1.6　分布関数

分布というものはまだ関数に比べて分かりにくいような気がする．そこで，関数の言葉で分布を語ることを考える．

定義 1.2 確率空間 $(\Omega, \mathcal{F}, P)$ 上で定義された確率変数 X に対してその分布関数 $F_X(t)$ を

$$F_X(t) := \mu_X((-\infty, t]) = P(X \leq t) \tag{1.11}$$

によって定義する．

命題 1.14 確率変数 X の分布関数 $F_X(t)$ は以下の性質を持つ．

(i)　$F_X(t)$ は $t \in \mathbf{R}$ について単調非減少．

(ii)　$F_X(t)$ は $t \in \mathbf{R}$ について右連続．

(iii)　$\lim_{t \to -\infty} F_X(t) = 0, \quad \lim_{t \to \infty} F_X(t) = 1$

証明　(i) $s < t$ のとき $F_X(s) \leq F_X(t)$ を言えばよい．明らかに，$\{\omega \in \Omega; X(\omega) \leq s\} \subset \{\omega \in \Omega; X(\omega) \leq t\}$ だから，

$$F_X(s) = P(X \leq s) \leq P(X \leq t) = F_X(t)$$

(ii) $t_n \searrow t$ のとき $F_X(t_n) \to F_X(t)$ を言えばよい．このとき $X^{-1}((-\infty, t_n]) = \{\omega \in \Omega; X(\omega) \leq t_n\}$ は n について単調に減少して，

$$\bigcap_{n=1}^{\infty} X^{-1}((-\infty, t_n]) = X^{-1}(\bigcap_{n=1}^{\infty}(-\infty, t_n]) = X^{-1}((-\infty, t])$$

なので，確率の連続性から

$$F_X(t) = P(X \leq t) = P(\bigcap_{n=1}^{\infty}\{X \leq t_n\}) = \lim_{n\to\infty} P(X \leq t_n) = \lim_{n\to\infty} F_X(t_n)$$

(iii) 上と同じように考えると，

$$\begin{aligned}\lim_{t\to-\infty} F_X(t) &= P(\cap_{n=1}^{\infty} X^{-1}((-\infty, -n])) = P(X^{-1}(\emptyset)) = 0\\ \lim_{t\to\infty} F_X(t) &= P(\cup_{n=1}^{\infty} X^{-1}((-\infty, n])) = P(X^{-1}(\mathbf{R})) = 1\end{aligned}$$

例 1.12 $\mu(\{0\}) = 1$ のとき，$\mu(A) = 1 \Leftrightarrow A \ni 0$ という分布が得られる．この μ を δ_0 とかき，$\{0\}$ に集中する質量をもつ**ディラック (Dirac) 分布**という．この分布関数は

$$F(t) = \begin{cases} 0 & t < 0; \\ 1 & t \geq 0 \end{cases}$$

となる．

例 1.13 サイコロの目の分布は

$$\mu(\{i\}) = \frac{1}{6}, \quad i = 1,2,3,4,5,6$$

で与えられる．この分布関数を求めてみよう。$x < 1$ のとき、

$$F(x) = \mu((-\infty, x]) = P(\text{ 目の数が } x \text{ 以下}) = 0$$

$1 \leq x < 2$ のとき

$$F(x) = \mu((-\infty, x] \cap \{1,2,3,4,5,6\}) = \mu(\{1\}) = \frac{1}{6}$$

$2 \leq x < 3$ のとき

$$F(x) = \mu((-\infty, x] \cap \{1,2,3,4,5,6\}) = \mu(\{1,2\}) = \frac{1}{3}$$

$3 \leq x < 4$ のとき

$$F(x) = \mu((-\infty, x] \cap \{1,2,3,4,5,6\}) = \mu(\{1,2,3\}) = \frac{1}{2}$$

$4 \leq x < 5$ のとき

$$F(x) = \mu((-\infty, x] \cap \{1,2,3,4,5,6\}) = \mu(\{1,2,3,4\}) = \frac{2}{3}$$

$4 \leq x < 5$ のとき

$$F(x) = \mu((-\infty, x] \cap \{1,2,3,4,5,6\}) = \mu(\{1,2,3,4,5\}) = \frac{5}{6}$$

$6 \leq x$ のとき

$$F(x) = \mu((-\infty, x] \cap \{1,2,3,4,5,6\}) = \mu(\{1,2,3,4,5,6\}) = 1$$

つまり、 $F(x)$ は $i = 1,2,3,4,5,6$ の各点で $\frac{1}{6}$ ずつジャンプする右連続な階段関数になる。

例 1.14 $a>0$ に対して μ が次で与えられるとき，パラメータ a の**コーシー (Cauchy) 分布**という．$A \in \mathcal{B}$ のとき，

$$\mu(A) = \frac{a}{\pi}\int_A \frac{1}{a^2+x^2}dx$$

この分布関数は

$$F(t) = \frac{1}{\pi}\left[\mathrm{Arctan}\frac{t}{a} + \frac{\pi}{2}\right]$$

練習問題 1.12 次の分布の分布関数を求めよ．

(i) 区間 $[a,b]$ 上の**一様分布**は

$$\mu(A) = \frac{1}{b-a}\int_{A\cap[a,b]} dx$$

である．

(ii) 2 項分布 $B(3, 0.5)$.

(iii) パラメータ 1 のポアソン分布

補遺　実は命題 1.14 の逆が言える．つまり，右連続な単調非減少関数 F が

$$\lim_{t\to-\infty} F(t) = 0, \quad \lim_{t\to\infty} F(t) = 1$$

を満たすとき，ある $(\mathbf{R}, \mathcal{B})$ 上の確率 μ_F が唯一つ存在して

$$\mu_F((-\infty, t]) = F(t) \quad \forall t \in \mathbf{R}$$

が成り立つことが知られている．すこし説明しよう．

言葉の準備がいる．いつものように Ω を勝手な集合 $(\neq \emptyset)$ とするとき，Ω の部分集合の族 $\mathcal{A}$ が Ω の加法族であるとは，

(i)　$\Omega \in \mathcal{A}$,

(ii)　$A \in \mathcal{A}$ ならば $A^c \in \mathcal{A}$

(iii)　$A, B \in \mathcal{A}$ ならば $A \cup B \in \mathcal{A}$

の 3 つが成り立つときに言う．Ω の σ-加法族は Ω の加法族になるが，逆は正しくない．

例 1.15

$$\mathcal{C}_5 := \{(a,b]; -\infty \le a < b < \infty\} \cup \{\mathbf{R}\} \cup \{\emptyset\}$$

とし，

$$\mathcal{A} := \{A_1 \cup \ldots \cup A_n; \{A_j\}_{j=1}^n \text{ は排反 }, A_1, \ldots, A_n \in \mathcal{C}_5\}$$

は $\mathbf{R}$ の加法族である．問題**??**の $\mathcal{C}_2$ と比べると，$\mathcal{C}_5 \supset \mathcal{C}_2$ だから，$\sigma(\mathcal{C}_5) = \mathcal{B}$ となるが，$\mathcal{C}_5$ 自身には $\cup_{n=1}^{\infty}(2n, 2n+1]$ というボレル集合は含まれていない．

定理 1.15 (カラテオドリの拡張定理[8])
Ω の加法族 $\mathcal{A}$ 上で定義された集合関数 Q が次の条件を満たすものとする．

[8]証明については，西尾真喜子著「確率論」2 章 §3 参照

(i)　$0 \le Q(A) \le 1\ \forall A \in \mathcal{A}, \quad Q(\Omega) = 1$

(ii)　$A_1, \ldots, A_n \in \mathcal{A}$ が排反ならば

$$Q(A_1 \cup \ldots \cup A_n) = \sum_{k=1}^n Q(A_k)$$

(iii)　$\{A_n\} \in \mathcal{A}$ が $A_1 \supset A_2 \supset \ldots$ かつ $\cap_{n\ge1} A_n = \emptyset$ のとき，

$$\lim_{n\to\infty} Q(A_n) = 0$$

このとき，$\sigma(\mathcal{A})$ 上に唯一つの確率 Q^* がとれて，

$$Q^*(A) = Q(A) \quad \forall A \in \mathcal{A} \tag{1.12}$$

が成り立つ．

F から 例 1.15 の $\mathcal{A}$ 上に Q を作ってみる．$A = [a_1, b_1) \cup [a_2, b_2) \cup \ldots \cup [a_n, b_n) \in \mathcal{A}$ とする．$[a_i, b_i)$ たちは排反である．このとき，

$$Q(A) := \sum_{i=1}^n \left[F(b_i) - F(a_i)\right]$$

と定める．$Q(\mathbf{R}) = 1, Q(\emptyset) = 0$ である．上の条件 (i), (ii) はすぐに分かる．(iii) が面倒だが，対偶を示す．$A_1 \supset A_2 \supset \ldots$ を $\mathcal{A}$ の要素で

$$\lim_{n\to\infty} Q(A_n) = \alpha > 0$$

として $\cap A_n \neq \emptyset$ をいう．

各 n について，

$$A_n = (a_1, b_1] \cup \ldots \cup (a_k, b_k] \quad \text{各区間は排反}$$

の形に書ける[9]．作り方から

$$Q(A_n) = \sum_{i=1}^k \big(F(b_i) - F(a_i)\big)$$

で，F の右連続性から，各 i について $a_i < a_i' < b_i$ となる a_i' を十分 a_i に近づけると

$$F(a_i') - F(a_i) < \frac{1}{k}2^{-n-1}\alpha$$

とできる．このとき，$A_n' = (a_1', b_1] \cup \ldots \cup (a_k', b_k]$ とおくと，これは排反な区間の和で，$Q(A_n \setminus A_n') \le 2^{-n-1}\alpha$ となっている．これから

$$\begin{aligned} Q(A_1' \cap \ldots \cap A_n') &= Q(A_n \setminus \{(A_1 \setminus A_1') \cup \ldots \cup (A_n \setminus A_n')\}) \\ &\ge Q(A_n) - \sum_{i=1}^n Q(A_i \setminus A_i') \\ &\ge (1 - 2^{-2} - 2^{-3} - \ldots 2^{-n-1})\alpha \ge \frac{\alpha}{2} \end{aligned}$$

何が言いたいかというと、これは $A_1' \cap \ldots \cap A_n' \neq \emptyset$ を意味している。A_n' の各半開区間 $(a_i', b_i]$ を $[a_i', b_i]$ としたものを A_n^* と書く．$A_n^* \supset A_n'$ だから $A_1^* \cap \ldots \cap A_n^*$ も空集合でない．しかもこの集合は A_1 に含まれる閉集合なのでコンパクト性から

$$A_1^* \cap A_2^* \cap \ldots \neq \emptyset \tag{1.13}$$

これは $\cap A_n$ に含まれるので，$\cap A_n \neq \emptyset$.

[9]正確にはこれらの半開区間は n ごとに違っていると考える方が自然なので，$a_1 = a_1^{(n)}, b_1 = b_1^{(n)}, \ldots, k = k(n)$ などと書くべきだが，記号が繁雑になるので，このように書いておこう．