確率統計の整理帳

tomixy

2025 年 7 月 25 日

目次

第 1 章

データの代表値

「データを 1 つの値で要約するならばこれ」といった指標を代表値という。

平均値

平均（`mean`）は、データをすべて足し合わせて、データの数で割ることで求まる。

> **平均**　N 個の観測値 $x_1, \ldots, x_N$ の総和をデータのサイズ N で割ったものを平均値という。
>
> $$\overline{x} := \frac{1}{N}\sum_{i=1}^{N} x_i$$

中央値

データを大きさの順に並べたときにちょうど中央に位置する値を中央値（`median`）という。

中央値　N 個の観測値 $x_1, \ldots, x_N$ を大きさの順に並べ替えたものを

$$x_{(1)} \leq \cdots \leq x_{(N)}$$

とするとき、中央値は次のように定義される。

$$\tilde{x} := \begin{cases} x_{(k+1)} & (N = 2k+1) \\ \dfrac{x_{(k)} + x_{(k+1)}}{2} & (N = 2k) \end{cases}$$

第 2 章

確率と集合

標本空間

実験や観測を行うときに起こりうるすべての結果を集合に閉じ込める

ref: スッキリわかる確率統計 p65

標本空間　起こりうるすべての結果からなる集合を標本空間といい、Ω や U などと表す

根元事象　標本空間 Ω の元、すなわち、標本空間の中の各結果を根元事象という

事象

実験や観測の結果、実際に起こったことを事象という

ref: スッキリわかる確率統計 p65

事象　標本空間の部分空間 Ω の部分集合を事象といい、A や B などと表す

起こりうる場合の集合（標本空間）から、実際に起こったものだけを取り出したものが事象であるので、事象は標本空間の部分集合といえる

空事象と全事象

ref: スッキリわかる確率統計 p65

空事象

決して起こらないことを事象の一つとみなしたものを空事象という

空事象　空集合である事象を空事象といい、$\emptyset$ と表す

事象とは、実際に起こったことだけを取り出した集合なので、決して起こらないことは空集合となる

全事象

一方、起こりうるすべての場合が起こるとき、標本空間そのものを事象とみなすことができる

全事象　起こりうるすべての場合からなる事象を全事象という

集合としてみれば、全事象は標本空間そのものなので、標本空間と同じく Ω や U で表す

事象と集合演算

ref: スッキリわかる確率統計 p65〜67

和事象と積事象

「A または B が起こる」という事象を**和事象**という

> **和事象**　事象 A と事象 B の和集合を $A \cup B$ と表し、事象 A と事象 B の**和事象**という

一方、「A と B が同時に起こる」という事象を**積事象**という

> **積事象**　事象 A と事象 B の共通部分を $A \cap B$ と表し、事象 A と事象 B の**積事象**という

余事象

「A が起こらない」という事象を**余事象**という

> **余事象**　標本空間の中で、A が起こらないという事象を A の**余事象**といい、A^c で表す

標本空間を全体集合とすると、余事象は補集合に対応する

> 余事象との和事象・積事象
>
> $$A \cup A^c = \Omega$$
> $$A \cap A^c = \emptyset$$

余事象の余事象

$$(A^c)^c = A$$

事象の差

「A は起こるが B は起こらない」という事象は、「A と B^c が同時に起こる」事象として表せる

事象の差　次の事象を、事象 A と B の差という

$$A - B = A \cap B^c$$

排反

事象 A と事象 B が同時に起こることがないとき、A と B は互いに排反であるという

ref: スッキリわかる確率統計 p67

排反　$A \cap B = \emptyset$ であるとき、事象 A と事象 B は排反、あるいは互いに素であるという

同様に確からしい

頻度的立場の確率は、次の等確率性（同様に確からしい）を前提としている

ref: スッキリわかる確率統計 p67

同様に確からしい　標本空間に属するどの根元事象も同じ程度に起こると期待されるとき、これらの事象は**同様に確からしい**という

以降、特に断りがなければ「同様に確からしい」事象を扱うことにする

第 3 章

基本的な確率計算

割合としての確率

確率を最も馴染みのある考え方でとらえると、確率とはある事象が起こる可能性であり、

ref: 数学図鑑 p28〜29

> 起こりうるすべての場合のうち、ある事象が起こる場合の割合

として計算できる

たとえば、起こりうるすべての場合が $n(U)$ 通り、事象 A が起こる場合が $n(A)$ 通りあるとすると、

$$P(A) = \frac{n(A)}{n(U)}$$

が、事象 A が起こる確率となる（ここで、U は標本空間である）

ここで注意が必要なのは、割り算は全体を「均等に」分けることを前提とした演算であることだ

標本空間に含まれるすべての場合の数で割ったものを確率とみなすには、どの事象も同程度に起こりうる（同様に確からしい）という仮定が必要になる

和事象の確率

A または B が起こる場合が $n(A \cup B)$ 通りあるとすると、その確率は、次の割合で表される

ref: 数学図鑑 p30〜31

$$P(A \cup B) = \frac{n(A \cup B)}{n(U)}$$

ここで、$n(A \cup B)$ は、$n(A)$ と $n(B)$ の和から、A と B の重なっている部分 $n(A \cap B)$ を引いたものとなる

$$n(A \cup B) = n(A) + n(B) - n(A \cap B)$$

よって、和事象 $A \cup B$ の確率は、

$$\begin{aligned} P(A \cup B) &= \frac{n(A \cup B)}{n(U)} \\ &= \frac{n(A) + n(B) - n(A \cap B)}{n(U)} \\ &= \frac{n(A)}{n(U)} + \frac{n(B)}{n(U)} - \frac{n(A \cap B)}{n(U)} \\ &= P(A) + P(B) - P(A \cap B) \end{aligned}$$

として求められる

⚓ **和事象の確率**　和事象 $A \cup B$ の確率は、

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

排反な事象と確率の加法定理

和事象の確率において、事象 A と B が互いに排反である（つまり、同時に起こることはない）ならば、$n(A \cap B) = 0$ となるので、

ref: 数学図鑑 p32

$$P(A \cup B) = P(A) + P(B)$$

が成り立つ

確率の加法定理　互いに排反な事象 A, B の和事象 $A \cup B$ の確率は、

$$P(A \cup B) = P(A) + P(B)$$

余事象の確率

[Todo 1:]

ref: 数学図鑑 p34

条件つき確率

事象 A が起こったときに事象 B が起こる確率を条件つき確率という

ref: 数学図鑑 p38〜39

条件つき確率　事象 A が起こったときに事象 B が起こる確率を $P(B|A)$ あるいは $P_A(B)$ と表し、これを A が起こったときの B が起こる条件つき確率という

条件つき確率では、標本空間「全体」ではなく、その一部分である「A が起きた場合」に限定して考える

その中で B も起こる割合だから、「A かつ B」の確率を「A」の中での割合でみればよい

$$P(B|A) = \frac{P(A \cap B)}{P(A)}$$

⚓ **条件つき確率**　事象 A が起こったときの事象 B が起こる条件つき確率は、

$$P(B|A) = \frac{P(A \cap B)}{P(A)}$$

積事象の確率と確率の乗法定理

A かつ B が起こる場合が $n(A \cap B)$ 通りあるとすると、その確率は、次のような割合で表される

ref: 数学図鑑 p30〜31、38〜39

$$P(A \cap B) = \frac{n(A \cap B)}{n(U)}$$

一方、積事象の確率は、次のように分けて考えることもできる

1. 全体のうち A が起こる（この確率は $P(A)$）
2. A が起こったとき、B が起こる（この確率は $P(B|A)$）

全体のうち A が起こる場合の数を $n(A)$、A が起こった場合のうち B が起こる場合の数を $n(B|A)$ とすると、場合の数の積の法則より、

$$n(A \cap B) = n(A) \cdot n(B|A)$$

よって、積事象 $A \cap B$ の確率は、

$$\begin{aligned} P(A \cap B) &= \frac{n(A \cap B)}{n(U)} \\ &= \frac{n(A) \cdot n(B|A)}{n(U)} \\ &= \frac{n(A)}{n(U)} \cdot \frac{n(B|A)}{n(U)} \\ &= P(A) \cdot P(B|A) \end{aligned}$$

という形で表すことができる

確率の乗法定理　積事象 $A \cap B$ の確率は、

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

ベイズの定理

[Todo 2:]

ref: 数学図鑑 p40～41

独立な試行の確率

2 つの試行が互いに他方の結果に影響を及ぼさないとき、これらの試行は独立であるという

ref: 数学図鑑 p33

このとき、A が起きたかどうかが B の起きやすさに影響しないので、

$$P(B|A) = P(B)$$

が成り立つ

よって、確率の乗法定理は次のように書き換えられる

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

独立な試行の確率　2 つの独立な試行 S, T を行うとき、「試行 S では事象 A が起こり、試行 T では事象 B が起こる」という事象を C とすると、事象 C の確率は、

$$P(C) = P(A) \cdot P(B)$$

反復試行の確率

[Todo 3:]

ref: 数学図鑑 p35〜37

第 4 章

さまざまな立場の確率

古典的立場の確率

サイコロを 1 回投げると、1 から 6 のいずれかの目が出る

ref: スッキリわかる確率統計 p61〜62

ここで、サイコロの目が 1 となる確率は、全体が 6 通りで、1 が出る場合は 1 通りしかないため、$\frac{1}{6}$ と考える

確率をこのように考えることを古典的立場あるいは組合せ的という

古典的立場による確率　全体で n 通りの場合があり、そのうちある事象 A が起こる場合の数が a 通りあるとき、事象 A の起こる確率を次のように定義する

$$P(A) = \frac{a}{n}$$

このように定義された確率を算術的確率あるいは先験的確率という

頻度的立場の確率

実際には、6 回サイコロを投げたときに、必ず 1 が 1 回出るわけではない

ref: スッキリわかる確率統計 p62〜63

サイコロを何回も（膨大な回数を繰り返し）投げれば、やがて 1 が出る確率は $\frac{1}{6}$ に近づいていく

一般的に述べると、n 回中 k 回だけ 1 が出た場合の割合 $\frac{k}{n}$ は、n が大きくなるにつれて一定値 $\frac{1}{6}$ に近づいていく
確率に対するこのような考え方を**頻度的立場**という

頻度的立場による確率　試行を n 回繰り返して行った場合に、ある事象 A の起こった回数を $k(n)$ とする

試行回数 n を増やしていくとき、割合 $\frac{k(n)}{n}$ が一定値 p に近づくならば、p を事象 A の起こる**確率**と定義する

$$P(A) = p = \lim_{n\to\infty} \frac{k(n)}{n}$$

このように定義される確率を**統計的確率**あるいは**経験的確率**という

この立場の客観性を保証するものは、多数回の試行あるいは大量データによる結果であり、理論的な根拠になっているものは**大数の法則**である

主観確率とベイズ統計学

確率を考えるにあたって、頻度的立場だけで十分とは言い切れない

ref: スッキリわかる確率統計 p63〜64

i. 本当にサイコロが均一な材料で作られ、完全な立方体になっているのか？

…もしそうでないなら、頻度的立場は意味がないのではないか？

ii. サイコロを投げて出る目は、投げた瞬間に決まっているのではないか？

…もしそうだとすると、サイコロを手にしたときから出る目は決まっているので、そもそも確率なんて存在しないのではないか？

iii. サイコロの目が出る確率なんて主観的なものでもよいのではないか？

…たとえば、100 回投げて 1 が 30 回出たら、その確率は $\frac{30}{100}$ としてもよいのではないか？

(iii) のような、「確率は主観的なものでよい」という立場の統計学はベイズ統計学と呼ばれている

この立場では、今までの情報、知識や経験などによって得られた確率を与え、これを主観確率という

主観確率では、全く起こっていない、あるいはほとんど起こっていない事象や実験ごとに統計的規則が変わってしまうような事象の分析も可能になる

公理論的立場の確率

ref: スッキリわかる確率統計 p64〜67

化学や物理では、ある現象を考えるとき、議論がしやすいように理想状態というものを考える
それと同じように、確率も理想化された状態で考えることにする

サイコロでいえば、そのサイコロの根拠（均一な材料か、完全な立方体なのか、etc.）を問うのではなく、最初から理想化されたサイコロを考えるようにする

そして、現実の問題と理想化された問題との間を統計的検定を使ってつなぐことにする

現実の問題 $\xleftrightarrow{\text{統計的検定}}$ 理想化された問題

確率を理想化された数学の世界で考えるために、確率をある公理を満たすものとして定義する

確率をこのように考える立場を公理論的立場という

第 5 章

確率の公理と確率空間

事象と集合の同一視

事象という概念を導入することにより、実験や観測による結果を集合に対応させることができた

ref: スッキリわかる確率統計 p67

$$\text{観測結果（事象）} \quad \longleftrightarrow \quad \text{集合}$$

いわば、確率のもととなる事柄を集合に閉じ込めたことになる

このような集合を使って、確率を定義することができる

定義をつくる：集合演算に関する閉性の保証

確率が定義される事象の全体 $\mathscr{A}$ を考える

ref: スッキリわかる確率統計 p68

このとき、確率を数学の世界に閉じ込めるため、$\mathscr{A}$ に集合の演算に関して閉じていることを要求する

具体的には、$\mathscr{A}$ に対して次の 3 つを要求する

i. 空事象 $\emptyset$ と全事象 Ω は $\mathscr{A}$ に含まれる

ii. 事象 A, B が $\mathscr{A}$ に属するとき、その和事象 $A \cup B$、積事象 $A \cap B$、また、A の余事象 A^c も $\mathscr{A}$ に属する

iii. $A_1, A_2, \ldots, A_n, \ldots$ が $\mathscr{A}$ に属せば、$\displaystyle\bigcup_{n=1}^{\infty} A_n$ も $\mathscr{A}$ に属する

ここで、$\displaystyle\bigcup_{n=1}^{\infty} A_n$ は、「$A_1, A_2, \ldots, A_n, \ldots$ のいずれかが起こる」という事象を表している

たとえば、サイコロを投げる試行において、

- 「いつかは 1 の目が出る」という事象を A
- 「n 回目に初めて 1 の目が出る」という事象を A_n

とすると、

$$A = \bigcup_{n=1}^{\infty} A_n$$

と表される

集合族と σ-加法族

$\mathscr{A}$ は事象の全体、つまり、部分集合の全体なので、集合の集まりである

このような「集合の集合」を集合族という

ref: スッキリわかる確率統計 p68〜69

そして、先ほどの 3 つの条件を満たす集合族 $\mathscr{A}$ を σ-加法族という

定義をつくる：どの立場の確率でも成り立つ性質の抽出

ある事象 A の確率 $P(A)$ が満たすべき条件を考える

ref: スッキリわかる確率統計 p69〜70

確率の値のとりうる範囲

確率 $P(A)$ は、事象 A が起こる可能性を表すので、

- 事象が全く起こらないとき：$P(A) = 0$
- 事象が必ず起こるとき：$P(A) = 1$

と考えることができる

そこで、確率 $P(A)$ は次の範囲の値をとりうるものとする

$$0 \leq P(A) \leq 1$$

しかし、この不等式だけでは、$P(A)$ が 0 や 1 になるのはどんな場合なのかを示すことはできない

「事象が必ず起こるときの確率は 1」「事象が全く起こらないときの確率は 0」という、直観的には当たり前の事実も確率の定義に含める必要がある

全事象の確率

事象 A がいつも起こるときは、事象 A は起こりうるすべての場合を含んでいることになるので、A は全事象である
そこで、全事象を Ω とし、次の条件を定義に追加する

$$P(\Omega) = 1$$

空事象の確率

同様に、「事象が全く起こらないときの確率は 0」という事実は、次のように表される

$$P(\emptyset) = 0$$

しかし、実はこの式は確率の定義の他の条件から導出できるので、定義には加えないことにする

排反な事象の確率

最後に、「互いに排反な事象は別々に起こる」という性質も、確率の定義に含めることが重要である

A または B が起こる確率は、

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

であるが、A と B が互いに排反であるときは、

$$P(A \cup B) = P(A) + P(B)$$

これと同様のことが、事象が増えても成り立つことを定義として要求する

$$P(A_1 \cup A_2 \cup \cdots \cup A_n \cdots) = P(A_1) + P(A_2) + \cdots + P(A_n) + \cdots$$

確率と確率空間の定義

以上の議論から、確率を次のように定義する

ref: スッキリわかる確率統計 p70

公理論的立場の確率と確率空間　標本空間 Ω の部分空間を A とし、$\mathscr{A}$ を σ-加法族とする

このとき、次の 3 つの性質を満たす関数 $P(\cdot)$ を事象 A の確率、あるいは $(\Omega, \mathscr{A})$ 上の確率という

i. 任意の $A \in \mathscr{A}$ に対して、$0 \leq P(A) \leq 1$

ii. $P(\Omega) = 1$

iii. 完全加法性：任意の互いに排反な事象 $A_1, \ldots, A_n, \ldots$ に対して、$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$

なお、3 つの組 $(\Omega, \mathscr{A}, P)$ を確率空間という

第 6 章

確率変数と確率分布

確率変数

離散的な例

ref: スッキリわかる確率統計 p75〜77

サイコロを投げて出る目の値を X とすると、X のとりうる値は $1, 2, 3, 4, 5, 6$ のいずれかである

X が出るという事象が同様に確からしいならば、X は確率 $\frac{1}{6}$ で、$1, 2, 3, 4, 5, 6$ のいずれかの値をとる

ここで、X が i となる確率を $P(X = i)$ と表すことにすると、

$$P(X = 1) = \cdots = P(X = 6) = \frac{1}{6}$$

が成り立つ

この場合、標本空間は $\Omega = \{1, 2, 3, 4, 5, 6\}$ であり、X は Ω 上の値をとる関数あるいは変数と考えられる

連続的な例

ルーレットを回して最初の位置から X 度のところで止まったとすると、X のとりうる値は $0 < X \leq 360$ を満たすすべての実数である

X 度のところで止まるという事象が同様に確からしいとすると、たとえば X が $30 \leq X < 90$ という値をとる確率は、区間幅 $90 - 30 = 60$ に比例することになる

そこで、X が $a \leq X < b$ という値をとる確率を $P(a \leq X < b)$ と表すことにすると、

$$P(30 \leq X < 90) = \frac{60}{360} = \frac{1}{6}$$

となる

この場合、標本空間は $\Omega = \{x \mid 0 < x \leq 360\}$ であり、X は Ω 上の値をとる関数あるいは変数と考えられる

確率変数の定義

このように、$X = k$ となる確率 $P(X = k)$ や、$a \leq X < b$ となる確率 $P(a \leq X < b)$ が定まっている変数 X を確率変数という

確率変数　試行の結果に応じていろいろな値をとる変数 X が考えられ、変数 X がある値をとる場合の確率が定まるとき、X を確率変数という
また、確率変数が離散的な値をとるときは離散型確率変数、連続的な値をとるときは連続型確率変数という

確率変数は、標本空間上の実数値関数ともいえる

離散型確率変数の確率分布

ref: スッキリわかる確率統計 p77〜79

確率関数　離散型確率変数 X が定数 x_i という値を取る確率を次のようにおく

$$P(X = x_i) = p_i$$

ただし、各確率 p_i は 0 以上の値であり、その総和は 1 であるとする

$$\sum_i p_i = 1, \quad p_i \geq 0$$

このとき、$P(X = x_i) = p_i$ は x_i の関数なので、それぞれの値に対する確率を関数 $f(x)$ を用いて、次のように表すことができる

$$f(x_i) = P(X = x_i)$$

この関数 $f(x)$ を X の**確率関数**といい、x_i と p_i との対応 $f(x_i) = p_i$ を X の**確率分布**という

連続型確率変数の確率分布

ref: スッキリわかる確率統計 p79〜81

確率密度関数　連続型確率変数 X に対して、次の条件を満たす関数 $f(x)$ の存在を仮定する

[Todo 4:]

Zebra Notes

Type	Number
todo	4