り立つ

ならば，A_i のうちいくつかを Ω に固定することによって，この等式が，I を任意の部分集合 $F \subseteq I$ で置き換えても成り立つことがわかる．すなわち，$(\mathfrak{A}_i)_{i\in I}$ が独立であることを示すためには，条件 $(*)$ を確かめれば十分である．確率変数の有限族の独立性についても，同じことがいえる．

命題 1.7 $(X_i)_{i\in I}$ を確率空間 $(\Omega, \mathfrak{A}, P)$ 上の確率変数の有限族（X_i の終域である可測空間 $(E_i, \mathfrak{B}_i)$ は，i ごとに異なってもよい）とし，各 $i \in I$ に対して，X_i の分布を μ_i と書く．次の 2 条件は同値である．

(a) $(X_i)_{i\in I}$ は独立である．
(b) $(\prod_{i\in I} E_i, \bigotimes_{i\in I} \mathfrak{B}_i)$ 値確率変数 $(X_i)_{i\in I}$ の分布は，積測度 $\bigotimes_{i\in I} \mu_i$ に等しい．

証明 $(X_i)_{i\in I}$ が独立であるための必要十分条件は，各 $i \in I$ に対する任意の $B_i \in \mathfrak{B}_i$ に対して

$$P(\text{すべての } i \in I \text{ に対して } X_i \in B_i) = \prod_{i\in I} P(X_i \in B_i)$$

であることである（注意 1.6）．確率変数 $(X_i)_{i\in I}$ の分布を μ と書けば，上式は，

$$\mu\left(\prod_{i\in I} B_i\right) = \prod_{i\in I} \mu_i(B_i)$$

と書き直せる．この等式が各 $i \in I$ に対する任意の $B_i \in \mathfrak{B}_i$ に対して成り立つことは，μ が積測度 $\bigotimes_{i\in I} \mu_i$ に等しいということにほかならない． □

命題 1.8 $(\Omega, \mathfrak{A}, P)$ を確率空間とし，Ω の部分集合 A に対して，$A^0 = A$，$A^1 = \Omega \setminus A$ と定める．$(\Omega, \mathfrak{A}, P)$ における事象の族 $(A_i)_{i\in I}$ が独立ならば，任意の写像 $b\colon I \to \{0,1\}$ に対して，$(A_i^{b(i)})_{i\in I}$ も独立である．

証明 有限部分集合 $F \subseteq I$ とその部分集合 $S \subseteq F$ に対する等式

$$P\left(\bigcap_{i\in F} A_i^{\chi_S(i)}\right) = \prod_{i\in F} P(A_i^{\chi_S(i)}) \qquad (*_{F,S})$$

を考える．各 $n \in \mathbb{N}$ に対して，n 元部分集合 $F \subseteq I$ とその部分集合 $S \subseteq F$ に対しては $(*_{F,S})$ が成り立つことを，n に関する帰納法で示す．$F \subseteq I$ を n 元部分集合とする．$(A_i)_{i\in I}$ は独立だから，$(*_{F,\emptyset})$ は成り立つ．また，$(*_{F,S})$（$S \subseteq F$）が成り立つとして，$i_0 \in F \setminus S$ を固定すると，帰納法の仮定より

$$P\left(\bigcap_{i\in F\setminus\{i_0\}} A_i^{\chi_S(i)}\right) = \prod_{i\in F\setminus\{i_0\}} P(A_i^{\chi_S(i)})$$

であり，等式 $(*_{F,S})$ が成り立つという仮定より

$$P\left(\bigcap_{i\in F} A_i^{\chi_S(i)}\right) = \prod_{i\in F} P(A_i^{\chi_S(i)})$$

だから，辺々引いて

$$P\left(\bigcap_{i\in F} A_i^{\chi_{S\cup\{i_0\}}(i)}\right) = \prod_{i\in F} P(A_i^{\chi_{S\cup\{i_0\}}(i)})$$

を得る．すなわち，$(*_{F,S\cup\{i_0\}})$ が成り立つ．よって，帰納法より，任意の部分集合 $S \subseteq F$ に対して $(*_{F,S})$ が成り立つ．以上で，帰納法が完成し，任意の有限部分集合 $F \subseteq I$ とその部分集合 $S \subseteq F$ に対して $(*_{F,S})$ が成り立つことが示された． □

命題 1.9 $(\Omega, \mathfrak{A}, P)$ を確率空間とし，$(\mathfrak{A}_i)_{i\in I}$ を $\mathfrak{A}$ の部分 σ-代数の族とする．$(I_j)_{j\in J}$ を I の分割とし，各 $j \in J$ に対して $\mathfrak{B}_j = \sigma(\bigcup_{i\in I_j} \mathfrak{A}_i)$ と置く．このとき，$(\mathfrak{A}_i)_{i\in I}$ が独立ならば，$(\mathfrak{B}_j)_{j\in J}$ も独立である．

証明 各 $j \in J$ に対して，$\bigcup_{i\in I_j} \mathfrak{A}_i$ の有限個の元の交叉全体のなす集合を $\mathfrak{P}_j$ と置くと，$\mathfrak{P}_j$ は $\mathfrak{B}_j$ を σ-代数として生成する二交叉族である．有限部分集合 $F \subseteq J$ を固定し，部分集合 $S \subseteq F$ に対する命題

$(*_{F,S})$ 任意の $B_j \in \mathfrak{B}_j$（$j \in S$）と $B_j \in \mathfrak{P}_j$（$j \in F \setminus S$）に対して，$P(\bigcap_{j\in F} B_j) = \prod_{j\in F} P(B_j)$ である．

を考える．$(\mathfrak{A}_i)_{i\in I}$ が独立であることより，$(*_{F,\emptyset})$ は成り立つ．また，$(*_{F,S})$（$S \subseteq F$）が成り立つとして，$j_0 \in F \setminus S$ を固定すると，

$$\left\{ B_{j_0} \in \mathfrak{B}_{j_0} \;\middle|\; \text{任意の } B_j \in \mathfrak{B}_j\ (j \in S) \text{ と } B_j \in \mathfrak{P}_j\ (j \in F \setminus (S \cup \{j_0\})) \text{ に対して，} P\left(\bigcap_{j\in F} B_j\right) = \prod_{j\in F} B_j \right\}$$

は $\mathfrak{P}_j$ を含む Dynkin 族だから，Dynkin 族補題より，$\mathfrak{B}_{j_0}$ 全体に等しい．すなわち，$(*_{F,S\cup\{j_0\}})$ が成り立つ．よって，帰納法より，$(*_{F,F})$ が成り立つ．任意の有限部分集合 $F \subseteq J$ に対してこれが成り立つから，$(\mathfrak{B}_j)_{j\in J}$ は独立である． □

1.3 Borel–Cantelli の補題

命題 1.10（Borel–Cantelli の補題） $(\Omega, \mathfrak{A}, P)$ を確率空間とし，$(A_i)_{i\in I}$ をこの確率空間における事象の可算族とする．

(1) $\sum_{i\in I} P(A_i) < \infty$ ならば，無限個の i に対して A_i が起こる確率は 0 である．
(2) $(A_i)_{i\in I}$ が独立であり，$\sum_{i\in I} P(A_i) = \infty$ ならば，無限個の i に対して A_i が起こる確率は 1 である．

証明 (1) 「無限個の i に対して A_i が起こる」という事象は，可測集合 $\bigcap_{\text{有限部分集合 } F \subseteq I} \bigcup_{i\in I\setminus F} A_i$ で表される．したがって，任意の有限部分集合 $F \subseteq I$ に対して，

$$P(\text{無限個の } i \text{ に対して } A_i \text{ が起こる}) \leq P\left(\bigcup_{i\in I\setminus F} A_i\right) \leq \sum_{i\in I\setminus F} P(A_i)$$

である．$\sum_{i\in I} P(A_i) < \infty$ とすると，上式の右辺は $F \to I$ のとき 0 に収束するから，無限個の i に対して A_i が起こる確率は 0 である．

(2) $(A_i)_{i\in I}$ が独立であり，$\sum_{i\in I} P(A_i) = \infty$ であるとする．まず，

$$\begin{aligned} 1 - P(\text{無限個の } i \text{ に対して } A_i \text{ が起こる}) &\leq P\left(\bigcup_{\text{有限部分集合 } F \subseteq I} \bigcap_{i\in I\setminus F} (\Omega \setminus A_i)\right) \\ &\leq \sum_{\text{有限部分集合 } F \subseteq I} P\left(\bigcap_{i\in I\setminus F} (\Omega \setminus A_i)\right) \end{aligned}$$

である．次に，有限部分集合 $F \subseteq I$ を固定すると，

$$
\begin{aligned}
P\left(\bigcap_{i \in I \setminus F} (\varOmega \setminus A_i)\right) &= \prod_{i \in I \setminus F} (1 - P(A_i)) && (*) \\
&\leq \prod_{i \in I \setminus F} e^{-P(A_i)} && (**) \\
&= \exp\left(-\sum_{i \in I \setminus F} P(A_i)\right) \\
&= 0
\end{aligned}
$$

である．ここで，等号 $(*)$ は $(A_i)_{i \in I}$ の独立性，命題 1.8，および注意 1.5 から，不等号 $(**)$ は実数 $t \in \mathbb{R}$ に対して $1 - t \leq e^{-t}$ であることから従う．よって，無限個の i に対して A_i が起こる確率は 1 である． □

1.4 有限次元実線型空間に値をとる確率変数

$(\varOmega, \mathfrak{A}, \mu)$ を測度空間, V を有限次元実線型空間とするとき, 可測関数 $f \colon \varOmega \to V$ に対する p **乗可積分性**（$p \in [1, \infty]$）が定義される[*3]．$(\varOmega, \mathfrak{A}, \mu)$ 上の p 乗可積分な V 値可測関数全体のなす線型空間を，$\mathscr{L}^p(\varOmega, \mathfrak{A}, \mu; V)$ と書く．$(\varOmega, \mathfrak{A}, \mu)$ が有限測度空間であるとき，$1 \leq p \leq q \leq \infty$ ならば $\mathscr{L}^q(\varOmega, \mathfrak{A}, \mu; V) \subseteq \mathscr{L}^p(\varOmega, \mathfrak{A}, \mu; V)$ である．

定義 1.11（期待値） $(\varOmega, \mathfrak{A}, P)$ を確率空間とし，V を有限次元実線型空間とする．$(\varOmega, \mathfrak{A}, P)$ 上の可積分な V 値確率変数 X について，X の**期待値**（expected value, expectation）あるいは**平均**（mean）を，

$$E[X] = \int_{\varOmega} X \, dP$$

と定める．

定義 1.12（共分散，分散） $(\varOmega, \mathfrak{A}, P)$ を確率空間とする．

(1) X, Y は確率空間 $(\varOmega, \mathfrak{A}, P)$ 上の実確率変数であり，X, Y, XY はいずれも可積分であるとする．このとき，X と Y の**共分散**（covariance）を，

$$\operatorname{Cov}[X, Y] = E[(X - E[X])(Y - E[Y])] = \int_{\varOmega} (X - E[X])(Y - E[Y]) \, dP$$

と定める．

(2) $(\varOmega, \mathfrak{A}, P)$ 上の 2 乗可積分な実確率変数 X に対して，X の**分散**（variance）を，$\operatorname{Var}[X] = \operatorname{Cov}[X, X]$ と定める．X の分散の正の平方根 $\sqrt{\operatorname{Var}[X]}$ を，X の**標準偏差**（standard deviation）という．

定義 1.13（相互共分散テンソル，共分散テンソル） $(\varOmega, \mathfrak{A}, P)$ を確率空間とし，V, W を有限次元実線型空間とする．

[*3] たとえば，可測関数 $f \colon \varOmega \to V$ が p 乗可積分であるとは，任意の $\phi \in V^*$ に対して $\phi \circ f \colon \varOmega \to \mathbb{R}$ が p 乗可積分であることと定義すればよい．$|-|$ を V 上のノルムとすると，これは，E 上の関数 $\omega \mapsto |f(\omega)|$ が p 乗可積分であることと同値である．

(1) X, Y は $(\Omega, \mathfrak{A}, P)$ 上のそれぞれ V, W に値をとる確率変数であり，任意の $s \in V^*$ と $t \in W^*$ に対して共分散 $\mathrm{Cov}[s(X), t(Y)]$ が定義されるとする．このとき，関数 $(s, t) \mapsto \mathrm{Cov}[s(X), t(Y)]$ は $V^* \times W^*$ 上の双線型形式だから，$V \otimes W$ の元を定める．この元を，X と Y の**相互共分散テンソル**（cross-covariance tensor）という．

(2) $(\Omega, \mathfrak{A}, P)$ 上の V 値確率変数 X について，X と X 自身の相互共分散テンソルが定義されるならば，これを，X の**共分散テンソル**（covariance tensor）という．[*4]

$V = \mathbb{R}^{d_1}$，$W = \mathbb{R}^{d_2}$ の場合，X と Y の相互共分散テンソルは $\mathbb{R}^{d_1} \otimes \mathbb{R}^{d_2}$ の元だが，これを標準的な基底に関して成分表示して $d_1 \times d_2$ 行列の形に表したものを，X と Y の**相互共分散行列**（cross-covariance matrix）という．すなわち，$X = (X_1, \ldots, X_{d_1})^{\mathrm{T}}$ と $Y = (Y_1, \ldots, Y_{d_2})^{\mathrm{T}}$ の相互共分散行列は，

$$E[(X - E[X])(Y - E[Y])^{\mathrm{T}}] = ((X_i - E[X_i])(Y_j - E[Y_j]))_{1 \leq i \leq d_1,\ 1 \leq j \leq d_2}$$

である．X と X 自身の相互共分散行列を，**共分散行列**（covariance matrix）という．

注意 1.14　有限次元実線型空間に値をとる確率変数 X と Y が 2 乗可積分ならば，X と Y の相互共分散テンソルが定義される．これ以外にも，たとえば，X と Y が可積分で互いに独立である場合には，相互共分散テンソルが定義される（後述の命題 1.18 を参照のこと）．

注意 1.15　X を有限次元実線型空間 V に値をとる確率変数とし，X の共分散テンソル $\Sigma \in V \otimes V$ が定義されるとすると，容易に確かめられるように，Σ は正値対称 2 階テンソルである．すなわち，任意の $s, t \in V^*$ に対して $\langle s \otimes t, \Sigma \rangle = \langle t \otimes s, \Sigma \rangle$ であり，任意の $t \in V^*$ に対して $\langle t \otimes t, \Sigma \rangle \geq 0$ である．

注意 1.16　有限次元実線型空間に値をとる確率変数 X が p 乗可積分であるかどうかや，X の期待値や共分散テンソルは，X の分布のみに依存する．そこで，本稿では，用語の濫用で，有限次元実線型空間上の確率測度 μ に対しても，「μ は p 乗可積分である[*5]」，「μ の期待値」，「μ の共分散テンソル」ということがある．また，有限次元実線型空間に値をとる確率変数 X と Y の相互共分散テンソルは，(X, Y) の分布にのみ依存する．

命題 1.17　$(\Omega, \mathfrak{A}, P)$ を確率空間，V を有限次元実線型空間とし，X, Y を $(\Omega, \mathfrak{A}, P)$ 上の互いに独立な V 値確率変数とする．X, Y の分布をそれぞれ μ, ν とすると，$X + Y$ の分布は畳み込み $\mu * \nu$ に等しい．

証明　$V \times V$ 値確率変数 (X, Y) の分布は積測度 $\mu \otimes \nu$ に等しく，$X + Y$ の分布はこれの写像 $(x, y) \mapsto x + y$ による像測度，すなわち畳み込み $\mu * \nu$ に等しい．　□

命題 1.18　$(\Omega, \mathfrak{A}, P)$ を確率空間，V, W を有限次元実線型空間とし，X, Y を $(\Omega, \mathfrak{A}, P)$ 上のそれぞれ V, W に値をとる可積分な確率変数とする．X と Y が独立ならば，これらの相互共分散テンソルは定義され，0 に等しい．

証明　X と Y の相互共分散テンソルが定義され 0 であるとは，任意の $s \in V^*$ と $t \in W^*$ に対して $s(X)$ と $t(Y)$ の共分散が定義され 0 であることにほかならない．また，X と Y が独立ならば，任意の $s \in V^*$ と $t \in W^*$ に対して，$s(X)$ と $t(Y)$ は独立である．したがって，主張は，$V = W = \mathbb{R}$ の場合に示せば十分で

[*4] 「相互共分散テンソル」と「共分散テンソル」は，本稿だけの用語である．ただし，すぐ下に述べる「相互共分散行列」と「共分散行列」は，一般的な用語である．

[*5] 関数ではなく測度に対して「p 乗可積分である」というのは違和感があるのだが，ほかに適当な語も思いつかないので，このようにいう．

ある．

X と Y を $(\Omega, \mathfrak{A}, P)$ 上の可積分な実確率変数とし，これらは独立であるとする．X, Y の分布をそれぞれ μ, ν と書くと，(X, Y) の分布は $\mu \otimes \nu$ である（命題 1.7）．よって，XY は可積分であり，

$$\begin{aligned}\mathrm{Cov}[X,Y] &= E[(X-E[X])(Y-E[Y])] \\ &= \int_{\mathbb{R}\times\mathbb{R}} (x-E[X])(y-E[Y])\,d(\mu\otimes\nu)(x,y) \\ &= \left(\int_{\mathbb{R}} (x-E[X])\,d\mu(x)\right)\left(\int_{\mathbb{R}} (y-E[Y])\,d\nu(y)\right) \\ &= 0\end{aligned}$$

が成り立つ．□

系 1.19　確率空間 $(\Omega, \mathfrak{A}, P)$ 上の可積分な実確率変数 X と Y が独立ならば，XY も可積分であり，$E[XY] = E[X]E[Y]$ が成り立つ．

証明　$V = W = \mathbb{R}$ の場合の命題 1.18 のいいかえにすぎない．□

命題 1.20（Markov の不等式）　$(\Omega, \mathfrak{A}, P)$ を確率変数とし，X をその上の可積分な $\mathbb{R}_{\geq 0}$ 値確率変数とする．任意の $a > 0$ に対して，

$$P(X \geq a) \leq \frac{1}{a}E[X]$$

である．

証明　X の期待値は

$$E[X] = \int_{\Omega} X\,dP \geq \int_{\{X\geq a\}} X\,dP \geq aP(X \geq a)$$

と評価できるから，$P(X \geq a) \leq (1/a)E[X]$ である．□

系 1.21（Chebyshev の不等式）　$(\Omega, \mathfrak{A}, P)$ を確率変数，X をその上の 2 乗可積分な実確率変数とし，X の標準偏差を σ と書く．任意の $k > 0$ に対して，

$$P(|X - E[X]| \geq k\sigma) \leq \frac{1}{k^2}$$

である．

証明　Markov の不等式（命題 1.20）より，

$$\begin{aligned}P(|X-E[X]| \geq k\sigma) &= P((X-E[X])^2 \leq k^2\sigma^2) \\ &\leq \frac{1}{k^2\sigma^2}E[(X-E[X])^2] \\ &= \frac{1}{k^2}\end{aligned}$$

である．□

1.5 特性関数

定義 1.22（特性関数） 有限次元実線型空間 V 上の確率測度 μ に対して，その**特性関数**（characteristic function）$\phi_\mu\colon V^* \to \mathbb{C}$ を，

$$\phi_\mu(t) = \int_V e^{i\langle t,x\rangle}\, d\mu(x) \qquad (t \in V^*)$$

と定める．

$(\Omega, \mathfrak{A}, P)$ を確率空間とし，V を有限次元実線型空間とする．$(\Omega, \mathfrak{A}, P)$ 上の V 値確率変数 X に対して，X の分布の特性関数を，単に X の**特性関数**といい，ϕ_X と書く．

$\mathbb{R}^d$ 上の確率測度 μ に対しては，特に断らなくても，標準的な同型 $(\mathbb{R}^d)^* \cong \mathbb{R}^d$ を通して，特性関数 ϕ_μ を $\mathbb{R}^d$ 上の関数とみなす．すなわち，

$$\phi_\mu(t_1, \dots, t_d) = \int_{\mathbb{R}^d} e^{i(t_1x_1+\dots+t_dx_d)}\, d\mu(x_1, \dots, x_d) \qquad ((t_1, \dots, t_d) \in \mathbb{R}^d)$$

とみなす．

命題 1.23 V, W を有限次元実線型空間とし，$A\colon V \to W$ を線型写像，$b \in W$ とする．μ を V 上の確率測度とし，その V から W へのアフィン写像 $x \mapsto Ax + b$ による像測度を ν と書く．このとき，ν の特性関数 ϕ_ν は，μ の特性関数 ϕ_μ を用いて

$$\phi_\nu(t) = e^{i\langle t,b\rangle}\phi_\mu(A^*t) \qquad (t \in W^*)$$

と表せる（A の双対線型写像を $A^*\colon W^* \to V^*$ と書いた）．

証明 $t \in W^*$ に対して，

$$\begin{aligned}
\phi_\nu(t) &= \int_W e^{i\langle t,y\rangle}\, d\nu(y) \\
&= \int_V e^{i\langle t,Ax+b\rangle}\, d\mu(x) \\
&= e^{i\langle t,b\rangle} \int_V e^{i\langle A^*t,x\rangle}\, d\mu(x) \\
&= e^{i\langle t,b\rangle}\phi_\mu(A^*t)
\end{aligned}$$

である． □

V と W を有限次元実線型空間とし，V' を V の開集合とする．写像 $f\colon V' \to W$ に対して，f の点 $x \in V'$ における方向 $v \in V$ に関する**方向微分**を，

$$D_vf(x) = \lim_{\epsilon\to 0} \frac{f(x+\epsilon v) - f(x)}{\epsilon}$$

と定義する（極限が存在する場合）．任意の点 $x \in V'$ において方向微分 $D_vf(x)$ が定義される場合，f の方向 v に関する**偏導関数** $D_vf\colon V' \to W$ が定まる．偏導関数をとる操作を繰り返したもの（高階偏導関数）を考えるとき，$D_{v_1} \cdots D_{v_k}f$ を $D^k_{v_1,\dots,v_k}f$ と略記することにする．$p \in \mathbb{N}$ に対して，p 階までのすべての高階偏導関数 $D^k_{v_1,\dots,v_k}f$（$k \in \{0, \dots, p\}$，v_1, …, $v_k \in V$）が V' 上で定義され連続であるとき，f は p **階連続微分可**

能であるという．写像 $f\colon V' \to W$ が p 階連続微分可能であるとき，任意の $k \in \{0, \dots, p\}$ と点 $x \in V'$ に対して，$(v_1, \dots, v_k) \mapsto D^k_{v_1,\dots,v_k} f(x)$ が V^k から W への対称 k 重線型形式であることが知られている．

命題 1.24　V を有限次元実線型空間とし，μ をその上の確率測度とする．μ の特性関数 ϕ_μ を考える．

(1) ϕ_μ は連続である．

(2) p を正の整数とする．μ が p 乗可積分ならば，ϕ_μ は p 階連続微分可能であり，任意の $k \in \{0, \dots, p\}$ と $t_1, \dots, t_k \in V^*$ に対して，

$$D^k_{t_1,\dots,t_k} \phi_\mu(t) = i^k \int_V t_1(x) \cdots t_k(x) e^{itx}\, d\mu(x) \qquad (t \in V^*)$$

である．

証明　(1)　Lebesgue の収束定理から従う．

(2)　μ が p 乗可積分であるとする．このとき，$k \in \{0, \dots, p\}$ と $t_1, \dots, t_k \in V^*$ に対して，V 上の関数 $x \mapsto t_1(x) \cdots t_k(x)$ が可積分であることに注意する．k に関する帰納法で主張を示す．$k = 0$ のときの主張は，特性関数の定義そのものである．$k \geq 1$ とし，$k - 1$ のときの主張が正しいとすると，微分と積分の順序交換に関する定理より，

$$\begin{aligned} D^k_{t_1,\dots,t_k} \phi_\mu(t) &= D_{t_k}\left(i^{k-1} \int_V t_1(x) \cdots t_{k-1}(x) e^{itx}\, d\mu(x) \right) \\ &= i^{k-1} \int_V t_1(x) \cdots t_{k-1}(x) D_{t_k} e^{itx}\, d\mu(x) \\ &= i^k \int_V t_1(x) \cdots t_k(x) e^{itx}\, d\mu(x) \qquad (t_1, \dots, t_k \in V^*) \end{aligned}$$

が成り立つ．すなわち，k のときの主張も正しい．これで，帰納法が完成した．□

次の系では，実線型空間 V の複素化を，$V_{(\mathbb{C})}$ と書く．

系 1.25　V を有限次元実線型空間とし，μ をその上の確率測度とする．μ の特性関数 ϕ_μ を考える．

(1) $\phi_\mu(0) = 1$ である．

(2) μ が可積分ならば，V^* から $\mathbb{C}$ への実線型写像 $t \mapsto D_t \phi_\mu(0)$ が定める $V_{(\mathbb{C})}$ の元は，μ の期待値の i 倍に等しい．

(3) μ が 2 乗可積分ならば，$V^* \times V^*$ から $\mathbb{C}$ への実双線型写像 $(s, t) \mapsto D^2_{s,t} \phi_\mu(0)$ が定める $(V \otimes V)_{(\mathbb{C})}$ の元は，μ の共分散テンソルの -1 倍に等しい．

証明　(1)　特性関数の定義から明らかである．

(2), (3)　命題 1.24 から従う．□

命題 1.26　有限次元実線型空間 V 上の確率測度 μ と ν について，これらの畳み込み $\mu * \nu$ の特性関数 $\phi_{\mu*\nu}$ は，それぞれの特性関数 ϕ_μ と ϕ_ν の積に等しい．

証明　$t \in V^*$ に対して,

$$\begin{aligned}\phi_{\mu*\nu}(t) &= \int_V e^{i\langle t,x\rangle}\,d(\mu*\nu)(x)\\ &= \int_{V\times V} e^{i\langle t,y+z\rangle}\,d(\mu\otimes\nu)(y,z)\\ &= \left(\int_V e^{i\langle t,y\rangle}\,d\mu(y)\right)\left(\int_V e^{i\langle t,z\rangle}\,d\nu(z)\right)\\ &= \phi_\mu(t)\phi_\nu(t)\end{aligned}$$

である. □

系 1.27　$(\Omega,\mathfrak{A},P)$ を確率空間, V を有限次元実線型空間とし, X, Y を $(\Omega,\mathfrak{A},P)$ 上の互いに独立な V 値確率変数とする. $X+Y$ の特性関数 ϕ_{X+Y} は, それぞれの特性関数 ϕ_X と ϕ_Y の積に等しい.

証明　$X+Y$ の分布は X と Y の分布の畳み込みに等しいから（命題 1.17）, 主張は命題 1.26 から従う. □

命題 1.28　$V_1,\dots,V_n$ を有限次元実線型空間とし, 各 $k\in\{1,\dots,n\}$ に対して, μ_k を V_k 上の確率測度とする. 積測度 $\mu_1\otimes\cdots\otimes\mu_n$ の特性関数は,

$$\phi_{\mu_1\otimes\cdots\otimes\mu_n}(t_1,\dots,t_n) = \phi_{\mu_1}(t_1)\cdots\phi_{\mu_n}(t_n) \qquad (t_k\in V_k^*)$$

と表される. ここで, $(V_1\times\cdots\times V_n)^*$ を自然な同型によって $V_1^*\times\cdots\times V_n^*$ と同一視している.

証明　$(t_1,\dots,t_n)\in V_1^*\times\cdots\times V_n^*$ に対して,

$$\begin{aligned}\phi_{\mu_1\otimes\cdots\otimes\mu_n}(t_1,\dots,t_n) &= \int_{V_1\times\cdots\times V_n}\prod_{k=1}^n e^{it_kx_k}\,d(\mu_1\otimes\cdots\otimes\mu_n)(x_1,\dots,x_n)\\ &= \prod_{k=1}^n\int_{V_k} e^{it_kx_k}\,d\mu_k(x_k)\\ &= \phi_{\mu_1}(t_1)\cdots\phi_{\mu_n}(t_n)\end{aligned}$$

である. □

本小節の以下の部分では, 特性関数が確率測度を特徴付けることを示す. そのために, 補題を用意する. $\sigma>0$ に対して, $\mathbb{R}$ 上の正値関数 $x\mapsto(2\pi\sigma^2)^{-1/2}e^{-x^2/2\sigma^2}$（$|{-}|$ は Euclid ノルムを表す）の Lebesgue 測度に関する積分値は 1 だから（Gauss 積分）, この正値関数を Lebesgue 測度に関する密度関数にもつ $\mathbb{R}$ 上の確率測度 ν_σ が存在する[*6]. 以下の補題および定理では, この記号 ν_σ を用いる.

補題 1.29　$\mathbb{R}^d$ 上の確率測度 $\nu_\sigma^{\otimes d}$ の特性関数 $\phi_{\nu_\sigma^{\otimes d}}$ は,

$$\phi_{\nu_\sigma^{\otimes d}}(t) = e^{-\sigma^2|t|^2/2} \qquad (t\in\mathbb{R}^d)$$

で与えられる（$|{-}|$ は Euclid ノルムを表す）.

証明　命題 1.28 と命題 1.23 より, 主張は, $d=1$ かつ $\sigma=1$ のときに示せば十分である. 関数 $\varPhi\colon\mathbb{C}\to\mathbb{C}$ を

$$\varPhi(z) = \int_{\mathbb{R}} e^{zx}\cdot\frac{1}{(2\pi)^{1/2}}e^{-x^2}\,dx \qquad (z\in\mathbb{C})$$

[*6] 4.1 節で定義する用語と記号を用いれば, ν_σ は, 中心正規分布 $N(0,\sigma^2)$ である.

と定めると，微分と積分の順序交換に関する定理より，Φ は正則である．$t \in \mathbb{R}$ に対しては

$$\Phi(t) = \int_{\mathbb{R}} e^{tx} \cdot \frac{1}{(2\pi)^{1/2}} e^{-x^2}\, dx = e^{t^2/2} \int_{\mathbb{R}} \frac{1}{(2\pi)^{1/2}} e^{-(x-t)^2}\, dx = e^{t^2/2}$$

だから，一致の定理より，$\Phi(z) = e^{z^2/2}$（$z \in \mathbb{C}$）である．特に，$\mathbb{R}$ 上の確率測度 ν_1 の特性関数 ϕ_{ν_1} は，$\phi_{\nu_1}(t) = \Phi(it) = e^{-t^2/2}$（$t \in \mathbb{R}$）で与えられる． □

補題 1.30　μ_1, μ_2 を $\mathbb{R}^d$ 上の確率測度とする．任意の $\sigma > 0$ に対して $\mu_1 * \nu_\sigma^{\otimes d} = \mu_2 * \nu_\sigma^{\otimes d}$ ならば，$\mu_1 = \mu_2$ である．

証明　μ を $\mathbb{R}^d$ 上の確率測度とする．$\sigma > 0$ に対して，$\nu_\sigma^{\otimes d}$ の Lebesgue 測度に関する密度関数

$$g_\sigma(x) = \frac{1}{(2\pi\sigma^2)^{d/2}} e^{-|x|^2/2\sigma^2} \qquad (x \in \mathbb{R}^d)$$

を考える．すると，容易に確かめられるように，コンパクト台連続関数 $f\colon \mathbb{R}^d \to \mathbb{R}$ に対して，$f * g_\sigma$ は $\sigma \to 0+$ のとき f に一様収束するから，

$$\begin{aligned}
\mu(f) &= \lim_{\sigma \to 0+} \mu(f * g_\sigma) \\
&= \lim_{\sigma \to 0+} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} f(x) g_\sigma(y-x)\, dx\, d\mu(y) \\
&= \lim_{\sigma \to 0+} \int_{\mathbb{R}^d} f(x) \left(\int_{\mathbb{R}^d} g_\sigma(y-x)\, d\mu(y) \right) dx \\
&= \lim_{\sigma \to 0+} \int_{\mathbb{R}^d} f(x) \left(\int_{\mathbb{R}^d} g_\sigma(x-y)\, d\mu(y) \right) dx \\
&= \lim_{\sigma \to 0+} \int_{\mathbb{R}^d} f\, d(\mu * \nu_\sigma^{\otimes d})
\end{aligned}$$

が成り立つ．よって，$\sigma > 0$ に対する $\mu * \nu_\sigma^{\otimes d}$ は，有界連続関数 $f\colon \mathbb{R}^d \to \mathbb{R}$ に対する $\mu(f)$ の値を一意に定め，したがって，μ を一意に定める． □

定理 1.31　有限次元実線型空間 V 上の確率測度 μ_1 と μ_2 が等しいための必要十分条件は，それらの特性関数 ϕ_{μ_1} と ϕ_{μ_2} が等しいことである．

証明　必要性は明らかだから，十分性を示す．一般性を失わず，$V = \mathbb{R}^d$ と仮定する．μ を $\mathbb{R}^d$ 上の確率測度とし，その特性関数 ϕ_μ を考える．すると，Tonelli の定理と補題 1.29 より，$\sigma > 0$ に対して

$$\begin{aligned}
\frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \phi_\mu(t) e^{-\sigma^2|t|^2} e^{-itx}\, dt &= \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}} e^{ity}\, d\mu(y) \right) e^{-\sigma^2|t|^2} e^{-itx}\, dt \\
&= \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} e^{-\sigma^2|t|^2} e^{-it(x-y)}\, dt\, d\mu(y) \\
&= \int_{\mathbb{R}^d} \frac{1}{(2\pi\sigma^2)^{d/2}} e^{|x-y|^2/2\sigma^2}\, d\mu(y) \qquad (x \in \mathbb{R}^d)
\end{aligned}$$

である（$|{-}|$ は Euclid ノルムを表す）．上式の最右辺は，$\mu * \nu_\sigma^{\otimes d}$ の Lebesgue 測度に関する密度関数にほかならない．したがって，$\mu * \nu_\sigma^{\otimes d}$ は，特性関数 ϕ_μ から一意に定まる．このことと補題 1.30 から，主張が従う． □

系 1.32　$(\Omega, \mathfrak{A}, P)$ を確率空間，$V_1, \ldots, V_n$ を有限次元実線型空間とし，各 $k \in \{1, \ldots, n\}$ に対して，X_k を $(\Omega, \mathfrak{A}, P)$ 上の V_k 値確率変数とする．次の 2 条件は同値である．

(a) $X_1, \ldots, X_n$ は独立である．

(b) $V_1 \times \cdots \times V_n$ 値確率変数 $X = (X_1, \ldots, X_n)$ の特性関数は，

$$\phi_X(t_1, \ldots, t_n) = \phi_{X_1}(t_1) \cdots \phi_{X_n}(t_n) \qquad (t_k \in V_k^*)$$

と表される．ここで，$(V_1 \times \cdots \times V_n)^*$ を自然な同型によって $V_1^* \times \cdots \times V_n^*$ と同一視している．

証明　各 X_k の分布を μ_k と書き，$X = (X_1, \ldots, X_n)$ の分布を μ と書く．$X_1, \ldots, X_n$ が独立であるための必要十分条件は，X の分布 μ が積測度 $\mu_1 \otimes \cdots \otimes \mu_n$ に等しいことであり（命題 1.7），これはさらに，μ と $\mu_1 \otimes \cdots \otimes \mu_n$ の特性関数が等しいことと同値である（定理 1.31）．よって，主張は命題 1.28 から従う．　□

2　確率変数の収束

2.1　L^p 収束，概収束，確率収束

$(\Omega, \mathfrak{A}, \mu)$ を測度空間，V を有限次元実線型空間とする．V 上のノルム $|-|$ を一つ固定するとき，$f \in \mathscr{L}^p(\Omega, \mathfrak{A}, \mu; V)$（$p \in [1, \infty]$）に対して

$$\|f\|_p = \begin{cases} \left(\displaystyle\int_V |f(x)|^p \, d\mu(x) \right)^{1/p} & (p \in [1, \infty)) \\ \operatorname*{ess\,sup}_{x \in V} |f(x)| & (p = \infty) \end{cases}$$

と定めると，これは $\mathscr{L}^p(\Omega, \mathfrak{A}, \mu; V)$ 上の半ノルムとなる．さらに，この半ノルムが定める位相は，V 上のノルム $|-|$ のとり方によらない．この位相を，L^p **位相**といい，L^p 位相に関する収束を，L^p **収束**という．$(\Omega, \mathfrak{A}, \mu)$ が有限測度空間であるとき，$1 \leq p \leq q \leq \infty$ ならば，1.4 節で述べたように $\mathscr{L}^q(\Omega, \mathfrak{A}, \mu; V) \subseteq \mathscr{L}^p(\Omega, \mathfrak{A}, \mu; V)$ であり，L^q 位相は L^p 位相が定める相対位相よりも細かい．

$(\Omega, \mathfrak{A}, \mu)$ を測度空間，E を位相空間とするとき，$(\Omega, \mathfrak{A}, \mu)$ 上の E 値可測写像のネット $(f_i)_{i \in I}$ が E 値可測写像 f に**概収束**（almost covergent）するとは，μ-無視可能な集合 $N \subseteq \Omega$ が存在して，$(f_i)_{i \in I}$ が f に $\Omega \setminus N$ 上で各点収束することをいう．確率空間を考えているときには，確率変数のネット $(X_i)_{i \in I}$ が確率変数 X に概収束することを，$X_i \xrightarrow{\text{a.s.}} X$ と書く．

$(\Omega, \mathfrak{A}, \mu)$ を測度空間，(E, d) を距離空間とするとき，$(\Omega, \mathfrak{A}, \mu)$ 上の E 値可測写像のネット $(f_i)_{i \in I}$ が E 値可測写像 f に**測度収束**（convergent in measure）するとは，任意の $\epsilon > 0$ に対して，

$$\lim_i \mu^*(\{\omega \in \Omega \mid d(f_i(\omega), f(\omega)) \geq \epsilon\}) = 0$$

となることをいう．（ここで，μ^* は，μ が定める外測度を表す．E が第二可算ならば，上式の μ^* の中身は Ω の可測集合であり，したがって，μ^* を μ に置き換えてよい．）確率空間を考えているときには，測度収束の代わりに**確率収束**（convergence in probability）といい，確率変数のネット $(X_i)_{i \in I}$ が確率変数 X に確率収束することを，$X_i \xrightarrow{\text{P}} X$ と書く．

命題 2.1　$(\Omega, \mathfrak{A}, P)$ を確率空間，V を有限次元実線型空間とし，$(X_i)_{i \in I}$ を $(\Omega, \mathfrak{A}, P)$ 上の可積分な V 値確率変数のネット，X を $(\Omega, \mathfrak{A}, P)$ 上の可積分な V 値確率変数とする．$(X_i)_{i \in I}$ が X に L^1 収束するならば，V 上のノルム $|-|$ を一つ固定して V をノルム空間とみなすとき，$(X_i)_{i \in I}$ は X に確率収束する[*7]．

[*7] 容易に確かめられるように，V 値確率変数の確率収束の定義は，V 上のノルム $|-|$ のとり方によらない．

証明　Markov の不等式（命題 1.20）より，$\epsilon > 0$ に対して

$$P(|X_i - X| \geq \epsilon) \leq \frac{1}{\epsilon} E[|X_i - X|]$$

である．よって，$(X_i)_{i \in I}$ が X に L^1 収束するならば，$(X_i)_{i \in I}$ は X に確率収束する．　□

命題 2.2　$(\Omega, \mathfrak{A}, P)$ を確率空間，(E, d) を距離空間とし，$(X_n)_{n \in \mathbb{N}}$ を $(\Omega, \mathfrak{A}, P)$ 上の E 値確率変数の列，X を $(\Omega, \mathfrak{A}, P)$ 上の E 値確率変数とする．

(1) (E, d) が可分距離空間である（特に，E は第二可算である）とする．このとき，$(X_n)_{n \in \mathbb{N}}$ が X に概収束するならば，$(X_n)_{n \in \mathbb{N}}$ は X に確率収束する．
(2) $(X_n)_{n \in \mathbb{N}}$ が X に確率収束するならば，$(X_n)_{n \in \mathbb{N}}$ は X に概収束する部分列をもつ．

証明　(1)　$(X_n)_{n \in \mathbb{N}}$ が X に概収束するとする．$\epsilon > 0$ を任意にとり，各 $n \in \mathbb{N}$ に対して，$A_n = \{d(X_n, X) \geq \epsilon\}$ と置く．すると，$(\bigcup_{i=n}^{\infty} A_i)_{n \in \mathbb{N}}$ は事象の減少列であり，その全体の交叉 $\bigcap_{n \in \mathbb{N}} \bigcup_{i=n}^{\infty} A_i$ は $N = \{(X_n)_{n \in \mathbb{N}}$ は X に収束しない$\}$ に含まれる．仮定より，$P(N) = 0$ だから，$n \to \infty$ のとき

$$P(A_n) \leq P\left(\bigcup_{i=n}^{\infty} A_i\right) \to P(N) = 0$$

である（命題 1.3 (2)）．よって，$(X_n)_{n \in \mathbb{N}}$ は X に確率収束する．

(2)　$(X_n)_{n \in \mathbb{N}}$ が X に確率収束するとすると，自然数の狭義単調増加列 $(n_k)_{k \in \mathbb{N}}$ であって，任意の $k \in \mathbb{N}$ に対して

$$P^*(d(X_{n_k}, X) \geq 2^{-k}) < 2^{-k}$$

（P^* は P が定める外測度を表す）を満たすものがとれる．さらに，各 $k \in \mathbb{N}$ に対して，$\{d(X_{n_k}, X) \geq 2^{-k}\}$ を含む事象 A_k を，$P(A_k) \leq 2^{-k}$ を満たすようにとれる．このとき，Borel–Cantelli の補題（命題 1.10 (1)）より，「無限個の $k \in \mathbb{N}$ に対して事象 A_k が起こる」確率は 0 である．すなわち，ほとんど確実に，有限個を除くすべての $k \in \mathbb{N}$ に対して $d(X_{n_k}, X) < 2^{-k}$ となる．よって，$(X_{n_k})_{n \in \mathbb{N}}$ は X に概収束する．　□

2.2　確率測度の空間上の弱位相

本節の以下の部分では，位相空間 E 上の確率測度全体のなす集合を，$\mathscr{P}(E)$ と書く．

定義 2.3（弱位相）　位相空間 E 上の確率測度全体のなす集合 $\mathscr{P}(E)$ 上の位相であって，有界連続関数 $f \colon E \to \mathbb{R}$ に対する $\mathscr{P}(E)$ から $\mathbb{R}$ への写像 $\mu \mapsto \mu(f)$ をすべて連続にする最小のものを，$\mathscr{P}(E)$ の**弱位相**（weak topology）という．弱位相に関する収束を，**弱収束**（weak convergence）という．

$(\Omega, \mathfrak{A}, P)$ を確率空間とし，E を位相空間とする．$(\Omega, \mathfrak{A}, P)$ 上の E 値確率変数のネット $(X_i)_{i \in I}$ と E 値確率変数 X について，X_i の分布のなすネットが X の分布に弱収束するとき，単に $(X_i)_{i \in I}$ は X に**弱収束**[*8] するという．

確率変数のネット $(X_i)_{i \in I}$ が確率変数 X に弱収束することを，$X_i \xrightarrow{\mathrm{d}} X$ と書く．また，X の分布が μ であるとき，このことを，$X_i \xrightarrow{\mathrm{d}} \mu$ とも書く．

[*8] 確率変数については，「弱収束」のほかに，「分布収束（convergence in distribution）」や「法則収束（convergence in law）」ともいう．すぐ下に述べる記号 $X_i \xrightarrow{\mathrm{d}} X$ の「d」は，「distribution」の頭文字である．

位相空間 E が**正規**（normal）であるとは，互いに交わらない任意の閉集合 $F_1, F_2 \subseteq E$ に対して，互いに交わらない開集合 $G_1, G_2 \subseteq E$ であって，$F_1 \subseteq G_1$ かつ $F_2 \subseteq G_2$ を満たすものが存在することをいう．また，E が**完全正規**（perfectly normal）であるとは，E が正規であり，かつ E の任意の閉集合が可算個の開集合の交叉として書けることをいう．

命題 2.4 完全正規空間 E 上の確率測度 μ_1 と μ_2 が等しいための必要十分条件は，任意の有界連続関数 $f\colon E \to \mathbb{R}$ に対して $\mu_1(f) = \mu_2(f)$ であることである．

証明 必要性は明らかだから，十分性を示す．μ を E 上の確率測度として，閉集合 $F \subseteq E$ を任意にとる．E は完全正規だから，E の開集合の減少列 $(G_k)_{k\in\mathbb{N}}$ であって $\bigcap_{k\in\mathbb{N}} G_k = F$ を満たすものがとれる．さらに，Urysohn の補題より，各 $k \in \mathbb{N}$ に対して，連続関数 $f_k\colon E \to [0,1]$ であって $f_k(F) \subseteq \{1\}$ かつ $f_k(E \setminus G_k) \subseteq \{0\}$ を満たすものがとれる．このとき，$(f_k)_{k\in\mathbb{N}}$ は χ_F に各点収束するから，Lebesgue の収束定理より，

$$\mu(F) = \lim_{k\to\infty} \mu(f_k)$$

である．したがって，閉集合 $F \subseteq E$ に対する $\mu(F)$ の値は，有界連続関数 $f\colon E \to \mathbb{R}$ に対する $\mu(f)$ の値から一意に定まる．一方で，位相空間上の二つの有限測度は，すべての閉集合に対する値が一致すれば，全体で一致する（Dynkin 族補題から従う）．以上の二つのことから，主張が従う． □

系 2.5 完全正規空間 E に対して，$\mathscr{P}(E)$ の弱位相は Hausdorff である．

証明 命題 2.4 より，任意の異なる 2 点 $\mu_1, \mu_2 \in \mathscr{P}(E)$ に対して，ある有界連続関数 $f\colon E \to \mathbb{R}$ が存在して，$\mu_1(f) \neq \mu_2(f)$ となる．有界連続関数 $f\colon E \to \mathbb{R}$ に対して，$\mathscr{P}(E)$ 上の関数 $\mu \mapsto \mu(f)$ は弱位相に関して連続であり，その終域である $\mathbb{R}$ は Hausdorff だから，$\mathscr{P}(E)$ は Hausdorff である． □

補題 2.6 $(\Omega, \mathfrak{A}, \mu)$ を σ-有限測度空間とする．可測関数 $f\colon \Omega \to \overline{\mathbb{R}}_{\geq 0}$ に対して，$A_t = f^{-1}([t,\infty])$ $(t \in \mathbb{R}_{\geq 0})$ と置くと，$\mathbb{R}_{\geq 0}$ 上の $\overline{\mathbb{R}}_{\geq 0}$ 値関数 $t \mapsto \mu(A_t)$ は可測であり，

$$\int_\Omega f\, d\mu = \int_{\mathbb{R}} \mu(A_t)\, dt$$

が成り立つ．

証明 $\widetilde{A} = \{(\omega, t) \in \Omega \times \mathbb{R} \mid t \leq f(\omega)\}$ と置き，μ と Lebesgue 測度の積測度を $\widetilde{\mu}$ と書く．$\widetilde{A}$ は $\Omega \times \mathbb{R}$ の（積可測構造に関する）可測集合であり，$A_t = \{\omega \in \Omega \mid (\omega, t) \in \widetilde{A}\}$（$t \in \mathbb{R}_{\geq 0}$）である．よって，Tonelli の定理より，$\mathbb{R}_{\geq 0}$ 上の $\overline{\mathbb{R}}_{\geq 0}$ 値関数 $t \mapsto \mu(A_t)$ は可測であり，主張の等式の両辺は，ともに $\widetilde{\mu}(\widetilde{A})$ に等しい． □

定理 2.7（Portmanteau の定理[*9]） E を位相空間とする．E 上の確率測度の列 $(\mu_n)_{n\in\mathbb{N}}$ と確率測度 μ に関する次の 4 条件について，(b) $\Longleftrightarrow$ (c) $\Longrightarrow$ (d) $\Longrightarrow$ (a) が成り立つ．さらに，E が完全正規空間ならば，4 条件は同値である．

(a) $(\mu_n)_{n\in\mathbb{N}}$ は μ に弱収束する．
(b) 任意の閉集合 $F \subseteq E$ に対して，$\limsup_{n\to\infty} \mu_n(F) \leq \mu(F)$ である．
(c) 任意の開集合 $G \subseteq E$ に対して，$\liminf_{n\to\infty} \mu_n(G) \geq \mu(G)$ である．

[*9] 「Portmanteau」は，人名ではなく，もともとは「両開き式の旅行用鞄」を意味し，現在では主に「鞄語・混成語（複数の語のそれぞれの一部を組み合わせて作られた語）」という意味で使われる単語である．

(d) 任意の Borel 集合 $A \subseteq E$ に対して，$\mu(\partial A) = 0$ ならば $\lim_{n\to\infty} \mu_n(A) = \mu(A)$ である．

証明 (a) $\Longrightarrow$ (b)（E が完全正規空間である場合） $(\mu_n)_{n\in\mathbb{N}}$ が μ に弱収束するとして，E の閉集合 F を任意にとる．E は完全正規だから，E の開集合の減少列 $(G_k)_{k\in\mathbb{N}}$ であって $\bigcap_{k\in\mathbb{N}} G_k = F$ を満たすものがとれる．さらに，Urysohn の補題より，各 $k \in \mathbb{N}$ に対して，連続関数 $f_k\colon E \to [0,1]$ であって $f_k(F) \subseteq \{1\}$ かつ $f_k(E \setminus G_k) \subseteq \{0\}$ を満たすものがとれる．仮定より，任意の $k \in \mathbb{N}$ に対して，

$$\limsup_{n\to\infty} \mu_n(F) \leq \limsup_{n\to\infty} \mu_n(f_k) = \mu(f_k) \leq \mu(G_k)$$

である．$k \to \infty$ とすれば，$\limsup_{n\to\infty} \mu_n(F) \leq \mu(F)$ を得る．

(b) $\Longleftrightarrow$ (c) 補集合をとればよい．

(b) かつ (c) $\Longrightarrow$ (d) (b) と (c) が成り立つとすると，任意の Borel 集合 $A \subseteq E$ に対して，

$$\begin{aligned}
\mu(\overline{A}) &\geq \limsup_{n\to\infty} \mu_n(\overline{A}) \\
&\geq \limsup_{n\to\infty} \mu_n(A) \\
&\geq \liminf_{n\to\infty} \mu_n(A) \\
&\geq \liminf_{n\to\infty} \mu_n(A^\circ) \\
&\geq \mu(A^\circ)
\end{aligned}$$

である．さらに，$\mu(\partial A) = 0$，すなわち $\mu(\overline{A}) = \mu(A^\circ)$ ならば，上式ですべての等号が成立するから，$\lim_{n\to\infty} \mu_n(A) = \mu(A)$ となる．

(d) $\Longrightarrow$ (a) (d) が成り立つとする．連続関数 $f\colon E \to [0,1]$ を任意にとり，$A_t = f^{-1}([t,1])$（$t \in [0,1]$）と置く．すると，各 $t \in [0,1]$ に対して

$$\partial A_t = A_t \setminus A_t^\circ \subseteq f^{-1}([t,1]) \setminus f^{-1}((t,1]) = f^{-1}(\{t\})$$

だから，異なる t に対する ∂A_t は交わらない．特に，$P(\partial A_t) > 0$ となる t は可算個である．したがって，仮定と Lebesgue の収束定理より，

$$\lim_{n\to\infty} \mu_n(f) = \lim_{n\to\infty} \int_0^1 \mu_n(A_t)\,dt \to \int_0^1 \mu(A_t)\,dt = \mu(f)$$

である（補題 2.6）．よって，$(\mu_n)_{n\in\mathbb{N}}$ は μ に弱収束する． □

注意 2.8 Portmanteau の定理（定理 2.7）の証明で，$(\mu_n)_{n\in\mathbb{N}}$ がネットではなく点列であることが必要なのは，(d) $\Longrightarrow$ (a) で Lebesgue の収束定理を使っている箇所のみである．したがって，定理の主張のうち，(b) $\Longleftrightarrow$ (c) $\Longrightarrow$ (d) の部分と E が完全正規空間である場合の (a) $\Longrightarrow$ (b) の部分は，確率測度のネットに対しても同様に成り立つ．

注意 2.9 Portmanteau の定理（定理 2.7）の状況で，さらに，E 上に距離 d が定まっているとする．このとき，定理の (a) $\Longrightarrow$ (b) の証明における連続関数 $f_k\colon E \to [0,1]$ の代わりに，

$$g_k(x) = \max\{1 - kd(x,F), 0\} \qquad (x \in E)$$

で定まる一様連続関数 $g_k\colon E \to [0,1]$ を用いても，同じ議論ができる．したがって，距離空間 (E,d) に対しては，Portmanteau の定理（定理 2.7）の 4 条件は，次の条件とも同値である．

(a′) 任意の有界一様連続関数 $f\colon E \to \mathbb{R}$ に対して，$\lim_{n\to\infty} \mu_n(f) = \mu(f)$ である．

命題 2.10（写像定理） E を完全正規空間，E' を位相空間，$\phi\colon E \to E'$ を可測写像とし，ϕ の不連続点全体を D_ϕ と書く．E 上の確率測度の列 $(\mu_n)_{n\in\mathbb{N}}$ が確率測度 μ に弱収束し，D_ϕ が μ-無視可能ならば，E' 上の確率測度の列 $(\phi_*\mu_n)_{n\in\mathbb{N}}$ は $\phi_*\mu$ に弱収束する．

証明 主張の仮定が成り立つとして，D_ϕ を含む Borel 集合 $N \subseteq E$ であって $\mu(N) = 0$ を満たすものをとる．E' の任意の閉集合 F' に対して，$(\mu_n)_{n\in\mathbb{N}}$ が μ に弱収束することと Portmanteau の定理（定理 2.7），$\mu(N) = 0$ であること，$\overline{\phi^{-1}(F')} \setminus N \subseteq \overline{\phi^{-1}(F')} \setminus D_\phi \subseteq \phi^{-1}(\overline{F'}) = \phi^{-1}(F')$ であることを順に用いると，

$$
\begin{aligned}
\limsup_{n\to\infty} \mu_n(\phi^{-1}(F')) &\le \limsup_{n\to\infty} \mu_n(\overline{\phi^{-1}(F')}) \\
&\le \mu(\overline{\phi^{-1}(F')}) \\
&= \mu(\overline{\phi^{-1}(F')} \setminus N) \\
&\le \mu(\phi^{-1}(F'))
\end{aligned}
$$

を得る．よって，ふたたび Portmanteau の定理（定理 2.7）より，$(\phi_*\mu_n)_{n\in\mathbb{N}}$ は $\phi_*\mu$ に弱収束する． □

確率収束と弱収束の関係を述べる．

命題 2.11 $(\Omega, \mathfrak{A}, P)$ を確率空間とし，(E, d) を距離空間とする．$(X_n)_{n\in\mathbb{N}}$ を $(\Omega, \mathfrak{A}, P)$ 上の E 値確率変数の列，X を E 値確率変数とする．

(1) $(\Omega, \mathfrak{A}, P)$ 上の E 値確率変数の列 $(X_n)_{n\in\mathbb{N}}$ が E 値確率変数 X に確率収束するならば，$(X_n)_{n\in\mathbb{N}}$ は X に弱収束する．

(2) $(\Omega, \mathfrak{A}, P)$ 上の E 値確率変数の列 $(X_n)_{n\in\mathbb{N}}$ が定値確率変数 $x_0 \in E$ に確率収束することと弱収束することとは同値である．

証明 (1) $(X_n)_{n\in\mathbb{N}}$ が X に確率収束するとする．$f\colon E \to \mathbb{R}$ を有界一様連続関数として，$\epsilon > 0$ を任意にとると，ある $\delta > 0$ が存在して，$d(x, y) \le \delta$ を満たす任意の 2 点 x, $y \in E$ に対して $|f(x) - f(y)| \le \epsilon$ となる．さらに，$(X_n)_{n\in\mathbb{N}}$ は X に確率収束するから，十分大きい n に対しては $P(d(X_n, X) > \delta) \le \epsilon$ となる．このような n に対しては，

$$
\begin{aligned}
\left| \int_\Omega (f(X_n) - f(X))\, dP \right| &\le \int_\Omega |f(X_n) - f(X)|\, dP \\
&= \int_{\{d(X_n, X) \le \delta\}} |f(X_n) - f(X)|\, dP + \int_{\{d(X_n, X) > \delta\}} |f(X_n) - f(X)|\, dP \\
&\le \epsilon + 2\left(\sup_{x\in E} |f(x)| \right) \epsilon
\end{aligned}
$$

が成り立つ．したがって，$n \to \infty$ のとき

$$
\int_\Omega f(X_n)\, dP \to \int_\Omega f(X)\, dP
$$

である．これが任意の有界一様連続関数 $f\colon E \to \mathbb{R}$ に対して成り立つから，注意 2.9 より，$(X_n)_{n\in\mathbb{N}}$ は X に弱収束する．

(2)　$(X_n)_{n\in\mathbb{N}}$ が定値確率変数 x_0 に弱収束するとは，分布の列 $((X_n)_*P)_{n\in\mathbb{N}}$ が x_0 に集中した Dirac 測度 δ_{x_0} に弱収束するということであり，Portmanteau の定理（定理 2.7）より，これはさらに，E の任意の開集合 G に対して

$$\liminf_{n\to\infty} P(X_n\in G) \geq \begin{cases} 1 & (x_0\in G) \\ 0 & (x_0\notin G) \end{cases}$$

であることと同値である．上式は，$x_0\in G$ のときは $\lim_{n\to\infty} P(X_n\in G)=1$ を意味し，$x_0\notin G$ のときは常に正しい．よって，条件は，$(X_n)_{n\in\mathbb{N}}$ が定値確率変数 x_0 に確率収束することと同値である．　□

2.3　Prokhorov 距離

定義 2.12（Prokhorov 距離）　(E,d) を距離空間とする．$\mu,\nu\in\mathscr{P}(E)$ に対して

$$\pi(\mu,\nu)=\inf\left\{ r\in\mathbb{R}_{\geq 0} \;\middle|\; \begin{array}{l} \text{任意の Borel 集合 } A\subseteq E \text{ に対して,} \\ \mu(A)\leq\nu(B_d(A;r))+r \text{ かつ } \nu(A)\leq\mu(B_d(A;r))+r \end{array} \right\}$$

と定めるとき，π を，d が定める **Prokhorov 距離**（Prokhorov metric）という．

Prokhorov 距離が確かに $\mathscr{P}(E)$ 上の距離であることは，容易に確かめられる．

補題 2.13　(E,d) を距離空間とし，d が定める Prokhorov 距離を π と書く．$\mu,\nu\in\mathscr{P}(E)$ と $r\in\mathbb{R}_{\geq 0}$ について，条件「任意の Borel 集合 A に対して，$\mu(A)\leq\nu(B_d(A;r))+r$ である」が成り立つならば，$\pi(\mu,\nu)\leq\epsilon$ である．

証明　B を E の Borel 集合とする．$A=E\setminus B_d(B;r)$ が主張の仮定の不等式を満たすとすると，

$$\begin{aligned} \nu(B) &\leq \nu(E\setminus B_d(A;r)) \\ &= 1-\nu(B_d(A;r)) \\ &\leq 1-\mu(A)+r \\ &= \mu(B_d(B;r))+r \end{aligned}$$

となる．よって，主張の仮定の下で，$\pi(\mu,\nu)\leq r$ が成り立つ．　□

定理 2.14　(E,d) を距離空間とし，d が定める Prokhorov 距離を π と書く．

(1) π が定める $\mathscr{P}(E)$ 上の位相は，弱位相よりも細かい．
(2) E が可分ならば，π が定める $\mathscr{P}(E)$ 上の位相は，弱位相に等しい．

証明　(1)　$\mathscr{P}(E)$ 上の点列 $(\mu_n)_{n\in\mathbb{N}}$ と $\mu\in\mathscr{P}(E)$ が $\lim_{n\to\infty}\pi(\mu_n,\mu)=0$ を満たすとして，$(\mu_n)_{n\in\mathbb{N}}$ が μ に弱収束することを示せばよい．

F を E の閉集合とすると，任意の $\epsilon>0$ に対して，n が十分大きければ $\mu(B_d(F;\epsilon))\geq\mu_n(F)-\epsilon$ となるから，

$$\mu(B_d(F;\epsilon))\geq\limsup_{n\to\infty}\mu_n(F)-\epsilon$$

である．F が閉集合であることより $F=\bigcap_{\epsilon>0}B_d(F;\epsilon)$ だから，上式で極限 $\epsilon\to 0+$ をとれば，

$$\mu(F)=\lim_{\epsilon\to 0+}\mu(B_d(F;\epsilon))\geq\limsup_{n\to\infty}\mu_n(F)$$

を得る（命題 1.3 (2)）．よって，Portmanteau の定理（定理 2.7）より，$(\mu_n)_{n\in\mathbb{N}}$ は μ に弱収束する．

(2)　E が可分であるとする．$\mathscr{P}(E)$ 上のネット $(\mu_i)_{i\in I}$ が $\mu\in\mathscr{P}(E)$ に弱収束するとして，$\lim_i \pi(\mu_i,\mu)=0$ となることを示せばよい．

$\epsilon>0$ を任意にとる．E は可分距離空間であり，特に Lindelöf だから，直径 ϵ 以下の開球からなる E の可算被覆がとれる．ここから，E の直径 ϵ 以下の Borel 集合の有限族 $(B_j)_{j\in J}$ であって，どの異なる二つも交わらず，かつ

$$\mu\left(E\setminus\bigcup_{j\in J}B_j\right)\leq\epsilon \tag{$*$}$$

を満たすものが作れる．部分集合 $J'\subseteq J$ に対して，E の開集合 $G_{J'}$ を

$$G_{J'}=B_d^\circ\left(\bigcup_{j\in J'}B_j;\epsilon\right)$$

と定める．$(\mu_i)_{i\in I}$ は μ に弱収束するから，Portmanteau の定理（定理 2.7）と注意 2.8 より，十分大きい i に対しては，任意の部分集合 $J'\subseteq J$ に対して

$$\mu_i(G_{J'})\geq\mu(G_{J'})-\epsilon \tag{$**$}$$

が成り立つ．このような i に対して，$\pi(\mu_i,\mu)\leq 2\epsilon$ であることを示そう．E の Borel 集合 A を任意にとり，これに対して，$J'=\{j\in J\mid A\cap B_j\neq\emptyset\}$ と定める．すると，$(B_j)_{j\in J}$ のとり方より

$$A\cap\bigcup_{j\in J}B_j\subseteq G_{J'}\subseteq B_d(A;2\epsilon)$$

だから，($*$), ($**$) と合わせて，

$$\mu(A)\leq\mu(G_{J'})+\epsilon\leq\mu_i(G_{J'})+2\epsilon\leq\mu_i(B_d(A;2\epsilon))+2\epsilon$$

を得る．よって，補題 2.13 より，$\pi(\mu_i,\mu)\leq 2\epsilon$ である．

以上で，$\lim_i \pi(\mu_i,\mu)=0$ となることが示された．□

系 2.15　E が可分距離化可能空間ならば，$\mathscr{P}(E)$ は弱位相に関して距離化可能である．

証明　定理 2.14 (2) の結果である．□

2.4　緊密性

定義 2.16（緊密性）　E を位相空間とする．確率測度の集合 $\mathscr{S}\subseteq\mathscr{P}(E)$ が**緊密**（tight）であるとは，任意の $\epsilon>0$ に対して，あるコンパクト集合 $K\subseteq E$ が存在して，任意の $\mu\in\mathscr{S}$ に対して $\mu(K)\geq 1-\epsilon$ を満たすことをいう．

一般に，位相空間 E の部分集合 S について，S が E において**相対コンパクト**（relatively compact）であるとは，S を含む E のコンパクト集合が存在することをいい，S が E において**相対点列コンパクト**（relatively sequentially compact）であるとは，S 上の任意の点列が E における収束部分列をもつことをいう．E が距離化可能であれば，これらの 2 条件は同値である．

以下，緊密性と弱位相に関する相対（点列）コンパクト性との関係を述べる．

補題 2.17　E を位相空間とし，確率測度の集合 $\mathscr{S} \subseteq \mathscr{P}(E)$ は $\mathscr{P}(E)$ において弱位相に関して相対点列コンパクトであるとする．$(G_n)_{n\in\mathbb{N}}$ を E の開集合の増加列であって $\bigcup_{n\in\mathbb{N}} G_n = E$ を満たすものとすると，任意の $\epsilon > 0$ に対して，ある $n_0 \in \mathbb{N}$ が存在して，任意の $n \geq n_0$ と $\mu \in \mathscr{S}$ に対して，$\mu(G_n) \geq 1 - \epsilon$ となる．

証明　主張が成り立たないとすると，ある $\epsilon > 0$ と $\mathscr{S}$ 上の点列 $(\mu_n)_{n\in\mathbb{N}}$ が存在して，任意の $n \in \mathbb{N}$ に対して $\mu_n(G_n) \geq 1 - \epsilon$ となる．$\mathscr{S}$ は相対点列コンパクトだから，$(\mu_n)_{n\in\mathbb{N}}$ の弱収束部分列 $(\mu_{n_k})_{k\in\mathbb{N}}$（$(n_k)_{k\in\mathbb{N}}$ は狭義単調増加自然数列）がとれる．その弱収束極限を ν と置くと，Portmanteau の定理（定理 2.7）より，任意の $n \in \mathbb{N}$ に対して，

$$\nu(G_n) \leq \liminf_{k\to\infty} \mu_{n_k}(G_n) \leq \liminf_{k\to\infty} \mu_{n_k}(G_{n_k}) \leq 1 - \epsilon$$

となる．ところが，上式で $n \to \infty$ とすると，$1 = \nu(E) \leq 1 - \epsilon$ となり矛盾する．よって，背理法より，主張が成り立つ．　□

補題 2.18　E を距離化可能空間とする．E の任意のコンパクト集合の可算族 $\mathfrak{K}_0$ に対して，E のコンパクト集合の可算族 $\mathfrak{K}$ であって，次の 3 条件を満たすものが存在する．

(i) $\mathfrak{K}_0 \subseteq \mathfrak{K}$ である．
(ii) $\mathfrak{K}$ は有限合併で閉じている（特に，$\emptyset \in \mathfrak{K}$ である）．
(iii) 閉集合 $F \subseteq E$ が $\mathfrak{K}$ のある元に含まれるとする．このとき，F を含む任意の開集合 $G \subseteq E$ に対して，$K \in \mathfrak{K}$ であって $F \subseteq K \subseteq G$ を満たすものが存在する．

証明　各 $C \in \mathfrak{K}_0$ はコンパクト距離化可能だから，第二可算であり，可算開基 $\mathfrak{O}_C$ がとれる．$\bigcup_{C\in\mathfrak{K}_0} \mathfrak{O}_C$ の元の閉包の有限合併全体のなす集合を $\mathfrak{K}$ と置く．明らかに，$\mathfrak{K}$ は E のコンパクト集合の可算族であり，条件 (i), (ii) を満たす．$\mathfrak{K}$ が条件 (iii) を満たすことを示す．閉集合 $F \subseteq E$ が $\mathfrak{K}$ のある元に含まれるとして，F を含む開集合 $G \subseteq E$ を任意にとる．F は $\mathfrak{K}$ のある元に含まれるから，有限部分族 $\mathfrak{K}_1 \subseteq \mathfrak{K}_0$ が存在して $F \subseteq \bigcup_{C\in\mathfrak{K}_1} C$ となる．さらに，各 $C \in \mathfrak{K}_1$ に対して，コンパクト集合 $F \cap C$ は C の開集合 $G \cap C$ に含まれ，$\mathfrak{O}_C$ は C の開基であり，C は正則だから，有限部分族 $\mathfrak{O}'_C \subseteq \mathfrak{O}_C$ が存在して

$$F \cap C \subseteq \bigcup_{O\in\mathfrak{O}'_C} O \subseteq \bigcup_{O\in\mathfrak{O}'_C} \overline{O} \subseteq G \cap C$$

となる．これらを用いて $K = \bigcup_{C\in\mathfrak{K}_1} \bigcup_{O\in\mathfrak{O}'_C} \overline{O}$ と置くと，$K \in \mathfrak{K}$ かつ $F \subseteq K \subseteq G$ である．よって，$\mathfrak{K}$ は条件 (iii) を満たす．　□

補題 2.19　正規空間 E の任意の有限開被覆 $(G_i)_{i\in I}$ に対して，それを細分する E の閉被覆 $(F_i)_{i\in I}$（すなわち，E の閉被覆 $(F_i)_{i\in I}$ であって，任意の $i \in I$ に対して $F_i \subseteq G_i$ を満たすもの）が存在する．

証明　$I = \{1, \dots, n\}$ とし，$n \in \mathbb{N}$ に関する帰納法で主張を示す．$n = 0$ のときは明らかである．$n \geq 1$ とし，$n - 1$ のとき主張は正しいとする．E の開集合 $G_1, \dots, G_n$ が E を被覆するとする．$G' = G_1 \cup \cdots \cup G_{n-1}$ と置くと，$E \setminus G_n$ と $E \setminus G'$ は互いに交わらない E の閉集合だから，E の正規性より，互いに交わらない E の開集合 U', U_n であって，$E \setminus G_n \subseteq U'$ かつ $E \setminus G' \subseteq U_n$ を満たすものがとれる．これらの補集合をそれぞれ F', F_n と置けば，(F', F_n) は (U', U_n) を細分する E の閉被覆である．さらに，正規空間 F' の開被覆 $(G_i \cap F')_{1\leq i\leq n-1}$ に帰納法の仮定を適用して，これを細分する F' の閉被覆 $(F_i)_{1\leq i\leq n-1}$ を得る．以上で，$(U_i)_{1\leq i\leq n}$ を細分する E の閉被覆 $(F_i)_{1\leq i\leq n}$ が構成できた．よって，主張は n のときも正しい．これで，帰納法が完成した．　□

定理 2.20（Prokhorov の定理） E を位相空間とする．確率測度の集合 $\mathscr{S} \subseteq \mathscr{P}(E)$ に関する次の 3 条件について，E が距離化可能ならば (a) $\Longrightarrow$ (b) が成り立ち，E が可分距離化可能ならば (b) $\Longleftrightarrow$ (c) が成り立ち，E がポーランド空間ならば 3 条件は同値である．

(a) $\mathscr{S}$ は緊密である．
(b) $\mathscr{S}$ は $\mathscr{P}(E)$ において弱位相に関して相対点列コンパクトである．
(c) $\mathscr{S}$ は $\mathscr{P}(E)$ において弱位相に関して相対コンパクトである．

証明 (b) $\Longleftrightarrow$ (c)（E が可分距離化可能である場合） $\mathscr{P}(E)$ の弱位相が距離化可能であること（系 2.15）から従う．

(b) $\Longrightarrow$ (a)（E がポーランド空間である場合） E の位相と整合する完備な距離 d をとる．E が可分であることより，各 $k \in \mathbb{N}$ に対して，d に関する半径 2^{-k} の開球の可算族 $\mathfrak{A}_k$ であって E を被覆するものがとれる．$\mathscr{S}$ が相対点列コンパクトであるとして，$\epsilon > 0$ を任意にとる．すると，補題 2.17 より，各 $k \in \mathbb{N}$ に対して，$\mathfrak{A}_k$ の元の有限合併として表せる集合 A_k であって，任意の $\mu \in \mathscr{S}$ に対して $\mu(A_k) \geq 1 - 2^{-k}\epsilon$ を満たすものがとれる．これらを用いて $K = \overline{\bigcap_{k=1}^{\infty} A_k}$ と置く．定義より $\bigcap_{k=1}^{\infty} A_k$ は d に関して全有界であり，d は完備だから，K はコンパクトである．また，K は，任意の $\mu \in \mathscr{S}$ に対して

$$\mu(K) \geq \mu\left(\bigcap_{k=0}^{\infty} A_k\right) \geq \mu\left(E \setminus \bigcup_{k=0}^{\infty}(E \setminus A_k)\right) \geq 1 - \sum_{k=0}^{\infty}(1 - \mu(A_k)) = 1 - 2\epsilon$$

を満たす．よって，$\mathscr{S}$ は緊密である．

(a) $\Longrightarrow$ (b)（E が距離化可能である場合） $\mathscr{S}$ が緊密であるとすると，各 $k \in \mathbb{N}$ に対して，コンパクト集合 $C_k \subseteq E$ であって，任意の $\mu \in \mathscr{S}$ に対して $\mu(C_k) \geq 1 - 2^{-k}$ を満たすものがとれる．$\mathfrak{K}_0 = \{C_k \mid k \in \mathbb{N}\}$ と置き，これに対して，E のコンパクト集合の可算族 $\mathfrak{K}$ であって補題 2.18 の条件を満たすものをとる．

$\mathscr{S}$ 上の点列 $(\mu_n)_{n\in\mathbb{N}}$ を任意にとる．$\mathfrak{K}$ が可算であることより積空間 $[0,1]^{\mathfrak{K}}$ は点列コンパクトだから，$(\mu_n)_{n\in\mathbb{N}}$ の部分列 $(\mu_{n_k})_{k\in\mathbb{N}}$（$(n_k)_{k\in\mathbb{N}}$ は狭義単調増加自然数列）であって，任意の $K \in \mathfrak{K}$ に対して極限

$$\alpha(K) = \lim_{k\to\infty} \mu_{n_k}(K) \in [0,1]$$

が存在するものがとれる．以下，$(\mu_{n_k})_{k\in\mathbb{N}}$ の弱収束極限を構成する．

主張 2.21 上記の α について，次が成り立つ．

(1) $K_1, K_2 \in \mathfrak{K}$ について，$K_1 \subseteq K_2$ ならば $\alpha(K_1) \leq \alpha(K_2)$ である．
(2) 任意の有限個の元 $K_1, \dots, K_m \in \mathfrak{K}$ に対して，$\alpha(K_1 \cup \cdots \cup K_m) \leq \alpha(K_1) + \cdots + \alpha(K_m)$ である．$K_1, \dots, K_m$ のどの二つも交わらないならば，等号が成立する．（特に，$\alpha(\emptyset) = 0$ である．）

主張 2.21 の証明 明らかである． //

主張 2.22 開集合 $G \subseteq E$ に対して

$$\beta(G) = \sup\{\alpha(K) \mid K \in \mathfrak{K} \text{ は } G \text{ に含まれる}\}$$

と定めると，次が成り立つ．

(1) $\beta(E) = 1$ である．
(2) E の任意の開集合族 $(G_i)_{i\in I}$ に対して，$\beta(\bigcup_{i\in I} G_i) \leq \sum_{i\in I} \beta(G_i)$ である．

主張 2.22 の証明　(1)　任意の $k \in \mathbb{N}$ に対して $\beta(E) \geq \alpha(C_k) \geq 1 - 2^{-k}$ だから，$\beta(E) = 1$ である．

(2)　$(G_i)_{i \in I}$ を E の開集合族とする．$K \in \mathfrak{K}$ が $\bigcup_{i \in I} G_i$ に含まれるとすると，K のコンパクト性より，有限部分集合 $I_0 \subseteq I$ であって $K \subseteq \bigcup_{i \in I_0} G_i$ を満たすものがとれる．これに対して，補題 2.19 より，$(G_i \cap K)_{i \in I_0}$ を細分する K の閉被覆 $(F_i)_{i \in I_0}$ がとれる．さらに，$\mathfrak{K}$ のとり方より，各 $i \in I$ に対して，$K_i \in \mathfrak{K}$ であって $F_i \subseteq K_i \subseteq G_i$ を満たすものがとれる．よって，主張 2.21 より，

$$\alpha(K) \leq \alpha\left(\bigcup_{i \in I_0} K_i\right) \leq \sum_{i \in I_0} \alpha(K_i) \leq \sum_{i \in I_0} \alpha(G_i) \leq \sum_{i \in I} \alpha(G_i)$$

である．これが $\bigcup_{i \in I} G_i$ に含まれる任意の $K \in \mathfrak{K}$ に対して成り立つから，$\beta(\bigcup_{i \in I} G_i) \leq \sum_{i \in I} \beta(G_i)$ である．　//

主張 2.23　部分集合 $A \subseteq E$ に対して

$$\gamma(A) = \inf\{\beta(G) \mid G \subseteq E \text{ は } A \text{ を含む開集合}\}$$

と定めると，次が成り立つ．

(1) γ は E 上の外測度である．すなわち，次の二つが成り立つ．
- 部分集合 $A_1, A_2 \subseteq E$ について，$A_1 \subseteq A_2$ ならば $\gamma(A_1) \leq \gamma(A_2)$ である（単調性）．
- E の部分集合の可算族 $(A_i)_{i \in I}$ について，$\gamma(\bigcup_{i \in I} A_i) \leq \sum_{i \in I} \gamma(A_i)$ である（可算劣加法性）．

(2) E の任意の Borel 集合は，γ-可測である．

主張 2.23 の証明　(1)　γ が単調であることと，$\gamma(\emptyset) = 0$ であることは明らかである．あとは，E の部分集合の列 $(A_i)_{i \in \mathbb{N}}$ について，

$$\gamma\left(\bigcup_{i=0}^{\infty} A_i\right) \leq \sum_{i=0}^{\infty} \gamma(A_i) \qquad (*)$$

$\epsilon > 0$ を任意にとる．すると，各 $i \in \mathbb{N}$ に対して，A_i を含む開集合 $G_i \subseteq E$ であって，$\beta(G_i) \leq \gamma(A_i) + 2^{-i}\epsilon$ を満たすものがとれる．このとき，主張 2.22 (2) より，

$$\gamma\left(\bigcup_{i=0}^{\infty} A_i\right) \leq \beta\left(\bigcup_{i=0}^{\infty} G_i\right) \leq \sum_{i=0}^{\infty} \beta(G_i) \leq \sum_{i=0}^{\infty} \gamma(G_i) + 2\epsilon$$

である．任意の $\epsilon > 0$ に対してこれが成り立つから，等式 ($*$) が成り立つ．

(2)　まず，任意の閉集合 $F \subseteq E$ と開集合 $G \subseteq E$ に対して

$$\beta(G) \geq \gamma(G \cap F) + \gamma(G \setminus F) \qquad (**)$$

であることを示す．$\epsilon > 0$ を任意にとる．すると，$G \setminus F$ に含まれる $K_2 \in \mathfrak{K}$ を，$\alpha(K_2) \geq \beta(G \setminus F) - \epsilon$ を満たすようにとれる．さらに，$G \setminus K_2$ に含まれる $K_1 \in \mathfrak{K}$ を，$\alpha(K_1) \geq \beta(G \setminus K_2) - \epsilon$ を満たすようにとれる．このとき，主張 2.21 (2) より，

$$\begin{aligned}
\beta(G) &\geq \alpha(K_1 \cup K_2) \\
&= \alpha(K_1) + \alpha(K_2) \\
&\geq \beta(G \setminus K_2) + \beta(G \setminus F) - 2\epsilon \\
&\geq \gamma(G \cap F) + \gamma(G \setminus F) - 2\epsilon
\end{aligned}$$

である．任意の $\epsilon > 0$ に対してこれが成り立つから，等式 (**) が成り立つ．

次に，閉集合 $F \subseteq E$ と部分集合 $A \subseteq E$ を任意にとる．A を含む任意の開集合 $G \subseteq E$ について，前段の結果より

$$\beta(G) \geq \gamma(G \cap F) + \gamma(G \setminus F) \geq \gamma(A \cap F) + \gamma(A \setminus F)$$

だから，$\gamma(A) \geq \gamma(A \cap F) + \gamma(A \setminus F)$ である．すなわち，F は γ-可測である．よって，E の任意の閉集合は γ-可測であり，したがって，E の任意の Borel 集合は γ-可測である． //

主張 2.22 (1) と主張 2.23 より，γ の Borel 集合族への制限は，E 上の確率測度である．$G \subseteq E$ を開集合とすると，G に含まれる任意の $K \in \mathfrak{K}$ に対して

$$\alpha(K) = \lim_{k\to\infty} \mu_{n_k}(K) \leq \liminf_{k\to\infty} \mu_{n_k}(G)$$

だから，

$$\gamma(G) = \beta(G) = \sup\{\alpha(K) \mid K \in \mathfrak{K} \text{ は } G \text{ に含まれる}\} \leq \liminf_{k\to\infty} \mu_{n_k}(G)$$

である．よって，Portmanteau の定理（定理 2.7）より，$(\mu_{n_k})_{k\in\mathbb{N}}$ はこの確率測度に弱収束する． □

2.5 Lévy の連続性定理

補題 2.24 μ を $\mathbb{R}^d$ 上の確率測度とし，$\delta > 0$ とすると，

$$\mu\left(\mathbb{R}^d \setminus \left[-\frac{2}{\delta}, \frac{2}{\delta}\right]^d\right) \leq \frac{2}{(2\delta)^d} \int_{[-\delta,\delta]^d} (1 - \phi_\mu(t))\, dt$$

が成り立つ（右辺が実数であることも主張に含む）．

証明 特性関数 ϕ_μ の $[-\delta, \delta]^d$ 上の平均値は，

$$\begin{aligned}
\frac{1}{(2\delta)^d} \int_{[-\delta,\delta]^d} \phi_\mu(t)\, dt &= \frac{1}{(2\delta)^d} \int_{[-\delta,\delta]^d} \int_{\mathbb{R}^d} e^{itx}\, d\mu(x)\, dt \\
&= \frac{1}{(2\delta)^d} \int_{[-\delta,\delta]^d} \prod_{k=1}^{d} \left(\int_{-\delta}^{\delta} e^{it_k x_k}\, dt_k\right) d\mu(x_1, \ldots, x_d) \\
&= \frac{1}{(2\delta)^d} \int_{[-\delta,\delta]^d} \prod_{k=1}^{d} \left(\frac{2\sin(\delta x_k)}{x_k}\right) d\mu(x_1, \ldots, x_d) \\
&= \int_{[-\delta,\delta]^d} \prod_{k=1}^{d} \left(\frac{\sin(\delta x_k)}{\delta x_k}\right) d\mu(x_1, \ldots, x_d)
\end{aligned}$$

と表せる（関数 $y \mapsto (\sin y)/y$ は，$y = 0$ にも連続に拡張して考えるものとする．以下同様）．また，容易に確かめられるように，$y \in \mathbb{R}$ に対して $(\sin y)/y \leq 1$ であり，$|y| \geq 1/2$ ならば $(\sin y)/y \leq 1/2$ である．よって，

$$\begin{aligned}
\frac{1}{(2\delta)^d} \int_{[-\delta,\delta]^d} (1 - \phi_\mu(t))\, dt &= \int_{\mathbb{R}^d} \left(1 - \prod_{k=1}^{d} \frac{\sin(\delta x_k)}{\delta x_k}\right) d\mu(x_1, \ldots, x_d) \\
&\geq \frac{1}{2}\mu\left(\mathbb{R}^d \setminus \left[-\frac{2}{\delta}, \frac{2}{\delta}\right]^d\right)
\end{aligned}$$

である． □

定理 2.25（Lévy の連続性定理） V を有限次元実線型空間とし，$(\mu_n)_{n\in\mathbb{N}}$ をその上の確率測度の列とする．特性関数の列 $(\phi_{\mu_n})_{n\in\mathbb{N}}$ が，関数 $\phi\colon V^*\to\mathbb{C}$ に各点収束するとする．このとき，次の 4 条件は同値である．

(a) V 上の確率測度の集合 $\{\mu_n \mid n\in\mathbb{N}\}$ は緊密である．
(b) $(\mu_n)_{n\in\mathbb{N}}$ は V 上のある確率測度 μ に弱収束する．
(c) ϕ は V 上のある確率測度 μ の特性関数に等しい．
(d) ϕ は連続である．
(e) ϕ は点 $0\in V^*$ において連続である．

さらに，これらの条件の下で，条件 (b) と (c) の確率測度 μ は一致する．

証明 (b) $\Longrightarrow$ (c) および最後の主張　$(\mu_n)_{n\in\mathbb{N}}$ が V 上の確率測度 μ に弱収束するならば，

$$\phi(t)=\lim_{n\to\infty}\phi_{\mu_n}(t)=\lim_{n\to\infty}\int_V e^{i\langle t,x\rangle}\,d\mu_n(x)=\int_V e^{i\langle t,x\rangle}\,d\mu(x)=\phi_\mu(t)\qquad(t\in V^*)$$

である．さらに，特性関数が確率測度を特徴付けること（定理 1.31）より，(b) と (c) の確率測度 μ は一致する．

(c) $\Longrightarrow$ (d)　一般に特性関数が連続であること（命題 1.24）の結果である．

(d) $\Longrightarrow$ (e)　明らかである．

(e) $\Longrightarrow$ (a)　一般性を失わず，$V=\mathbb{R}^d$ であるとし，特性関数や ϕ を $\mathbb{R}^d$ 上の関数とみなす．ϕ は連続関数列 $(\phi_{\mu_n})_{n\in\mathbb{N}}$ の各点収束極限だから，可測である．ϕ が点 $0\in\mathbb{R}^d$ において連続であるとすると，任意の $\epsilon>0$ に対して，十分小さい $\delta>0$ をとれば，

$$\frac{1}{(2\delta)^d}\int_{[-\delta,\delta]^d}|1-\phi(t)|\,dt<\epsilon \tag{$*$}$$

となる．$(\phi_{\mu_n})_{n\in\mathbb{N}}$ は ϕ に各点収束するから，Lebesgue の収束定理より，不等式 ($*$) は，ϕ を十分大きい n に対する ϕ_{μ_n} に置き換えても成り立つ．このような n に対しては，補題 2.24 より，

$$\mu_n\left(\mathbb{R}^d\setminus\left[-\frac{2}{\delta},\frac{2}{\delta}\right]^d\right)<2\epsilon \tag{$**$}$$

である．さらに，必要に応じて δ を小さくとり直すことで，不等式 ($**$) が残る有限個の n に対しても成り立つようにできる．よって，$\{\mu_n \mid n\in\mathbb{N}\}$ は緊密である．

(a) $\Longrightarrow$ (b)　$\{\mu_n \mid n\in\mathbb{N}\}$ が緊密であるとすると，Prokhorov の定理（定理 2.20）より，$(\mu_n)_{n\in\mathbb{N}}$ の弱収束部分列がとれる．その弱収束極限を μ と置くと，$\phi=\phi_\mu$ である（上記の (b) $\Longrightarrow$ (c) の証明を参照のこと）．

$(\mu_n)_{n\in\mathbb{N}}$ が μ に弱収束しないと仮定して，矛盾を導こう．このとき，ある有界連続関数 $f\colon E\to\mathbb{R}$ が存在して，$(\mu_n(f))_{n\in\mathbb{N}}$ は $\mu(f)$ に収束しない．一方で，任意の $n\in\mathbb{N}$ に対して $|\mu_n(f)|\leq\|f\|_\infty$ だから，$(\mu_n(f))_{n\in\mathbb{N}}$ の収束部分列 $(\mu_n(f))_{n\in S_1}$（S_1 は $\mathbb{N}$ の無限部分集合）がとれる．この部分列は，$\mu(f)$ とは異なる値に収束する．さらに，ふたたび $\{\mu_n \mid n\in\mathbb{N}\}$ の緊密性と Prokhorov の定理（定理 2.20）より，$(\mu_n)_{n\in S_1}$ の弱収束部分列 $(\mu_n)_{n\in S_2}$（S_2 は S_1 の無限部分集合）がとれる．その弱収束極限を ν と置くと，各点収束の意味で

$$\phi_\nu=\lim_{n\in S_2,\,n\to\infty}\phi_{\mu_n}=\phi=\phi_\mu$$

だから（ふたたび (b) $\Longrightarrow$ (c) を用いた），$\nu=\mu$ である（定理 1.31）．ところが一方で，

$$\nu(f)=\lim_{n\in S_2,\,n\to\infty}\mu_n(f)\neq\mu(f)$$

であり，これは矛盾である．よって，背理法より，$(\mu_n)_{n\in\mathbb{N}}$ は μ に弱収束する． □

系 2.26　有限次元実線型空間 V 上の確率測度の列 $(\mu_n)_{n\in\mathbb{N}}$ と確率測度 μ に対して，次の 2 条件は同値である．

(a) $(\mu_n)_{n\in\mathbb{N}}$ は μ に弱収束する．
(b) 特性関数の列 $(\phi_{\mu_n})_{n\in\mathbb{N}}$ は ϕ_μ に各点収束する．

証明　Lévy の連続性定理（定理 2.25）の (b) $\iff$ (c) と最後の主張から従う． □

3　大数の法則

3.1　大数の弱法則

定理 3.1　$(X_i)_{i\in\mathbb{N}_{>0}}$ を確率空間 $(\Omega,\mathfrak{A},P)$ 上の 2 乗可積分な実確率変数の独立列とする．この列が

$$\lim_{n\to\infty}\frac{1}{n^2}\sum_{i=1}^{n}\mathrm{Var}[X_i]=0$$

を満たすならば，$n\to 0$ のとき

$$\frac{1}{n}\sum_{i=1}^{n}X_i-\frac{1}{n}\sum_{i=1}^{n}E[X_i]\xrightarrow{\mathrm{p}}0$$

である．

証明　$\epsilon>0$ を任意にとる．$(1/n)\sum_{i=1}^{n}X_i$ の平均は $(1/n)\sum_{i=1}^{n}E[X_i]$，分散は $(1/n^2)\sum_{i=1}^{n}\mathrm{Var}[X_i]$ だから，Chebyshev の不等式（系 1.21）より，

$$P\left(\frac{1}{n}\left|\sum_{i=1}^{n}X_i-\sum_{i=1}^{n}E[X_i]\right|\geq\epsilon\right)\leq\frac{1}{\epsilon^2n^2}\sum_{i=1}^{\infty}\mathrm{Var}[X_i]=0$$

である．主張の仮定が成り立つとすると，$n\to\infty$ のとき上式の右辺は 0 に収束する．すなわち，主張の確率収束が成り立つ． □

系 3.2（大数の弱法則）　$(X_i)_{i\in\mathbb{N}_{>0}}$ を確率空間 $(\Omega,\mathfrak{A},P)$ 上の 2 乗可積分な実確率変数の独立同分布列とし，これらの共通の期待値を μ と書く．すると，$n\to\infty$ のとき

$$\frac{1}{n}\sum_{i=1}^{n}X_i\xrightarrow{\mathrm{p}}\mu$$

である． □

3.3 節で，大数の弱法則よりも強い主張である，大数の強法則を示す．

3.2　大数の強法則の証明の準備

本小節では，大数の強法則の証明に用いる命題・補題のうち，議論の本筋とは比較的独立しているものをまとめる．

補題 3.3（Kronecker の補題） $(a_i)_{i\in\mathbb{N}_{>0}}$ を実数列とし，$(b_i)_{i\in\mathbb{N}_{>0}}$ を正の実数の増加列とする．$\sum_{i=1}^{n} a_i/b_i$ が $n\to\infty$ のとき（有限値に）収束するならば，

$$\lim_{n\to\infty}\frac{1}{b_n}\sum_{i=1}^{n}a_i = 0$$

である．

証明 $A_n=\sum_{i=1}^{n}a_i/b_i$ と置き，これが $n\to\infty$ のとき $A\in\mathbb{R}$ に収束するとする．Abel の総和公式より

$$\begin{aligned}\frac{1}{b_n}\sum_{i=1}^{n}a_i &= \frac{1}{b_n}\sum_{i=1}^{n}\frac{a_i}{b_i}b_i\\ &= \frac{1}{b_n}\left(A_nb_n-\sum_{i=1}^{n-1}A_i(b_{i+1}-b_i)\right)\\ &= A_n-\frac{1}{b_n}\sum_{i=1}^{n-1}A_i(b_{i+1}-b_i)\end{aligned}$$

だから，主張を示すためには，

$$\lim_{n\to\infty}\frac{1}{b_n}\sum_{i=1}^{n-1}A_i(b_{i+1}-b_i)=A \tag{$*$}$$

をいえばよい．任意の $\epsilon>0$ に対して，正の整数 n_0 が存在して，任意の整数 $i\geq n_0$ に対して $|A_i-A|\leq\epsilon$ となる．このような n_0 をとると，任意の整数 $n>n_0$ に対して

$$\begin{aligned}\left|\frac{1}{b_n}\sum_{i=1}^{n-1}A_i(b_{i+1}-b_i)\right| &= \left|\frac{1}{b_n}\sum_{i=1}^{n-1}(A_i-A)(b_{i+1}-b_i)-\frac{b_1}{b_n}A\right|\\ &\leq \frac{1}{b_n}\sum_{i=1}^{n-1}|A_i-A|(b_{i+1}-b_i)+\frac{b_1}{b_n}A\\ &\leq \frac{1}{b_n}\sum_{i=1}^{n_0}|A_i-A|(b_{i+1}-b_i)+\frac{b_n-b_{n_0+1}}{b_n}\epsilon+\frac{b_1}{b_n}A\end{aligned}$$

となり，n が十分大きければ，上式の最右辺は 2ϵ 以下となる．これで，($*$) が示された． □

補題 3.4 $(\Omega,\mathfrak{A},\mu)$ を測度空間とする．$f\colon\Omega\to\overline{\mathbb{R}}_{\geq 0}$ を可測関数とし，$C>0$ とすると，

$$C\sum_{i=1}^{\infty}\mu(f^{-1}([iC,\infty]))\leq\int_{\Omega}f\,d\mu\leq C\sum_{i=0}^{\infty}\mu(f^{-1}((iC,\infty]))$$

である．

証明 可測関数 $g_+,g_-\colon E\to\overline{\mathbb{R}}_{\geq 0}$ を

$$\begin{aligned}g_+(x)&=\inf\{iC\mid i\in\mathbb{N},\ f(x)\leq iC\},\\ g_-(x)&=\sup\{iC\mid i\in\mathbb{N},\ f(x)\geq iC\}\end{aligned}$$

と定めると，$g_-\leq f\leq g_+$ である．$g_+=C\sum_{i=0}^{\infty}\chi_{f^{-1}((iC,\infty])}$，$g_-=C\sum_{i=1}^{\infty}\chi_{f^{-1}([iC,\infty])}$ と書けることに注意して，この不等式の各辺を積分すれば，主張の不等式を得る． □

3.3 大数の強法則

補題 3.5（Kolmogorov の不等式） $X_1, \ldots, X_n$ を確率空間 $(\Omega, \mathfrak{A}, P)$ 上の 2 乗可積分な実確率変数とし，これらは独立であり，これらの期待値はすべて 0 であるとする．各 $i \in \{1, \ldots, n\}$ に対して，$S_i = X_1 + \cdots + X_i$ と置く．このとき，任意の $\epsilon > 0$ に対して，

$$P\left(\max_{1 \leq i \leq n} |S_i| \geq \epsilon\right) \leq \frac{1}{\epsilon^2} \mathrm{Var}[S_n]$$

が成り立つ．

証明 各 $i \in \{1, \ldots, n\}$ に対して

$$A_i = \{|S_1|, \ldots, |S_{i-1}| < \epsilon,\ |S_i| \geq \epsilon\}$$

と置くと，$\{\max_{1 \leq i \leq n} |S_i| \geq \epsilon\} = \coprod_{i=1}^n A_i$ であり，A_i 上では $\chi_{A_i} S_i^2 \geq \epsilon^2$ だから，

$$P\left(\max_{1 \leq i \leq n} |S_i| \geq \epsilon\right) = \sum_{i=1}^n P(A_i) \leq \frac{1}{\epsilon^2} \sum_{i=1}^n E[\chi_{A_i} S_i^2] \tag{$*$}$$

である．また，各 $i \in \{1, \ldots, n\}$ に対して，$\chi_{A_i} S_i$ と $S_n - S_i$ は独立だから，

$$\begin{aligned}
E[\chi_{A_i} S_n^2] &= E[\chi_{A_i} S_i^2] + 2E[\chi_{A_i} S_i (S_n - S_i)] + E[\chi_{A_i} (S_n - S_i)^2] \\
&= E[\chi_{A_i} S_i^2] + 2E[\chi_{A_i} S_i] E[S_n - S_i] + E[\chi_{A_i} (S_n - S_i)^2] \\
&= E[\chi_{A_i} S_i^2] + E[\chi_{A_i} (S_n - S_i)^2] \\
&\geq E[\chi_{A_i} S_i^2]
\end{aligned}$$

である（系 1.19）．したがって，

$$\sum_{i=1}^n E[\chi_{A_i} S_i^2] \leq \sum_{i=1}^n E[\chi_{A_i} S_n^2] \leq E[S_n^2] = \mathrm{Var}[S_n] \tag{$**$}$$

である．$(*)$ と $(**)$ より，主張の不等式が成り立つ． □

補題 3.6 $(X_i)_{i \in \mathbb{N}_{>0}}$ を確率空間 $(\Omega, \mathfrak{A}, P)$ 上の 2 乗可積分な実確率変数の独立列とする．この列が $\sum_{i=1}^\infty E[X_i] < \infty$ かつ $\sum_{i=1}^\infty \mathrm{Var}[X_i] < \infty$ を満たすならば，確率変数列 $(\sum_{i=1}^n X_i)_{n \in \mathbb{N}_{>0}}$ は概収束する．

証明 $\sum_{i=1}^\infty E[X_i] < \infty$ だから，$X_i - E[X_i]$ を改めて X_i と置き直すことで，一般性を失わず，$E[X_i] = 0$（$i \in \mathbb{N}_{>0}$）と仮定する．Kolmogorov の不等式（補題 3.5）より，任意の $\epsilon > 0$ と整数 $1 \leq m \leq n$ に対して，

$$P\left(\max_{m \leq b \leq n} |X_m + \cdots + X_b| \geq \epsilon\right) \leq \frac{1}{\epsilon^2} \sum_{i=m}^n \mathrm{Var}[X_i]$$

である．さらに，a, b を $m \leq a \leq b \leq n$ を満たす整数とすると $|X_a + \cdots + X_b| \leq |X_m + \cdots + X_b| + |X_m + \cdots + X_{a-1}|$ だから，上式から

$$P\left(\max_{m \leq a \leq b \leq n} |X_a + \cdots + X_b| \geq 2\epsilon\right) \leq \frac{1}{\epsilon^2} \sum_{i=m}^n \mathrm{Var}[X_i] \leq \frac{1}{\epsilon^2} \sum_{i=m}^\infty \mathrm{Var}[X_i]$$

を得る．ここで，$n \to \infty$ とすると

$$P\left(\sup_{m \le a \le b} |X_a + \cdots + X_b| \ge 3\epsilon\right) \le P\left(\bigcup_{n=m}^{\infty}\left\{\max_{m \le a \le b \le n} |X_a + \cdots + X_b| \ge 2\epsilon\right\}\right) \le \frac{1}{\epsilon^2}\sum_{i=m}^{\infty} \mathrm{Var}[X_i]$$

となり（命題 1.3 (1)），続けて $m \to \infty$ とすると，$\sum_{i=1}^{\infty} \mathrm{Var}[X_i] < \infty$ より

$$P\left(\lim_{m\to\infty}\sup_{m \le a \le b} |X_a + \cdots + X_b| \ge 3\epsilon\right) = P\left(\bigcap_{m=1}^{\infty}\left\{\sup_{m \le a \le b} |X_a + \cdots + X_b| \ge 3\epsilon\right\}\right) = 0$$

となる（命題 1.3 (2)）．これが任意の $\epsilon > 0$ に対して成り立つから，

$$P\left(\lim_{m\to\infty}\sup_{m \le a \le b} |X_a + \cdots + X_b| > 0\right) = 0$$

である．すなわち，$(\sum_{i=1}^{n} X_i)_{n\in\mathbb{N}_{>0}}$ は，ほとんど確実に Cauchy 列である．よって，主張の概収束が成り立つ． □

定理 3.7 $(X_i)_{i\in\mathbb{N}_{>0}}$ を確率空間 $(\Omega, \mathfrak{A}, P)$ 上の 2 乗可積分な確率変数の独立列とする．この列が

$$\sum_{i=1}^{\infty} \frac{\mathrm{Var}[X_i]}{i^2} < \infty$$

を満たすならば，$n \to \infty$ のとき

$$\frac{1}{n}\sum_{i=1}^{n} X_i - \frac{1}{n}\sum_{i=1}^{n} E[X_i] \xrightarrow{\text{a.s.}} 0$$

である．

証明 主張の仮定が成り立つとすると，確率変数列 $((X_i - E[X_i])/i)_{i\in\mathbb{N}_{>0}}$ について，

$$E\left[\frac{X_i - E[X_i]}{i}\right] = 0 \quad (i \in \mathbb{N}_{>0}), \qquad \sum_{i=1}^{\infty} \mathrm{Var}\left[\frac{X_i - E[X_i]}{i}\right] = \sum_{i=1}^{\infty} \frac{\mathrm{Var}[X_i]}{i^2} < \infty$$

である．このとき，補題 3.6 より，確率変数列 $(\sum_{i=1}^{n}(X_i - E[X_i])/i)_{n\in\mathbb{N}_{>0}}$ は概収束する．よって，Kronecker の補題（補題 3.3）より，主張の概収束が成り立つ． □

注意 3.8 Kronecker の補題（補題 3.3）より，定理 3.7 の仮定 $\sum_{i=1}^{\infty} \mathrm{Var}[X_i]/i^2 < \infty$ は，定理 3.1 の仮定 $\lim_{n\to\infty}(1/n^2)\sum_{i=1}^{n} \mathrm{Var}[X_i] = 0$ よりも強い．

定理 3.9（大数の強法則） $(X_i)_{i\in\mathbb{N}_{>0}}$ を確率空間 $(\Omega, \mathfrak{A}, P)$ 上の 2 乗可積分な実確率変数の独立同分布列とし，これらの共通の期待値を μ と置く．すると，$n \to \infty$ のとき

$$\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{\text{a.s.}} 0$$

である．

証明 各 $i \in \mathbb{N}_{>0}$ に対して，確率変数 Y_i を，

$$Y_i = \begin{cases} X_i & (|X_i| \le i) \\ 0 & (\text{otherwise}) \end{cases}$$

と定める．各 Y_i は有界だから，特に 2 乗可積分である．また，X_i の共通の分布を ν と書くと，

$$
\begin{aligned}
\sum_{i=1}^{\infty} \frac{\mathrm{Var}[Y_i - E[Y_i]]}{i^2} &\leq \sum_{i=1}^{\infty} \frac{E[Y_i^2]}{i^2} \\
&\leq 4 \sum_{i=1}^{\infty} \frac{E[Y_i^2]}{(i+1)^2} \\
&= 4 \sum_{i=1}^{\infty} \frac{1}{(i+1)^2} \int_{|x| \leq i} x^2 \, d\nu(x) \\
&\leq 4 \int_0^{\infty} \frac{1}{y^2} \int_{|x| \leq i} x^2 \, d\nu(x) \, dy \\
&= 4 \int_{\mathbb{R}} \left(\int_{|x|}^{\infty} \frac{1}{y^2} \, dy \right) x^2 \, d\nu(x) \\
&= 4 \int_{\mathbb{R}} |x| \, d\nu(x) \\
&< \infty
\end{aligned}
$$

である．したがって，定理 3.7 より，$n \to \infty$ のとき

$$
\frac{1}{n} \sum_{i=1}^{n} (Y_i - E[Y_i]) \xrightarrow{\text{a.s.}} 0 \tag{$*$}
$$

である．

確率変数の列 $(X_i)_{i \in \mathbb{N}_{>0}}$ が同分布で可積分であることと補題 3.4 より

$$
\sum_{i=1}^{\infty} P(X_i \neq Y_i) = \sum_{i=1}^{\infty} P(|X_i| > i) = \sum_{i=1}^{\infty} P(|X_1| > i) < \infty
$$

だから，Borel–Cantelli の補題（命題 1.10 (1)）より，無限個の i に対して $X_i \neq Y_i$ となる確率は 0 である．したがって，$n \to \infty$ のとき

$$
\frac{1}{n} \sum_{i=1}^{n} (X_i - Y_i) \xrightarrow{\text{a.s.}} 0 \tag{$**$}
$$

である．また，$(X_i)_{i \in \mathbb{N}_{>0}}$ が同分布で可積分であることと Lebesgue の収束定理より

$$
\lim_{i \to \infty} E[Y_i] = \lim_{i \to \infty} E[\chi_{\{|X_1| \leq i\}} X_1] = E[X_1] = \mu
$$

だから，

$$
\lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} E[Y_i] = \mu \tag{$***$}
$$

である（Cesàro 平均の性質）．以上，($*$), ($**$), ($***$) より，主張の概収束が従う． □

3.4 大数の強法則の逆

次の命題は，ある意味で，大数の法則の逆を主張している．

命題 3.10　$(X_i)_{i\in\mathbb{N}_{>0}}$ を確率空間 $(\Omega,\mathfrak{A},P)$ 上の実確率変数の独立同分布列とし，これらの確率変数は可積分でないとする．このとき，ほとんど確実に，

$$\limsup_{n\to\infty}\frac{1}{n}\left|\sum_{i=1}^{n}X_i\right|=\infty$$

が成り立つ．

証明　$C\geq 0$ とすると，$(X_i)_{i\in\mathbb{N}_{>0}}$ が同分布であり可積分でないことと補題 3.4 より，

$$\sum_{i=1}^{\infty}P(|X_i|\geq iC)=\sum_{i=1}^{\infty}P(|X_1|\geq iC)=\infty$$

である．したがって，Borel–Cantelli の補題（命題 1.10 (2)）より，ほとんど確実に無限個の i に対して $|X_i|\geq iC$ である．特に，ほとんど確実に $\limsup_{n\to\infty}|X_n|/n\geq C$ である．これが任意の $C\geq 0$ に対して成り立つから，ほとんど確実に

$$\limsup_{n\to\infty}\frac{|X_n|}{n}=\infty \tag{$*$}$$

である．一方で，

$$\begin{aligned}\frac{|X_n|}{n}&=\frac{1}{n}\left|\sum_{i=1}^{n}X_i-\frac{n-1}{n}\frac{1}{n-1}\sum_{i=1}^{n-1}X_i\right|\\&\leq\frac{1}{n}\left|\sum_{i=1}^{n}X_i\right|+\frac{n-1}{n}\frac{1}{n-1}\left|\sum_{i=1}^{n-1}X_i\right|\end{aligned}$$

だから，上極限をとって

$$\limsup_{n\to\infty}\frac{|X_n|}{n}\leq 2\limsup_{n\to\infty}\frac{1}{n}\left|\sum_{i=1}^{n}X_i\right| \tag{$**$}$$

を得る．$(*)$ と $(**)$ より，主張が成り立つ．　□

4　中心極限定理

4.1　正規分布

V を有限次元実線型空間とする．対称 2 階テンソル $\Sigma\in V\otimes V$ が**正値**であるとは，任意の $t\in V^*$ に対して，$\langle t\otimes t,\Sigma\rangle\geq 0$ であることをいう．これに加えて，$\langle t\otimes t,\Sigma\rangle=0$ となるのが $t=0$ のときのみであるならば，Σ は**非退化**であるという．

命題 4.1　V を有限次元実線型空間とする．任意の $\mu\in V$ と正値対称 2 階テンソル $\Sigma\in V\otimes V$ に対して，関数 $\gamma_{\mu,\Sigma}\colon V^*\to\mathbb{C}$ を

$$\gamma_{\mu,\Sigma}(t)=\exp(i\langle t,\mu\rangle-\langle t\otimes t,\Sigma\rangle)\qquad(t\in V^*)$$

と定めると，$\gamma_{\mu,\Sigma}$ を特性関数にもつ V 上の確率測度 $\nu_{\mu,\Sigma}$ が一意に存在する．さらに，この $\nu_{\mu,\Sigma}$ は，任意の $p\in[1,\infty)$ に対して p 乗可積分であり，その期待値は μ，共分散テンソルは Σ である．

証明 <u>一意性</u>　特性関数が確率測度を特徴付ける（定理 1.31）ことの結果である．

<u>存在</u>　V 上の確率測度 ν が $\gamma_{0,\Sigma}$ を特性関数にもつとすると，その写像 $v \mapsto v+\mu$ による像は，$\gamma_{\mu,\Sigma}$ を特性関数にもつ（命題 1.23）．したがって，主張は，$\mu=0$ の場合に示せば十分である．さらに，$\Sigma \in V\otimes V$ は対称 2 階テンソルだから，V の基底 $(e_1,\dots,e_d)$ と $k\in\{0,\dots,d\}$ が存在して，$\Sigma=\sum_{i=1}^k e_i\otimes e_i$ と書ける．このとき，$(e_1,\dots,e_d)$ の双対基底が定める同型を通して $\gamma_{0,\Sigma}$ を $\mathbb{R}^d$ 上の関数とみなしたものは，

$$\gamma_k^{(d)}(t_1,\dots,t_d)=\exp\left(-\sum_{i=1}^k t_i^2\right)=\prod_{i=1}^k e^{-t_i^2} \qquad ((t_1,\dots,t_d)\in\mathbb{R}^d)$$

で与えられる関数 $\gamma_k^{(d)}$ となる．以上より，結局，任意の $k\in\{0,\dots,d\}$ に対して，$\gamma_k^{(d)}$ を特性関数にもつ $\mathbb{R}^d$ 上の確率測度が存在することを示せばよい．

Lebesgue 測度に関して密度関数 $x\mapsto(2\pi)^{1/2}e^{-x^2/2}$ をもつ $\mathbb{R}$ 上の確率測度を ν と書くと，ν の特性関数は $t\mapsto e^{-t^2/2}$ である（補題 1.29）．また，点 $0\in\mathbb{R}$ に集中した $\mathbb{R}$ 上の Dirac 測度 δ_0 の特性関数は，定数関数 1 である．よって，$\mathbb{R}^d$ 上の積測度 $\nu^{\otimes k}\otimes\delta_0^{\otimes(d-k)}$ の特性関数は，上記の $\gamma_k^{(d)}$ となる（命題 1.28）．これで，主張が示された．

<u>最後の主張</u>　前段の記号で，$\mathbb{R}^d$ 上の確率測度 $\nu^{\otimes k}\otimes\delta_0^{\otimes(d-k)}$ は，任意の $p\in[1,\infty)$ に対して p 乗可積分である．よって，確率測度 $\nu_{\mu,\Sigma}$ についても同様である．また，

$$D_t\gamma_{\nu,\Sigma}(0)=i\langle t,\mu\rangle \quad (t\in V^*), \qquad D_{s,t}^2\gamma_{\nu,\Sigma}(0)=-\langle s\otimes t,\Sigma\rangle \quad (s,\,t\in V^*)$$

だから，$\nu_{\mu,\Sigma}$ の期待値は μ，共分散テンソルは Σ である（系 1.25）．　□

定義 4.2（正規分布）　V を有限次元実線型空間とする．$\mu\in V$ と正値対称 2 階テンソル $\Sigma\in V\otimes V$ に対して，命題 4.1 で定まる V 上の確率測度 $\nu_{\mu,\Sigma}$ を，期待値 μ，共分散テンソル Σ の**正規分布**（normal distribution）といい，記号 $N(\mu,\Sigma)$ で表す[*10]．$\mu=0$ のとき，$N(0,\Sigma)$ を，共分散テンソル Σ の**中心正規分布**（centered normal distribution）という．Σ が非退化であるとき，正規分布 $N(\mu,\Sigma)$ は**非退化**（non-degenerate）であるという．

$V=\mathbb{R}^d$ の場合は，$\mu\in\mathbb{R}^d$ と d 次実正値対称行列 Σ に対して，期待値 μ，共分散行列 Σ の正規分布 $N(\mu,\Sigma)$ が定まる．期待値 0，共分散行列 I_d（d 次単位行列）の正規分布 $N(0,I_d)$ を，**d 次元標準正規分布**（d-dimensional standard normal distribution）という．

注意 4.3　命題 4.1 の証明からわかるように，d 次元標準正規分布 $N(0,I_d)$ は，Lebesgue 測度に関して密度関数

$$f_{0,I_d}(x)=\frac{1}{(2\pi)^{d/2}}\exp\left(-\frac{1}{2}|x|^2\right) \qquad (x\in\mathbb{R}^d)$$

をもつ確率測度である（$|{-}|$ は Euclid ノルムを表す）．$\mu\in\mathbb{R}^d$ とし，Σ を d 次実正値対称行列とすると，$N(0,I_d)$ のアフィン写像 $x\mapsto\Sigma^{1/2}x+\mu$ による像は，正規分布 $N(\mu,\Sigma)$ である（命題 1.23 を用いて確かめられる）．以上のことと変数変換公式より，非退化な正規分布 $N(\mu,\Sigma)$ は，Lebesgue 測度に関して密度関数

$$f_{\mu,\Sigma}(x)=\frac{1}{(2\pi)^{d/2}(\det\Sigma)^{1/2}}\exp\left(-\frac{1}{2}(x-\mu)^{\mathrm{T}}\Sigma^{-1}(x-\mu)\right) \qquad (x\in\mathbb{R}^d)$$

をもつ確率測度である．

[*10] この記号は，主に確率変数の分布を扱う文脈で，たとえば「X は正規分布 $N(\mu,\Sigma)$ に従う」というように使われる．一方で，この正規分布に関する Borel 集合 A の関する測度を $N(\mu,\Sigma)(A)$ と書くことは，あまりない．

4.2 中心極限定理

補題 4.4 複素数列 $(z_n)_{n\in\mathbb{N}_{>0}}$ が $z\in\mathbb{C}$ に収束するならば，

$$\lim_{n\to\infty}\left(1+\frac{z_n}{n}\right)^n = e^z$$

である．

証明 各 $n\in\mathbb{N}_{>0}$ に対して，

$$\left(1+\frac{z_n}{n}\right)^n = \sum_{k=0}^{n}\binom{n}{k}\frac{z_n^k}{n^k} = \sum_{k=0}^{\infty}\frac{n}{n}\frac{n-1}{n}\cdots\frac{n-k+1}{n}\cdot\frac{z_n^k}{k!}$$

である．$C = \sup_{n\in\mathbb{N}_{>0}}|z_n|$ と置くと，上式の最右辺の各項の絶対値は $C^k/k!$ で上から抑えられ，$\sum_{k=0}^{\infty}C^k/k! = e^C < \infty$ である．よって，Lebesgue の収束定理より，上式は $n\to\infty$ のとき $\sum_{k=0}^{\infty}z^k/k! = e^z$ に収束する． $\square$

定理 4.5（中心極限定理） $(\Omega,\mathfrak{A},P)$ を確率空間，V を有限次元実線型空間とし，$(X_i)_{i\in\mathbb{N}_{>0}}$ を $(\Omega,\mathfrak{A},P)$ 上の 2 乗可積分な V 値確率変数の独立同分布列とする．X_i の共通の期待値は 0 であるとし，共通の共分散テンソルを $\varSigma\in V\otimes V$ と書く．すると，$n\to\infty$ のとき

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}X_i \xrightarrow{\mathrm{d}} N(0,\varSigma)$$

である．

証明 X_i の共通の分布を μ とする．特性関数 ϕ_μ は 2 階連続微分可能であり，

$$\phi_\mu(0)=1,\qquad D_t\phi_\mu(0)=0\quad(t\in V^*),\qquad D^2_{s,t}\phi_\mu(0) = -\langle s\otimes t,\varSigma\rangle\quad(s,\,t\in V^*)$$

だから（系 1.25），Taylor の定理より，

$$\phi_\mu(t) = 1-\frac{1}{2}\langle\varSigma,t\otimes t\rangle + R(t)\quad(t\in V^*),\qquad \lim_{t\to 0}\frac{R(t)}{|t|^2}=0$$

と書ける．ここで，$|{-}|$ は任意に固定した V^* 上のノルムである．これを用いると，$T_n = (1/\sqrt{n})\sum_{i=1}^{n}X_i$ の特性関数は，

$$\phi_{T_n}(t) = \phi_\mu\left(\frac{1}{\sqrt{n}}t\right)^n = \left(1-\frac{1}{2n}\langle t\otimes t,\varSigma\rangle + R\left(\frac{1}{\sqrt{n}}t\right)\right)^n\qquad(t\in V^*)$$

と表せる（系 1.27）．補題 4.4 より，上式は，$n\to\infty$ のとき $\exp(-(1/2)\langle t\otimes t,\varSigma\rangle) = \gamma_{0,\varSigma}(t)$ に収束する（$\gamma_{0,\varSigma}$ は，命題 4.1 で定義したものである）．よって，Lévy の連続性定理の系（系 2.26）より，T_n の分布は $N(0,\varSigma)$ に弱収束する． $\square$

参考文献

[1] P. Billingsley, *Convergence of Probability Measures*, 2nd edition, Wiley, 1999.

[2] D. L. Cohn, *Measure Theory*, 2nd edition, Springer, 2013.

[3] 鈴木大慈,「確率数理工学補足資料 大数の法則と中心極限定理」, 2017.
https://ibis.t.u-tokyo.ac.jp/suzuki/lecture/2017/probth/Supplementary1.pdf