確率・統計入門

森 立平

2025-03-28

目次

前書き

これは確率論と統計学の入門書である。確率論を数学的に取り扱うには通常は測度論とルベーグ積分を用いる。本書では測度論を学ぶ前の数学専攻の学生を対象に確率論と統計学の基礎を解説する。測度論とルベーグ積分を省略するため、しばしば積分と極限の交換などの等式を証明なしに用いる。後で測度論を学んだ後にぜひ振り返って欲しい。

第 I 部

確率論

第 1 章

はじめに

1.1　なぜ確率論と統計学を学ぶか？

This is a book created from markdown and executable code.

See Knuth (1984) for additional discussion of literate programming.

1.2　本書の構成

1.3　その他の参考文献

第 2 章

集合論

第 3 章

確率空間

3.1 確率論を数学的に定式化するには

確率は身近に現れる (感じられる) ものであるが、それを数学的に定式化することは自明な問題ではない。実際に確率論には複数の数学的定式化が存在する。その中で圧倒的に一般的なのが**測度論的確率論**と呼ばれる定式化である。測度というのは集合の「面積」のようなものであり、確率を測度と捉えるのが測度論的確率論である。これは多くの人間の直感にも自然なものであろう。

まず、確率を考える集合について考えよう。例えばコインを投げて表もしくは裏が出る確率を考えたいときは

$$\Omega = \{\mathrm{H}, \mathrm{T}\}$$

という集合になる。明日の天気の確率を考えたいときは

$$\Omega = \{\text{"晴れ"}, \text{"雨"}, \text{"曇り"}\}$$

という集合になる。この Ω の部分集合に確率を与える関数 $P\colon 2^\Omega \to \mathbb{R}$ を定義しよう。偏りのないコインの場合は以下のようになる。

$$P(\emptyset) = 0 \qquad P(\{\mathrm{H}\}) = \frac{1}{2} \qquad P(\{\mathrm{T}\}) = \frac{1}{2} \qquad P(\{\mathrm{H}, \mathrm{T}\}) = 1$$

天気の場合は例えば以下のようになる。

$$\begin{aligned} P(\emptyset) &= 0 & P(\{晴\}) &= 0.7 & P(\{雨\}) &= 0.2 & P(\{雪\}) &= 0.1 \\ P(\{晴, 雨\}) &= 0.9 & P(\{雨, 雪\}) &= 0.3 & P(\{雪, 晴\}) &= 0.8 \\ P(\{晴, 雨, 雪\}) &= 1. \end{aligned}$$

このように集合 Ω の部分集合に確率を与えることを考える。Ω の要素一つに確率を与えれば十分であるようにも思えるが、Ω が連続的な場合には Ω の一つの要素の確率は 0 になってしまうことが多い。例えば明日の正午の気温が $10\pi°C$ になる確率は 0 であろう。このことから、Ω の要素ではなく部分集合に確率を与えると考える。そのために測度論が適しているわけである。

3.2 有限集合上の確率空間

確率を考える集合を Ω とする。例えば明日の天気の確率を考えたいときは

$$\Omega = \{\text{"晴れ"}, \text{"雨"}, \text{"曇り"}\}$$

とする。この Ω のことを**標本空間**という。また、Ω の部分集合のことを**事象**という。そして、事象に確率を与える関数 $P\colon 2^\Omega \to \mathbb{R}_{\geq 0}$ を**確率測度**という。確率測度は以下の条件を満たす。

1. $P(\Omega) = 1.$
2. $\forall A, B \subseteq \Omega, \quad A \cap B = \emptyset \implies P(A \cup B) = P(A) + P(B).$

一つ目の条件は全体の確率が 1 であるという意味の条件である。二つ目の条件は排反な二つの事象の確率はそれぞれの事象の確率の和であるという意味の条件である。この二つ目の条件を**有限加法性**という。例えば

$$P(\{\text{"晴れ"}, \text{"雨"}\}) = P(\{\text{"晴れ"}\}) + P(\{\text{"雨"}\})$$

という等式は「晴れもしくは雨になる確率 $=$ 晴れになる確率 $+$ 雨になる確率」という意味の等式になる。よって有限加法性が自然な条件であることが分かるだろう。また、これらのことから、P は各要素 $\omega \in \Omega$ に対する確率 $P(\{\omega\})$ から一意に定まることが分かる。この標本空間と確率測度のペア (Ω, P) を**確率空間**という。

3.3 可算無限集合上の確率

標本空間 Ω が可算無限集合のときも、同様に確率測度を定義することもできるが、ここではより強い以下の条件を考える。

1. $P(\Omega) = 1.$
2. $\forall (A_n \subseteq \Omega)_{n\geq 0}, \quad \forall i \neq j, A_i \cap A_j = \emptyset \implies P\left(\bigcup_{n\geq 0} A_n\right) = \sum_{n\geq 0} P(A_n).$

この二つ目の条件を**完全加法性**もしくは **σ-加法性**という。ここで

$$\bigcup_{n\geq 0} A_n = \{\omega \in \Omega \mid \exists n \geq 0, \omega \in A_n\}$$

は事象列 $(A_n)_{n\geq 0}$ の並び順に依存しない。よって、P の完全加法性が成り立つとき

$$\sum_{n\geq 0} P(A_n)$$

も事象列 $(A_n)_{n\geq 0}$ の並び順に依存しない。そもそも $P(A_n) \geq 0$ なので、無限和 $\sum_{n\geq 0} P(A_n)$ が存在するときこれは絶対収束することから、この無限和は事象列の並び順に依存しないことが分かる。完全加法性ではなく有限加法性だけを使って確率論を構築する試みもあるが、確率測度の連続性などの性質が失なわれるため標準的な確率論では完全加法性を課す。

3.4 すべての部分集合を可測にはできない

より一般に Ω が非可算無限集合の場合を考えよう。例えば $\Omega = [0,1)$ の場合が考えられる。このとき、すべての $A \subseteq \Omega$ に確率を与えることができるのだろうか？ ごく自然な性質を持つような確率測度が存在しないことが、選択公理を認めると示される。

定理 3.1 (非可測集合の存在)**.** $\Omega = [0,1)$ とする。また、集合の平行移動を

$$A + x \coloneqq \{a + x - \lfloor a + x \rfloor \mid a \in A\}$$

と定義する。このとき、

$$\forall x \in \Omega,\ A \subseteq \Omega, \quad P(A + x) = P(A)$$

を満たす確率測度 $P\colon 2^{[0,1)} \to \mathbb{R}_{\geq 0}$ は存在しない。

証明. Ω 上の同値関係を $x \sim y \overset{\text{def}}{\iff} x - y \in \mathbb{Q}$ と定義する。選択公理より、この同値関係の同値類から一つずつ要素を含む集合 V が存在する (Vitali 集合)。任意の $x \in [0,1)$ について、ある $z \in V$ が**唯一**存在して $x \sim z$ である。よって、任意の $x \in [0,1)$ について、ある $z \in V$ と $q \in \mathbb{Q}$ が**唯一**存在して $x = z + q$ であることから

$$[0,1) = \bigcup_{q \in \mathbb{Q}} (V + q)$$

であり、右辺は互いに排反である。よって条件を満たす P が存在すると仮定すると、

$$\begin{aligned} 1 = P([0,1)) &= P\left(\bigcup_{q \in \mathbb{Q}} (V + q)\right) \\ &= \sum_{q \in \mathbb{Q}} P(V + q) \qquad \text{(完全加法性)} \\ &= \sum_{q \in \mathbb{Q}} P(V) \qquad \text{(平行移動不変性)} \end{aligned}$$

ここで $P(V)$ をどのように定めても、それを無限回足して 1 にすることはできない。よって P は存在しない。 □

3.5 確率論の公理

定理 3.1 より、

1. 選択公理
2. 確率測度の完全加法性
3. $[0,1)$ 上の平行移動不変な確率測度
4. Ω のすべての部分集合に確率を与える

のどれかを諦めないといけない。標準的な確率論では 4 を諦める。以下に確率空間の厳密な定義を述べる。

定義 3.1 (完全加法族)**.** Ω を集合とする。$\mathcal{F} \subseteq 2^{\Omega}$ が以下を満たすとする。

1. $\Omega \in \mathcal{F}$.
2. $\forall A \in \mathcal{F}, \quad \Omega \setminus A \in \mathcal{F}$.
3. $\forall (A_n \in \mathcal{F})_{n \geq 0}, \quad \bigcup_{n \geq 0} A_n \in \mathcal{F}$.

このとき、$\mathcal{F}$ を Ω 上の完全加法族もしくは σ-加法族という。

定義 3.2 (確率空間)**.** Ω を集合とし、$\mathcal{F} \subseteq 2^{\Omega}$ を Ω 上の完全加法族とする。

また、$P\colon \mathcal{F} \to \mathbb{R}_{\geq 0}$ が以下を満たすとする。

1. $P(\Omega) = 1$.
2. $\forall (A_n \in \mathcal{F})_{n \geq 0}, \quad \forall i \neq j,\, A_i \cap A_j = \emptyset \implies P\left(\bigcup_{n \geq 0} A_n\right) = \sum_{n \geq 0} P(A_n)$.

このとき、$(\Omega, \mathcal{F}, P)$ を確率空間という。また、Ω を標本空間、$\mathcal{F}$ を事象集合、P を確率測度という。

このように事象集合 $\mathcal{F}$ の元についてのみ確率が与えられる。定理 3.1 の証明では選択公理を用いたが、実際に選択公理を認めないと定理 3.1 は成立しない。

以降、Ω が非可算無限集合である場合も含めて、$\mathcal{F} = 2^{\Omega}$ だと思うことにする。これは正しくない場合もあるのだが、選択公理を使わない限り矛盾は導かれないので、問題になることはほとんどない。

よって、以降は (Ω, P) を確率空間とする。

第 4 章

確率変数と確率分布

第 5 章

期待値、分散、モーメント

第 6 章

大数の法則と集中不等式

第 7 章

正規分布と中心極限定理

第 8 章

複数の確率変数、多変量正規分布

第 9 章

サノフの定理、KL ダイバージェンス

第 II 部

統計学

第 10 章

ベイズ推定

第 11 章

仮説検定

参考文献

Knuth, Donald E. 1984. "Literate Programming." *Comput. J.* 27 (2): 97–111. https://doi.org/10.1093/comjnl/27.2.97.