確率論の基礎; 大数の法則と中心極限定理

Basics of Probability Theory; Law of Large Numbers and Central Limit Theorem

平場　誠示

2024 年 12 月 20 日

目 次

1　確率論の基本 (Basics of Probability Theory) 1

1.1　確率空間と確率変数 1

1.2　期待値, 平均値 2

1.3　大数の法則 3

1.4　大数の強法則の証明 5

1.5　特性関数と分布の収束 8

1.6　中心極限定理 12

1.7　特性関数の性質 13

1.8　Lévy の反転公式 14

1.9　Lebesgue-Stielties 測度 15

1.10　測度の弱収束 16

2　大偏差原理 (Large Deviation Principle) 19

3　測度の拡張定理と応用 (Extension Theomre & Its Applications) 22

3.1　無限次元直積確率空間 23

3.2　Kolmogorov の拡張定理 23

3.3　無限個の独立性に関する話題 25

確率論は統計学の一部であるが, 統計手法の根拠となる理論を与える学問である. その基礎となるのが, 「大数の法則」と「中心極限定理」である. 本テキストでは, それらの証明を与え, 更に, 「大偏差原理」についても言及する. 文献の [4] を元に, CLT の証明は [5] を, 大偏差原理は [1] を参考にした,

1 確率論の基本 (Basics of Probability Theory)

確率論は測度論が分からないと理解できないと言われるが, 確かにそういう面もあることは否めない. しかし扱う対象によってはそれ程, 詳しい知識が無くても理解できる分野がある. 本章では確率論を学ぶに際して必要な定義や性質について, 測度論の知識を仮定せずに理解できるよう, 最小限の事柄について解説する.

1.1 確率空間と確率変数

確率論とは, 必ず, ある適当な確率空間 $(\Omega, \mathcal{F}, P)$ があり, その上で定義された, ランダムな量＝確率変数 $X = X(\omega)$ の様々な性質 (確率) を調べて行こうとする学問である.

ここで $(\Omega, \mathcal{F}, P)$ が**確率空間 (probability space)** とは次の性質をみたすものをいう.

- $\Omega \neq \emptyset$ はある集合 (元を $\omega \in \Omega$, また, Ω の全部分集合族を 2^Ω で表す).
- $\mathcal{F}$ $(\subset 2^\Omega)$ は Ω 上の **σ-加法族 (σ-additive class)** または, **σ-加法族 (σ-field)**;

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

  (2) $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$

  (3) $A_n \in \mathcal{F}$ $(n = 1, 2, \dots) \Rightarrow \bigcup A_n \in \mathcal{F}$

- $P = P(d\omega)$ は可測空間 $(\Omega, \mathcal{F})$ 上の**確率測度 (probability measure)**, i.e., 全測度 1 の測度; $P : \mathcal{F} \to [0, 1]$ は集合関数で次をみたす.

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

  (2) $A_n \in \mathcal{F}$ $(n = 1, 2, \dots)$ が互いに素 $\Rightarrow P(\bigcup A_n) = \sum P(A_n)$ (**σ-加法性**)

$A \in \mathcal{F}$ を事象 (event) という.

問 1.1 $(\Omega, \mathcal{F}, P)$ を確率空間とするとき, 以下が成立することを示せ.

(1) σ-加法族は可算個の集合演算に関して閉じている. 即ち,
$\mathcal{F}$ を σ-加法族とし, $A, B, A_n \in \mathcal{F}$ とする. 次も $\mathcal{F}$ に属する.

$$\emptyset, \quad A \cap B, \quad A \setminus B, \quad A \triangle B := (A \setminus B) \cup (B \setminus A), \quad \bigcap_{n=1}^{\infty} A_n.$$

これから $\overline{\lim} A_n = \limsup A_n := \bigcap_{N \geq 1} \bigcup_{n \geq N} A_n$, $\underline{\lim} A_n = \liminf A_n := \bigcup_{N \geq 1} \bigcap_{n \geq N} A_n \in \mathcal{F}$.
($\overline{\lim} = \inf \sup$, $\underline{\lim} = \sup \inf$ と覚えると良い.)

(2) $P(\emptyset) = 0$, $\{A_k\}_{k=1}^n \subset \mathcal{F}$ が互いに素 $\Rightarrow P(\bigcup_{k=1}^n A_k) = \sum_{k=1}^n P(A_k)$ (**有限加法性**).

(3) $A, B \in \mathcal{F}; A \subset B \Rightarrow P(A) \leq P(B)$ (**単調性**).

(4) $A_n \in \mathcal{F}$, $A_n \uparrow \Rightarrow P\left(\bigcup A_n\right) = \lim_{n \to \infty} P(A_n)$.

(5) $A_n \in \mathcal{F}$, $A_n \downarrow \Rightarrow P\left(\bigcap A_n\right) = \lim_{n \to \infty} P(A_n)$. 上のと合わせて**確率の単調連続性**という.

(6) $A_n \in \mathcal{F}\ (n \geq 1) \Rightarrow P\left(\bigcup A_n\right) \leq \sum P(A_n)$ (**劣加法性**).

(7) (**Borel-Cantelli の補題**)
$A_n \in \mathcal{F}\ (n \geq 1)$, $\sum P(A_n) < \infty \Rightarrow P\left(\limsup_{n\to\infty} A_n\right) = 0$, i.e., $P\left(\liminf_{n\to\infty} A_n^c\right) = 1$.

この確率空間 $(\Omega, \mathcal{F}, P)$ 上で定義された関数 $X = X(\omega) : \Omega \to \mathbf{R}$ が $\{X \leq a\} := \{\omega \in \Omega; X(\omega) \leq a\} \in \mathcal{F}$ $({}^\forall a \in \mathbf{R})$. をみたすとき**確率変数 (random variable)** という. 特に X が可算個の値 $S = \{a_j\}_{j\geq 1} \subset \mathbf{R}$ しかとらないとき, この条件は $\{X = a_j\} \in \mathcal{F}$ $({}^\forall j \geq 1)$ となる.

X_k を確率空間 $(\Omega, \mathcal{F}, P)$ 上の実数値確率変数とする $(k = 1, 2, \ldots, n)$. このとき $\{X_k\}_{k=1}^n$ が**独立 (independent)** であるとは

$$P(X_1 \leq a_1, \cdots, X_n \leq a_n) = P(X_1 \leq a_1) \cdots P(X_n \leq a_n) \quad ({}^\forall a_k \in \mathbf{R}, k = 1, \ldots, n).$$

さらに上で n が無限のとき $\{X_n\}_{n\geq 1}$ が独立であるとは ${}^\forall N \geq 1$ に対して $\{X_n\}_{n=1}^N$ が独立なときをいう. 特に X_k が可算個の値 $S = \{a_j\}_{j\geq 1}$ しかとらないとき, 上の式は次のように変えても良い:

$$P(X_1 = b_1, \cdots, X_n = b_n) = P(X_1 = b_1) \cdots P(X_n = b_n) \quad (b_k \in S, k = 1, \ldots, n).$$

また $\mu(A) = \mu_X(A) = P(X \in A)$ を X の**分布 (distribution)** といい, $F(x) = F_X(x) = P(X \leq x)$ を X の**分布関数 (distribution function)** という.

1.2 期待値, 平均値

一般に確率空間 $(\Omega, \mathcal{F}, P)$ 上の確率変数 X の**期待値 (expectation)** or **平均値 (mean)** は

$$EX = E[X] := \int X dP = \int_\Omega X(\omega) P(d\omega)$$

と確率測度 P での Lebesgue 積分として定義される. しかしここではルベーグ積分論が苦手な人にも理解しやすいよう, 確率変数 X は $\overline{\mathbf{Z}} := \mathbf{Z} \cup \{\pm\infty\}$ に値をとるものとする. このとき期待値 EX は次のように定義される.

(1) $X \geq 0$ のとき $EX := \sum_{n=0}^{\infty} nP(X = n) + \infty P(X = \infty)$.
$(P(X = \infty) = 0$ なら $\infty P(X = \infty) = 0$ とする. 当然, $P(X = \infty) > 0$ なら $EX = \infty$.)

(2) X: 一般のときは $X^+ := X \vee 0$, $X^- := (-X) \vee 0$ とおき (このとき $X^\pm \geq 0$, $X = X^+ - X^-$ となる → 確かめよ.) $EX := EX^+ - EX^-$ とおく. 但し, $\infty - \infty$ となるときは定義されないとする.

この定義を形式的に表せば $EX = \sum_{n\in\overline{\mathbf{Z}}} nP(X = n)$ で, 一般の関数 $f : \overline{\mathbf{Z}} \to \mathbf{R}$ に対しても, 形式的に $Ef(X) = \sum_{n\in\overline{\mathbf{Z}}} f(n)P(X = n)$ と定義する. (厳密には上のように $\sum_{n;f(n)>0}$ と $\sum_{n;f(n)<0}$ で分けて定義する.)

確率変数 X に対して, その分散を $V(X) := E[(X - EX)^2] = E[X^2] - (EX)^2$ で定義する (最後の等号を確かめよ). これから $(EX)^2 \leq E[X^2]$ も成り立つ.

定理 1.1 (Chebyshev の不等式) $p \geq 1$ とする．このとき $^{\forall}a > 0$ に対し，

$$P(|X| \geq a) \leq \frac{E[|X|^p]}{a^p}.$$

[証明] $P(|X| \geq a) = P(|X|^p \geq a^p)$ より $p = 1$ として示せば十分.

$$E|X| = \sum_{n\geq 1} nP(|X| = n) \geq \sum_{n\geq a} nP(|X| = n) \geq a \sum_{n\geq a} P(|X| = n) = aP(|X| \geq a).$$

より一般的に,

$$E|X| = \int_\Omega |X|dP \geq \int_{\{|X|\geq a\}} |X|dP \geq aP(|X| \geq a).$$

■

定理 1.2 $X_1, \dots, X_n$ を $\overline{\mathbf{Z}}$ に値をとる確率変数として, $E[X_k^2] < \infty$ $(k = 1, \dots, n)$ とする．このとき $X_1, \dots, X_n$ が独立なら, $E[X_j X_k] = E[X_j]E[X_k]$ $(j \neq k)$．さらに平均 $E[X_k] = 0$ なら

$$E\left[\left(\sum_{k=1}^n X_k\right)^2\right] = \sum_{k=1}^n E[X_k^2].$$

[証明] (1) $j \neq k$ なら独立性から $P(X_j = m, X_k = n) = P(X_j = m)P(X_k = n)$ より

$$E[X_j X_k] = \sum_{m,n} mnP(X_j = m, X_k = n) = \sum_{m,n} mnP(X_j = m)P(X_k = n) = E[X_j]E[X_k].$$

(2) 展開式 $\left(\sum_{k=1}^n X_k\right)^2 = \sum_{k=1}^n X_k^2 + \sum_{j\neq k} X_j X_k$ と (1) より $j \neq k$ なら $E[X_j X_k] = E[X_j]E[X_k] = 0$ となることから明らか. ■

1.3 大数の法則

コイン投げで, 投げる回数を増やしていくと表の出る割合が 1/2 に近づいていくという現象がある．これが大数の法則 **(LLN=Law of Large Numbers)** の典型的な例であるが, このとき 1 回毎にコインを投げるという行為は独立である.

数式化するには n 回目にコインを投げて, 表なら $X_n = 1$, 裏なら $X_n = 0$ と決める．このとき確率平均は $EX_n = 1/2$ (ちなみに分散は $V(X_n) = 1/4$ で有界)．このとき n 回まで投げて, 表の出た回数の算術平均は $\frac{1}{n}\sum_{k=1}^n X_k$ で, 大数の法則とはこれが $n \to \infty$ のとき, 確率平均の 1/2 に「近づく」ということである.

定理 1.3 (大数の弱法則 (Weak Law of Large Numbers)) $X_1, X_2, \dots$ を独立な確率変数で，平均一定 $EX_n = m$，分散が有界 $v := \sup_n V(X_n) < \infty$ であるとする．このとき次が成り立つ：$^{\forall}\varepsilon > 0$ に対して,

$$\lim_{n\to\infty} P\left(\left|\frac{1}{n}\sum_{k=1}^n X_k - m\right| \geq \varepsilon\right) = 0, \quad \text{i.e.,} \quad \lim_{n\to\infty} P\left(\left|\frac{1}{n}\sum_{k=1}^n X_k - m\right| < \varepsilon\right) = 1.$$

[証明] $\{X_n\}$ が独立であるから $\{\widetilde{X}_n = X_n - m\}$ も独立となる (確かめよ). よって

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

から, X_n の代わりに $\widetilde{X}_n$ を考えることにより，初めから $m = 0$, i.e., $E[X_n] = 0$ として良い．このとき $V(X_n) = E[X_n^2]$ で, 前命題から

$$E\left[\left(\sum_{k=1}^{n} X_k\right)^2\right] = \sum_{k=1}^{n} E[X_k^2] = \sum_{k=1}^{n} V(X_k) \le n \sup_n V(X_n) = nv.$$

よって $^\forall\varepsilon > 0$ に対して,

$$\begin{aligned} P\left(\left|\frac{1}{n}\sum_{k=1}^{n} X_k\right| \ge \varepsilon\right) &= P\left(\left|\sum_{k=1}^{n} X_k\right| \ge \varepsilon n\right) \le \frac{E[(\sum_{k=1}^{n} X_k)^2]}{\varepsilon^2 n^2} \\ &\le \frac{nv}{\varepsilon^2 n^2} = \frac{v}{\varepsilon^2 n} \to 0 \quad (n \to \infty). \end{aligned}$$

■

上の定理と同じ条件のもとで, もっと強い結果が成り立つ. それが次の定理である.

定理 1.4 (大数の強法則 (Strong Law of Large Numbers)) $X_1, X_2, \dots$ を独立な確率変数で，平均一定 $EX_n = m$，分散が有界 $v := \sup_n V(X_n) < \infty$ であるとする．このとき次が成り立つ：

$$P\left(\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} X_k = m\right) = 1.$$

注意 1.1 一般に, 平均が一定でない場合でも, 次が成り立つ.

$$P\left(\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n}(X_k - EX_k) = 0\right) = 1.$$

これの証明は簡単ではないが, 次の部分節で述べる. もう少し強い条件のもとでは次のように簡単に証明できる.

[$\sup E[X_n^4] < \infty$ **のもとでの定理 1.4 の証明**] X_n の代わりに $X_n - m$ を考えることにより, $m = 0$, i.e., $E[X_n] = 0$ として示せばよい．まず $\left(\sum_{k=1}^{n} X_k\right)^4$ の展開式を考えるのだが，独立性と平均が 0, さらに $E[X^2] \le (E[X^4])^{1/2}$ に注意して,

$$E\left[\left(\sum_{k=1}^{n} X_k\right)^4\right] = \sum_{k=1}^{n} E[X_k^4] + \sum_{i\ne j, 1\le i,j\le n} E[X_i^2]E[X_j^2] \le n^2 \sup_k E[X_k^4]$$

をえるから，単調収束定理 (or Fubini の定理) を用いて

$$E\left[\sum_{n=1}^{\infty}\left(\frac{1}{n}\sum_{k=1}^{n} X_k\right)^4\right] = \sum_{n=1}^{\infty} \frac{1}{n^4} E\left[\left(\sum_{k=1}^{n} X_k\right)^4\right] \le \sum_{n=1}^{\infty} \frac{1}{n^2} \sup_k E[X_k^4] < \infty$$

をえる．これは $P\left(\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} X_k = 0\right) = 1$ を意味する. ■

さらに重要な話題として, 次の中心極限定理 **(CLT=Central Limit Theorem)** がある.

定理 1.5 (CLT) 確率変数列 $\{X_n\}$ を独立同分布 (independent identically distributed = i.i.d.) とする. その平均を $EX_1 = m$, 分散を $V(X_1) = v$ とすると $\dfrac{1}{\sqrt{n}}\sum_{k=1}^{n}(X_k - m)$ の分布は平均 0, 分散 v の正規分布 $N(0,v)$ に収束する, i.e., 任意の $a < b$ に対し,

$$\lim_{n\to\infty} P\left(a < \frac{1}{\sqrt{n}}\sum_{k=1}^{n}(X_k - m) \le b\right) = \frac{1}{\sqrt{2\pi v}}\int_a^b e^{-\frac{x^2}{2v}}\,dx.$$

言い換えれば, $\dfrac{1}{\sqrt{nv}}\sum_{k=1}^{n}(X_k - m)$ の分布は平均 0, 分散 1 の正規分布 $N(0,1)$ に収束する,

ここで, 独立性と分布の性質について少し述べておく. $\mathcal{B}^1 = \mathcal{B}(\mathbf{R}^1)$ を 1 次元 Borel 集合体とする. 実確率変数列 $X_1, \dots, X_n$ に対し, $\boldsymbol{X} = (X_1, \dots, X_n)$ として, その結合分布を $\mu_{\boldsymbol{X}}(A_1 \times \cdots \times A_n) = P(X_1 \in A_1, \dots, X_n \in A_n)$ $(A_i \in \mathcal{B}^1)$ と定義する.

定理 1.6 実確率変数列 $X_1, \dots, X_n$ が独立なら, $\boldsymbol{X} = (X_1, \dots, X_n)$ として,

$$\mu_{\boldsymbol{X}} = \bigotimes_{i=1}^{n}\mu_{X_i} \quad \text{i.e.,} \quad \mu_{\boldsymbol{X}}(A_1 \times \cdots \times A_n) = \mu_{X_1}(A_1)\cdots\mu_{X_n}(A_n).$$

これは半直線 $(-\infty, a]$ の全体で生成される σ 加法族が $\mathcal{B}^1$ となることから容易に分かる.

定理 1.7 実確率変数 X, Y が独立なら, 有界 Borel 関数 $f(x,y)$ に対し,

$$E[f(X,Y)] = E\left[E[f(x,Y)]|_{x=X}\right] = E\left[E[f(X,y)]|_{y=Y}\right].$$

これは分布を用いて表せば,

$$\int_{\mathbf{R}^2} f(x,y)\mu_{(X,Y)}(dx,dy) = \int_{\mathbf{R}^2} f(x,y)\mu_X(dx)\mu_Y(dy)$$

となるので, 上の結果から明らか.

例 1.1 実確率変数 X, Y が独立なら,

$$P(X < Y) = \int_{\mathbf{R}} P(x < Y)\mu_X(dx).$$

1.4 大数の強法則の証明

2 つの収束概念について述べておく.

一般に確率変数 X_n, X に対して, $^\forall\varepsilon > 0$, $P(|X_n - X| \ge \varepsilon) \to 0$ $(n \to \infty)$ のとき, $X_n \to X$ in pr. と表し, X_n が X に**確率収束**するという. また $P(X_n \to X) = 1$ のとき, $X_n \to X$, P-a.s. と表し, X_n が X に**概収束**するという. (a.s. は almost surely の略)

問 1.2 "**概収束** $\Longrightarrow$ **確率収束**", i.e., $X_n \to X$, P-a.s. $\Longrightarrow X_n \to X$ in pr. を示せ.

(**ヒント** $P(X_n \to X) = 1 \iff$

$$P\left(\bigcap_{k\ge 1}\bigcup_{N\ge 1}\bigcap_{n\ge N}\left\{|X_n - X| < \frac{1}{k}\right\}\right) = 1 \iff P\left(\bigcup_{k\ge 1}\bigcap_{N\ge 1}\bigcup_{n\ge N}\left\{|X_n - X| \ge \frac{1}{k}\right\}\right) = 0$$

$$\iff {}^{\forall}k \ge 1, \lim_{N\to\infty} P\left(\bigcup_{n\ge N}\left\{|X_n - X| \ge \frac{1}{k}\right\}\right) = P\left(\bigcap_{N\ge 1}\bigcup_{n\ge N}\left\{|X_n - X| \ge \frac{1}{k}\right\}\right) = 0$$

$$\Longrightarrow {}^{\forall}k \ge 1, \lim_{N\to\infty} P\left(|X_N - X| \ge \frac{1}{k}\right) \le \lim_{N\to\infty} P\left(\bigcup_{n\ge N}\left\{|X_n - X| \ge \frac{1}{k}\right\}\right) = 0$$

これは確率収束と同値 (即ち, $1/k$ を $\varepsilon > 0$ に置き換えられる. 何故か？)

注意 1.2 一般に, 上の問の逆は成り立たない. 即ち, 確率収束しても概収束しない例が作れる. しかし, 確率収束していれば, 適当な部分列が概収束するようにとれることは知られている.

問 1.3 「確率収束なら適当な部分列に対して概収束」, i.e., ”$X_n \to X$ in pr. $\Rightarrow {}^{\exists}\{n_k\}; X_{n_k} \to X$ a.s.“を示せ.

(**ヒント** 確率収束の仮定より, 次が分る (何故か?) ${}^{\exists}\{n_k\};\ P\left(|X_{n_k} - X| \ge \frac{1}{2^k}\right) \le \frac{1}{2^k}$. 和が収束することから Borel-Cantelli の補題が使えて $P\left(\bigcup_{N\ge 1}\bigcap_{k\ge N}\left\{|X_{n_k} - X| < \frac{1}{2^k}\right\}\right) = 1$. これから容易に分る.)

さてこれから大数の強法則の話に入ろう. 一応, 再び, 定理を述べておく.

定理 1.8 (大数の強法則 (Strong Law of Large Numbers)) $X_1, X_2, \ldots$ を独立な確率変数で，平均一定 $EX_n = m$，分散が有界 $v := \sup_n V(X_n) < \infty$ であるとする．このとき次が成り立つ：

$$P\left(\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n} X_k = m\right) = 1.$$

[証明の流れ] まず $EX_n = 0$ として示せば良く, $\overline{S}_n = \sum_{k=1}^{n}(X_k/k)$ に対し,

(1) **Kolmogorov の最大不等式**より $\sup_{k\ge n}|\overline{S_k} - \overline{S_n}| \to 0\ (n\to\infty)$ in pr.

(2) 「確率収束なら適当な部分列に対して概収束」を用いれば,
確率 1 で $\{\overline{S_n}\}$ が Cauchy 列, 故に収束列.

(3) **Kronecker の補題**より $\frac{1}{n}\sum_{k=1}^{n} X_k \to 0$ P-a.s.

という手順で示す. そのため Kronecker の補題と Kolmogorov の最大不等式を先に与え, 証明する.

補題 1.1 (Kronecker の補題) 数列 $\{x_n\}$ と $\{a_n\}; 0 < a_n \uparrow \infty$ に対し,

$$\lim_{n\to\infty}\sum_{k=1}^{n}\frac{x_k}{a_k} \quad \text{exists} \quad \Longrightarrow \quad \lim_{n\to\infty}\frac{1}{a_n}\sum_{k=1}^{n} x_k = 0$$

[証明] $s_0 = 0$, $s_n = \sum_{k=1}^{n}(x_k/a_k) \to s$ とする.

$$\frac{1}{a_n}\sum_{k=1}^{n} x_k = \sum_{k=1}^{n}\frac{a_k}{a_n}\frac{x_k}{a_k} = \sum_{k=1}^{n}\frac{a_k}{a_n}(s_k - s_{k-1}) = s_n - \sum_{k=1}^{n-1}\frac{a_{k+1}-a_k}{a_n}s_k.$$

結局, 題意は

$$s_n \to s \quad \text{なら} \quad \frac{1}{a_n}\sum_{k=1}^{n-1}(a_{k+1}-a_k)s_k \to s$$

の証明に帰着する. これは $s^* = \sup_m |s_m| < \infty$ と $^\forall\varepsilon > 0, {}^\exists N; {}^\forall k \geq N, |s_k - s| < \varepsilon$ より, $n > N$ に対し, 和を N で分けて

$$\begin{aligned}
&\left|\frac{1}{a_n}\sum_{k=1}^{n-1}(a_{k+1}-a_k)s_k - s\right| \quad \left(s = \frac{1}{a_n}\sum_{k=N}^{n-1}(a_{k+1}-a_k)s + \frac{a_N}{a_n}s \quad \text{として}\right)\\
\leq\ & \frac{1}{a_n}\sum_{k=N}^{n-1}(a_{k+1}-a_k)|s_k - s| + \frac{1}{a_n}\sum_{k=1}^{N-1}(a_{k+1}-a_k)(\sup_m |s_m|) + \frac{a_N}{a_n}|s|\\
\leq\ & \varepsilon\frac{a_n - a_N}{a_n} + s^*\frac{a_N - a_1}{a_n} + \frac{a_N}{a_n}|s|\\
\to\ & \varepsilon \quad (n \to \infty)
\end{aligned}$$

よって $\varepsilon > 0$ の任意性から極限は 0. ■

補題 1.2 (Kolmogorov の最大不等式) $\{X_n\}$ を独立な確率変数列で, 平均 $EX_n = 0$ とする. 部分和 $S_n = \sum_{k=1}^{n} X_n$ に対し,

$$a > 0 \quad \Longrightarrow \quad a^2 P(\max_{1\leq n\leq N}|S_n| \geq a) \leq E[|S_N|^2; \max_{1\leq n\leq N}|S_n| \geq a] \leq E[|S_N|^2]$$

[証明] $A_k = \{|S_k| \geq a, |S_1| < a, \ldots, |S_{k-1}| < a\}$, $S^{(k+1)} = X_{k+1} + \cdots + X_N$ とおくと, $S^{(k+1)}$ と $S_k 1_{A_k}$ は独立で $E[S_k S^{(k+1)}; A_k] = E[S_k 1_{A_k}]E[S^{(k+1)}] = 0$. また $A = \bigcup_{k=1}^{N} A_k$ (素和). よって

$$\begin{aligned}
E[|S_N|^2; \max_{1\leq n\leq N}|S_n| \geq a] &= \sum_{k=1}^{N} E[(S_k + S^{(k+1)})^2; A_k]\\
&= \sum_{k=1}^{N} E[S_k^2 + 2S_k S^{(k+1)} + (S^{(k+1)})^2; A_k]\\
&\geq \sum_{k=1}^{N} E[S_k^2; A_k]\\
&\geq \sum_{k=1}^{N} a^2 P(A_k) \quad (A_k \text{ 上 } |S_k| \geq a \text{ より})\\
&= a^2 P(\max_{1\leq n\leq N}|S_n| \geq a)
\end{aligned}$$

■

[大数の強法則 (定理 1.4) の証明] まず X_n の代わりに $\widetilde{X}_n = X_n - EX_n$ を考えると $\{\widetilde{X}_n\}$ も独立で $V(\widetilde{X}_n) = V(X_n) \leq v$ より, 初めから $EX_n = 0$ として示せば十分. このとき独立性から

$E[X_nX_m] = E[X_n]E[X_m] = 0$ $(m \neq n)$, また $E[X_n^2] = V(X_n) \le v$. そこで $\overline{S}_n = \sum_{k=1}^n \frac{X_k}{k}$ とおくと, Kolmogorov の最大不等式より, 任意の $a > 0$ に対し,

$$a^2 P(\max_{n<k\le N} |\overline{S_k} - \overline{S_n}| \ge a) \le E[|\overline{S}_N - \overline{S_n}|^2] = \sum_{k=n+1}^N \frac{E[X_k^2]}{k^2} \le \sum_{k>n} \frac{v}{k^2}.$$

順に $N \to \infty, n \to \infty$ として,

$$\lim_{n\to\infty} P(\sup_{k>n} |\overline{S_k} - \overline{S_n}| \ge a) = 0, \quad \text{i.e.,} \quad \sup_{k>n} |\overline{S_k} - \overline{S_n}| \to 0 \ (n \to \infty) \text{ in pr.}$$

よって適当な部分列 $\{n_j\} \subset \mathbf{N}; n_j \uparrow \infty$ をとれば,

$$\lim_{j\to\infty} \sup_{k\ge n_j} |\overline{S_k} - \overline{S}_{n_j}| = 0 \quad P\text{-a.s.}$$

これから $n, m \ge n_j$ なら $|\overline{S_n} - \overline{S_m}| \le |\overline{S_n} - \overline{S}_{n_j}| + |\overline{S_m} - \overline{S}_{n_j}| \to 0$ $(j \to \infty)$ P-a.s., 即ち, $\{\overline{S_n}\}$ は Cauchy 列となり, $\lim_{n\to\infty} \sum_{k=1}^n \frac{X_k}{k} = \lim_{n\to\infty} \overline{S_n}$ が確率 1 で存在する. 従って Kronecker の補題を用いれば, $\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^n X_k = 0$ P-a.s. をえる. ■

上の証明から次が成り立つ.

系 1.1 $\{X_n\}$ を平均 0 の独立な確率変数列とする. $\sum_{k=1}^\infty V(X_k) < \infty$ なら極限 $\lim_{n\to\infty} \sum_{k=1}^n X_k$ は確率 1 で存在する.

補足 $\{X_n\}$ に独立同分布を仮定すると, 可積分 $E|X_1| < \infty$ という条件の下で, LLN は証明できる. (文献 [2], [3] 参照)

系 1.2 大数の法則の定理の条件の下, 任意の $\delta > 0$ に対し, 次も成り立つ.

$$\lim_{n\to\infty} \frac{1}{\sqrt{n^{1+\delta}}} \sum_{k=1}^n (X_k - EX_k) = 0 \quad P\text{-a.s.}$$

証明は大数の法則の証明で $\overline{S_n}$ として $\sum_{k=1}^n (X_k/\sqrt{k^{1+\delta}})$ を考えれば良い.

では上で δ を 0 としたときにはどうなるのだろうか？この疑問に対する (適当な条件の下での) 答えを与えるのが次々節の中心極限定理であるが, その証明には. 特性関数と呼ばれる, 確率測度の Fourier 変換を用いて, それが収束すると分布も収束するという事実が用いられる. そこで先に特性関数について述べ, その収束と分布の収束に関する定理を述べてから, 中心極限定理とその証明を与えようと思う.

1.5 特性関数と分布の収束

確率変数 X に対し, 次の関数 $\varphi = \varphi_X : \mathbf{R}^1 \to \mathbf{C}$ を X の**特性関数 (c.f.=characteristic function)** という.

$$\varphi(z) = \varphi_X(z) := E\left[e^{izX}\right] \quad (z \in \mathbf{R}^1).$$

またこのとき X の**分布 (distribution)** $\mu(A) = \mu_X(A) := P(X \in A)$ に対し, 次のようにも表される.

$$\varphi(z) = \int_{\mathbf{R}} e^{izx} \mu(dx)$$

また $\mathbf{R}^1$ 上の確率測度 μ が与えられたとき (これを単に**分布 (dist.)** というが), 上で定義される $\varphi(z)$ を μ の特性関数という.

まず正規分布を定義しておく.

$\mathbf{R}$ 上の分布 $\mu(dx) = g(x)dx$ が

$$g(x) = \frac{1}{\sqrt{2\pi v}} \exp\left[-\frac{(x-m)^2}{2v}\right]$$

で与えられるとき, これを平均 m, 分散 v の**正規分布 (normal dist.)**, あるいは**ガウス分布 (Gauss dist.)** といい, それを表す記号として $N(m, v)$ を用いる (正規分布をもつ確率変数として用いることもある).

この分布の特性関数は次のようになる.

$$\varphi(z) = \int_{-\infty}^{\infty} e^{izx} \frac{1}{\sqrt{2\pi v}} \exp\left[-\frac{(x-m)^2}{2v}\right] dx = \exp\left[imz - \frac{vz^2}{2}\right].$$

問 1.4 上の特性関数の計算を確かめよ.

テント関数 区間 $(-1, 1)$ の外では 0 で, 間は 3 点 $(-1, 0), (0, 1), (1, 0)$ を線分で結んだグラフをもつ関数を $T(x)$ とする, i.e.,

$$T(x) = \frac{1}{2}(|x+1| + |x-1| - 2|x|).$$

また $-\infty < a < b < \infty, h > 0$ に対し, 区間 (a, b) 上の高さ h の**テント関数**を

$$T_{a,b;h}(x) = hT\left(\frac{2}{b-a}\left(x - \frac{a+b}{2}\right)\right)$$

とおく. さらに高さ $h > 1$ のテント関数から高さ 1 以上の部分を切り取ってできる台形関数 $D_{a,b;h}$ を次で定める.

$$D_{a,b;h}(x) = \min\{T_{a,b;h}(x), 1\} = T_{a,b;h}(x) - T_{a+(b-a)/(2h), b-(b-a)/(2h); h-1}(x).$$

このテント関数は次のような分布に現れる.

問 1.5 U, V を独立な確率変数でともに $[0, a]$ $(a > 0)$ 上の一様分布に従うとき, $X = U - V$ の分布の密度関数は $T_{-a,a;1/a}$ となることを示せ. またその特性関数は次で与えれることを示せ.

$$\varphi_X(z) = \frac{2(1 - \cos az)}{a^2 z^2}.$$

（ヒント） 有界な Borel 関数 f に対し,

$$E[f(X)] = \int_{-a}^{a} f(x) T_{-a,a;1/a}(x) dx$$

を示せば良い．このとき, (U,V) の結合分布が独立性から, それぞれの分布の積になることを用いる．即ち,

$$P(U\in du, V\in dv)=P(U\in du)P(V\in dv)=\frac{1}{a}1_{[0,a]}(u)du\frac{1}{a}1_{[0,a]}(v)dv.$$

これにより, 上の計算は次に帰着する．

$$\frac{1}{a^2}\int_{\mathbf{R}}1_{[0,a]}(v)1_{[0,a]}(x+v)dv=T_{-a,a;1/a}(x)$$

後半は, $U,-V$ の特性関数の積となることを用いる．

命題 1.1 確率変数 X の特性関数 $\varphi(z)$ に対し,

$$E[T(X)]=\frac{1}{\pi}\int_{-\infty}^{\infty}\varphi(z)\frac{1-\cos z}{z^2}dz,$$

$$E[T_{a,b;h}(X)]=\frac{2h}{\pi(b-a)}\int_{-\infty}^{\infty}\varphi(z)e^{-i(a+b)z/2}\frac{1-\cos\frac{(b-a)z}{2}}{z^2}dz.$$

証明 まず前問より,

$$\int_{-\infty}^{\infty}e^{izx}T(x)dx=\frac{2(1-\cos z)}{z^2}.$$

さらに次が成り立つことが示せる．(要は、Fourier 逆変換)

$$\int_{-\infty}^{\infty}e^{izx}\frac{1-\cos z}{z^2}dz=\pi T(x). \tag{1.1}$$

これを認めれば, x へ X を代入し, 期待値をとると, Fubini の定理を用いて

$$E[T(X)]=\frac{1}{\pi}E\left[\int_{-\infty}^{\infty}e^{izX}\frac{1-\cos z}{z^2}dz\right]=\frac{1}{\pi}\int_{-\infty}^{\infty}\varphi_X(z)\frac{1-\cos z}{z^2}dz.$$

後半の $E[T_{a,b;h}(X)]$ については変数変換により, 容易に示せる．最後に (1.1) を示そう．$(1-\cos z)/z^2$ が偶関数であることと

$$\cos zx(1-\cos z)=\cos zx-\frac{1}{2}(\cos z(x+1)+\cos z(x-1))$$

と変数変換により, (1.1) の左辺は次のようになる．

$$\begin{aligned}\int_{-\infty}^{\infty}\cos zx\frac{1-\cos z}{z^2}dz &= \frac{1}{2}\int_{-\infty}^{\infty}\frac{1-\cos z(x+1)}{z^2}dz+\frac{1}{2}\int_{-\infty}^{\infty}\frac{1-\cos z(x-1)}{z^2}dz\\ &\quad -\int_{-\infty}^{\infty}\frac{1-\cos zx}{z^2}dz\\ &= \int_{-\infty}^{\infty}\frac{1-\cos z}{z^2}dz\left(\frac{1}{2}(|x+1|+|x-1|)-|x|\right).\end{aligned}$$

これと等式 $\displaystyle\int_{-\infty}^{\infty}\frac{1-\cos z}{z^2}dz=\pi$ から (1.1) を得る． ■

問 1.6 (i) 部分積分を用いて積分 $I(t)=\displaystyle\int_0^{\infty}e^{-tz}\sin z dz \quad (t>0)$ を求めよ． $1/(1+t^2)$

(ii) 等式 $\displaystyle\int_0^{\infty}I(t)dt=\int_0^{\infty}\frac{\sin z}{z}dz$ を示し, その積分の値を求めよ． $\pi/2$

(iii) 部分積分を用いて証明の最後の等式 $\displaystyle\int_{-\infty}^{\infty}\frac{1-\cos z}{z^2}dz=\pi$ を確かめよ．

確率変数 X, Y が同分布であるとは任意の $a \in \mathbf{R}$ に対し, $P(X > a) = P(Y > a)$ が成り立つときをいい, 記号で $X \overset{(\mathrm{d})}{=} Y$ と表す. ($X = Y$ in the sense of distribution の意)

定理 1.9 確率変数 X, Y の特性関数 φ_X, φ_Y に対し, $\varphi_X(z) = \varphi_Y(z)$ ($z \in \mathbf{R}$) なら X と Y の分布は一致する, i.e., $X \overset{(\mathrm{d})}{=} Y$.

証明 仮定と前の命題から任意のテント関数 $T_{a,b;h}$ に対し, $E[T_{a,b;h}(X)] = E[T_{a,b;h}(Y)]$, 故に任意の台形関数 $D_{a,b;h}$ に対しても $E[D_{a,b;h}(X)] = E[D_{a,b;h}(Y)]$. そこで $\lim_{h\to\infty} D_{a,b;h}(x) = I_{(a,b)}(x)$ に注意して Lebesgue の収束定理より $P(a < X < b) = P(a < Y < b)$. $b \to \infty$ として $X \overset{(\mathrm{d})}{=} Y$ を得る. ■

定理 1.10 確率変数 X と確率変数列 $\{X_n\}$ の特性関数をそれぞれ $\varphi(z)$, $\{\varphi_n(z)\}$ とする.

$$\lim_{n\to\infty} \varphi_n(z) = \varphi(z) \quad (z \in \mathbf{R}^1) \quad [\text{各点収束}]$$

なら, 任意の $a \in \mathbf{R}; P(X = a) = 0$ に対し, $\lim_{n\to\infty} P(X_n > a) = P(X > a)$.

証明 仮定と $|\varphi_n(z)| \le 1$, さらに Lebesgue の収束定理により,

$$\lim_{n\to\infty} \int_{-\infty}^{\infty} \varphi_n(z) e^{-i(a+b)z/2} \frac{1 - \cos((b-a)z/2)}{z^2} dz = \int_{-\infty}^{\infty} \varphi(z) e^{-i(a+b)z/2} \frac{1 - \cos((b-a)z/2)}{z^2} dz.$$

従って命題 1.1 より, $\lim_{n\to\infty} E[T_{a,b;h}(X_n)] = E[T_{a,b;h}(X)]$. よって任意の台形関数 $D_{a,b;h}$ に対しても $\lim_{n\to\infty} E[D_{a,b;h}(X_n)] = E[D_{a,b;h}(X)]$. さらに $h > 1, a < b$ に対し,

$$I_{(a,b)}(x) \ge D_{a,b;h}(x) \ge I_{[a+(b-a)/(2h), b-(b-a)/(2h)]}(x) \quad (x \in \mathbf{R})$$

に注意すると

$$\begin{aligned} \liminf_{n\to\infty} P(a < X_n < b) &\ge \lim_{n\to\infty} E[D_{a,b;h}(X_n)] \\ &= E[D_{a,b;h}(X)] \ge P\left(a + \frac{b-a}{2h} \le X \le b - \frac{b-a}{2h}\right). \end{aligned}$$

ここで $h \to \infty, b \to \infty$ とすれば, $^{\forall} a \in \mathbf{R}$ に対し,

$$\liminf_{n\to\infty} P(X_n > a) \ge P(X > a).$$

また $h \to \infty, a \to -\infty$ とし, b を a に直せば, $^{\forall} a \in \mathbf{R}$ に対し, $\liminf_{n\to\infty} P(X_n < a) \ge P(X < a)$. これからさらに $\limsup_{n\to\infty} P(X_n > a) \le 1 - \liminf_{n\to\infty} P(X_n < a) \le 1 - P(X < a) = P(X \ge a)$ となるので $^{\forall} a \in \mathbf{R}; P(X = a) = 0$ に対し,

$$\limsup_{n\to\infty} P(X_n > a) \le P(X > a).$$

従って $\lim_{n\to\infty} P(X_n > a) = P(X > a)$ を得る. ■

1.6 中心極限定理

定理 1.11 (CLT) 確率変数列 $\{X_n\}$ を独立同分布 (i.i.d.) とする. 平均を $EX_1 = m$, 分散を $V(X_1) = v$ とすると $\dfrac{1}{\sqrt{nv}}\displaystyle\sum_{k=1}^{n}(X_k - m)$ の分布は平均 0, 分散 1 の正規分布 $N(0,1)$ に収束する, i.e., 任意の $a < b$ に対し,

$$\lim_{n\to\infty} P\left(a < \frac{1}{\sqrt{nv}}\sum_{k=1}^{n}(X_k - m) \le b\right) = \frac{1}{\sqrt{2\pi}}\int_a^b e^{-\frac{x^2}{2}}\,dx.$$

まず証明の前に必要な補題を与えておく.

補題 1.3 $EX = 0, V(X) = E(X^2) = 1$ なる確率変数 X に対し,

$$\varphi_X\left(\frac{z}{\sqrt{n}}\right) - \left(1 - \frac{z^2}{2n}\right) = o\left(\frac{1}{n}\right) \quad (n \to \infty).$$

証明 $g(z)$ を

$$e^{iz} - 1 - iz + \frac{z^2}{2} = z^2 g(z)$$

で定義すると $|g(z)| \le 1$, $\displaystyle\lim_{z\to 0} g(z) = 0$ をみたす. 実際, テイラーの定理により,

$$^{\exists}\theta \in (0,1); e^{iz} - 1 - iz = -\frac{z^2}{2}e^{i\theta z}$$

を用いれば $|g(z)| \le 1$, $\to 0$ $(z \to 0)$ は容易に分かる. そこで

$$\exp\frac{izX}{\sqrt{n}} = 1 + \frac{izX}{\sqrt{n}} - \frac{z^2X^2}{2n} + \frac{z^2X^2}{n}g\left(\frac{zX}{\sqrt{n}}\right)$$

において両辺の期待値をとれば,

$$\varphi_X\left(\frac{z}{\sqrt{n}}\right) = 1 - \frac{z^2}{2n} + E\left[\frac{z^2X^2}{n}g\left(\frac{zX}{\sqrt{n}}\right)\right].$$

ここで 最後の期待値については

$$X^2\left|g\left(\frac{zX}{\sqrt{n}}\right)\right| \le X^2, \quad \lim_{n\to\infty} g\left(\frac{zX}{\sqrt{n}}\right) = 0$$

より, Lebesgue の収束定理が適応できて, $n \to \infty$ で 0 に収束する. 従って求める結果を得る. ■

[CLT の証明]

$\widetilde{X}_n = (X_n - m)/\sqrt{v}$ とすると $E\widetilde{X}_n = 0, V(\widetilde{X}_n) = 1$ で $\{\widetilde{X}_n\}$ は i.i.d. となるので, $m = 0, v = 1$ のときに示せば良い. $Y_n := (\sum_{k=1}^{n} X_k)/\sqrt{n}$ に対し, その特性関数は $\{X_k\}$ が i.i.d. であることから,

$$\varphi_n(z) = E\left[\exp\left(\frac{iz}{\sqrt{n}}\sum_{k=1}^{n}X_k\right)\right] = \prod_{k=1}^{n}\varphi_{X_k}\left(\frac{z}{\sqrt{n}}\right) = \varphi_{X_1}\left(\frac{z}{\sqrt{n}}\right)^n.$$

さらに上の補題により, 各 $z \in \mathbf{R}$ に対し,

$$\lim_{n\to\infty}\varphi_n(z) = \lim_{n\to\infty}\left(1 - \frac{z^2}{2n} + o\left(\frac{1}{n}\right)\right)^n = \exp[-z^2/2].$$

ここで最後の等号については

$$\left(1-\frac{z^2}{2n}+o\left(\frac{1}{n}\right)\right)^n=\left(1-\frac{z^2}{2n}\right)^n+R_n(z)$$

で $R_n(z)$ を定義すると $|R_n(z)|=o(1)$ $(n\to\infty)$ が示せる (下の問). 従って, $\varphi_n(z)$ が正規分布 $N(0,1)$ の特性関数 $\varphi(z)=\exp[-z^2/2]$ に各点収束するので前定理 (定理 1.10) より, 正規分布が質点を持たないことと併せて, 定理の証明が終わる. ■

問 1.7 証明の最後の $|R_n(z)|=o(1)$ $(n\to\infty)$ を示せ.

1.7 特性関数の性質

命題 1.2 $\mathbf{R}$ 上の分布 μ の特性関数 $\varphi=\varphi_\mu$ に対し, 次が成り立つ.

(1) $\varphi(0)=1$, $|\varphi(z)|\le 1$, $\overline{\varphi(z)}=\varphi(-z)$.

(2) φ は $\mathbf{R}$ 上の一様連続関数.

(3) **[正定符号性]** ${}^\forall n\ge 1, {}^\forall\alpha_k\in\mathbf{C}, {}^\forall z_k\in\mathbf{R}$ $(k=1,\dots,n)$ に対し, $\displaystyle\sum_{j,k=1}^n\alpha_j\overline{\alpha_k}\varphi(z_j-z_k)\ge 0$.

(証明) (1) は容易. (2) ${}^\forall z,h\in\mathbf{R}$ に対し, $|e^{i(z+h)x}-e^{izx}|\le|e^{ihx}-1|\to 0$ $(h\to 0)$ かつ $|e^{ihx}-1|\le 2$ なので, Lebesgue の収束定理から,

$$\sup_z|\varphi(z+h)-\varphi(z)|\le\int|e^{ihx}-1|\mu(dx)\to 0\quad(h\to 0).$$

(3) について.

$$\sum_{j,k=1}^n\alpha_j\overline{\alpha_k}\varphi(z_j-z_k)=\int\sum_{j,k=1}^n\alpha_j\overline{\alpha_k}e^{i(z_j-z_k)x}\mu(dx)=\int\left|\sum_{j=1}^n\alpha_je^{iz_jx}\right|^2\mu(dx)\ge 0.$$

■

定理 1.12 特性関数 φ に対し, 次が成り立つ. 但し, $L^1(d\mu)=L^1(\mathbf{R},\mathcal{B},\mu)$ とする.

(1) $x\in L^1(d\mu)$ なら $\varphi\in C^1$ で, $\varphi'(z)=i\displaystyle\int xe^{izx}\mu(dx)$.

(2) ${}^\exists\varphi''(0)$ なら $x^2\in L^1(d\mu)$.

(証明) (1) 次に注意すれば, Lebesgue の収束定理から容易に示せる. $h\ne 0$ に対し,

$$\frac{e^{ix(z+h)}-e^{ixz}}{h}=i\frac{x}{h}\int_0^h e^{ix(z+s)}ds \text{ より, } \left|\frac{e^{ix(z+h)}-e^{ixz}}{h}\right|\le\frac{|x|}{|h|}\int_0^{|h|}|e^{ix(z+s)}|ds=|x|.$$

(2) $h\ne 0$ に対し,

$$\psi_h(z):=(\varphi(z+h)+\varphi(z-h)-2\varphi(z))/h^2$$

(対称差) とおく. このとき

$$\psi_h(z)=\int_{\mathbf{R}}e^{izx}\left(\frac{i\sin(hx/2)}{h/2}\right)^2\mu(dx)\tag{1.2}$$

と表せる. $\lim_{h\to 0}\psi_h(0)=\varphi''(0)$ より, Fatou の補題を用いて,

$$|\varphi''(0)|=\lim_{h\to 0}\int_{\mathbf{R}}\left(\frac{\sin(hx/2)}{h/2}\right)^2\mu(dx)\ge\int_{\mathbf{R}}x^2\mu(dx).$$

■

問 1.8 上の証明で式 (1.2) と $\lim_{h\to 0}\psi_h(0) = \varphi''(0)$ を示せ.

後半ヒント $\varphi(z\pm h) = \varphi(z) + \varphi'(z)h + \varphi''(z)h^2/2 + o(h^2)$ $(h \to 0)$ を示して, 用いる.

1.8 Lévy の反転公式

L^1 関数の Fourier 変換に対しては逆変換ができるためには可積分であるという条件が必要だった. 次の定理はそれをさらに一般化したもので, 条件無しで成り立つ.

定理 1.13 (Lévy の反転公式) $\mathbf{R}$ 上の分布 μ とその特性関数 φ に対し, $\mu(\{a\}) = \mu(\{b\}) = 0$ なら

$$\mu((a,b)) = \frac{1}{2\pi}\lim_{T\to\infty}\int_{-T}^{T}\frac{e^{-iza}-e^{-izb}}{iz}\varphi(z)dz.$$

より一般的には次が成り立つ.

$$\mu((a,b)) = \frac{1}{2\pi}\lim_{T\to\infty}\int_{-T}^{T}\frac{e^{-iza}-e^{-izb}}{iz}\varphi(z)dz - \frac{1}{2}[\mu(\{a\})+\mu(\{b\})].$$

(証明) まず $z \neq 0$ に対し, $|(e^{-iza}-e^{-izb})/iz| \le (b-a)$ に注意して, Fubini の定理から

$$\int_{-T}^{T}\frac{e^{-iza}-e^{-izb}}{iz}\varphi(z)dz = \int_{\mathbf{R}}\mu(dx)\int_{-T}^{T}\frac{e^{iz(x-a)}-e^{iz(x-b)}}{iz}dz.$$

上の最後の z による積分を $J(T,x,a,b)$ とおくと奇関数は消えるので

$$J(T,x,a,b) = 2\int_0^T\frac{\sin(x-a)z}{z}dz - 2\int_0^T\frac{\sin(x-b)z}{z}dz.$$

ここで, よく知られているように (→ 前の問 1.6)

$$\int_0^\infty\frac{\sin z}{z}dz = \frac{\pi}{2}\quad により,\quad \int_0^\infty\frac{\sin zx}{z}dz = \begin{cases}\pi/2 & (x>0)\\ 0 & (x=0)\\ -\pi/2 & (x<0).\end{cases}$$

これにより

$$\lim_{T\to\infty}J(T,x,a,b) = \begin{cases}0 & (x<a \text{ or } b<x)\\ \pi & (x=a \text{ or } x=b)\\ 2\pi & (a<x<b).\end{cases}$$

一方, sin の形から $|J(T,x,a,b)| \le 4\displaystyle\int_0^\pi\frac{\sin z}{z}dz$. 従って Lebesgue の収束定理から

$$\lim_{T\to\infty}\int_{\mathbf{R}}J(T,x,a,b)\mu(dx) = \pi\int_{\mathbf{R}}1_{\{a,b\}}(x)\mu(dx) + 2\pi\int_{\mathbf{R}}1_{(a,b)}(x)\mu(dx).$$

これから求める式を得る. ■

問 1.9 上の $|J(T,x,a,b)| \le 4\displaystyle\int_0^\pi\frac{\sin z}{z}dz$. を示せ.

(ヒント) $x>0$ なら $^\forall T>0$ に対し,

$$\int_0^T\frac{\sin xt}{t}dt = \int_0^{Tx}\frac{\sin z}{z}dz \le \int_0^\pi\frac{\sin z}{z}dz.$$

定理 1.14 (一意性定理) $\mathbf{R}$ 上の分布 μ, ν それぞれの特性関数 φ_μ, φ_ν に対し, $\varphi_\mu = \varphi_\nu$ なら $\mu = \nu$.

(証明) $(a,b); \mu(\{a\}) = \mu(\{b\}) = \nu(\{a\}) = \nu(\{b\}) = 0$ なる区間の全体を $\mathcal{I}$ とする. 反転公式から $\mu = \nu$ on $\mathcal{I}$ で, これを満たさない区間は高々可算個であることに注意すれば (次の問参照), 任意の半開区間 $(a,b]$ を上から近似することにより; $\mu((a,b]) = \nu((a,b])$ も成り立つ. 更に, 半開区間の素な有限和全体 $\mathcal{A}$ (これは加法族) でも一致しているので, 一致する集合全体が, $\mathcal{A}$ を含み、しかも単調族なので, $m(\mathcal{A}) = \sigma(\mathcal{A}) = \mathcal{B}^1$ を含み (ここで単調族定理を用いた) $\mu = \nu$ on $\mathcal{B}^1$ を得る. ■

問 1.10 $\mathbf{R}$ 上の分布 μ に対し, $\mu(\{a\}) > 0$ なる $a \in \mathbf{R}$ は高々可算個しかないことを示せ.

分布と特性関数に関する, この他の重要な話題については簡単に紹介だけしておこう。その前に, 分布関数についての定義と結果を与えておく.

1.9 Lebesgue-Stielties 測度

定義 確率空間 $(\Omega, \mathcal{F}, P)$ 上の実数値確率変数 X に対し, $F(x) := P(X \leq x)$ とおき, これを X の分布関数 (distribution function) という.

(1) このとき $F : \mathbf{R} \to [0,1]$ は次を満たす.

(1) 単調増加, i.e., $x < y \Rightarrow F(x) \leq F(y)$.

(2) 右連続で, 左極限をもつ, i.e., $F(x) = F(x+) := \lim_{y \to x, y > x} F(y)$, $^\exists F(x-) := \lim_{y \to x, y < x} F(y)$.

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

定義 上の問の 3 つの性質を満たす関数 $F : \mathbf{R} \to [0,1]$ が与えられたとき, この $F(x)$ を単に **$\mathbf{R}$ 上の分布関数**という.

ではこの分布関数が与えられたとき, $F(x) = P(X \leq x)$ を満たすような確率空間と確率変数が存在するか？という逆問題が考えられるがその答えを肯定的に与えるのが次の定理である.

定理 [Lebesgue-Stielties 測度 (Lebesgue-Stielties measure)]

分布関数 $F : \mathbf{R} \to [0,1]$ に対し, $^{\exists 1} \mu : \mathcal{B}^1 \to [0,1]$ 分布; $\mu((-\infty, x]) = F(x)$.

定義 この分布 μ on $(\mathbf{R}, \mathcal{B}^1)$ を分布関数 $F(x)$ による **Lebesgue-Stielties 測度 (分布)** という. またこれによる積分が定義できるが, それを

$$\int_{\mathbf{R}} f(x) dF(x) := \int_{\mathbf{R}} f(x) \mu(dx); \quad \textbf{Lebesgue-Stielties 積分}\text{という.}$$

注 F の値域を $[0,1]$ から $\mathbf{R}$ に変えても (このとき分布関数とは呼ばないが), 同様な結果が成り立つので, 一般的に μ を Lebesgue-Stielties 測度という呼び方をする.

1.10 測度の弱収束

$\mathcal{P}(\mathbf{R})$ を $\mathbf{R}$ 上の分布全体とする. $\mu_n, \mu \in \mathcal{P}(\mathbf{R})$ として, μ_n が μ に収束 (弱収束) するというのを次で定義する.

$$\mu_n \to \mu \overset{\text{def}}{\iff} {}^\forall f \in C_b(\mathbf{R}), \langle \mu_n, f\rangle \to \langle \mu, f\rangle.$$

ここで $C_b(\mathbf{R})$ は $\mathbf{R}$ 上の有界連続関数全体で, $\langle \mu, f\rangle = \displaystyle\int f d\mu$ である.

また確率変数 X_n, X に対し, その分布が収束するとき, **法則収束**するという. 即ち,

$$X_n \to X \text{ in law} \overset{\text{def}}{\iff} {}^\forall f \in C_b(\mathbf{R}),\ E[f(X_n)] \to E[f(X)].$$

問 1.11 概収束なら法則収束が, 確率収束でも法則収束が成り立つことを, それぞれ直接, 示せ.

前半は, f が有界連続なので Lebesgue の収束定理より明らか. 後半は, まず $\limsup$ の性質より, ${}^\exists\{n_k\}; \lim E[f(X_{n_k})] = \limsup Ef(X_n)]$. 更に, 確率収束なら概収束する部分列がとれるので, ${}^\exists\{Y_j\} \subset \{X_{n_k}\}; Y_j \to X$. 前半の結果から $\limsup Ef(X_n)] = \lim E[f(Y_j)] = E[f(X)]$ となり, $\liminf$ についても同様なので, 示せる

ちなみに, 分布の収束を定義するのに, 単純に考えると, ${}^\forall A \in \mathcal{B}, \mu_n(A) \to \mu(A)$ とするのが, 自然だと思うかもしれない. しかし, 実際には, これでは条件が厳し過ぎて, 役に立たないのである. それを表しているのが以下に述べる結果である.

定理 1.15 $\mathbf{R}$ 上の分布 $\mu_n.\mu$ に対し, 次は全て同値.

(1) $\mu_n \to \mu$

(2) ${}^\forall U \subset \mathbf{R}$: 開集合, $\liminf \mu_n(U) \geq \mu(U)$.

(3) ${}^\forall F \subset \mathbf{R}$: 閉集合, $\limsup \mu_n(F) \leq \mu(F)$.

(4) ${}^\forall A \in \mathcal{B}^1; \mu(\partial A) = 0$, $\lim \mu_n(A) = \mu(A)$.

(5) F_n, F を μ_n, μ それぞれの分布関数とする. ${}^\forall x$: F の連続点, i.e., $F(x-) = F(x)$, $F_n(x) \to F(x)$.

(6) ${}^\forall f \in C_c(\mathbf{R}), \langle \mu_n, f\rangle \to \langle \mu, f\rangle$. 但し, $C_c(\mathbf{R})$ は compact な台をもつ連続関数全体で, 台 (support) とは $\operatorname{supp} f = \overline{\{f \neq 0\}}$ である.

証明は, (1) $\Rightarrow$ (2) は, 定義関数 1_U を下から非負の連続関数で, 各点近似するもの $0 \leq h_k \uparrow 1_U$ を作れば良い. 実際, $\mu_n(h_k) \leq \mu_n(U)$ で, $\lim_{n\to\infty}\langle \mu_n, h_k\rangle = \langle \mu, h_k\rangle \uparrow \langle \mu, 1_U\rangle = \mu(U)$ なので, 最初の不等式で、$\liminf_{n\to\infty}$ を施してから, $k \to 0$ とすれば良い. また, h_k は集合までの距離関数が連続であることを用いて, U を $1/k$ だけ縮めた閉集合 $U_k = \{x \in U; d(x, U^c) \geq 1/k\}$ に対し, $h_k = d(x, U^c)/(d(x, U_k) + d(x, U^c))$ とおけば良い. 但し, $d(x, A) = \inf_{y\in A} d(x, y)$ である.

(2) $\iff$ (3) は補集合を, (2),(3) $\Rightarrow$ (4) は, 内核と閉包 $A^o, \overline{A}$ を用いれば容易. (4) $\Rightarrow$ (5) は F の連続点 x に対し, $A = (-\infty, x]$ を考えれば良い. (5) $\Rightarrow$ (6) は $f \in C_c(\mathbf{R})$ の単関数近似 f_k が, 階段関数 (定義関数が区間で表されるもの) となり, しかもその端点は, F の連続点であるようにできることを用いれば示せる. 実際, f_k の作り方より, $k \geq 1$, $\langle \mu_n, f_k\rangle \to \langle \mu, f_k\rangle$ を得て, 更に, f_k を $\|f - f_k\|_\infty \to 0 (k \to \infty)$ となるようにとれるので, $|\langle \mu_n, f\rangle - \langle \mu, f\rangle| \leq \langle \mu_n, |f - f_k|\rangle + |\langle \mu_n, f_k\rangle - \langle \mu, f_k\rangle| + \langle \mu, |f - f_k|\rangle \leq 2\|f - f_k\|_\infty + |\langle \mu_n, f_k\rangle - \langle \mu, f_k\rangle|$ において, $n \to \infty$, $k \to \infty$ とすれば良い. (6) $\Rightarrow$ (1) は $f \in C_b(\mathbf{R})$ を $C_c(\mathbf{R})$ の元で, 広義一様 (即ち, ${}^\forall k \geq 1, [-k, k]$ で一様) に近似してやれば良い. 具体的には $g_k \in C_b(\mathbf{R}); 0 \leq g_k \leq 1$ を $[-k, k]$ で 1, $[-k-1, k+1]^c$ で 0 となる連続関数として, 十分大きな $K > 1$ を固定して, fg_K を考えれば良い. そ

の際, (6) から $^\forall\varepsilon>0, {}^\exists K; \sup_{n\ge1}\mu_n([-K,K]^c)<\varepsilon$ が得られることを用いる. 実際, これは $1_{[-k,k]}\le g_k\le 1_{[-k-1,k+1]}$ と (6) より, $\mu([-k,k])\le\langle\mu,g_k\rangle=\lim_{n\to\infty}\langle\mu_n,g_k\rangle\le\varlimsup_{n\to\infty}\mu_n([-k-1,k+1])$ に注意すれば, $^\forall\varepsilon>0, {}^\exists K_0; \varlimsup_{n\to\infty}\mu_n([-K_0-1,K_0+1]^c)\le\mu([-K_0,K_0]^c)<\varepsilon$ を得て, 更に, これから, $^\exists K\ge K_0; \sup_{n\ge1}\mu_n([-K,K]^c)<\varepsilon$ が言える. ■

これでも, まだ, 分布の収束の定義の妥当性には不十分だと思うかもしれない. しかし, 更なる結果として, 次の特性関数の収束との関係があり, これにより, その定義が, 必要かつ十分であると思って良いのではないだろうか。

定理 1.16 φ_n,φ: $\mu_n,\mu\in\mathcal{P}(\mathbf{R})$ の特性関数とする. $\mu_n\to\mu$ なら $\varphi_n\to\varphi$ (広義一様)

ちなみに, 「広義一様 (収束)」とは, 任意の cpt (コンパクト) 集合 = 有界閉集合上で一様 (収束) である.

$f(x)=e^{izx}$ が有界連続関数なので, 各点収束するのは明らかで, 広義一様をいうためには, 最後に述べる定理が必要となる. ($\{\mu_n\}$ が相対 cpt なので, 緊密という結果を.)

定理 1.17 (Lévy の連続性定理) φ_n: $\mu_n\in\mathcal{P}(\mathbf{R})$ の特性関数とする. $^\exists\varphi;\varphi_n\to\varphi$ (各点収束) かつ, φ が原点で連続なら $^\exists\mu$: $\mathbf{R}$ 上の分布; φ は μ の特性関数. $\mu_n\to\mu$, しかも $\varphi_n\to\varphi$ (広義一様).

系 1.3 (Glivenko の定理) φ_n,φ: $\mu_n,\mu\in\mathcal{P}(\mathbf{R})$ の特性関数. $\varphi_n\to\varphi$ (各点収束) なら, $\mu_n\to\mu$.

Lévy の定理で, φ の原点での連続性の条件は落とせない. 例えば, $N(0,n)$ の特性関数は $\varphi_n(z)=\exp(-nz^2/2)\to 1_{\{0\}}(z)$ を満たし, 極限は原点で連続で無いので特性関数ではない.

この定理の証明には, 次の, 「確率測度のある集まり $\Lambda\subset\mathcal{P}(\mathbf{R})$ に対し, **緊密 (tight)** $\Longleftrightarrow$ **相対 cpt (コンパクト) (relatively compact)** 」であることを用いる.

定理 1.18 $\Lambda\subset\mathcal{P}(\mathbf{R})$: 緊密 $\overset{\text{def}}{\Longleftrightarrow}$ $^\forall\varepsilon>0, {}^\exists K_\varepsilon\subset\mathbf{R}$: compact; $^\forall\mu\in\Lambda, \mu(K_\varepsilon)>1-\varepsilon$. $\Longleftrightarrow$ Λ: 相対コンパクト, i.e., 任意の列に対し, 収束する部分列が存在する. $^\forall\{\mu_n\}\subset\Lambda$, $^\exists\{n_k\},\mu\in\mathcal{P}(\mathbf{R});\mu_{n_k}\to\mu$

この「相対コンパクト」は厳密には, "相対点列コンパクト", 即ち, Λ の閉包 $\overline{\Lambda}$ が点列コンパクトということである.

(証明) 「相対コンパクトなら緊密」は簡単で, もし緊密でないとすると, $^\exists\varepsilon_0>0, {}^\forall K$: cpt に対し, $^\exists\mu_K\in\Lambda;\mu_K(K)<1-\varepsilon_0$. $K=[-n,n]$ として, $\mu_n=\mu_K$ とする. 仮定により, $^\exists\mu_{n_k}\to{}^\exists\mu\in\mathcal{P}(\mathbf{R})$. $n_k\ge n$ とすると, $\mu_{n_k}([-n,n])\le\mu_{n_k}([-n_k,n_k])<1-\varepsilon$. $k\to\infty$ とすると, $\mu((-n,n))\le\liminf\mu_{n_k}([-n,n])\le1-\varepsilon$. 所が, $n\ge1$ は任意なので, $\mu(\mathbf{R})\le1-\varepsilon$ となり, 矛盾.

問題は逆の「緊密なら相対コンパクト」で, 定理の証明でも用いるのだが, 概略のみを. まず有理数 r_k でのみ分布関数列 $F_n(r_k)=\mu_n((-\infty,r_k])\in[0,1]$ を考え, 各 $k=1,2,\dots$ に対し, その収束部分列 $\{F_{n_j^k}(r_k)\}$ を $\{n_j^{k+1}\}\subset\{n_j^k\}$ と順次取り出せば, 対角成分 $n_j=n_j^j$; $F_{n_j}(r_k)$ は $^\forall k$ に対し, 収束するので, その極限を $F(r_k):=\lim_{j\to\infty}F_{n_j}(r_k)$ とする. さらに $^\forall x\in\mathbf{R}$ に対し, $F(x)=\inf_{r>x;r\in\mathbf{Q}}F(r)$ として拡張すれば, 非減少, かつ, 右連続で左極限をもち, 更に, 緊密性より, $F(-\infty)=0, F(\infty)=1$ が示せる. 従って, $F(x)$ が分布関数となり, しかも, 連続点での F_{n_j} の収束も示せるので証明が終わる. ■

ここで, 次で用いる事実を 1 つ述べておく.

・$\{\mu_n\} \subset \mathcal{P}(\mathbf{R})$ が相対 cpt で, 任意の収束部分列が同じ極限 μ に収束するなら, 元の μ_n 自身が μ に収束する, i.e., ${}^\exists\mu \in \mathcal{P}(\mathbf{R}); {}^\forall\{n_k\}, \mu_{n_k} \to \mu \Longrightarrow \mu_n \to \mu$.

実際, そうでないとすると, ${}^\exists g \in C_b(\mathbf{R}^d); \langle\mu_n, g\rangle \not\to \langle\mu, g\rangle$. これから, ${}^\exists\{n_k\}; {}^\exists\langle\mu_{n_k}, g\rangle \to {}^\exists\alpha \neq \langle\mu, g\rangle$ が言えるが, 相対 cpt の仮定から, ${}^\exists\{n_{k_j}\}; \mu_{n_{k_j}} \to {}^\exists\widetilde{\mu}$ で, $\widetilde{\mu} \neq \mu$ となる. 実際, $\langle\widetilde{\mu}, g\rangle = \lim\langle\mu_{n_{k_j}}, g\rangle \neq \langle\mu, g\rangle$ なので. しかし, これは仮定の $\widetilde{\mu} = \mu$ に反する. ■

[Lévy の連続性定理の証明] 仮定から, $\{\mu_n\}$ が緊密であることが示せるので, 相対コンパクト性より, ${}^\exists\mu_{n_k} \to {}^\exists\mu \in \mathcal{P}(\mathbf{R})$. これから, $\varphi = \varphi_\mu$ で, 再び相対コンパクト性から, $\mu_n \to \mu$ も言える. 実際, そうでないとすると上の結果の対偶から, ${}^\exists\mu_{n_k} \to {}^\exists\widetilde{\mu} \neq \mu$ となり, $\varphi_{\widetilde{\mu}} = \varphi_\mu$, 即ち, 一意性定理より, $\widetilde{\mu} = \mu$ となってしまい矛盾

緊密性については, まず, 一般に, ${}^\forall\nu \in \mathcal{P}(\mathbf{R}), {}^\forall L > 0$ に対し,

$$\nu([-2L.2L]^c) \leq L\int_{-1/L}^{1/L}(1 - \varphi_\nu(z))dz$$

が成り立つ. 実際, Fubini により, ($|1 - e^{izx}| \leq 2$ に注意)

$$\begin{aligned}(\text{右辺}) &= L\int_{-1/L}^{1/L}dz\int_{\mathbf{R}}(1 - e^{izx})\nu(dx) = \int_{\mathbf{R}}\nu(dx)L\int_{-1/L}^{1/L}(1 - e^{izx})dz \\ &= \int_{\mathbf{R}}\nu(dx)2\int_0^1(1 - \cos(zx/L))dz = \int_{\mathbf{R}}2\left(1 - \frac{\sin x/L}{x/L}\right)\nu(dx).\end{aligned}$$

ここで, $1 - \dfrac{\sin x/L}{x/L} \geq 0$ と $|x| \geq 2L$ なら $\dfrac{\sin x/L}{x/L} \leq \dfrac{1}{2}$ より, 求める式を得る.

そこで, φ: conti. at 0 より,

$$\lim_{L\to\infty} L\int_{-1/L}^{1/L}(1 - \varphi(z))dz = \lim_{h\to 0}\frac{1}{h}\int_{-h}^{h}(1 - \varphi(z))dz = 2(1 - \varphi(0)) = 0.$$

最後の等号は, $\varphi(0) = \lim\varphi_n(0) = 1$ による. 従って, $\varphi_n \to \varphi$ と, $|1 - \varphi_n(z)| \leq 2$ に注意して, 上の結果を組み合わせれば, ${}^\forall\varepsilon > 0, {}^\exists L > 1; \limsup\mu_n([-2L, 2L]^c) < \varepsilon$ を得る. 従って, ${}^\exists N; a$
$n \geq N, \mu_n([-2L, 2L]^c) < \varepsilon$ で, 更に, 必要なら L を大きく取り直すことにより, ${}^\forall n < N$ についても成り立つようにできる. ■

最後に, 「$\mu_n \to \mu$ なら $\varphi_n \to \varphi$ (広義一様)」について.

$\{\mu, \mu_n\}$: 相対コンパクトより, 緊密となるので, ${}^\forall\varepsilon > 0, {}^\exists K \subset \mathbf{R}$: cpt; $\mu(K), \mu_n(K) > 1 - \varepsilon$. このとき,

$${}^\exists\delta > 0; |{}^\forall h| < \delta, \sup_{x\in K}|e^{ihx} - 1| < \varepsilon$$

が成り立つ. これより, 同じ h に対し, $\sup_{z\in\mathbf{R}}|\varphi_n(z + h) - \varphi_n(z)| < 3\varepsilon$ が成り立つ. そこで, 任意の cpt $C \subset \mathbf{R}$ に対し, δ-稠密 $\{z_1, \ldots, z_k\}$, i.e., ${}^\forall z \in C, {}^\exists j; |z - z_j| < \delta$ をとり, $N \geq 1$ を ${}^\forall n \geq N, |\varphi_n(z_j) - \varphi(z_j)| < \varepsilon$ ととれば, ${}^\forall z \in C$ に対し一様に次が成り立つ.

$$|\varphi_n(z) - \varphi(z)| \leq |\varphi_n(z) - \varphi_n(z_j)| + |\varphi_n(z_j) - \varphi(z_j)| + |\varphi(z_j) - \varphi(z)| < 7\varepsilon.$$

■

2 大偏差原理 (Large Deviation Principle)

$\{X_n\}$: i.i.d., $EX_1 = m \in \mathbf{R}, V(X) = v > 0$ とする.
$S_n = \sum_{k=1}^n X_k$ とおき, $a > m$ に対し, $P(S_n > an) \sim ?$ を調べたい.
大数の法則より, $S_n/n \to m$, a.s. なので, $P(S_n > an) \to 0$ であるが, 更に, 中心極限定理より,

$$P\left(S_n > \sqrt{n}a + mn\right) = P\left(\frac{S_n - mn}{\sqrt{n}} > a\right) \to \frac{1}{\sqrt{2\pi v}} \int_a^\infty e^{-x^2/(2v)} dx$$

なので, a を $\sqrt{n}(a-m)$ とすれば, ラフには,

$$P\left(S_n > na\right) \sim \frac{1}{\sqrt{2\pi v}} \int_{\sqrt{n}(a-m)}^\infty e^{-x^2/(2v)} dx$$

これの $n \to \infty$ での減少のオーダーとその係数を調べたい.
明らかに中心（平均）から大きく離れた部分の確率を調べることになるので, 「大偏差」という. 次の定理が成り立つ.

定理 2.1 (大偏差原理；Cramér の定理) $\{X_n\}$: i.i.d, で, $^\forall t \in \mathbf{R}, E[e^{t|X_1|}] < \infty$ とする. $\psi(t) = E[e^{tX_1}]$ とおく. このとき, $S_n = \sum_{k=1}^n X_k$ とおくと, $^\forall a > EX_1$ に対し,

$$\lim_{n\to\infty} \frac{1}{n} \log P(S_n \geq an) = -I(a), \text{ 即ち, } P(S_n \geq an) \sim e^{-nI(a)}.$$

ここで, $I(x) = \sup_{t\in\mathbf{R}}(xt - \log\psi(t))$ で, $\mathbf{R}$ において, 下半連続な凸関数で, 次を満たす.

$$\lim_{x\to\pm\infty} I(x) = \infty, \quad I(x) \geq 0 = I(EX_1) \quad (x \in \mathbf{R}).$$

Cramér は, もっと弱い条件 $^\exists t_0 > 0; E[e^{t_0|X_1|}] < \infty$ の下で示した.
ちなみに,
・I: **下半連続 (lower semi-conti.)** at $x \iff {}^\forall\varepsilon > 0, {}^\exists\delta > 0; |{}^\forall y - x| < \delta, I(y) > I(x) - \varepsilon$. (下から ε 幅で抑えられる, $I(x)$ の.) これより, $I(x) \leq \liminf_{y\to x} I(y)$ が成立.
・I: **凸関数 (convex)** on 区間 $I \iff {}^\forall p, q \geq 0; p + q = 1, {}^\forall x, y \in I, I(px + qy) \leq pI(x) + qI(y)$.

まず, 道具として, 次の有用な定理を挙げておく.

定理 2.2 [劣加法定理 (sub-additive theorem)] 実数列 $\{a_n\}$ が劣加法性 $a_{m+n} \leq a_m + a_n$ をもつなら, $^\exists \lim(a_n/n) = \inf(a_n/n)$. 逆に, 実数列 $\{b_n\}$ が優加法性 $b_{m+n} \geq b_m + b_n$ をもつなら, $^\exists \lim(b_n/n) = \sup(b_n/n)$.

[証明] $\liminf(a_n/n) \geq \inf(a_n/n)$ より, $\limsup(a_n/n) \leq a_m/m$ が $^\forall m \geq 1$ で成り立つことを示せば良い. m を 1 つ固定し, n を m で割った結果を $n = km + r$ と表す. $0 \leq r < m$. 劣加法性より, $a_n \leq ka_m + a_r$. n で割って,

$$\frac{a_n}{n} \leq \frac{km}{km + r} \cdot \frac{a_m}{m} + \frac{a_r}{n}.$$

両辺で, $\limsup_{n\to\infty}$ をとると, $k = k_n \to \infty$ より, 求める結果を得る. 後半も同様である. ■

まず, $a > EX_1$ のとき, 定理の極限の存在について議論しよう.

$\{X_n\}$ が同分布で, 独立ということから,

$$P(S_m \geq ma)P(S_n \geq na) = P(S_{m+n} - S_n \geq ma, S_n \geq na) \leq= P(S_{m+n} \geq (m+n)a).$$

$p_n = \log P(S_n \geq na)$ とおくと, 優加法性をもつことになるので, 劣加法定理より, $^\exists \lim(p_n/n) = \sup(p_n/n) =: -\widetilde{I}(a)$ とおけば, $0 \leq \widetilde{I}(a) \leq \infty$.

さて, 大偏差原理の証明には, 分布関数に対応した確率測度が存在し, さらに, それを分布としてもつ可算無限個の独立な確率変数（と確率空間）が構成できることを用いる. それには, Lebesgue-Stielties 測度の存在定理と Kolmogorov の拡張定理が必要となる.

取り敢えず, それについてはまた後で述べることにして, 先に証明をしてしまおう.

[大偏差原理の証明] $I(a) = \sup_t\{at - \log\psi(t)\}$ とする. $^\forall a > EX_1$ に対し, $\widetilde{I}(a) = I(a)$, 即ち, $\lim_{n\to\infty} \frac{1}{n} \log P(S_n \geq an) = -I(a)$ を示すのだが, まず, X_1 の代わりに $X_1 - a$ を考えることにより, $\psi(t)$ は $e^{-ta}\psi(t)$ に, $I(a)$ は $I(0)$ となるので, $a = 0 > EX_1$ とし, $\rho := \inf_{t\in\mathbf{R}} \psi(t)$ として, $I(0) = -\log\rho$ で, 次を示せば良い. 但し, $\log 0 = -\infty$ である.

$$\lim_{n\to\infty} \frac{1}{n} \log P(S_n \geq 0) = \log\rho.$$

簡単のため $X_1 = X$ と書く. 今, 仮定の $^\forall t \in \mathbf{R}, E[e^{t|X|}] < \infty$ により, $^\forall n \geq 1, ^\forall t \in \mathbf{R}$, $E[|X|^n e^{t|X|}] < \infty$ (実際, $x > 0$ なら $x^n \leq n! e^x$ より.) 従って, 収束定理により (より詳しくは, 後述の微分と積分の交換定理; 定理 2.3 により), $\psi(t)$ は C^∞ で, $\psi'(t) = E[Xe^{tX}]$, $\psi''(t) = E[X^2e^{tX}] > 0$. ($EX < 0$ より, $P(X = 0) < 1$, i.e., $P(X \neq 0) > 0$ に注意.) よって, $\psi(t)$ は $t \in \mathbf{R}$ で, 狭義凸で, $\psi'(0) = EX < 0$. 以下, $X = X_1$ の分布によって場合分けして考えよう.

(Case 1) $P(X \leq 0) = 1, P(X = 0) \geq 0$ のとき. ψ は減少関数で, 収束定理により, $\psi(t) \downarrow P(X = 0) = \rho \geq 0$ $(t \uparrow \infty)$. また, $\rho > 0$ なら, $P(S_n \geq 0) = P(X_1 = 0, \cdots, X_n = 0) = \rho^n$ となり, 求める式の左辺は $\log\rho$ となる. $\rho = 0$ でも, $P(S_n \geq 0) = 0$ で, $\log 0 = -\infty$ なので, $-\infty = -\infty$ として求める結果を得る. (この計算は上に含まれると思っても良い.)

(Case 2) $P(X < 0) > 0, P(X > 0) > 0$ のとき. $\psi(\pm\infty) = \infty$ となる. 実際, 確率の連続性より, $^\exists\delta > 0; P(X \geq \delta) > 0$ となるので, $0 < t \to \infty$ なら $\psi(t) \geq E[e^{tX}; X \geq \delta] \geq e^{\delta t}P(X \geq \delta) \to \infty$. $0 > t \to -\infty$ も同様. また, ψ が狭義凸で, $\psi'(0) = EX < 0$ だったので, $^\exists\tau > 0; \psi'(\tau) = 0, \psi(\tau) = \inf\psi = \rho > 0$. そこで, X の分布関数 $F(x) = P(X \leq x)$ に対し, **クラメール変換 (Cramér transform)**

$$\widehat{F}(x) := \frac{1}{\rho}\int_{-\infty}^x e^{\tau y}dF(y) = \frac{1}{\rho}E[e^{\tau X}; X \leq x]$$

とおくと, $\widehat{F}(\infty) = \psi(\tau)/\rho = 1$ より, これも分布関数となる. これに対応する i.i.d. の確率変数列 $\{\widehat{X}_n\}$ に対し, $\widehat{S}_n = \sum_{k=1}^n \widehat{X}_k$ とおくと,

$$E\widehat{X}_1 = 0, \quad \widehat{\sigma}^2 := V(\widehat{X}_1) \in (0,\infty), \quad P(S_n \geq 0) = \rho^n E\left[e^{-\tau\widehat{S}_n} 1_{\{\widehat{S}_n \geq 0\}}\right] \tag{2.1}$$

を満たす.

これを一旦, 認めて, $^\forall M > 0$,

$$e^{-\tau M\widehat{\sigma}\sqrt{n}} P\left(0 \leq \frac{\widehat{S}_n}{\widehat{\sigma}\sqrt{n}} < M\right) \leq E\left[e^{-\tau\widehat{S}_n} 1_{\{\widehat{S}_n \geq 0\}}\right] \leq 1$$

一方, $\widehat{S}_n$ に対し, CLT が成り立つので, 上の確率は $n \to \infty$ で, $\int_0^M e^{-x^2/2}dx/\sqrt{2\pi} > 0$ に近づく. よって, 各辺で log をとり, n で割り, $\liminf$ を施すと, はさみうちにより,

$$\liminf_{n\to\infty} \frac{1}{n} \log E\left[e^{-\tau \widehat{S}_n} 1_{\{\widehat{S}_n \geq 0\}}\right] = 0$$

これと (2.1) の最後の式より, 求める結果を得る. 即ち,

$$\lim_{n\to\infty} \frac{1}{n} \log P(S_n \geq 0) = \log \rho.$$

(2.1) を示す. $\widehat{\psi}(t) := E[e^{t\widehat{X}_1}]$ とおくと,

$$\widehat{\psi}(t) = \int_{\mathbf{R}} e^{tx} d\widehat{F}(x) = \frac{1}{\rho} \int_{\mathbf{R}} e^{tx} e^{\tau x} dF(x) = \frac{1}{\rho} \psi(t+\tau).$$

よって $\widehat{\psi}$ も C^∞ 級で, $E\widehat{X}_1 = \widehat{\psi}'(0) = \psi'(\tau)/\rho = 0$, $V(\widehat{X}_1) = \widehat{\psi}''(0) = \psi''(\tau)/\rho \in (0,\infty)$ となる. また, $d\widehat{F}(x) = \frac{1}{\rho} e^{\tau x} dF(x)$ より, $dF(x) = \rho e^{-\tau x} d\widehat{F}(x)$ なので,

$$P(S_n \geq 0) = \int_{\{\sum_{k=1}^n x_k \geq 0\}} dF(x_1) \cdots dF(x_n) = \rho^n \int_{\{\sum_{k=1}^n x_k \geq 0\}} e^{-\tau \sum_{k=1}^n x_k} d\widehat{F}(x_1) \cdots d\widehat{F}(x_n)$$

これより (2.1) の最後の式を得る.

最後に, $I(x) = \sup_{t\in\mathbf{R}}\{xt - \log\psi(t)\} = -\sup_{n\geq 1}([\log P(S_n \geq an)]/n)$ の性質を調べる. 一般に, 線形関数の族の上限は, 凸関数で, 連続関数の族の上限は, 下半連続となるので (→ 確かめよ.) 線形連続関数 $x \mapsto xt - \log\psi(t)$ の上限である $I(x)$ は下半連続な凸関数である.

$I(x) \to \infty$ $(x \to \pm\infty)$ について, まず, $x \to \infty$ のとき, そうでないとすると, $^\exists L > 0, {}^\exists x_n \geq n; I(x_n) < L$, i.e., $^\forall t \in \mathbf{R}, x_n t - \log\psi(t) < L$. そこで $t = 2L/x_n$ ととると, $2L - \log\psi(2L/x_n) < L$. $x_n \to \infty$ と ψ の連続性から, $0 < L < \log\psi(2L/x_n) \to \log\psi(0) = \log 1 = 0$ となり, 矛盾. $x \to -\infty$ の時も同様. 次に, $I(x) \geq 0 = I(EX)$ について. $I(x) \geq -\log\psi(0) = 0$ $(t = 0$ とした) で, 更に, $-\log x$ が下に凸なので, Jensen の不等式 (次の問) から, $^\forall t \in \mathbf{R}$, $\log\psi(t) = \log E[e^{tX}] \geq E[\log e^{tX}] = tEX$, i.e., $tEX - \log\psi(t) \leq 0$. また, $I(x) \geq 0$ だったので, $0 \leq I(EX) = \sup_t(tEX - \log\psi(t)) \leq 0$. 故に, $I(EX) = 0$. ■

定理 2.3 (微分と積分の交換定理) $(X, \mathcal{F}, \mu)$ を一般の測度空間とする, $t \in (a,b)$ に対し, 可積分関数 $f_t = f_t(x) \in L^1(d\mu)$ があり, μ-a.e $x \in X$ に対し, $f_t(x)$ が $t \in (a,b)$ について可微分とする. このとき, $\sup_{t\in(a,b)} |\partial_t f_t| \in L^1$ なら,

$$d_t \int_X f_t(x)\mu(dx) = \int_X \partial_t f_t(x)\mu(dx).$$

特に, $(X, \mathcal{F}) = (\mathbf{R}, \mathcal{B}^1)$ で, $\mu = \mu_X$: X の分布なら,

$$d_t E[f_t(X)] = E[\partial_t f_t(X)].$$

ここで, $d_t = d/dt, \partial_t = \partial/\partial t$ である.

(証明) 平均値の定理と Lebesgue の収束定理より明らか. 実際, $t, t+h \in (a,b); h \neq 0$ に対し, 平均値の定理より, a.e. x について,

$$^\exists\theta \in (0,1);\ \frac{1}{h}(f_{t+h}(x) - f_t(x)) = \partial_t f_{t+\theta h}(x).$$

よって, 仮定から, これは h にも t にも無関係な可積分関数で抑えられるので, 収束定理が使えて, $h \to 0$ の極限と積分の交換ができる. ■

問 2.1 (イエンセンの不等式 (Jensen's ineq.)) $-\infty \le a < b \le \infty$ に対し, 区間 $I = (a,b)$ 上の凸関数 f と I に値をとる可積分な確率変数 X で, 更に $f(X)$ も可積分とするなら,

$$f(EX) \le E[f(X)], \quad \text{i.e.,} \quad f\left(\int_{\mathbf{R}} x\mu_X(dx)\right) \le \int_{\mathbf{R}} f(x)\mu_X(dx).$$

凸関数はそれ以下の一次関数の上限で表される, i.e., $f(x) = \sup\{cx+d; {}^\forall y \in I, cy+d \le f(y)\}$ と表されることを用いれば, $cEX + d \le E[f(X)]$ より, 明らか. 実際, $a < s < t < u < b$ に対し, 凸性より,

$$\frac{f(t)-f(s)}{t-s} \le \frac{f(u)-f(t)}{u-t}$$

($t = ps+qu; p,q \ge 0, p+q = 1$ に対し, $f(t) \le pf(s)+qf(u)$ で, $p = (u-t)/(u-s), q = (t-s)/(u-s)$ による. $p(f(t)-f(s)) \le q(f(u)-f(t))$ で, $(u-s)$ 倍して $(u-t)(t-s)$ で割れば良い.) $\alpha_t = \sup_{s<t}($左辺$)$ とすると, $\alpha_t(u-t) \le f(u)-f(t)$, i.e., $f(u) \ge \alpha_t(u-t)+f(t)$ $(u > t)$ これは α_t の定義より, $u \le t$ でも成り立ち, 結局, 全ての $a < u < b$ で成り立つ. 従って $a < t < b$ も任意となる. また明らかに $t = u$ なら等しくなるので, 一次関数の上限として表される. もしくは, 直接, $f(u) \ge \alpha_t(u-t)+f(t)$ に $t = EX, u = X$ を代入し, 平均をとれば, 求める式を得る. ■

問 2.2 連続関数の族の上限で与えられる関数は下半連続となることを示せ.
${}^\forall t \in T$, $f_t(x)$: conti. in x, なら $g = \sup_{t\in T} f_t$: lower semi-conti.

${}^\forall x$: fixed. 上限の性質より, ${}^\forall\varepsilon > 0, {}^\exists t \in T; g(x)-\varepsilon/2 < f_t(x)$ $(\le g(x))$. f_t: conti. at x より, ${}^\exists\delta > 0; |{}^\forall y - x| < \delta, |f_t(y)-f_t(x)| < \varepsilon/2$. よって, $f_t(x)-\varepsilon/2 < f_t(y)$. これらより, $g(y) \ge f_t(y) > f_t(x)-\varepsilon/2 > g(x)-\varepsilon$ を得るので下半連続. ■

3 測度の拡張定理と応用 (Extension Theomre & Its Applications)

本節では測度の拡張定理を用いて, 確率論で特に必要となる 3 つの重要な定理について紹介する. まず拡張定理の確率バージョンを述べておく.

[確率測度の拡張定理] $\mathcal{A}$ を Ω 上の加法族とする. $\mathcal{A}$ 上の加法的集合関数 $P_0 : \mathcal{A} \to [0,1]$ が $P_0(\Omega) = 1$ をみたし, $\mathcal{A}$ 上 σ-加法的

$$A_n \in \mathcal{A}\ (n = 1,2,\dots) \text{ disjoint かつ } \bigcup_{n=1}^\infty A_n \in \mathcal{A} \implies P_0(\bigcup_{n=1}^\infty A_n) = \sum_{n=1}^\infty P_0(A_n) \tag{3.1}$$

であるなら, $(\Omega, \sigma(\mathcal{A}))$ 上の確率測度 P で $P|_{\mathcal{A}} = P_0$ なるものが唯一つ存在する. この P は実際には, P_0 による**外測度 outer measure**) として与えられる;

$$P(A) = \inf\left\{\sum_{n=1}^\infty P_0(A_n); A_n \in \mathcal{A}, \bigcup_{n=1}^\infty A_n \supset A\right\}.$$

ここで, $A \subset \Omega$ は任意の部分集合として, 定義できる.
上の (3.1) 「$\mathcal{A}$ 上 σ-加法的」と同値な条件を述べておく.

$$A_n \in \mathcal{A} \downarrow \emptyset \implies \lim_{n\to\infty} P_0(A_n) = 0 \tag{3.2}$$

(3.3) $$A_n \in \mathcal{A} \downarrow, \lim_{n\to\infty} P_0(A_n) > 0 \implies \bigcap_{n\geq 1} A_n \neq \emptyset$$

実際, 応用においては (3.3) をチェックすることが多い.

3.1 無限次元直積確率空間

$(\Omega_n, \mathcal{F}_n, P_n)$ を確率空間とする. n 個の有限次元確率空間を次のようにおく.

$$(\Omega^{(n)}, \mathcal{F}^{(n)}, P^{(n)}) := (\prod_{k=1}^n \Omega_k, \bigotimes_{k=1}^n \mathcal{F}_k, \bigotimes_{k=1}^n P_k)$$

$\Omega := \prod_{n=1}^\infty \Omega_n$ とおく.

$$\mathcal{A} = \left\{ A = A_n \times \Omega_{n+1} \times \Omega_{n+2} \times \cdots ; A_n \in \mathcal{F}^{(n)}, n = 1, 2, \dots \right\}$$

とおくと $\mathcal{A}$ は Ω 上の加法族となる. この元 $A = A_n \times \Omega_{n+1} \times \Omega_{n+2} \times \cdots$ に対し, $P_0(A) = P^{(n)}(A_n)$ とおけば, $P_0(\emptyset) = 0$ をみたす $\mathcal{A}$ 上の加法的集合関数となる. さらに拡張定理の条件 (3.3) をみたすことが示せる. 従って $^{\exists 1} P$ 確率測度 on $\sigma(\mathcal{A})$; $P = P_0$ on $\mathcal{A}$. 即ち, 各 n に対し,

$$P(A_1 \times \cdots \times A_n \times \Omega_{n+1} \times \Omega_{n+2} \times \cdots) = P_1(A_1) \cdots P_n(A_n) \quad (A_k \in \mathcal{F}_k)$$

をみたす $\sigma(\mathcal{A})$ 上の確率測度が一意的に存在する. このとき $\bigotimes_{n=1}^\infty \mathcal{F}_n := \sigma(\mathcal{A})$, $\bigotimes_{n=1}^\infty P_n := P$ と表し, $(\prod_{n=1}^\infty \Omega_n, \bigotimes_{n=1}^\infty \mathcal{F}_n, \bigotimes_{n=1}^\infty P_n)$ を $(\Omega_n, \mathcal{F}_n, P_n)$ $(n = 1, 2, \dots)$ の**無限次元直積確率空間 (infinite-dimensional product probability space)** という.

[独立確率変数の構成]

この結果により, 与えられた分布を持つ可算無限個の独立な確率変数を具体的に構成することが出来る. (しかも, 同分布でも, 全て異なる分布でも可能.) 例えば, 各 $n \geq 1$ に対し, μ_n が $(\mathbf{R}_+ = [0, \infty), \mathcal{B}(\mathbf{R}_+))$ 上の分布として与えられたとする. これの無限直積空間が存在するので, それを $(\Omega, \mathcal{F}, P)$ として, 更に, $\omega = (\omega_n) \in \Omega = \mathbf{R}_+^\infty$ に対し, 確率変数列 X_n を $X_n(\omega) = \omega_n$ と定義すれば, その分布は μ_n で, しかも独立となる. 実際,

$$P(X_1 \in A_1, X_2 \in A_2) = P(A_1 \times A_2 \times \mathbf{R}_+^\infty) = \mu_1(A_1)\mu_2(A_2)$$

上の定理は, $\Omega_n = \mathbf{R}$ の場合には, 次のように拡張できる.

3.2 Kolmogorov の拡張定理

各 $n \geq 1$ に対し, $(\mathbf{R}^n, \mathcal{B}^n, P_n)$ を n 次元確率空間とする. これが次の両立条件

$$P_n(A) = P_{n+1}(A \times \mathbf{R}) \quad (A \in \mathcal{B}^n)$$

をみたすとする.

$$\mathcal{B}^\infty = \mathcal{B}(\mathbf{R}^\infty) := \sigma(\bigcup_{n\geq 1} (\mathcal{B}^n \times \mathbf{R}^\infty))$$

を無限次元 Borel 集合体といい, このとき $P(A_n\times\mathbf{R}^\infty)=P_n(A_n)$ $(A_n\in\mathcal{B}^n)$ をみたす $(\mathbf{R}^\infty,\mathcal{B}^\infty)$ 上の確率測度 P が一意的に存在する.

実際, 上と同様に $\mathcal{A}=\{A=A_n\times\mathbf{R}^\infty;A_n\in\mathcal{B}^n,n=1,2,\dots\}$ とおき, この元 $A=A_n\times\mathbf{R}^\infty$ に対し, $P_0(A)=P_n(A_n)$ とおけば, $P_0(\emptyset)=0$ をみたす $\mathcal{A}$ 上の加法的集合関数で, 拡張定理の条件 (3.3) をみたすことが示せる. 従って $P(A_n\times\mathbf{R}^\infty)=P_n(A_n)$ $(A_n\in\mathcal{B}^n)$ をみたす $\mathcal{B}^\infty=\sigma(\mathcal{A})$ 上の確率測度 P が一意的に存在する.

次の 2 つの定理は一般の測度空間で成り立つが, 確率論バージョンを述べておく.

定理 3.1 (近似定理 (Approximating Theorem)) $(\Omega,\mathcal{F},P)$ を確率空間とする. $\mathcal{A}\subset\mathcal{F}$ が加法族なら,

$$^\forall A\in\sigma(\mathcal{A}),{}^\exists A_n\in\mathcal{A};\lim_{n\to\infty}P(A\triangle A_n)=0.$$

$(A\triangle B=(A\setminus B)\cup(B\setminus A)$: 対称差である.)

(証明) 定理の条件を満たす $A\subset\Omega$ の全体を $\mathcal{G}$ とすると, これは $\mathcal{A}$ を含む σ 加法族となることが分る. 従って, $\sigma(\mathcal{A})\subset\mathcal{G}$ となり, 題意を得る. ■

問 3.1 上の $\mathcal{G}$ が σ 加法族となることを示せ.

可算和については, $^\forall\varepsilon>0$ をとり, 各 $A_n\in\mathcal{G}$ に対し, その近似列で, $B_n\in\mathcal{A};P(A_n\triangle B_n)<\varepsilon/2^{n+1}$ なるものを 1 つずつ固定する. $A:=\bigcup A_n$ を有限和 $\bigcup_{n\le N}A_n$ で近似し, (上方連続性により, 差の確率を $\varepsilon/2$ で抑える.) 更にそれが, $C_N:=\bigcup_{n\le N}B_n\in\mathcal{A}$ で近似できるので, $A\in\mathcal{G}$ を得る. 実際,

$$A\triangle\left(\bigcup_{n=1}^N B_n\right)\subset\left(A\setminus\left(\bigcup_{n=1}^N A_n\right)\right)\cup\bigcup_{n=1}^N(A_n\triangle B_n)$$

より, 次が成り立つ.

$$P\left(A\triangle\left(\bigcup_{n=1}^N B_n\right)\right)\le P\left(A\setminus\left(\bigcup_{n=1}^N A_n\right)\right)+\sum_{n=1}^N P(A_n\triangle B_n)<\varepsilon.$$

■

系 3.1 $(\mathbf{R}^d,\mathcal{B}^d,P)$ を確率空間とする. $A\in\mathcal{B}^d$ に対し, $^\exists C_n$ 有界閉集合, $^\exists G_n$ 開集合; $C_k\subset A\subset G_k$, $\lim P(A\setminus C_n)=\lim P(G_n\setminus A)=0$.

(証明) $\mathcal{A}$ を基本矩形 $\prod_{k=1}^d(a_k,b_k]$ の有限和全体とすると, 加法族で, $\sigma(\mathcal{A})=\mathcal{B}^d$. 測度の拡張定理の証明にあるように, $P(A)$ は外測度でもあるので,

$$P(A)=\inf\left\{\sum_{n\ge1}P(A_n);A_n\in\mathcal{A},\bigcup_{n\ge1}A_n\supset A\right\}$$

より, $^\forall\varepsilon>0$ に対し, $\bigcup A_n$ で外から $\varepsilon/2$ 近似できる. 各 A_n は明らかに開区間の直積の有限和で $\varepsilon/2^{n+1}$ 近似できるので, その可算和も, 開集合で $\varepsilon/2$ 近似できることになる. 従って A が開集合で ε 近似できる. 補集合を考えれば, 閉集合近似できるので, それを更に, 有界閉集合で近似すれば良い. $(\{|x|\le n\}$ との共通部分を考えれば良い.) ■

3.3 無限個の独立性に関する話題

本節では, 各 $n \geq$ に対し, X_n を実数値確率変数, $\mathcal{F}_n$ を σ 加法族 $\mathcal{F}$ の部分 σ 加法族とする.

Borel-Cantelli の補題で, 事象列の確率の和が有限であれば, 上極限の確率が 0 となることは証明したが,

$$A_n \in \mathcal{F}, \sum P(A_n) < \infty \Longrightarrow P(\limsup A_n) = 0$$

では, この和が無限の時は, 確率が正となる, あるいは, 1 になるなどということは言えるのだろうか？これに対する 1 つの答えが次の結果である.

定理 3.2 (Borel-Cantelli の定理)
$A_n \in \mathcal{F}$ が独立のとき, $\sum P(A_n) = \infty$ なら $P(\limsup A_n) = 1$.

[証明] $P(\liminf A_n^c) = \lim_{N\to\infty} P(\bigcap_{n\geq N} A_n^c)$ で, $\{A_n^c\}$ も独立となることに注意して, $P(\bigcap_{n\geq N} A_n^c) = \prod_{n\geq N}(1 - P(A_n))$. ここで, $1 - x \leq e^{-x}$ と $\sum_{n\geq N} P(A_n) = \infty$ より, $P(\liminf A_n^c) = 0$ を得る. ■

σ 加法族 $\bigcap_{N\geq 1} \sigma\left(\bigcup_{n\geq N} \mathcal{F}_n\right)$ の元を**末尾事象 (tail event)** といい, これについて可測な確率変数を**末尾関数 (tail function)** という. また, $\mathcal{F}_n = \sigma(X_n) = X_n^{-1}\mathcal{B}^1)$ のとき, $\{X_n\}$ **に関する末尾事象, 末尾関数**という.

問 3.2 $A_n \in \mathcal{F}_n$ に対し, $\limsup A_n, \liminf A_n$ は末尾事象である. また $\{X_n \to 0\}$ も $\{X_n\}$ に関して, そうであることを示せ.

ラフに言えば, 有限個の $\mathcal{F}_1, \ldots, \mathcal{F}_n$ や $X_1, \ldots, X_n$ に無関係な事象, 確率変数をいう.

定理 3.3 (Kolmogorov の 0-1 法則)
$\mathcal{F}_n \subset \mathcal{F}$ が独立な部分 σ 加法族のとき, 末尾事象の確率は全て, 0 か 1.

(証明) まず $P(A) = 0$ or $1 \iff P(A)^2 = P(A) \iff A$ が自分自身と独立に注意する. これを示すのに, 近似定理を上手く用いて, 末尾事象 A と独立な近似事象 A_ε がとれることを使う.

$$\mathcal{G}_n = \sigma\left(\bigcup_{k\leq n} \mathcal{F}_k\right), \quad \mathcal{A} = \bigcup_{n\geq 1} \mathcal{G}_n$$

とおくと, $\mathcal{A}$ は加法族で, $\sigma(\mathcal{A}) = \sigma\left(\bigcup_{k\geq 1} \mathcal{F}_k\right)$ を満たす. $A \in \sigma(\mathcal{A})$ でもあるので, 近似定理より, ${}^\forall\varepsilon > 0, A_\varepsilon \in \mathcal{A}; P(A \triangle A_\varepsilon) < \varepsilon$. $\mathcal{A}$ の定義より, ${}^\exists n; A_\varepsilon \in \mathcal{G}_n$ また, $A \in \sigma\left(\bigcup_{k\geq n+1} \mathcal{F}_k\right)$ なので, A_ε とは独立. これから, A 自身が A と独立, i.e., $P(A) = P(A \cap A) = P(A)^2$ を得るので, $P(A) = 0$ or 1 となる. 実際, 一般に $A \subset (A \cap B) \cup (A \triangle B) \subset B \cup (A \triangle B)$ を用いて,

$$P(A) \leq P(A \cap A_\varepsilon) + P(A \triangle A_\varepsilon) < P(A)P(A_\varepsilon) + \varepsilon \leq P(A)(P(A) + P(A \triangle A_\varepsilon)) + \varepsilon < P(A)^2 + 2\varepsilon.$$

同様にして,

$$P(A) \geq P(A \cap A_\varepsilon) - P(A \triangle A_\varepsilon) > P(A)P(A_\varepsilon) - \varepsilon \geq P(A)(P(A) - P(A \triangle A_\varepsilon)) - \varepsilon > P(A)^2 - 2\varepsilon.$$

■

参考文献

[1] 熊谷　隆「確率論」　共立出版　(2003)

[2] 舟木　直久　著「確率論」　朝倉書店　(2004)

[3] 伊藤　雄二　著「確率論」　朝倉書店　(2002)

[4] 西尾　真喜子「確率論」　実教出版　(1978 初版, 1985 第 5 版)

[5] 志賀　徳造「ルベーグ積分から確率論」　共立出版　(2000)