確率統計学 I

杉浦　誠

平成 22 年 1 月 27 日

目 次

これは

- 浅野, 江島, 李 共著 基本統計学 森北出版

を教科書として作った講義ノートです。多くの部分で、

- 舟木 直久 著 確率論 朝倉書店

を参考にしています。(むしろ 3 節, 4 節では後者のほうが教科書に近い存在です。)

3　大数の法則

3.1　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 3.1　*(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。

(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) **(**r **次平均収束)**　$r \geq 1$ として、X_n が X に r 次平均収束 *(convergence in the mean of order* r*)* 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。

注意 3.2 確率変数がなす空間上に確率収束, r 次平均収束が定める位相は、それぞれ距離付け可能である。*(*前者は演習問題 **4***(1)* を参照せよ。後者は定理 *2.13(ii)* と注意 *2.5(3) (*教科書では *§2.13)* を参照せよ。*)* 概収束は距離付けできない。*(*演習問題 **4***(2)* を参照せよ。*)*

定理 3.3 *(1)*　X_n が X に概収束すれば、確率収束する。

(2)　X_n が X に r 次平均収束すれば、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

$$\Big\{\lim_{n\to\infty} X_n = X\Big\} = \bigcap_{j=1}^{\infty}\bigcup_{n=1}^{\infty}\bigcap_{m=n}^{\infty} A_{m,j} \tag{3.1}$$

と表すことができる。ただし、

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

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、$\Big\{\lim\limits_{n\to\infty} X_n = X\Big\} \in \mathcal{B}$ 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、(3.1) により $\bigcup\limits_{n=1}^{\infty}\bigcap\limits_{m=n}^{\infty} A_{m,j} \supset \bigcup\limits_{n=1}^{\infty}\bigcap\limits_{m=n}^{\infty} A_{m,j+1} \supset \cdots \supset \Big\{\lim\limits_{n\to\infty} X_n = X\Big\}$ となるので、

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

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

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

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

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

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 $1/j\le\varepsilon$ とすれば

$$\Big\{|X_n-X|<\frac{1}{j}\Big\}\subset\{|X_n-X|<\varepsilon\}$$

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

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

証明: まず、次に注意する。

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

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

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

定理 3.3(2) の証明: 仮定より、$E[|X_n-X|^r]\to 0$ となることに注意する。よって、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 3.5 定理 *3.3(1), (2)* の逆は、必ずしも成立しない。また、概収束と r 次平均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* 測度としてそれを例示する。

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

  $$X_{1,1},\ X_{2,1},\ X_{2,2},\ X_{3,1},\ X_{3,2},\ X_{3,3},\ X_{4,1},\ \dots$$

  のように並べた列を考える。この確率変数は $X\equiv 0$ に r 次平均収束の意味で収束するが、概収束しない。*(*この証明は演習問題 **3***(1)* とする。*)*
- 概収束する *(*従って確率収束する*)* が、r 次平均収束しない例
  $X_n(\omega)=n1_{(0,\frac{1}{n})}(\omega)$, $\omega\in[0,1]$ を考えると、これは $X\equiv 0$ に概収束するが、r 次平均収束しない。*(*この証明も演習問題 **3***(2)* とする。*)*

定理 3.6 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、r 次平均収束すれば *(*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。

証明: 各 $k \in \boldsymbol{N}$ 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

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

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

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

であるから、Borel-Cantelli の定理 (教科書 p.21) により、$P\Big(\bigcap_{l=1}^{\infty} \bigcup_{k=l}^{\infty} C_k\Big) = 0.$ ここで、

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

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

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。 □

3.2 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 $S_n/n = \sum\limits_{i=1}^{n} X_i/n$ の収束について議論する。

定義 3.7 ある数列 $\{a_n\}$ に対し、
(1) $S_n/n - a_n$ 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) $S_n/n - a_n$ 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 3.8 $\{X_n\}$ が組ごとに独立、つまりどの組 i, j *($i \neq j$)* をとっても X_i と X_j は独立で、

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

ならば、大数の弱法則を満たす。ただし、$V(X) = E[(X - E[X])^2]$ は X の分散を表す。

証明: $m_n = E[X_n]$ とし、$a_n = \frac{1}{n}\sum_{j=1}^{n} m_j$, $Y_n = S_n/n - a_n = \frac{1}{n}\sum_{j=1}^{n}(X_j - m_j)$ とする。このとき、$\varepsilon > 0$ は任意として、Chebyshev の不等式により

$$\begin{aligned} P(|Y_n| \geq \varepsilon) &\leq \frac{1}{\varepsilon^2} E[{Y_n}^2] = \frac{1}{\varepsilon^2 n^2} \sum_{i,j=1}^{n} E[(X_i - m_i)(X_j - m_j)] \\ &= \frac{1}{\varepsilon^2 n^2} \sum_{j=1}^{n} E[(X_j - m_j)^2] \leq \frac{1}{\varepsilon^2 n} \sup_j V(X_j) \ \to 0 \quad (n \to \infty) \end{aligned}$$

となり、弱法則が従う。ここで、2 行目第 1 の等号は $i \neq j$ のとき X_i と X_j は独立であるから

$$E[(X_i - m_i)(X_j - m_j)] = E[X_i - m_i]E[X_j - m_j] = (E[X_i] - m_i)(E[X_j] - m_j) = 0$$

となることを用いた。 □

3.3 大数の強法則

定理 3.9 (Kolmogorov の不等式) $\{X_n\}$ 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

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

とおく。S_k を確率過程のように考え $|S_k|$ がいつはじめて a 以上になるか、そのような 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

$$A_k^*=\{\,\omega\in\Omega\,;\,j=1,2,\ldots,k-1\text{ に対しては }|S_j|<a\text{ で、かつ }|S_k|\ge a\}\tag{3.2}$$

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、(3.2) 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 は $X_{k+1},X_{k+2},\cdots,X_n$ に依存しているので、$\{X_n\}$ は独立だから、$S_k\cdot 1_{A_k^*}$ と S_n-S_k は独立な確率変数である。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 3.10 (Kolmogorov の第 1 定理) $\{X_n\}$ を独立な確率変数列で、

$$\sum_{n=1}^{\infty}\frac{1}{n^2}V(X_n)<\infty\tag{3.3}$$

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: $\forall n \in \boldsymbol{N}$ 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、

$$P(A(\varepsilon)) = 1 \tag{3.4}$$

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、(3.4) 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 $j \in \boldsymbol{N}$ 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> (3.4) 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、$\forall l \in \boldsymbol{N}$ に対して

$$A(\varepsilon)^{\mathrm{c}} = \bigcap_{N=1}^{\infty}\bigcup_{n=N}^{\infty}\{|Y_n| \geq \varepsilon\} \subset \bigcup_{m=l}^{\infty} B_m(\varepsilon) \tag{3.5}$$

だから、(3.4), すなわち $P(A(\varepsilon)^{\mathrm{c}}) = 0$ を示すためには

$$\sum_{m=1}^{\infty} P(B_m(\varepsilon)) < \infty \tag{3.6}$$

を示せばよい。実際、Borel-Cantelli の定理により $P\left(\bigcap_{l=1}^{\infty}\bigcup_{m=l}^{\infty} B_m(\varepsilon)\right) = 0$ であるが、(3.5) 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> (3.6) 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 3.8) 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 $m_k \in \boldsymbol{N}$ は $2^{m_k - 1} < k \leq 2^{m_k}$ となるようにとる。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 (3.3) より (3.6) が示された。 □

$\{X_n\}$ の分布が同じならば、定理 3.10 の仮定 (3.3)、特に $E[X_n^2] < \infty$ は不要になる。

定理 3.11 (Kolmogorov の第 2 定理) $\{X_n\}$ を独立で同分布をもつ確率変数列で、期待値 $m = E[X_n]$ は有限であるとする。*(*特に、期待値の定義から $E[|X_n|] < \infty$ に注意。*)* 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 m に概収束する。

証明: $\forall n \in \boldsymbol{N}$ に対して $E[X_n] = 0$ 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し、その分布関数を F_X と書く。

<u>1st step</u> (番号 k に依存した cut-off の導入) $Z_k = X_k 1_{(0,k]}(|X_k|) - \tilde{m}_k$, $\tilde{m}_k = E[X_k 1_{(0,k]}(|X_k|)]$ とおくと、$\{Z_k\}$ は定理 3.10 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 3.10 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 では、次の Lebesgue の優収束定理を用いるため、復習しておく。(証明は略す[1]。)

定理 3.12 (Lebesgue の優収束定理) 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 $n \in \boldsymbol{N}$ に対して $|X_n| \le Y$ を満たすならば

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

が成立する。

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

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

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、1st step により

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

[1]期待値を Lebesgue 積分論の書き方で、$E[X] = \displaystyle\int_{\Omega} X(\omega)\,dP(\omega)$ となることに注意すれば、関数解析学 II で習う定理そのものとなる。

3rd step　$P(\#\{k \in \boldsymbol{N}; |X_k| > k\} < \infty) = 1$ 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

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

に注意する。ただし、2 行目の一つ目の不等号は $\sum_{j=1}^{\infty} j1_{(j,j+1]}(|x|) \leq \sum_{j=1}^{\infty} |x|1_{(j,j+1]}(|x|) = |x|$ となることを用いた。よって、Borel-Cantelli の定理から $\#\{k; |X_k| > k\} < \infty$ a.s. がいえた。　□

注意 3.13 *(1)* 独立で同分布をもつ確率変数 $\{X_n\}$ が $E[X_n^4] < \infty$ を満たすときには、大数の強法則の証明は比較的容易である *(*演習問題 **7***)*。
(2) 大数の法則は一般に期待値が存在しない確率変数列に対しては成り立たない *(*演習問題 **8***)*。
(3) 定理 *3.11* の記号の下、$m \notin [a,b]$ なら $P(a \leq \frac{1}{n}\sum_{k=1}^{n} X_k \leq b)$ は 0 に収束する。もし、$E[e^{tX}] < \infty$ $(\forall t \in \boldsymbol{R})$ 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、応用例も多く盛んに研究されている。

4　特性関数と中心極限定理

4.1　特性関数

定義 4.1 *(1)* 複素数値関数 Z が可測である *(*確率変数である*)* とは、その実部 $X = \mathrm{Re}\, Z$, 虚部 $Y = \mathrm{Im}\, Z$ がともに可測である *(*確率変数である*)* ときにいう。ここで、$Z = X + iY$, $i = \sqrt{-1}$ である。
(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 4.2　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、

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

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

定義 4.3　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。

$$\phi_X(t) = E[e^{itX}], \qquad t \in \boldsymbol{R}$$

命題 4.4　(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、$|\phi_X(t)| \leq 1$ 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 4.2 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) $|s-t|<\delta$ のとき $|\phi_X(s)-\phi_X(t)| \le E[|e^{itX}(e^{i(s-t)X}-1)|] = E[|e^{i(s-t)X}-1|]$. ここで、$|e^{i(s-t)X}-1| \le |e^{i(s-t)X}|+1 = 2$ であるから、Lebesgue の優収束定理により、$\delta \to 0$ のとき、$E[|e^{i(s-t)X}-1|] \to E[|e^0-1|] = 0$ となり主張を得る。 □

命題 4.5 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 4.6 *(1) X が二項分布 $B(n,p)$ に従うとき, $q = 1-p$ とすると,*

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

(2) X が Poisson 分布 $P(\lambda)$ に従うとき

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

例 4.7 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、

$$\phi_X(t) = E[e^{itX}] = \int_{-\infty}^{\infty} e^{itx}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}\,dx = e^{-\frac{1}{2}t^2}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{1}{2}(x-it)^2}\,dx \tag{4.1}$$

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 $\boldsymbol{C}$ 上で正則な関数だから Cauchy の積分定理により $\int_{C_R} e^{-\frac{1}{2}z^2}\,dz = 0$ 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

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

同様に $C_{R,4}$ は $z = -R+iy$ と考えて

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

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、$\sqrt{2\pi} - \int_{-\infty}^{\infty} e^{-\frac{1}{2}(x-it)^2}\,dx = 0$ となるので、(4.1) に代入して $\phi_X(t) = e^{-\frac{1}{2}t^2}\frac{1}{\sqrt{2\pi}}\sqrt{2\pi} = e^{-\frac{1}{2}t^2}$ を得る。 □

系 4.8　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 4.5 を適用して

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

例 4.9　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。

$$\phi_X(t)=E[e^{itX}]=\int_{-\infty}^{\infty}e^{itx}\frac{1}{\pi}\frac{1}{1+x^2}\,dx \tag{4.2}$$

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 $z\neq -i$ で正則だから留数定理により

$$\frac{1}{2\pi i}\int_{C_R}\frac{e^{itz}}{1+z^2}\,dz=\frac{1}{2\pi i}\int_{C_R}\frac{g(z)}{z-i}\,dz=g(i)=\frac{e^{-t}}{2i}. \tag{4.3}$$

一方、

$$\int_{C_R}\frac{e^{itz}}{1+z^2}\,dz=\sum_{n=1}^{2}\int_{C_{R,n}}\frac{e^{itz}}{1+z^2}\,dz \tag{4.4}$$

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

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

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、(4.3), (4.4) 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 (4.2) 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 4.4(iii) による。

$t<0$ のとき、(4.2) 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。　$\square$

命題 4.10　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 $\phi_X^{(k)}(t)=i^kE[X^ke^{itX}]$ となる。

証明: 演習問題 **10** とする。 □

注意 Cauchy 分布の特性関数は $t=0$ で微分可能ではない。これは Cauchy 分布は期待値を持たないことに関係する。

定義 4.11 p 次元確率変数 $\boldsymbol{X}=(X_1,\dots,X_p)'$ に対して、次の $\boldsymbol{R}^p$ 上の関数 $\phi_{\boldsymbol{X}}(\boldsymbol{t})$ を $\boldsymbol{X}$ の特性関数という。

$$\phi_{\boldsymbol{X}}(\boldsymbol{t})=E[e^{i\boldsymbol{t}'\boldsymbol{X}}]=E[\exp\Big\{i\sum_{j=1}^p t_jX_j\Big\}],\qquad \boldsymbol{t}=(t_1,\dots,t_p)'\in\boldsymbol{R}^p$$

命題 4.12 p 次元確率変数 $\boldsymbol{X}$ と p 次正方行列 A と p 次元ベクトル $\boldsymbol{b}$ に対して $\phi_{A\boldsymbol{X}+\boldsymbol{b}}(\boldsymbol{t})=e^{i\boldsymbol{t}'\boldsymbol{b}}\phi_{\boldsymbol{X}}(A'\boldsymbol{t})$.

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

例 4.13 $\boldsymbol{X}=(X_1,\dots,X_p)'$ を p 次元正規分布 $N(\boldsymbol{m},\boldsymbol{\Sigma})$ に従うとする。$\boldsymbol{m}=(m_1,\dots,m_p)'\in\boldsymbol{R}^p$, $\boldsymbol{\Sigma}=(\sigma_{ij})$ 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 $P'\boldsymbol{\Sigma}P=D$ なるようにとる。このとき、$\boldsymbol{Y}=(Y_1,\dots,Y_p)'=P'(\boldsymbol{X}-\boldsymbol{m})$ とすると、定理 2.22 のようにして $\boldsymbol{Y}$ が $N((0,\dots,0)',D)$ に従うことがわかる。(証明は演習問題 **12** とする。) よって、その密度関数を考えれば、$Y_1,\dots,Y_p$ は独立で各 Y_j は正規分布 $N(0,\lambda_{jj})$ に従うことがわかる。よって、定理 2.11 より

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

従って、$\boldsymbol{X}=P\boldsymbol{Y}+\boldsymbol{m}$ に命題 4.12 を適用して

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

4.2 特性関数と分布

定義 4.14 確率変数 X に対して、それが決める $\boldsymbol{R}$ 上の確率測度を μ_X と書く:

$$\mu_X(A)=P(X\in A),\qquad A\in\mathcal{B}(\boldsymbol{R}) \tag{4.5}$$

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

X の分布関数 $F_X(x)$ に対して $F_X(x)=P(X<x)=\mu_X((-\infty,x))$ に注意する。これより、$F_X(x)=F_Y(x)$ $(\forall x\in\boldsymbol{R})$ であることと $\mu_X=\mu_Y$, 即ち、$\mu_X(A)=\mu_Y(A)$ $(\forall A\in\mathcal{B}(\boldsymbol{R}))$ であることが同値となる。

命題 4.15 可測関数 $f(x)$ が $f\geq 0$ もしくは $\int_{\boldsymbol{R}}|f(x)|\mu_X(dx)<\infty$ を満たせば、$E[f(X)]=\int_{\boldsymbol{R}}f(x)\mu_X(dx)$.

略証: $f(x)$ が階段関数 $f(x)=\sum a_i 1_{A_i}(x)$ の場合は

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

と示すことができ、一般の場合は前期に §2.7 で期待値を定義するのに行ったように $f(x)\geq 0$ なら単調収束定理を用い、一般の実数値の場合は $f(x)=f^+(x)-f^-(x)$ として、さらに、複素数値の場合は $f(x)=\operatorname{Re}f(x)+i\operatorname{Im}f(x)$ として証明できる。 □

この節の目標は次を示すことにある。

定理 4.16 *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 $\phi_X(t)=\phi_Y(t)$, $\forall t\in\boldsymbol{R}$, であれば、$\mu_X=\mu_Y$ *(*あるいは $F_X(x)=F_Y(x)$, $\forall x\in\boldsymbol{R}$, *)* となる。

補題 4.17 *Dirichlet* 積分

$$f_T(\alpha) = \int_0^T \frac{\sin \alpha t}{t}\, dt, \qquad T > 0,\ \alpha \in \boldsymbol{R} \tag{4.6}$$

について、*(1)* $\sup_{T,\alpha} |f_T(\alpha)| < \infty$, *(2)* $\lim_{T\to\infty} f_T(\alpha) = \frac{\pi}{2}\left\{1_{(0,\infty)}(\alpha) - 1_{(-\infty,0)}(\alpha)\right\}$ 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。$|\sin u| \le |u|$ ($|u| \le \frac{\pi}{2}$) より、$\sup_{|\alpha T| \le \frac{\pi}{2}} |f_T(\alpha)| \le \frac{\pi}{2}$. 一方、$M \to \infty$ のときは、部分積分により $\int_{\pi/2}^M \frac{\sin u}{u}\, du$ は (条件) 収束することが示せるので、$\frac{\sin u}{u}$ が遇関数であることに注意すれば証明は容易である。詳しい証明は演習問題 **13** とする。
(2) $\alpha = 0$ のときは明らか。$\alpha > 0$ のときは、次に補題 4.18 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。 □

補題 4.18 $\displaystyle\int_0^\infty \frac{\sin x}{x}\, dx = \frac{\pi}{2}.$

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。)

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

Cauchy の積分定理より

$$0 = \int_{C_{\varepsilon,R}} \frac{e^{iz}}{z}\, dz = \sum_{n=1}^4 \int_{C_{\varepsilon,R,n}} \frac{e^{iz}}{z}\, dz. \tag{4.7}$$

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 4.17(1) により $\varepsilon \to +0,\ R \to \infty$ のとき右辺は $2i \int_0^\infty \frac{\sin x}{x}\, dx$ 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 の優収束定理により右辺は 0 に収束する。$C_{\varepsilon,R,4}$ も同様に

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

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 の優収束定理により右辺は $-i\int_0^\pi d\theta = -i\pi$ に収束する。よって、(4.7) 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。 □

定理 4.20 の証明においては、Fubini の定理を用いるため、その主張のみを述べておく。(詳しくは Lebesgue 積分論の講義で勉強すること。)

定理 4.19 (Fubini の定理) μ_1, μ_2 を $\boldsymbol{R}$ 上の測度とする。関数 $f(x,y)$ が *2* 変数関数として *Borel* 可測で、$f(x,y) \geq 0$ または $\displaystyle\int_{\boldsymbol{R}}\int_{\boldsymbol{R}} |f(x,y)|\,\mu_1(dx)\mu_2(dy) < \infty$ を満たせば、積分の順序を交換できる。

$$\int_{\boldsymbol{R}}\left(\int_{\boldsymbol{R}} f(x,y)\,\mu_1(dx)\right)\mu_2(dy) = \int_{\boldsymbol{R}}\left(\int_{\boldsymbol{R}} f(x,y)\,\mu_2(dy)\right)\mu_1(dx)$$

定理 4.20 *(***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、$a, b \in \boldsymbol{R}$ *(*$a < b$*)* が F の連続点ならば、次が成立する。

$$F(b) - F(a) = \frac{1}{2\pi}\lim_{T\to\infty}\int_{-T}^{T}\frac{e^{-itb} - e^{-ita}}{-it}\phi(t)\,dt. \tag{4.8}$$

ここで x が F_X の連続点であるとは $\displaystyle\lim_{y\to x} F_X(y) = F_X(x)$ 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 4.15 により

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

である。$g(x) = e^{-itx}$ に平均値の定理を用いて

$$e^{-itb} - e^{-ita} = g'(\xi)(b-a) = -ite^{-it\xi}(b-a) \quad (a < \xi < b)$$

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

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

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 4.17 の $f_T(\alpha)$ を用いると、

$$\text{右辺} = 2\int_{\boldsymbol{R}}\left(f_T(x-a) - f_T(x-b)\right)\mu_X(dx)$$

となる。ここで、補題 4.17(1), (2) に注意して Lebesgue の優収束定理を用いると、$T \to \infty$ のとき、

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

となる。よって、a, b が F の連続点だから、

$$(4.8)\ \text{の右辺} = \frac{1}{2\pi}\pi\left[1 - F(a+0) - F(a) - (1 - F(b+0)) + F(b)\right] = F(b) - F(a)$$

を得る。 □

定理 4.16 の証明: F_X と F_Y の連続点の合併集合を 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 $\boldsymbol{R}$ で稠密となる。定理 4.20 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、分布関数は左連続であるから、$\forall x \in \boldsymbol{R}$ 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、$\mu_X = \mu_Y$ である。 □

例 4.21 $X_1,\dots,X_n$ を独立で、各 X_j が *Poisson* 分布 $P(\lambda_j)$ に従うとする。このとき、$Y=X_1+\cdots+X_n$ は *Poisson* 分布 $P(\lambda_1+\cdots+\lambda_n)$ に従う。実際、

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

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 *4.6(2)* を用いた。よって、再び例 *4.6(2)* により右辺が *Poisson* 分布 $P(\lambda_1+\cdots+\lambda_n)$ の特性関数とわかるから、定理 *4.16* により主張を得る。

例 4.22 $X_1,X_2,\dots$ 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 *4.9* 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。

4.3 法則収束と弱収束

定義 4.23 (法則収束) 任意の $f\in C_b(\boldsymbol{R})$ 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)* するという[2]。ここで、$C_b(\boldsymbol{R})$ は $\boldsymbol{R}$ 上の有界連続関数全体を表す。$X_n\Longrightarrow X$ と表す。

法則収束は、分布関数の間の距離として距離付け可能となる。例えば、Durrett 著 Probability: Theory and Examples を参照のこと。

定理 4.24 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: <u>1st step</u> まず、$\{X_n\}$ が X に概収束する場合を考える。このとき、$f\in C_b(\boldsymbol{R})$ 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 $|f(x)|\le M$ $(\forall x\in\boldsymbol{R})$ 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 の優収束定理 (定理 3.12) 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。
<u>2nd step</u> $\{X_n\}$ が X に確率収束するとし、$f\in C_b(\boldsymbol{R})$ に対し、$a_n=E[f(X_n)]$ とおく。まず、$|f(x)|\le M$ $(x\in\boldsymbol{R})$ 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 3.6 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 $f\in C_b(\boldsymbol{R})$ に対して成立するから、$\{X_n\}$ は X に法則収束する。 □

例 4.25 定理 *4.24* の逆は、必ずしも成立しない。実際、確率変数列 $\{X_n\}$ を独立で各 $n\in\boldsymbol{N}$ に対し $P(X_n=1)=P(X_n=-1)=1/2$ なるとする。このとき、$\forall f\in C_b(\boldsymbol{R})$ に対して

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

[2]教科書と定義が異なることに注意せよ。同値性は定理 4.27 で示す。

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

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

となり、$\{X_n\}$ は X_1 に確率収束しないことがわかる。

このように、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 4.26 μ_n, $n = 1, 2, \ldots$, と μ を *(*$\boldsymbol{R}$ 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、任意の有界連続関数 $f \in C_b(S)$ 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、

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

となるが、ここで、$a \in \boldsymbol{R}$ を任意とし、(グラフを書く)

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

とすると、$\int_{\boldsymbol{R}} f_n(x)\mu(dx) = \int_{\boldsymbol{R}} f_n(x)\nu(dx)$ であるが、$f_n(x) \to 1_{(-\infty,a)}(x)$ $(n \to \infty)$ と $0 \leq f_n(x) \leq 1$ ($x \in \boldsymbol{R}$) に注意して Lebesgue の優収束定理を用いると、$\int_{\boldsymbol{R}} 1_{(-\infty,a)}(x)\mu(dx) = \int_{\boldsymbol{R}} 1_{(-\infty,a)}(x)\nu(dx)$, 即ち、$\mu((-\infty,a)) = \nu((-\infty,a))$ となるからである。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 4.27 確率変数列 $\{X_n\}$ が 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 $F_Y(x) = P(Y < x)$ と表す。
(1) $\{X_n\}$ は X に法則収束する。
(2) F_X の任意の連続点 x に対して $\lim_{n\to\infty} F_{X_n}(x) = F_X(x)$ が成立する。

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

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

で定める (グラフを書く)。このとき、

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

となることに注意する。まず、

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

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

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

を得る。よって、$\delta \to +0$ として、

$$\limsup_{n\to\infty} F_{X_n}(x) \le \lim_{\delta\to+0} F_X(x+\delta) = F_X(x) \tag{4.9}$$

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。同様に、

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

で $f_\delta^- \in C_b(\boldsymbol{R})$ より

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

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。これと、(4.9) 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) $\Rightarrow$ (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 $\boldsymbol{R}$ 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 $a, b \in \boldsymbol{R}$ $(a < b)$ 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 $f \in C_b(\boldsymbol{R})$ が任意に与えられたとして点列 $a = a_0 < a_1 < \cdots < a_K = b$ を

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

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

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

とおく。$\|f\|_\infty = \sup_{x\in\boldsymbol{R}} |f(x)|$ 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、$\lim\limits_{n\to\infty}E[f(X_n)]=E[f(X)]$ となる。 $\square$

距離空間 S 上の確率測度の列 μ_n については次が成り立つ。主張のみを述べておく。

定理 4.28 S は完備可分距離空間とする。このとき、確率測度の列 $\{\mu_n\}$ と確率測度 μ に対して次の *4* 条件 *(1)–(4)* はどの *2* つも同値である。
(1) $\{\mu_n\}$ は μ に弱収束する。
(2) S の任意の開集合 G に対して $\liminf\limits_{n\to\infty}\mu_n(G)\ge\mu(G)$ が成立する。
(3) S の任意の閉集合 F に対して $\limsup\limits_{n\to\infty}\mu_n(F)\le\mu(F)$ が成立する。
(4) S の部分集合 $A\in\mathcal{B}(S)$ が $\mu(\partial A)=0$ を満たせば、$\lim\limits_{n\to\infty}\mu_n(A)=\mu(A)$ が成立する。ただし、∂A は A の境界を表す。

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 4.29 *(***Prohorov の定理***)* 確率変数の族 $\{X_\alpha\}$ は対して次の条件 *(1), (2)* は同値である。
(1) $\{X_\alpha\}$ は点列 *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、$\{X_{\alpha_{n'}}\}$ 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

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

とできる。

確率測度の族 $\{\mu_\alpha\}$ に対して (2) のように $\inf\limits_\alpha\mu_\alpha([-M,M])\ge1-\varepsilon$ とできるとき、$\{\mu_\alpha\}$ は **tight** であるという。(2) の条件は $\{X_\alpha\}$ に対応する確率測度の族が tight となることにほかならない。定理 4.29 は完備可分距離空間 S 上の確率測度の族に対しても $[-M,M]$ を S の compact 集合とすることで成立する (むしろこの場合を Prohorov の定理とよぶ)。

この定理を証明するために次の補題を準備する。

補題 4.30 *(***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 $\{F_{n_k}(x)\}$ と左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

$$\lim_{k\to\infty} F_{n_k}(x) = F(x) \tag{4.10}$$

とできる。

証明: 1st step 有理数全体 $\boldsymbol{Q} = \{x_1, x_2, \cdots\}$ 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、$\{F_{n_k}(x_j)\}_{k\ge j}$ は $\{F_{j,n}(x_j)\}_{n\ge 1}$ 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、$F_{n_k}(x_i) \le F_{n_k}(x_j)$ となるから $\tilde{F}(x_i) \le \tilde{F}(x_j)$ となる。

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

$$F(x) = \sup\{\tilde{F}(y); y \in \boldsymbol{Q}, y < x\} \tag{4.11}$$

とおく。このとき、F が単調増加であることは明らか。また、$F(x)$ は左連続となる。(演習問題とする。) x を F の連続点とし、(4.10) を示す。$\varepsilon > 0$ とし、$z_1, z_2, z_3 \in \boldsymbol{Q}$ を

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

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、(4.11) より $k \to \infty$ のとき

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

だから、k が十分大ならば

$$F(x) - \varepsilon < F_{n_k}(z_1) \le F_{n_k}(x) \le F_{n_k}(z_2) < F(x) + \varepsilon,$$

即ち、$|F_{n_k}(x) - F(x)| < \varepsilon$ が成立するから、(4.10) は成立する。 □

定理 4.29 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 $n \in \boldsymbol{N}$ に対して、

$$P(X_{X_{\alpha_n}} \in [-n, n]) < 1 - \varepsilon \tag{4.12}$$

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、(4.12) により

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

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

(2) $\Rightarrow$ (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 4.30) により、$F_{X_{\alpha_n}}$ の部分列 $F_{X_{\alpha_{n_k}}}$ と左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 $y < -M,\ M < x$ ととれば

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

となるので、$F(x)$ は分布関数となっている。よって、$F_X(x) = F(x)$ となる確率変数 X が存在する。 □

4.4 特性関数と法則収束

定理 4.31 $\{X_n\}$ 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、$\forall t \in \boldsymbol{R}$ に対して $\phi_n(t)$ は $\phi_X(t)$ に収束する。
(2) $\forall t \in \boldsymbol{R}$ 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 4.29 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

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

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、

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

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 $M > 0$ があって、

$$\left|I^{(1)}(M)\right| < \frac{\varepsilon}{2}$$

とできる。この M に対して、$\lim_{n\to\infty} \phi_n(s) = \phi(s)$ $(\forall s \in \boldsymbol{R})$ かつ $|\phi(s) - \phi_n(s)| \le |\phi(s)| + |\phi_n(s)| \le 2$ だから、Lebesgue の優収束定理により

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

となる。すなわち、ある n_0 があって、

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

とできる。以上より、$\inf_{n\ge n_0} P(|X_n| \le M) \ge 1 - \varepsilon$ となるので、定理 4.29 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 4.16 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。 □

4.5 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 4.32 *(**中心極限定理**)* $\{X_n\}$ は独立で同分布をもつ確率変数列で、$E[X_n] = m,\ V(X_n) = \sigma^2 > 0$ を満たすとする。このとき、

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

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

$$\lim_{n\to\infty} P\left(a \le U_n \le b\right) = \frac{1}{\sqrt{2\pi}}\int_a^b e^{-\frac{y^2}{2}}\,dy, \qquad -\infty < a < b < \infty. \tag{4.13}$$

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 4.10 により、$\phi(t)$ は C^2-級であるから Taylor の定理により、

$$\phi\Big(\frac{t}{\sqrt{n}}\Big) = \phi(0) + \frac{t}{\sqrt{n}}\phi'(0) + \frac{1}{2}\Big(\frac{t}{\sqrt{n}}\Big)^2\phi''\Big(\theta\frac{t}{\sqrt{n}}\Big), \qquad 0 < \theta < 1$$

とできる。ここで、$\phi(0) = 1,\ \phi'(0) = E[Z_1] = 0$ であり、$\phi''(\theta\frac{t}{\sqrt{n}}) = -E[Z_1^2 e^{i\theta\frac{t}{\sqrt{n}}}]$ であるが、$|Z_1^2 e^{i\theta\frac{t}{\sqrt{n}}}| \le Z_1^2$ であるから、Lebesgue の優収束定理により

$$\lim_{n\to\infty}\phi''\Big(\theta\frac{t}{\sqrt{n}}\Big) = -E[Z_1^2] = -1.$$

したがって、

$$\lim_{n\to\infty}\left(\phi\Big(\frac{t}{\sqrt{n}}\Big)\right)^n = \lim_{n\to\infty}\left(1 + \frac{1}{n}\frac{t^2}{2}\phi''\Big(\theta\frac{t}{\sqrt{n}}\Big)\right)^n = e^{-\frac{t^2}{2}}$$

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 4.31(2) により、$\{U_n\}$ は Y に法則収束する。(4.13) は Y の分布関数 $F_Y(x)$ が連続だから、定理 4.27 により成立する。 □

系 4.33 *(***de Moivre-Laplace の定理***)* 成功率 p の *Bernoulli* 試行列 $X_1, X_2, \ldots$ を考える*:* $X_1, X_2, \ldots$ は独立な確率変数列で $P(X_n = 1) = p, P(X_n = 0) = 1 - p.$ このとき、$S_n = \sum_{k=1}^n X_k$ とおくと、

$$\lim_{n\to\infty} P\left(a \leq \frac{S_n - np}{\sqrt{np(1-p)}} \leq b\right) = \frac{1}{\sqrt{2\pi}}\int_a^b e^{-\frac{y^2}{2}}\,dy, \qquad -\infty < a < b < \infty. \tag{4.14}$$

証明: $E[X_k] = p,\ V(X_k) = p(1-p)$ であるから、定理 4.32 により従う。 □

注意 4.34 *(1)* 系 *4.33* で S_n は二項分布 $B(n,p)$ に従う。このことから、*(4.14)* は $B(n,p)$ に従う確率変数 S_n の、n が十分大きいときの確率分布の近似値を与えている *(cf.* 例 *4.35)*。
(2) *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、直接 *Stirling* の公式を用いて直接分布を計算することで *(4.14)* 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。

例 4.35 正確に作られたサイコロを *720* 回投げて、*6* の目の出る回数が *130* 回以上 *150* 回以下となる確率を、中心極限定理を用いて求めよう。

S で *6* の目の出る回数をあらわすと、S は二項分布 $B(720, \frac{1}{6})$ に従う。このとき、求める確率は $P(130 \leq S \leq 150)$ となる。今、$E[S] = 120,\ V(S) = 720\cdot\frac{1}{6}(1-\frac{1}{6}) = 100$ だから、中心極限定理により、$Z := \dfrac{S-120}{\sqrt{100}}$ は正規分布で近似できる。したがって、

$$\begin{aligned} P(130 \leq S \leq 150) &= P\Big(\frac{130-120}{\sqrt{100}} \leq \frac{S-120}{\sqrt{100}} \leq \frac{150-120}{\sqrt{100}}\Big) = P(1 \leq Z \leq 3) \\ &= P(Z > 1) - P(Z > 3) = 0.1586 - 0.001349 = 0.157251 \end{aligned}$$

より 0.1573 となる。

注意 4.36 *2* 項分布の正規近似は n が大きければかなり良い近似であるが、n が小さいときも**半整数補正**を用いることにより、近似度がよくなることが知られている。これは、上の例題で $P(a \leq S \leq b)$ を求める際、$P(a - 0.5 \leq S \leq b + 05)$ と補正してから計算する方法である。このことは、二項分布のグラフを柱状グラフとして考えて近似するとすればよい近似となることは直感的に明らかであろう。例 *4.35* では、

$$\begin{aligned} P(130 - 0.5 \leq S \leq 150 + 0.5) &= P\Big(\frac{129.5-120}{\sqrt{100}} \leq \frac{S-120}{\sqrt{100}} \leq \frac{150.5-120}{\sqrt{100}}\Big) = P(0.95 \leq Z \leq 3.05) \\ &= P(Z > 0.95) - P(Z > 3.05) = 0.1710 - 0.001144 = 0.169856 \end{aligned}$$

より 0.1699 となる。一方、*Maple* を用いて厳密な値を求めると[3]

$$P(130 \leq S \leq 150) = P(S \leq 150) - P(S \leq 129) = 0.9984997678 - 0.8292544225 = 0.1692453453$$

となる。

例 4.37 $X_1, X_2, \ldots$ を *Poisson* 分布 $Po(1)$ に従う独立な確率変数列とする。このとき、例 *4.21* により、$S_n := X_1 + \cdots + X_n$ は *Poisson* 分布 $Po(n)$ に従う。ここで、$E[X_1] = 1,\ V(X_1) = 1$ であるから、中心極限定理により $\dfrac{S_n - n}{\sqrt{n}}$ は正規分布で近似できる。特に、

$$\lim_{n\to\infty} P\left(\frac{S_n - n}{\sqrt{n}} \leq 0\right) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^0 e^{-\frac{y^2}{2}}\,dy = \frac{1}{2}$$

であるが、$P(S_n \leq n) = \sum\limits_{k=0}^n e^{-n}\dfrac{n^k}{k!}$ であるから、

$$\lim_{n\to\infty} e^{-n}\left(1 + n + \frac{n^2}{2!} + \cdots + \frac{n^n}{n!}\right) = \frac{1}{2}$$

となることがわかった。

[3] S が $B(n,p)$ に従うとき、$P(S \leq k)$ を求めるには「stats[statevalf,dcdf,binomiald[n,p]](k); 」とする。

5 Dynkin 族定理

定義 5.1 *(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)* $A, B \in \mathcal{P} \Rightarrow A \cap B \in \mathcal{P}$

の *2* 条件を満たすときにいう。

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

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

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

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

の *3* 条件を満たすときにいう。

集合 S の部分集合族 $\mathcal{C}$ に対して、$\mathcal{C}$ を含む最小の Dynkin 族を $\mathcal{L}(\mathcal{C})$ と表す。$\mathcal{D}_\alpha$ $(\alpha \in A)$ が Dynkin 族であれば $\bigcap_{\alpha \in A} \mathcal{D}_\alpha$ も Dynkin 族となる (証明は演習問題とする) ことから、$\{\mathcal{D}_\alpha\}_{\alpha \in A}$ をすべての $\mathcal{C}$ を含む Dynkin 族とし、$\mathcal{L}(\mathcal{C}) = \bigcap_{\alpha \in A} \mathcal{D}_\alpha$ とすればよい。

定理 5.2 *(***Dynkin 族定理***)* $\mathcal{P}$ が π 族のとき $\mathcal{L}(\mathcal{P}) = \sigma(\mathcal{P})$ となる[4]。

証明: σ-加法族は Dynkin 族となる (証明は演習問題とする) 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-加法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 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) $B_n \in \mathcal{G}_A$, $B_n \subset B_{n+1}$ $(\forall n \in \boldsymbol{N})$ 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 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-加法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) $A_n \in \mathcal{L}(\mathcal{P})$ $(n \in \boldsymbol{N})$ 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。 □

定理 5.3 $F_X(x) = F_Y(x)$ *(*$\forall x \in \boldsymbol{R}$*)* ならば、$\mu_X(A) = \mu_Y(A)$ *(*$\forall A \in \mathcal{B}(\boldsymbol{R})$*)* となる。

証明: $\mathcal{J} = \{(-\infty, x); x \in (-\infty, \infty]\}$ とし、$\mathcal{A} = \{A \in \mathcal{B}(\boldsymbol{R}); \mu_X(A) = \mu_Y(A)\}$ とする。このとき、$\mathcal{J}$ が π 族になることは明らか。$\mathcal{A}$ が Dynkin 族となることは測度の性質より容易に証明できる (演習問題とする)。$x \in (-\infty, \infty)$ 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、$\mathcal{J} \subset \mathcal{A}$ となる。よって、定理 5.2 により $\sigma(\mathcal{J}) = \mathcal{L}(\mathcal{J}) \subset \mathcal{A}$. ところで、$\mathcal{J}$ を含む最小の σ-加法族 $\sigma(\mathcal{J})$ は Borel 集合族 $\mathcal{B}(\boldsymbol{R})$ と一致するので、$\mathcal{B}(\boldsymbol{R}) \subset \mathcal{A}$ となり主張を得る。 □

[4] $\sigma(\mathcal{P})$ は $\mathcal{P}$ を含む最小の σ-加法族であった。