立教大学　数理統計学序論１・数理統計学１

今野 良彦

大阪公立大学

2025 年 9 月 20 日

ガイダンス

- 日程: 9/20, 10/4, 10/25, 11/8, 12/20, 1/10, 1/17
- 講義のホームページ:
  `https://mcm-www.jwu.ac.jp/~konno/stat_rikkyo.html`

この講義で学ぶこと

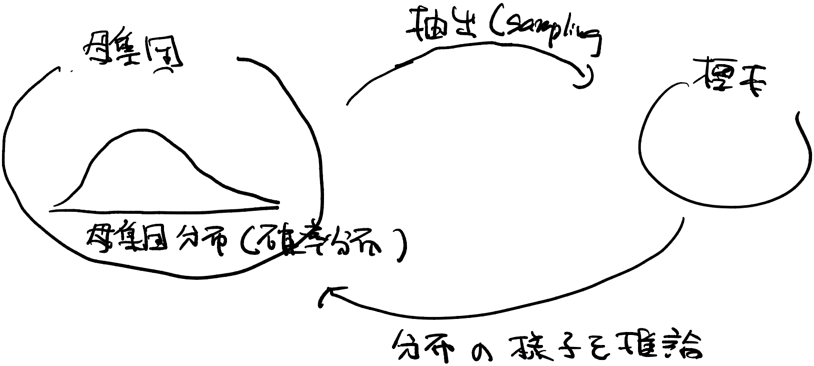
- 統計学: データ解析のための方法を研究する学問 (竹村 (2020, p.1)).
  - 記述統計学: 調査や実験によって得られたデータの整理・解釈を助けるための統計的手法のついての学問.
  - 数理統計学 (推測統計学): 統計的推測の論理を数学的に整理した学問.

統計的推測の例: 新薬の効果

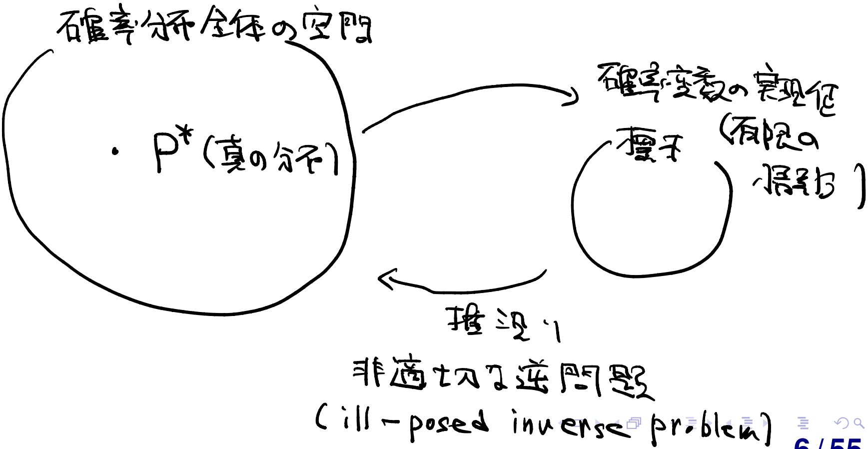
- 臨床実験では, 治験参加者（モニター）に対する新薬の有効性を調べる（個体差がある）.
- 人への有効性を知りたい.

一部を調べて全体の様子を推論する.

統計的推測の考え方 (1)

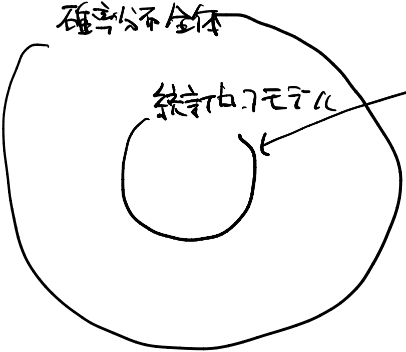


統計的推測の考え方 (2)


確率分布全体の空間
P* (真の分布)
確率変数の実現値
(有限の情報)
標本
推測
非適切な逆問題
(ill-posed inverse problem)

統計的モデル (Statistical models)

統計的モデルの導入



分布に名前付けてる
= 母数

母数の範囲
= 母数空間

P*(真の分布)を探す ⇔ P* に対応する母数(真の母数)を探す.

母数空間

- 有限次元母数空間をもつ統計的モデル $\Rightarrow$ 母数モデル.
- 無限次元母数空間をもつ統計的モデル $\Rightarrow$ ノンパラメトリックモデル（言葉使いが変であるが, 統計学では伝統的な用語なので, それを踏襲）.
- 基本的な議論を確率論の言葉を使って行う.

講義の構成

- 統計的推測理論に必要な確率論の道具立ての準備.
- 統計的推測理論を数学的に整理.

注意

- 確率論: 順問題 ··· 生成メカニズムからその結果を記述.
- 数理統計学: 逆問題 ··· 結果（データ）から生成メカニズムを推論.

数理統計学の隣接分野

- Data Science
- Machine Learning
- Pattern Recogniton
- Signal Processing
- Data mining
- Biostatistics
- Econometrics
- Pychometric

などなど.

1.1 確率

- 試行: 確率論で扱う行為（実験, 調査など）.
- 標本空間: 試行の起こりうる結果すべての集めた集合. この講義では, Ω と記す.
- 事象: $Omega$ の部分集合（Ω が連続濃度のときは注意が必要）.
- 事象には Ω と $\emptyset$(空事象) も含める.
- 事象すべてを集めた集合族を $\mathcal{A}$ と記す.

例 さいころを 1 回投げる試行.

- $\Omega = \{1, 2, 3, 4, 5, 6\}$.
- $\{1, 2, 3\}$ など.

1.1 確率 (2)

定義 1.1 Ω を空でない集合とし, $\mathcal{A}$ を Ω の部分集合族とする. $\mathcal{A}$ が次の 3 条件をみたすとき, σ 加法族と呼ばれる.

(1) $\Omega \in \mathcal{A}$.

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

(3) $A_n \in \mathcal{A}\,(n = 1, 2, \ldots) \Rightarrow \bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$.

ただし $A^c = \{\omega \in \Omega;\ \omega \notin A\}$ である. Ω と $\mathcal{A}$ の組 $(\Omega, \mathcal{A})$ を可測空間と呼ぶ

1.1 確率 (2)

補題 1.2 $(\Omega, \mathcal{A})$ を可測空間とする. このとき以下が成立する.

(1) $\emptyset \in \mathcal{A}$.

(2) $A_n \in \mathcal{A}\,(n = 1, 2, \ldots) \Rightarrow \bigcap_{n=1}^{\infty} A_n \in \mathcal{A}$.

証明: (1) $\emptyset = \Omega^c \in \mathcal{A}$.

(2) $A_n^c \in \mathcal{A}\,(n = 1, 2, \ldots)$ と De Morgan の法則から

$$\bigcap_{n=1}^{\infty} A_n = \bigcap_{n=1}^{\infty} (A_n^c)^c = \left(\bigcup_{n=1}^{\infty} A_n^c\right)^c \in \mathcal{A}.$$

最後は定義 1.1(3) を用いた. □

1.1 確率 (3)

注意 1.3 $\mathcal{C}$ を Ω の部分集合族とする. 集合族 $\mathcal{C}$ は σ 加法性をみたしてなくともよい. このとき, 集合族 $\sigma[\mathcal{C}]$ を

$$\sigma[\mathcal{C}] := \bigcap\{\mathcal{A};\ \mathcal{A} \supset \mathcal{C},\ \mathcal{A}\ \text{は}\ \sigma\ \text{加法族}\}$$

で定める. すると $\sigma[\mathcal{C}]$ は σ 加法族となることを確かめることができる. さらに $\mathcal{G}$ を $\mathcal{A}$ を含む σ 加法族としたとき

$$\sigma[\mathcal{C}] \subseteq \mathcal{G}$$

となることが直にわかる. すなわち $\sigma[\mathcal{C}]$ は $\mathcal{C}$ を含む最小 (包含関係の意味) の σ 加法族となる. □

1.1 確率 (4)

定義 1.4 $\Omega = \mathbb{R}$ とし

$$\mathcal{O} = \{O \subset \mathbb{R};\ O \text{ は } \mathbb{R} \text{ の開集合}\}$$

とする. $\sigma[\mathcal{O}]$ を $\mathbb{R}$ の **Borel** 集合族と呼び, $\mathcal{B}(\mathbb{R})$ と記す. また

$$\mathcal{C} = \{(-\infty, x) \subset \mathbb{R};\ x \in \mathbb{R}\}$$

とする. このとき $\mathcal{C}$ は $\mathcal{O}$ の真部分集合であるが, $\sigma[\mathcal{C}] = \sigma[\mathcal{O}]$ となる [1].

[1] $\sigma[\mathcal{C}] \subset \sigma[\mathcal{O}]$ は明らかであるが, 逆の包含関係も示すことができる. この逆の証明は, 定理 **1.16** と Euclid 位相の事実を用いて証明ができる.

1.1 確率 (5)

定義 1.6 $(\Omega, \mathcal{A})$ を可測空間とする. $\mathcal{A}$ 上の関数

$$\mathrm{Pr} : \mathcal{A} \ni A \mapsto \mathrm{Pr}(A) \in [0, 1]$$

が次の 2 条件をみたすとき, $(\Omega, \mathcal{A})$ 上の確率測度[2] と呼ばれる.

(1) $\mathrm{Pr}(\Omega) = 1$ である.

(2) 互いに排反[3] な事象列 $A_n \in \mathcal{A}\,(n = 1, 2, \ldots)$ に対して

$$\mathrm{Pr}\Bigl(\bigcup_{n=1}^{\infty} A_n\Bigr) = \sum_{n=1}^{\infty} \mathrm{Pr}(A_n) := \lim_{N\to\infty} \sum_{n=1}^{N} \mathrm{Pr}(A_n)$$

をみたす.

これらの 3 つの組 $(\Omega, \mathcal{A}, \mathrm{Pr})$ を確率空間という.

[2] 簡単に Ω 上の確率測度ともいう.

[3] $m \neq n$ ならば, $A_m \cap A_n = \emptyset$ が成立していること.

1.1 確率 (6)

補題 1.7 $(\Omega, \mathcal{A}, \mathrm{Pr})$ を確率空間とする. このとき以下が成立する.

(1) $\mathrm{Pr}(\emptyset) = 0$ である.

(2) $A \in \mathcal{A}$ に対して, $\mathrm{Pr}(A^c) = 1 - \mathrm{Pr}(A)$ となる.

(3) $N \in \mathbb{N}$ とする. $\{A_n\}_{n=1}^N \subset \mathcal{A}$ が互いに排反ならば

$$\mathrm{Pr}\left(\bigcup_{n=1}^{N} A_n\right) = \sum_{n=1}^{N} \mathrm{Pr}(A_n)$$

となる.

(4) $A, B \in \mathcal{A}, A \subset B \Rightarrow \mathrm{Pr}(B \setminus A) = \mathrm{Pr}(B) - \mathrm{Pr}(A)$ となる. よって, $A \subset B \Rightarrow \mathrm{Pr}(A) \leq \mathrm{Pr}(B)$ が成立する.

1.1 確率 (7)

補題 1.7 の続き

(5) $A_n \in \mathcal{A}$ が $A_n \subset A_{n+1}$ $(n = 1, 2, \ldots)$ をみたすならば

$$\lim_{n\to\infty} \Pr(A_n) = \Pr\left(\bigcup_{n=1}^{\infty} A_n\right)$$

となる.

(6) $A_n \in \mathcal{A}$ が $A_n \supset A_{n+1}$ $(n = 1, 2, \ldots)$ をみたすならば

$$\lim_{n\to\infty} \Pr(A_n) = \Pr\left(\bigcap_{n=1}^{\infty} A_n\right)$$

となる.

(7) (Boole の定理/ユニオン・バウンド) $A_n \in \mathcal{A}$ $(n = 1, 2, \ldots)$ に対して, $\Pr\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} \Pr(A_n)$ となる.

1.1 確率 (8)

(3) 証明: $F_i := A_i\,(i = 1, 2, \ldots, N)$ と $F_i = \emptyset\,(i \geq N+1)$ とおくと $\{F_n\}_{n=1}^{\infty}$ は互いに排反な事象列で $\bigcup_{n=1}^{\infty} F_n = \bigcup_{i=1}^{N} A_i$ となる. 定義 1.6(2) を用いると

$$\Pr\left(\bigcup_{i=1}^{N} A_i\right) = \Pr\left(\bigcup_{n=1}^{\infty} F_n\right) = \sum_{i=1}^{\infty} \Pr(F_i) = \sum_{i=1}^{N} \Pr(A_i) + \sum_{i=N+1}^{\infty} \Pr(\emptyset)$$
$$= \sum_{i=1}^{N} \Pr(A_i)$$

を得る. よって, (3) は示された. □

1.1 確率 (9)

定義 1.9 $(\Omega, \mathcal{A}, \mathrm{Pr})$ を確率空間とし, $A, B \in \mathcal{A}$ とする.

(1) (独立性):

$$A \text{ と } B \text{ は独立} \Leftrightarrow \mathrm{Pr}(A \cap B) = \mathrm{Pr}(A)\mathrm{Pr}(B).$$

(2) (条件付き確率) $\mathrm{Pr}(B) > 0$ のとき, B を与えたときの A の条件付き確率 $\mathrm{Pr}(A|B)$ を

$$\mathrm{Pr}(A|B) := \frac{\mathrm{Pr}(A \cap B)}{\mathrm{Pr}(B)}$$

で定める.

(3) $m \geq 3$ とし, $A_j \in \mathcal{A}\,(j = 1, 2, \ldots, m)$ とする. 事象 $A_1, A_2, \ldots, A_m$ は互いに独立であるとは, $2 \leq \forall k \leq m$ と $\forall 1 \leq j_1 < j_2 < \cdots < j_k \leq m$ に対して

$$\mathrm{Pr}(A_{j_1} \cap A_{j_2} \cap \cdots \cap A_{j_k}) = \mathrm{Pr}(A_{j_1})\mathrm{Pr}(A_{j_2}) \cdots \mathrm{Pr}(A_{j_k})$$

1.1 確率 (10)

注意 1.10 (1) $A, B \in \mathcal{A}$ とし, $\Pr(B) > 0$ とする. A と B が独立のとき

$$\Pr(A|B) = \Pr(A)$$

が成立する.
(2) $\Pr(B) > 0$ のとき

$$\Pr(A \cap B) = \Pr(B)\Pr(A|B)$$

となる. これを乗法の公式という.
(3) $\Pr(B) > 0$ のとき, 関数 $\Pr(\cdot|B) : \mathcal{A} \to [0, 1]$ は定義 1.6(1) – (2) をみたす. すなわち $(\Omega, \mathcal{A})$ 上の確率測度である.
(4) $\Omega = \{b_1, b_2, b_3, b_4\}$ とし

$$\Pr(\{b_j\}) = \frac{1}{4} \qquad (j = 1, 2, 3, 4)$$

とする

1.1 確率 (3)

注意 1.10(11) の続き　いま

$$A_j = \{b_j,\ b_4\} \qquad (j = 1,\ 2,\ 3)$$

とする. すると

$$A_1 \cap A_2 = A_1 \cap A_3 = A_2 \cap A_3 = A_1 \cap A_2 \cap A_3 = \{b_4\}$$

である. したがって

$$\Pr(A_1 \cap A_2) = \Pr(A_1)\Pr(A_2);\ \Pr(A_1 \cap A_3) = \Pr(A_1)\Pr(A_3);$$
$$\Pr(A_2 \cap A_3) = \Pr(A_2)\Pr(A_3)$$

である. しかし

$$\frac{1}{4} = \Pr(A_1 \cap A_2 \cap A_3) \neq \Pr(A_1)\Pr(A_2)\Pr(A_3) = \frac{1}{8}$$

である.

1.1 確率 (12)

補題 1.11(全確率の法則) $N \in \mathbb{N}$ とする. $A_1, A_2, \ldots, A_N \in \mathcal{A}$ は Ω の分割とする. すなわち, すべての $m \neq n\,(m, n \in \{1, 2, \ldots, N\})$ に対し, $A_m \cap A_n = \emptyset$ で, $\bigcup_{n=1}^{N} A_n = \Omega$ が成立している. このとき任意の $B \in \mathcal{A}$ に対して

$$\mathrm{Pr}(B) = \mathrm{Pr}\left(\bigcup_{n=1}^{N}(A_n \cap B)\right) = \sum_{n=1}^{N} \mathrm{Pr}(A_n \cap B)$$

となる.

1.1 確率 (13)

証明: $B = B \cap \Omega = B \cap \left(\bigcup_{n=1}^{N} A_n\right) = \bigcup_{n=1}^{N} (A_n \cap B)$ と $\{A_i \cap B\}_{i=1}^{n}$ は互いに排反であることに注意して, 定義 1.6(2) を用いればよい. □

定理 1.12(Bayes の定理) (1) $N \in \mathbb{N}$ とする. $A_1, A_2, \ldots, A_N, B \in \mathcal{A}$ とし, $A_1, A_2, \ldots, A_N$ は Ω の分割とする. $j = 1, 2, \ldots, N$ に対して, $\mathrm{Pr}(B) > 0$, $\mathrm{Pr}(A_j) > 0$ のとき

$$\mathrm{Pr}(A_j | B) = \frac{\mathrm{Pr}(A_j)\mathrm{Pr}(B | A_j)}{\sum_{n=1}^{N} \mathrm{Pr}(A_n)\mathrm{Pr}(B | A_n)}$$

が成立する.

(2) $A, B \in \mathcal{A}$ とし, $\mathrm{Pr}(A) > 0$, $\mathrm{Pr}(B) > 0$ とする. このとき, $\mathrm{Pr}(A | B)\mathrm{Pr}(B) = \mathrm{Pr}(B | A)\mathrm{Pr}(A)$ が成立する.

1.1 確率 (14)

□

例 1.13 病気 D に対する検査の結果を $+$ と $-$ とし, 確率が以下であるとする.

$$\Pr(+|\,D) = 0.9, \quad \Pr(-|\,D^c) = 0.9, \quad \Pr(D) = 0.01.$$

Bayes の定理を用いて, 検査で $+$ の人が本当に D である確率を求めると

$$\Pr(D|+) = \frac{\Pr(+\cap D)}{\Pr(+)} = \frac{\Pr(D)\Pr(+|\,D)}{\Pr(D)\Pr(+|\,D) + \Pr(D^c)\Pr(+|\,D^c)}$$
$$= \frac{0.01 \times 0.9}{0.01 \times 0.9 + (1-0.01) \times \{1-0.9\}} \approx 0.83$$

となる.

1.2 確率変数 (1)

定義 1.14 $(\Omega, \mathcal{A})$ と $(\mathbb{X}, \mathcal{B})$ を可測空間とする. 写像 $X : \Omega \to \mathbb{X}$ は $(\Omega, \mathcal{A})$ から $(\mathbb{X}, \mathcal{B})$ への可測写像であるとは

$$X^{-1}(B) := \{\omega \in \Omega : X(\omega) \in B\} \in \mathcal{A} \qquad (\forall B \in \mathcal{B})$$

をみたすときをいう.

注意 1.15 (1) $d \geq 2\,(d \in \mathbb{N})$ とする. $(\mathbb{X}, \mathcal{B}) = (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ のとき, X は確率空間 $(\Omega, \mathcal{A}, \mathrm{Pr})$ 上の確率ベクトルと呼ばれる. 定義されている確率空間に誤解がないときには, 簡単に確率ベクトルということもある.

(2) $d = 1$ のとき, X は確率空間 $(\Omega, \mathcal{A}, \mathrm{Pr})$ 上の 確率変数と呼ばれる. 定義されている確率空間に誤解がないときには, 簡単に確率変数と簡単にいうこともある. □

1.2 確率変数 (2)

定理 1.16 $(\Omega, \mathcal{A})$ と $(\mathbb{X}, \mathcal{B})$ を可測空間とし, $X : \Omega \to \mathbb{X}$ を写像とし, $\mathcal{C}$ を $\mathbb{X}$ の集合族とする. $\forall C \in \mathcal{C}$ に対して, $\{\omega \in \Omega;\ X(\omega) \in C\} \in \mathcal{A}$ であり, $\mathcal{C}$ が $\mathcal{B}$ を生成する [4] とき, X は可測となる.

証明: $B \in \mathcal{B}$ に対して, $\{\omega \in \Omega : X(\omega) \in B\}$ を $\{X \in B\}$ と書くことにする. $\{B\}, \{B_n\}_{n=1}^{\infty} \subset \mathcal{B}$ に対して

$$\left\{X \in \bigcup_{n=1}^{\infty} B_n\right\} = \bigcup_{n=1}^{\infty}\{X \in B_n\}, \quad \{X \in B^c\} = \{X \in B\}^c$$

となる. したがって, 集合族 $\mathcal{D} := \{B \in \mathcal{B};\ \{X \in B\} \in \mathcal{A}\}$ は σ 加法族 (必要ならば, $\mathbb{X}$ も加える) となる. よって, $\mathcal{C} \subset \mathcal{D}$ であり, $\mathcal{C}$ は $\mathcal{B}$ を生成するので, 最小性から $\mathcal{B} \subset \mathcal{D}$ となることから X は確率変数であることがわかる.

1.2 確率変数 (3)

注意 1.17 (1). $(\mathbb{X}, \mathcal{B}) = (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ に対して, 定理 1.16 における C の選択として, $\{(-\infty, r] : r \in \mathbb{R}\}$ と $\{(-\infty, q] : q \in \mathbb{Q}\}$ などがある.

(2). $d \geq 2\,(d \in \mathbb{N})$ とする. $(\mathbb{X}, \mathcal{B}) = (\mathbb{R}^d, \mathcal{B}(\mathbb{R}))$ に対して, 定理 1.16 における C の選択として

$$\{(a_1, b_1) \times (a_2, b_2) \times \cdots \times (a_d, b_d) : \\ -\infty < a_i < b_i < \infty\,(i = 1, 2, \ldots, d)\}$$

がある. □

1.2 確率変数 (4)

定理 1.18 $(\Omega, \mathcal{A}), (\mathbb{X}, \mathcal{B}), (\mathbb{Y}, \mathcal{C})$ を可測空間とする. $X : \Omega \to \mathbb{X}$ と $f : \mathbb{X} \to \mathbb{Y}$ は可測写像のとき, $f(X) : \Omega \to \mathbb{Y}$ は可測写像となる.

証明: 資料を参照せよ. □

定理 .19 $n \in \mathbb{N}$ とし, $X_1, X_2, \ldots, X_n$ は確率変数とし, $f : \mathbb{R}^n \to \mathbb{R}$ は可測とする. このとき, $f(X_1, X_2, \ldots, X_n)$ は確率変数となる.

証明: 定理 **1.16** から $(X_1, X_2, \ldots, X_n)$ は確率ベクトルであることを示せばよい. $A_1, A_2, \ldots, A_n \in \mathcal{B}(\mathbb{R})$ に対して

$$\{(X_1, X_2, \ldots, X_n) \in A_1 \times A_2 \times \cdots \times A_n\} = \bigcap_{i=1}^{n} \{X_i \in A_i\} \in \mathcal{A}$$

1.2 確率変数 (5)

定理 1.19 の証明の続き さらに, 集合族 $\{A_1 \times A_2 \times \cdots \times A_n;\ A_i \in \mathcal{B}(\mathbb{R})\ (i = 1, 2, \ldots, n)\}$ は $\mathcal{B}(\mathbb{R}^n)$ を生成[5] する. したがって, 定理 1.16 から $(\mathbb{X}, \mathcal{B}) = (\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ に対して, $(X_1, X_2, \ldots, X_n)$ は確率ベクトルであることがわかる. □

[5] この事実の証明には, 測度論の議論が必要なので, 信じることにしよう.

1.2 確率変数 (6)

定理 1.20 $X_1, X_2, \ldots, X_n$ は確率変数のとき, $X_1 + X_2 + \cdots + X_n$ も確率変数となる.

証明: 定理 **1.19** から $\mathbb{R}^n$ 上の実数値関数 $f(x_1, x_2, \ldots, x_n) = x_1 + x_2 + \cdots + x_n$ が可測であることを示せばよい. $\forall r \in \mathbb{R}$ に対して

$$\{(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n : x_1 + x_2 + \cdots + x_n < r\}$$

は $\mathbb{R}^n$ の開集合となるので

$$\{(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n : x_1 + x_2 + \cdots + x_n < r\} \in \mathcal{B}(\mathbb{R}^n)$$

となる. 集合族 $\{(-\infty, r) : r \in \mathbb{R}\}$ は $\mathcal{B}(\mathbb{R})$ を生成[6]するので, f は可測である. □

[6] この事実の証明には, 測度論の議論が必要となる. 信じることにする. 丁寧な証明は

1.2 確率変数 (7)

注意 1.21 同様の議論から, X_1, X_2 は確率変数のとき, X_1X_2 も確率変数であることがわかる. さらに, $X_2 \neq 0$ のとき, X_1/X_2 も確率変数であることもわかる.

定理 1.21 $X_1, X_2, \ldots$ は確率変数列のとき

$$\inf_n X_n \quad \sup_n X_n \quad \limsup_n X_n \quad \liminf_n X_n$$

も確率変数となる.

証明: $r \in \mathbb{R}$ に対して

$$\{\omega \in \Omega;\ \inf_n X_n(\omega) < r\} = \bigcup_{n=1}^{\infty} \{\omega \in \Omega;\ X_n(\omega) < r\} \in \mathcal{A}$$

から $\inf_n X_n$ は確率変数であることがわかる. 残りの証明は資料を参照せよ.

1.2 確率変数 (8)

注意 1.23 **(1).** $X_1, X_2, \ldots\ldots$ を確率変数列とする. 定理 **1.22** と定理 **1.19** から

$$\begin{aligned}\Omega_0 :=&\{\omega \in \Omega : \lim_{n\to\infty} X_n(\omega) \text{ は存在 }\}\\ =&\{\omega \in \Omega : \limsup_{n\to\infty} X_n(\omega) - \liminf_{n\to\infty} X_n(\omega) = 0\}\end{aligned}$$

は可測集合となる. $\Pr(\Omega_0) = 1$ のとき, 確率変数列 $\{X_n\}_{n=1}^{\infty}$ はほとんど確実に収束するといい, a.s. と記す. $\mathbb{R} \cup \{\pm\infty\}$ に値を取る確率変数を X_∞ とおくと

$$X_\infty(\omega) := \lim_{n\to\infty} X_n(\omega) \qquad \text{(a.s.) } \omega \in \Omega$$

と書ける.

1.2 確率変数 (9)

注意 1.23 の続き (2). 任意の $a, b \in \mathbb{R}\ (a < b)$ とし, $\{-\infty\} \cup (-\infty, a), (a, b), (b, \infty) \cup \{+\infty\}$ によって生成される σ 加法族を $\mathcal{B}(\mathbb{R} \cup \{\pm\infty\})$ と書く. 集合 $D \subset \Omega$ が定義域で, $\mathbb{R} \cup \{\pm\infty\}$ を終域ともつ関数 X は可測であるとは, $\forall B \in \mathcal{B}(\mathbb{R} \cup \{\pm\infty\})$ に対して

$$X^{-1}(B) = \{\omega \in \Omega : X(\omega) \in B\} \in \mathcal{A}$$

をみたすときをいう. □

1.2 確率変数 (10)

定義 1.24 (1) 確率空間 $(\Omega, \mathcal{A}, \mathsf{Pr})$ 上の確率変数 X に対して

$$\mathsf{F}^X(x) := \mathsf{Pr}(X \leq x) := \mathsf{Pr}(\{\omega \in \Omega;\ X(\omega) \leq x\}) \qquad (x \in \mathbb{R})$$

を X の累積分布関数 (cumulative distribution function(c.d.f.)) という. また

$$\mathsf{P}^X(B) := \mathsf{Pr}(X \in B) := \mathsf{Pr}(\{\omega \in \Omega;\ X(\omega) \in B\}) \qquad (B \in \mathcal{B}(\mathbb{R}))$$

を X の分布という. したがって, $\mathsf{P}^X((-\infty, x]) = \mathsf{F}^X(x)$ である.
(2) 確率変数 X, Y を $(\Omega, \mathcal{A}, \mathsf{Pr})$ の確率変数とし, それぞれの c.d.f. を $\mathsf{F}^X, \mathsf{F}^Y$ とする. このとき, $\mathsf{F}^X(x) = \mathsf{F}^Y(x)\ (\forall x \in \mathbb{R})$ が成立するとき, X と Y の分布は同じであるという. これを $X \overset{d}{=} Y$ と書く.
(3) 確率変数 X が c.d.f. F を持つとき, $X \sim \mathsf{F}$ と書く. □

1.2 確率変数 (11)

注意 1.25 と 1.26 (1) 文脈から, どの確率変数の確率分布または累積分布関数であることがわかる場合には, 簡単に P または F と書くこともある.
(2) P^X は可測空間 $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ 上の確率測度である. すなわち P^X は定義 **1.6** をみたすことがわかる.
(3) c.d.f. F が与えられると

$$\mathsf{F}(x) = \mathsf{P}((-\infty, x]) \qquad (x \in \mathbb{R})$$

をみたす $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ 上の確率測度 P が一意的に定まることが知られている. このことにより X の c.d.f. と分布を同一視する. さらに $X \sim \mathsf{F}$ とし, c.d.f. F から定まる確率測度を P としたとき, $X \sim \mathsf{P}$ とも書く.
(4) 確率変数 X と Y の分布が同じとき,
$\mathsf{P}^X(B) = \mathsf{P}^Y(B)\ (\forall B \in \mathcal{B}(\mathbb{R}))$ も成立する.

1.2 確率変数 (12)

定理 1.27 確率変数 X を確率空間 $(\Omega, \mathcal{A}, \mathsf{Pr})$ 上の確率変数とし, F を X の c.d.f. とする. このとき F は次をみたす.

(1) F は非減少関数; $x < y \Rightarrow \mathsf{F}(x) \leq \mathsf{F}(y)$.
(2) $\lim_{x\to\infty} \mathsf{F}(x) = 1$; $\lim_{x\to-\infty} \mathsf{F}(x) = 0$.
(3) F は右連続関数; $\lim_{y\to x+0} \mathsf{F}(y) = \mathsf{F}(x)$.
(4) F の不連続点は高々可算個.

証明: 確率変数 X の分布を P とする. すなわち
$\mathsf{P}(B) := \mathsf{Pr}(X \in B)\ (\forall B \in \mathcal{B}(\mathbb{R}))$ である.
(1) $(-\infty, x] \subset (-\infty, y]$ に注意して, 補題 1.7(4) を P に適用すればよい.
(2) $A_n = (-\infty, n]$ と $A_n = (-\infty, -n]$ として, P に補題 1.7(6)(7) を適用すればよい.
(3) $A_n = \left(-\infty, x + \dfrac{1}{n}\right]$ として, P に補題 ??(7) を適用すればよい.

1.2 確率変数 (13)

定理 1.28 確率変数 X を確率空間 $(\Omega, \mathcal{A}, \mathrm{Pr})$ 上の確率変数とし, X の c.d.f. を F とする. このとき次の (1) – (3) は同値である.

(1) F は $\mathbb{R}$ 上の連続関数.

(2) $\mathsf{F}(x) = \mathsf{F}(x-)$ $(\forall x \in \mathbb{R})$. ただし, $\mathsf{F}(x-) := \sup_y\{\mathsf{F}(y);\ y < x\}$ とした.

(3) $\mathrm{Pr}(X = x) = 0$ $(\forall x \in \mathbb{R})$.

証明: 資料を参照. □

1.2 確率変数 (14)

定義 1.30 確率変数 X を確率空間 $(\Omega, \mathcal{A}, \mathsf{Pr})$ 上の確率変数とし，その c.d.f. と分布を F と P と書く．
(1) X が高々可算個の集合 $\{x_1, x_2, \ldots\}$ 上にしか値を取らないとき，X を離散型であるという．この場合には

$$\mathsf{p}(x) := \mathsf{Pr}(X = x) = \mathsf{F}(x) - \mathsf{F}(x-) \qquad (x \in \{x_1, x_2, \ldots\})$$

で X の分布が特徴付けられる[7]．その p を X の確率関数 (probability mass function(p.m.f.)) と呼ぶ．

[7] $\mathsf{p}(x) = 0\,(x \notin \{x_1, x_2, \ldots\})$ となるので，p は $\mathbb{R}$ 上の関数であり，$0 \le \mathsf{p}(x) \le 1$ となる．

1.2 確率変数 (15)

定義 1.30 の続き (2) $\Pr(X = x) = 0 \ (\forall x \in \mathbb{R})$ のとき, X を連続型であるという. さらにある非負値関数 p で

$$\mathsf{F}(x) = \int_{-\infty}^{x} \mathsf{p}(t)\,\mathrm{d}t \qquad (\forall x \in \mathbb{R})$$

をみたすものが存在するとき, p を X の確率密度関数 (probability density function(p.d.f.)) という. 特に F がほとんどいたるところ[8]で微分可能ならば, ほとんどいたるところで

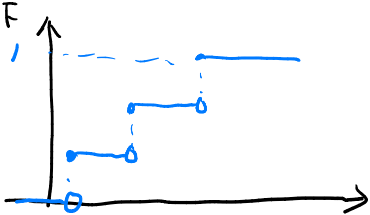
$$\dot{\mathsf{F}}(x) = \frac{\mathrm{d}\mathsf{F}}{\mathrm{d}x}(x) = \mathsf{p}(x)$$

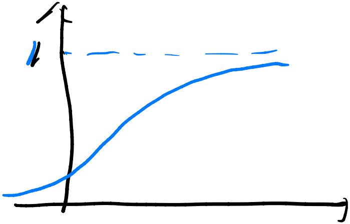
となる.

[8] この授業では, $\mathbb{R}$ から可算個の点を除いた集合上で微分可能と理解して差し支えない.

1.2 確率変数 (16)

例1 分布関数 F の例





$$p(x) := \Pr(X = x)$$
$$= F(x) - F(x-1).$$

$$p(x) = \Pr(X = x)$$
$$= F(x) - F(x-) = 0.$$

1.2 確率変数 (17)

注意 1.31 確率変数 X の分布を P とする.
(1) X を離散型とし, $S = \{x \in \mathbb{R};\ \mathsf{p}(x) > 0\}$ とおく[9]. すると任意の Borel 集合 $B \subset \mathbb{R}$ に対して

$$\mathsf{P}(B) = \mathsf{Pr}(X \in B) = \sum_{x \in B \cap S} \mathsf{p}(x)$$

が成り立つことを示すことができる.
(2) X を連続型とし, その p.d.f. p が定義されるとする. このとき, 任意の Borel 集合 $B \subset \mathbb{R}$ に対して

$$\mathsf{P}(B) = \mathsf{Pr}(X \in B) = \int_B \mathsf{p}(x)\,\mathrm{d}x$$

が成り立つことを示すことができる. また,

$$\mathsf{Pr}(B) = \int_B \mathsf{p}_1(x)\,\mathrm{d}x = \int_B \mathsf{p}_2(x)\,\mathrm{d}x \qquad (\forall B \in \mathcal{B}(\mathbb{R}))$$

1.2 確率変数 (18)

例 1.32 (1) $\Omega = \{0, 1\}$, $\mathcal{A} = 2^{\Omega}$, $\mathsf{Pr}(\{0\}) = \mathsf{Pr}(\{1\}) = 1/2$ とし

$$X : \Omega \ni \omega \mapsto X(\omega) = \omega \in \mathbb{R}$$

と定義すれば

$$\mathsf{Pr}(X = 0) = \mathsf{Pr}(X = 1) = \frac{1}{2}$$

となる. このとき, X の c.d.f. F と p.m.f. p はそれぞれ

$$\mathsf{F}(x) = \begin{cases} 0 & (x < 0) \\ \frac{1}{2} & (0 \leq x < 1) \\ 1 & (x \geq 1) \end{cases} \quad ; \qquad \mathsf{p}(x) = \begin{cases} \frac{1}{2} & (x = 0,\ 1) \\ 0 & (\text{その他の場合}) \end{cases}$$

となる.

1.2 確率変数 (19)

例 1.32 の続き (2)

$\Omega = [0, 1]$, $\mathcal{A} = \mathcal{B}(\mathbb{R}) \cap [0, 1] := \{B \cap [0, 1];\ B \in \mathcal{B}(\mathbb{R})\}$ とし

$$\Pr((a, b]) = b - a \qquad (0 \le a < b \le 1)$$

とする. さらに

$$X : \Omega \ni \omega \mapsto X(\omega) = \omega \in \mathbb{R}$$

と定義すれば, X の c.d.f. F と p.d.f. p はそれぞれ

$$\mathsf{F}(x) = \begin{cases} 0 & (x < 0) \\ x & (0 \le x \le 1) \\ 1 & (x > 1) \end{cases} \quad ; \quad \mathsf{p}(x) = \begin{cases} 1 & (0 \le x \le 1) \\ 0 & (\text{その他の場合}) \end{cases}$$

となる. □

1.3 分位点関数 (1)

定義 1.33 c.d.f. F に対して, 分位点関数 (quantile function) $\mathsf{F}^{-} : (0, 1) \to \mathbb{R}$ を

$$\mathsf{F}^{-}(y) := \inf\{x \in \mathbb{R} : \mathsf{F}(x) \geq y\} \qquad (0 < y < 1)$$

で定義する. また, $y \in (0, 1)$ に対して, $\mathsf{F}^{-1}(y)$ を F の y 分位点とよぶ. $1/2$ 分位点をメディアン (median) と呼ぶ.

注意 1.34, 1.36 (1) c.d.f. F が連続かつ $\operatorname{supp} \mathsf{F} = \{x \in \mathbb{R}; 0 < \mathsf{F}(x) < 1\}$ 上で狭義単調増加のとき, F^{-} は F の逆関数になる.

(2) F が連続 $\Rightarrow$ $\mathsf{F}(\mathsf{F}^{-}(y)) = y \ (0 < \forall y < 1)$ となる.

(3) F が不連続ならば, $\mathsf{F}(\mathsf{F}^{-}(y)) = y$ とは限らない. 反例は例 1.39.

□

1.3 分位点関数 (2)

例 1.39 $\mathrm{Pr}(X=0)=\mathrm{Pr}(X=1)=\dfrac{1}{2}$ とする. このとき, X の c.d.f. は

$$\mathsf{F}(x)=\begin{cases} 0 & (x<0) \\ \dfrac{1}{2} & (0\le x<1) \\ 1 & (x\ge 1) \end{cases}$$

となる. よって, F の分位点関数は

$$\mathsf{F}^{-}(y)=\begin{cases} 0 & \left(0<y\le \dfrac{1}{2}\right) \\ 1 & \left(\dfrac{1}{2}<y<1\right) \end{cases}$$

となる. この場合, $y\neq\dfrac{1}{2}$ のとき

1.3 分位点関数 (3)

定理 1.40 分位点関数は以下の性質 **(1)** ~ **(3)** をもつ.
(1) F^{-} は非減少である.
(2) F^{-} は左連続である. すなわち, $\{y_n\}_{n=1}^{\infty}$ を任意の非減少列で $\lim_{n\to\infty} y_n = y$ としたとき

$$\lim_{n\to\infty} \mathsf{F}^{-}(y_n) = \mathsf{F}^{-}(y)$$

が成立する.
(3) $\mathsf{F}^{-}(y) \leq x \Leftrightarrow y \leq \mathsf{F}(x)$ が成立する.
証明: 資料を参照. □

1.3 分位点関数 (4)

系 1.41 F を c.d.f. とする. このとき, $U \sim \mathrm{U}(0, 1)$ に対して

$$X := \mathsf{F}^{-}(U) \sim \mathsf{F}$$

となる.
証明: 定理 **1.40(3)** から, $\forall x \in \mathbb{R}$ に対して

$$\{X \le x\} \Leftrightarrow \{U \le \mathsf{F}(x)\}$$

となる. よって

$$\Pr(X \le x) = \Pr(U \le \mathsf{F}(x)) = \mathsf{F}(x)$$

を得る. □

1.4 主な 1 次元分布 (1)

Bernoulli 分布 $0 \leq \theta \leq 1$ とする. 確率変数 X は母数 θ の Bernoulli 分布に従うとは, X の p.m.f. $\mathrm{p}_\theta(\cdot)$ が

$$\mathrm{p}_\theta(x) = p(x) = \begin{cases} \theta^x(1-\theta)^{1-x} & (x = 0, 1) \\ 0 & (\text{その他の場合}) \end{cases}$$

のときをいう. このことを $X \sim \mathrm{Ber}(\theta)$ と記す.

注意 1.42 $\theta = 0$ のとき, $0^0 = 1$ と定めていることに注意せよ. □

1.4 主な 1 次元分布 (2)

2 項分布 $n \in \mathbb{N}$, $0 \leq \theta \leq 1$ とする. 確率変数 X は母数 (n, θ) の 2 項分布に従うとは, X の p.m.f. $\mathsf{p}_\theta(\cdot)$ が

$$\mathsf{p}_{\theta,n}(x) = \mathsf{p}(x) = \begin{cases} \begin{pmatrix} n \\ x \end{pmatrix} \theta^x (1-\theta)^{n-x} & (x = 0, 1, \ldots, n) \\ 0 & (\text{その他の場合}) \end{cases}$$

のときをいう. ただし

$$\begin{pmatrix} n \\ x \end{pmatrix} = \frac{n!}{x!\,(n-x)!}, \qquad 0! = 1$$

である. このことを $X \sim \mathsf{Bino}(n, \theta)$ と記す.

1.4 主な 1 次元分布 (3)

$\boxed{\text{Poisson 分布}}$ $\theta > 0$ とする. 確率変数 X は母数 θ の Poisson 分布に従うとは, X の p.m.f. $\mathrm{p}_\theta(\cdot)$ が

$$\mathrm{p}_\theta(x) = \mathrm{p}(x) = \begin{cases} e^{-\theta}\dfrac{\theta^x}{x!} & (x = 0,\ 1,\ 2,\ \ldots) \\ 0 & (\text{その他の場合}) \end{cases}$$

のときをいう. このことを $X \sim \mathrm{Po}(\theta)$ と記す. □

1.4 主な 1 次元分布 (4)

正規分布 $\mu \in \mathbb{R}, 0 < \sigma < \infty$ とする. 確率変数 X は平均 μ, 分散 σ^2 の正規分布[10] に従うとは, X の p.d.f. $\mathsf{p}_{\mu,\sigma^2}(\cdot)$ が

$$\mathsf{p}_{\mu,\sigma^2}(x) = \mathsf{p}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \qquad (-\infty < x < \infty)$$

のときをいう. ただし $\exp(x) = e^x$ である. このことを $X \sim \mathrm{N}(\mu, \sigma^2)$ と記す. $\mu = 0, \sigma = 1$ のときの分布を標準正規分布という.

[10] 母数 (μ, σ^2) を平均と分散となぜ呼ぶかは第 ?? 章で判明する.

1.4 主な 1 次元分布 (5)

注意 1.44 よく知られている事実として

$$\int_0^\infty e^{-t^2}\,\mathrm{d}t = \sqrt{\pi}$$

がある. $\mathbb{R} \ni t \mapsto e^{-t^2} \in (0, \infty)$ は偶関数であることに注意して, $t = z/\sqrt{2}$ と変換すると

$$\int_{-\infty}^\infty e^{-z^2/2}\,\mathrm{d}z = \sqrt{2\pi}$$

となることがわかる. さらに, $z = (x-\mu)/\sigma$ と変数変換をすると

$$\int_{-\infty}^\infty \mathsf{p}_{\mu,\sigma^2}(x)\,\mathrm{d}x = 1$$

となることが確認できる. □

1.4 主な 1 次元分布 (6)

$\boxed{\text{ガンマ分布}}$ $\alpha > 0, \beta > 0$ とする. 確率変数 X は母数 (α, β) のガンマ分布に従うとは, X の p.d.f. $\mathsf{p}_{\alpha,\beta}(\cdot)$ が

$$\mathsf{p}_{\alpha,\beta}(x) = \mathsf{p}(x) = \begin{cases} \dfrac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta} & (x > 0) \\ 0 & (\text{その他の場合}) \end{cases}$$

のときをいう. ただし, $\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x}\, \mathrm{d}x$ $(\alpha > 0)$ である. このことを $X \sim \mathsf{Ga}(\alpha, \beta)$ と記す. $\lambda > 0$ とし, $\alpha = 1, \beta = \dfrac{1}{\lambda}$ とおく. このとき, $\mathsf{Ga}(1, 1/\lambda)$ を母数 λ の指数分布といい, $\mathsf{Exp}(\lambda)$ と書く. また, $p \in \mathbb{N}$ とし, $\alpha = \dfrac{p}{2}, \beta = 2$ とおく. このとき, $\mathsf{Ga}(p/2, 2)$ を自由度 p の χ^2 分布といい, χ^2_p と記す. □

本日の講義のまとめ

- 確率空間, 確率を定義し, 基本的な計算（補題 1.7）を照会した.
- 確率変数を定義した.
- 確率変数の分布と累積分布関数を導入した.
- 累積分布関数の基本的な性質を説明した.
- 離散型確率変数と連続型確率変数を導入した.
- 分位点関数を定義した.
- 代表的な 1 次元分布を紹介した.

□