# 確率統計学 I

杉浦　誠

2020 年 11 月 24 日

## 目次

これは

- 浅野, 江島, 李 共著 基本統計学 森北出版
  を教科書として作った講義ノートです。多くの部分で、
- 舟木 直久 著 確率論 朝倉書店
- Durrett, R.: Probablity Theory and Examples, 4th ed., (2010), Cambridge Series in Statistical and Probabilistic Mathematics.
  を参考にしています。(むしろ後者の 2 冊ほうが教科書に近い存在です。)

を教科書として作った講義ノートです。

[10 月 7 日]

この授業では確率論を測度論に基づいて述べていく。シラバスに書いたように、確率統計学 I および解析学 I, II, IV, 関数解析学 I, II を履修済みあるいは同時履修しているものとし、複素関数論やフーリエ解析、ルベーグ積分論の詳細な説明はそちらの授業で理解できているものとして述べていく。

参考書として「舟木 直久 著 確率論 朝倉書店」をあげておく。以下 [F] として引用します。

# 1 確率空間と確率変数

## 1.1 確率測度と事象

**定義 1.1** 集合 $\Omega(\neq \emptyset)$ の部分集合からなる集合を $\Omega$ 上の集合族という。$\Omega$ 上の集合族 $\mathcal{F}$ が次の条件

(i) $\Omega \in \mathcal{F}$

(ii) $A \in \mathcal{F} \Longrightarrow A^{\mathrm{c}} \in \mathcal{F}$ （$A^{\mathrm{c}}$ は $A$ の補集合を表す。）

(iii) $A_1, A_2, \dots, A_n, \dots \in \mathcal{F} \Longrightarrow \bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$

を満たすとき、この $\mathcal{F}$ を $\Omega$ 上の $\sigma$ 集合族という。また、$(\Omega, \mathcal{F})$ を可測空間という。

確率論では、$\sigma$ 集合族 $\mathcal{F}$ の元を事象という。特に、$\Omega$ を全事象と、$\emptyset = \Omega^{\mathrm{c}}$ を空事象という。

**定義 1.2** $(\Omega, \mathcal{F})$ を可測空間とする。次の条件を満たす $P : \mathcal{F} \to [0,1]$ を $(\Omega, \mathcal{F})$ 上の確率測度という。

(i) $P(\Omega) = 1$

(ii) $A_n \in \mathcal{F}$, $n = 1, 2, \dots,$ が任意の $m \neq n$ なる組に対し $A_m \cap A_n = \emptyset$ を満たせば (このとき $A_1, A_2, \dots$ は互いに排反であるという)、$P\Big(\bigcup\limits_{n=1}^{\infty} A_n\Big) = \sum\limits_{n=1}^{\infty} P(A_n)$.

このとき、$(\Omega, \mathcal{F}, P)$ を確率空間という。また、$\mathcal{F}$ の元を事象、$A \in \mathcal{F}$ に対し $P(A)$ を事象 $A$ の確率という。

以下の定理 1.1, 定理 1.2, 定理 1.3 の証明については、以下の url にある昨年度の前期の講義ノートを参照していただくとして、証明は省略する。

`http://www.math.u-ryukyu.ac.jp/~sugiura/2019/prob2019a.pdf`

**定理 1.1** $(\Omega, \mathcal{F})$ を可測空間、$\{B_n\} \subset \mathcal{F}$ とすると次が成り立つ。

(0) $\emptyset \in \mathcal{F}$, (1) $\displaystyle\bigcup_{i=1}^{n} B_i \in \mathcal{F}$, (2) $\displaystyle\bigcap_{i=1}^{n} B_i \in \mathcal{F}$, (3) $\displaystyle\bigcap_{n=1}^{\infty} B_n \in \mathcal{F}$, (4) $B_1 \backslash B_2 := B_1 \cap B_2^{\mathrm{c}} \in \mathcal{F}$

**定理 1.2** $(\Omega, \mathcal{F}, P)$ を確率空間とする。

(1) $P(\emptyset) = 0$

(2) $B_i \in \mathcal{F}$, $i = 1, \dots, n,$ が互いに排反であれば、$\displaystyle P\Big(\bigcup_{i=1}^{n} B_i\Big) = \sum_{i=1}^{n} P(B_i)$.

(3) $A, B \in \mathcal{F}, A \subset B \Longrightarrow P(A) \leq P(B)$

(4) $B \in \mathcal{F} \Longrightarrow P(B^{\mathrm{c}}) = 1 - P(B)$

(5) $A, B \in \mathcal{F}$ に対し $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

**定理 1.3** $A_1, A_2, \dots \in \mathcal{F}$ に対し $P(\bigcup_{n=1}^{\infty} A_n) \leq \sum_{n=1}^{\infty} P(A_n)$.

次の定理 1.4 (2) は一般の測度空間では成立しない ($cf.$ 問題 1.1)。

**定理 1.4** $(\Omega, \mathcal{F}, P)$ を確率空間、$\{B_n\} \subset \mathcal{F}$ とする。

(1)　$B_1 \subset B_2 \subset \cdots \subset B_n \subset \cdots \implies P(\bigcup_{n=1}^\infty B_n) = \lim_{n\to\infty} P(B_n)$.
(2)　$B_1 \supset B_2 \supset \cdots \supset B_n \supset \cdots \implies P(\bigcap_{n=1}^\infty B_n) = \lim_{n\to\infty} P(B_n)$.

**証明:** (1) $A_1 = B_1$, $A_n = B_n \backslash \bigcup_{k=1}^{n-1} B_k$ $(n \geq 2)$ とおくと、$A_1, A_2, \ldots$ は互いに排反で $\bigcup_{k=1}^n A_k = B_n$ $(n \in \mathbf{N})$, $\bigcup_{k=1}^\infty A_k = \bigcup_{n=1}^\infty B_n$. よって、定義 1.2 (ii), 定理 1.2 (2) により、

$$P(\bigcup_{n=1}^\infty B_n) = P(\bigcup_{k=1}^\infty A_k) = \sum_{k=1}^\infty P(A_k) = \lim_{n\to\infty} \sum_{k=1}^n P(A_k) = \lim_{n\to\infty} P(\bigcup_{k=1}^n A_k) = \lim_{n\to\infty} P(B_n).$$

(2) $B_1^{\rm c} \subset \cdots \subset B_n^{\rm c} \subset \cdots$ であるから、定理 1.2 (4), de Morgan の法則と前半より

$$P(\bigcap_{n=1}^\infty B_n) = 1-P((\bigcap_{n=1}^\infty B_n)^{\rm c}) = 1-P(\bigcup_{n=1}^\infty B_n^{\rm c}) = 1-\lim_{n\to\infty} P(B_n^{\rm c}) = \lim_{n\to\infty}(1-P(B_n^{\rm c})) = \lim_{n\to\infty} P(B_n). \qquad \square$$

**問題 1.1** $\mu$ を $\mathbf{R}$ 上の Lebesgue 測度とする。このとき、$\mathbf{R}$ の区間の列 $\{A_n\}$ で $\bigcap_{n=1}^\infty A_n = \emptyset$ であるが、$\lim_{n\to\infty} \mu(A_n) \neq 0$ となるものを例示し、その性質を持つことを示せ。

次の定理は Borel 集合族および Borel 関数を導入するのに用いる。

**定理 1.5** $\Omega$ を集合とし $\mathcal{A}$ をその部分集合族とする。$\mathcal{A}$ を含む $\Omega$ 上の $\sigma$ 集合族はいくつもある。これに index を付け、$\mathcal{F}_\lambda$ $(\lambda \in \Lambda)$ と表す。このとき、次が成り立つ。
(1)　$\bigcap_{\lambda\in\Lambda} \mathcal{F}_\lambda$ は $\sigma$ 集合族である。
(2)　$\bigcap_{\lambda\in\Lambda} \mathcal{F}_\lambda$ は $\mathcal{A}$ を含む $\sigma$ 集合族の中で最小の $\sigma$ 集合族である。

**証明:** (1) は問題 1.2 とする。(2) まず、$\mathcal{A}$ を含む $\Omega$ 上の $\sigma$ 集合族として、$\Omega$ のべき集合があることに注意する (よって $\Lambda$ は空ではない)。(1) より、$\mathcal{F} := \bigcap_{\lambda\in\Lambda} \mathcal{F}_\lambda$ が $\mathcal{A}$ を含む $\sigma$ 集合族の中で最小であることを示せばよい。今、$\mathcal{F}$ より小さい (真部分集合となる)、$\mathcal{A}$ を含む $\sigma$ 集合族 $\mathcal{F}_0$ があったとする。$\mathcal{F}_0$ は $\mathcal{A}$ を含む $\sigma$ 集合族なので、$\mathcal{F}_\lambda$ のいずれかである。従って、$\mathcal{F}_0 \supset \bigcap_{\lambda\in\Lambda} \mathcal{F}_\lambda = \mathcal{F}$ であるが、これは $\mathcal{F}_0$ が $\mathcal{F}$ の真部分集合であることに矛盾する。　$\square$

**問題 1.2** 定理 1.5 (1) を証明せよ。

確率論では通常の測度論に加えて、次の事象の独立性を考察する。

**定義 1.3**　事象の列 $\{B_n\} \subset \mathcal{F}$ (有限個でも無限個でもよい) に対し、

$$P(B_{n_1} \cap B_{n_2} \cap \cdots \cap B_{n_l}) = P(B_{n_1})P(B_{n_2})\cdots P(B_{n_l}) \qquad \text{for all } n_1 < n_2 < \cdots < n_l \tag{1.1}$$

を満たすとき、事象の列 $\{B_n\}$ は独立であるという。

**問題 1.3** (1)　$\Omega = \{1,2,3,4\}$, $P(\{k\}) = 1/4$ $(k = 1,2,3,4)$ において、$A = \{1,2\}$, $B = \{1,3\}$, $C = \{2,3\}$ とおくと、事象 $A, B, C$ に対しどの 2 つの事象の組も独立であるが、$A, B, C$ が独立とならないことを示せ。
(2)　$\Omega = \{1,2,3,4\}$, $P(\{1\}) = \dfrac{1}{\sqrt{2}} - \dfrac{1}{4}$, $P(\{2\}) = \dfrac{3}{4} - \dfrac{1}{\sqrt{2}}$, $P(\{3\}) = P(\{4\}) = 1/4$ において、$A = \{1,2\}$, $B = \{2,3\}$, $C = \{2,4\}$ とおくと、$P(A\cap B\cap C) = P(A)P(B)P(C)$ となるが、事象 $A, B, C$ に対しどの 2 つの事象の組も独立とならないことを示せ。

**定理 1.6** 事象の列 $B_1, B_2, \ldots, B_n$ が独立であれば、任意の $k = 1, \ldots, n$ に対して、$B_1^{\rm c}, \ldots, B_k^{\rm c}, B_{k+1}, \ldots, B_n$ も独立となる。

**証明:** $B_1, B_2, \ldots, B_n$ が独立ならば、$B_1^c, B_2, \ldots, B_n$ も独立となることを示せばよい。実際、これより $B_2, B_3, \ldots, B_n, B_1^c$ が独立となるから、$B_2^c, B_3, \ldots, B_n, B_1^c$ は独立となり、これを繰り返すことで主張を得る。

このため任意の $1 \leq n_1 < n_2 < \cdots < n_l \leq n$ に対し (1.1) に相当する式を示せばよが、$n_1 \geq 2$ の場合は $B_1^c$ が関与しないため (1.1) は明らかに成り立つ。$n_1 = 1$ の場合:

$$P(B_1^c \cap B_{n_2} \cap \cdots \cap B_{n_l}) = P(B_1^c)P(B_{n_2}) \cdots P(B_{n_l}) \tag{1.2}$$

を示せばよい。まず、

$$P(B_1 \cap B_{n_2} \cap \cdots \cap B_{n_l}) + P(B_1^c \cap B_{n_2} \cap \cdots \cap B_{n_l}) = P(B_{n_2} \cap \cdots \cap B_{n_l})$$

に注意する。ここで、$B_1, B_2, \cdots, B_n$ の独立性を用いると

$$P(B_1 \cap B_{n_2} \cap \cdots \cap B_{n_l}) = P(B_1)P(B_{n_2}) \cdots P(B_{n_l}), \quad P(B_{n_2} \cap \cdots \cap B_{n_l}) = P(B_{n_2}) \cdots P(B_{n_l})$$

となる。よって、

$$P(B_1^c \cap B_{n_2} \cap \cdots \cap B_{n_l}) = \{1 - P(B_1)\}P(B_{n_2}) \cdots P(B_{n_l}) = P(B_1^c)P(B_{n_2}) \cdots P(B_{n_l}). \qquad \square$$

## 1.2 確率変数

**定義 1.4** 可測空間 $(\Omega, \mathcal{F})$ において、$\Omega$ 上の実数値関数 $X : \Omega \to \mathbf{R}$ が $\forall a \in \mathbf{R}$ に対して

$$\{\omega; X(\omega) \leq a\} \in \mathcal{F}$$

を満たすとき $X$ を $(\Omega, \mathcal{F})$ 上の確率変数という。

**例 1.7** 1 回のコイン投げでは、$\Omega = \{H, T\}$, $\mathcal{F} = 2^\Omega = \{\emptyset, \{H\}, \{T\}, \{H, T\}\}$ であった。このとき、

$$X(\omega) = \begin{cases} 1 & (\omega = H) \\ 0 & (\omega = T) \end{cases} \text{ とすると、} \{\omega; X(\omega) \leq a\} = \begin{cases} \Omega & (1 \leq a) \\ \{T\} & (0 \leq a < 1) \\ \emptyset & (a < 0) \end{cases} \text{ となるので、}$$

$\forall a \in \mathbf{R}$ に対して $\{\omega; X(\omega) \leq a\} \in \mathcal{F}$ がわかる。従って、$X$ は確率変数である。

**定理 1.8** $X$ は $(\Omega, \mathcal{F})$ 上の確率変数 $\iff$ $\forall a \in \mathbf{R}$ に対して $\{\omega; X(\omega) < a\} \in \mathcal{F}$.

**証明:** $(\Longrightarrow)$ $\{\omega; X(\omega) < a\} = \displaystyle\bigcup_{n=1}^{\infty} \{\omega; X(\omega) \leq a - \frac{1}{n}\}$ より従う。

$(\Longleftarrow)$ $\{\omega; X(\omega) \leq a\} = \displaystyle\bigcap_{n=1}^{\infty} \{\omega; X(\omega) < a + \frac{1}{n}\}$ と定理 1.2 (2) より従う。 $\square$

ここで、$\sigma$ 集合族の定義から

$$\{\omega; X(\omega) \leq a\} \in \mathcal{F} \iff \{\omega; X(\omega) > a\} \in \mathcal{F},$$
$$\{\omega; X(\omega) < a\} \in \mathcal{F} \iff \{\omega; X(\omega) \geq a\} \in \mathcal{F}.$$

これと定理 1.8 より、一般に次が成立する。

**定理 1.9** $X$ は $(\Omega, \mathcal{F})$ 上の確率変数 $\iff$ 任意の区間 $I$ に対して $\{\omega; X(\omega) \in I\} \in \mathcal{F}$.

**証明:** $(\Longleftarrow)$ は明らか。$(\Longrightarrow)$ について、$I = (-\infty, b), (-\infty, b], (a, \infty), [a, \infty)$ の場合はすでに示した。$I = (a, b), [a, b], (a, b], [a, b)$ の場合に $\{\omega; X(\omega) \in I\} \in \mathcal{F}$ を示せばよい。例えば、$I = (a, b)$ のときは、

$$\{\omega; X(\omega) \in (a, b)\} = \{\omega; X(\omega) < b\} \cap \{\omega; X(\omega) > a\}$$

より従う。他は演習問題とする。 $\square$

**定義 1.5** 定理 1.5 より、$\mathbf{R}$ 上の区間 $(a,b]$ 全体からなる族を含む最小の $\sigma$ 集合族が存在する。この集合族を $\mathbf{R}$ の Borel 集合族といい、$\mathcal{B}(\mathbf{R})$ と表す。Borel 集合族に属する集合を Borel 集合という。

**補題 1.10** 高々可算個の元からなる集合 $E$ は Borel 集合である。また、開集合 $G$ も Borel 集合である。

**証明:** 一点からなる集合 $\{x\}$ は $\{x\} = \bigcap_{n=1}^{\infty}\left(x - \dfrac{1}{n}, x\right]$ となるから $\{x\} \in \mathcal{B}(\mathbf{R})$. よって、$E$ は一点からなる集合の高々可算和であるから $E \in \mathcal{B}(\mathbf{R})$. 次に開集合 $G$ の各点 $x \in G$ に対して、$a_x, b_x \in \mathbf{Q}$ があって、$x \in (a_x, b_x] \subset G$ とできる。このとき、$\bigcup_{x\in G}(a_x, b_x] = G$ であるが、有理数体 $\mathbf{Q}$ は可算集合なので、この左辺は可算和となる。したがって、定義 1.1 (iii) により $G$ は Borel 集合となる。 $\square$

**補題 1.11** 可測空間 $(\Omega, \mathcal{F})$ において、関数 $X : \Omega \to \mathbf{R}$ によって与えられる $\mathbf{R}$ の部分集合族 $\{A \subset \mathbf{R}; X^{-1}(A) \in \mathcal{F}\}$ は $\sigma$ 集合族である。ただし、$X^{-1}(A) = \{\omega \in \Omega; X(\omega) \in A\}$ とする。

**問題 1.4** $X^{-1}(\mathbf{R}) = \Omega$, $X^{-1}(A^{\mathrm{c}}) = (X^{-1}(A))^{\mathrm{c}}$, $X^{-1}(\bigcup_{n=1}^{\infty} A_n) = \bigcup_{n=1}^{\infty} X^{-1}(A_n)$ となることを示し、補題 1.11 を証明せよ。

**定理 1.12** $X$ が $(\Omega, \mathcal{F})$ 上の確率変数 $\iff$ 任意の Borel 集合 $B \in \mathcal{B}(\mathbf{R})$ に対して $\{\omega; X(\omega) \in B\} \in \mathcal{F}$.

**証明:** $(\Longleftarrow)$ $B = (-\infty, a]$ ととれば明らか。
$(\Longrightarrow)$ $X$ を確率変数とし、$\mathcal{A} = \{A \subset \mathbf{R}; X^{-1}(A) \in \mathcal{F}\}$ とおく。定理 1.9 より $(a,b] \in \mathcal{A}$ であり、補題 1.11 より $\sigma$ 集合族である。従って、$\mathcal{B}(\mathbf{R})$ は区間 $(a,b]$ 全体からなる集合族を含む最小の $\sigma$ 集合族であるから $\mathcal{A} \supset \mathcal{B}(\mathbf{R})$ となる。よって $B \in \mathcal{B}(\mathbf{R})$ なら $B \in \mathcal{A}$ となり、$X^{-1}(B) \in \mathcal{F}$ を得る。 $\square$

**定義 1.6** 可測空間 $(\mathbf{R}, \mathcal{B}(\mathbf{R}))$ において、関数 $f : \mathbf{R} \to \mathbf{R}$ が任意の $a \in \mathbf{R}$ に対して $\{x; f(x) \le a\} \in \mathcal{B}(\mathbf{R})$ を満たすとき、$f$ は $(\mathbf{R}, \mathcal{B}(\mathbf{R}))$ 上の可測関数、あるいは Borel 可測関数であるという。

**注意 1.1** $f$ が $(\mathbf{R}, \mathcal{B}(\mathbf{R}))$ 上で可測であることは、$f$ が $(\mathbf{R}, \mathcal{B}(\mathbf{R}))$ 上の確率変数であることは同値である。

**定理 1.13** 高々可算個の点を除いて連続な関数 $f : \mathbf{R} \to \mathbf{R}$ は Borel 可測である。特に、連続関数は Borel 可測関数である。

**証明:** $D_f$ は高々可算個の点からなる集合で、$f$ は $\mathbf{R}\backslash D_f$ で連続とする。$\forall a \in \mathbf{R}$ に対して $\{x; f(x) > a\}$ が Borel 集合であることを示せばよい。$C_{a,1} := \{x \in D_f; f(x) > a\}$ は高々可算なのでこれは Borel 集合。また、$C_{a,2} := \{x \in \mathbf{R}\backslash D_f; f(x) > a\}$ は開集合 $G$ を用いて $C_{2,a} = G\backslash D_f$ と表せる。実際、$\forall x \in C_{a,2}$ に対して、$\delta'_x > 0$ を $y \in (x - \delta'_x, x + \delta'_x)\backslash D_f$ ならば $|f(y) - f(x)| < f(x) - a$ となるように選べば、このとき $f(y) > a$ なので $(x - \delta'_x, x + \delta'_x)\backslash D_f \subset C_{a,2}$. よって、$G = \bigcup_{x\in C_{a,2}}(x - \delta'_x, x + \delta'_x)$ とおくと、$G$ は $\mathbf{R}$ の開集合で、$C_{a,2} = G\backslash D_f$ と表せるので補題 1.10 より $C_{a,2} \in \mathcal{B}(\mathbf{R})$. よって、$\{x; f(x) < a\} = C_{a,1} \cup C_{a,2} \in \mathcal{B}(\mathbf{R})$. $\square$

**定理 1.14** $X$ が $(\Omega, \mathcal{F})$ 上の確率変数、$f : \mathbf{R} \to \mathbf{R}$ が Borel 可測関数のとき、$f(X(\omega))$ は $(\Omega, \mathcal{F})$ 上の確率変数である。

**証明:** $\forall a \in \mathbf{R}$ とし、$B = \{x; f(x) \le a\}$ とおくと、$f$ は可測だから $B \in \mathcal{B}(\mathbf{R})$. よって、定理 1.12 により $\{\omega; f(X(\omega)) \le a\} = \{\omega; X(\omega) \in B\} \in \mathfrak{B}$ を得る。 $\square$

**例 1.15** 定理 1.13、定理 1.14 より、$X$ が確率変数であるとき $aX + b$ ($a, b$ は定数), $X^a$, $\sqrt{X}$, $\log X$, $e^X$, $\cos X$ などは、すべて確率変数である。

**Message:** 今回紹介した内容で、その証明が理解できなくても構いません。特に 1.2 節については定理 1.14 の帰結として、例 1.15 が得られることを承知おきください。ただし、問題 1.1–問題 1.4 はそれほど難しくないはずなので、必ず解いておいてください。

[10 月 14 日]

## 1.3 分布関数

**定義 1.7** (1) 確率空間 $(\Omega, \mathcal{F}, P)$ 上の確率変数 $X$ に対して、$F_X(x) = P(X \le x) = P(\{\omega; X(\omega) \le x\})$ で定義される関数 $F_X : \mathbf{R} \to [0,1]$ を確率変数 $X$ の分布関数という。
(2) 確率変数 $X, Y$ に対して、その分布関数が一致するとき、即ち $F_X(x) = F_Y(x)$ $(\forall x \in \mathbf{R})$ となるとき、$X$ と $Y$ は同じ分布に従うという。(「分布」の定義は定義 3.5 で行う。注意 3.2, 定理 3.13 も参照のこと。)

**定理 1.16** $P(a < X \le b) = F_X(b) - F_X(a)$, $a, b \in \mathbf{R}$ $(a < b)$

**証明:** $A = \{X \le a\}$, $B = \{a < X \le b\}$ とすると、$A \cap B = \emptyset$, $A \cup B = \{X \le b\}$ より、
$F_X(b) = P(A \cup B) = P(A) + P(B) = F_X(a) + P(a < X \le b)$. □

**定理 1.17** $F(x)$ を確率変数 $X$ の分布関数とする。このとき、次が成立する。

(1) $F$ は単調非減少である、即ち、$x < y$ ならば $F(x) \le F(y)$.
(2) 右連続関数である、即ち、$F(a+0) := \lim_{x \to a+0} F(x) = F(a)$ $(\forall a \in \mathbf{R})$.
(3) $F(-\infty) := \lim_{x \to -\infty} F(x) = 0$, $F(\infty) := \lim_{x \to \infty} F(x) = 1$.

**証明:** (1) $x < y$ より $\{X \le x\} \subset \{X \le y\}$. 従って、定理 1.2 (3) により $F(x) = P(X \le x) \le P(X \le y) = F(y)$ を得る。
(2) $\{a_n\}$ を単調減少で $a$ に収束する任意の数列とし、$C_n = \{X \le a_n\}$ とおく。このとき、$\bigcap_{n=1}^{\infty} C_n = \{X \le a\}$ となる。実際、「$\supset$」は $\forall n$ で $C_n \supset \{X \le a\}$ だから明らか。$\omega \in \{X > a\}$ とすると $X(\omega) > a$ で $a_n \to a + 0$ $(n \to \infty)$ より、ある $N$ があって $n \ge N \Longrightarrow a_n - a < X(\omega) - a$, これより $X(\omega) > a_N$, 即ち $\omega \notin C_N \supset \bigcap_{n=1}^{\infty} C_n$ となり「$\subset$」が従う。よって、$\{C_n\}$ は単調減少あったらから、定理 1.4 (2) より $\lim_{n \to \infty} F(a_n) = \lim_{n \to \infty} P(C_n) = P(X \le a) = F(a)$ となる。
(3) $\{x_n\}$ を単調減少で $-\infty$ に発散する任意の数列とする。このとき $B_n = \{X \le x_n\}$ とおくと、$\{B_n\}$ は単調減少な事象の列で、また $X$ は実数値なので、$\bigcap_{n=1}^{\infty} \{X \le x_n\} = \emptyset$. 従って、定理 1.4 (2) と命題 1.2 (1) により、$\lim_{n \to \infty} F(x_n) = \lim_{n \to \infty} P(B_n) = P(\emptyset) = 0$ となり、前半の主張を得る。後半は、問題 1.5 とする。 □

**問題 1.5** 任意の単調増加で $\infty$ に発散する数列 $\{x_n\}$ に対し、$\lim_{n \to \infty} F(x_n) = 1$ を示せ。

**注意 1.2** 関数 $F : \mathbf{R} \to \mathbf{R}$ が定理 1.17 の (1), (2), (3) を満たせば、$F(x)$ を分布関数としてもつ確率変数 $X$ が (無限個) 存在することが知られている。(*cf.* [F] p.47, p.83.)

**定義 1.8** 確率空間 $(\Omega, \mathcal{F}, P)$ 上の確率変数 $X$ に対して、$P(X \in E) = 1$ となる高々可算個の元からなる集合 $E \subset \mathbf{R}$ が存在するとき、$X$ を離散型確率変数という。(補題 1.10 より $E$ は Borel 集合となることに注意。)

離散型分布については前期に学んだ。連続型分布については以下の注意を与える。もし絶対連続型なら前期のように扱える。そうでない場合はあまり考えない。

**定義 1.9** $F_X(x)$ を、確率空間 $(\Omega, \mathcal{F}, P)$ 上の確率変数 $X$ の分布関数とする。$(\mathbf{R}, \mathcal{B}(\mathbf{R}))$ 上の非負可測関数 $f(x)$ が存在して

$$F_X(x) = \int_{-\infty}^{x} f(t)\, dt \qquad (\forall x \in \mathbf{R})$$

と表せるとき、$F$ を絶対連続な分布関数, $X$ を絶対連続型確率変数という。また、この $f(x)$ を $X$ の密度関数 (density function) という。

**注意 1.3** (1) $X$ が絶対連続型なら $F_X(x)$ は連続である。
(2) $F_X(x)$ が $x$ について連続であっても、密度関数 $f(x)$ が存在するとは限らない ($cf$. 例 1.18)。

**例 1.18 (Cantor 集合と Cantor 関数)** 参考図書: 伊藤清三 著 ルベーグ積分入門 裳華房 pp.41–45.

$[0,1]$ の閉部分集合の減少列 $K_n$, $n=1,2,\ldots,$ を次のように定め、$K=\bigcap_{n=1}^{\infty} K_n$ とおく。

$$K_1=\Big[0,\frac{1}{3}\Big]\cup\Big[\frac{2}{3},1\Big]=[0,1]\backslash\Big(\frac{1}{3},\frac{2}{3}\Big),\qquad K_n=K_{n-1}\backslash\bigcup_{k=0}^{3^{n-1}-1}\Big(\frac{k}{3^{n-1}}+\frac{1}{3^n},\frac{k}{3^{n-1}}+\frac{2}{3^n}\Big),\ n=2,3,\ldots.$$

ここで、各 $K_n$ はコンパクトなので $K\neq\emptyset$ であることに注意する。この $K$ を **Cantor 集合**という。

**定理 1.19** Cantor 集合 $K$ は連続濃度をもつ長さ 0 の集合となる。

**証明:** 各 $K_n$ は $2^n$ 個の長さ $1/3^n$ の区間の和集合なので、$K_n$ の長さは $2^n/3^n$ となる。特に、$K$ の長さは 0 となる。

$K$ が連続濃度をもつことを示すためには $K$ から $[0,1]$ への全射 $F$ を定義すればよい。

各 $n\in\mathbf{N}$ に対して、$F_n:\mathbf{R}\to[0,1]$ を $F_n(0)=0$ で $K_n$ 上でのみ傾き $(3/2)^n$ で増大し、それ以外では傾き 0 となる折れ線と定義する。例えば、

$$F_1(x)=\frac{3}{2}x,\ 0\le x\le\frac{1}{3},\quad F_1(x)=\frac{1}{2},\ \frac{1}{3}<x<\frac{2}{3},\quad F_1(x)=\frac{3}{2}\Big(x-\frac{2}{3}\Big)+\frac{1}{2},\ \frac{1}{3}\le x\le 1,$$

となる。このとき、構成法より $|F_{n+1}(x)-F_n(x)|\le 1/2^n$ となることに注意する。よって、$m>n$ とすると、

$$|F_m(x)-F_n(x)|\le\sum_{k=n}^{m-1}|F_{k+1}(x)-F_k(x)|\le\sum_{k=n}^{m-1}\frac{1}{2^k}=2\Big(\frac{1}{2^n}-\frac{1}{2^m}\Big)\to 0,\qquad m,n\to\infty$$

となり、$\{F_n(x)\}$ は一様収束の意味で Cauchy 列をなすので、連続関数 $F(x)$ が存在して一様収束する。

構成法より $x\notin K_n$ のとき $m\ge n$ なら $F_m(x)=F_n(x)=k/2^n$ となる $k=0,1,\ldots,2^n$ があるので、$F(x)$ は $K_n$ の補集合の各開区間では定数関数となる。すなわち、$F(x)$ は $K$ 上でのみ増加する連続関数なので、$F$ が $K$ から $[0,1]$ への全射となっていることが示された。 $\square$

**注意 1.4** 定理 1.19 の証明で定義した $F(x)$ を **Cantor 関数**という。$F(x)$ は $F(x)\equiv 0,\ x<0,\ F(x)\equiv 1,$ $x>1$ と拡張すると、$F(x)$ はある確率変数の分布関数となるが、密度関数は持たない。$K$ の外で定数なので、$F'(x)=0$ となる。すなわち $K$ の Lebesgue 測度が 0 であるから、$F'=0$ a.e. となるので、$\displaystyle\int_0^1 F'(t)\,dt=0,$ 一方 $F(1)-F(0)=1$ となるためである。

**問題 1.6** $[0,1]$ 上の連続関数の列 $\{F_n(x)\}$ が一様収束の意味で Cauchy 列をなせば、$[0,1]$ 上の連続関数 $F(x)$ が存在して $\{F_n(x)\}$ は $F(x)$ に一様収束することを、実数の完備性を用いてから証明せよ。(ヒント: 微積の教科書に書かれているはずです。)

次に多次元確率変数を考える。

**定義 1.10** $X_1(\omega),\ldots,X_n(\omega)$ を確率空間 $(\Omega,\mathcal{F},P)$ 上の確率変数とするとき、

$$\boldsymbol{X}(\omega)=(X_1(\omega),\ldots,X_n(\omega))\tag{1.3}$$

を $n$ 次元確率変数または $n$ 次元確率ベクトルという。ここに、$\boldsymbol{X}=(X_1,\ldots,X_n)$ とも略記される。

**定義 1.11** 任意の $(x_1,\ldots,x_n)\in\mathbf{R}^n$ に関し、(1.3) の多次元確率変数 $\boldsymbol{X}=(X_1,\ldots,X_n)$ の同時分布関数を次式で定義する。

$$F_{\boldsymbol{X}}(x_1,x_2,\ldots,x_n)=P(X_1\le x_1,X_2\le x_2,\cdots,X_n\le x_n).$$

**問題 1.7** 確率変数 $X, Y$ は連続型 ($P(X \le x, Y \le y)$ が $x, y$ について連続) で、$x_0, y_0$ をそれぞれ $X, Y$ の中央値とする: $P(X \le x_0) = P(Y \le y_0) = 1/2$. このとき、$P((X - x_0)(Y - y_0) > 0) = 2P(X > x_0, Y > y_0)$ となることを示せ。

**定義 1.12** $\mathbf{R}^n$ のすべての区間 $(a_1, b_1] \times (a_2, b_2] \times \cdots \times (a_n, b_n]$, $a_i < b_i$, $i = 1, 2, \cdots, n$, を含む最小の $\sigma$ 集合族を、$n$ 次元 Borel 集合族といい、$\mathcal{B}(\mathbf{R}^n)$ と表す。$\mathcal{B}(\mathbf{R}^n)$ の元を Borel 集合という。また、関数 $\varphi : \mathbf{R}^n \to \mathbf{R}$ が任意の $B \in \mathcal{B}(\mathbf{R})$ に対して、$\varphi^{-1}(B) \in \mathcal{B}(\mathbf{R}^n)$ を満たすとき、$\varphi$ は $(\mathbf{R}^n, \mathcal{B}(\mathbf{R}^n))$ 上の可測関数、あるいは、単に Borel 可測であるという。

次の定理は定理 1.14 の拡張である。証明は省略する。

**定理 1.20** $X_1, \ldots, X_n$ が $(\Omega, \mathcal{F})$ 上の確率変数、$f : \mathbf{R}^n \to \mathbf{R}$ が $(\mathbf{R}^n, \mathcal{B}(\mathbf{R}^n))$ 上の可測関数のとき、$f(X_1(\omega), \cdots, X_n(\omega))$ は $(\Omega, \mathcal{F})$ 上の確率変数である。

多次元の場合も離散分布, 絶対連続分布が考えられる。絶対連続型の場合のみ確認しておこう: $\mathbf{R}^n$ 上の可測関数 $f(x_1, \ldots, x_n)$ が存在して

$$F_{\mathbf{X}}(x_1, \ldots, x_n) = \int_{-\infty}^{x_n} \int_{-\infty}^{x_{n-1}} \cdots \int_{-\infty}^{x_1} f(t_1, t_2, \ldots, t_n)\, dt_1 dt_2 \cdots dt_n$$

と表せるとき、$F$ を絶対連続な分布関数, $\boldsymbol{X} = (X_1, \ldots, X_n)$ を絶対連続型確率変数という。また、この $f(x_1, \ldots, x_n)$ を $\boldsymbol{X}$ の同時密度関数という。$B \subset \mathbf{R}^n$ が Borel 集合のとき、

$$P((X_1, \ldots, X_n) \in B) = \int\int \cdots \int_B f(t_1, t_2, \ldots, t_n)\, dt_1 dt_2 \cdots dt_n \tag{1.4}$$

となることが後に学ぶ Dynkin 族定理 (定理 3.12, 定理 3.13 も参照のこと) を用い証明できる。

## 1.4 確率変数の独立性

**定義 1.13** (1) 確率変数の列 $X_1, X_2, \ldots, X_n$ が独立であるとは、任意の $x_1, x_2, \ldots, x_n \in \mathbf{R}$ に対して

$$P(X_1 \le x_1, X_2 \le x_2, \cdots, X_n \le x_n) = P(X_1 \le x_1) P(X_2 \le x_2) \cdots P(X_n \le x_n) \tag{1.5}$$

となるときにいう。
(2) 無限個の確率変数の族 $\{X_\lambda\}$ が独立であるとは、その任意の有限部分列 $X_{\lambda_1}, \ldots, X_{\lambda_q}$ が独立であるときにいう。

確率変数の列 $X_1, X_2, \ldots, X_n$ が独立であることと、任意の Borel 集合 $B_1, \ldots, B_n \subset \mathbf{R}$ に対して

$$P(X_1 \in B_1, X_2 \in B_2, \cdots, X_n \in B_n) = P(X_1 \in B_1) P(X_2 \in B_2) \cdots P(X_n \in B_n) \tag{1.6}$$

となることが同値である。後に学ぶ Dynkin 族定理 (定理 3.12, 定理 3.13 も参照のこと) を用い証明できる。特に、離散型確率変数であれば

$$P(X_1 = x_1, X_2 = x_2, \cdots, X_n = x_n) = P(X_1 = x_1) P(X_2 = x_2) \cdots P(X_n = x_n) \qquad \forall x_1, x_2, \ldots, x_n \in \mathbf{R}$$

と同値となる。絶対連続型確率変数の場合は次の定理で見るように同時密度関数がそれぞれの一次元周辺密度関数の積に分解できればよい。

**定理 1.21** 絶対連続型確率変数の列 $X_1, \ldots, X_n$ が独立であることための必要十分条件は、$(X_1, \ldots, X_n)$ の同時密度関数を $f(x_1, \ldots, x_n)$, $X_j$ の周辺密度関数を $f_{X_j}(x)$ $(j = 1, \ldots, n)$ とするとき

$$f(x_1, x_2, \ldots, x_n) = f_{X_1}(x_1) f_{X_2}(x_2) \cdots f_{X_n}(x_n), \qquad \forall x_1, x_2, \ldots, x_n \in \mathbf{R} \tag{1.7}$$

が成立することである。

**証明:** 絶対連続型であるから (1.5) を密度関数で表すと

$$\int_{-\infty}^{x_n} \cdots \int_{-\infty}^{x_1} f(t_1, t_2, \ldots, t_n)\, dt_1 \cdots dt_n = \Big(\int_{-\infty}^{x_1} f_{X_1}(t)\, dt\Big) \times \cdots \times \Big(\int_{-\infty}^{x_n} f_{X_n}(t)\, dt\Big). \tag{1.8}$$

これを $x_1, \cdots, x_n$ でそれぞれ 1 回ずつ偏微分すれば (1.7) を得る。逆に、(1.7) を各 $x_j$ に関して $(-\infty, x_j]$ の範囲で定積分すれば (1.8) を得る。 □

**例 1.22 ($d$ 次元正規分布)** $\boldsymbol{m} = (m_1, \ldots, m_d)'$ と $d$ 次正定値対称行列 $\boldsymbol{\Sigma} = (\sigma_{ij})$ に対して、$d$ 次元確率変数 $\boldsymbol{X} = (X_1, \ldots, X_d)'$ が次の密度関数をもつとき、この分布を $d$ 次元正規分布 $\mathrm{N}(\boldsymbol{m}, \boldsymbol{\Sigma})$ という。

$$f(x_1, \ldots, x_d) = \frac{1}{(2\pi)^{d/2} (\det \boldsymbol{\Sigma})^{1/2}} e^{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{m})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{m})}, \qquad \boldsymbol{x} = (x_1, \ldots, x_d)' \tag{1.9}$$

ここに、$A'$ は行列 $A$ の転置を表す。例えば $\boldsymbol{m}, \boldsymbol{x}$ は縦ベクトルとなる。

このとき、$E[X_k] = m_k$, $\mathrm{Cov}(X_k, X_l) = \sigma_{kl}$, $k, l = 1, 2, \ldots, d$ となる。

**証明:** $\boldsymbol{\Sigma}$ は正定値対称行列なので、線形代数の対角化に関する定理から、直交行列 $P = (p_{ij})$ と対角成分がすべて正の対角行列 $D = (\lambda_{ij})$ がとれて、$P'\boldsymbol{\Sigma}P = D$ とできる。このとき、$\boldsymbol{Y} = (Y_1, \ldots, Y_d)' = P'(\boldsymbol{X} - \boldsymbol{\mu})$ とし、この密度関数 $g(y_1, \ldots, y_d)$ を求めよう。まず、$\boldsymbol{y} = (y_1, \ldots, y_d)' = P'(\boldsymbol{x} - \boldsymbol{\mu})$ とすると、

$$(\boldsymbol{x} - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{x} - \boldsymbol{\mu}) = \boldsymbol{y}'P'\boldsymbol{\Sigma}^{-1}P\boldsymbol{y} = \boldsymbol{y}'D^{-1}\boldsymbol{y} = \sum_{j=1}^{d} \frac{1}{\lambda_{jj}} y_j^2.$$

また、$\boldsymbol{x} = P\boldsymbol{y} + \boldsymbol{\mu}$ より $\frac{\partial \boldsymbol{x}}{\partial \boldsymbol{y}} = \det P = \pm 1$, $\det \boldsymbol{\Sigma} = \det D = \prod_{j=1}^{d} \lambda_{jj}$. よって、

$$g(y_1, \ldots, y_d) = \frac{1}{(2\pi)^{d/2}(\lambda_{11} \cdots \lambda_{dd})^{1/2}} e^{-\frac{1}{2}\sum_{j=1}^{d} \frac{1}{\lambda_{jj}} y_j^2} \Big| \frac{\partial \boldsymbol{x}}{\partial \boldsymbol{y}} \Big| = \prod_{j=1}^{d} \frac{1}{(2\pi\lambda_{jj})^{1/2}} e^{-\frac{1}{2}\frac{y_j^2}{\lambda_{jj}}}.$$

これは、$Y_1, \ldots, Y_d$ が独立で、各 $Y_j$ は正規分布 $\mathrm{N}(0, \lambda_{jj})$ に従うことを意味している。特に、

$$E[Y_i] = 0,\ E[Y_i^2] = \lambda_{ii},\ 1 \le i \le d, \qquad E[Y_iY_j] = E[Y_i]E[Y_j] = 0,\ 1 \le i < j \le d.$$

今、$\boldsymbol{X} = P\boldsymbol{Y} + \boldsymbol{\mu}$ より、$X_k = \sum_{i=1}^{d} p_{ki}Y_i + \mu_k$ なので、

$$E[X_k] = \sum_{i=1}^{d} p_{ki}E[Y_i] + \mu_k = \mu_k,$$

$$\mathrm{Cov}(X_k, X_l) = E[(X_k - \mu_k)(X_l - \mu_l)] = \sum_{i=1}^{d}\sum_{j=1}^{d} p_{ki}p_{lj}E[Y_iY_j] = \sum_{i=1}^{d} p_{ki}p_{li}\lambda_{ii} = \sigma_{kl}$$

を得る。最後の等号は $\boldsymbol{\Sigma} = PDP'$ であるが、$DP'$ の $(i, l)$-成分が $\lambda_{ii}p_{li}$ であるから、$PDP'$ の $(k, l)$-成分が $\sum_{i=1}^{d} p_{ki}\lambda_{ii}p_{li}$ となることを用いた。 □

**問題 1.8** $\mathbf{0} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$, $\boldsymbol{\Sigma} = \begin{pmatrix} 3 & -1 & 1 \\ -1 & 3 & 1 \\ 1 & 1 & 3 \end{pmatrix}$ とし、$\boldsymbol{X} = \begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix}$ が 3 次元正規分布 $N(\mathbf{0}, \boldsymbol{\Sigma})$ に従うとする。

(1) $\boldsymbol{\Sigma}^{-1}$ を求め、$(X_1, X_2, X_3)$ の密度関数を述べよ。ヒント: (1.9) を書き下すだけです。

(2) $\boldsymbol{\Sigma}$ を対称行列 $P$ を用いて対角化せよ。$P'\boldsymbol{\Sigma}P = D$ を対角行列とし、$P = (p_{ij})$, $D = (d_{jj})$ と表す。

(3) $\boldsymbol{Y} = (Y_1, Y_2, Y_3)'$ を $Y_j = p_{j1}X_1 + p_{j2}X_2 + p_{j3}X_3$, $j = 1, 2, 3$ とおく。$(Y_1, Y_2, Y_3)$ の密度関数を求めよ。

(4) $X_1, X_2, X_3$ を $Y_1, Y_2, Y_3$ で表し、$V(X_1)$, $V(X_3)$, $\mathrm{Cov}(X_1, X_2)$, $\mathrm{Cov}(X_1, X_3)$ を求めよ。

**Message:** 今回も前期に行った内容を少し精密化した話でした。問題 1.5–問題 1.8 はそれほど難しくないはずなので、必ず解いておいてください。例 1.22 は少し複雑そうに見えますが、その具体例である問題 1.8 を解くことで理解しやすくなると思います。

[10 月 21 日]

今回は、前期に学んだ期待値について、Lebesgue 積分を用いて厳密な定義を行う。Lebesgue 積分についてはその構成法のごく一部のみを紹介するにとどめるので、関数解析学 II の授業で深く学んでおいてほしい。

## 1.5 Lebesgue 積分

Lebesgue 積分論に基づいて $\mathbf{R}^n$ 上の Borel 可測関数 $\varphi(\boldsymbol{x})$ と $n$ 次元確率変数 $\boldsymbol{X} = (X_1, \cdots, X_n)$ に対して、$\displaystyle\int_\Omega \varphi(\boldsymbol{X}(\omega))\,dP(\omega)$ の定義の概略を述べる。

<u>1st step</u> $\varphi(\boldsymbol{x})$ が単関数のとき、即ち、$\varphi(\boldsymbol{x}) = \sum_{k=1}^N a_k 1_{A_k}(\boldsymbol{x})$ と表せるとき。ここで、$a_1, \ldots, a_N \in \mathbf{R}$, $A_1, \ldots, A_N \in \mathcal{B}(\mathbf{R}^n)$ とし、$1_A(\boldsymbol{x})$ は $A$ の定義関数を表す。

$$1_A(\boldsymbol{x}) = \begin{cases} 1 & (\boldsymbol{x} \in A) \\ 0 & (\boldsymbol{x} \notin A) \end{cases}$$

このとき、次のように定義する。(本来は well-defined かどうか調べなければならないが省略する。)

$$\int_\Omega \varphi(\boldsymbol{X}(\omega))\,dP(\omega) = \sum_{k=1}^N a_k P(\boldsymbol{X} \in A_k). \tag{1.10}$$

<u>2nd step</u> $\varphi(\boldsymbol{x}) \geq 0$ のとき。各 $q \in \mathbf{N}$ に対して、

$$A_{q,i} = \left\{\boldsymbol{x} \in \mathbf{R}^n\,;\, \frac{i}{2^q} \leq \varphi(\boldsymbol{x}) < \frac{i+1}{2^q}\right\} \quad (i = 0, 1, \ldots, q2^q - 1), \qquad A_{q,q2^q} = \left\{\boldsymbol{x} \in \mathbf{R}^n\,;\, \varphi(\boldsymbol{x}) \geq q\right\}$$

とし、$\displaystyle\varphi_q(\boldsymbol{x}) = \sum_{i=0}^{q2^q} \frac{i}{2^q} 1_{A_{q,i}}(\boldsymbol{x})$ と定める。このとき、$\varphi_q(\boldsymbol{x})$ は単関数で (1.10) により $\displaystyle\int_\Omega \varphi_q(\boldsymbol{X}(\omega))\,dP(\omega)$ は定義されるが、$A_{q,i} = A_{q+1,2i} \cup A_{q+2,2i+1}$, $A_{q+1,2i} \cap A_{q+2,2i+1} = \emptyset$ $(i = 0, 1, \ldots, q2^q - 1)$ となるので、

$$\frac{i}{2^q} P(\boldsymbol{X} \in A_{q,i}) \leq \frac{2i}{2^{q+1}} P(\boldsymbol{X} \in A_{q+1,2i}) + \frac{2i+1}{2^{q+1}} P(\boldsymbol{X} \in A_{q+1,2i+1})$$

となり、$\displaystyle\int_\Omega \varphi_q(\boldsymbol{X}(\omega))\,dP(\omega) \leq \int_\Omega \varphi_{q+1}(\boldsymbol{X}(\omega))\,dP(\omega)$ を得る。よって、$\displaystyle\left\{\int_\Omega \varphi_q(\boldsymbol{X}(\omega))\,dP(\omega)\right\}$ は単調増加列なので、その $q \to \infty$ とした極限は無限大かもしれないが定義される。これを用いて、次のように定める。

$$\int_\Omega \varphi(\boldsymbol{X}(\omega))\,dP(\omega) = \lim_{q\to\infty} \int_\Omega \varphi_q(\boldsymbol{X}(\omega))\,dP(\omega). \tag{1.11}$$

<u>3rd step</u> 一般の $\varphi(\boldsymbol{x})$ について。

$$\varphi^+(\boldsymbol{x}) = \max\{\varphi(\boldsymbol{x}), 0\}, \qquad \varphi^-(\boldsymbol{x}) = \max\{-\varphi(\boldsymbol{x}), 0\} \tag{1.12}$$

とおく。このとき、$\varphi(\boldsymbol{x}) = \varphi^+(\boldsymbol{x}) - \varphi^-(\boldsymbol{x})$, $|\varphi(\boldsymbol{x})| = \varphi^+(\boldsymbol{x}) + \varphi^-(\boldsymbol{x})$ となることに注意する。これを用いて、$\displaystyle\int_\Omega |\varphi(\boldsymbol{X}(\omega))|\,dP(\omega) < \infty$ のとき (この判定は 2nd step より可能)、各 $* = +, -$ に対して 2nd step の定義と $0 \leq \varphi^*(\boldsymbol{x}) \leq |\varphi(\boldsymbol{x})|$ より $\displaystyle 0 \leq \int_\Omega \varphi^*(\boldsymbol{X}(\omega))\,dP(\omega) \leq \int_\Omega |\varphi(\boldsymbol{X}(\omega))|\,dP(\omega) < \infty$ となることに注意して

$$\int_\Omega \varphi(\boldsymbol{X}(\omega))\,dP(\omega) = \int_\Omega \varphi^+(\boldsymbol{X}(\omega))\,dP(\omega) - \int_\Omega \varphi^-(\boldsymbol{X}(\omega))\,dP(\omega) \tag{1.13}$$

と定める。

**注意 1.5** 上記の定義は、$\mu(B) = P(\boldsymbol{X} \in B)$ $(B \in \mathcal{B}(\mathbf{R}^n))$ によって定義される確率測度 $\mu$ ($\boldsymbol{X}$ の分布という) に対して、$\displaystyle\int_{\mathbf{R}^n} \varphi(\boldsymbol{x})\,d\mu(\boldsymbol{x})$ を定義したことに相当する。特に、$n = 1$ のときこの積分を確率変数 $X$ の分布関数 $F(x)$ を用いて、$\displaystyle\int_{-\infty}^{\infty} \varphi(x)\,dF(x)$ と表すことがある。

以下の Lebesgue 積分論の定理 1.23, 1.27 が成立することに注意する。(証明は関数解析学で勉強のこと。)

**定理 1.23** (1) 積分は線形性や単調性を有する。(期待値の形で定理**??**で述べる。)
(2) $\varphi(\boldsymbol{x}) \geq 0,\ \displaystyle\int_\Omega \varphi(\boldsymbol{X}(\omega))\,dP(\omega) = 0 \Longrightarrow P(\varphi(\boldsymbol{X}) = 0) = 1.$

**注意 1.6** $P(\varphi(\boldsymbol{X}) = 0) = 1$ のとき、ほとんどいたるところ (almost everywhere) $\varphi(\boldsymbol{X}) = 0$ といい、「$\varphi(\boldsymbol{X}) = 0$ a.e.」と表す。(ほとんど確かに (almost surely) といい、「$\varphi(\boldsymbol{X}) = 0$ a.s.」と表すこともある。)

**定理 1.24** (1) **(単調収束定理)** $(R, \mathfrak{A}, \mu)$ を測度空間とする。可測関数の列 $\{f_q(x)\}$ が非負値で単調増加であるとする, $0 \leq f_1(x) \leq \cdots \leq f_q(x) \leq \cdots$ a.e. このとき、次が成立する。

$$\lim_{q\to\infty} \int_R f_q(x)\,d\mu(x) = \int_R \Big(\lim_{q\to\infty} f_q(x)\Big)\,d\mu(x).$$

(2) **(Lebesgue の収束定理)** $(R, \mathfrak{A}, \mu)$ を測度空間とする。可測関数の列 $\{f_q(x)\}$ に対して、ある可測関数 $M(x)$ で $\displaystyle\int_R M(x)\,d\mu(x) < \infty$ を満たすもの (可積分という) が存在して、$|f_q(x)| \leq M(x)$ a.e. $(q = 1, 2, \ldots)$ を満たすとるとする。このとき、$f(x) = \displaystyle\lim_{q\to\infty} f_q(x)$ a.e. が存在して可測であれば、$f(x)$ も可積分であり、次が成立する。

$$\lim_{q\to\infty} \int_R f_q(x)\,d\mu(x) = \int_R f(x)\,d\mu(x).$$

## 1.6 期待値の定義

§1.5 で準備した積分を用いて、$\varphi \geq 0$ あるいは $\displaystyle\int_\Omega |\varphi(\boldsymbol{X}(\omega))|\,dP(\omega) < \infty$ のとき、期待値 $E[\varphi(\boldsymbol{X})]$ を

$$E[\varphi(\boldsymbol{X})] = \int_\Omega \varphi(\boldsymbol{X}(\omega))\,dP(\omega)$$

で定義する。次の定理が成立する。証明は省略する。

**定理 1.25** $\varphi, \psi : \mathbf{R}^n \to \mathbf{R}$ を Borel 可測関数、$a, b, c$ を定数とする。このとき、次が成立する。
(1) $\varphi(\boldsymbol{X}) = c$ a.e. のとき、$E[\varphi(\boldsymbol{X})] = c.$
(2) $\varphi(\boldsymbol{X}) \leq \psi(\boldsymbol{X})$ a.e. のとき、$E[\varphi(\boldsymbol{X})] \leq E[\psi(\boldsymbol{X})].$
(3) $|E[\varphi(\boldsymbol{X})]| \leq E[|\varphi(\boldsymbol{X})|].$
(4) $E[|\varphi(\boldsymbol{X})|] < \infty,\ E[|\psi(\boldsymbol{X})|] < \infty$ のとき、$E[a\varphi(\boldsymbol{X}) + b\psi(\boldsymbol{X})] = aE[\varphi(\boldsymbol{X})] + bE[\psi(\boldsymbol{X})].$

次に、$\boldsymbol{X}$ が離散型、絶対連続型のときの計算方法に関する定理を与える。$\varphi : \mathbf{R}^n \to \mathbf{R}$ を Borel 可測関数とする。

**定理 1.26** $\boldsymbol{X}$ が離散型確率変数のとき、そのとりうる値を $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots$ とすれば、

$$E[\varphi(\boldsymbol{X})] = \sum_{k=1}^\infty \varphi(\boldsymbol{x}_k) P(\boldsymbol{X} = \boldsymbol{x}_k). \tag{1.14}$$

**証明:** $\varphi \geq 0$ とする。§1.5 の 2nd step のように $\varphi_q(x)$ を定義すると、

$$\begin{aligned} E[\varphi_q(\boldsymbol{X})] &= \sum_{i=0}^{q2^q} \frac{i}{2^q} P(\boldsymbol{X} \in A_{q,i}) = \sum_{i=0}^{q2^q} \frac{i}{2^q} \sum_{k:\boldsymbol{x}_k \in A_{q,i}} P(\boldsymbol{X} = \boldsymbol{x}_k) \\ &= \sum_{i=0}^{q2^q} \sum_{k:\boldsymbol{x}_k \in A_{q,i}} \varphi_q(\boldsymbol{x}_k) P(\boldsymbol{X} = \boldsymbol{x}_k) = \sum_{k=1}^\infty \varphi_q(\boldsymbol{x}_k) P(\boldsymbol{X} = \boldsymbol{x}_k) \end{aligned}$$

ここで、$\{\varphi_q\}$ は非負値単調増加列なので、$\lim_{q\to\infty} \varphi_q(\boldsymbol{x}) = \varphi(\boldsymbol{x})$ に注意して単調収束定理 (個数測度として) を用いれば (1.14) を得る。一般の場合は、(1.13) として定義するのだから、$\varphi \geq 0$ の場合より成立する。 $\square$

**定理 1.27** $\boldsymbol{X}$ が絶対連続型確率変数のとき、その密度関数を $f(x_1, \ldots, x_p)$ とすれば、

$$E[\varphi(\boldsymbol{X})] = \int \cdots \int_{\mathbf{R}^p} \varphi(x_1, \ldots, x_p) f(x_1, \ldots, x_p)\, dx_1 \cdots dx_p. \tag{1.15}$$

**証明:** 記号を簡単にするため 1 次元確率変数として記す。定理 1.26 の証明と同様に $\varphi \geq 0$ とし、§1.5 の 2nd step の $\varphi_q(x)$ を用いると、

$$\begin{aligned} E[\varphi_q(X)] &= \sum_{i=0}^{q2^q} \frac{i}{2^q} P(X \in A_{q,i}) = \sum_{i=0}^{q2^q} \frac{i}{2^q} \int_{A_{q,i}} f(x)\, dx \\ &= \sum_{i=0}^{q2^q} \int_{A_{q,i}} \varphi_q(x) f(x)\, dx = \int_{\mathbf{R}} \varphi_q(x) f(x)\, dx \end{aligned}$$

ここで、$\{\varphi_q(x)f(x)\}$ は非負値単調増加列なので、単調収束定理を用いて (1.15) を得る。一般の $\varphi$ に対しては、(i) と同様に (1.13) として定義するのだから明らか。 $\square$

**注意 1.7** $\mathbf{R}^p$ 上の関数 $f(x_1, \ldots, x_p)$ は通常の意味で Riemann 積分可能であれば、$f(x_1, \ldots, x_p)$ は Lebesgue 可測で、Riemann 積分の意味で定義された積分 $\int \cdots \int_{\mathbf{R}^p} f(x_1, \ldots, x_p)\, dx_1 \cdots dx_p$ の値と、Lebesgue 積分の意味で定義されたそれの値は一致することが知られている。従って、都合のいい方で解釈し式変形し、計算できることに注意する。(詳しくは関数解析学で勉強ください。)

**注意 1.8** 2 次元確率変数 $(X, Y)$ が $X$ は離散型で取り得る値が $x_1, x_2, \ldots$ で、$Y$ は絶対連続型である非負 Borel 関数 $f(x, y)$ を用いて

$$P(X = x_k, Y \leq a) = \int_{-\infty}^{a} f(x_k, y)\, dy \qquad a \in \mathbf{R},\ k = 1, 2, \ldots$$

と表される場合

$$E[\varphi(X, Y)] = \sum_{k=1}^{\infty} \int_{-\infty}^{\infty} \varphi(x_k, y) f(x_k, y)\, dy$$

となることが、定理 1.26, 1.27 と同様に証明できる。

このような厳密な定義を行うことで、次のような期待値と独立性の関係式が証明できる。

**定理 1.28** 確率変数 $X_1, \ldots, X_n$ は独立であるとし、$\varphi_1, \cdots, \varphi_n$ は Borel 可測であるとする。このとき、もし、任意の $j$ で $E[|\varphi_j(X_j)|] < \infty$ であれば、$E[|\varphi_1(X_1)\varphi_2(X_2)\cdots\varphi_n(X_n)|] < \infty$ であり、次が成立する。

$$E[\varphi_1(X_1)\varphi_2(X_2)\cdots\varphi_n(X_n)] = E[\varphi_1(X)]E[\varphi_2(X_2)]\cdots E[\varphi_n(X_n)]. \tag{1.16}$$

**証明:** 簡単のため $n = 2$ の場合に証明する。$X_1 = X, X_2 = Y, \varphi_1 = \varphi, \varphi_2 = \psi$ と記す。

<u>1st step</u> $\varphi(x), \psi(y)$ が単関数 $\varphi(x) = \sum_{i=1}^{M} a_i 1_{A_i}(x),\ \psi(y) = \sum_{j=1}^{N} b_i 1_{B_i}(y)$ のとき。$1_A(x) 1_B(y)$ も $x \in A$ かつ $y \in B$ のときのみ 1 になる定義関数だから、$\varphi(x)\psi(y) = \sum_{i=1}^{M}\sum_{j=1}^{N} a_i b_j 1_{A_i}(x) 1_{B_j}(y)$ も単関数で、独立性の定義 1.13 とそのあとの注意により、

$$\begin{aligned} E[\varphi(X)\psi(Y)] &= \sum_{i=1}^{M}\sum_{j=1}^{N} a_i b_j P(X \in A_i, Y \in B_j) = \sum_{i=1}^{M}\sum_{j=1}^{N} a_i b_j P(X \in A_i) P(Y \in B_j) \\ &= \sum_{i=1}^{M} a_i P(X \in A_i) \sum_{j=1}^{N} b_j P(Y \in B_j) = E[\varphi(X)]E[\psi(Y)] \end{aligned}$$

<u>2nd step</u>　$\varphi \geq 0, \psi \geq 0$ のとき。§1.5 の 2nd step と同様に、各 $q \in \mathbf{N}$ に対して単関数 $\varphi_q, \psi_q$ を定義する。このとき、1st step により

$$E[\varphi_q(X)\psi_q(Y)] = E[\varphi_q(X)]E[\psi_q(Y)]$$

となる。ここで、単調収束定理 (定理 1.27 (1)) により $q \to \infty$ とすることで (1.16) を得る。特に、$E[|\varphi(X)|] < \infty$ かつ $E[|\psi(Y)|] < \infty$ であれば、$E[|\varphi(X)\psi(Y)|] = E[|\varphi(X)|]E[|\psi(X)|] < \infty$ となることもわかった。

<u>3rd step</u>　一般の $\varphi, \psi$ の場合。$(\varphi\psi)(x, y) = \varphi(x)\psi(y)$ と書くものとする。このとき、(1.12) のように、$\varphi^+, \varphi^-, \psi^+, \psi^-, (\varphi\psi)^+, (\varphi\psi)^-$ を定める。このとき、

$$(\varphi\psi)^+(x, y) = \varphi^+(x)\psi^+(y) + \varphi^-(x)\psi^-(y), \qquad (\varphi\psi)^-(x, y) = \varphi^+(x)\psi^-(y) + \varphi^-(x)\psi^+(y)$$

に注意する。これより、

$$\begin{aligned}
&E[\varphi(X)]E[\psi(Y)] = (E[\varphi^+(X)] - E[\varphi^-(X)])(E[\psi^+(Y)] - E[\psi^-(Y)]) \\
&\quad = (E[\varphi^+(X)]E[\psi^+(Y)] + E[\varphi^-(X)]E[\psi^-(Y)]) - (E[\varphi^+(X)]E[\psi^-(Y)] + E[\varphi^-(X)]E[\psi^+(Y)]) \\
&\quad = (E[\varphi^+(X)\psi^+(Y)] + E[\varphi^-(X)\psi^-(Y)]) - (E[\varphi^+(X)\psi^-(Y)] + E[\varphi^-(X)\psi^+(Y)]) \\
&\quad = E[(\varphi\psi)^+(X, Y)] - E[(\varphi\psi)^-(X, Y)] = E[\varphi(X)\psi(Y)] \qquad \square
\end{aligned}$$

以下、$L^2(\Omega)$ で、確率変数 $X$ で $E[X^2] < \infty$ を満たすもの全体を表すとする。

**定理 1.29 (Cauchy-Schwarz の不等式)** $X, Y \in L^2(\Omega)$ のとき、$E[XY]$ は定義され、

$$|E[XY]|^2 \leq E[X^2]E[Y^2] \tag{1.17}$$

が成立する。等号成立のための必要十分条件は、$(a, b) \neq (0, 0)$ なる定数 $a, b$ が存在して $aX + bY = 0$ a.e. となることである。

**証明:** $|XY| \leq \frac{1}{2}(X^2 + Y^2)$ より、$E[|XY|] \leq \frac{1}{2}(E[X^2] + E[Y^2]) < \infty$ となり $E[XY]$ は定義される。$E[X^2] = 0$ のとき、定理 1.25, 注意 1.6 より $X = 0$ a.e. なので (1.17) は $0 = 0$ として明らかに成立する。$E[X^2] > 0$ のとき、

$$0 \leq E[(tX + Y)^2] = t^2E[X^2] + 2tE[XY] + E[Y^2] = E[X^2]\Big\{t + \frac{E[XY]}{E[X^2]}\Big\}^2 - \frac{(E[XY])^2}{E[X^2]} + E[Y^2] \tag{1.18}$$

が任意の $t$ に対して成立するので、$t = -E[XY]/E[X^2]$ することで (1.17) は成立することがわかる。等号成立のための必要十分条件は省略する。　$\square$

この定理を用いて $X, Y \in L^2(\Omega)$ なる確率変数 $X, Y$ の共分散 $\mathrm{Cov}(X, Y)$ が定義できる:

$$\mathrm{Cov}(X, Y) = E[(X - \mu_1)(Y - \mu_2)] = E[XY] - \mu_1E[X] - \mu_2E[Y] + \mu_1\mu_2 = E[XY] - \mu_1\mu_2,$$

ただし、$\mu_1 = E[X], \mu_2 = E[Y]$ とした。

特に、$X$ と $Y$ が独立なら $\mathrm{Cov}(X, Y) = 0$ が成立する。逆が成立しないことは次の問題 1.9 からわかる。ただし、正規分布に従うときは共分散が 0 であれば独立となる ($cf.$ 問題 1.10)。

**問題 1.9** 2 次元確率変数 $(X, Y)$ の同時密度関数が $f(x, y) = 1/\pi$, $(x^2 + y^2 \leq 1)$, $= 0$ (その他) で与えられるとする。このとき、$\mathrm{Cov}(X, Y) = 0$ であるが、$X$ と $Y$ が独立とならないことを示せ。

**問題 1.10** $\boldsymbol{X} = (X_1, \ldots, X_d)'$ が $d$ 次元正規分布 $\mathrm{N}(\boldsymbol{m}, \boldsymbol{\Sigma})$ に従うとき、$X_1, \ldots, X_d$ が独立であるための必要十分条件は $\boldsymbol{\Sigma}$ が対角行列であることである。これを例 1.22 を用いて証明せよ。

**Message:** 今回は Lebesgue 積分の単関数、非負値、一般の場合と定義される流れを見ておいていただければと思っています。問題 1.9, 問題 1.10 はそれほど難しくないなので、解いておいてください。

[10 月 28 日]

# 2 大数の法則

## 2.1 確率変数の極限

$(\Omega, \mathcal{B}, P)$ を確率空間とする。

この節では、$(\Omega, \mathcal{B}, P)$ 上の確率変数列 $\{X_n\}$ の確率変数 $X$ への収束について述べる。

**定義 2.1** (1) **(概収束)** $X_n$ が $X$ に概収束 (almost surely convergence) するとは、$P$-a.a. $\omega$ に対して $X_n(\omega) \to X(\omega)$ $(n \to \infty)$ であるとき、つまり

$$P\Big(\lim_{n\to\infty} X_n = X\Big) = 1$$

あるいは、更に正確に言えば

$$P\Big(\Big\{\omega \in \Omega\,;\, \lim_{n\to\infty} X_n(\omega) = X(\omega)\Big\}\Big) = 1$$

であるときにいう。$X_n \to X$ a.s. と表す。($X_n \to X$ a.e. とも表す。)

(2) **(確率収束)** $X_n$ が $X$ に確率収束 (convergence in probability) するとは、任意の $\varepsilon > 0$ に対して、

$$\lim_{n\to\infty} P(|X_n - X| \geq \varepsilon) = 0$$

のときにいう。$X_n \to X$ in prob. と表す。

(3) **($L^r$-収束)** $r \geq 1$ として、$X_n$ が $X$ に $L^r$-収束するとは、

$$\lim_{n\to\infty} E[|X_n - X|^r] = 0$$

のときにいう。$X_n \to X$ in $L^r$ と表す。$r$ 次平均収束 (convergence in the mean of order $r$) ともいう。

**注意 2.1** 確率変数がなす空間上に確率収束, $L^r$-収束が定める位相は、それぞれ距離付け可能である。(前者は [F] p.59 を参照のこと、後者は $L^r$-ノルムが定める距離である。$r=1$ の場合は自明なので、$r=2$ の場合を問題 2.2 とします。) 概収束は距離付けできない ($cf.$ [F] p.59)。

**問題 2.1** ((3) は少し面倒です。) (1) $X_n \to X$ a.s. かつ $Y_n \to Y$ a.s. (概収束) のとき、$X_n + Y_n \to X + Y$ a.s. , $X_n \cdot Y_n \to X \cdot Y$ a.s. を示せ。[*1]

(2) $X_n \to X$ in prob. かつ $Y_n \to Y$ in prob. のとき、$X_n + Y_n \to X + Y$ in prob. を示せ。

(3) $X_n \to X$ in prob. かつ $Y_n \to Y$ in prob. のとき、$X_n Y_n \to XY$ in prob. を示せ。

**問題 2.2** $X \in L^r$ に対して、$\|X\|_r = (E[|X|^r])^{1/r}$ とおく。これを $L^r$-ノルムという。

(1) Cauchy-Shwarz の不等式 (定理 1.29) を用いて、$X, Y \in L^2$ に対して $\|X+Y\|_2 \leq \|X\|_2 + \|Y\|_2$ を示せ。

(2) $X_n \to X$ in $L^2$ かつ $Y_n \to Y$ in $L^2$ のとき、$X_n + Y_n \to X + Y$ in $L^2$ を示せ。

**問題 2.3** $a > 0$ する。確率変数列 $\{X_n\}$ が独立で一様分布 $U(0,a)$ に従うとき、$Y_n = \max\{X_1, X_2, \cdots, X_n\}$ とおく。$\{Y_n\}$ は $a$ に確率収束することを示せ。

**問題 2.4** $\{X_n\}$ は独立で同じ分布に従う確率変数列で、その密度関数が $f(x) = 2x1_{(0,1]}(x)$ であるとする。このとき、$Y_n = \max\{X_1, X_2, \cdots, X_n\}$ とおく。

(1) $Y_n$ の分布関数 $F_{Y_n}(y) = P(Y_n \leq y)$ と密度関数 $f_{Y_n}(y)$ を求めよ。

(2) $r \geq 1$ に対し、$E[|Y_n - 1|^r]$ を求め、$Y_n \to 1$ in $L^r$ を示せ。

---

[*1] 一部の問題の解答は昨年度の確率統計学 II の演習問題にあります。WebClass にそれを upload します。参照ください。

次に定義 2.1 で定めた収束の強弱の関係を調べる。

**定理 2.1** (1) $X_n$ が $X$ に概収束すれば、確率収束する。
(2) $X_n$ が $X$ に $L^r$-収束すれば、確率収束する。

**証明:** (1) $X_n$ が $X$ に収束するような $\omega$ の集合は

$$\left\{\lim_{n\to\infty} X_n = X\right\} = \bigcap_{j=1}^{\infty}\bigcup_{n=1}^{\infty}\bigcap_{m=n}^{\infty} A_{m,j} \tag{2.1}$$

と表すことができる。ただし、

$$A_{m,j} = \left\{|X_m - X| < \frac{1}{j}\right\}$$

である。$X_n \to X$ a.s. であるから、この事象の確率は 1 である。(仮定より $\forall m, j$ に対して $A_{m,j} \in \mathcal{B}$ であるから、$\left\{\lim\limits_{n\to\infty} X_n = X\right\} \in \mathcal{B}$ となることに注意する。) ここで、$A_{m,j} \supset A_{m,j+1}$ $(\forall m, j)$ であるから、(2.1) により $\bigcup\limits_{n=1}^{\infty}\bigcap\limits_{m=n}^{\infty} A_{m,j} \supset \bigcup\limits_{n=1}^{\infty}\bigcap\limits_{m=n}^{\infty} A_{m,j+1} \supset \cdots \supset \left\{\lim\limits_{n\to\infty} X_n = X\right\}$ となるので、

$$P\Big(\bigcup_{n=1}^{\infty}\bigcap_{m=n}^{\infty} A_{m,j}\Big) = 1 \qquad (\forall j \in \mathbf{N})$$

である。さらに、$B_{n,j} = \bigcap\limits_{m=n}^{\infty} A_{m,j}$ とすると、$B_{n,j} \subset B_{n+1,j}$ $(\forall n, j)$ だから、

$$\lim_{n\to\infty} P\left(B_{n,j}\right) = P\Big(\bigcup_{n=1}^{\infty} B_{n,j}\Big) = 1$$

となる。ここで、$B_{n,j} \subset A_{n,j}$ であるから、以上より $\forall j \in \mathbf{N}$ に対して、

$$\lim_{n\to\infty} P\left(A_{n,j}\right) = \lim_{n\to\infty} P\Big(|X_n - X| \leq \frac{1}{j}\Big) = 1$$

であることがわかった。ここで、$\forall \varepsilon > 0$ が与えられたとき、$j$ を十分大きくとって $1/j < \varepsilon$ とすれば

$$\left\{|X_n - X| \leq \frac{1}{j}\right\} \subset \{|X_n - X| < \varepsilon\}$$

だから

$$\lim_{n\to\infty} P\big(|X_n - X| < \varepsilon\big) = 1$$

が得られ、余事象を考えれば、$X_n$ が $X$ に確率収束していることがわかる。(2) の証明には次を必要とする。

**命題 2.2 (チェビシェフ (Chebyshev) の不等式)** $r > 0,\ \lambda > 0$ と確率変数 $Y$ について次の不等式が成立する。

$$P\left(|Y| \geq \lambda\right) \leq \frac{1}{\lambda^r} E[|Y|^r]$$

**証明:** まず、次に注意する。

$$1_{[\lambda,\infty)}(|Y|) \leq \Big(\frac{|Y|}{\lambda}\Big)^r 1_{[\lambda,\infty)}(|Y|) \leq \frac{|Y|^r}{\lambda^r}$$

であるから ($1_A$ は定義関数、即ち、$1_A(x) = 1$ $(x \in A)$, $1_A(x) = 0$ $(x \notin A)$ なる関数)、両辺の期待値をとって

$$P(|Y| \geq \lambda) = E[1_{[\lambda,\infty)}(|Y|)] \leq E\Big[\frac{|Y|^r}{\lambda^r}\Big] = \frac{1}{\lambda^r} E[|Y|^r]. \qquad \square$$

**定理 2.1(2) の証明:** 仮定と Chebyshev の不等式により

$$P(|X_n - X| \geq \varepsilon) \leq \frac{1}{\varepsilon^r} E[|X_n - X|^r] \to 0 \qquad (n \to \infty)$$

となる。 $\square$

**例 2.3** 定理 2.1(1), (2) の逆は、必ずしも成立しない。また、概収束と $L^r$-収束の間に強弱の関係はない。$\Omega = [0,1]$, $\mathcal{B}$ をそれ上の Borel 集合全体, $P$ を Lebesgue 測度として、以下問題としてそれを例示する。

**問題 2.5** ($L^r$-収束する (従って確率収束する) が、概収束しない例)
$X_{n,k}(\omega) = 1_{[\frac{k-1}{n},\frac{k}{n})}(\omega)$, $\omega \in [0,1]$, $k = 1,\dots,m$, $m = 1,2,\dots$ とおき、これを

$$X_1 = X_{1,1},\ X_2 = X_{2,1},\ X_3 = X_{2,2},\ X_4 = X_{3,1},\ X_5 = X_{3,2},\ X_6 = X_{3,3},\ X_7 = X_{4,1},\ \dots$$

のように並べた列 $\{X_n\}$ を考える。このとき、$E[|X_n|^r]$ を求め、$X_n \to 0$ in $L^r$ を示せ。一方、$X_n$ が概収束しないことを示せ。

**問題 2.6** (概収束する (従って確率収束する) が、$L^r$-収束しない例)
$X_n(\omega) = n1_{(0,\frac{1}{n})}(\omega)$, $\omega \in [0,1]$ を考えると、これは $X_n \to 0$ a.s. である (概収束する) が、$L^r$-収束しないことを示せ。注意: 定理 2.4 より $X_n \to X$ in $L^r$ ならば 部分列 $\{X_{n_k}\}$ があって $X_{n_k} \to X$ a.s. となるが、$X_n \to 0$ a.s. なので $X = 0$ となることがわかる。

**定理 2.4** $X_n$ が $X$ に確率収束するならば、適当に部分列を選んで概収束するようにできる。特に、$L^r$-収束すれば (確率収束するから)、適当に部分列を選んで概収束するようにできる。

**定理 2.5 (Borel-Cantelli の定理)** $\{B_n\} \subset \mathcal{B}$ に対し、$\sum_{n=1}^{\infty} P(B_n) < \infty$ ならば $P(\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty} B_k) = 0.$

**証明:** $\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty} B_k \subset \bigcup_{k=n}^{\infty} B_k$ $(\forall n)$ より、

$$0 \le P(\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty} B_k) \le P(\bigcup_{k=n}^{\infty} B_k) \le \sum_{k=n}^{\infty} P(B_k).$$

ここで、$\sum_{k=1}^{\infty} P(B_k) < \infty$ より $\sum_{k=n}^{\infty} P(B_k) \to 0$ $(n \to \infty)$. よって、$P(\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty} B_k) = 0.$ □

**定理 2.4 の証明:** 各 $k \in \mathbf{N}$ に対して、$X_n$ は $X$ に確率収束するから、$\varepsilon = \frac{1}{2^k}$ として、ある $N_k$ があって

$$n \ge N_k \quad \Longrightarrow \quad P\Big(|X_n - X| \ge \frac{1}{2^k}\Big) \le \frac{1}{2^k}$$

とできる。特に、ある番号の列 $n_1 < n_2 < \cdots < n_k < \cdots$ があって ($n_1 = N_1$, $n_k = \max\{N_k, n_{k-1}+1\}$, $k \ge 2$ とせよ)、$P\Big(|X_{n_k} - X| \ge \dfrac{1}{2^k}\Big) \le \dfrac{1}{2^k}$ とできる。
この $X_{n_k}$ が $X$ に概収束することを示す。$C_k = \Big\{|X_{n_k} - X| \ge \dfrac{1}{2^k}\Big\}$ とおくと、

$$\sum_{k=1}^{\infty} P(C_k) \le \sum_{k=1}^{\infty} \frac{1}{2^k} = 1 < \infty$$

であるから、Borel-Cantelli の定理により、$P\Big(\bigcap_{l=1}^{\infty}\bigcup_{k=l}^{\infty} C_k\Big) = 0.$ ここで、

$$\omega \in \Big(\bigcap_{l=1}^{\infty}\bigcup_{k=l}^{\infty} C_k\Big)^{\mathrm{c}} = \bigcup_{l=1}^{\infty}\bigcap_{k=l}^{\infty} C_k{}^{\mathrm{c}} \text{とすると} \exists l \in \mathbf{N} \text{ such that } \forall k \ge l \text{ に対し } |X_{n_k}(\omega) - X(\omega)| < \frac{1}{2^k}$$

$$\text{すなわち} \lim_{k\to\infty} X_{n_k}(\omega) = X(\omega)$$

となる。これは、$X_{n_k}$ は $X$ に概収束することを意味している。 □

[11 月 4 日]

## 2.2　大数の弱法則

確率空間 $(\Omega, \mathcal{B}, P)$ 上の確率変数列 $\{X_n\}$ に対して、その平均 $S_n/n = \sum_{i=1}^{n} X_i/n$ の収束について議論する。

**定義 2.2** ある数列 $\{c_n\}$ に対し、
(1)　$S_n/n - c_n$ が 0 に確率収束するとき、大数の弱法則 (weak law of large numbers) が成立すると、
(2)　$S_n/n - c_n$ が 0 に概収束するとき、大数の強法則 (strong law of large numbers) が成立するという。

**定理 2.6** $X_1, X_2, \ldots$ が無相関、つまりどの組 $i, j$ $(i \neq j)$ をとっても $\mathrm{Cov}(X_i, X_j) = 0$ で (特に、$X_i, X_j$ が独立なら無相関に注意)、

$$\sup_n V(X_n) < \infty$$

ならば、数列 $\{c_n\}$ が存在し $S_n/n - c_n$ は 0 に $L^2$-収束する。特に、大数の弱法則を満たす。$V(X) = E[(X - E[X])^2]$ は $X$ の分散を、$\mathrm{Cov}(X, Y) = E[(X - E[X])(Y - E[Y])]$ は $X$ と $Y$ の共分散を表す。

**証明:** $L^2$-収束することが示されれば、大数の弱法則は定理 2.1 から従う。$m_n = E[X_n]$ とし、$c_n = \frac{1}{n}\sum_{j=1}^{n} m_j$ とすると、

$$\begin{aligned} E\Big[\Big(\frac{S_n}{n} - c_n\Big)^2\Big] &= \frac{1}{n^2} E\Big[\Big\{\sum_{j=1}^{n}(X_j - m_j)\Big\}^2\Big] = \frac{1}{n^2}\sum_{i,j=1}^{n} E[(X_i - m_i)(X_j - m_j)] \\ &= \frac{1}{n^2}\Big\{\sum_{j=1}^{n} V(X_j) + 2\sum_{1 \le i < j \le n} \mathrm{Cov}(X_i, X_j)\Big\} \ \le \frac{1}{n}\sup_j V(X_j) \ \to 0 \quad (n \to \infty) \end{aligned}$$

となり、$L^2$-収束することがわかる。　　□

**例 2.7 (株式投資)**　ある株価の月ごとの成長率が確率変数で $X_1, X_2, \ldots$ ($n$ ヶ月目に $n-1$ ヶ月目に比べて $X_n$ 倍になる) と表せるとする。この株の株価は $n$ ヵ月後には元値の $Y_n = \prod_{j=1}^{n} X_j$ 倍になる。$Y_n$ が長期的にどうなるか予想したい。ここでは、簡単のため $X_1, X_2, \ldots$ を区間 $(a, b)$ $(0 < a < 1 < b)$ の値をとる i.i.d. とする。(i.i.d. は独立で同分布に従う independently, identically distributed の略。) $Y_n$ の対数を取ると、

$$\log Y_n = \sum_{j=1}^{n} \log X_j$$

で $\log X_1, \log X_2, \ldots$ は i.i.d で有界 (従って分散が存在する) なので、定理 2.6 より $\forall \varepsilon > 0$ に対して

$$P\Big(\Big|\frac{1}{n}\log Y_n - l\Big| \le \varepsilon\Big) \to 1, \qquad \text{ただし } l = E[\log X_1], \qquad \text{すなわち、}$$

$$P\left(e^{(l-\varepsilon)n} \le Y_n \le e^{(l+\varepsilon)n}\right) \to 1 \tag{2.2}$$

となる。$\varepsilon > 0$ は任意に小さくとれるから、これより月ごとの平均的な成長率は $e^l$ となる。

一方、単純に $Y_n$ の平均をとると独立性より

$$E[Y_n] = E[X_1]\cdots E[X_n] = m^n, \qquad \text{ただし } m = E[X_1]$$

となり、ここから「月ごとの平均的な成長率は $m$」と思ってしまいそうだが、$e^l$ のほうが正しいことは (2.2) から明らかである。

例えば、$P(X_1 = 1.3) = 3/5$, $P(X_1 = 0.6) = 2/5$ の場合を考えると、

$$l = E[\log X_1] = \frac{3}{5}\log 1.3 + \frac{2}{5}\log 0.6 = -0.0469\cdots, \qquad m = E[X_1] = \frac{3}{5}1.3 + \frac{2}{5}0.6 = 1.02$$

となり $e^l < 1 < m$. 従ってこの場合 $m > 1$ を平均的な成長率と勘違いして投資すると、(2.2) により資産は指数的に減衰してしまう。(Jensen の不等式 "$\varphi(x)$ が下に凸のとき、$\varphi(E[X]) \le E[\varphi(X)]$" により、一般に $e^l \le m$ となることが証明できる。)

次は、任意の連続関数が有界閉集合上では多項式により一様に近似されることを意味している。定理 2.4 と同様に証明できるので、ここで扱う。

**定理 2.8 (Bernstein の多項式近似定理)** $f(x)$ を $[0,1]$ 上の連続関数とするとき、次が成立する。

$$\lim_{n\to\infty}\max_{0\le p\le 1}\left|f(p)-\sum_{k=0}^{n} f\Big(\frac{k}{n}\Big)\binom{n}{k}p^k(1-p)^{n-k}\right| = 0 \tag{2.3}$$

絶対値の中の第 2 項は $p$ の $n$ 次多項式となっているが、これを Bernstein の多項式ということがある。

**証明:** $0 \le p \le 1$ を任意にとり固定する。$X_1, X_2, \dots$ を i.i.d. で、各 $n$ で $P(X_n = 1) = p$, $P(X_n = 0) = 1-p$ を満たすとする。このとき、$S_n = \sum_{k=1}^{n} X_k$ とおくと、$S_n$ は二項分布 $B(n,p)$ に従うので、

$$E\Big[f\Big(\frac{S_n}{n}\Big)\Big] = \sum_{k=0}^{n} f\Big(\frac{k}{n}\Big)P(S_n = k) = \sum_{k=0}^{n} f\Big(\frac{k}{n}\Big)\binom{n}{k}p^k(1-p)^{n-k}. \tag{2.4}$$

一方、$\forall\delta > 0$ に対して、Chebyshev の不等式により

$$\begin{aligned} P\Big(\Big|\frac{S_n}{n}-p\Big| \ge \delta\Big) &= P(|S_n - np| \ge n\delta) \le \frac{1}{(n\delta)^2}E[|S_n-np|^2] = \frac{1}{(n\delta)^2}V(S_n) \\ &= \frac{np(1-p)}{(n\delta)^2} = \frac{1}{n\delta^2}\Big\{-\Big(p-\frac{1}{2}\Big)^2+\frac{1}{4}\Big\} \le \frac{1}{4n\delta^2}, \end{aligned}$$

ここで、$V(S_n)$ は $S_n$ の分散であり $np(1-p)$ となることを用いた。よって、$\|f\|_\infty = \sup_{x\in[0,1]}|f(x)|$, $u_f(\delta) = \sup_{|x-y|<\delta}|f(x)-f(y)|$ とおくと、

$$\begin{aligned} \Big|f(p) - E\Big[f\Big(\frac{S_n}{n}\Big)\Big]\Big| &= \Big|E\Big[f(p) - f\Big(\frac{S_n}{n}\Big)\Big]\Big| \le E\Big[\Big|f(p)-f\Big(\frac{S_n}{n}\Big)\Big|\Big] \\ &= E\Big[\Big|f(p)-f\Big(\frac{S_n}{n}\Big)\Big|1_{\{|\frac{S_n}{n}-p|\ge\delta\}}\Big] + E\Big[\Big|f(p)-f\Big(\frac{S_n}{n}\Big)\Big|1_{\{|\frac{S_n}{n}-p|<\delta\}}\Big] \\ &\le 2\|f\|_\infty P\Big(\Big|\frac{S_n}{n}-p\Big| \ge \delta\Big) + u_f(\delta)P\Big(\Big|\frac{S_n}{n}-p\Big| < \delta\Big) \le \frac{\|f\|_\infty}{2n\delta^2} + u_f(\delta). \end{aligned}$$

ここで、$f(x)$ は $[0,1]$ で連続であるから一様連続なので、$\lim_{\delta\to 0} u_f(\delta) = 0$. よって、任意の $\forall\varepsilon > 0$ に対してある $\delta > 0$ があって、$u_f(\delta) < \varepsilon/2$. 次に $n$ を $n > \|f\|_\infty/(\varepsilon\delta^2)$ とすれば、

$$\Big|f(p) - E\Big[f\Big(\frac{S_n}{n}\Big)\Big]\Big| < \frac{\varepsilon}{2}+\frac{\varepsilon}{2} = \varepsilon.$$

ここで $n$ は $p$ に依存していないので (2.4) とあわせて、(2.3) は示された。 □

**問題 2.7** (1) $\{X_k\}$ は独立な確率変数列で、各 $X_k$ の密度関数が $f_{X_k}(x) = \dfrac{3}{4k}\Big(1-\dfrac{x^2}{k^2}\Big)1_{(-k,k)}(x)$ であるとする。$a > 0$ を定数とし、$Y_n = \dfrac{1}{n^a}\sum\limits_{k=1}^{n} X_k$ とするとき、$E[|Y_n|^2]$ を求め、$Y_n \to 0$ in $L^2$ となる $a > 0$ の範囲を求めよ。

(2) $a, b, c$ を定数とする。確率変数列 $\{X_n\}$ が、各 $k \in \mathbf{N}$ に対して $E[X_k] = a$, $V(X_k) = b$, $\mathrm{Cov}(X_k, X_{k+1}) = c$, $\mathrm{Cov}(X_k, X_{k+p}) = 0$ $(p \ge 2)$ を満たすとし、$S_n = \sum\limits_{k=1}^{n} X_k$ とおく。$E[S_n]$ と $V(S_n)$ を $a, b, c$ と $n$ の式で表し、$\frac{1}{n}S_n$ が $a$ に $L^2$-収束することを示せ。

もう少し詳しく大数の弱法則を調べるため、以下の Lebesgue 積分の道具 (定理 2.9–2.11) を導入する。証明は関数解析学 II で学習するものとして略す[*2]。(関数解析学 I,II の講義の教科書を調べてください。)

**定理 2.9 (単調収束定理)** 非負値の確率変数列 $\{X_n\}$ が単調増加 $0 \le X_1 \le X_2 \le \cdots \le X_n \le \cdots$ であれば、次が成立する。

$$\lim_{n\to\infty} E[X_n] = E[\big(\lim_{n\to\infty} X_n\big)].$$

**定理 2.10 (Lebesgue の収束定理)** 確率変数列 $\{X_n\}$ が $X$ に概収束し、かつ非負確率変数 $Y$ で可積分 $(E[Y] < \infty)$ なものが存在し任意の $n \in \mathbf{N}$ に対して $|X_n| \le Y$ を満たすならば次が成立する。

$$\lim_{n\to\infty} E[X_n] = E[X].$$

**定理 2.11 (Fubini の定理)** $(R_i, \mathfrak{A}_i, \mu_i)$, $i = 1, 2$, を二つの $\sigma$-有限な測度空間とする。関数 $f(x,y)$ がこの直積測度空間の関数として可測[*3]で、$f(x,y) \ge 0$ または $\displaystyle\int_{R_1\times R_2} |f(x,y)|\, d(\mu_1 \otimes \mu_2)(x,y) < \infty$ を満たせば、次が成立する。

$$\int_{R_1\times R_2} f(x,y)\, d(\mu_1\otimes\mu_2)(x,y) = \int_{R_2}\bigg(\int_{R_1} f(x,y)\, d\mu_1(x)\bigg) d\mu_2(y) = \int_{R_1}\bigg(\int_{R_2} f(x,y)\, d\mu_2(y)\bigg) d\mu_1(x).$$

**定理 2.12** $X_1, X_2, \ldots$ は組ごとに独立とし、ある $b_n > 0$, $b_n \to \infty$ $(n \to \infty)$ があって、$n \to \infty$ のとき、

$$\text{(a)} \quad \sum_{k=1}^{n} P(|X_k| > b_n) \to 0, \qquad\qquad \text{(b)} \quad \frac{1}{{b_n}^2}\sum_{k=1}^{n} E[{X_k}^2 1_{\{|X_k|\le b_n\}}] \to 0$$

とする。このとき、$S_n = \sum\limits_{k=1}^{n} X_k$, $a_n = \sum\limits_{k=1}^{n} E[X_k 1_{\{|X_k|\le b_n\}}]$ とすると、$\dfrac{S_n - a_n}{b_n}$ は 0 に確率収束する。

**証明:** $\tilde{S}_n = \sum_{k=1}^{n} X_k 1_{\{|X_k|\le b_n\}}$ とすると、$\forall\varepsilon > 0$ に対して、

$$P\Big(\Big|\frac{S_n - a_n}{b_n}\Big| > \varepsilon\Big) \le P(S_n \ne \tilde{S}_n) + P\Big(\Big|\frac{\tilde{S}_n - a_n}{b_n}\Big| > \varepsilon\Big).$$

ここで、$\{S_n = \tilde{S}_n\} \supset \bigcap_{k=1}^{n}\{X_k 1_{\{|X_k|\le b_n\}} = X_k\} = \bigcap_{k=1}^{n}\{|X_k| \le b_n\}$ より、

$$P(S_n \ne \tilde{S}_n) \le P\Big(\bigcup_{k=1}^{n}\{|X_k| \le b_n\}^{\mathrm{c}}\Big) \le \sum_{k=1}^{n} P(|X_k| > b_n) \to 0, \qquad \text{((a) による)}.$$

一方、$a_n = E[\tilde{S}_n]$ であるから、Chebyshev の不等式により

$$\begin{aligned} P\Big(\Big|\frac{\tilde{S}_n - a_n}{b_n}\Big| > \varepsilon\Big) &\le \frac{1}{\varepsilon^2} E\Big[\Big|\frac{\tilde{S}_n - a_n}{b_n}\Big|^2\Big] = \frac{1}{\varepsilon^2 b_n^2} V(\tilde{S}_n) = \frac{1}{\varepsilon^2 b_n^2}\sum_{k=1}^{n} V(X_k 1_{\{|X_k|\le b_n\}}) \\ &\le \frac{1}{\varepsilon^2 b_n^2}\sum_{k=1}^{n} E[X_k^2 1_{\{|X_k|\le b_n\}}] \to 0, \qquad \text{((b) による)}. \qquad \Box \end{aligned}$$

**定理 2.13** $X_1, X_2, \ldots$ は i.i.d. で、

$$xP(|X_1| > x) \to 0 \qquad (x \to \infty) \tag{2.5}$$

とする。このとき、$S_n = \sum\limits_{k=1}^{n} X_k$, $c_n = E[X_1 1_{\{|X_1|\le n\}}]$ とすると、$\dfrac{S_n}{n} - c_n$ は 0 に確率収束する。

---

[*2] 期待値を Lebesgue 積分論の書き方で、$E[X] = \int_\Omega X(\omega)\, dP(\omega)$ となることに注意せよ。

[*3] 例えば、$R_2 = \mathbf{R}$ で $\mathfrak{A}_2$ をその Borel 集合族とするとき、$f(x,y)$ が $\forall y$ を固定すると $x$ について $\mathfrak{A}_1$-可測で $\forall x$ を固定すると $y$ について右連続であれば、$f(x,y)$ は直積測度空間で可測となる (*cf*. 伊藤清三: ルベーグ積分入門 (1963), pp.68–69)。

**証明:** $X_1, X_2, \ldots$ は i.i.d. なので、定理 2.12 の $a_n$ に対して $a_n = nc_n$ となることに注意する。よって、定理 2.12 の条件 (a), (b) を $b_n = n$ に対して示せばよい。(a) は

$$\sum_{k=1}^{n} P(|X_k| > n) = nP(|X_1| > n)$$

だから (2.5) より明らか。(b) のために次の補題を準備する。

**補題 2.14** $Y \geq 0,\, p > 0$ とすると、$E[Y^p] = \displaystyle\int_0^\infty py^{p-1}P(Y > y)\,dy.$

**証明:**

$$\begin{aligned}(\text{右辺}) &= \int_0^\infty py^{p-1}\Big(\int_\Omega 1_{(y,\infty)}(Y(\omega))\,dP(\omega)\Big)dy = \int_\Omega\Big(\int_0^\infty py^{p-1}1_{(-\infty,Y(\omega))}(y)\,dy\Big)dP(\omega) \\ &= \int_\Omega\Big(\int_0^{Y(\omega)} py^{p-1}\,dy\Big)dP(\omega) = \int_\Omega Y(\omega)^p dP(\omega) = (\text{左辺}),\end{aligned}$$

ここで、第 2 の等号において、$py^{p-1}1_{(y,\infty)}(Y(\omega)) = py^{p-1}1_{(-\infty,Y(\omega))}(y) \geq 0$ に注意して Fubini の定理 (定理 2.11) を用いた。 $\square$

**定理 2.13 の証明の続き:** $Y_n = |X_1|1_{\{|X_1|\leq n\}}$ とすると、$Y_n \geq 0$ より補題 2.14 から

$$E[Y_n^2] = \int_0^\infty 2yP(Y_n > y)\,dy = \int_0^n 2yP(Y_n > y)\,dy.$$

ここで、第 2 の等号は $P(Y_n > n) = 0$ より $P(Y_n > y) = 0$ $(y \geq n)$ となることを用いた。よって、

$$\begin{aligned}\frac{1}{n^2}\sum_{k=1}^{n} E[X_k^2 1_{\{|X_k|\leq n\}}] &= \frac{1}{n}E[X_1^2 1_{\{|X_1|\leq n\}}] = \frac{1}{n}E[Y_n^2] \\ &= \frac{1}{n}\int_0^n 2yP(Y_n > y)\,dy \leq \frac{1}{n}\int_0^n 2yP(|X_1| > y)\,dy\end{aligned}$$

となるが、次の問題 2.8(2) から、(2.5) より定理 2.12 の条件 (b) が成り立つことがわかる。 $\square$

**問題 2.8** (1) 数列 $\{a_n\}$ が $\displaystyle\lim_{k\to\infty} a_k = \alpha$ を満たせば、$\displaystyle\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n} a_k = \alpha$ となることを示せ。

(2) $\varphi(x)$ が任意の有界閉区間で積分可能で $\displaystyle\lim_{x\to\infty}\varphi(x) = 0$ を満たせば、$\displaystyle\lim_{n\to\infty}\frac{1}{n}\int_0^n \varphi(x)\,dx = 0$ となることを示せ。(ともに微積の問題です。)

上の定理 2.13 について、$X_1, X_2, \ldots$ が i.i.d. であれば $\dfrac{S_n}{n} - c_n$ が 0 に確率収束ような $c_n$ が存在するための必要条件でもあることが知られている (*cf*. Feller, W.: An Introduction to Probability Theory and Its Applications, vol.II, (1971) pp.234–6)。

定理 2.12 は他にも応用がある。興味のある人は昨年度の確率統計学 II の講義ノートを参照していただくとして、ここでは省略する。

`http://www.math.u-ryukyu.ac.jp/~sugiura/2020/prob2019b_text.pdf`

[11 月 11 日]

前回の定理 2.13 の応用として次の例を証明しよう。この例から必ずしも期待値を持たなくても大数の弱法則が成り立つことがわかる ($cf$. 注意 2.4)。

**例 2.15** $X_1, X_2, \dots$ は i.i.d. で、$P(X_1 = (-1)^j j) = \frac{K}{j^2 \log j}$, $j = 2, 3, \dots,$ を満たすとする。ただし $K = 1/\Big(\sum_{j=2}^{\infty} \frac{1}{j^2 \log j}\Big)$ とする。このとき、$E[|X_1|] = \infty$ であるが、ある定数 $c$ が存在して $\dfrac{S_n}{n}$ は $c$ に確率収束する (即ち大数の弱法則が成立する) ことを示せ。

**証明:** 期待値が存在しないことは $E[|X_1|] = \displaystyle\sum_{j=2}^{\infty} \frac{K}{j \log j} \geq \int_2^{\infty} \frac{K}{t \log t}\, dt = \infty$ よりわかる。次に

$$xP(|X_1| > x) = x \sum_{j=\lfloor x+1 \rfloor}^{\infty} \frac{K}{j^2 \log j} \leq x \int_x^{\infty} \frac{K}{t^2 \log t}\, dt$$

であるが、ロピタルの定理を用いればこの右辺 $\to 0$ がわかるので、定理 2.13 により $\dfrac{S_n}{n} - c_n \to 0$ in prob. を得る。ここで、$c_n = E[X_1 1_{\{|X_1| \leq n\}}] = K \displaystyle\sum_{j=2}^{n} \frac{(-1)^j}{j \log j}$ であるが、$\left\{\dfrac{1}{n \log n}\right\}$ は単調減少で 0 に収束するので、$c = K \displaystyle\sum_{j=2}^{\infty} \frac{(-1)^j}{j \log j}$ とおくとこの右辺は次の問題 2.9 より収束し、$c_n \to c$ となる。よって、$\dfrac{S_n}{n} \to c$ in prob. が示された。 □

**問題 2.9** $\{a_n\}$ が単調減少で 0 に収束するとき、交代級数 $\displaystyle\sum_{n=1}^{\infty} (-1)^{n-1} a_n$ は収束することを示せ。(微積の問題です。)

**注意 2.2** 定理 2.13 の仮定の下 ($X_1, X_2, \dots$ は i.i.d. とする)、$m \notin [a, b]$ なら $P(a \leq \frac{1}{n} \sum_{k=1}^{n} X_k \leq b)$ は 0 に収束する。もし、$E[e^{tX_1}] < \infty$ ($\forall t \in \mathbf{R}$) であれば、この収束は指数的に速く減衰する。その収束の速さを決定するのが Cramér の定理である。これを**大偏差原理** (large deviation principle) といい、Varadhan により整備され、応用例も多く盛んに研究されている ($cf$. 直接計算できる例として問題 2.11)。

次の 2 つの問題は Stirling の公式 $n! \sim \sqrt{2\pi} n^{n+\frac{1}{2}} e^{-n}$, ただし、$a_n \sim b_n$ とは $\displaystyle\lim_{n\to\infty} \frac{a_n}{b_n} = 1$ を表す。Stirling の公式の証明は昨年度の前期の講義ノート p.41 にあります。参照ください。

**問題 2.10** $0 < p < 1$ とする。$X_n$ が二項分布 $B(3n, p)$ に従うとする。$\displaystyle\lim_{n\to\infty} \frac{1}{n} \log P(X_n = n)$ を求めよ。

**問題 2.11** $X_1, X_2, \dots$ は i.i.d. で、$P(X_1 = 1) = P(X_1 = -1) = 1/2$ として $S_n = \sum_{k=1}^{n} X_k$ とする。このとき $0 < \forall a < 1$ に対して $\displaystyle\lim_{n\to\infty} \frac{1}{2n} \log P\Big(\frac{S_{2n}}{2n} \geq a\Big) = -\psi(a)$ となることを、以下に従って示せ。ただし $\psi(a) = \frac{1}{2}\{(1+a)\log(1+a) + (1-a)\log(1-a)\}$ とする。

(0) $\psi(a) > 0$ ($0 < a < 1$) を示せ。

(1) Stirling の公式 $n! \sim \sqrt{2\pi} n^{n+\frac{1}{2}} e^{-n}$ を用いて次を示せ。ただし $a_n \sim b_n$ とは $\displaystyle\lim_{n\to\infty} \frac{a_n}{b_n} = 1$ を表す。

$$P(S_{2n} = 2l) = \binom{2n}{n+l} \frac{1}{2^{2n}} \sim (\pi n)^{-\frac{1}{2}} \Big(1 + \frac{l}{n}\Big)^{-n-l-\frac{1}{2}} \Big(1 - \frac{l}{n}\Big)^{-n+l-\frac{1}{2}}, \ l = 1, \dots, n.$$

(2) $l_n \in \{1, \dots, n\}$ が $\displaystyle\lim_{n\to\infty} \frac{l_n}{n} = a$ を満たすとき、$\displaystyle\lim_{n\to\infty} \frac{1}{2n} \log P(S_{2n} = 2l_n) = -\psi(a)$ を示せ。

(3) $l_n = \lceil na \rceil$ のとき $P(S_{2n} = 2l_n) \leq P(S_{2n} \geq 2na) \leq nP(S_{2n} = 2l_n)$ を示し、証明を完成させよ。ただし $\lceil na \rceil$ は $na$ 以上の最小の整数を表す。

注意: これより $P\big(\frac{S_{2n}}{2n} \geq a\big)$ はおおよそ $e^{-2n\psi(a)}$ の速さで減衰することがわかる。

## 2.3　大数の強法則

**定理 2.16 (Kolmogorov の不等式)** $X_1, X_2, \ldots$ を独立な確率変数列で、$\forall n$ に対して $E[X_n] = 0$ かつ $V(X_n) < \infty$ とする。このとき、任意の $a > 0$ に対して

$$P\left(\max_{1\le k\le n}\left|\sum_{j=1}^{k} X_j\right| \ge a\right) \le \frac{1}{a^2}\sum_{j=1}^{n} V(X_j)$$

が成立する。

**証明:** $S_k = \sum_{j=1}^{k} X_j$ とし、評価したい事象を

$$A^* = \left\{\omega \in \Omega\,;\ \max_{1\le k\le n} |S_k(\omega)| \ge a\right\}$$

とおく。$S_k$ の $k$ を時刻のように考え、$|S_k|$ がはじめて $a$ 以上になる $k$ に着目して、$A^*$ を互いに排反な事象に分ける。すなわち、$k = 1, 2, \ldots, n$ に対して

$$A_k^* = \{\,\omega \in \Omega\,;\ j = 1, 2, \ldots, k-1 \text{ に対しては } |S_j(\omega)| < a \text{ で、かつ } |S_k(\omega)| \ge a\} \tag{2.6}$$

とおくと、$A^* = \bigcup_{k=1}^{n} A_k^*$ (互いに排反) となる。したがって、

$$P(A^*) = \sum_{k=1}^{n} P(A_k^*) = \sum_{k=1}^{n} E[1_{A_k^*}] \le \sum_{k=1}^{n} \frac{1}{a^2} E[S_k^2 \cdot 1_{A_k^*}]$$

となる。最後の不等号は $\omega \in A_k^*$ ならば $S_k(\omega)^2 \ge a^2$ となることを用いた。ここで、

$$S_n^2 = (S_k + (S_n - S_k))^2 = S_k^2 + 2S_k(S_n - S_k) + (S_n - S_k)^2 \ge S_k^2 + 2S_k(S_n - S_k)$$

に注意すると、

$$E[S_n^2 \cdot 1_{A_k^*}] - E[S_k^2 \cdot 1_{A_k^*}] \ge 2E[S_k(S_n - S_k) \cdot 1_{A_k^*}]$$

ここで、(2.6) より事象 $A_k^*$ は $X_1, \cdots, X_k$ のみによって決まっており、一方 $S_n - S_k = \sum_{j=k+1}^{n} X_j$ なので、$\{X_n\}$ は独立だから $S_k \cdot 1_{A_k^*}$ と $S_n - S_k$ は独立となる。したがって、

$$E[S_k(S_n - S_k) \cdot 1_{A_k^*}] = E[S_k \cdot 1_{A_k^*}]E[S_n - S_k] = E[S_k \cdot 1_{A_k^*}] \sum_{j=k+1}^{n} E[X_j] = 0.$$

以上より、

$$\begin{aligned} P(A^*) &\le \sum_{k=1}^{n} \frac{1}{a^2} E[S_k^2 \cdot 1_{A_k^*}] \le \sum_{k=1}^{n} \frac{1}{a^2} E[S_n^2 \cdot 1_{A_k^*}] = \frac{1}{a^2} E[S_n^2 \cdot 1_{A^*}] \le \frac{1}{a^2} E[S_n^2] \\ &= \frac{1}{a^2} V(S_n) = \frac{1}{a^2} V(X_1 + X_2 + \cdots + X_n) = \frac{1}{a^2} \sum_{j=1}^{n} V(X_j) \end{aligned}$$

最後の等号では再び $X_1, \ldots, X_n$ が独立であることを用いた。　□

**定理 2.17 (Kolmogorov の第 1 定理)** $X_1, X_2, \ldots$ が独立な確率変数列で、

$$\sum_{n=1}^{\infty} \frac{1}{n^2} V(X_n) < \infty \tag{2.7}$$

を満たせば、大数の強法則が成立、すなわち、$\displaystyle \frac{1}{n}\sum_{j=1}^{n}(X_j - E[X_j])$ は 0 に概収束する。

**証明:** $\forall n \in \mathbf{N}$ に対して $E[X_n] = 0$ と仮定してよい。実際、$X_n - E[X_n]$ を $X_n$ とみなせばよい。$Y_n = \frac{1}{n}\sum_{j=1}^n X_j = \frac{1}{n}S_n$ と書くこととする。

<u>1st step</u> $\forall \varepsilon > 0$ に対して、

$$A(\varepsilon) = \bigcup_{N=1}^{\infty}\bigcap_{n=N}^{\infty}\{|Y_n| < \varepsilon\}$$

とおき、

$$P(A(\varepsilon)) = 1 \tag{2.8}$$

が示されれば、定理の主張が示される。実際、$A = \bigcap_{j=1}^{\infty} A(1/j)$ とおけば、(2.8) より各 $j = 1, 2, \ldots$ について $P(A(1/j)) = 1$ だから、$P(A) = 1$. ここで、$\omega \in A$ とすると、任意の $j \in \mathbf{N}$ に対して $\omega \in A(1/j)$ だから $N = N(\omega, j)$ が存在して $n \geq N$ ならば $|Y_n(\omega)| < 1/j$ である。したがって、$\omega \in A$ ならば $\lim_{n\to\infty} Y_n(\omega) = 0$ となり、証明は完了する。

<u>2nd step</u> (2.8) を示す。そのために

$$B_m(\varepsilon) = \bigcup_{n=2^{m-1}}^{2^m-1}\{|Y_n| \geq \varepsilon\} = \left\{\max_{2^{m-1}\leq n<2^m}|Y_n| \geq \varepsilon\right\}$$

とおく。このとき、$\forall l \in \mathbf{N}$ に対して

$$A(\varepsilon)^{\mathrm{c}} = \bigcap_{N=1}^{\infty}\bigcup_{n=N}^{\infty}\{|Y_n| \geq \varepsilon\} \subset \bigcup_{m=l}^{\infty} B_m(\varepsilon) \tag{2.9}$$

だから、(2.8), すなわち $P(A(\varepsilon)^{\mathrm{c}}) = 0$ を示すためには

$$\sum_{m=1}^{\infty} P(B_m(\varepsilon)) < \infty \tag{2.10}$$

を示せばよい。実際、Borel-Cantelli の定理により $P\left(\bigcap_{l=1}^{\infty}\bigcup_{m=l}^{\infty} B_m(\varepsilon)\right) = 0$ であるが、(2.9) より $A(\varepsilon)^{\mathrm{c}} \subset \bigcap_{l=1}^{\infty}\bigcup_{m=l}^{\infty} B_m(\varepsilon)$ となるから従う。

<u>3rd step</u> (2.10) を示すため、$S_n = \sum_{j=1}^n X_j (= nY_n)$ として、

$$\begin{aligned}
P(B_m(\varepsilon)) &= P\left(\max_{2^{m-1}\leq k<2^m}\frac{1}{k}|S_k| \geq \varepsilon\right) \leq P\left(\max_{2^{m-1}\leq k<2^m}|S_k| \geq \varepsilon 2^{m-1}\right) \\
&\leq P\left(\max_{1\leq k\leq 2^m}|S_k| \geq \varepsilon 2^{m-1}\right) \leq \frac{1}{\varepsilon^2 2^{2m-2}}\sum_{k=1}^{2^m} V(X_k)
\end{aligned}$$

ただし 1 行目の不等号では $2^{m-1} \leq k$ を、最後の不等号は Kolmogorov の不等式 (定理 2.6) を用いた。したがって、

$$\begin{aligned}
\sum_{m=1}^{\infty} P(B_m(\varepsilon)) &\leq \frac{4}{\varepsilon^2}\sum_{m=1}^{\infty}\frac{1}{2^{2m}}\sum_{k=1}^{2^m} V(X_k) = \frac{4}{\varepsilon^2}\sum_{m=1}^{\infty}\frac{1}{2^{2m}}\sum_{k=1}^{\infty} 1_{[1,2^m]}(k)V(X_k) \\
&= \frac{4}{\varepsilon^2}\sum_{k=1}^{\infty} V(X_k)\sum_{m=1}^{\infty} 1_{[k,\infty)}(2^m)\frac{1}{2^{2m}} = \frac{4}{\varepsilon^2}\sum_{k=1}^{\infty} V(X_k)\sum_{m=m_k}^{\infty}\frac{1}{2^{2m}} \leq \frac{16}{3\varepsilon^2}\sum_{k=1}^{\infty} V(X_k)\frac{1}{k^2}
\end{aligned}$$

ただし $m_k = \lceil \log_2 k \rceil$ とする ($\lceil a \rceil$ は $a$ 以上の最小の整数を表す)。このとき、$2^{m_k-1} < k \leq 2^{m_k}$ であるから 2 行目の不等号は

$$\sum_{m=m_k}^{\infty}\frac{1}{2^{2m}} = \frac{1/2^{2m_k}}{1-1/4} = \frac{4}{3}\frac{1}{(2^{m_k})^2} \leq \frac{4}{3}\frac{1}{k^2}$$

となることを用いた。よって、仮定 (2.7) より (2.10) が示された。 □

[11 月 18 日]

$\{X_n\}$ の分布が同じならば、定理 2.17 の仮定 (2.7)、特に $E[X_n^2] < \infty$ は不要になる。

**定理 2.18 (Kolmogorov の第 2 定理)** $X_1, X_2, \ldots$ は i.i.d. で、$E[|X_1|] < \infty$ とする。このとき、大数の強法則が成立、すなわち、$\displaystyle\frac{1}{n}\sum_{j=1}^{n} X_j$ は $E[X_1]$ に概収束する。

**証明:** $E[X_1] = 0$ と仮定してよい。$X$ を $X_n$ と共通の分布をもつ確率変数とする。

<u>1st step</u> (番号 $k$ に依存した cut-off の導入) $Z_k = X_k 1_{(0,k]}(|X_k|) - \tilde{m}_k$, $\tilde{m}_k = E[X_k 1_{(0,k]}(|X_k|)]$ とおくと、$\{Z_k\}$ は定理 2.17 の仮定を満たす。実際、$\{Z_k\}$ は独立であり、

$$
\begin{aligned}
\sum_{k=1}^{\infty} \frac{1}{k^2} V(Z_k) &= \sum_{k=1}^{\infty} \frac{1}{k^2} (E[\{X_k 1_{(0,k]}(|X_k|)\}^2] - \tilde{m}_k^2) \le \sum_{k=1}^{\infty} \frac{1}{k^2} E[X_k^2 1_{(0,k]}(|X_k|)] \\
&= \sum_{k=1}^{\infty} \frac{1}{k^2} \sum_{j=1}^{k} E[X^2 1_{(j-1,j]}(|X|)] = \sum_{j=1}^{\infty} E[X^2 1_{(j-1,j]}(|X|)] \sum_{k=j}^{\infty} \frac{1}{k^2} \\
&\le E[X^2 1_{(0,1]}(|X|)] \sum_{k=1}^{\infty} \frac{1}{k^2} + \sum_{j=2}^{\infty} E[X^2 1_{(j-1,j]}(|X|)] \frac{1}{j-1} \\
&\le \sum_{k=1}^{\infty} \frac{1}{k^2} + \sum_{j=2}^{\infty} 2E[|X| 1_{(j-1,j]}(|X|)] \le \sum_{k=1}^{\infty} \frac{1}{k^2} + 2E[|X|] < \infty
\end{aligned}
$$

となる。ここで 3 行目の不等号は

$$
\sum_{k=j}^{\infty} \frac{1}{k^2} \le \int_{j-1}^{\infty} \frac{dx}{x^2} = \frac{1}{j-1},
$$

4 行目の最初の不等号は $j \ge 2$ のとき $j-1 < |x| \le j$ であれば

$$
x^2 \frac{1}{j-1} = |x| \frac{|x|}{j-1} \le |x| \frac{j}{j-1} \le 2|x|
$$

となることを用いた。したがって、$E[Z_k] = 0$ だから定理 2.17 から

$$
\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} Z_k = 0 \quad \text{a.s.}
$$

が示された。

<u>2nd step</u> $|X 1_{(0,k]}(X)| \le |X|$ $(\forall k \in \mathbf{N})$ で $E[|X|] < \infty$ なので、Lebesgue の収束定理 (定理 2.10) により

$$
\tilde{m}_k = E[X 1_{(0,k]}(|X|)] \to E[X] = 0 \quad (k \to \infty)
$$

がわかる。したがって、$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \tilde{m}_k = 0$ となり ($cf.$ 問題 2.8(1))、1st step により

$$
\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} X_k 1_{(0,k]}(|X_k|) = 0 \quad \text{a.s.}
$$

<u>3rd step</u> $P(\#\{k \in \mathbf{N}; |X_k| > k\} < \infty) = 1$ を示す。これがいえれば、a.a. $\omega$ に対して有限個の $k$ を除いて $X_k = X_k 1_{(0,k]}(|X_k|)$ だから 2nd step から結論が得られる。そこで、まず

$$
\begin{aligned}
\sum_{k=1}^{\infty} P(|X_k| > k) &= \sum_{k=1}^{\infty} \sum_{j=k}^{\infty} P(j < |X| \le j+1) = \sum_{j=1}^{\infty} \sum_{k=1}^{j} P(j < |X| \le j+1) \\
&= \sum_{j=1}^{\infty} jP(j < |X| \le j+1) = E\Big[\sum_{j=1}^{\infty} j 1_{(j,j+1]}(|X|)\Big] \le E[|X|] < \infty
\end{aligned}
$$

に注意する。ただし、2 行目の一つ目の不等号は $\sum_{j=1}^{\infty} j1_{(j,j+1]}(|x|) \le \sum_{j=1}^{\infty} |x|1_{(j,j+1]}(|x|) \le |x|$ となることを用いた。よって、Borel-Cantelli の定理から $P(\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty}\{|X_k|>k\}) = 0$ であるが、$\{\omega\in\Omega\,;\#\{k\in \mathbf{N};|X_k(\omega)|>k\}<\infty\} \supset \bigcup_{n=1}^{\infty}\bigcap_{k=n}^{\infty}\{|X_k|\le k\} = (\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty}\{|X_k|>k\})^{\mathrm{c}}$ となり主張は示された。 □

**注意 2.3** 定理 2.18 は $X_1, X_2, \ldots$ が組ごとに独立であれば成立することが知られている。(*cf*. Durrett, R.: Probablity Theory and Examples, 4th ed. pp.73–75.) ここでは、*Kolmogorov* の不等式などマルチンゲール理論につながる考え方があるためこの証明法を用いた。

**定理 2.19** $X_1, X_2, \ldots$ は i.i.d. で $E[|X_1|] = \infty$ となるとき、

$$P\left(\limsup_{n\to\infty}\Big|\frac{1}{n}\sum_{k=1}^{n}X_k\Big| = \infty\right) = 1. \tag{2.11}$$

**定理 2.20 (Borel-Cantelli の第 2 定理)** $\{B_n\}$ が独立で $\sum\limits_{n=1}^{\infty}P(B_n)=\infty$ ならば $P(\bigcap\limits_{n=1}^{\infty}\bigcup\limits_{k=n}^{\infty}B_k) = 1.$

**証明:** $(\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty}B_k)^c = \bigcup_{n=1}^{\infty}\bigcap_{k=n}^{\infty}B_k^{\mathrm{c}}$ で $\{\bigcap_{k=n}^{\infty}B_k^{\mathrm{c}}\}$ は $n$ について単調増加、また、$\{\bigcap_{k=n}^{N}B_k^{\mathrm{c}}\}$ は $N$ について単調減少で $\bigcap_{k=n}^{\infty}B_k^{\mathrm{c}} = \bigcap_{N=n}^{\infty}\bigcap_{k=n}^{N}B_k^{\mathrm{c}}$ なので、

$$P((\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty}B_k)^c) = \lim_{n\to\infty}P(\bigcap_{k=n}^{\infty}B_k^{\mathrm{c}}) = \lim_{n\to\infty}\lim_{N\to\infty}P(\bigcap_{k=n}^{N}B_k^{\mathrm{c}}). \tag{2.12}$$

次に、仮定と定理 1.7 (2) により $B_n^{\mathrm{c}}, \ldots, B_N^{\mathrm{c}}$ は独立であることと、$P(B_k^{\mathrm{c}}) = 1-P(B_k) \le e^{-P(B_k)}$ となること[*4]を用いて、

$$0 \le P(\bigcap_{k=n}^{N}B_k^{\mathrm{c}}) = \prod_{k=n}^{N}P(B_k^{\mathrm{c}}) \le \prod_{k=n}^{N}e^{-P(B_k)} = e^{-\sum_{k=n}^{N}P(B_k)}.$$

ここで、$\sum_{k=n}^{\infty}P(B_k)=\infty$ だから、(右辺)$\to 0$ $(N\to\infty)$. 以上より、(2.12) より $P((\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty}B_k)^{\mathrm{c}}) = 0$ となるから主張を得る。 □

**定理 2.19 の証明:** $X$ を $X_n$ と同じ分布をもつ確率変数とする。

<u>1st step</u> $M>0$ とし、$B_n^M = \{|X_n|\ge Mn\}$ とすると、

$$\begin{aligned}
&\sum_{n=0}^{\infty}P(B_n^M) = \sum_{n=0}^{\infty}P(|X|\ge Mn) = \sum_{n=0}^{\infty}\sum_{k=n}^{\infty}P(Mk\le|X|<M(k+1)) \\
&= \sum_{k=0}^{\infty}\sum_{n=0}^{k}P(Mk\le|X|<M(k+1)) = \sum_{k=0}^{\infty}(k+1)P\Big(k\le\frac{|X|}{M}<k+1\Big) = \sum_{k=0}^{\infty}E\Big[(k+1)1_{\{k\le\frac{|X|}{M}<k+1\}}\Big] \\
&\ge \sum_{k=0}^{\infty}E\Big[\frac{|X|}{M}1_{\{k\le\frac{|X|}{M}<k+1\}}\Big] = E\Big[\frac{|X|}{M}\Big] = \frac{E[|X|]}{M} = \infty.
\end{aligned}$$

ここで、2 行目の二つ目の等号は $P(Mk\le|X|<M(k+1))$ が $n$ によらないことを用いた。

<u>2nd step</u> $\{B_n^M\}$ は独立な事象の列だから Borel-Cantelli の第 2 定理により $P\left(\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty}B_k^M\right) = 1$ であり、よって $P\left(\bigcap_{M=1}^{\infty}\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty}B_k^M\right) = 1.$ $\omega\in\bigcap_{M=1}^{\infty}\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty}B_k$ とすると、$\forall M,n\in\mathbf{N}$ に対してある $k\ge n$ があって $\omega\in B_k^M$、i.e., $\frac{|X_k(\omega)|}{k}\ge M$ となるから、

$$\limsup_{n\to\infty}\frac{|X_n(\omega)|}{n}\ge M.$$

これが、$\forall M\in\mathbf{N}$ に対して成り立つので、$\limsup_{n\to\infty}\frac{|X_n(\omega)|}{n}=\infty$.

[*4] $0\le x\le 1$ のとき $0\le 1-x\le e^{-x}$ を用いた。

次に、数列 $\{a_n\}$ に対して

$$\limsup_{n\to\infty}\Big|\frac{1}{n}\sum_{k=1}^{\infty}a_k\Big|<\infty \qquad \Longrightarrow \qquad \limsup_{n\to\infty}\frac{|a_n|}{n}<\infty$$

となることに注意する。これは、

$$|a_n|=|(a_1+\cdots+a_{n-1}+a_n)-(a_1+\cdots+a_{n-1})|\le|a_1+\cdots+a_{n-1}+a_n|+|a_1+\cdots+a_{n-1}|$$

となることからすぐにわかる。この対偶を $a_n=X_n(\omega)$ に対して用いると、

$$\bigcap_{M=1}^{\infty}\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty}B_k^M\subset\Big\{\limsup_{n\to\infty}\frac{|X_n|}{n}=\infty\Big\}\subset\Big\{\limsup_{n\to\infty}\Big|\frac{1}{n}\sum_{k=1}^{n}X_k\Big|=\infty\Big\}$$

となり、(2.11) が成り立つことがわかった。 □

**注意 2.4** 例 2.15 の例は大数の弱法則を満たすが、$E[|X_1|]=\infty$ となるため定理 2.19 より大数の強法則を満たさない。

# 3 特性関数と中心極限定理

特性関数の理論は Fourier 変換の確率変数版といえる。Fourier 変換はいたるところで応用されている。例えば CT スキャンはこの理論の上に成り立っている。この授業では時間の都合で扱えないが、解析学 IV は Fourier 級数や Fourier 変換を扱うのでそこでぜひ受講して勉強しておいてほしい。

## 3.1 特性関数

**定義 3.1** (1) 複素数値関数 $Z$ が可測 (複素数値確率変数) であるとは、その実部 $X=\operatorname{Re}Z$, 虚部 $Y=\operatorname{Im}Z$ がともに可測 (確率変数) であるときにいう。ここで、$Z=X+iY$, $i=\sqrt{-1}$ である。以下、単に確率変数といえば、実数値確率変数を表すものとする。
(2) 複素数値確率変数 $Z$ に対して、$E[|\operatorname{Re}Z|]<\infty$ かつ $E[|\operatorname{Im}Z|]<\infty$ のとき、$Z$ の期待値を

$$E[Z]=E[\operatorname{Re}Z]+iE[\operatorname{Im}Z]$$

と定める。

**命題 3.1** 複素数値確率変数 $Z$ に対して、$|E[Z]|\le E[|Z|]$ が成立する。

**証明:** $E[|Z|]<\infty$ のとき、$|\operatorname{Re}Z|\le|Z|$, $|\operatorname{Im}Z|\le|Z|$ より $E[Z]$ が定義されることに注意する。$\alpha=E[Z]$, $\tilde{Z}=\dfrac{Z}{|Z|}1_{\{Z\neq0\}}$ とする。このとき、

$$|E[Z]|=|\alpha|=\frac{\overline{\alpha}}{|\alpha|}\alpha=E\Big[\frac{\overline{\alpha}}{|\alpha|}Z\Big]=E\Big[|Z|\frac{\overline{\alpha}}{|\alpha|}\tilde{Z}\Big]=E\Big[|Z|\operatorname{Re}\Big(\frac{\overline{\alpha}}{|\alpha|}\tilde{Z}\Big)\Big]+iE\Big[|Z|\operatorname{Im}\Big(\frac{\overline{\alpha}}{|\alpha|}\tilde{Z}\Big)\Big]$$

であるが、左辺は実数なので、(右辺の虚部)$=0$ となる。ここで、$\operatorname{Re}\Big(\frac{\overline{\alpha}}{|\alpha|}\tilde{Z}\Big)\le\Big|\frac{\overline{\alpha}}{|\alpha|}\tilde{Z}\Big|\le1$ であるから、(右辺の実部)$\le E[|Z|]$ となり、主張を得る。 □

**定義 3.2** 確率変数 $X$ に対して、次の関数 $\phi_X(t)$ を $X$ の特性関数 (characteristic function) という。

$$\phi_X(t)=E[e^{itX}]=E[\cos(tX)]+i\,E[\sin(tX)], \qquad t\in\mathbf{R}.$$

**命題 3.2** (i) 任意の確率変数 $X$ の特性関数はつねに存在する。
(ii) すべての実数 $t$ に対して、$|\phi_X(t)| \le 1$ である。
(iii) $\phi_X(0) = 1$ かつ $\phi_X(t) = \overline{\phi_X(-t)}$ である。
(iv) $t$ の関数として、$\phi_X(t)$ は一様連続である。

**証明:** (i), (ii) $|e^{itX}|^2 = |\cos tX + i\sin tX|^2 = \cos^2 tX + \sin^2 tX = 1$ と命題 3.1 より明らか。
(iii) $\phi_X(0) = E[e^0] = E[1] = 1$, $\overline{\phi_X(-t)} = \overline{E[e^{-itX}]} = E[\overline{e^{-itX}}] = E[e^{itX}] = \phi_X(t)$
(iv) 0 に収束する任意の数列 $\{h_n\}$ に対し $\sup_{s\in\mathbf{R}} |\phi_X(s+h_n) - \phi_X(s)| \to 0$ を示せばよい。ここで、(右辺)$\le \sup_s E[|e^{isX}(e^{ih_nX}-1)|] = E[|e^{ih_nX}-1|]$. よって $|e^{ih_nX}-1| \le |e^{ih_nX}|+1 = 2$ で $E[2] = 2 < \infty$ であるから、Lebesgue の収束定理 (定理 2.10) により $E[|e^{ih_nX}-1|] \to E[|e^0-1|] = 0$ となり主張を得る。 □

**命題 3.3** 確率変数 $X$ と定数 $a, b$ に対して $\phi_{aX+b}(t) = e^{itb}\phi_X(at)$.

**証明:** $\phi_{aX+b}(t) = E[e^{iatX}e^{itb}] = e^{itb}E[e^{iatX}] = e^{itb}\phi_X(at)$. □

**例 3.4** (1) $X$ が二項分布 $B(n,p)$, $0<p<1$ に従うとき、

$$\phi_X(t) = E[e^{itX}] = \sum_{k=0}^{n} e^{itk}\binom{n}{k}p^k(1-p)^{n-k} = \sum_{k=0}^{n}\binom{n}{k}(pe^{it})^k(1-p)^{n-k} = (e^{it}p+1-p)^n.$$

(2) $X$ が Poisson 分布 $P(\lambda)$, $\lambda>0$, に従うとき

$$\phi_X(t) = E[e^{itX}] = \sum_{k=0}^{\infty} e^{itk}\frac{\lambda^k}{k!}e^{-\lambda} = \sum_{k=0}^{\infty}\frac{(e^{it}\lambda)^k}{k!}e^{-\lambda} = e^{e^{it}\lambda}e^{-\lambda} = e^{\lambda(e^{it}-1)}.$$

**問題 3.1** 確率変数 $X$ の密度関数が $f_X(x)$ であるとき、その特性関数 $\phi_X(t)$ は

$$\phi_X(t) = E[e^{itX}] = E[\cos(tX)] + i\,E[\sin(tX)] = \int_{-\infty}^{\infty} f_X(x)\cos(tx)\,dx + i\int_{-\infty}^{\infty} f_X(x)\sin(tx)\,dx$$

となる。この右辺の 2 つの積分を計算することで、次の確率変数 $X$ に対して特性関数 $\phi_X(t)$ を求めよ。
(1) $X$ の密度関数が $f_X(x) = (1-|x|)1_{(-1,1)}(x)$ で与えられるとき。
(2) $X$ の密度関数が $f_X(x) = \dfrac{1}{2}e^{-|x|}$ で与えられるとき。
(3) $X$ が指数分布 $Ex(\lambda)$ $(\lambda>0)$ に従うとき。

**問題 3.2** (1) $X$ が負の二項分布 $NB(\alpha,p)$, $\alpha>0$, $0<p<1$, に従うとき、$X$ の特性関数 $\phi_X(t)$ を求めよ。
(2) $X$ が二項分布 $B\Big(2n, \dfrac{1}{2}\Big)$ に従うとし $Y = X - n$ とおく。$Y$ の特性関数 $\phi_Y(t)$ を $\cos t$ を用いて表せ。

[11 月 25 日]

**命題 3.5** 確率変数 $X$ が $E[|X|^k] < \infty$ を満たせば、その特性関数 $\phi_X(t)$ は $C^k$-級で $\phi_X^{(k)}(t) = E[(iX)^k e^{itX}]$ となる。

**証明:** (帰納法を用いる) $k = 0$ のとき、$\phi_X(t)$ が $t$ について連続であることは命題 3.2 で示した。
$k$ のとき成り立つと仮定する。$k+1$ のとき、まず任意の 0 に収束する数列 $\{h_n\}$ に対し

$$\lim_{n\to\infty} \frac{\phi_X^{(k)}(t+h_n) - \phi_X^{(k)}(t)}{h_n} = \lim_{n\to\infty} E\Big[(iX)^k \frac{e^{i(t+h_n)X} - e^{itX}}{h_n}\Big] = E[(iX)^{k+1} e^{itX}] \tag{3.1}$$

を示す。$|e^{i(t+h)x} - e^{itx}| = \Big|\int_t^{t+h} ixe^{ix\theta}\, d\theta\Big| \le \Big|\int_t^{t+h} |ixe^{ix\theta}|\, d\theta\Big| \le |hx|$ より、

$$\Big|(iX)^k \frac{e^{i(t+h)X} - e^{itX}}{h}\Big| \le |X|^{k+1} \qquad \text{と仮定} \qquad E[|X|^{k+1}] < \infty$$

から、Lebesgue の収束定理 (定理 2.10) と $\displaystyle\lim_{h\to 0} \frac{e^{i(t+h_n)X} - e^{itX}}{h_n} = iXe^{itX}$ となることから (3.1) は成立する。
$\phi^{(k+1)}(t)$ が $t$ について連続であることは命題 3.2(iv) と同様に示すことができる。 □

**例 3.6** $X$ が標準正規分布 $N(0,1)$ に従うとき、その特性関数は $\phi_X(t) = e^{-\frac{1}{2}t^2}$ となる。

**問題 3.3** 以下に従って例 3.6 を証明せよ。
(1) $E[|X|] < \infty$ に注意して命題 3.5 を用いると $\phi_X(t) = E[e^{itX}]$ は $C^1$-級で $\phi_X'(t) = E[iXe^{itX}]$ となる。これを用いて、$\phi_X(t) = \displaystyle\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \cos(tx)\, dx$ と $\phi_X'(t) = \displaystyle\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} (-x)\sin(tx)\, dx$ を導け。
(2) $\phi_X'(t) = -t\phi_X(t)$ を導け。ヒント: 部分積分を用いよ。
(3) (2) の微分方程式を $\phi_X(0) = 1$ に注意して解くことにより、$\phi_X(t) = e^{-\frac{1}{2}t^2}$ を示せ。

**系 3.7** $X$ が正規分布 $N(m, \sigma^2)$ に従うときその特性関数は $\phi_X(t) = e^{imt - \frac{1}{2}\sigma^2 t^2}$ となる。

**証明:** $Z = \frac{X-m}{\sigma}$ とすると、$Z$ は標準正規分布に従う。よって、$X = \sigma Z + m$ に命題 3.3 を適用して

$$\phi_X(t) = \phi_{\sigma Z + m}(t) = e^{imt}\phi_Z(\sigma t) = e^{imt} e^{-\frac{1}{2}(\sigma t)^2} = e^{imt - \frac{1}{2}\sigma^2 t^2} \qquad \square$$

例 3.6 は解析学 I で学んだ Cauchy の積分定理を用いても導出できる。詳しくは昨年度の確率統計学 II の講義ノート pp.13-14 を参照ください。次の例も解析学で学んだ留数定理を用いて証明できる ($cf.$ 昨年度の講義ノート pp.14-15) が後に定理 3.21 の応用として別証明を与えることとして、ここでは省略する。

**例 3.8** $X$ が Cauchy 分布に従う、すなわち、その密度関数が $f(x) = \dfrac{1}{\pi}\dfrac{1}{1+x^2}$ $(-\infty < x < \infty)$ のとき、その特性関数は $\phi_X(t) = e^{-|t|}$ となる。

**注意 3.1** Cauchy 分布の特性関数は $t = 0$ で微分可能ではない。実際、Cauchy 分布は平均を持たない ($cf.$ 命題 3.5)。

**定義 3.3** $d$ 次元確率変数 $\boldsymbol{X} = (X_1, \ldots, X_d)'$ (縦ベクトル, $A'$ は $A$ の転置を表す) に対して、次の $\mathbf{R}^d$ 上の関数 $\phi_{\boldsymbol{X}}(\boldsymbol{t})$ を $\boldsymbol{X}$ の特性関数という。

$$\phi_{\boldsymbol{X}}(\boldsymbol{t}) = E[e^{i\boldsymbol{t}'\boldsymbol{X}}] = E[\exp\Big\{i\sum_{j=1}^d t_j X_j\Big\}], \qquad \boldsymbol{t} = (t_1, \ldots, t_d)' \in \mathbf{R}^d$$

**命題 3.9** $d$ 次元確率変数 $\boldsymbol{X}$ と $d$ 次正方行列 $A$ と $d$ 次元ベクトル $\boldsymbol{b}$ に対して $\phi_{A\boldsymbol{X}+\boldsymbol{b}}(\boldsymbol{t}) = e^{i\boldsymbol{t}'\boldsymbol{b}}\phi_{\boldsymbol{X}}(A'\boldsymbol{t})$.

**証明:** $\phi_{A\boldsymbol{X}+\boldsymbol{b}}(\boldsymbol{t}) = E[e^{i\boldsymbol{t}'A\boldsymbol{X}}e^{i\boldsymbol{t}'\boldsymbol{b}}] = e^{i\boldsymbol{t}'\boldsymbol{b}}E[e^{i(A'\boldsymbol{t})'\boldsymbol{X}}] = e^{i\boldsymbol{t}'\boldsymbol{b}}\phi_{\boldsymbol{X}}(A'\boldsymbol{t})$. □

**例 3.10** $\boldsymbol{X} = (X_1,\dots,X_d)'$ を $d$ 次元正規分布 $N(\boldsymbol{m},\boldsymbol{\Sigma})$ に従うとする。$\boldsymbol{m} = (m_1,\dots,m_d)' \in \mathbf{R}^d$, $\boldsymbol{\Sigma} = (\sigma_{ij})$ は正定値対称行列であった。このとき、$\phi_{\boldsymbol{X}}(\boldsymbol{t}) = e^{i\boldsymbol{t}'\boldsymbol{m}-\frac{1}{2}\boldsymbol{t}'\boldsymbol{\Sigma}\boldsymbol{t}}$ となる。

**証明:** 例 1.22 と同様に直交行列 $P = (p_{ij})$ と対角成分がすべて正の対角行列 $D = (\lambda_{ij})$ を $P'\boldsymbol{\Sigma}P = D$ なるようにとり、$\boldsymbol{Y} = (Y_1,\dots,Y_d)' = P'(\boldsymbol{X}-\boldsymbol{m})$ とすると、$Y_1,\dots,Y_d$ は独立で各 $Y_j$ は正規分布 $N(0,\lambda_{jj})$ に従う。よって、

$$\begin{aligned}\phi_{\boldsymbol{Y}}(\boldsymbol{t}) &= E[e^{it_1Y_1}\cdots e^{it_dY_d}] = E[e^{it_1Y_1}]\cdots E[e^{it_dY_d}] \\ &= e^{-\frac{1}{2}\lambda_{11}t_1^2}\cdots e^{-\frac{1}{2}\lambda_{dd}t_d^2} = e^{-\frac{1}{2}\boldsymbol{t}'D\boldsymbol{t}}.\end{aligned}$$

従って、$\boldsymbol{X} = P\boldsymbol{Y}+\boldsymbol{m}$ に命題 3.9 を適用して

$$\phi_{\boldsymbol{X}}(\boldsymbol{t}) = e^{i\boldsymbol{t}'\boldsymbol{m}}\phi_{\boldsymbol{Y}}(P'\boldsymbol{t}) = e^{i\boldsymbol{t}'\boldsymbol{m}}e^{-\frac{1}{2}(P'\boldsymbol{t})'D(P'\boldsymbol{t})} = e^{i\boldsymbol{t}'\boldsymbol{m}}e^{-\frac{1}{2}\boldsymbol{t}'PDP'\boldsymbol{t}} = e^{i\boldsymbol{t}'\boldsymbol{m}-\frac{1}{2}\boldsymbol{t}'\boldsymbol{\Sigma}\boldsymbol{t}}. \quad \square$$

命題 3.5 と同様に次が成立する。証明は同様なので省略する。

**命題 3.11** $d$ 次元確率変数 $\boldsymbol{X} = (X_1,\dots,X_d)'$ が $E[|\boldsymbol{X}|^k] < \infty$ を満たせば、その特性関数 $\phi_{\boldsymbol{X}}(\boldsymbol{t})$ は $C^k$-級である。特に、$k=2$ であれば $\dfrac{\partial^2}{\partial t_k \partial t_l}\phi_{\boldsymbol{X}}(\boldsymbol{t}) = i^2E[X_kX_le^{i\boldsymbol{t}'\boldsymbol{X}}]$, $1 \le k,l \le d$ となる。

## 3.2 分布と Dynkin 族定理

$\sigma$-集合族に関連して、Dynkin 族の概念を導入する。

**定義 3.4** (1) 集合 $S$ の部分集合族 $\mathcal{P}$ が $\pi$ 族であるとは、

(a) $S \in \mathcal{P}$, (b) $A, B \in \mathcal{P} \Longrightarrow A \cap B \in \mathcal{P}$

の 2 条件を満たすときにいう。

(2) 集合 $S$ の部分集合族 $\mathcal{D}$ が Dynkin 族であるとは、

(a) $S \in \mathcal{D}$

(b) $A, B \in \mathcal{D}$ で $A \supset B \Longrightarrow A \backslash B \in \mathcal{D}$

(c) $A_n \in \mathcal{D}$, $A_n \subset A_{n+1}$ $(\forall n \in \mathbf{N}) \Longrightarrow \bigcup_{n=1}^\infty A_n \in \mathcal{D}$

の 3 条件を満たすときにいう。

集合 $S$ の部分集合族 $\mathcal{C}$ に対して、$\mathcal{C}$ を含む最小の Dynkin 族を $\mathcal{L}(\mathcal{C})$ と表す。$\mathcal{D}_\lambda$ $(\lambda \in \Lambda)$ が Dynkin 族であれば $\bigcap_{\lambda\in\Lambda}\mathcal{D}_\lambda$ も Dynkin 族となる ($cf.$ 問題 3.4) ことから、$\{\mathcal{D}_\lambda\}_{\lambda\in\Lambda}$ をすべての $\mathcal{C}$ を含む Dynkin 族とし、$\mathcal{L}(\mathcal{C}) = \bigcap_{\lambda\in\Lambda}\mathcal{D}_\lambda$ とすればよい。実際、最小性は $\bigcap_{\lambda\in\Lambda}\mathcal{D}_\lambda \subset \mathcal{D}_\lambda$ $(\forall\lambda \in \Lambda)$ と、この左辺が $\mathcal{C}$ を含む最小の Dynkin 族であるから、ある $\lambda_0 \in \Lambda$ があって $\mathcal{D}_{\lambda_0} = \bigcap_{\lambda\in\Lambda}\mathcal{D}_\lambda$ となることからわかる。

**問題 3.4** (1) 各 $\lambda \in \Lambda$ に対して $\mathcal{D}_\lambda$ が $S$ 上の Dynkin 族のとき、$\cap_{\lambda\in\Lambda}\mathcal{D}_\lambda$ は $S$ 上の Dynkin 族をなすことを示せ。

(2) $\mathcal{D}$ が $S$ 上の $\sigma$-集合族であれば、$S$ 上の Dynkin 族となることを示せ。

**定理 3.12 (Dynkin 族定理)** $\mathcal{P}$ が $\pi$ 族のとき $\mathcal{L}(\mathcal{P}) = \sigma(\mathcal{P})$ となる[*5]。

[*5] $\sigma(\mathcal{P})$ は $\mathcal{P}$ を含む最小の $\sigma$-集合族であった。

**証明:** $\sigma$-集合族は Dynkin 族となる ($cf.$ 問題 3.4) ので、その最小性により $\mathcal{L}(\mathcal{P}) \subset \sigma(\mathcal{P})$ となる。
$\mathcal{L}(\mathcal{P}) \supset \sigma(\mathcal{P})$ を示す。このためには、$\mathcal{L}(\mathcal{P})$ が $\sigma$-集合族となることを示せばよい。
1st step　$A \in \mathcal{P}$ を任意に固定し、$\mathcal{G}_A = \{B; A \cap B \in \mathcal{L}(\mathcal{P})\}$ とおく。このときが $\mathcal{G}_A$ が $\mathcal{P}$ を含む $S$ 上の Dynkin 族となることを示す。
$\pi$ 族の定義から $B \in \mathcal{P}$ であれば $A \cap B \in \mathcal{P} \subset \mathcal{L}(\mathcal{P})$. よって、$B \in \mathcal{G}_A$, 即ち、$\mathcal{P} \subset \mathcal{G}_A$. 特に (a) $S \in \mathcal{P} \subset \mathcal{G}_A$ を得る。
(b) $B_1, B_2 \in \mathcal{G}_A$, $B_1 \supset B_2$ とすると、$A \cap B_1, A \cap B_2 \in \mathcal{L}(\mathcal{P})$ で $A \cap B_1 \supset A \cap B_2$ より、$A \cap (B_1 \backslash B_2) = (A \cap B_1) \backslash (A \cap B_2) \in \mathcal{L}(\mathcal{P})$. よって、$B_1 \backslash B_2 \in \mathcal{G}_A$.
(c) $B_n \in \mathcal{G}_A$, $B_n \subset B_{n+1}$ $(\forall n \in \mathbf{N})$ とすると、$A \cap B_n \in \mathcal{L}(\mathcal{P})$, $A \cap B_n \subset A \cap B_{n+1}$ より $A \cap \left(\bigcup_{n=1}^{\infty} B_n\right) = \bigcup_{n=1}^{\infty} A \cap B_n \in \mathcal{L}(\mathcal{P})$. よって、$\bigcup_{n=1}^{\infty} B_n \in \mathcal{G}_A$.
特に、$\mathcal{L}(\mathcal{P}) \subset \mathcal{G}_A$ となり、$\forall A \in \mathcal{P}$, $\forall B \in \mathcal{L}(\mathcal{P})$ に対して $A \cap B \in \mathcal{L}(\mathcal{P})$ となる。
2nd step　$A \in \mathcal{L}(\mathcal{P})$ を任意に固定し、$\mathcal{G}'_A = \{B; A \cap B \in \mathcal{L}(\mathcal{P})\}$ とする。
1st step の最後の注意により $\mathcal{P} \subset \mathcal{G}'_A$ となり、特に (a) $S \in \mathcal{P} \subset \mathcal{G}'_A$ を得る。
(b), (c) は 1st step とまったく同様に示せるので、これより $\mathcal{G}'_A$ は $\mathcal{P}$ を含む $S$ 上の Dynkin 族となる。
特に $\forall A, B \in \mathcal{L}(\mathcal{P})$ ならば $A \cap B \in \mathcal{L}(\mathcal{P})$ となり、$\mathcal{L}(\mathcal{P})$ 自身が $\pi$ 族になっていることがわかる。
3rd step　$\mathcal{L}(\mathcal{P})$ が $\sigma$-集合族となることを示す。
(i) $S \in \mathcal{L}(\mathcal{P})$ は明らか。(ii) $A \in \mathcal{L}(\mathcal{P})$ ならば、(i) より $S \in \mathcal{L}(\mathcal{P})$ であるから、$A^{\mathrm{c}} = S \backslash A$ より $A^{\mathrm{c}} \in \mathcal{L}(\mathcal{P})$.
(iii) $A_n \in \mathcal{L}(\mathcal{P})$ $(n \in \mathbf{N})$ とする。このとき、$B_n = \bigcup_{k=1}^{n} A_k$ とおくと、(ii) と $\mathcal{L}(\mathcal{P})$ が $\pi$ 族となることから、$B_n = \left(\bigcap_{k=1}^{n} A_k{}^{\mathrm{c}}\right)^{\mathrm{c}} \in \mathcal{L}(\mathcal{P})$ を得る。したがって、(c) より $\bigcup_{k=1}^{\infty} A_k = \bigcup_{k=1}^{\infty} B_k \in \mathcal{L}(\mathcal{P})$ を得る。
以上より、証明は完了した。　□

**問題 3.5** 定理 3.12 の証明で 2nd step の証明を述べよ。すなわち、$A \in \mathcal{L}(\mathcal{P})$ を任意に固定し $\mathcal{G}'_A = \{B; A \cap B \in \mathcal{L}(\mathcal{P})\}$ とするとき、$\mathcal{G}'_A$ が $\mathcal{P}$ を含む $S$ 上の Dynkin 族となることを示せ。

**定義 3.5** 確率変数 $X$ に対して、それが定める $\mathbf{R}$ 上の確率測度を $\mu_X$ と書く:

$$\mu_X(A) = P(X \in A), \qquad A \in \mathcal{B}(\mathbf{R}). \tag{3.2}$$

ここで $\mathcal{B}(\mathbf{R})$ は $\mathbf{R}$ の Borel 集合族である。この $\mu_X$ を $X$ の分布 (distribution) という。

**注意 3.2** $\mu_X$ は $X$ から一意的に定まる。逆に $(\mathbf{R}, \mathcal{B}(\mathbf{R}))$ 上の確率測度 $\mu$ に対して、適当に確率空間 $(\Omega, \mathcal{F}, P)$ を定めれば、$\mu$ をその分布としてもつ確率変数が (無限個) 構成できる ($cf.$ [F] p.47, p.83.)。注意 1.2 も参照ください。

$X$ の分布関数 $F_X(x)$ に対して $F_X(x) = P(X \le x) = \mu_X((-\infty, x])$ に注意する。次が成立する。

**定理 3.13** $X, Y$ を確率変数とする。$X, Y$ の分布が一致する: $\mu_X = \mu_Y$, 即ち、$\mu_X(A) = \mu_Y(A)$ $(\forall A \in \mathcal{B}(\mathbf{R}))$ であることと $F_X(x) = F_Y(x)$ $(\forall x \in \mathbf{R})$ であることとは同値とである。

**証明:** $(\Longrightarrow)$ $A = (-\infty, x]$ ととればよい。
$(\Longleftarrow)$ $\mathcal{J} = \{(-\infty, x]; x \in \mathbf{R}\} \cup \{(-\infty, \infty)\}$ とし、$\mathfrak{A} = \{A \in \mathcal{B}(\mathbf{R}); \mu_X(A) = \mu_Y(A)\}$ とする。このとき、$\mathcal{J}$ が $\mathbf{R}$ 上の $\pi$ 族になることは明らか。$\mathfrak{A}$ が $\mathbf{R}$ 上の Dynkin 族となることは測度の性質より容易に証明できる (問題 3.6)。$x \in (-\infty, \infty)$ のとき仮定より、$\mu_X((-\infty, x]) = F_X(x) = F_Y(x) = \mu_Y((-\infty, x])$ で、$x = \infty$ のとき $\mu_X((-\infty, \infty)) = \mu_Y((-\infty, \infty)) = 1$. よって、$\mathcal{J} \subset \mathfrak{A}$ となる。よって、定理 3.12 により $\sigma(\mathcal{J}) = \mathcal{L}(\mathcal{J}) \subset \mathfrak{A}$.
ところで、$\mathcal{J}$ を含む最小の $\sigma$-集合族 $\sigma(\mathcal{J})$ は $\forall a, b \in \mathbf{R}: a < b$ に対して $(a, b] = (-\infty, b] \cap (-\infty, a]^c \in \sigma(\mathcal{J})$ となるので、Borel 集合族 $\mathcal{B}(\mathbf{R})$ ($cf.$ 定義 1.5) と一致し、$\mathcal{B}(\mathbf{R}) \subset \mathfrak{A}$ となり主張を得る。　□

**問題 3.6** 定理 3.13 の証明で定めた $\mathfrak{A}$ が $\mathbf{R}$ 上の Dynkin 族となることを示せ。

[12 月 9 日]

**命題 3.14** 可測関数 $f(x)$ が $f \geq 0$ もしくは $\int_{\mathbf{R}} |f(x)|\mu_X(dx) < \infty$ を満たせば、$E[f(X)] = \int_{\mathbf{R}} f(x)\mu_X(dx)$.

**略証:** (§1.5 Lebesgue 積分の節の各 step に従って考察する。)
1st step: $f(x)$ が階段関数 $f(x) = \sum a_i 1_{A_i}(x)$ の場合は次のようにと示すことができる。

$$E[f(X)] = \sum_i a_i P(X \in A_i) = \sum_i a_i \mu_X(A_i) = \int_{\mathbf{R}} f(x)\mu_X(dx)$$

2nd step の $f(x) \geq 0$ の場合はそこで定義された単関数による増加列 $\{f_q(x)\}$ を考え単調収束定理 (定理 1.24) を用いて証明できる。3rd step の $f(x)$ が実数値の場合は $f(x) = f^+(x) - f^-(x)$ として、さらに、複素数値の場合は $f(x) = \mathrm{Re}\, f(x) + i\, \mathrm{Im}\, f(x)$ として証明できる。 □

**定理 3.15** $(\mathbf{R}, \mathcal{B}(\mathbf{R}))$ 上の確率測度 $\mu, \nu$ について、$\forall f \in C_b(\mathbf{R})$ に対して $\int_{\mathbf{R}} f(x)\mu(dx) = \int_{\mathbf{R}} f(x)\nu(dx)$ であれば、$\mu = \nu$ となる。ここで、$C_b(\mathbf{R})$ は $\mathbf{R}$ 上の有界連続関数全体を表す。

**証明:** ここで、$a \in \mathbf{R}$ を任意とし、(グラフを書く)

$$f_n(x) = 1 \quad (x \leq a), \qquad = 1 - n(x-a) \quad (a < x < a + \frac{1}{n}), \qquad = 0 \quad (a \leq x)$$

とすると、$f_n \in C_b(\mathbf{R})$ より、$\int_{\mathbf{R}} f_n(x)\mu(dx) = \int_{\mathbf{R}} f_n(x)\nu(dx)$ であるが、$\forall x \in \mathbf{R}$ に対して $f_n(x) \to 1_{(-\infty,a]}(x)$ $(n \to \infty)$ かつ $0 \leq f_n(x) \leq 1$ となるので Lebesgue の収束定理 (定理 2.10) を用いると、$\int_{\mathbf{R}} 1_{(-\infty,a]}(x)\mu(dx) = \int_{\mathbf{R}} 1_{(-\infty,a]}(x)\nu(dx)$, 即ち、$\mu((-\infty,a]) = \nu((-\infty,a])$ となる。よって、定理 3.13 の証明と全く同様にして $\mu = \nu$ を得る。 □

**問題 3.7** 定理 3.12 を用いて等式 (1.4) が成り立つことを $n = 2$ のときに示せ。

## 3.3 特性関数と分布

**定理 3.16** 2 つの特性関数が一致すれば、それらは同一の分布の特性関数である。即ち、もし $\phi_X(t) = \phi_Y(t)$, $\forall t \in \mathbf{R}$, であれば、$\mu_X = \mu_Y$ (あるいは $F_X(x) = F_Y(x)$, $\forall x \in \mathbf{R}$) となる。

これを証明するため、いくつか準備を必要とする。

次の補題は解析学で学んだ Cauchy の積分定理を用いて証明できる ($cf$. 昨年度の講義ノート pp.18-19)。

**補題 3.17** $\int_0^\infty \frac{\sin x}{x}\, dx = \frac{\pi}{2}$. すなわち、$f(M) = \int_0^M \frac{\sin t}{t}\, dt$ とすると、$\lim_{M\to\infty} f(M) = \frac{\pi}{2}$.

**証明:** $t > 0$ のとき $\int_0^\infty e^{-ut}\, du = \frac{1}{t}$ より $f(M) = \int_0^M \frac{\sin t}{t}\, dt = \int_0^M \int_0^\infty e^{-ut} \sin t\, du\, dt$ となる。同様に、$\int_0^M \int_0^\infty \left|e^{-ut} \sin t\right| du\, dt = \int_0^M \frac{|\sin t|}{t}\, dt < \infty$ ($cf$. 問題 3.8) より、Fubini の定理 (定理 2.11) から、$\int_0^M \int_0^\infty e^{-ut} \sin t\, du\, dt = \int_0^\infty \int_0^M e^{-ut} \sin t\, dt\, du$. 部分積分することにより

$$\int_0^M e^{-ut} \sin t\, dt = \frac{\varphi_M(u)}{1+u^2}, \quad ただし \quad \varphi_M(u) = 1 - e^{-Mu}\cos M - ue^{-Mu}\sin M \tag{3.3}$$

を得る。さらに、$|\varphi_M(u)| \leq 3$ $(u \geq 0)$ であるから Lebesgue の収束定理 (定理 2.10) を用いて

$$\lim_{M\to\infty} f(M) = \int_0^\infty \lim_{M\to\infty} \frac{\varphi_M(u)}{1+u^2}\, du = \int_0^\infty \frac{1}{1+u^2}\, du = \frac{\pi}{2}. \quad \square$$

**問題 3.8** (1)　補題 3.17 で $M > 0$ に対して $|f(M)| < \infty$ と $|\varphi_M(u)| \leq 3$ $(u \geq 0)$ を示せ。ヒント: $|\sin u| \leq |u|$ $(|u| \leq \frac{\pi}{2})$ を用いよ。
(2)　補題 3.17 を用いて $\sup_{M>0} |f(M)| < \infty$ を示せ。また、$\displaystyle\int_0^\infty \left|\frac{\sin t}{t}\right| dt = \infty$ となることを示せ。

**補題 3.18** $f_T(\alpha) = \displaystyle\int_0^T \frac{\sin \alpha t}{t}\, dt$, $T > 0$, $\alpha \in \mathbf{R}$ について、(1) $\sup_{T,\alpha} |f_T(\alpha)| < \infty$,

(2) $\displaystyle\lim_{T\to\infty} f_T(\alpha) = \frac{\pi}{2}\left\{1_{(0,\infty)}(\alpha) - 1_{(-\infty,0)}(\alpha)\right\} = \begin{cases} \frac{\pi}{2} & \alpha > 0 \\ 0 & \alpha = 0 \\ -\frac{\pi}{2} & \alpha < 0 \end{cases}$ が成り立つ。

**証明:** (1)　$u = \alpha t$ とおくと、$f_T(\alpha) = \displaystyle\int_0^{\alpha T} \frac{\sin u}{u}\, du$ となる。よって、問題 3.8 より主張は従う。
(2) $\alpha = 0$ のとき $f_T(0) = 0$ より明らか。$\alpha > 0$ のときは補題 3.17 より成立する。$\alpha < 0$ のときは $-s = t$ とすれば $\alpha > 0$ の場合に帰着できる。　□

**定理 3.19 (Lévy の反転公式)** 確率変数 $X$ の分布関数 $F$ と特性関数 $\phi(t)$ について次が成立する。

$$\frac{1}{2}\{F(b) + F(b-0) - F(a) - F(a-0)\} = \frac{1}{2\pi}\lim_{T\to\infty}\int_{-T}^T \frac{e^{-itb} - e^{-ita}}{-it}\phi(t)\, dt, \quad a, b \in \mathbf{R}, a < b. \tag{3.4}$$

特に、$a, b$ がともに $F$ の連続点であれば、(3.4) の右辺 $= F(b) - F(a)$ となる。ここで $x$ が $F_X$ の連続点であるとは $\lim_{y\to x} F_X(y) = F_X(x)$ となる $x$ のことである。

**証明:** $X$ の分布を $\mu_X$ とすると、命題 3.14 により

$$\int_{-T}^T \frac{e^{-itb} - e^{-ita}}{-it}\phi(t)\, dt = \int_{-T}^T\int_{\mathbf{R}} \frac{e^{-itb} - e^{-ita}}{-it} e^{itx}\, \mu_X(dx)\, dt$$

である。ここで、

$$|e^{-itb} - e^{-ita}| = \left|\int_a^b (-ite^{-it\theta})\, d\theta\right| \leq \int_a^b |-ite^{-it\theta}|\, d\theta = |it(b-a)|$$

となるから、$\left|\frac{e^{-itb} - e^{-ita}}{-it}\right| \leq |b-a|$ より、$\displaystyle\int_{-T}^T\int_{\mathbf{R}} \left|\frac{e^{-itb} - e^{-ita}}{-it} e^{itx}\right| \mu_X(dx)dt \leq |b-a| \cdot 2T < \infty$ に注意して Fubini の定理 (定理 2.11) を用いて、

$$右辺 = \int_{\mathbf{R}}\left(\int_{-T}^T \frac{e^{-itb} - e^{-ita}}{-it} e^{itx}\, dt\right)\mu_X(dx) = \int_{\mathbf{R}}\left(\int_{-T}^T \frac{\sin t(x-a) - \sin t(x-b)}{-t}\, dt\right)\mu_X(dx)$$

となる。ここで、$e^{i\xi} = \cos\xi + i\sin\xi$ と

$$\int_{-T}^T \frac{\cos t(x-a) - \cos t(x-b)}{-it}\, dt = 0$$

を用いた。これは被積分関数が $t$ の奇関数であるためである。したがって、補題 3.18 の $f_T(\alpha)$ を用いると、

$$右辺 = 2\int_{\mathbf{R}} (f_T(x-a) - f_T(x-b))\, \mu_X(dx)$$

となる。ここで、補題 3.18(1), (2) に注意して Lebesgue の収束定理 (定理 2.10) を用いると、$T \to \infty$ のとき、

$$\begin{aligned} &2\int_{\mathbf{R}} \frac{\pi}{2}\left[1_{(0,\infty)}(x-a) - 1_{(-\infty,0)}(x-a) - 1_{(0,\infty)}(x-b) + 1_{(-\infty,0)}(x-b)\right] \mu_X(dx) \\ &= \pi\int_{\mathbf{R}} \left[1_{(a,\infty)}(x) - 1_{(-\infty,a)}(x) - 1_{(b,\infty)}(x) + 1_{(-\infty,b)}(x)\right] \mu_X(dx) \end{aligned}$$

$$
\begin{aligned}
&= \pi\left[\mu_X((a,\infty)) - \mu_X((-\infty,a)) - \mu_X((b,\infty)) + \mu_X((-\infty,b))\right] \\
&= \pi\left[1 - F(a) - F(a-0) - (1-F(b)) + F(b-0)\right]
\end{aligned}
$$

となる。$a, b$ が $F$ の連続点であれば、$F(b-0) = F(b)$ かつ $F(a-0) = F(a)$ であるから

$$
(3.4)\text{ の右辺} = \frac{1}{2\pi}\pi\left[1 - F(a) - F(a-0) - (1-F(b)) + F(b-0)\right] = F(b) - F(a) \tag{3.5}
$$

を得る。 □

**定理 3.16 の証明:** $F_X$ と $F_Y$ の連続点の共通部分を $R_c$ とする。$R_c$ の補集合は高々可算個の点からなるので、$R_c$ は $\mathbf{R}$ で稠密となる。定理 3.19 により、$a, b \in R_c$ であれば $F_X$, $F_Y$ の両方の連続点なので、

$$
F_X(b) - F_X(a) = F_Y(b) - F_Y(a).
$$

よって $\{a_n\} \subset R_c$, $a_n \to -\infty$ ととれば、$F_X(b) = F_Y(b)$ となる。今、分布関数は右連続であるから、$\forall x \in \mathbf{R}$ に対して、$\{b_n\} \subset R_c$ を $b_n \to x+0$ と選べば、$F_X(x) = F_Y(x)$ となる。よって、定理 3.13 より $\mu_X = \mu_Y$ である。 □

**例 3.20** $X_1, \ldots, X_n$ を独立で、各 $X_j$ が Poisson 分布 $Po(\lambda_j)$ に従うとする。このとき、$Y = X_1 + \cdots + X_n$ は Poisson 分布 $Po(\lambda_1 + \cdots + \lambda_n)$ に従う。実際、

$$
\phi_Y(t) = E[e^{itX_1} \cdots e^{itX_n}] = E[e^{itX_1}] \cdots E[e^{itX_n}] = e^{\lambda_1(e^{it}-1)} \cdots e^{\lambda_n(e^{it}-1)} = e^{(\lambda_1 + \cdots + \lambda_n)(e^{it}-1)}
$$

ここで 2 つ目の等号は $X_1, \cdots, X_n$ の独立性を、次の等号は例 3.4(2) を用いた。よって、再び例 3.4(2) により右辺が Poisson 分布 $Po(\lambda_1 + \cdots + \lambda_n)$ の特性関数とわかるから、定理 3.16 により主張を得る。

**問題 3.9** (1) $0 < p < 1$ とする。$X_1, \ldots, X_n$ を独立で、各 $X_j$ が負の二項分布 $NB(\alpha_j, p)$, $\alpha_j > 0$, に従うとき、$Y = X_1 + \cdots + X_n$ は負の二項分布 $NB(\alpha_1 + \cdots + \alpha_n, p)$ に従うことを示せ。($cf$. 問題 3.2 (1).)
(2) $X_1, \ldots, X_n$ はを独立で、各 $X_k$ は正規分布 $N(m_k, v_k)$ に従う確率変数とし、$a_1, \ldots, a_n$ を定数とする。このとき、$Y = \sum_{k=1}^{n} a_k X_k$ とおくと、$Y$ は正規分布 $N\left(\sum_{k=1}^{n} a_k m_k, \sum_{k=1}^{n} {a_k}^2 v_k\right)$ に従うことを示せ。

注意: 以降、正規分布表はアップロードしたものを用いてください。

**問題 3.10** (1) $X_1, X_2, X_3$ は独立で、それぞれ正規分布 $N(1,1)$, $N(2,2)$, $N(3,3)$ に従うとし、$Y = 2X_1 - 3X_2 + X_3$ とおく。このとき、$Y$ は正規分布に従うことを示し、確率 $P(-10.2 \leq Y \leq 9.9)$ を求めよ。
(2) $X_1, X_2, \cdots, X_n$ は独立でそれぞれ正規分布 $N(2,4)$ に従うとき、$\overline{X} = \frac{1}{n}\sum_{k=1}^{n} X_k$ に対して、$P(1.5 \leq \overline{X} \leq 2.5) \geq 0.9$ となるような最小の $n$ を求めよ。ただし、$\displaystyle\int_{1.645}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2}\, dt = 0.05$ を用いよ。

**問題 3.11** $f(x) = \dfrac{1}{x\sqrt{2\pi}} e^{-\frac{1}{2}(\log x)^2} 1_{(0,\infty)}(x)$ とし、$f_a(x) = (1 + a\sin(2\pi \log x)) f(x)$, $-1 \leq a \leq 1$, とする。
(1) $f_a(x)$ が密度関数となることを示せ。
(2) $f_a(x)$ を密度関数とする確率変数を $X_a$ とするとき、$\forall n \in \mathbf{N}$ に対して、$E[|X_a|^n] < \infty$ を示し、$E[{X_a}^n]$ が $-1 \leq a \leq 1$ によらないことを確かめよ。

**注意** この例より、$E[X^n] = E[Y^n]$ $(\forall n \in \mathbf{N})$ であっても、$\phi_X(t) = \phi_Y(t)$ とは限らない (即ち、分布が一致するとは限らない) ことがわかる。

**問題 3.12** $\displaystyle\int_0^1\int_0^1 \left|\frac{x^2 - y^2}{(x^2+y^2)^2}\right| dxdy = \infty$ であり、$\displaystyle\int_0^1\left(\int_0^1 \frac{x^2 - y^2}{(x^2+y^2)^2}\, dx\right)dy \neq \int_0^1\left(\int_0^1 \frac{x^2 - y^2}{(x^2+y^2)^2}\, dy\right)dx$ となることを示せ。ヒント: $\dfrac{\partial}{\partial x}\left(\dfrac{x}{x^2+y^2}\right)$ を計算せよ。

**注意** この例より、Fubini の定理 (定理 2.11) で可積分条件 $\displaystyle\int_{R_1 \times R_2} |f(x,y)|\, d(\mu_1 \otimes \mu_2)(x,y) < \infty$ が成立しない場合には、必ずしも積分の順序が交換できないことがわかる。

[12 月 16 日]

**定理 3.21** 確率変数 $X$ の特性関数 $\phi(t)$ が $\displaystyle\int_{-\infty}^{\infty} |\phi(t)|\, dt < \infty$ を満たせば、$X$ は絶対連続型でその密度関数 $f_X(x)$ は次で与えられる。

$$f_X(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-itx}\phi(t)\, dt. \tag{3.6}$$

**証明:** 定理 3.19 の記号を用いると $\left|\dfrac{e^{-itb}-e^{-ita}}{-it}\right| \le |b-a|$ であったから、(3.4) の右辺の被積分関数は $(-\infty,\infty)$ で積分可能なので、(3.5) より $F(x)$ の連続点とは限らずに $\forall a, b\ (a<b)$ に対して、

$$\frac{1}{2}\{F(b)+F(b-0)-F(a)-F(a-0)\} = \frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{e^{-itb}-e^{-ita}}{-it}\phi(t)\, dt \le \frac{b-a}{2\pi}\int_{-\infty}^{\infty}|\phi(t)|\, dt$$

となる。ここで、$b_n$ を $F(x)$ の連続点として、$b=b_n$ として上式に代入し、$b_n \to a+0$ とすると、

$$\frac{1}{2}\{F(a)-F(a-0)\} \le 0$$

となり、$F(x)$ は非減少だから $F(a)=F(a-0)$, 即ち、連続であることがわかる。さらに、

$$F(x+h)-F(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{e^{-it(x+h)}-e^{-itx}}{-it}\phi(t)\, dt = \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{x}^{x+h} e^{-ity}\phi(t)\, dy\, dt$$

であるが、$h>0$ に対し $\displaystyle\int_{-\infty}^{\infty}\int_{x-h}^{x+h}|e^{-ity}\phi(t)|\, dy\, dt \le 2h\int_{-\infty}^{\infty}|\phi(t)|\, dt < \infty$ であるから、Fubini の定理 (定理 2.11) により $\forall h \in \mathbf{R}$ に対し

$$F(x+h)-F(x) = \frac{1}{2\pi}\int_{x}^{x+h}\int_{-\infty}^{\infty} e^{-ity}\phi(t)\, dt\, dy$$

ここで、$\displaystyle\int_{-\infty}^{\infty}|\phi(t)|\, dt < \infty$ であるから、命題 3.2(iv) と同様に Lebesgue の収束定理 (定理 2.10) により $\displaystyle\int_{-\infty}^{\infty} e^{-ity}\phi(t)\, dt$ は $y$ について連続となる。従って、$F(x)$ は微分可能で $F'(x) = \dfrac{1}{2\pi}\displaystyle\int_{-\infty}^{\infty} e^{-itx}\phi(t)\, dt$ より (3.6) を得る。□

**例 3.8 の証明:** 問題 3.2(2) より $f(x) = \frac{1}{2}e^{-|x|}$ を密度関数とするとその特性関数は $\phi(t) = \frac{1}{1+t^2}$ であった。ここで、$\displaystyle\int_{-\infty}^{\infty}|\phi(t)|\, dt = \pi < \infty$ より、定理 3.21 から

$$\frac{1}{2}e^{-|x|} = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-itx}\frac{1}{1+t^2}\, dt$$

を得る。ここで、$x$ を $-t$, $t$ を $x$ と読み替えることで Cauchy 分布の特性関数 $\phi_X(t)$

$$\phi_X(t) = \int_{-\infty}^{\infty} e^{itx}\frac{1}{\pi}\frac{1}{1+x^2}\, dx = 2\cdot\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-i(-t)x}\frac{1}{1+x^2}\, dx = 2\cdot\frac{1}{2}e^{-|-t|} = e^{-|t|}$$

を得る。　□

**例 3.22** $X_1, X_2, \ldots$ を独立な確率変数列で Cauchy 分布に従う、すなわち、その密度関数が $f(x) = \dfrac{1}{\pi}\dfrac{1}{1+x^2}$ $(-\infty < x < \infty)$ だとする。例 3.8 により、その特性関数は $\phi_{X_i}(t) = e^{-|t|}$ となる。このとき、$Y_n = \frac{1}{n}(X_1+X_2+\cdots+X_n)$ とおくと、

$$\phi_{Y_n}(t) = E[e^{i\frac{t}{n}X_1}\cdots e^{i\frac{t}{n}X_n}] = E[e^{i\frac{t}{n}X_1}]\cdots E[e^{i\frac{t}{n}X_n}] = e^{-|\frac{t}{n}|}\cdots e^{-|\frac{t}{n}|} = e^{-|t|}$$

となり、$Y_n$ も同じ Cauchy 分布に従うことがわかる。Cauchy 分布は期待値を持たないことに注意すると、これは、期待値を持たない独立確率変数列で大数の法則が成り立たない例となっている。

**問題 3.13** $Y$ の密度関数 $g(y)$ が $g(0) = \dfrac{1}{2\pi}$, $g(y) = \dfrac{1}{2\pi}\dfrac{2(1-\cos y)}{y^2}$ $(y \neq 0)$ のとき、$Y$ の特性関数を求めよ。ヒント: 問題 3.1(1) を用いて、例 3.8 の証明と同様に定理 3.21 を用いよ。

## 3.4 法則収束と弱収束

**定義 3.6 (法則収束)** 任意の $f \in C_b(\mathbf{R})$ に対して、

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

が成立するとき、$X_n$ が $X$ に法則収束 (convergence in law) または分布収束 (covergence in distribution) するといい、$X_n \to X$ in law と表す。ここで $C_b(\mathbf{R})$ は $\mathbf{R}$ 上の有界連続関数全体を表す。

法則収束は、分布の間の距離として距離付け可能となる。(*cf*. 例えば 小谷真一著 測度と確率 pp.206–207.)

**定理 3.23** 確率変数列 $\{X_n\}$ が $X$ に確率収束すれば、法則収束する。

**証明:** 1st step　まず、$\{X_n\}$ が $X$ に概収束する場合を考える。このとき、$f \in C_b(\mathbf{R})$ に対して、$f(X_n)$ は $f(X)$ に概収束し $f$ は有界だからある $M$ があって $|f(x)| \leq M$ $(\forall x \in \mathbf{R})$ とできるので、$|f(X_n(\omega))| \leq M$ $(\omega \in \Omega)$ とできる。したがって、Lebesgue の収束定理 (定理 2.10) により

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

となり、$X_n$ が $X$ に法則収束する。
2nd step　$\{X_n\}$ が $X$ に確率収束するとし、$f \in C_b(\mathbf{R})$ に対し、$a_n = E[f(X_n)]$ とおく。まず、$|f(x)| \leq M$ $(x \in \mathbf{R})$ であれば、$|a_n| \leq M$ であるから、その任意の部分列は収束部分列を持つことに注意する。ここで、もし $\{a_n\}$ が $a = E[f(X)]$ に収束しないとすると、ある部分列 $\{a_{n'}\}$ があって $a$ 以外に収束する。一方、$\{a_{n'}\}$ に対応する確率変数列 $\{X_{n'}\}$ に対して、定理 2.4 により、その部分列 $\{X_{n''}\}$ を選んで $X$ に概収束するようにできる。したがって、1st step により $\{a_{n''}\}$ は $a$ に収束する。これは $\{a_{n'}\}$ が $a$ 以外に収束することに矛盾する。よって、$\lim_{n\to\infty} E[f(X_n)] = E[f(X)]$ となる。これは任意の $f \in C_b(\mathbf{R})$ に対して成立するから、$\{X_n\}$ は $X$ に法則収束する。　□

**例 3.24** 定理 3.23 の逆は、必ずしも成立しない (*cf*. 問題 3.14 にも注意のこと)。実際、確率変数列 $\{X_n\}$ を独立で各 $n \in \mathbf{N}$ に対し $P(X_n = 1) = P(X_n = -1) = 1/2$ なるとする。このとき、$\forall f \in C_b(\mathbf{R})$ に対して

$$E[f(X_n)] = \frac{1}{2}f(1) + \frac{1}{2}f(-1)$$

となるので、これを $\lim_{n\to\infty} E[f(X_n)] = E[f(X_1)]$ と解釈すれば、$\{X_n\}$ は $X_1$ に法則収束する (実はどの $X_k$ にも収束するといえる。)。一方、$0 < \varepsilon < 1$ とすると、$n \geq 2$ のとき、

$$P(|X_n - X_1| \geq \varepsilon) = P(X_1 = 1, X_n = -1) + P(X_1 = -1, X_n = 1) = \frac{1}{2}\cdot\frac{1}{2} + \frac{1}{2}\cdot\frac{1}{2} = \frac{1}{2}$$

となり、$\{X_n\}$ は $X_1$ に確率収束しないことがわかる。

**問題 3.14** $X_n \to a$ in law ($a$ は定数) ならば、$X_n \to a$ in prob. を示せ。

特に、法則収束では、確率変数としては極限は一意的でなくなる。しかし、極限となる分布は一意となる。まず、分布の弱収束を導入する。

**定義 3.7** $\mu_n$, $n = 1, 2, \ldots,$ と $\mu$ を ($\mathbf{R}$ より一般とし) 距離空間 $S$ 上の分布 (確率測度) とする。$\mu_n$ が $\mu$ に弱収束するとは、$\forall f \in C_b(S)$ に対して

$$\lim_{n\to\infty} \int_S f(x)\mu_n(dx) = \int_S f(x)\mu(dx)$$

が成立するときにいう。

確率変数列 $\{X_n\}$ が $X$ に法則収束することは、対応する分布の列 $\{\mu_{X_n}\}$ が $\mu_X$ に弱収束することと同値である。このとき、極限 $\mu$ は一意的である。実際、もう一つの極限を $\nu$ とすると、

$$\int_{\mathbf{R}} f(x)\,\mu(dx) = \int_{\mathbf{R}} f(x)\,\nu(dx), \qquad \forall f \in C_b(\mathbf{R})$$

となるが、これが $\mu = \nu$ を意味することは定理 3.15 で示した。

このため、確率変数の法則収束は確率測度の弱収束として説明したほうが自然である。しかし、この授業では測度の扱いに慣れていないことを配慮しできるだけ確率変数の言葉で述べていく。

**定理 3.25** 確率変数 $X_1, X_2, \ldots$ と $X$ について次は同値である。ただし、確率変数 $Y$ に対応する分布関数を $F_Y(x) = P(Y \le x)$ と表す。
(1) $X_n \to X$ in law.
(2) $F_X$ の任意の連続点 $x$ に対して $\lim_{n\to\infty} F_{X_n}(x) = F_X(x)$ が成立する。

**証明:** <u>(1) $\Rightarrow$ (2)</u> $x$ を $F_X$ の連続点とする。関数 $1_{(-\infty,x]}(y)$ を上下から近似する連続関数列 $f_\delta^+, f_\delta^- \in C_b(\mathbf{R})$, $\delta > 0$ を

$$f_\delta^+(y) = \begin{cases} 1 & y \le x \\ 1 - \frac{1}{\delta}(y - x) & x < y < x + \delta \\ 0 & y \ge x + \delta \end{cases}, \qquad f_\delta^-(y) = \begin{cases} 1 & y \le x - \delta \\ 1 - \frac{1}{\delta}(y - (x - \delta)) & x - \delta < y < x \\ 0 & y \ge x \end{cases}$$

で定める (グラフを書く)。このとき、

$$1_{(-\infty,x-\delta]}(y) \le f_\delta^-(y) \le 1_{(-\infty,x]}(y) \le f_\delta^+(y) \le 1_{(-\infty,x+\delta]}(y), \qquad y \in \mathbf{R}$$

となることに注意する。まず、

$$F_{X_n}(x) = P(X_n \le x) = E[1_{(-\infty,x]}(X_n)] \le E[f_\delta^+(X_n)]$$

で $f_\delta^+ \in C_b(\mathbf{R})$ より $\{X_n\}$ は $X$ に法則収束するから、

$$\limsup_{n\to\infty} F_{X_n}(x) \le \lim_{n\to\infty} E[f_\delta^+(X_n)] = E[f_\delta^+(X)] \le E[1_{(-\infty,x+\delta]}(X)] = P(X \le x + \delta) = F_X(x + \delta)$$

を得る。よって、$\delta \to +0$ として、

$$\limsup_{n\to\infty} F_{X_n}(x) \le \lim_{\delta\to+0} F_X(x + \delta) = F_X(x) \tag{3.7}$$

を得る。同様に、

$$F_{X_n}(x) = E[1_{(-\infty,x]}(X_n)] \ge E[f_\delta^-(X_n)]$$

で $f_\delta^- \in C_b(\mathbf{R})$ より

$$\liminf_{n\to\infty} F_{X_n}(x) \ge \lim_{n\to\infty} E[f_\delta^-(X_n)] = E[f_\delta^-(X)] \ge E[1_{(-\infty,x-\delta]}(X)] = P(X \le x - \delta) = F_X(x - \delta)$$

を得る。よって、$\delta \to +0$ として、

$$\liminf_{n\to\infty} F_{X_n}(x) \ge \lim_{\delta\to+0} F_X(x - \delta) = F_X(x)$$

を得る。最後の等号は $x$ が $F_X$ の連続点であることを用いた。これと、(3.7) をあわせて

$$\lim_{n\to\infty} F_{X_n}(x) = F_X(x)$$

となることがわかった。

(2) ⇒ (1)　まず、$F_X$ の不連続点は高々可算個しかないこと、したがって、連続点が $\mathbf{R}$ 上稠密に存在することに注意する。まず、$\varepsilon > 0$ を任意にとる。$F_X$ の連続点 $a, b \in \mathbf{R}$ $(a < b)$ を

$$F_X(a) \le \varepsilon, \qquad 1-\varepsilon \le F_X(b)$$

と選べる。特に、条件 (2) により、ある $N$ があって

$$n \ge N \quad \Longrightarrow \quad F_{X_n}(a) \le 2\varepsilon,\ 1-2\varepsilon \le F_{X_n}(b)$$

とできる。次に $\delta > 0$ と $f \in C_b(\mathbf{R})$ が任意に与えられたとして点列 $a = a_0 < a_1 < \cdots < a_K = b$ を

- 各 $a_j$ $(1 \le j \le K-1)$ は $F_X$ の連続点
- $\displaystyle\max_{a_{j-1}\le x\le a_j} |f(x) - f(a_j)| \le \delta$ $(1 \le j \le K)$

を満たすようにとる。第 2 の条件は、連続関数 $f$ は有界閉区間 $[a,b]$ 上で一様連続だから可能となる。このとき

$$h_f(x) = \sum_{j=1}^K f(a_j) 1_{(a_{j-1},a_j]}(x)$$

とおく。$\|f\|_\infty = \sup_{x\in\mathbf{R}} |f(x)|$ と表すと、$y \notin (a,b]$ のとき $|f(y) - h_f(y)| = |f(y)| \le \|f\|_\infty$ だから、$n \ge N$ であれば、

$$\begin{aligned}
|E[f(X_n)] - E[h_f(X_n)]| &\le \sum_{j=1}^K E[|f(X_n) - h_f(X_n)| 1_{(a_{j-1},a_j]}(X_n)] + E[|f(X_n) - h_f(X_n)| 1_{(a,b]^c}(X_n)] \\
&\le \sum_{j=1}^K \delta P(X_n \in (a_{j-1}, a_j]) + \|f\|_\infty P(X_n \notin (a,b]) \\
&= \delta P(X_n \in (a_0, a_K]) + \|f\|_\infty (F_{X_n}(a) + 1 - F_{X_n}(b)) \\
&\le \delta + 4\varepsilon \|f\|_\infty.
\end{aligned}$$

同様に

$$\begin{aligned}
|E[f(X)] - E[h_f(X)]| &\le \delta P(X \in (a_0, a_K]) + \|f\|_\infty (F_X(a) + 1 - F_X(b)) \\
&\le \delta + 2\varepsilon \|f\|_\infty.
\end{aligned}$$

一方、各 $a_j$ は $F_X$ の連続点だから (2) の仮定より $F_{X_n}(a_j) \to F_X(a_j)$ $(n \to \infty)$ となるので

$$\begin{aligned}
E[h_f(X_n)] &= \sum_{j=1}^K E[h_f(X_n) 1_{(a_{j-1},a_j]}(X_n)] = \sum_{j=1}^K f(a_j)(F_{X_n}(a_j) - F_{X_n}(a_{j-1})) \\
&\to \sum_{j=1}^K f(a_j)(F_X(a_j) - F_X(a_{j-1})) = E[h_f(X)] \qquad (n \to \infty)
\end{aligned}$$

よって、三角不等式を用いて

$$|E[f(X_n)] - E[f(X)]| \le |E[f(X_n)] - E[h_f(X_n)]| + |E[h_f(X_n)] - E[h_f(X)]| + |E[h_f(X)] - E[f(X)]|$$

としから、上の評価を用いると

$$\limsup_{n\to\infty} |E[f(X_n)] - E[f(X)]| \le 2\delta + 6\varepsilon \|f\|_\infty$$

がわかる。左辺は $\varepsilon, \delta$ によらないので、$\varepsilon, \delta > 0$ が任意だったことに注意して、$\delta \to +0$, $\varepsilon \to +0$ とすると

$$\limsup_{n\to\infty} |E[f(X_n)] - E[f(X)]| \le 0.$$

よって、$\displaystyle\lim_{n\to\infty} E[f(X_n)] = E[f(X)]$ となる。　□

[1 月 6 日]

**注意 3.3** 定理 3.25 (2) で任意の点 $x$ に対してを $\lim_{n\to\infty} F_{X_n}(x) = F_X(x)$ は一般には成立しない。実際、$X_n = 1/n$ (定数), $n = 1, 2, \ldots,$ に対して、数列として $X_n \to 0$ となるから、特に $X_n \to 0$ in law も示せる。一方、$F_{X_n}(x) = \begin{cases} 0 & x < 1/n \\ 1 & x \geq 1/n \end{cases}$ であり、あえて $X = 0$ (定数) とすると $F_X(x) = \begin{cases} 0 & x < 0 \\ 1 & x \geq 0 \end{cases}$ である。これより、$F_{X_n}(0) = 0, n = 1, 2, \ldots,$ となり、$F_X(0) = 1$ となる。

**例 3.26** $\alpha > 0$ とする。$X_1, X_2, \ldots$ が i.i.d. でその密度関数は $f(x) = \alpha(x+1)^{-\alpha+1} 1_{(0,\infty)}(x)$ であるとする (Parate 分布という)。このとき、$Y_n = n^{-1/\alpha} \max\{X_1, X_2, \cdots, X_n\}$ とおくと、$\{Y_n\}$ は次の分布関数 $F_Z(z)$ をもつ確率変数 $Z$ に法則収束する。ただし。$F_Z(z) = 0$ $(z \leq 0)$, $F_Z(z) = e^{-z^{-\alpha}}$ $(z > 0)$ である (Fréchet 分布という)。

**証明:** $F_Z$ は $\mathbf{R}$ 上で連続なので、$\forall z \in \mathbf{R}$ に対して $\lim_{n\to\infty} P(Y_n \leq z) = F_Z(z)$ を示せばよい。$z \leq 0$ のとき、$\lim_{n\to\infty} P(Y_n \leq z) = 0$ は明らか。$z > 0$ のとき、$Y_n$ の分布関数は

$$\begin{aligned}
P(Y_n \leq z) &= P(X_1 \leq n^{1/\alpha} z, \cdots, X_n \leq n^{1/\alpha} z) = P(X_1 \leq n^{1/\alpha} z) \times \cdots \times P(X_n \leq n^{1/\alpha} z) \\
&= \left( \int_0^{n^{1/\alpha} z} \alpha (t+1)^{-\alpha-1}\, dt \right)^n = \left( 1 - (n^{1/\alpha} z + 1)^{-\alpha} \right)^n \\
&= \left( 1 - \frac{1}{n} \left( z + n^{-1/\alpha} \right)^{-\alpha} \right)^n \to e^{-z^{-\alpha}} \qquad (n \to \infty)
\end{aligned}$$

となるので、$\{Y_n\}$ は $Z$ に法則収束する。 □

**問題 3.15** $X_1, X_2, \cdots$ を i.i.d. とするとき、次を示せ。
(1) $\alpha > 0$ とする。$X_1$ が Beta 分布 $B(1, \alpha)$ に従うとき、$Y_n = n^{1/\alpha}(\max\{X_1, X_2, \cdots, X_n\} - 1)$ は次の分布関数 $F_Z(z)$ をもつ確率変数 $Z$ に法則収束する。ただし。$F_Z(z) = e^{-(-z)^{-\alpha}}$ $(z < 0)$, $F_Z(z) = 1$ $(z \geq 0)$ である (Weibull 分布という)。
(2) $X_1$ が指数分布 $Ex(1)$ に従うとき、$Y_n = \max\{X_1, X_2, \cdots, X_n\} - \log n$ は次の分布関数 $F_Z(z)$ をもつ確率変数 $Z$ に法則収束する。ただし。$F_Z(z) = e^{-e^{-z}}$ $(z \in \mathbf{R})$ とする (Gumbel 分布という)。

**注意 3.4 (極値理論, Fisher-Tippet の定理)** $X_1, X_2, \ldots$ を i.i.d. とし、その最大値 $M_n = \max\{X_1, \cdots, X_n\}$ を考える。このとき、a.s. には定数でない確率変数 $Y$ と定数 $c_n > 0$, $d_n \in \mathbf{R}$ が存在して、$\dfrac{M_n - d_n}{c_n} \to Y$ in law であれば、$Y$ の分布は Fréchet 分布、Gumbel 分布、Weibull 分布のいずれかであることが知られている。

**問題 3.16** $X_n$ が幾何分布 $Ge(1/n)$ に、$Y$ が指数分布 $Ex(1)$ に従うとき、$\dfrac{1}{n} X_n$ が $Y$ に法則収束することを示せ。

**問題 3.17** $X_n$ はその密度関数が $f_{X_n}(x) = 1 - (-1)^{[2nx]}$ $(0 < x < 1)$, $= 0$ (その他) で与えられる確率変数とする。このとき、$X_n$ は $(0, 1)$ 上の一様分布に法則収束するが、$0 < \forall x < 1$ に対して $f_{X_n}(x)$ は収束しないことを確認せよ。ここで、$[u]$ は $u$ の整数部分 ($u$ 以上の最小の整数) とする。

次の定理は法則収束の位相における、点列 compact 性のための必要十分条件となる。

**定理 3.27 (Prohorov の定理)** 確率変数の族 $\{X_\alpha\}$ に対して次の条件 (1), (2) は同値である。
(1) $\{X_\alpha\}$ は法則収束の定める位相について点列 compact, 即ち、$\{X_\alpha\}$ の任意の部分列 $\{X_{\alpha_n}\}$ について、さらにその部分列 $\{X_{\alpha_{n_k}}\}$ と確率変数 $X$ がとれて、$\{X_{\alpha_{n_k}}\}$ は $X$ に法則収束するようにできる。

(2) 任意の $\varepsilon > 0$ に対してある $M > 0$ があって

$$\inf_{\alpha} P(X_\alpha \in [-M, M]) \geq 1 - \varepsilon$$

とできる。すなわち、$\{X_\alpha\}$ に対応する確率測度の族が **tight** となる。

この定理を証明するために次の補題を準備する。

**補題 3.28 (Helly の選出定理)** 分布関数の列 $\{F_n(x)\}$ が与えられたとき、その部分列 $\{F_{n_k}(x)\}$ と右連続な単調増加関数 $F(x)$ が存在して ($F(x)$ は分布関数になるとは限らない)、$F$ の任意の連続点 $x$ において

$$\lim_{k \to \infty} F_{n_k}(x) = F(x) \tag{3.8}$$

とできる。

**証明:** 1st step 有理数全体 $\mathbf{Q} = \{x_1, x_2, \cdots\}$ と番号付け並べる。$F_n(x) = F_{0,n}(x)$ と書く。$\{F_{0,n}(x_1)\} \subset [0, 1]$ であるから、Bolzano-Weierstrass の定理により部分列 $\{F_{1,n}(x_1)\}$ があって $n \to \infty$ のとき $\tilde{F}(x_1)$ に収束するとできる。次に $\{F_{1,n}(x_2)\} \subset [0, 1]$ であるから、再び Bolzano-Weierstrass の定理により部分列 $\{F_{2,n}(x_2)\}$ があって $n \to \infty$ のとき $\tilde{F}(x_2)$ に収束するとできる。これを繰り返し、各 $j = 1, 2, \ldots$ に対して

- $\{F_{j+1,n}(x)\}$ は $\{F_{j,n}(x)\}$ の部分列であり
- $\{F_{j,n}(x_j)\}$ は $\tilde{F}(x_j)$ に収束する

とできる。このとき、$F_{n_k}(x) = F_{k,k}(x)$ と定めると、各 $j = 1, 2, \ldots$ に対して、$\{F_{n_k}(x_j)\}_{k \geq j}$ は $\{F_{j,n}(x_j)\}_{n \geq 1}$ の部分列であるから、$\{F_{n_k}(x_j)\}$ は $k \to \infty$ のとき $\tilde{F}(x_j)$ に収束することがわかる。(これを対角線論法という。) また、$x_i < x_j$ のとき、$F_{n_k}(x_i) \leq F_{n_k}(x_j)$ となるから $\tilde{F}(x_i) \leq \tilde{F}(x_j)$ となる。

2nd step 1st step で構成した $\tilde{F}(x)$ $(x \in \mathbf{Q})$ に対して、$F(x)$ $(x \in \mathbf{R})$ を

$$F(x) = \inf\{\tilde{F}(y); y \in \mathbf{Q} \cap (x, \infty)\} \tag{3.9}$$

とおく。このとき、$F$ が単調増加であることは明らか。また、$F(x)$ は右連続となる。(証明は各自試みよ。)
$x$ を $F$ の連続点とし、(3.8) を示す。$\varepsilon > 0$ とし、$z_1, z_2, z_3 \in \mathbf{Q}$ を

- $z_1 < z_2 < x < z_3$
- $F(x) - \varepsilon < F(z_1) \leq F(z_2) \leq F(x) \leq F(z_3) < F(x) + \varepsilon$

を満たすようにとる。これは $x$ が連続点だから可能である。しかも、(3.9) より $k \to \infty$ のとき

$$F_{n_k}(z_2) \to \tilde{F}(z_2) \geq F(z_1), \qquad F_{n_k}(z_3) \to \tilde{F}(z_3) \leq F(z_3)$$

だから、$k$ が十分大ならば

$$F(x) - \varepsilon < F_{n_k}(z_2) \leq F_{n_k}(x) \leq F_{n_k}(z_3) < F(x) + \varepsilon,$$

即ち、$|F_{n_k}(x) - F(x)| < \varepsilon$ が成立するから、(3.8) は成立する。 □

**定理 3.27 の証明:** (1) ⇒ (2) もし、(2) が成立しなければ、ある $\varepsilon > 0$ が存在して、$\forall M > 0$ に対して、

$$\inf_{\alpha} P(X_\alpha \in [-M, M]) < 1 - \varepsilon$$

とできる。すなわち、$\{X_\alpha\}$ の部分列 $\{X_{\alpha_n}\}$ があって、各 $n \in \mathbf{N}$ に対して、

$$P(X_{\alpha_n} \in [-n, n]) < 1 - \varepsilon \tag{3.10}$$

とできる。一方、(1) により、$\{X_{\alpha_n}\}$ の部分列 $\{X_{\alpha_{n_k}}\}$ と確率変数 $X$ がとれて、$\{X_{\alpha_{n_k}}\}$ は $X$ に法則収束する。よって、$x$ を $F_X$ の連続点とすると、$\lim_{k \to \infty} F_{X_{\alpha_{n_k}}}(x) = F_X(x)$ となる。しかし、$\{x_m\}$, $\{y_m\}$ を

$\lim_{m\to\infty} x_m = -\infty,\ \lim_{m\to\infty} y_m = \infty$ かつ各 $x_m, y_m$ がともに $F_X$ の連続点になるように選べば、各 $m$ に対して $k$ を十分大きくすれば $-n_k < x_m,\ y_m < n_k$ とでき、(3.10) により

$$\begin{aligned} F_X(y_m) - F_X(x_m) &= \lim_{k\to\infty}(F_{X_{\alpha_{n_k}}}(y_m) - F_{X_{\alpha_{n_k}}}(x_m)) = \lim_{k\to\infty} P(x_m < X_{\alpha_{n_k}} \le y_m) \\ &\le \liminf_{k\to\infty} P(-n_k \le X_{\alpha_{n_k}} \le n_k) \le 1-\varepsilon \end{aligned}$$

となり、$\lim_{m\to\infty}\{F_X(y_m) - F_X(x_m)\} \le 1-\varepsilon$. これは分布関数が $\lim_{y\to-\infty} F_X(y) = 0,\ \lim_{x\to\infty} F_X(x) = 1$ を満たすことに矛盾する。

<u>(2) ⇒ (1)</u>　$F_{X_\alpha}$ の任意の部分列 $F_{X_{\alpha_n}}$ が与えられたとき、Helly の選出定理 (補題 3.28) により、$F_{X_{\alpha_n}}$ の部分列 $F_{X_{\alpha_{n_k}}}$ と右連続な単調増加関数 $F$ が存在して、$F$ の任意の連続点 $x$ に対して

$$\lim_{k\to\infty} F_{X_{\alpha_{n_k}}}(x) = F(x)$$

とできる。ここで、$\varepsilon > 0$ に対して、$M > 0$ を

$$\inf_k P(X_{\alpha_{n_k}} \in [-M, M]) \ge 1-\varepsilon$$

なるようにとると、$F(x)$ の連続点 $x, y$ を $x < -M,\ M < y$ ととれば

$$\begin{aligned} F(y) - F(x) &= \lim_{k\to\infty}(F_{X_{\alpha_{n_k}}}(y) - F_{X_{\alpha_{n_k}}}(x)) = \lim_{k\to\infty} P(X_{\alpha_{n_k}} \in (x, y]) \\ &\ge \inf_k P(X_{\alpha_{n_k}} \in [-M, M]) > 1-\varepsilon \end{aligned}$$

となるので、$0 \le F \le 1$ に注意すると、これは $\lim_{x\to-\infty} F(x) = 0,\ \lim_{y\to\infty} F(y) = 1$ を意味する。よって、$F(x)$ は分布関数なので、$F_X(x) = F(x)$ となる確率変数 $X$ が存在する。　□

## 3.5　特性関数と法則収束

**定理 3.29** $\{X_n\}$ を確率変数列とし、$X_n$ の特性関数を $\phi_n(t)$ とする。このとき、
(1)　$\{X_n\}$ が $X$ に法則収束するならば、$\forall t \in \mathbf{R}$ に対して $\phi_n(t)$ は $\phi_X(t)$ に収束する。
(2)　$\forall t \in \mathbf{R}$ に対して $\phi_n(t) \to \phi(t)$ が成り立ち、$\phi(t)$ が $t = 0$ で連続ならば、$\phi(t)$ はある確率変数 $X$ の特性関数であって、$\{X_n\}$ は $X$ に法則収束する。

**証明:** (1) は $f(x) = e^{itx}$ の実部 $\cos x$, 虚部 $\sin x$ は共に有界連続関数だから、法則収束の定義より明らか。(2) を示す。

<u>1st step</u>　定理 3.27 を用いて $\{X_n\}$ が点列 compact であることを示す。$X_n$ の分布を $\mu_n$ とかく。まず、

$$|a| \ge 1 \implies |\sin a| \le |a| \cdot \sin 1 \implies \frac{\sin a}{a} \le \sin 1 \implies 1 - \frac{\sin a}{a} \ge 1 - \sin 1$$

より、$c = 1/(1-\sin 1) > 0$ とすると、$M > 0$ に対して $\frac{\sin a}{a} \le 1$ $(\forall a \in \mathbf{R})$ に注意して

$$\begin{aligned} P(|X_n| \ge M) &\le E\Big[c\Big(1 - \frac{\sin\frac{X_n}{M}}{\frac{X_n}{M}}\Big)1_{\{|\frac{X_n}{M}| \ge 1\}}\Big] \le cE\Big[\Big(1 - \frac{\sin\frac{X_n}{M}}{\frac{X_n}{M}}\Big)\Big] \\ &= c\int_{\mathbf{R}}\Big(1 - \frac{\sin\frac{x}{M}}{\frac{x}{M}}\Big)\mu_n(dx) = c\int_{\mathbf{R}}\Big(1 - \frac{1}{2}\int_{-1}^{1} e^{it\frac{x}{M}}\,dt\Big)\mu_n(dx) \\ &= c\Big(\int_{\mathbf{R}}\mu_n(dx) - \frac{1}{2}\int_{\mathbf{R}}\int_{-1}^{1} e^{ix\frac{t}{M}}\,dt\,\mu_n(dx)\Big) = c\Big(1 - \frac{1}{2}\int_{-1}^{1}\phi_n\Big(\frac{t}{M}\Big)\,dt\Big) \end{aligned}$$

とできる。ここで、2 行目の第 2 の等号は

$$\frac{1}{2}\int_{-1}^{1} e^{itx}\,dt = \Big[\frac{1}{2ix}e^{itx}\Big]_{t=-1}^{1} = \frac{e^{ix} - e^{-ix}}{2ix} = \frac{\sin x}{x},$$

3 行目の最後の等号は $|e^{ix\frac{t}{M}}| \leq 1$ に注意して Fubini の定理 (定理 2.11) を用いた。次に最後の式を $I_n(M)$ とし、$s = t/M$ と変換し、

$$\begin{aligned}
I_n(M) &= c\left(1 - \frac{1}{2}\int_{-1/M}^{1/M} \phi_n(s)M\,ds\right) = \frac{c}{2}\int_{-1/M}^{1/M}(1-\phi_n(s))M\,ds \\
&= \frac{c}{2}\int_{-1/M}^{1/M}(1-\phi(s))M\,ds + c\frac{M}{2}\int_{-1/M}^{1/M}(\phi(s)-\phi_n(s))\,ds \\
&\equiv I^{(1)}(M) + I_n^{(2)}(M)
\end{aligned}$$

とおく。$\varepsilon > 0$ が任意に与えられたとする。$\phi(s)$ は $s = 0$ で連続かつ $\phi(0) = 1$ だから、ある $M > 0$ を

$$|s| < \frac{1}{M} \text{ ならば } |\phi(s) - 1| < \frac{\varepsilon}{2c} \text{ とでき、} \qquad \left|I^{(1)}(M)\right| < \frac{\varepsilon}{2}$$

となる。この $M$ に対して、$\displaystyle\lim_{n\to\infty} \phi_n(s) = \phi(s)$ $(\forall s \in \mathbf{R})$ かつ $|\phi(s) - \phi_n(s)| \leq |\phi(s)| + |\phi_n(s)| \leq 2$ だから、Lebesgue の収束定理 (定理 2.10) により

$$\lim_{n\to\infty} I_n^{(2)}(M) = 0$$

となる。すなわち、ある $n_0$ があって、

$$n \geq n_0 \quad \Longrightarrow \quad |I_n^{(2)}(M)| < \frac{\varepsilon}{2}$$

とできる。以上より、$\inf_{n\geq n_0} P(|X_n| \leq M) \geq 1 - \varepsilon$ となるので、定理 3.27 により $\{X_n\}$ が点列 compact であることがわかった。

<u>2nd step</u> $\{X_n\}$ の任意の部分列 $\{X_{n'}\}$ が与えられたとき、1st step によりその部分列 $\{X_{n''}\}$ と確率変数 $X$ が存在して、$\{X_{n''}\}$ は $X$ に法則収束する。このとき、(1) により $\phi_{n''}(t)$ は $\phi_X(t)$ に収束するので、$\phi_X(t) = \phi(t)$ となる。もし、部分列のとり方により異なる確率変数 $Y$ に法則収束するとすれば、$\phi_X(t) = \phi_Y(t)$ となる。このとき、定理 3.16 により、$\mu_X = \mu_Y$ となる。これは、$\mu_{X_n}$ 自身が $\mu_X$ に弱収束すること、すなわち、$\{X_n\}$ 自身が $X$ に法則収束することを示している。 □

[ 1 月 13 日]

## 3.6 中心極限定理

この節では中心極限定理 (central limit theorem) を扱う。

**定理 3.30 (中心極限定理)** $X_1, X_2, \dots$ が i.i.d. で、$E[X_n] = m,\ V(X_n) = \sigma^2 > 0$ を満たすとする。このとき、

$$U_n = \frac{1}{\sqrt{n}\sigma}\sum_{k=1}^{n}(X_k - m)$$

とおけば $\{U_n\}$ は標準正規分布 $N(0,1)$ に従う確率変数 $Y$ に法則収束する。特に、次が成立する。

$$\lim_{n\to\infty} P\left(a \le U_n \le b\right) = \frac{1}{\sqrt{2\pi}}\int_a^b e^{-\frac{y^2}{2}}\,dy, \qquad -\infty < a < b < \infty. \tag{3.11}$$

**証明:** $Z_n = \dfrac{X_n - m}{\sigma}$ とすると、$E[Z_n] = 0,\ V(Z_n) = 1,\ U_n = \frac{1}{\sqrt{n}}\sum_{k=1}^{n} Z_k$ で、$U_n$ の特性関数は

$$\phi_{U_n}(t) = E[e^{i\frac{t}{\sqrt{n}}\sum_{k=1}^{n} Z_k}] = \prod_{k=1}^{n} E[e^{i\frac{t}{\sqrt{n}} Z_k}] = \Big(\phi\Big(\frac{t}{\sqrt{n}}\Big)\Big)^n$$

である。ここで、$\phi(t)$ は $Z_1$ の特性関数である。($Z_k$ は同じ分布をもつので、$\phi(t) = \phi_{Z_k}(t)$ に注意。) ここで、$E[Z_1^2] < \infty$ と命題 3.5 により、$\phi(t)$ は $C^2$-級であるから $\phi(t)$ の実部、虚部を分けて、それを $\phi_1(t), \phi_2(s)$ とし、それぞれに Taylor の定理を用いると、

$$\phi\Big(\frac{t}{\sqrt{n}}\Big) = \phi(0) + \frac{t}{\sqrt{n}}\phi'(0) + \frac{1}{2}\Big(\frac{t}{\sqrt{n}}\Big)^2\Big\{\phi_1''\Big(\theta_1\frac{t}{\sqrt{n}}\Big) + i\phi_2''\Big(\theta_2\frac{t}{\sqrt{n}}\Big)\Big\}, \qquad \theta_1, \theta_2 \in (0,1)$$

とできる。ここで、$\phi(0) = 1,\ \phi'(0) = E[Z_1] = 0$ であり、$\phi_1''(\theta_1\frac{t}{\sqrt{n}}) = -E[Z_1^2\cos\theta_1\frac{tZ_1}{\sqrt{n}}],\ \phi_2''(\theta_2\frac{t}{\sqrt{n}}) = -E[Z_1^2\sin\theta_2\frac{tZ_1}{\sqrt{n}}]$, であるが、$|Z_1^2\cos\theta\frac{tZ_1}{\sqrt{n}}| \le Z_1^2,\ |Z_1^2\sin\theta\frac{tZ_1}{\sqrt{n}}| \le Z_1^2$ であるから、Lebesgue の収束定理 (定理 2.10) により

$$\lim_{n\to\infty}\phi_1''\Big(\theta_1\frac{t}{\sqrt{n}}\Big) = -E[Z_1^2] = -1, \qquad \lim_{n\to\infty}\phi_2''\Big(\theta_2\frac{t}{\sqrt{n}}\Big) = 0.$$

したがって、

$$\lim_{n\to\infty}\Big(\phi\Big(\frac{t}{\sqrt{n}}\Big)\Big)^n = \lim_{n\to\infty}\Big(1 + \frac{1}{n}\frac{t^2}{2}\Big\{\phi_1''\Big(\theta_1\frac{t}{\sqrt{n}}\Big) + i\phi_2''\Big(\theta_2\frac{t}{\sqrt{n}}\Big)\Big\}\Big)^n = e^{-\frac{t^2}{2}}$$

となる。ここで、$Y$ を $N(0,1)$ に従う確率変数とすると、$\phi_Y(t) = e^{-\frac{t^2}{2}}$ であるから、定理 3.29(2) により、$\{U_n\}$ は $Y$ に法則収束する。(3.11) は $Y$ の分布関数 $F_Y(x)$ が連続だから、定理 3.25 により成立する。 □

**注意 3.5** 大数の弱法則や強法則は $X_1, X_2, \dots$ が同じ分布に従えば組ごとに独立であれば成立した (*cf*. 定理 2.6, 注意 2.2, 注意 2.3)。しかし、中心極限定理は組ごとに独立という仮定の下では成立しないことが知られている (*cf*. Durrett, R.: Probablity Theory and Examples, 4th ed. p.127, Example3.4.5)。

**問題 3.18** $X_n$ は二項分布 $B\Big(n, \dfrac{1}{2}\Big)$ に従うとし、$Z_n = \dfrac{X_{2n} - n}{\sqrt{n}}$ とおく。

(1) $Z_n$ の特性関数 $\phi_{Z_n}(t) = E[e^{itZ_n}]$ を余弦関数 $\cos x$ を用いた式で表せ。

(2) 定数 $v$ を適切に定め、$\{Z_n\}$ は正規分布 $N(0, v)$ に従う確率変数 $Z$ に法則収束することを示せ。

ヒント: Taylor の公式 $\cos x = 1 - \frac{1}{2}x^2\cos\theta x,\ 0 < \theta < 1$, を用いよ。

**問題 3.19** $X_1, X_2, \dots$ は i.i.d. で各 $X_k$ は一様分布 $U(-1,1)$ に従うとし、$Z_n = \dfrac{1}{\sqrt{n}} \displaystyle\sum_{k=1}^{n} X_k$ とおく。

(1) $Z_n$ の特性関数 $\phi_{Z_n}(t) = E[e^{itZ_n}]$ を正弦関数 $\sin x$ を用いた式で表せ。

(2) 定数 $v$ を適切に定め、$\{Z_n\}$ は正規分布 $N(0,v)$ に従う確率変数 $Z$ に法則収束することを示せ。

ヒント: Taylor の公式 $\sin x = x - \frac{1}{3!}x^3 \cos\theta x,\ 0 < \theta < 1,$ を用いよ。

**問題 3.20** $X_1, X_2, \dots$ は i.i.d. で、各 $X_k$ の密度関数は $f(x) = (1-|x|)1_{[-1,1]}(x)$ であるとし、$Y_n = \dfrac{1}{n} \displaystyle\sum_{k=1}^{n} X_k$ とおく。

(1) 問題 3.1(1) を用いて、$T_n = \sqrt{n} Y_n$ とするとき、$T_n$ の特性関数 $\phi_{T_n}(t) = E[e^{itT_n}]$ および $\displaystyle\lim_{n\to\infty} \phi_{T_n}(t)$ を求めよ。さらに、適切に $v$ を定め、$\{T_n\}$ は正規分布 $N(0,v)$ に従う確率変数に法則収束することを示せ。

ヒント: Taylor の公式 $\cos x = 1 - \dfrac{1}{2!}x^2 + \dfrac{1}{4!}x^4 \cos\theta x\ (0 < \theta < 1)$ を用いよ。

(3) (2) より、$n$ が十分大きければ $\sqrt{n} Y_n$ は近似的に $N(0,v)$ に従う。これを用いて、$P(|Y_n| < 0.02) \geq 0.95$ となる最小の $n$ を求めよ。ただし、標準正規分布の上側 2.5% 点は $u(0.025) = 1.960$ である。

**系 3.31 (de Moivre-Laplace の定理)** 成功率 $p$ の Bernoulli 試行列 $X_1, X_2, \dots$ を考える: $X_1, X_2, \dots$ は独立な確率変数列で $P(X_n = 1) = p, P(X_n = 0) = 1 - p$. このとき、$S_n = \sum_{k=1}^{n} X_k$ とおくと、

$$\lim_{n\to\infty} P\left(a \leq \frac{S_n - np}{\sqrt{np(1-p)}} \leq b\right) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{y^2}{2}}\, dy, \qquad -\infty < a < b < \infty. \tag{3.12}$$

**証明:** $E[X_k] = p,\ V(X_k) = p(1-p)$ であるから、定理 3.30 により従う。 □

**注意 3.6** de Moivre-Laplace の定理は $S_n$ が $B(n,p)$ に従うことから、特性関数を用いる関数解析的な方法を用いなくても、Stirling の公式を用いて直接分布を計算することで (3.12) を示すことができる。詳しくは 福島正俊著 確率論 裳華房 p.17– を参照のこと (*cf*. 問題 3.18, 3.22)。

**問題 3.21** 次の確率を中心極限定理を応用し、正規分布表を用いて求めよ。ただし、半整数補正を行うこと。

(1) 公正なサイコロを 720 回投げて、6 の目の出る回数が 130 回以上 140 回以下となる確率を求めよ。

(2) ある都市の一月あたりの事故件数は Poisson 分布 $Po(225)$ に従うという。この都市で一月に 242 件以上の事故が起こる確率を中心極限定理を応用することにより求めよ。

(3) ある都市の一月あたりの事故件数は負の二項分布 $NB\Big(150, \dfrac{5}{6}\Big)$ に従うという。この都市で一月に 41 件以上の事故が起こる確率を求めよ。

**問題 3.22** Stirling の公式を用いて次の極限を求めよ。

(1) $X_n$ が二項分布 $B\Big(n, \dfrac{1}{3}\Big)$ に従うとき、$\displaystyle\lim_{n\to\infty} \sqrt{n} P(X_{3n} = n)$.

(2) $X_n$ が負の二項分布 $NB\Big(3n+1, \dfrac{3}{4}\Big)$ に従うとき、$\displaystyle\lim_{n\to\infty} \sqrt{n} P(X_n = n)$.

[1 月 20 日]

## 3.7 多次元中心極限定理と適合度の検定

この節では $d$ 次元の中心極限定理を紹介し、その応用として適合度の検定を紹介する。

**定理 3.32** $d$ 次元確率変数 $\boldsymbol{X}_1, \boldsymbol{X}_2, \dots$ が i.i.d. で、各 $k \in \mathbf{N}$ に対し $\boldsymbol{X}_k = (X_{k,1}, \cdots, X_{k,d})'$ と表し $E[|X_{k,j}|^2] < \infty$, $j = 1, \dots, d$ と仮定する。このとき、

$$\boldsymbol{U}_n = (U_{n,1}, \cdots, U_{n,d}) \qquad U_{n,j} = \frac{1}{\sqrt{n}} \sum_{k=1}^{n} (X_{k,j} - m_j), \quad j = 1, \dots, d, \qquad n \in \mathbf{N}$$

とおけば $\{\boldsymbol{U}_n\}$ は $d$ 次元正規分布 $N(\mathbf{0}, \boldsymbol{\Sigma})$ に従う確率変数 $\boldsymbol{Y}$ に法則収束する。ただし、$\mathbf{0} = (0, \cdots, 0)'$, $m_j = E[X_{1,j}]$ であり、$\boldsymbol{\Sigma} = (\mathrm{Cov}(X_{1,i}, X_{1,j}))$ は $\boldsymbol{X}_1$ の共分散行列を表す。

**証明の概略:** §§6.2–6.6 の定理などの主張は一般の $d$ 次元に拡張できる。また、定理 3.30 と同様に

$$\lim_{n\to\infty} \phi_{\boldsymbol{U}_n}(\boldsymbol{t}) = e^{-\frac{1}{2}\boldsymbol{t}'\boldsymbol{\Sigma}\boldsymbol{t}}, \qquad \boldsymbol{t} \in \mathbf{R}^d$$

が証明でき、例 3.10 より右辺は $N(\mathbf{0}, \boldsymbol{\Sigma})$ の特性関数であるから、上記のことより成立する。 □

**注意 3.7** 例 3.10 では $\boldsymbol{\Sigma}$ は正定値対称行列であったが、定理 3.32 は $\boldsymbol{\Sigma}$ が半正定値でも成立する。

直交行列 $P = (p_{jk})$ と対角成分がすべて非負の対角行列 $D = (\lambda_{jk})$ を $P'\boldsymbol{\Sigma}P = D$ となるようにとる。このとき、$D^{1/2} = (\sqrt{\lambda_{jk}})$ とし、$A = PD^{1/2}P'$ とすると $A$ は対称行列 ($A' = A$ を満たし) で $A^2 = \boldsymbol{\Sigma}$ となる。ここで、独立に標準正規分布 $N(0,1)$ に従う確率変数 $Z_1, \dots, Z_d$ をとり、

$$\boldsymbol{Y} = A\boldsymbol{Z}, \quad \text{ただし } \boldsymbol{Z} = (Z_1, \cdots, Z_d)'$$

とする。このとき、$\phi_{\boldsymbol{Z}}(\boldsymbol{t}) = \prod_{j=1}^{d} \phi_{Z_j}(t_j) = e^{-\frac{1}{2}(t_1^2 + \cdots + t_d^2)} = e^{-\frac{1}{2}\boldsymbol{t}'\boldsymbol{t}}$, $\boldsymbol{t} = (t_1, \cdots, t_d)' \in \mathbf{R}^d$ に注意して、命題 3.9 を用いると、

$$\phi_{\boldsymbol{Y}}(\boldsymbol{t}) = \phi_{\boldsymbol{Z}}(A'\boldsymbol{t}) = e^{-\frac{1}{2}(A'\boldsymbol{t})'A'\boldsymbol{t}} = e^{-\frac{1}{2}\boldsymbol{t}'AA'\boldsymbol{t}} = e^{-\frac{1}{2}\boldsymbol{t}'\boldsymbol{\Sigma}\boldsymbol{t}}$$

となる。このとき $\boldsymbol{Y}$ は正規分布 $N(\mathbf{0}, \boldsymbol{\Sigma})$ に従うという。($A$ が $A^2 = \boldsymbol{\Sigma}$ を満たせば $\boldsymbol{Y}$ の分布は定理 3.16 より $A$ の取り方によらない。) ただし、$\boldsymbol{\Sigma}$ が正則でない場合、$\boldsymbol{Y}$ は $\mathbf{R}^d$ 上の関数としては密度関数をもたない。

**問題 3.23** 次の $\Sigma$ に対して、それぞれ $\Sigma$ を直交行列 $P$ によって対角化し、$\Sigma = A^2$ を満たす半正定値対称行列 $A$ を直交行列 $P$ を用いて定めよ。

(1) $\Sigma = \begin{pmatrix} 2 & 0 & -2 \\ 0 & 2 & 0 \\ -2 & 0 & 2 \end{pmatrix}$ (2) $\Sigma = \begin{pmatrix} 2 & 1 & -1 \\ 1 & 2 & 1 \\ -1 & 1 & 2 \end{pmatrix}$ (3) $\Sigma = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$

**定理 3.33** 確率変数 $Y_1, Y_2, \dots$ が i.i.d. で、$Y_k$ のとりうる値が $\bigcup_{j=1}^{d} A_j$ と分解できるとする。ただし、各 $A_j$ は Borel 集合で $p_j = P(Y_1 \in A_j) > 0$ とし、$i \neq j$ ならば $A_i \cap A_j = \emptyset$ とする。ただし、$\sum_{j=1}^{n} p_j = 1$ に注意する。このとき、$X_{k,j} = 1_{\{Y_k \in A_j\}}$ とし、

$$T_n = \sum_{j=1}^{d} \frac{(\sum_{k=1}^{n} X_{k,j} - np_j)^2}{np_j}$$

とおけば $\{T_n\}$ は自由度 $d-1$ の $\chi^2$ 分布 ($cf.$ 前期の定義 3.1 と定理 3.5) に法則収束する。

**証明:** <u>1st step</u>　$\widetilde{X}_{k,j}=\frac{1}{\sqrt{p_j}}X_{k,j}$ とおくと、$E[X_{k,j}]=E[X_{k,j}^2]=p_j$, $i\neq j$ ならば $E[X_{k,i}X_{k,j}]=0$ に注意して、

$$E[\widetilde{X}_{k,j}]=\sqrt{p_j},\quad V(\widetilde{X}_{k,j})=1-p_j,\quad i\neq j \text{ ならば } \mathrm{Cov}(\widetilde{X}_{k,i},\widetilde{X}_{k,j})=-\sqrt{p_ip_j}$$

を得る。よって、$\boldsymbol{U}_n=(U_{n,1},\cdots,U_{n,d})'$ を $U_{n,j}=\dfrac{1}{\sqrt{n}}\displaystyle\sum_{k=1}^n(\widetilde{X}_{k,j}-\sqrt{p_j})=\dfrac{\sum_{k=1}^n X_{k,j}-np_j}{\sqrt{np_j}}$ と定めると、定理 3.32 より $\{\boldsymbol{U}_n\}$ は $N(\boldsymbol{0},\boldsymbol{\Sigma})$ に法則収束する。ただし、$\boldsymbol{\Sigma}=\left(\delta_{ij}-\sqrt{p_ip_j}\right)_{1\leq i,j\leq d}$, また $\delta_{jj}=1$, $i\neq j$ なら $\delta_{ij}=0$ である。ここで、注意 3.7 のように $\boldsymbol{Z}$, $A$ を定めると、

$$\lim_{n\to\infty}E[\tilde{f}(\boldsymbol{U}_n)]=E[\tilde{f}(A\boldsymbol{Z})],\qquad \tilde{f}\in C_b(\mathbf{R}^d),\tag{3.13}$$

となる。ここで、$T_n=|\boldsymbol{U}_n|^2$ であるから、$\forall f\in C_b(\mathbf{R})$ に対して $\tilde{f}(x_1,\cdots,x_d)=f(\sum_{j=1}^d x_j^2)=f(|\boldsymbol{x}|^2)$ と選ぶことで、(3.13) は

$$\lim_{n\to\infty}E[f(T_n)]=E[f(|A\boldsymbol{Z}|^2)],\qquad f\in C_b(\mathbf{R}),$$

と書き換えられる。これは $\{T_n\}$ が $|A\boldsymbol{Z}|^2$ に法則収束することを意味している。

<u>2nd step</u>　$|A\boldsymbol{Z}|^2$ が自由度 $d-1$ の $\chi^2$ 分布に従うことを示す。

$\boldsymbol{\Sigma}_2=I-\boldsymbol{\Sigma}=\left(\sqrt{p_ip_j}\right)_{1\leq i,j\leq d}$ とする ($I$ は単位行列)。このとき、

$$\begin{aligned}|\boldsymbol{Z}|^2&=\boldsymbol{Z}'\boldsymbol{Z}=\boldsymbol{Z}'(\boldsymbol{\Sigma}+\boldsymbol{\Sigma}_2)\boldsymbol{Z}=\boldsymbol{Z}'\boldsymbol{\Sigma}\boldsymbol{Z}+\boldsymbol{Z}'\boldsymbol{\Sigma}_2\boldsymbol{Z}\\&=(A\boldsymbol{Z})'A\boldsymbol{Z}+\sum_{i,j=1}^d\sqrt{p_ip_j}Z_iZ_j=|A\boldsymbol{Z}|^2+(\boldsymbol{p}_1'\boldsymbol{Z})^2.\end{aligned}\tag{3.14}$$

ただし、$\boldsymbol{p}_1=(\sqrt{p_1},\cdots,\sqrt{p_d})'$ とした。次に $|\boldsymbol{p}_1|=1$ に注意して、グラム・シュミットの正規直交化法を用いて $\boldsymbol{p}_2,\ldots,\boldsymbol{p}_d$ を $P=(\boldsymbol{p}_1\cdots\boldsymbol{p}_d)$ が直交行列となるように定め、$\widetilde{\boldsymbol{Z}}=(\widetilde{Z}_1,\cdots,\widetilde{Z}_d)'=P'\boldsymbol{Z}$ とする。このとき、$\widetilde{Z}_1=\boldsymbol{p}_1'\boldsymbol{Z}$ であり、$|\widetilde{\boldsymbol{Z}}|^2=\widetilde{\boldsymbol{Z}}'\widetilde{\boldsymbol{Z}}=\boldsymbol{Z}'PP'\boldsymbol{Z}=|\boldsymbol{Z}|^2$ より、(3.14) より

$$|A\boldsymbol{Z}|^2=|\boldsymbol{Z}|^2-(\boldsymbol{p}_1'\boldsymbol{Z})^2=|\widetilde{\boldsymbol{Z}}|^2-\widetilde{Z}_1^2=\sum_{j=2}^d\widetilde{Z}_j^2$$

となる。一方、$\widetilde{\boldsymbol{Z}}$ の同時密度関数を計算することにより、$\widetilde{Z}_1,\cdots,\widetilde{Z}_d$ は独立でそれぞれ標準正規分布 $N(0,1)$ に従うことがわかる。以上より、$|A\boldsymbol{Z}|^2$ が自由度 $d-1$ の $\chi^2$ 分布に従うことがわかった。 □

**注意 3.8** 定理 3.33 は、測定値がある法則や分布に適合しているかどうかの検定に用いられる。

$n$ 個のデータが $d$ 個の階級 $A_1,\ldots,A_d$ に分類できたとき、階級 $A_j$ が出現する母比率を $p_j$ とし、次の仮説を考える。

$$\text{帰無仮説 H}_0{:}(p_1,\cdots,p_d)=(p_1^0,\cdots,p_d^0),\qquad \text{対立仮説 H}_1{:}(p_1,\cdots,p_d)\neq(p_1^0,\cdots,p_d^0)$$

このとき、階級 $A_j$ の出現回数を $n_j$ とし (定理 3.33 では $\sum_{k=1}^n X_{k,j}$ に対応する)、統計量

$$T_n=\sum_{j=1}^d\frac{(n_j-np_j^0)^2}{np_j^0}\tag{3.15}$$

を考える。このとき、もし $\mathrm{H}_0$ が正しければ $T_n$ は自由度 $d-1$ の $\chi^2$ 分布 (以下、$\chi^2_{d-1}$ 分布と略記する) に法則収束するから

$$\lim_{n\to\infty}P(T_n\leq a)=P(Y\leq a),\qquad \forall a>0,$$

となる。ただし、$Y$ は $\chi^2_{d-1}$ 分布する確率変数を表す。一方、$\mathrm{H}_1$ が正しいとき $p_j\neq p_j^0$ であれば大数の法則より $\frac{1}{n}\sum_{k=1}^n X_{k,j}$ は $p_j^0$ と異なる値に確率収束するから、

$$\frac{(n_j-np_j^0)^2}{np_j^0}=\frac{(\sum_{k=1}^n X_{k,j}-np_j^0)^2}{np_j^0}=n\frac{(\frac{1}{n}\sum_{k=1}^n X_{k,j}-p_j^0)^2}{p_j^0}\to\infty\quad\text{in prob.}$$

となり、$\lim_{n\to\infty} P(T_n \le a) = 0,\ \forall a > 0,$ となる。

以上より、もし有意水準 $\alpha$ で検定するのであれば、(3.15) に実現値を代入した値を $t$ とするとき、

$$\begin{aligned}&\text{もし } t < \chi^2_{d-1}(\alpha) \text{ であれば } \mathrm{H}_0 \text{ を採択し、}\\ &\text{もし } t \ge \chi^2_{d-1}(\alpha) \text{ であれば } \mathrm{H}_0 \text{ を棄却する。}\end{aligned}$$

ここで、$\chi^2_{d-1}(\alpha)$ は $\chi^2_{d-1}$ 分布の上側 $\alpha$ 点、即ち、$P(Y \ge k) = \alpha$ となるとき $k = \chi^2_{d-1}(\alpha)$ とした値であり、例えば [TS] 新訂 確率統計 p.162 の $\chi^2$ 分布表から求めることができる。

ただし、$np_j^0 < 5$ となる階級があるときは、隣接の階級と合併し、すべての階級で $np_j^0 \ge 5$ となるように分けなおす。

**例 3.34** あるサイコロを使ってゲームをすることになったが、その前にこのサイコロが正しく作られているか調べることにした。このサイコロを 120 回投げたとき、次の結果を得た。

| 目の数 | 1 | 2 | 3 | 4 | 5 | 6 | 計 |
|---|---|---|---|---|---|---|---|
| 出た回数 | 27 | 12 | 14 | 28 | 24 | 15 | 120 |

このサイコロの各目の出る確率は等しいか、有意水準 5% で検定せよ。

**解:** $j$ の目の出る確率を $p_j$ とし、次のように仮説を設定する。

帰無仮説 $\mathrm{H}_0 : (p_1, p_2, p_3, p_4, p_5, p_6) = \left(\frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}\right)$

対立仮説 $\mathrm{H}_1 : (p_1, p_2, p_3, p_4, p_5, p_6) \ne \left(\frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}\right)$

有意水準 5% であるから、棄却域は自由度が $6-1$ であることに注意して

$$T \ge \chi^2_5(0.05) = 11.07$$

である。ここで、実現値を代入すると $120/6 = 20$ より、

$$t = \frac{(27-20)^2}{20} + \frac{(12-20)^2}{20} + \cdots + \frac{(15-20)^2}{20} = 12.7$$

この値は棄却域に入るから、$\mathrm{H}_0$ は棄却される。従って、このサイコロは正しく作られているとはいえない。□

**問題 3.24** $\chi^2$ 分布表を用いて、以下の問いに答えよ。

(1) 子供の出生率と生まれ月の関連を調べるために、無作為標本によって次の月別の出生数を得た。

| 生まれ月 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 計 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 人数 | 90 | 89 | 81 | 85 | 92 | 93 | 102 | 109 | 116 | 101 | 94 | 88 | 1140 |

このとき、出生率は生まれ月に関連するといえるか、有意水準 5% で検定せよ。

(2) ある 1 枚の硬貨の対称性を調べるために、表が初めて出現するまで硬貨を投げ、次の結果を得た。このとき、硬貨は対称に作られているといえるか、有意水準 5% で検定せよ。

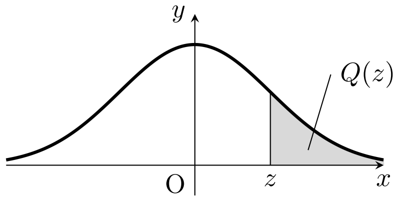
| $X$ | 1 | 2 | 3 | 4 | 5 | $\ge 6$ | 計 |
|---|---|---|---|---|---|---|---|
| 度数 | 132 | 66 | 36 | 10 | 11 | 1 | 256 |

ただし $X$ は表が出現するまでの試行回数であり $\ge 6$ は 5 回目までに表が出なかったことを意味する。

# 正規分布表、カイ 2 乗分布表

正規分布表　$Q(z) = P(Z > z),\ Z \sim N(0,1)$

Excel の cell に「`=1-NORMSDIST(z)`」として作成した。




| | 0.00 | 0.01 | 0.02 | 0.03 | 0.04 | 0.05 | 0.06 | 0.07 | 0.08 | 0.09 |
|---|---|---|---|---|---|---|---|---|---|---|
| 0.0 | 0.5000 | 0.4960 | 0.4920 | 0.4880 | 0.4840 | 0.4801 | 0.4761 | 0.4721 | 0.4681 | 0.4641 |
| 0.1 | 0.4602 | 0.4562 | 0.4522 | 0.4483 | 0.4443 | 0.4404 | 0.4364 | 0.4325 | 0.4286 | 0.4247 |
| 0.2 | 0.4207 | 0.4168 | 0.4129 | 0.4090 | 0.4052 | 0.4013 | 0.3974 | 0.3936 | 0.3897 | 0.3859 |
| 0.3 | 0.3821 | 0.3783 | 0.3745 | 0.3707 | 0.3669 | 0.3632 | 0.3594 | 0.3557 | 0.3520 | 0.3483 |
| 0.4 | 0.3446 | 0.3409 | 0.3372 | 0.3336 | 0.3300 | 0.3264 | 0.3228 | 0.3192 | 0.3156 | 0.3121 |
| 0.5 | 0.3085 | 0.3050 | 0.3015 | 0.2981 | 0.2946 | 0.2912 | 0.2877 | 0.2843 | 0.2810 | 0.2776 |
| 0.6 | 0.2743 | 0.2709 | 0.2676 | 0.2643 | 0.2611 | 0.2578 | 0.2546 | 0.2514 | 0.2483 | 0.2451 |
| 0.7 | 0.2420 | 0.2389 | 0.2358 | 0.2327 | 0.2296 | 0.2266 | 0.2236 | 0.2206 | 0.2177 | 0.2148 |
| 0.8 | 0.2119 | 0.2090 | 0.2061 | 0.2033 | 0.2005 | 0.1977 | 0.1949 | 0.1922 | 0.1894 | 0.1867 |
| 0.9 | 0.1841 | 0.1814 | 0.1788 | 0.1762 | 0.1736 | 0.1711 | 0.1685 | 0.1660 | 0.1635 | 0.1611 |
| 1.0 | 0.1587 | 0.1562 | 0.1539 | 0.1515 | 0.1492 | 0.1469 | 0.1446 | 0.1423 | 0.1401 | 0.1379 |
| 1.1 | 0.1357 | 0.1335 | 0.1314 | 0.1292 | 0.1271 | 0.1251 | 0.1230 | 0.1210 | 0.1190 | 0.1170 |
| 1.2 | 0.1151 | 0.1131 | 0.1112 | 0.1093 | 0.1075 | 0.1056 | 0.1038 | 0.1020 | 0.1003 | 0.0985 |
| 1.3 | 0.0968 | 0.0951 | 0.0934 | 0.0918 | 0.0901 | 0.0885 | 0.0869 | 0.0853 | 0.0838 | 0.0823 |
| 1.4 | 0.0808 | 0.0793 | 0.0778 | 0.0764 | 0.0749 | 0.0735 | 0.0721 | 0.0708 | 0.0694 | 0.0681 |
| 1.5 | 0.0668 | 0.0655 | 0.0643 | 0.0630 | 0.0618 | 0.0606 | 0.0594 | 0.0582 | 0.0571 | 0.0559 |
| 1.6 | 0.0548 | 0.0537 | 0.0526 | 0.0516 | 0.0505 | 0.0495 | 0.0485 | 0.0475 | 0.0465 | 0.0455 |
| 1.7 | 0.0446 | 0.0436 | 0.0427 | 0.0418 | 0.0409 | 0.0401 | 0.0392 | 0.0384 | 0.0375 | 0.0367 |
| 1.8 | 0.0359 | 0.0351 | 0.0344 | 0.0336 | 0.0329 | 0.0322 | 0.0314 | 0.0307 | 0.0301 | 0.0294 |
| 1.9 | 0.0287 | 0.0281 | 0.0274 | 0.0268 | 0.0262 | 0.0256 | 0.0250 | 0.0244 | 0.0239 | 0.0233 |
| 2.0 | 0.0228 | 0.0222 | 0.0217 | 0.0212 | 0.0207 | 0.0202 | 0.0197 | 0.0192 | 0.0188 | 0.0183 |
| 2.1 | 0.0179 | 0.0174 | 0.0170 | 0.0166 | 0.0162 | 0.0158 | 0.0154 | 0.0150 | 0.0146 | 0.0143 |
| 2.2 | 0.0139 | 0.0136 | 0.0132 | 0.0129 | 0.0125 | 0.0122 | 0.0119 | 0.0116 | 0.0113 | 0.0110 |
| 2.3 | 0.0107 | 0.0104 | 0.0102 | 0.0099 | 0.0096 | 0.0094 | 0.0091 | 0.0089 | 0.0087 | 0.0084 |
| 2.4 | 0.0082 | 0.0080 | 0.0078 | 0.0075 | 0.0073 | 0.0071 | 0.0069 | 0.0068 | 0.0066 | 0.0064 |
| 2.5 | 0.0062 | 0.0060 | 0.0059 | 0.0057 | 0.0055 | 0.0054 | 0.0052 | 0.0051 | 0.0049 | 0.0048 |
| 2.6 | 0.0047 | 0.0045 | 0.0044 | 0.0043 | 0.0041 | 0.0040 | 0.0039 | 0.0038 | 0.0037 | 0.0036 |
| 2.7 | 0.0035 | 0.0034 | 0.0033 | 0.0032 | 0.0031 | 0.0030 | 0.0029 | 0.0028 | 0.0027 | 0.0026 |
| 2.8 | 0.0026 | 0.0025 | 0.0024 | 0.0023 | 0.0023 | 0.0022 | 0.0021 | 0.0021 | 0.0020 | 0.0019 |
| 2.9 | 0.0019 | 0.0018 | 0.0018 | 0.0017 | 0.0016 | 0.0016 | 0.0015 | 0.0015 | 0.0014 | 0.0014 |
| 3.0 | 0.0013 | 0.0013 | 0.0013 | 0.0012 | 0.0012 | 0.0011 | 0.0011 | 0.0011 | 0.0010 | 0.0010 |

自由度 $n$ の $\chi^2$ 分布の上側 $\alpha$ 点: $\chi^2_n(\alpha)$

| $n$ \ $\alpha$ | 0.975 | 0.950 | 0.050 | 0.025 |
|---|---|---|---|---|
| 1 | 0.0010 | 0.0039 | 3.8415 | 5.0239 |
| 2 | 0.0506 | 0.1026 | 5.9915 | 7.3778 |
| 3 | 0.2158 | 0.3518 | 7.8147 | 9.3484 |
| 4 | 0.4844 | 0.7107 | 9.4877 | 11.1433 |
| 5 | 0.8312 | 1.1455 | 11.0705 | 12.8325 |
| 6 | 1.2373 | 1.6354 | 12.5916 | 14.4494 |
| 7 | 1.6899 | 2.1673 | 14.0671 | 16.0128 |
| 8 | 2.1797 | 2.7326 | 15.5073 | 17.5345 |
| 9 | 2.7004 | 3.3251 | 16.9190 | 19.0228 |
| 10 | 3.2470 | 3.9403 | 18.3070 | 20.4832 |
| 11 | 3.8157 | 4.5748 | 19.6751 | 21.9200 |
| 12 | 4.4038 | 5.2260 | 21.0261 | 23.3367 |
| 13 | 5.0088 | 5.8919 | 22.3620 | 24.7356 |
| 14 | 5.6287 | 6.5706 | 23.6848 | 26.1189 |
| 15 | 6.2621 | 7.2609 | 24.9958 | 27.4884 |
| 16 | 6.9077 | 7.9616 | 26.2962 | 28.8454 |
| 17 | 7.5642 | 8.6718 | 27.5871 | 30.1910 |
| 18 | 8.2307 | 9.3905 | 28.8693 | 31.5264 |
| 19 | 8.9065 | 10.1170 | 30.1435 | 32.8523 |
| 20 | 9.5908 | 10.8508 | 31.4104 | 34.1696 |
| 21 | 10.2829 | 11.5913 | 32.6706 | 35.4789 |
| 22 | 10.9823 | 12.3380 | 33.9244 | 36.7807 |
| 23 | 11.6886 | 13.0905 | 35.1725 | 38.0756 |
| 24 | 12.4012 | 13.8484 | 36.4150 | 39.3641 |
| 25 | 13.1197 | 14.6114 | 37.6525 | 40.6465 |
| 26 | 13.8439 | 15.3792 | 38.8851 | 41.9232 |
| 27 | 14.5734 | 16.1514 | 40.1133 | 43.1945 |
| 28 | 15.3079 | 16.9279 | 41.3371 | 44.4608 |
| 29 | 16.0471 | 17.7084 | 42.5570 | 45.7223 |
| 30 | 16.7908 | 18.4927 | 43.7730 | 46.9792 |
| 31 | 17.5387 | 19.2806 | 44.9853 | 48.2319 |
| 32 | 18.2908 | 20.0719 | 46.1943 | 49.4804 |
| 33 | 19.0467 | 20.8665 | 47.3999 | 50.7251 |
| 34 | 19.8063 | 21.6643 | 48.6024 | 51.9660 |
| 35 | 20.5694 | 22.4650 | 49.8018 | 53.2033 |
| 36 | 21.3359 | 23.2686 | 50.9985 | 54.4373 |
| 37 | 22.1056 | 24.0749 | 52.1923 | 55.6680 |
| 38 | 22.8785 | 24.8839 | 53.3835 | 56.8955 |
| 39 | 23.6543 | 25.6954 | 54.5722 | 58.1201 |
| 40 | 24.4330 | 26.5093 | 55.7585 | 59.3417 |

**注意.** Excel で cell に「=CHIINV($\alpha$,n)」として作成した。