

# 確率論のノート

hupom

2026年2月9日

## 概要

確率論および統計の知識などを勉強しながら書いていきます。当然に間違いが含まれます。  
また二段組のスタイルを検討しながら書くことも合わせて行うため、バージョンによっては組版がぐちゃぐ  
ちゃの可能性もあります。

## 1 はじめに：確率とは何か

### 1.1 確率とモデル

確率論とは「偶然によって結果が決まる不確実な現象を記述するための手法」(楠岡)です。ここで問題になるのは、数式によって現象を記述するのはあくまで数学モデルでしかない、ということです。通常、現実の現象とその数学モデルを区別することは少ないので、こと確率論においては、数学モデルが現実の現象をよく記述しているかというのは判定が難しい問題になります。

確率論に基づく数学モデルを「確率モデル」と呼びます。例を用いて考えましょう。例えば「サイコロの1の目の出る確率は $\frac{1}{6}$ である」という確率モデルを立て、実際にサイコロを振ると2の目が出ました。このとき、確率モデルが「良い」モデルだったのか、それとも「悪い」モデルだったのか、これだけでは判断できません。確率モデルの検証のためには「確率モデルを前提とすると一体何が言えるのか」を理論的に考える必要があり、それがまさに「確率論」です。

確率モデルの含むパラメータを決定したり、モデルそのものの信頼性を調べる「確率モデルの検証」を行うのが「(数理)統計学」になります。

確率論の知識がなくては統計学はできませんし、統計学の知識がなくては確率論は現実の役に立ちません。本稿では可能な限り確率論と統計学の両軸で知識をまとめていくことを意識しています。

### 1.2 確率とは何か

「確率とは何か」というのは今なお哲学的な議論がありますが、数学はこの問い合わせに答えるのは避け、確率の満たすべき公理を定め、それを満たすものはなんであれ「確率」である、というKolmogorovの導入した公

理主義的立場を取ります。

厳密に公理主義の立場に立つなら無定義用語は説明しないのですが、イメージのために説明をつけます。雑に述べると、まず「サイコロを1回投げると1の目が出た」のような事象(event)を“すべて”集めたデカい集合 $\Omega$ を取ります。次に $\Omega$ の中でいい感じの事象を集めた集合 $\mathcal{F} \subset \Omega$ を考えます。最後に $\mathcal{F}$ の各元 $A \in \mathcal{F}$ にその「確率」 $P(A)$ を対応させるいい感じの関数 $P: \mathcal{F} \rightarrow \mathbb{R}$ を考え、この三つ組 $(\Omega, \mathcal{F}, P)$ を「確率空間」と呼びます。

上に現れる $P$ が「確率」に相当するのですが、要は三つ組 $(\Omega, \mathcal{F}, P)$ であってKolmogorovの公理系を満たすものは全部「確率」を考えられる確率空間であるとしてしまうのが公理主義的な立場です。

以上の話は $\Omega$ が有限集合であるような場合(例えばサイコロ振りなど)では直観によく整合します。が、 $\Omega$ が無限集合の場合には、素朴に確率を考えることはもう難しくなります。確率の現代的な定式化はそうした場合も含む理論を構築するために、次章以降に述べるようなやや複雑な理路をたどります。

## 2 確率論

### 2.1 確率の定義

「いい感じの $\mathcal{F}$ を定式化するのが $\sigma$ -加法族になります。

**定義 2.1.** 空でない集合 $\Omega$ に対し、その部分集合族 $\mathcal{F} \subset 2^\Omega$ が次を満たすとき、 $\sigma$ -加法族という。

1.  $\emptyset \in \mathcal{F}$ .
2.  $A \in \mathcal{F}$ ならば $A^c \in \mathcal{F}$ .
3.  $A_1, A_2, \dots \in \mathcal{F}$ ならば $\bigcup_{i \in \mathbb{N}} A_i \in \mathcal{F}$ .

組 $(\Omega, \mathcal{F})$ を「可測空間」といい、 $\mathcal{F}$ の元を $\mathcal{F}$ -可測集

合, あるいは  $\mathcal{F}$ -可測な集合という.

3つ目の条件は  $A_1 \cup A_2 = A_1 \cup A_2 \cup \emptyset \cup \emptyset \cup \dots$  のように考えることで, 有限個の和集合についても閉じていることに注意する. 要は  $\sigma$ -加法族とは,  $\Omega$  のうち補集合や(可算個の)和集合をとる操作ができる集合ということになる.

これがなぜ良いのか? 集合の補集合は「事象の否定」に対応し, 集合の和は「事象の和」に対応します. つまり  $\sigma$ -加法族は「確率」を考えたい集合の集まりとして満たしてほしい感じの性質を定式化したものになっています.

この  $\mathcal{F}$  の上でいい感じの写像を考えます.

**定義 2.2.** 測度空間  $(\Omega, \mathcal{F})$  に対し, 写像  $P: \mathcal{F} \rightarrow [0, 1]$  が次の2条件を満たすとき,  $(\Omega, \mathcal{F})$  上の確率測度あるいは単に確率といいう:

1.  $P(\Omega) = 1$ .
2. 互いに素な集合  $A_1, A_2, \dots \in \mathcal{F}$  に対し

$$P\left(\bigcup_{i \in \mathbb{N}} A_i\right) = \sum_{i=1}^{\infty} P(A_i).$$

三つ組  $(\Omega, \mathcal{F}, P)$  を確率空間と呼ぶ. また  $A \in \mathcal{F}$  を事象と呼び,  $P(A)$  を事象  $A$  の確率といいう.

2番目の条件を可算加法性といいます. 以上が確率の現代数学における抽象的な定義です.

**定義 2.3.** 二つの事象  $A, B$  について

$$P(A \cap B) = P(A)P(B)$$

がなりたつとき,  $A$  と  $B$  は独立であるといいう.

## 2.2 確率変数

まずは定義を与えます.

**定義 2.4.** 確率空間  $(\Omega, \mathcal{F}, P)$  と測度空間  $(S, \mathcal{G})$  に対し, 写像

$$X: \Omega \rightarrow S$$

であって, 任意の  $G \in \mathcal{G}$  に対し

$$X^{-1}(G) = \{\omega \in \Omega \mid X(\omega) \in G\} \in \mathcal{F}$$

がなりたつものを, ( $S$ -値の) 確率変数といいう.

確率変数には「確率空間に具体的な問題を設定する」という役割があります. 例えば, 確率変数はギャンブルのようなものだと考えると分かりやすいです.

**例 2.5.** サイコロを1回投げて, 奇数なら一万円もら

い, 偶数なら一万円払うギャンブルを考える. これは

$$X(\omega) = \begin{cases} +10000 & (\omega = 1, 3, 5) \\ -10000 & (\omega = 2, 4, 6) \end{cases}$$

という確率変数と考えられる.

なお実数値の確率変数における  $(S, \mathcal{G})$  とは何か? というのが気になるかもしれません. 技術的には  $S = \mathbb{R}$  のときは  $\mathcal{G}$  として  $\mathbb{R}$  のすべての開区間を含む最小の  $\sigma$ -集合族 (Borel 集合族といいう) を設定するのですが, まあ「実数値で測れるのだな」と考えてもらって問題ありません.

## 2.3 確率分布と確率密度関数

以下では  $(\Omega, \mathcal{F}, P)$  上の  $\mathbb{R}$ -値の確率変数  $X: \Omega \rightarrow \mathbb{R}$  を考えます. ギャンブル(確率変数)そのものをいちいち考えるより, 代わりにそのギャンブルで得られる損得とその確率のみに注目する方が合理的です. この考え方を定式化したものが確率分布です.

$p_X(a) = P(X^{-1}(\{a\})) = P(\{\omega \in \Omega \mid X(\omega) = a\})$  とおく. また  $X$  の値の集合  $\text{Im } X$  を

$$\text{Im } X = \{a \in \mathbb{R} \mid p_X(a) > 0\}$$

と定める.

**定義 2.6 (離散確率変数の確率分布).** 値の集合が  $\text{Im } X = \{a_1, a_2, \dots\}$  という離散集合のとき,  $X$  を離散確率変数といい,

$$p_i = p_X(a_i)$$

によって与えられる  $(p_i)_{i=1}^{\infty}$  を  $X$  の確率分布といいう.

$X$  が一般の実数値を取り得る場合は, 上述のような「一点」の確率が決まつても確率測度全体は定まらない. そこで次の「確率分布関数」という道具を用いる.

**定義 2.7.**  $(\Omega, \mathcal{F}, P)$  上の実数値確率変数  $X$  に対し

$$\begin{aligned} F_X(x) &:= P(X^{-1}([-\infty, x])) \\ &= P(\{\omega \mid X(\omega) \leq x\}) \end{aligned}$$