確率統計学 II

杉浦　誠

2021 年 12 月 15 日

目次

確率統計学 II では [AEL] にしたがって **1** 大数の法則, **2** 中心極限定理, について解説します。
(内容的には [F] や [D] を教科書として作成した講義ノートと言えます。)
講義ノートは数ページずつ WebClass にアップロードしていきます。全体は

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

に置きますが、最新版の WebClass のものを参照ください。

参考文献

[AEL] 浅野, 江島, 李 共著 基本統計学 森北出版

[F] 舟木 直久 著 確率論 朝倉書店

[D] Durrett, R.: Probablity Theory and Examples, 4th ed., (2010), Cambridge Series in Statistical and Probabilistic Mathematics.

授業は学年歴の水曜日授業日の 2 限に複合棟 412 室で行います。振替日の 11 月 26 日 (金) も授業を行う予定です。出席を希望する人は感染対策をして出席ください。

ただし、**発熱等体調に不良がある人、新型コロナウイルスに感染した若しくは感染したおそれ等がある人は大学に来てはいけません**。特に、下記の中間テスト、期末テストの日にそうなった場合は連絡ください。後日追試を行います。

また、WebClass に講義ノートを掲載するので、必ずしも講義に出席する必要はありません。毎回、授業終了後に次回分以降の講義ノートを少しづつ WebClass にアップロードしていく予定ですが、その資料の「説明/注意点」の欄にその日の授業でどこまで進んだかを記載して行く予定です。WebClass での質問を受け付けます。

成績について　中間テスト (12 月 22 日) と期末テスト (2 月 2 日) の合計点数で判定します。6 割以上の得点で合格です。テストの範囲は遅くとも二週間前には WebClass を通じて連絡します。確率統計学 II の受講者は中間テストと期末テストを必ず受けて下さい。

5　大数の法則

5.1　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 5.1 (1) **(概収束)** X_n が X に概収束 (almost surely convergence) するとは、P-a.a. ω 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 5.1 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 5.2 とします。) 概収束は距離付けできない ($cf.$ [F] p.59)。

問題 5.1 ((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。

(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 5.2 $X \in L^r$ 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)　$X, Y \in L^2$ に対して $\|X + Y\|_2 \leq \|X\|_2 + \|Y\|_2$ を示せ。

　　ヒント: Cauchy-Schwarz の不等式 $(E[|XY|])^2 \leq E[|X|^2]E[|Y|^2]$ を用いよ。

(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 5.3 $a > 0$ 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 5.4 $\{X_n\}$ 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。

次に定義 5.1 で定めた収束の強弱の関係を調べる。

定理 5.1 (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 ω の集合は

$$\left\{\lim_{n\to\infty} X_n = X\right\} = \bigcap_{j=1}^{\infty}\bigcup_{n=1}^{\infty}\bigcap_{m=n}^{\infty} A_{m,j} \tag{5.1}$$

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、(5.1) 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、

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

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、

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

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

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

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 5.2 (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。

$$P(|Y| \ge \lambda) \le \frac{1}{\lambda^r} E[|Y|^r]$$

証明: まず、次に注意する。

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

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

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

定理 5.1(2) の証明: 仮定と Chebyshev の不等式により

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

となる。 $\square$

例 5.3 定理 5.1(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 5.5 (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 5.6 (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 5.4 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 5.4 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 5.5 (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 5.4 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。 □

5.2 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 5.2 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 5.6 $X_1, X_2, \ldots$ 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 5.1 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 5.7 (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 5.6 より $\forall \varepsilon > 0$ に対して

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

$$P\left(e^{(l-\varepsilon)n} \le Y_n \le e^{(l+\varepsilon)n}\right) \to 1 \tag{5.2}$$

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 (5.2) 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、(5.2) 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 5.4 と同様に証明できるので、ここで扱う。

定理 5.8 (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。

$$\lim_{n\to\infty}\max_{0\le p\le 1}\left|f(p)-\sum_{k=0}^{n} f\Big(\frac{k}{n}\Big)\binom{n}{k}p^k(1-p)^{n-k}\right| = 0 \tag{5.3}$$

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、

$$E\Big[f\Big(\frac{S_n}{n}\Big)\Big] = \sum_{k=0}^{n} f\Big(\frac{k}{n}\Big)P(S_n = k) = \sum_{k=0}^{n} f\Big(\frac{k}{n}\Big)\binom{n}{k}p^k(1-p)^{n-k}. \tag{5.4}$$

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 (5.4) とあわせて、(5.3) は示された。 □

問題 5.7 (1) $\{X_k\}$ は独立な確率変数列で、各 X_k の密度関数が $f_{X_k}(x) = \frac{3}{4k}\Big(1-\frac{x^2}{k^2}\Big)1_{(-k,k)}(x)$ であるとする。$a > 0$ を定数とし、$Y_n = \frac{1}{n^a}\sum_{k=1}^{n}X_k$ 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、$S_n = \sum_{k=1}^{n}X_k$ 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 5.9–5.11) 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[*1]。(関数解析学 I,II の講義の教科書を調べてください。)

定理 5.9 (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 5.10 (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 5.11 (Fubini の定理) $(R_i, \mathfrak{A}_i, \mu_i)$, $i = 1, 2$, を二つの σ-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[*2]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。

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

定理 5.12 $X_1, X_2, \ldots$ 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、$S_n = \sum_{k=1}^{n} X_k$, $a_n = \sum_{k=1}^{n} E[X_k 1_{\{|X_k|\le b_n\}}]$ 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、

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

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、

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

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

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

定理 5.13 $X_1, X_2, \ldots$ は i.i.d. で、

$$xP(|X_1| > x) \to 0 \qquad (x \to \infty) \tag{5.5}$$

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

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

[*2] 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)。

注意 5.2 定理 5.13 の仮定は、$\dfrac{S_n}{n} - c_n$ 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 ($cf.$ Feller, W.: An Introduction to Probability Theory and Its Applications, vol.II, (1971) pp.234–6)。

証明: $X_1, X_2, \ldots$ は i.i.d. なので、定理 5.12 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 5.12 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 (5.5) より明らか。(b) のために次の補題を準備する。

補題 5.14 $Y \geq 0$, $p > 0$ 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 5.11) を用いた。 □

定理 5.13 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 5.14 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 5.8(2) から、(5.5) より定理 5.12 の条件 (b) が成り立つことがわかる。 □

問題 5.8 (1) 数列 $\{a_n\}$ が $\lim\limits_{k\to\infty} a_k = \alpha$ を満たせば、$\lim\limits_{n\to\infty}\dfrac{1}{n}\sum\limits_{k=1}^{n} a_k = \alpha$ 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 $\lim\limits_{x\to\infty}\varphi(x) = 0$ を満たせば、$\lim\limits_{n\to\infty}\dfrac{1}{n}\displaystyle\int_0^n \varphi(x)\,dx = 0$ となることを示せ。(ともに微積の問題です。)

定理 5.15 $X_1, X_2, \ldots$ が i.i.d. で $E[|X_1|] < \infty$ であれば、$S_n = \sum\limits_{k=1}^{n} X_k$, $m = E[X_1]$ とすると、$\dfrac{S_n}{n}$ は m に確率収束する。

証明: $E[|X_1|] < \infty$ より $|X_1| < \infty$ a.s. であるから、$x \to \infty$ のとき $|X_1|1_{\{|X_1|>x\}} \to 0$ a.s. となる。よって、$\big||X_1|1_{\{|X_1|>x\}}\big| \leq |X_1|$ かつ $E[|X_1|] < \infty$ より Lebesgue の収束定理から

$$xP(|X_1| > x) = xE[1_{\{|X_1|>x\}}] \leq E[|X_1|1_{\{|X_1|>x\}}] \to E[0] = 0, \qquad x \to \infty.$$

よって、定理 5.13 より $\dfrac{S_n}{n} - c_n \to 0$ in prob. ただし、$c_n = E[X_1 1_{\{|X_1|\leq n\}}]$. 一方、$\big|X_1 1_{\{|X_1|\leq n\}}\big| \leq |X_1|$ かつ $E[|X_1|] < \infty$ より Lebesgue の収束定理から

$$c_n = E[X_1 1_{\{|X_1|\leq n\}}] \to E[X_1] = m, \qquad n \to \infty$$

となり問題 5.1 (2) から主張は従う。 □

前回の定理 5.13 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 5.5)。

例 5.16 $X_1, X_2, \ldots$ は i.i.d. で、$P(X_1 = (-1)^j j) = \frac{K}{j^2 \log j}$, $j = 2, 3, \ldots,$ 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 5.13 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、$c = K \sum_{j=2}^{\infty} \dfrac{(-1)^j}{j \log j}$ とおくとこの右辺は次の問題 5.9 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 5.9 $\{a_n\}$ 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。)

定理 5.12 の応用として、次の例を見ておこう。

例 5.17 (サンクトペテルスブルグのパラドックス) $X_1, X_2, \ldots$ を i.i.d. で $P(X_1 = 2^i) = 1/2^i$, $i = 1, 2, \ldots,$ となるとする。このとき、$E[X_1] = \infty$ であり、$S_n = \sum_{k=1}^{n} X_k$ とおくと、$\forall \varepsilon > 0$ に対して次が成立する。

$$P\Big(\Big|\frac{S_n}{n \log_2 n} - 1\Big| \leq \varepsilon\Big) \to 1, \qquad n \to \infty. \tag{5.6}$$

この X_k は、公正なコインを表が出るまで投げ続け、i 回目に表が初めて出たとき 2^i 円受け取る宝くじを表す確率変数と考えられる。この宝くじはいくらの価値があるかであるが、$E[X_k] = \infty$ よりいくら出しても購入する価値がありそうである。しかし、この宝くじで 2 億円以上獲得するためには、$2^{28} = 268,435,456$ より 28 回目以降に初めて表が出る必要がある。その確率は 1.3 億分の 1 以下である。したがって、それほどの価値があるとは思えない。これに対して (5.6) は n が十分大きければ、n 本のセットで $n \log_2 n$ 円の価値があることを表している。例えば 2^{28} 本売るのであれば、一本あたり 28 円となる。

証明: $b_n = n \log_2 n$ とし $c_n = \lfloor \log_2 b_n \rfloor$ とする ($\lfloor a \rfloor$ は a の整数部分を表す)。このとき、

$$c_n \leq \log_2 b_n < c_n + 1 \text{ より } 2^{c_n} \leq b_n < 2^{c_n+1}$$

に注意する。よって、$n \to \infty$ のとき、

$$\text{(a)} \qquad \sum_{k=1}^{n} P(|X_k| > b_n) = nP(X_1 \geq 2^{c_n+1}) = n \sum_{i=c_n+1}^{\infty} \frac{1}{2^i} = n \frac{1/2^{c_n+1}}{1 - 1/2} = \frac{n}{2^{c_n}} < \frac{n}{2^{-1} b_n} = \frac{2}{\log_2 n} \to 0,$$

$$\text{(b)} \qquad \frac{1}{{b_n}^2} \sum_{k=1}^{n} E[{X_k}^2 1_{\{|X_k| \leq b_n\}}] = \frac{n}{{b_n}^2} E[{X_1}^2 1_{\{X_1 \leq b_n\}}] = \frac{n}{{b_n}^2} \sum_{i=1}^{c_n} (2^i)^2 \frac{1}{2^i} = \frac{n}{{b_n}^2} \frac{2(2^{c_n} - 1)}{2 - 1}$$

$$\leq \frac{2n2^{c_n}}{{b_n}^2} \leq \frac{2nb_n}{{b_n}^2} = \frac{2}{\log_2 n} \to 0.$$

よって、定理 5.12 より $a_n = \sum_{k=1}^{n} E[X_k 1_{\{|X_k| \le b_n\}}]$ とすると、$\dfrac{S_n - a_n}{b_n}$ は 0 に確率収束する。ここで、

$$\frac{a_n}{b_n} = \frac{n}{b_n} E[X_1 1_{\{|X_1| \le b_n\}}] = \frac{n}{b_n} \sum_{i=1}^{c_n} 2^i \frac{1}{2^i} = \frac{nc_n}{b_n} = \frac{\lfloor \log_2(n \log_2 n) \rfloor}{\log_2 n} = \frac{\lfloor \log_2 n + \log_2 \log_2 n \rfloor}{\log_2 n} \to 1.$$

最後の極限は対数関数の性質

$$\lim_{n\to\infty} \log_2 n = \infty \qquad \text{かつ} \qquad \lim_{n\to\infty} \frac{\log_2 \log_2 n}{\log_2 n} = 0$$

に注意すれば容易に示せる。以上より (5.6) を得る。 $\square$

問題 5.10 $p_j = 1/\{2^j j(j+1)\}$, $j = 1, 2, \ldots$, $p_0 = 1 - \sum_{j=1}^{\infty} p_j$ とし、$P(X_1 = -1) = p_0$, $P(X_1 = 2^j - 1) = p_j$, $j = 1, 2, \ldots$, とする。このとき、$E[X_1] = 0$ となるが、$\dfrac{S_n}{n/\log_2 n}$ は -1 に確率収束することを示せ。ヒント: $b_n = \dfrac{n}{\log_2 n}$ として定理 5.12 を用いよ ($cf.$ 例 5.17, 難しいので脚注[*3]に略解を述べておく。)

注意 5.3 定理 5.13 の仮定の下 ($X_1, X_2, \ldots$ は i.i.d. とする)、$m \notin [a, b]$ なら $P(a \le \frac{1}{n} \sum_{k=1}^{n} X_k \le b)$ 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 5.12)。

次の 2 つの問題は Stirling の公式 $n! \sim \sqrt{2\pi} n^{n+\frac{1}{2}} e^{-n}$ を用いる。$a_n \sim b_n$ とは $\lim_{n\to\infty} \dfrac{a_n}{b_n} = 1$ を表す。

問題 5.11 $0 < p < 1$ とする。X_n が二項分布 $B(3n, p)$ に従うとする。$\lim_{n\to\infty} \dfrac{1}{n} \log P(X_n = n)$ を求めよ。

問題 5.12 $X_1, X_2, \ldots$ 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 $\lim_{n\to\infty} \dfrac{1}{2n} \log P\Big(\dfrac{S_{2n}}{2n} \ge a\Big) = -\psi(a)$ 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 $\lim_{n\to\infty} \dfrac{a_n}{b_n} = 1$ を表す。

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

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

(3) $l_n = \lceil na \rceil$ のとき $P(S_{2n} = 2l_n) \le P(S_{2n} \ge 2na) \le nP(S_{2n} = 2l_n)$ 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 $P\big(\frac{S_{2n}}{2n} \ge a\big)$ はおおよそ $e^{-2n\psi(a)}$ の速さで減衰することがわかる。

[*3] **問題 5.10 の略解:** $E[X_k] = -p_0 + \sum_{j=1}^{\infty} (2^j - 1) p_j = -p_0 + \sum_{j=1}^{\infty} (\frac{1}{j(j+1)} - p_j) = -1 + \sum_{j=1}^{\infty} (\frac{1}{j} - \frac{1}{j+1}) = 0$. $b_n = n/\log_2 n$, $c_n = \lfloor \log_2(b_n + 1) \rfloor$ とし $2^{c_n} \le b_n + 1 < 2^{c_n+1}$ に注意する。$n \to \infty$ のとき、定理 5.12(a) は $\sum_{k=1}^{n} P(|X_k| > b_n) = n \sum_{j=c_n+1}^{\infty} p_j \le \frac{n}{(c_n+1)^2} \sum_{j=c_n+1}^{\infty} \frac{1}{2^j} = \frac{n}{2^{c_n}(c_n+1)^2} \le \frac{n}{2^{-1} b_n (\log_2 b_n)^2} = \frac{2\log_2 n}{(\log_2 n - \log_2 \log_2 n)^2} \to 0$. $0 < \alpha < 1$ とする。$(b) := \frac{n}{b_n^2} E[X_1^2 1_{\{|X_1| \le b_n\}}] = \frac{n}{b_n^2}(p_0 + \sum_{j=1}^{c_n} (2^j - 1)^2 p_j) \le \frac{n}{b_n^2}(1 + \sum_{j=1}^{c_n} \frac{2^j}{j^2}) < \frac{n}{b_n^2}(1 + \int_1^{c_n+1} \frac{2^{t+1}}{t^2} dt)$ で $\int_1^{c_n+1} \frac{2^t}{t^2} dt = \{\int_1^{\alpha(c_n+1)} + \int_{\alpha(c_n+1)}^{c_n+1}\} \frac{2^t}{t^2} dt \le 2^{\alpha(c_n+1)}(1 - \frac{1}{\alpha(c_n+1)}) + \frac{2^{c_n+1}}{c_n+1}(\frac{1}{\alpha} - 1)$ より $(b) \le \frac{n}{b_n^2} + \frac{2 \cdot 2^{\alpha} n}{b_n^{2-\alpha}} + \frac{n}{b_n^2} \frac{4b_n}{\log_2 b_n}(\frac{1}{\alpha} - 1) \le \frac{(\log_2 n)^2}{n} + \frac{2(\log_2 n)^{2-\alpha}}{n^{1-\alpha}} + \frac{4\log_2 n}{\log_2 \frac{n}{\log_2 n}}(\frac{1}{\alpha} - 1)$. よって、$\forall \varepsilon > 0$ に対し α を 1 に十分近くとっておけば右辺の第 3 項 $< \varepsilon/2$ $(\forall n \ge 4)$ とでき、この α を固定して $n \to \infty$ とすると右辺の第 1 項、第 2 項はともに $\to 0$ となり、$(b) \to 0$ がわかった。よって、$a_n = nE[X_1 1_{\{|X_1| \le b_n\}}]$ とおくと定理 5.12 より $\frac{S_n - a_n}{b_n}$ は 0 に確率収束する。ここで、$\frac{a_n}{b_n} = \frac{n}{b_n}(\sum_{j=c_n+1}^{\infty} p_j - \frac{1}{c_n+1})$ で $0 \le \frac{n}{b_n} \sum_{j=c_n+1}^{\infty} p_j \le \frac{n}{b_n} \frac{2^{-c_n}}{(c_n+1)^2} \le \frac{2n}{b_n^2} \to 0$, $\frac{n}{b_n(c_n+1)} = \frac{\log_2 n}{\lfloor \log_2(\frac{n}{\log_2 n} + 1) \rfloor + 1} \to 1$ となり、$\frac{a_n}{b_n} \to -1$ が示され、主張が従う。

5.3 大数の強法則

定理 5.18 (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

$$A_k^* = \{\,\omega \in \Omega\,;\, j = 1, 2, \ldots, k-1 \text{ に対しては } |S_j(\omega)| < a \text{ で、かつ } |S_k(\omega)| \ge a\} \tag{5.7}$$

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、(5.7) 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 5.19 (Kolmogorov の第 1 定理) $X_1, X_2, \ldots$ が独立な確率変数列で、

$$\sum_{n=1}^{\infty} \frac{1}{n^2} V(X_n) < \infty \tag{5.8}$$

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、

$$P(A(\varepsilon)) = 1 \tag{5.9}$$

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、(5.9) 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> (5.9) 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

$$A(\varepsilon)^{\mathrm{c}} = \bigcap_{N=1}^{\infty}\bigcup_{n=N}^{\infty}\{|Y_n| \geq \varepsilon\} \subset \bigcup_{m=l}^{\infty} B_m(\varepsilon) \tag{5.10}$$

だから、(5.9), すなわち $P(A(\varepsilon)^{\mathrm{c}}) = 0$ を示すためには

$$\sum_{m=1}^{\infty} P(B_m(\varepsilon)) < \infty \tag{5.11}$$

を示せばよい。実際、Borel-Cantelli の定理により $P\Big(\bigcap_{l=1}^{\infty}\bigcup_{m=l}^{\infty} B_m(\varepsilon)\Big) = 0$ であるが、(5.10) 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> (5.11) 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 5.6) 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 (5.8) より (5.11) が示された。 □

$\{X_n\}$ の分布が同じならば、定理 5.19 の仮定 (5.8)、特に $E[X_n^2] < \infty$ は不要になる。

定理 5.20 (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。
1st step (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 5.19 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 5.19 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。
2nd step $|X 1_{(0,k]}(X)| \le |X|$ $(\forall k \in \mathbf{N})$ で $E[|X|] < \infty$ なので、Lebesgue の収束定理 (定理 5.10) 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 5.8(1))、1st step により

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

3rd step $P(\#\{k \in \mathbf{N}; |X_k| > k\} < \infty) = 1$ を示す。これがいえれば、a.a. ω 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

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

に注意する。ただし、2 行目の一つ目の不等号は $\sum_{j=1}^{\infty} j 1_{(j,j+1]}(|x|) \le \sum_{j=1}^{\infty} |x| 1_{(j,j+1]}(|x|) \le |x|$ 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 5.4 定理 5.20 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 5.21 $X_1, X_2, \ldots$ は i.i.d. で $E[|X_1|] = \infty$ となるとき、

$$P\left(\limsup_{n\to\infty}\Big|\frac{1}{n}\sum_{k=1}^{n} X_k\Big| = \infty\right) = 1. \tag{5.12}$$

定理 5.22 (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、

$$P((\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty} B_k)^c) = \lim_{n\to\infty} P(\bigcap_{k=n}^{\infty} B_k^{\mathrm{c}}) = \lim_{n\to\infty}\lim_{N\to\infty} P(\bigcap_{k=n}^{N} B_k^{\mathrm{c}}). \tag{5.13}$$

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、$P(B_k^{\mathrm{c}}) = 1 - P(B_k) \leq e^{-P(B_k)}$ となること[*4]を用いて、

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

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、(5.13) 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 5.21 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、$B_n^M = \{|X_n| \geq Mn\}$ とすると、

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

ここで、2 行目の二つ目の等号は $P(Mk \leq |X| < M(k+1))$ 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 $k \geq n$ があって $\omega \in B_k^M$、i.e., $\frac{|X_k(\omega)|}{k} \geq M$ となるから、

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

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

次に、数列 $\{a_n\}$ に対して

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

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

となることに注意する。これは、

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

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、

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

となり、(5.12) が成り立つことがわかった。 □

注意 5.5 例 5.16 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 5.21 より大数の強法則を満たさない。

5.4 点推定

定義 5.3 (1) 統計量 $T(X_1, X_2, \ldots, X_n)$ が母数 θ の**不偏推定量** (unbiased estimator) であるとは $E[T(X_1, X_2, \ldots, X_n)] = \theta$ が成り立つことである。
(2) 統計量 $T(X_1, X_2, \ldots, X_n)$ が母数 θ の**一致推定量** (consistent estimator) であるとは、$\forall \varepsilon > 0$ に対して

$$\lim_{n\to\infty} P\left(|T(X_1, X_2, \ldots, X_n) - \theta| \geq \varepsilon\right) = 0$$

が成り立つときにいう。

例 5.23 $X_1, X_2, \ldots, X_n$ を無作為標本、すなわち、i.i.d. となる確率変数列で $E[|X_1|] < \infty$ とする。標本平均 $\overline{X} = \dfrac{1}{n}(X_1 + \cdots + X_n)$ は母平均 $\mu = E[X_1]$ の不偏推定量であり、また一致推定量でもある。実際、$E[\overline{X}] = \mu$ より不偏推定量である。また、大数の弱法則 (定理 5.15) より一致推定量であることもわかる。

問題 5.13 $X_1, X_2, \ldots, X_n$ を無作為標本とし、$U^2 = \dfrac{1}{n-1} \displaystyle\sum_{i=1}^{n} (X_i - \overline{X})^2$ とおく。
(1) $E[X_1^2] < \infty$ のとき、U^2 が母分散 σ^2 の不偏推定量であることを示せ。これより U^2 を不偏分散という。
(2) $X_1, X_2, \ldots, X_n$ が正規母集団 $N(\mu, \sigma^2)$ からの無作為標本とするとき、$E[(U^2 - \sigma^2)^2]$ を求め、U^2 が一致推定量でもあることを (定理 5.6(大数の弱法則) の証明と同様にして) 示せ。
ヒント: (2) には $\dfrac{n-1}{\sigma^2} U^2$ が自由度 $n-1$ の χ^2 分布に従うことを用いよ。

定義 5.4 二つの統計量 $T_1(X_1, X_2, \ldots, X_n), T_2(X_1, X_2, \ldots, X_n)$ が母数 θ の不偏推定量であるとする。このとき、$V(T_1) < V(T_2)$ が成り立つといい、T_1 が T_2 より**有効**であるという。特に、不偏推定量の中で分散が最小となるものを**有効推定量**という。

例題 5.24 $X_1, X_2, \ldots, X_n$ を i.i.d. で一様分布 $\mathrm{U}(0, \theta)$ に従っており、θ が未知母数とする。このとき、

$$T_1(X_1, \ldots, X_n) = \frac{2}{n}(X_1 + \cdots + X_n), \qquad T_2(X_1, \ldots, X_n) = \frac{n+1}{n} \max\{X_1, \cdots, X_n\}$$

がともに θ の不偏推定量となることを示せ。また、T_1, T_2 のどちらが有効であるか調べよ。

解: $E[T_1] = \dfrac{2}{n}(E[X_1] + \cdots + E[X_n]) = \dfrac{2}{n} \cdot n\dfrac{\theta}{2} = \theta$. 一方、$Y = \max\{X_1, \cdots, X_n\}$ とすると、

$$P(Y \leq y) = P(X_1 \leq y, \cdots, X_n \leq y) = P(X_1 \leq y) \cdots P(X_n \leq y) = \left(\frac{y}{\theta}\right)^n \quad (0 \leq y < \theta)$$

$P(Y \leq y) = 0\ (y < 0)$, $P(Y \leq y) = 1\ (y \geq 1)$ より、$f_Y(y) = \dfrac{ny^{n-1}}{\theta^n}\ (0 < y < \theta), = 0$ (その他). 従って、

$$E[T_2] = \frac{n+1}{n}\int_0^\theta y\frac{ny^{n-1}}{\theta^n}\,dy = \frac{n+1}{n}n\frac{1}{n+1}\theta = \theta.$$

以上より、T_1, T_2 ともに不偏推定量である。

次に、$V(T_1) = \dfrac{4}{n^2}(V(X_1) + \cdots + V(X_n)) = \dfrac{4}{n^2}\cdot n\Big(\dfrac{\theta^2}{3} - \dfrac{\theta^2}{2^2}\Big) = \dfrac{\theta^2}{3n}$.

一方、$E[Y^2] = \displaystyle\int_0^\theta y^2\frac{ny^{n-1}}{\theta^n}\,dy = \frac{n}{n+2}\theta^2$ より

$$V(T_2) = \Big(\frac{n+1}{n}\Big)^2 E[Y^2] - (E[T_2])^2 = \frac{(n+1)^2}{n^2}\frac{n}{n+2}\theta^2 - \theta^2 = \frac{\theta^2}{n(n+2)}.$$

よって、$n \geq 2$ のとき $V(T_1) > V(T_2)$ すなわち T_2 は T_1 より有効となる。($n = 1$ のとき $T_1 = T_2$.) □

定理 5.25 (Cramér-Rao) $X_1, \ldots, X_n$ が i.i.d. で、θ が未知母数とし、その密度関数を $f(x|\theta)$ または確率関数を $p(x|\theta) = P(X = x)$ とする。このとき $T(X_1, \ldots, X_n)$ が母数 θ の不偏推定量であれば、適切な条件下で

$$V(T) \geq \frac{1}{nI(\theta)} \tag{5.14}$$

である。ここで、$I(\theta)$ は Fisher 情報量とよばれ次で定義される。

$$I(\theta) = E\Big[\Big(\frac{\partial \log f(X|\theta)}{\partial\theta}\Big)^2\Big] \quad (\text{密度関数があるとき}) \qquad = E\Big[\Big(\frac{\partial \log p(X|\theta)}{\partial\theta}\Big)^2\Big] \quad (\text{離散型のとき})$$

証明: 密度関数を持つときのみ示す。密度関数であることと、T が不偏推定量であるから、

$$\theta = \int_{-\infty}^\infty\cdots\int_{-\infty}^\infty T\prod_{j=1}^n f(x_j|\theta)\,dx_1\cdots dx_n, \qquad 1 = \int_{-\infty}^\infty\cdots\int_{-\infty}^\infty \prod_{j=1}^n f(x_j|\theta)\,dx_1\cdots dx_n$$

が成立する。上の二式を θ に関する微分が積分の記号内で行えるとし、この両辺を θ で微分し、

$$1 = \int_{-\infty}^\infty\cdots\int_{-\infty}^\infty T\frac{\partial}{\partial\theta}\Big\{\prod_{j=1}^n f(x_j|\theta)\Big\}\,dx_1\cdots dx_n, \qquad 0 = \int_{-\infty}^\infty\cdots\int_{-\infty}^\infty \frac{\partial}{\partial\theta}\Big\{\prod_{j=1}^n f(x_j|\theta)\Big\}\,dx_1\cdots dx_n$$

を得る。これから $\dfrac{\partial}{\partial\theta}g(\theta) = \Big(\dfrac{\partial}{\partial\theta}\log g(\theta)\Big)g(\theta)$ に注意して

$$\begin{aligned}1 &= \int_{-\infty}^\infty\cdots\int_{-\infty}^\infty (T-\theta)\Big(\frac{\partial}{\partial\theta}\log\Big\{\prod_{j=1}^n f(x_j|\theta)\Big\}\Big)\prod_{j=1}^n f(x_j|\theta)\,dx_1\cdots dx_n \\ &= E\Big[(T-\theta)\Big(\frac{\partial}{\partial\theta}\log\Big\{\prod_{j=1}^n f(X_j|\theta)\Big\}\Big)\Big].\end{aligned}$$

ここで Cauchy-Schwarz の不等式 ($(E[XY])^2 \leq E[X^2]E[Y^2]$) を用いて

$$\Big|E\Big[(T-\theta)\Big(\frac{\partial}{\partial\theta}\log\Big\{\prod_{j=1}^n f(X_j|\theta)\Big\}\Big)\Big]\Big|^2 \leq E\left[(T-\theta)^2\right] E\Big[\Big(\frac{\partial}{\partial\theta}\log\Big\{\prod_{j=1}^n f(X_j|\theta)\Big\}\Big)^2\Big]. \tag{5.15}$$

ここで $Y_i = \dfrac{\partial}{\partial\theta}\log f(X_i|\theta)$ とすると、$I(\theta) = E[Y_i^2]$ で

$$E[Y_i] = \int_{-\infty}^\infty\Big\{\frac{\partial}{\partial\theta}\log f(x|\theta)\Big\}f(x|\theta)\,dx = \int_{-\infty}^\infty\frac{\partial}{\partial\theta}f(x|\theta)\,dx = \frac{\partial}{\partial\theta}\Big(\int_{-\infty}^\infty f(x|\theta)\,dx\Big) = 0.$$

$Y_1, \ldots, Y_n$ は独立であるから

$$E\Big[\Big(\frac{\partial}{\partial\theta}\log\Big\{\prod_{j=1}^n f(X_j|\theta)\Big\}\Big)^2\Big] = E[(Y_1 + \cdots + Y_n)^2] = \sum_{i=1}^n E[Y_i^2] + 2\sum_{1\leq i<j\leq n} E[Y_iY_j] = nI(\theta)$$

となり、$V(T) = E[(T-\theta)^2]$ に注意して (5.14) を得る。 □

注意 5.6 (5.14) で等号成立はある定数 c に対して

$$\frac{\partial}{\partial\theta}\log\Big\{\prod_{j=1}^{n} f(X_j|\theta)\Big\} = c(T-\theta) \qquad \text{a.e.}$$

が成立するときである。これは (5.15) で用いた Cauchy-Schwarz の不等式における等号成立の条件による。

注意 5.7 Fisher 情報量は密度関数 $f(x|\theta)$ もしくは確率関数が $p(x|\theta)$ が "よい" 関数ならば、以下のような表示も持つ。

$$I(\theta) = -E\Big[\frac{\partial^2 \log f(X|\theta)}{\partial\theta^2}\Big] \quad (\text{密度関数があるとき}) \qquad = -E\Big[\frac{\partial^2 \log p(X|\theta)}{\partial\theta^2}\Big] \quad (\text{離散型のとき})$$

以下、密度関数を持つときのみ説明する。$f(x|\theta)$ が θ について 2 回微分可能で積分 $\displaystyle\int_a^b f(x|\theta)\,dx$ が θ について微分と積分が交換できれば、$\displaystyle\int_{-\infty}^{\infty} f(x|\theta)\,dx = 1$ より

$$\int_{-\infty}^{\infty} \frac{\partial^2}{\partial\theta^2} f(x|\theta)\,dx = \frac{\partial^2}{\partial\theta^2}\left(\int_{-\infty}^{\infty} f(x|\theta)\,dx\right) = 0.$$

よって、$\displaystyle\frac{\partial}{\partial\theta}\log f(x|\theta) = \frac{1}{f(x|\theta)}\frac{\partial f}{\partial\theta}(x|\theta),\ \frac{\partial^2}{\partial\theta^2}\log f(x|\theta) = \frac{1}{f(x|\theta)}\frac{\partial^2 f}{\partial\theta^2}(x|\theta) - \Big(\frac{1}{f(x|\theta)}\frac{\partial f}{\partial\theta}(x|\theta)\Big)^2$ より、

$$\begin{aligned}
E\Big[\frac{\partial^2}{\partial\theta^2}\log f(x|\theta)\Big] &= E\Big[\frac{1}{f(X|\theta)}\frac{\partial^2}{\partial\theta^2} f(X|\theta)\Big] - E\Big[\Big\{\frac{1}{f(X|\theta)}\frac{\partial f}{\partial\theta}(X|\theta)\Big\}^2\Big] \\
&= \int_{-\infty}^{\infty}\Big\{\frac{1}{f(x|\theta)}\frac{\partial^2}{\partial\theta^2} f(x|\theta)\Big\} f(x|\theta)\,dx - E\Big[\Big\{\frac{1}{f(X|\theta)}\frac{\partial f}{\partial\theta}(X|\theta)\Big\}^2\Big] \\
&= 0 - E\Big[\Big\{\frac{\partial}{\partial\theta}\log f(X|\theta)\Big\}^2\Big] = -I(\theta)
\end{aligned}$$

となる。

例 5.26 正規母集団 $\mathrm{N}(\mu,\sigma^2)$ からの無作為標本 $X_1,\ldots,X_n$ とする。このとき、例 5.23 より標本平均 $\overline{X}$ は μ の不偏推定量であるが、これは有効推定量にもなっている。ただし、σ^2 は既知とする。

証明: $V(\overline{X}) = \dfrac{1}{n}\sigma^2$. 一方、$\mathrm{N}(\mu,\sigma^2)$ の密度関数 $f(x|\mu) = \dfrac{1}{\sqrt{2\pi}\sigma}\exp\Big\{-\dfrac{(x-\mu)^2}{2\sigma^2}\Big\}$ に対して、Fisher 情報量は

$$I(\mu) = E\Big[\Big(\frac{\partial \log f(X|\mu)}{\partial\mu}\Big)^2\Big] = E\Big[\Big(\frac{X-\mu}{\sigma^2}\Big)^2\Big] = \frac{1}{\sigma^2}$$

であるから、Cramér-Rao の不等式の等号が成立する。よって、$\overline{X}$ は有効推定量である。 □

問題 5.14 $n \geq 2$ とする。$X_1, X_2, \ldots, X_n$ を i.i.d. で指数分布 $\mathrm{Ex}(1/\lambda)$ に従っており、母平均 λ の不偏推定量 $\overline{X}$ と $T = c\min\{X_1,\cdots,X_n\}$ を考える。

(1) 定数 c を定めよ。 (2) $\overline{X}$ と T のどちらが有効か調べよ。 (3) $\overline{X}$ が有効推定量であることを示せ。

問題 5.15 正規母集団 $\mathrm{N}(\mu,\sigma^2)$ からの無作為標本 $X_1,\ldots,X_n$ とする。このとき、問題 5.13 より不偏分散 U^2 は母分散 σ^2 の不偏推定量であるが、σ^2 の Fisher 情報量を求め、これが Cramér-Rao の定理の (5.14) の等号を満たすかどうか調べよ。ただし、μ は既知とする。

定義 5.5 未知母数 θ の母集団からの標本 $X_1,\ldots,X_n$ に対し、その密度関数を $f(x_1,\cdots,x_n|\theta)$ または確率関数を $p(x_1,\cdots,x_n|\theta) = P(X_1 = x_1,\cdots,X_n = x_n)$ とする。このとき標本 $X_1,\ldots,X_n$ が与えられた条件下での θ の**尤度関数** (likelihood function) を次で定義する。

$$L(\theta) = f(x_1,\cdots,x_n|\theta) \quad (\text{密度関数があるとき}) \qquad = p(x_1,\cdots,x_n|\theta) \quad (\text{離散型のとき})$$

また、尤度関数 $L(\theta)$ を最大にする $\theta = \widehat{\theta}(x_1, \cdots, x_n)$ が存在するとき、すなわち、

$$L(\widehat{\theta}(x_1, \cdots, x_n)) = \max_{\theta} L(\theta)$$

のとき $\widehat{\theta}(X_1, \cdots, X_n)$ を θ の**最尤推定量** (maximum likelihood estimator) という。

例題 5.27 $X_1, X_2, \ldots, X_n$ が Bernoulli 試行の結果で、$P(X_i = 1) = p$, $P(X_i = 0) = 1 - p$ とする。p の最尤推定量を求めよ。

解: p の尤度関数は

$$L(p) = P(X_1 = x_1, \cdots, X_n = x_n) = p^{x_1 + \cdots + x_n}(1-p)^{n-(x_1+\cdots+x_n)}.$$

$\overline{x} = (x_1 + \cdots + x_n)/n$ と表し、$L(p)$ の対数をとって微分すると

$$\frac{\partial}{\partial p}\log L(p) = \frac{\partial}{\partial p}\Big\{n\overline{x}\log p + n(1-\overline{x})\log(1-p)\Big\} = \frac{n\overline{x}}{p} - \frac{n(1-\overline{x})}{1-p} = \frac{n(\overline{x}-p)}{p(1-p)}.$$

従って、$\dfrac{\partial}{\partial p}\log L(p) = 0$ を解いて $p = \overline{x}$. このとき、$L(p)$ は最大となるので、

$$\widehat{p} = \frac{X_1 + \cdots + X_n}{n}$$

が p の最尤推定量となる。 □

例題 5.28 $X_1, X_2, \ldots, X_n$ が正規母集団 $\mathrm{N}(\mu, \sigma^2)$ からの無作為標本のとき、母平均 μ と母分散 σ^2 の最尤推定量を求めよ。

解: 尤度関数は

$$L(\mu, \sigma^2) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x_i-\mu)^2}{2\sigma^2}} = \Big(\frac{1}{2\pi\sigma^2}\Big)^{n/2} \exp\Big\{-\sum_{i=1}^{n}\frac{(x_i-\mu)^2}{2\sigma^2}\Big\}.$$

対数をとると $\log L(\mu, \sigma^2) = \dfrac{n}{2}\log 2\pi\sigma^2 - \displaystyle\sum_{i=1}^{n}\frac{(x_i-\mu)^2}{2\sigma^2}$ より

$$\frac{\partial}{\partial \mu}\log L(\mu, \sigma^2) = \sum_{i=1}^{n}\frac{x_i - \mu}{\sigma^2}, \qquad \frac{\partial}{\partial \sigma^2}\log L(\mu, \sigma^2) = -\frac{n}{2}\frac{1}{\sigma^2} + \sum_{i=1}^{n}\frac{(x_i-\mu)^2}{2(\sigma^2)^2}.$$

従って、$\dfrac{\partial}{\partial \mu}\log L(\mu, \sigma^2) = 0$, $\dfrac{\partial}{\partial \sigma^2}\log L(\mu, \sigma^2) = 0$ を解いて $\mu = \overline{x} = \dfrac{x_1 + \cdots + x_n}{n}$, $\sigma^2 = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n}(x_i - \overline{x})^2$. このとき、$L(\mu, \sigma^2)$ は最大となるので、

$$\widehat{\mu} = \overline{X} = \frac{X_1 + \cdots + X_n}{n}, \qquad \widehat{\sigma^2} = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2$$

が μ, σ^2 の最尤推定量となる。 □

注意 5.8 例 5.23, 問題 5.13 より上記の $\widehat{\mu}, \widehat{\sigma^2}$ は一致推定量になっていることがわかる。一般に、ある分布のクラスにおいて最尤推定量が一致推定量となることが証明できる。更に、中心極限定理を用いることで、最尤推定量が漸近的に正規分布となることが示され、その共分散行列は Fisher 情報量の逆行列となる (*cf*. 稲垣宣生 著 数理統計学 裳華房, 2003, pp.261–)。このことは損保数理で用いられる。

問題 5.16 (1) $X_1, X_2, \ldots, X_n$ を i.i.d. で Poisson 分布 $\mathrm{Po}(\lambda)$ に従うとする。このとき、λ の最尤推定量を求めよ。また、それが有効推定量となることを示せ。
(2) $X_1, X_2, \ldots, X_n$ を i.i.d. でガンマ分布 $\Gamma(\alpha, 1/\beta)$ に従うとする。ここで、α は既知で β は未知パラメーターとする。このとき、β の最尤推定量を求めよ。また、それが有効推定量となるか調べよ。

例題 5.29 $X_1, X_2, \ldots, X_n$ が i.i.d. で、一様分布 $\mathrm{U}(0, \theta)$ に従うとき、未知母数 θ の最尤推定量を求めよ。

解: サンプルの実現値 $x_1, \ldots, x_n$ が与えられたとき、尤度関数 $L(\theta)$ は $0 \leq x_1 \leq \theta, \ \cdots, \ 0 \leq x_x \leq \theta$ でなければ $L(\theta) = 0$ となる。よって、$x_1 > 0, \ldots, x_n > 0$ であり

$$L(\theta) = \frac{1}{\theta^n} \quad (\max\{x_1, \cdots, x_n\} \leq \theta \text{のとき}) \qquad = 0 \quad (\text{その他})$$

となる。従って、$\theta = \max\{x_1, \cdots, x_n\}$ のとき、$L(\theta)$ は最大値をとる。これより

$$\widehat{\theta} = \max\{X_1, \ldots, X_n\}$$

が θ の最尤推定量となる。 □

6 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。

6.1 特性関数

定義 6.1 (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 $X = \mathrm{Re}\, Z$, 虚部 $Y = \mathrm{Im}\, Z$ 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、$E[|\mathrm{Re}\, Z|] < \infty$ かつ $E[|\mathrm{Im}\, Z|] < \infty$ のとき、Z の期待値を

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

と定める。

命題 6.1 複素数値確率変数 Z に対して、$|E[Z]| \leq E[|Z|]$ が成立する。

証明: $E[|Z|] < \infty$ のとき、$|\mathrm{Re}\, Z| \leq |Z|$, $|\mathrm{Im}\, Z| \leq |Z|$ より $E[Z]$ が定義されることに注意する。$\alpha = E[Z]$, $\tilde{Z} = \dfrac{Z}{|Z|} 1_{\{Z \neq 0\}}$ とする。このとき、

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

であるが、左辺は実数なので、(右辺の虚部)$= 0$ となる。ここで、$\mathrm{Re}\Big(\frac{\overline{\alpha}}{|\alpha|}\tilde{Z}\Big) \leq \Big|\frac{\overline{\alpha}}{|\alpha|}\tilde{Z}\Big| \leq 1$ であるから、(右辺の実部)$\leq E[|Z|]$ となり、主張を得る。 □

定義 6.2 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 6.2 (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 6.1 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 5.10) により $E[|e^{ih_nX} - 1|] \to E[|e^0 - 1|] = 0$ となり主張を得る。 □

命題 6.3 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 6.4 (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 6.1 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 6.2 (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。

命題 6.5 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 6.2 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

$$\lim_{n\to\infty}\frac{\phi_X^{(k)}(t+h_n) - \phi_X^{(k)}(t)}{h_n} = \lim_{n\to\infty} E\Big[(iX)^k\frac{e^{i(t+h_n)X} - e^{itX}}{h_n}\Big] = E[(iX)^{k+1}e^{itX}] \tag{6.1}$$

を示す。$|e^{i(t+h)x} - e^{itx}| = \Big|\displaystyle\int_t^{t+h} ixe^{ix\theta}\,d\theta\Big| \le \Big|\displaystyle\int_t^{t+h}|ixe^{ix\theta}|\,d\theta\Big| \le |hx|$ 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 5.10) と $\lim_{h\to 0}\dfrac{e^{i(t+h_n)X}-e^{itX}}{h_n}=iXe^{itX}$ となることから (6.1) 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 6.2(iv) と同様に示すことができる。 □

例 6.6 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。

証明: Cauchy の積分定理を用いる。まず、

$$\phi_X(t)=E[e^{itX}]=\int_{-\infty}^{\infty}e^{itx}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}\,dx=e^{-\frac{1}{2}t^2}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}e^{-\frac{1}{2}(x-it)^2}\,dx \tag{6.2}$$

に注意する。最後の等号は $itx-\frac{1}{2}x^2=-\frac{1}{2}(x-it)^2-\frac{1}{2}t^2$ による。右辺の積分を求めるため、$R>0$ とし次の 4 つの線分からなる閉曲線 C_R を考える。(図示せよ。)

$$C_{R,1}:-R\to R,\quad C_{R,2}:R\to R-it,\quad C_{R,3}:R-it\to -R-it,\quad C_{R,4}:-R-it\to -R.$$

ここで、$e^{-\frac{1}{2}z^2}$ は複素平面 $\mathbf{C}$ 上で正則な関数だから Cauchy の積分定理により $\displaystyle\int_{C_R}e^{-\frac{1}{2}z^2}\,dz=0$ となる。一方、

$$\int_{C_R}e^{-\frac{1}{2}z^2}\,dz=\sum_{n=1}^{4}\int_{C_{R,n}}e^{-\frac{1}{2}z^2}\,dz$$

であるが、線積分を用いて計算すると、$R\to\infty$ のとき、

$$\int_{C_{R,1}}e^{-\frac{1}{2}z^2}\,dz=\int_{-R}^{R}e^{-\frac{1}{2}x^2}\,dx\to\int_{-\infty}^{\infty}e^{-\frac{1}{2}x^2}\,dx=\sqrt{2\pi}.$$

$C_{R,2}$ は $z=R+iy$ と考え $dz=i\,dy$ に注意して

$$\begin{aligned}\left|\int_{C_{R,2}}e^{-\frac{1}{2}z^2}\,dz\right|&=\left|\int_0^{-t}e^{-\frac{1}{2}(R+iy)^2}i\,dy\right|\le\int_0^{|t|}\left|e^{-\frac{1}{2}(R^2-y^2)+iRy}i\right|\,dy\\&=\int_0^{|t|}e^{-\frac{1}{2}(R^2-y^2)}\,dy\le|t|e^{-\frac{1}{2}(R^2-t^2)}\to 0.\end{aligned}$$

同様に $C_{R,4}$ は $z=-R+iy$ と考えて

$$\left|\int_{C_{R,4}}e^{-\frac{1}{2}z^2}\,dz\right|=\left|\int_{-t}^{0}e^{-\frac{1}{2}(-R+iy)^2}i\,dy\right|\le|t|e^{-\frac{1}{2}(R^2-t^2)}\to 0.$$

最後に $C_{R,3}$ は $z=x-it$ と考え $dz=i\,dx$ に注意して

$$\int_{C_{R,3}}e^{-\frac{1}{2}z^2}\,dz=\int_R^{-R}e^{-\frac{1}{2}(x-it)^2}\,dx=-\int_{-R}^{R}e^{-\frac{1}{2}(x-it)^2}\,dx\to-\int_{-\infty}^{\infty}e^{-\frac{1}{2}(x-it)^2}\,dx.$$

よって、$\sqrt{2\pi}-\displaystyle\int_{-\infty}^{\infty}e^{-\frac{1}{2}(x-it)^2}\,dx=0$ となるので、(6.2) に代入して $\phi_X(t)=e^{-\frac{1}{2}t^2}\dfrac{1}{\sqrt{2\pi}}\sqrt{2\pi}=e^{-\frac{1}{2}t^2}$ を得る。 □

問題 6.3 以下に従って例 6.6 を証明せよ。(上記の別証明です。)
(1) $E[|X|]<\infty$ に注意して命題 6.5 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 6.7 X が正規分布 $N(m,\sigma^2)$ に従うときその特性関数は $\phi_X(t)=e^{imt-\frac{1}{2}\sigma^2t^2}$ 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 6.3 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 6.8 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。

証明: 留数定理を用いる。

$$\phi_X(t) = E[e^{itX}] = \int_{-\infty}^{\infty} e^{itx}\frac{1}{\pi}\frac{1}{1+x^2}\,dx \tag{6.3}$$

に注意する。

<u>1st step</u> $t > 0$ とする。$R > 1$ とし次の 2 つの曲線からなる閉曲線 C_R を考える。(図示せよ。)

$$C_{R,1}: \text{実軸上を} -R \to R, \qquad C_{R,2}: \text{半円 } |z| = R,\ \mathrm{Im}\,z \geq 0 \text{ 上を } R \to -R.$$

ここで、$g(z) = \dfrac{e^{itz}}{z+i}$ は $z \neq -i$ で正則だから Cauchy の積分表示により

$$\frac{1}{2\pi i}\int_{C_R}\frac{e^{itz}}{1+z^2}\,dz = \frac{1}{2\pi i}\int_{C_R}\frac{g(z)}{z-i}\,dz = g(i) = \frac{e^{-t}}{2i}. \tag{6.4}$$

一方、

$$\int_{C_R}\frac{e^{itz}}{1+z^2}\,dz = \sum_{n=1}^{2}\int_{C_{R,n}}\frac{e^{itz}}{1+z^2}\,dz \tag{6.5}$$

であるが、$R \to \infty$ のとき、

$$\int_{C_{R,1}}\frac{e^{itz}}{1+z^2}\,dz = \int_{-R}^{R}\frac{e^{itx}}{1+x^2}\,dx \to \int_{-\infty}^{\infty}\frac{e^{itx}}{1+x^2}\,dx$$

$C_{R,2}$ は $z = Re^{i\theta}$, $0 \leq \theta \leq \pi$, と考え $dz = Rie^{i\theta}\,d\theta$ に注意して

$$\begin{aligned}\Big|\int_{C_{R,2}}\frac{e^{itz}}{1+z^2}\,dz\Big| &= \Big|\int_0^{\pi}\frac{e^{itRe^{i\theta}}}{1+R^2e^{2i\theta}}Rie^{i\theta}\,d\theta\Big| \leq \int_0^{\pi}\frac{\left|e^{itR(\cos\theta+i\sin\theta)}\right|}{|R^2e^{2i\theta}+1|}R\,d\theta \\ &= \int_0^{\pi}\frac{e^{-tR\sin\theta}}{|R^2e^{2i\theta}+1|}R\,d\theta \leq \int_0^{\pi}\frac{e^{-tR\sin\theta}}{R^2-1}R\,d\theta \leq \pi\frac{R}{R^2-1} \to 0.\end{aligned}$$

ここで、2 行目の最初の不等号は $|R^2e^{2i\theta}+1| \geq |R^2e^{2i\theta}| - 1 = R^2 - 1$ を、二つ目の不等号は $0 \leq \theta \leq \pi$ のとき $\sin\theta \geq 0$ となるから、$t > 0$ より $e^{-tR\sin\theta} \leq 1$ となることを用いた。よって、(6.4), (6.5) より

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

であるから、両辺を $2i$ 倍して (6.3) に代入して $\phi_X(t) = \dfrac{1}{\pi}\displaystyle\int_{-\infty}^{\infty}\frac{e^{itx}}{1+x^2}\,dx = e^{-t}$ を得る。

<u>2nd step</u> $t = 0$ のとき $\phi_X(0) = 1$ は命題 6.2(iii) による。

$t < 0$ のとき、(6.3) で $y = -x$ と変数変換すると、

$$\phi_X(t) = \int_{-\infty}^{\infty}e^{it(-y)}\frac{1}{\pi}\frac{1}{1+(-y)^2}\,dy = \int_{-\infty}^{\infty}e^{i(-t)y}\frac{1}{\pi}\frac{1}{1+y^2}\,dy = e^{-(-t)} = e^{-|t|}.$$

3 つ目の等号は $-t > 0$ に注意して 1st step の結果を用いた。 $\square$

注意 6.1 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 6.5)。

定義 6.3 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 6.9 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: $\phi_{A\boldsymbol{X}+\boldsymbol{b}}(\boldsymbol{t}) = E[e^{i\boldsymbol{t}'A\boldsymbol{X}}e^{i\boldsymbol{t}'\boldsymbol{b}}] = e^{i\boldsymbol{t}'\boldsymbol{b}}E[e^{i(A'\boldsymbol{t})'\boldsymbol{X}}] = e^{i\boldsymbol{t}'\boldsymbol{b}}\phi_{\boldsymbol{X}}(A'\boldsymbol{t})$. $\square$

例 6.10 $\boldsymbol{X} = (X_1, \ldots, X_d)'$ を d 次元正規分布 $N(\boldsymbol{m}, \boldsymbol{\Sigma})$ に従うとする。$\boldsymbol{m} = (m_1, \ldots, m_d)' \in \mathbf{R}^d$, $\boldsymbol{\Sigma} = (\sigma_{ij})$ 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: 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、$\boldsymbol{Y} = (Y_1, \ldots, Y_d)' = P'(\boldsymbol{X} - \boldsymbol{m})$ とすると、$Y_1, \ldots, Y_d$ 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 6.9 を適用して

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

命題 6.5 と同様に次が成立する。証明は同様なので省略する。

命題 6.11 d 次元確率変数 $\boldsymbol{X} = (X_1, \ldots, X_d)'$ 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 $\dfrac{\partial^2}{\partial t_k \partial t_l}\phi_{\boldsymbol{X}}(\boldsymbol{t}) = i^2 E[X_k X_l e^{i\boldsymbol{t}'\boldsymbol{X}}]$, $1 \leq k, l \leq d$ となる。

6.2 分布と Dynkin 族定理

σ-集合族に関連して、Dynkin 族の概念を導入する。

定義 6.4 (1) 集合 S の部分集合族 $\mathcal{P}$ が π 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 S 上の 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 S 上の 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 6.4) 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 6.4 (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 σ-集合族であれば、S 上の Dynkin 族となることを示せ。

定理 6.12 (Dynkin 族定理) $\mathcal{P}$ が π 族のとき $\mathcal{L}(\mathcal{P}) = \sigma(\mathcal{P})$ となる[*5]。

証明: σ-集合族は Dynkin 族となる (*cf*. 問題 6.4) 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 σ-集合族となることを示せばよい。
<u>1st step</u>　$A \in \mathcal{P}$ 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。
π 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。
<u>2nd step</u>　$A \in \mathcal{L}(\mathcal{P})$ 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 π 族になっていることがわかる。
<u>3rd step</u>　$\mathcal{L}(\mathcal{P})$ が σ-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 π 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 6.5 定理 6.12 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 6.5 確率変数 X に対して、それが定める $\mathbf{R}$ 上の確率測度を μ_X と書く:

$$\mu_X(A) = P(X \in A), \qquad A \in \mathcal{B}(\mathbf{R}). \tag{6.6}$$

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

注意 6.2 μ_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 μ に対して、適当に確率空間 $(\Omega, \mathcal{F}, P)$ を定めれば、μ 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.)。

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 6.13 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 π 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 6.6)。$x \in (-\infty, \infty)$ 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 6.12 により

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

$\sigma(\mathcal{J}) = \mathcal{L}(\mathcal{J}) \subset \mathfrak{A}$. ところで、$\mathbf{R}$ 上の Borel 集合は $(a, b]$, $a < b$, を含む最小の σ-集合族であった。よって、$\mathcal{J}$ を含む最小の σ-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、$\sigma(\mathcal{J})$ と Borel 集合族 $\mathcal{B}(\mathbf{R})$ は一致し、$\mathcal{B}(\mathbf{R}) \subset \mathfrak{A}$ となり主張を得る。 $\square$

問題 6.6 定理 6.13 の証明で定めた $\mathfrak{A}$ が $\mathbf{R}$ 上の Dynkin 族となることを示せ。

命題 6.14 可測関数 $f(x)$ が $f \geq 0$ もしくは $\displaystyle\int_{\mathbf{R}} |f(x)| \mu_X(dx) < \infty$ を満たせば、$E[f(X)] = \displaystyle\int_{\mathbf{R}} f(x) \mu_X(dx)$.

略証: (前期の授業で説明した §4.1 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 5.9) 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。 $\square$

定理 6.15 $(\mathbf{R}, \mathcal{B}(\mathbf{R}))$ 上の確率測度 μ, ν について、$\forall f \in C_b(\mathbf{R})$ に対して $\displaystyle\int_{\mathbf{R}} f(x) \mu(dx) = \int_{\mathbf{R}} f(x) \nu(dx)$ 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、$\displaystyle\int_{\mathbf{R}} f_n(x) \mu(dx) = \int_{\mathbf{R}} f_n(x) \nu(dx)$ 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 5.10) を用いると、$\displaystyle\int_{\mathbf{R}} 1_{(-\infty, a]}(x) \mu(dx) = \int_{\mathbf{R}} 1_{(-\infty, a]}(x) \nu(dx)$, 即ち、$\mu((-\infty, a]) = \nu((-\infty, a])$ となる。よって、定理 6.13 の証明と全く同様にして $\mu = \nu$ を得る。 $\square$

問題 6.7 $(X_1, \cdots, X_n)$ が絶対連続型確率変数とし、$f(x_1, \ldots, x_n)$ をその密度関数とすると、

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

と表せる。このとき、

$$P((X_1, \ldots, X_n) \in B) = \int\int \cdots \int_B f(t_1, t_2, \ldots, t_n)\, dt_1 dt_2 \cdots dt_n, \qquad B \in \mathcal{B}(\mathbf{R}^n)$$

となることを、Dynkin 族定理 (定理 6.12, 定理 6.13 も参照のこと) を用いて示せ。ただし、$\mathbf{R}^n$ 上の Borel 集合族 $\mathcal{B}(\mathbf{R}^n)$ は $(a_1, b_1] \times \cdots \times (a_n, b_n]$, $a_k < b_k$, $k = 1, \ldots, n$ を含む最小の σ-集合族であった。また、難しければ $n = 2$ の場合に示せばよい。

6.3 特性関数と分布

定理 6.16 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。

補題 6.17 $\displaystyle\int_0^\infty \frac{\sin x}{x}\,dx = \frac{\pi}{2}$. すなわち、$\displaystyle f(M) = \int_0^M \frac{\sin t}{t}\,dt$ とすると、$\displaystyle\lim_{M\to\infty} f(M) = \frac{\pi}{2}$.

証明: $t > 0$ のとき $\displaystyle\int_0^\infty e^{-ut}\,du = \frac{1}{t}$ より $\displaystyle f(M) = \int_0^M \frac{\sin t}{t}\,dt = \int_0^M\!\!\int_0^\infty e^{-ut}\sin t\,du\,dt$ となる。同様に、$\displaystyle\int_0^M\!\!\int_0^\infty \left|e^{-ut}\sin t\right|\,du\,dt = \int_0^M \frac{|\sin t|}{t}\,dt < \infty$ ($cf.$ 問題 6.8) より、Fubini の定理 (定理 5.11) から、$\displaystyle\int_0^M\!\!\int_0^\infty e^{-ut}\sin t\,du\,dt = \int_0^\infty\int_0^M e^{-ut}\sin t\,dt\,du$. 部分積分することにより

$$\int_0^M e^{-ut}\sin t\,dt = \frac{\varphi_M(u)}{1+u^2},\quad \text{ただし}\quad \varphi_M(u) = 1 - e^{-Mu}\cos M - ue^{-Mu}\sin M \tag{6.7}$$

を得る。さらに、$|\varphi_M(u)| \le 3$ ($u \ge 0$) であるから Lebesgue の収束定理 (定理 5.10) 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 6.8 (1) 補題 6.17 で $M > 0$ に対して $|f(M)| < \infty$ と $|\varphi_M(u)| \le 3$ ($u \ge 0$) を示せ。ヒント: $|\sin u| \le |u|$ ($|u| \le \frac{\pi}{2}$) を用いよ。

(2) 補題 6.17 を用いて $\displaystyle\sup_{M>0}|f(M)| < \infty$ 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 6.17 は解析学で学んだ Cauchy の積分定理を用いて証明できるので、述べておく。

補題 6.17 の別証明: $0 < \varepsilon < R$ とし $\operatorname{Im} z \ge 0$ に含まれる 4 つの曲線からなる閉曲線 $C_{\varepsilon,R}$ を考える。(図示せよ。)

$$\begin{aligned} C_{\varepsilon,R,1} &: \text{実軸上を}\varepsilon \to R, & C_{\varepsilon,R,2} &: \text{半円 } |z| = R,\ \operatorname{Im} z \ge 0 \text{ 上を } R \to -R,\\ C_{\varepsilon,R,3} &: \text{実軸上を} -R \to -\varepsilon, & C_{\varepsilon,R,4} &: \text{半円 } |z| = \varepsilon,\ \operatorname{Im} z \ge 0 \text{ 上を } -\varepsilon \to \varepsilon \end{aligned}$$

Cauchy の積分定理より

$$0 = \int_{C_{\varepsilon,R}} \frac{e^{iz}}{z}\,dz = \sum_{n=1}^4 \int_{C_{\varepsilon,R,n}} \frac{e^{iz}}{z}\,dz. \tag{6.8}$$

ここで、

$$\int_{C_{\varepsilon,R,1}} \frac{e^{iz}}{z}\,dz + \int_{C_{\varepsilon,R,3}} \frac{e^{iz}}{z}\,dz = \int_\varepsilon^R \frac{e^{ix}}{x}\,dx + \int_{-R}^{-\varepsilon} \frac{e^{ix}}{x}\,dx = \int_\varepsilon^R \frac{e^{ix}-e^{-ix}}{x}\,dx = 2i\int_\varepsilon^R \frac{\sin x}{x}\,dx$$

であるが、補題 6.18(1) により $\varepsilon \to +0$, $R \to \infty$ のとき右辺は $\displaystyle 2i\int_0^\infty \frac{\sin x}{x}\,dx$ に収束する。$C_{\varepsilon,R,2}$ は $z = Re^{i\theta}$, $0 \le \theta \le \pi$, と考え

$$\left|\int_{C_{\varepsilon,R,2}} \frac{e^{iz}}{z}\,dz\right| = \left|\int_0^\pi \frac{e^{iRe^{i\theta}}}{Re^{i\theta}} Rie^{i\theta}\,d\theta\right| \le \int_0^\pi \left|e^{iR(\cos\theta + i\sin\theta)}\right|\,d\theta = \int_0^\pi e^{-R\sin\theta}\,d\theta$$

であるが、$R \to \infty$ のとき、$e^{-R\sin\theta} \to 0$, $|e^{-R\sin\theta}| \le 1$ $(0 < \theta < \pi)$ となるから、Lebesgue の収束定理により右辺は 0 に収束する。$C_{\varepsilon,R,4}$ も同様に

$$\int_{C_{\varepsilon,R,4}} \frac{e^{iz}}{z}\,dz = \int_\pi^0 \frac{e^{i\varepsilon e^{i\theta}}}{\varepsilon e^{i\theta}} \varepsilon ie^{i\theta}\,d\theta = -i\int_0^\pi e^{i\varepsilon e^{i\theta}}\,d\theta$$

であるが、$\varepsilon \to +0$ のとき、$e^{i\varepsilon e^{i\theta}} \to 1$, $|e^{i\varepsilon e^{i\theta}}| = e^{-\varepsilon\sin\theta} \le 1$ $(0 < \theta < \pi)$ となるから、Lebesgue の収束定理により右辺は $\displaystyle -i\int_0^\pi d\theta = -i\pi$ に収束する。よって、(6.8) と組み合わせて

$$0 = 2i\int_0^\infty \frac{\sin x}{x}\,dx - i\pi$$

であるから、与式を得る。 $\square$

補題 6.18 $f_T(\alpha) = \int_0^T \frac{\sin \alpha t}{t}\, dt$, $T > 0$, $\alpha \in \mathbf{R}$ について、(1) $\sup_{T,\alpha} |f_T(\alpha)| < \infty$,

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

証明: (1) $u = \alpha t$ とおくと、$f_T(\alpha) = \int_0^{\alpha T} \frac{\sin u}{u}\, du$ となる。よって、問題 6.8 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 6.17 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 6.19 (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。

$$\frac{1}{2}\{F(b) + F(b-0) - F(a) - F(a-0)\} = \frac{1}{2\pi}\lim_{T\to\infty}\int_{-T}^{T} \frac{e^{-itb} - e^{-ita}}{-it}\phi(t)\, dt, \quad a, b \in \mathbf{R}, a < b. \tag{6.9}$$

特に、a, b がともに F の連続点であれば、(6.9) 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 μ_X とすると、命題 6.14 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、

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

となるから、$\left|\frac{e^{-itb}-e^{-ita}}{-it}\right| \le |b-a|$ より、$\int_{-T}^{T}\int_{\mathbf{R}}\left|\frac{e^{-itb} - e^{-ita}}{-it} e^{itx}\right| \mu_X(dx)dt \le |b-a| \cdot 2T < \infty$ に注意して Fubini の定理 (定理 5.11) 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 6.18 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 6.18(1), (2) に注意して Lebesgue の収束定理 (定理 5.10) を用いると、$T \to \infty$ のとき、

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

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

$$(6.9)\ の右辺 = \frac{1}{2\pi}\pi\left[1 - F(a) - F(a-0) - (1 - F(b)) + F(b-0)\right] = F(b) - F(a) \tag{6.10}$$

を得る。 $\square$

定理 6.16 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 6.19 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 6.13 より $\mu_X = \mu_Y$ である。 $\square$

例 6.20 $X_1, \ldots, X_n$ 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 6.4(2) を用いた。よって、再び例 6.4(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 6.16 により主張を得る。

問題 6.9 (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 6.2 (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 6.10 (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 $P(-10.2 \le Y \le 9.9)$ 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、$P(1.5 \le \overline{X} \le 2.5) \ge 0.9$ 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 6.11 $f(x) = \dfrac{1}{x\sqrt{2\pi}} e^{-\frac{1}{2}(\log x)^2} 1_{(0,\infty)}(x)$ とし、$f_a(x) = (1 + a\sin(2\pi \log x)) f(x)$, $-1 \le a \le 1$, 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 $-1 \le a \le 1$ 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 6.12 $\displaystyle\int_0^1\int_0^1 \left|\frac{x^2 - y^2}{(x^2+y^2)^2}\right| dxdy = \infty$ 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 5.11) 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。

定理 6.21 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。

$$f_X(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-itx}\phi(t)\, dt. \tag{6.11}$$

証明: 定理 6.19 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、(6.9) の右辺の被積分関数は $(-\infty,\infty)$ で積分可能なので、(6.10) 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 5.11) 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 6.2(iv) と同様に Lebesgue の収束定理 (定理 5.10) 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 (6.11) を得る。□

例 6.8 の証明: 問題 6.2(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 6.21 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 6.22 $X_1, X_2, \ldots$ 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 6.8 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 6.13 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 6.1(1) を用いて、例 6.8 の証明と同様に定理 6.21 を用いよ。

6.4 法則収束と弱収束

定義 6.6 (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 6.23 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 $|f(x)| \le M$ $(\forall x \in \mathbf{R})$ とできるので、$|f(X_n(\omega))| \le M$ $(\omega \in \Omega)$ とできる。したがって、Lebesgue の収束定理 (定理 5.10) 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、$|f(x)| \le M$ $(x \in \mathbf{R})$ であれば、$|a_n| \le M$ 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 5.4 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。　$\square$

例 6.24 定理 6.23 の逆は、必ずしも成立しない ($cf.$ 問題 6.14 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、$n \ge 2$ のとき、

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

となり、$\{X_n\}$ は X_1 に確率収束しないことがわかる。

問題 6.14 $X_n \to a$ 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 6.7 μ_n, $n = 1, 2, \ldots$, と μ 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。μ_n が μ 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 μ_X に弱収束することと同値である。このとき、極限 μ は一意的である。実際、もう一つの極限を ν 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 6.15 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 6.25 確率変数 $X_1, X_2, \dots$ 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: $\underline{(1) \Rightarrow (2)}$ 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、

$$\limsup_{n\to\infty} F_{X_n}(x) \le \lim_{\delta\to+0} F_X(x+\delta) = F_X(x) \tag{6.12}$$

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、(6.12) 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

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

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 ε, δ 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。　　□

注意 6.3 定理 6.25 (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 6.26 $\alpha > 0$ 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。 $\square$

問題 6.15 $X_1, X_2, \cdots$ を i.i.d. とするとき、次を示せ。
(1) $\alpha > 0$ とする。X_1 がベータ分布 $\mathrm{B{\scriptstyle ETA}}(1, \alpha)$ 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 6.4 (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 6.16 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。

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 6.27 (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 6.28 (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

$$\lim_{k\to\infty} F_{n_k}(x) = F(x) \tag{6.13}$$

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

$$F(x) = \inf\{\tilde{F}(y); y \in \mathbf{Q} \cap (x, \infty)\} \tag{6.14}$$

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、(6.13) 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、(6.14) 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、(6.13) は成立する。　□

定理 6.27 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、

$$P(X_{\alpha_n} \in [-n, n]) < 1 - \varepsilon \tag{6.15}$$

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、(6.15) により

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

となり、$\lim_{m\to\infty}\{F_X(y_m) - F_X(x_m)\} \leq 1 - \varepsilon$. 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。

(2) ⇒ (1)　F_{X_α} 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 6.28) 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

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

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

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

となるので、$0 \leq F \leq 1$ 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。　□

6.5　特性関数と法則収束

定理 6.29 $\{X_n\}$ 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。
1st step　定理 6.27 を用いて $\{X_n\}$ が点列 compact であることを示す。X_n の分布を μ_n とかく。まず、

$$|a| \geq 1 \implies |\sin a| \leq |a| \cdot \sin 1 \implies \frac{\sin a}{a} \leq \sin 1 \implies 1 - \frac{\sin a}{a} \geq 1 - \sin 1$$

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

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

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 5.11) を用いた。次に最後の式を $I_n(M)$ とし、$s = t/M$ と変換し、

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

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、$\lim_{n\to\infty} \phi_n(s) = \phi(s)$ $(\forall s \in \mathbf{R})$ 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 5.10) 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 6.27 により $\{X_n\}$ が点列 compact であることがわかった。

2nd step $\{X_n\}$ 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 6.16 により、$\mu_X = \mu_Y$ となる。これは、μ_{X_n} 自身が μ_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。 $\square$

問題 6.17 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 の特性関数を求め、X_n が $(0,1)$ 上の一様分布に法則収束することを示せ。また、$0 < \forall x < 1$ 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。

6.6 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 6.30 (中心極限定理) $X_1, X_2, \ldots$ 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。

$$\lim_{n\to\infty} P\left(a \leq U_n \leq b\right) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{y^2}{2}}\, dy, \qquad -\infty < a < b < \infty. \tag{6.16}$$

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

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

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 6.5 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 5.10) 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 6.29(2) により、$\{U_n\}$ は Y に法則収束する。(6.16) は Y の分布関数 $F_Y(x)$ が連続だから、定理 6.25 により成立する。 □

注意 6.5 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 5.6, 注意 5.3, 注意 5.4)。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 6.18 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 6.19 $X_1,X_2,\dots$ 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 6.20 $X_1,X_2,\dots$ 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 6.1(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、$P(|Y_n|<0.02)\ge 0.95$ 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 6.31 (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、

$$\lim_{n\to\infty}P\left(a\le\frac{S_n-np}{\sqrt{np(1-p)}}\le b\right)=\frac{1}{\sqrt{2\pi}}\int_a^b e^{-\frac{y^2}{2}}\,dy,\qquad -\infty<a<b<\infty. \tag{6.17}$$

証明: $E[X_k]=p,\ V(X_k)=p(1-p)$ であるから、定理 6.30 により従う。 □

注意 6.6 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 (6.17) 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 6.18, 6.22)。

問題 6.21 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 6.22 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、$\lim\limits_{n\to\infty} \sqrt{n}P(X_{3n} = n)$.

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

6.7 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 6.32 d 次元確率変数 $\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots$ 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 $E[|X_{k,j}|^2] < \infty$, $j = 1, \ldots, d$ と仮定する。このとき、

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

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 6.30 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 6.10 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 6.7 例 6.10 では $\boldsymbol{\Sigma}$ は正定値対称行列であったが、定理 6.32 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 $Z_1, \ldots, Z_d$ 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 6.9 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 6.16 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 6.23 次の Σ に対して、それぞれ Σ 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 6.33 確率変数 $Y_1, Y_2, \ldots$ 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 χ^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、

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

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 6.32 より $\{\boldsymbol{U}_n\}$ は $N(\boldsymbol{0}, \boldsymbol{\Sigma})$ に法則収束する。ただし、$\boldsymbol{\Sigma} = \left(\delta_{ij} - \sqrt{p_i p_j}\right)_{1 \leq i,j \leq d}$, また $\delta_{jj} = 1$, $i \neq j$ なら $\delta_{ij} = 0$ である。ここで、注意 6.7 のように $\boldsymbol{Z}$, A を定めると、

$$\lim_{n \to \infty} E[\tilde{f}(\boldsymbol{U}_n)] = E[\tilde{f}(A\boldsymbol{Z})], \qquad \tilde{f} \in C_b(\mathbf{R}^d), \tag{6.18}$$

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、(6.18) 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 χ^2 分布に従うことを示す。
$\boldsymbol{\Sigma}_2 = I - \boldsymbol{\Sigma} = \left(\sqrt{p_i p_j}\right)_{1 \leq i,j \leq d}$ とする (I は単位行列)。このとき、

$$\begin{aligned} |\boldsymbol{Z}|^2 &= \boldsymbol{Z}'\boldsymbol{Z} = \boldsymbol{Z}'(\boldsymbol{\Sigma} + \boldsymbol{\Sigma}_2)\boldsymbol{Z} = \boldsymbol{Z}'\boldsymbol{\Sigma}\boldsymbol{Z} + \boldsymbol{Z}'\boldsymbol{\Sigma}_2\boldsymbol{Z} \\ &= (A\boldsymbol{Z})'A\boldsymbol{Z} + \sum_{i,j=1}^d \sqrt{p_i p_j} Z_i Z_j = |A\boldsymbol{Z}|^2 + (\boldsymbol{p}_1'\boldsymbol{Z})^2. \end{aligned} \tag{6.19}$$

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、(6.19) 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 χ^2 分布に従うことがわかった。 □

注意 6.8 定理 6.33 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。

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

このとき、階級 A_j の出現回数を n_j とし (定理 6.33 では $\sum_{k=1}^n X_{k,j}$ に対応する)、統計量

$$T_n = \sum_{j=1}^d \frac{(n_j - np_j^0)^2}{np_j^0} \tag{6.20}$$

を考える。このとき、もし H_0 が正しければ T_n は自由度 $d-1$ の χ^2 分布 (以下、χ^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 χ^2_{d-1} 分布する確率変数を表す。一方、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、$\lim_{n\to\infty} P(T_n \leq a) = 0,\ \forall a > 0,$ となる。

以上より、もし有意水準 α で検定するのであれば、(6.20) に実現値を代入した値を t とするとき、

もし $t < \chi^2_{d-1}(\alpha)$ であれば H_0 を採択し、もし $t \geq \chi^2_{d-1}(\alpha)$ であれば H_0 を棄却する。

ここで、$\chi^2_{d-1}(\alpha)$ は χ^2_{d-1} 分布の上側 α 点、即ち、$P(Y \geq k) = \alpha$ 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 χ^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 $np_j^0 \geq 5$ となるように分けなおす。

例 6.34 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。

$\mathrm{H}_0 : (p_1, p_2, p_3, p_4, p_5, p_6) = \left(\frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}\right), \quad \mathrm{H}_1 : (p_1, p_2, p_3, p_4, p_5, p_6) \neq \left(\frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}\right).$

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

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

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、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 6.24 χ^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。

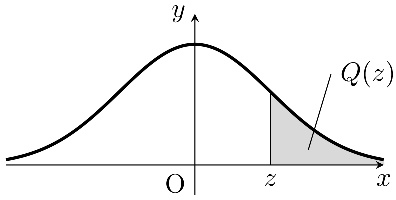
X	1	2	3	4	5	≥ 6	計
度数	132	66	36	10	11	1	256

ただし X は表が出現するまでの試行回数であり ≥ 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 χ^2 分布の上側 α 点: $\chi^2_n(\alpha)$

n \ α	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(α,n)」として作成した。

定理 6.21 の応用として computer tomography (計算機断層撮影) について紹介する。そのため定理 6.21 を2次元に拡張し通常の Fourier 変換の式で述べる。

定理 6.21′ $f(x,y)$ を "よい" 関数とし、

$$\phi(s,t) = \iint_{\mathbf{R}^2} e^{-i(sx+ty)} f(x,y)\,dxdy \tag{6.21}$$

とする。このとき、$\displaystyle\iint_{\mathbf{R}^2} |\phi(s,t)|\,dsdt < \infty$ を満たせば、

$$f(x,y) = \frac{1}{(2\pi)^2} \iint_{\mathbf{R}^2} e^{i(xs+yt)} \phi(s,t)\,dsdt. \tag{6.22}$$

$f(x,y)$ である物体 (例えば人体) の X 線を透過しない度合いの分布を表すとする。これを y 軸とのなす角が θ の遠方から撮影すると、x 軸と角 θ をなすスクリーン上 (これを u 軸とする) に透過できなかった割合を表す影ができる。これを $(Rf)(u,\theta)$ で表すと、u 軸と垂直な座標軸を v 軸とすると、

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} \tag{6.23}$$

の関係にある。よって、

$$(Rf)(u,\theta) = \int_{-\infty}^{\infty} f(u\cos\theta - v\sin\theta, u\sin\theta + v\cos\theta)\,dv, \qquad -\infty < u < \infty, \quad -\frac{\pi}{2} < \theta \leq \frac{\pi}{2}, \tag{6.24}$$

となる。目標は (6.24) のデータから $f(x,y)$ を復元することである。

まず、$(Rf)(u,\theta)$ の Fourier 変換を行う。すなわち、e^{-iru}, $r \in \mathbf{R}$, を掛けて u について積分すると、

$$\begin{aligned}
\psi(r,\theta) &= \int_{-\infty}^{\infty} e^{-iru} (Rf)(u,\theta)\,du \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-iru} f(u\cos\theta - v\sin\theta, u\sin\theta + v\cos\theta)\,dvdu \\
&= \iint_{\mathbf{R}^2} e^{-ir(x\cos\theta + y\sin\theta)} f(x,y)\,dxdy
\end{aligned}$$

となる。最後の等号は (6.23) を用いて (u,v) を (x,y) に置換した。この 3 行目の積分は (6.21) の $\phi(s,t)$ で $s = x\cos\theta, t = y\sin\theta$ としたものである。このため (6.22) の右辺で $s = r\cos\theta, t = r\sin\theta$, $r \in \mathbf{R}$, $-\pi/2 < \theta \leq \pi/2$ と置換する (極座標と異なり $r \in \mathbf{R}$ に注意) と、(6.22) より

$$f(x,y) = \frac{1}{(2\pi)^2} \int_{-\pi/2}^{\pi/2} \int_{-\infty}^{\infty} e^{i(xr\cos\theta + yr\sin\theta)} \phi(x\cos\theta, y\sin\theta) |r|\,drd\theta.$$

以上より

$$\begin{aligned}
f(x,y) &= \frac{1}{(2\pi)^2} \int_{-\pi/2}^{\pi/2} \int_{-\infty}^{\infty} e^{i(xr\cos\theta + yr\sin\theta)} \psi(r,\theta) |r|\,drd\theta \\
&= \frac{1}{(2\pi)^2} \int_{-\pi/2}^{\pi/2} \int_{-\infty}^{\infty} e^{i(xr\cos\theta + yr\sin\theta)} \left\{ \int_{-\infty}^{\infty} e^{-iru} (Rf)(u,\theta)\,du \right\} |r|\,drd\theta.
\end{aligned}$$

を得る。 □

現実問題としてこの積分の計算を行うのは大変困難で、計算機と Fourier 変換を求める技術 (高速 Fourier 変換) の発達が必要であった。

参考書 杉山 健一 著 フーリエ解析講義 理論と応用 講談社サイエンティフィック