

数学 A 1 章「場合の数と確率」 2 節「確率」 カンペ

2024 年 5 月 17 日

TAIYO

授業で使うスライド資料*¹ は作成したが、何を話すか忘れてしまいそうなのでこのようなレジュメも用意した。指導案の授業計画は高校生視点での展開なので、ここでは先公が話すことすることを書いた。自分のための単なるメモであり、高校生に配布するということを全く想定していない。

1 確率の意味と用語の確認

1.1 試行と事象

■**試行とは？ 事象とは？** 同じ状態のもとで繰り返すことができ、結果が偶然によって決まる実験や観測などを**試行**という。試行の結果として起こる事柄を**事象**という。例えば、「硬貨を 1 回投げる」とか「彼女とおみくじを引く」とかが試行で、その結果の「表が出る」や「俺は吉で彼女は凶」が事象である。

■**全事象と根元事象** この先、事象を集合を用いて表す。

具体例として、サイコロを 1 回投げるという試行を扱う。出る目は $1, 2, \dots, 6$ のいずれかなので、これらを要素としてもつ集合

$$U = \{1, 2, 3, 4, 5, 6\} \quad (1)$$

を定めておく。 U を**全事象**という。各々の試行について、あらかじめ全事象 U を決めておくことで、様々な事象を U の部分集合として表せるので便利である。例えば、偶数の目が出るという事象 A は $A = \{2, 4, 6\}$ と書き表され、もちろん $A \subset U$ である。 U の部分集合として要素を持たない集合 \varnothing を選んでもよく、これを**空事象**という。 U の 1 つの要素からなる集合

$$\{1\}, \{2\}, \{3\}, \{4\}, \{5\}, \{6\} \quad (2)$$

は**根元事象**と呼ばれる。

*¹ スライドは暫定的に <https://github.com/sodesudesu/suA.git> に置いとく。今のところ private のまま。

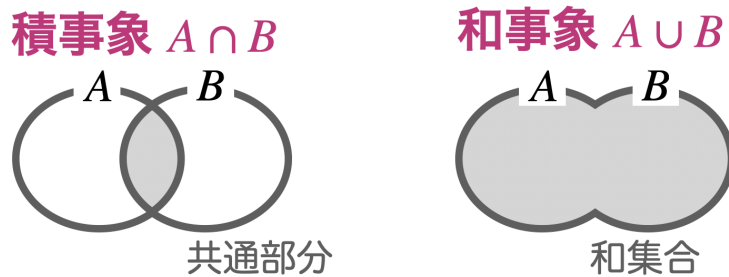


図 1: 積事象 $A \cap B$ のベン図 (左) と和事象 $A \cup B$ のベン図 (右)。

当然だが、全事象の選び方には任意性がある。(1) では全事象としてサイコロの出る目を指定したが、テーブル上のサイコロが着地する位置だとかサイコロが静止するまでの時間、さらにはこれらを組み合わせた (ものすごい凝った) 全事象を考えることもできる。キリがない。(1) は、サイコロ 1 個を投げるという操作に伴う雑多な情報の中から、出る目にだけ注目しました！ という宣言だと捉えられる。

■積事象と和事象 2つの事象 A, B について、 A と B がともに起こるという事象を**積事象**といい、共通部分 $A \cap B$ で表される。 A または^{*2} B が起こるという事象を**和事象**といい、和集合 $A \cup B$ で表される。

例 1 (サイコロを 1 回投げる試行での積事象と和事象) サイコロを 1 回投げる。偶数の目が出るという事象を A 、4 以上の目が出るという事象を B とすると、これらは

$$A = \{2, 4, 6\}, \quad (3)$$

$$B = \{4, 5, 6\} \quad (4)$$

という集合で表される。これらの積事象 $A \cap B$ と和事象 $A \cup B$ はそれぞれ、

$$A \cap B = \{4, 6\}, \quad (5)$$

$$A \cup B = \{2, 4, 5, 6\} \quad (6)$$

となる。ベン図を描いて確かめよ。

^{*2} 図 1 を見るとわかるように、 $A \cup B$ は $A \cap B$ で示される領域も含んでいる。つまり、「 A または B 」は「 A だけ」と「 B だけ」と「 A も B も両方」という 3 つの場合を含んだ表現である。国語の授業で習う「または」には 3 つ目の意味は含まれていないので (「法律により 10 年以下の懲役、もしくは 1000 万円以下の罰金、またはその両方が科せられます。NO MORE 映画泥棒」がいい例)、数学の「または」は日常語の「または」よりも欲張りだ。

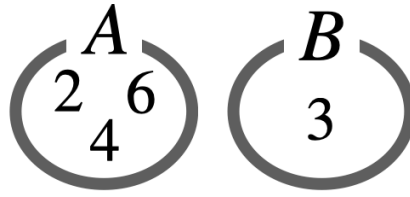


図 2: 事象 $A = \{2, 4, 6\}$ と事象 $B = \{3\}$ のベン図。 $A \cap B = \emptyset$ つまり、ベン図でこれらの集合は交わらない。3 つ以上の事象についても、ベン図を描いて交わらないならば、それらは互いに排反である。

■**排反事象** サイコロを 1 回投げる試行について（再び）考えよう。事象 A : 「偶数の目が出る」と事象 B : 「3 の目が出る」の積事象 $A \cap B$ は、

$$A \cap B = \emptyset \text{ (空事象)} \quad (7)$$

である（図 2）。つまり、事象 A と事象 B は同時には起こらない。この時、事象 A と事象 B は互いに排反である、または、**互いに排反事象**であるという。

1.2 確率の計算と同様に確からしい根元事象について注意

ついに確率の計算が始まる。

■**確率の計算** ある 1 つの試行について、根元事象のどれが起こることも同じ程度に期待される時、これらの根元事象は**同様に確からしい**という。

全事象の**どの根元事象も同様に確からしい**とき、事象 A の確率 $P(A)$ は

$$P(A) = \frac{\text{事象 } A \text{ の要素の個数 } n(A)}{\text{全事象 } U \text{ の要素の個数 } n(U)} \quad (8)$$

で計算される。

例 2（「どの根元事象も同様に確からしい」に注意せよ！） 3 枚のカードがあり、それぞれ 0, 1, 2 と書かれている。カードを裏返しにしてよくきり、1 枚選ぶ。数字を確認したあと、そのカードは戻してよくシャッフルし、再び 1 枚選んで数字を確認する。選んだ 2 つのカードに書かれた数字の和が 3 となる確率はいくらか。

(答え?) 2つの数字の和の候補は 0, 1, 2, 3, 4 だから、全事象は、

$$\{0, 1, 2, 3, 4\} \quad (9)$$

である。したがって、

$$\text{和が3となる確率} = \frac{\text{事象「和が3」の要素数}}{\text{全事象の要素数}} = \frac{1}{5} \quad (10)$$

となる。

この解答は一見もっともらしいが、根本的な誤りがある。反省すべきは (9) のように根元事象を

$$\{0\}, \{1\}, \{2\}, \{3\}, \{4\} \quad (11)$$

としてしまった点である。確率の求め方 (8) を使えるのは、根元事象が同様に確からしい場合であった。(11) は同様に確からしい根元事象ではない。なぜならば、(1枚目のカードの数字, 2枚目のカードの数字) と表すことにすると、

| 和 | 0 | 1 | 2 | 3 | 4 |
|------------|--------|------------------|----------------------------|------------------|--------|
| 2枚の 組合せ | (0, 0) | (0, 1) (1, 0) | (0, 2) (1, 1) (2, 0) | (1, 2) (2, 1) | (2, 2) |

のように 0~4 の各和にカードの選ばれ方が対応しており、それぞれの場合の数が異なっているからだ。表の下側に現れた 2枚の組合せ (0, 0), (0, 1), ..., (2, 2) は、どれが起こることも同様に期待できる。よって、今考えている試行において根元事象として選ぶべきは、2つの数字の和ではなく、2枚のカードの組合せであった。

こっちが正しい (答え) 根元事象を 2枚のカードの組合せとして、(上の表を見ながら) 2つの数字の和が 3 となる確率を正しく計算すると、

$$\text{和が3となる確率} = \frac{2}{9} \quad (12)$$

となる。

くり返しになるが、確率の定義 (8) は、**全事象のどの根元事象も同様に確からしいときに限って使える** (やや不自由な) ものである。つまり我々は確率を計算する際、まず、同様に確からしい根元事象を探さなければならない。上の例では、

だめ：根元事象は $\{0\}, \{1\}, \{2\}, \{3\}, \{4\}$ だ！

OK：根元事象は $\{(0, 0)\}, \{(0, 1)\}, \dots, \{(2, 2)\}$ だ！

であった。

複数個のサイコロや複数個の硬貨などが登場する場合に、名前をつけてそれらを区別したり、(例 2 のような) 同じ操作を繰り返し行う場合に「1 回目」「2 回目」・・・のように回数でそれらを区別したりするのが「同様に確からしい根元事象探し」の常套手段である。

例 3 (複数個の硬貨に名前をつけて区別) 2 枚の硬貨を同時に投げる時、2 枚とも表が出る確率を計算せよ。

(答え) 硬貨が 2 枚登場するので、1 つを硬貨 A、もう 1 方を硬貨 B と名付けて区別する。(硬貨 A の表裏, 硬貨 B の表裏) と表すと、根元事象は

$$\{(\text{表}, \text{表})\}, \{(\text{表}, \text{裏})\}, \{(\text{裏}, \text{表})\}, \{(\text{裏}, \text{裏})\} \quad (13)$$

のように選べば、これらは同様に確からしい ($\{2$ 枚とも表 $\}, \{1$ 枚は表、1 枚は裏 $\}, \{2$ 枚とも裏 $\}$, は同様に確からしいと言えない)。したがって、2 枚とも表が出る確率は、

$$P(2 \text{ 枚とも表}) = \frac{1}{4} \quad (14)$$

である。

1.3 確率の基本性質

■確率の基本性質その 1 1.1 節で述べたように、事象 A は全事象 U の部分集合で表すことができるので、 A の要素の個数 $n(A)$ について、

$$n(\emptyset) \leq n(A) \leq n(U) \quad (15)$$

が成り立つ。各辺を $n(U)$ で割ると、

$$\frac{n(\emptyset)}{n(U)} \leq \frac{n(A)}{n(U)} \leq \frac{n(U)}{n(U)} \quad (16)$$

すなわち、

$$0 \leq P(A) \leq 1 \quad (17)$$

が成り立つことがわかる。ただし、最左辺の変形には、空事象 \emptyset は要素を持たない集合で表されるから、 $n(\emptyset) = 0$ となることを利用した。中辺の変形には確率の定義 (8) を使った。(16) と (17) を見比べると、 $P(A) = 0$ は $A = \emptyset$ の時に、 $P(A) = 1$ は $A = U$ の時に成り立つことがわかる。

■確率の基本性質その 2 (確率の加法定理) 事象 A と事象 B は互いに排反である とする。これらの和事象 $n(A \cup B)$ の要素の個数は、

$$n(A \cup B) = n(A) + n(B) \quad (18)$$

と書ける。図 2 から明らかなように、2 つの集合が交わらないので各集合の要素数を単純に足し合わせたものが和事象 $A \cup B$ の要素数になるのだ。

確率の定義 (8) と (18) を用いることで、和事象 $A \cup B$ の確率 $P(A \cup B)$ は、

$$\begin{aligned} P(A \cup B) &= \frac{n(A \cup B)}{n(U)} \\ &= \frac{n(A) + n(B)}{n(U)} \\ &= \frac{n(A)}{n(U)} + \frac{n(B)}{n(U)} \end{aligned} \quad (19)$$

となるので、

$$P(A \cup B) = P(A) + P(B) \quad (20)$$

という表式が得られる。(20) は**確率の加法定理** と呼ばれる。

(20) は、事象 A と事象 B が互いに排反であるときに限って使えるということに注意せよ。 A と B が互いに排反でない場合への拡張は次節でおこなう。

例 4 (確率の加法定理の問題) 2 つのサイコロを同時に投げる時、出た目の和が 5 の倍数になる確率を計算せよ。

(答え) 2 つの目の和としてとり得るのは、2~12 の整数である。この内、5 の倍数は 5 と 10 であるから、事象 A 「出た目の和が 5 になる」と事象 B 「出た目の和が 10 になる」という 2 つの排反な事象の確率をそれぞれ計算し、最後に足し合わせればよい。(1 つ目のサイコロの目, 2 つ目のサイコロの目) のように表すことにすれば、

$A = \{(1, 4), (2, 3), (3, 2), (4, 1)\}$, $B = \{(4, 6), (5, 5), (6, 4)\}$ なので、確率の定義 (8) より、

$$P(A) = \frac{n(A)}{n(U)} = \frac{4}{36}, \quad (21)$$

$$P(B) = \frac{n(B)}{n(U)} = \frac{3}{36} \quad (22)$$

と求められる。確率の加法定理 (18) より、

$$P(A \cup B) = P(A) + P(B) = \frac{4}{36} + \frac{3}{36} = \frac{7}{36} \quad (23)$$

となる。

2 様々な確率の計算

2.1 和事象と余事象の確率

■**和事象の確率** 前節で強調したように、確率の加法定理 (20) は互いに排反な事象の和事象を計算するものであった。ここでは、互いに排反とは限らない事象についても和事象を計算することが目標である。

互いに排反ではない 2 つの事象 A , B について、これらの和事象の要素の個数は

$$n(A \cup B) = n(A) + n(B) - n(A \cap B) \quad (24)$$

と表される (図 3 を参照、もしくは教科書やノートの集合のページを復習せよ)。確率の定義 (8) とあわせて、

$$\begin{aligned} P(A \cup B) &= \frac{n(A \cup B)}{n(U)} \\ &= \frac{n(A) + n(B) - n(A \cap B)}{n(U)} \\ &= \frac{n(A)}{n(U)} + \frac{n(B)}{n(U)} - \frac{n(A \cap B)}{n(U)} \end{aligned} \quad (25)$$

であるから、和事象の確率の公式

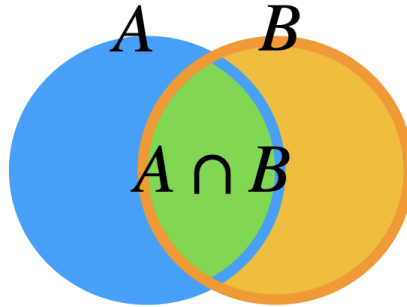


図 3: 互いに排反でない事象 A と事象 B のベン図。和事象 $A \cup B$ の要素の個数はどう表されるか。各集合の要素の個数を $n(A) + n(B)$ のように足し算してしまうと、共通部分 $A \cap B$ (図の緑色の部分) を余分に数え上げることになる。したがって、2つの集合の要素数の和 $n(A) + n(B)$ から、重複して数えてしまった $n(A \cap B)$ を引いたものが、和事象 $A \cup B$ の要素の個数となる： $n(A \cup B) = n(A) + n(B) - n(A \cap B)$ 。

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \quad (26)$$

が導かれる。

事象 A と事象 B が互いに排反 (図 2 のように、ベン図で描けば互いに交わらない) であれば、 $A \cap B = \emptyset$ つまり $n(A \cap B) = 0$ だから、 $P(A \cap B) = \frac{n(A \cap B)}{n(U)} = 0$ となるので、(26) の右辺第 3 項が消えて、

$$P(A \cup B) = P(A) + P(B) \quad (27)$$

が得られる。この表式は、6 ページで紹介した確率の加法定理 (18) である。つまり、確率の加法定理は $A \cap B = \emptyset$ (A と B は互いに排反事象) という特殊な場合として、和事象の確率の公式 (26) に含まれている。

例 5 (倍数と確率) 1 から 100 までの数字が書かれた 100 枚のカードがある。カードをよくシャッフルして 1 枚取り出す時、その番号が 4 の倍数または 6 の倍数である確率を計算せよ。

(答え) 取り出したカードの数字が 4 の倍数である事象を A 、6 の倍数である事象を B とする。ここで、 $A \cup B$ は 4 の倍数であり 6 の倍数でもある (つまり 12 の倍数) カードを引くという事象である。各事象の要素の個数は $n(A) = 25$, $n(B) = 16$, $n(A \cup B) = 8$

であるから、各事象の確率は、

$$P(A) = \frac{25}{100}, P(B) = \frac{16}{100}, P(A \cup B) = \frac{8}{100} \quad (28)$$

と計算できる。もちろん $n(U) = 100$ と確率の定義 (8) を使った。和事象の確率の公式 (26) より、

$$\begin{aligned} P(A \cup B) &= P(A) + P(B) - P(A \cap B) \\ &= \frac{25}{100} + \frac{16}{100} - \frac{8}{100} = \frac{33}{100} \end{aligned} \quad (29)$$

が得られる。

■余事象の確率 事象 A に対して、 A が起こらないという事象を**余事象**といい、補集合 \bar{A} で表す。図 4 を見ると明らかなように、

$$n(\bar{A}) = n(U) - n(A) \quad (30)$$

という関係が成り立つ。これまでと同様に、 $n(\cdot)$ は各集合の要素の個数を表す。確率の定義 (8) に (30) を適用すれば、

$$\begin{aligned} P(\bar{A}) &= \frac{n(\bar{A})}{n(U)} \\ &= \frac{n(U) - n(A)}{n(U)} \end{aligned} \quad (31)$$

となるので、余事象の確率の公式

$$P(\bar{A}) = 1 - P(A) \quad (32)$$

が導かれる。もちろん、(32) の $P(\bar{A})$ と $P(A)$ を移行すれば*3、

$$P(A) = 1 - P(\bar{A}) \quad (33)$$

という表式も得られる。

*3 (32) の A を \bar{A} に置き換えて、補集合の性質 $\overline{\bar{A}} = A$ を利用しても、(33) を得られる。

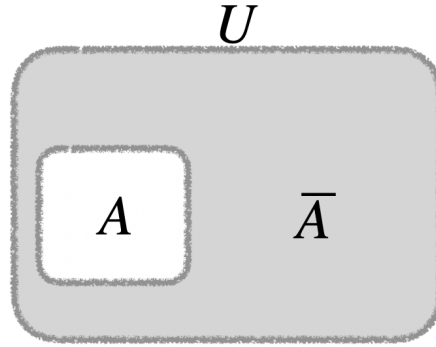


図 4: 事象 A とその余事象 \bar{A} のベン図。

例 6 (ド・メレの問題) 次のどちらの賭けの勝率が高いか。

- サイコロを 4 回投げて、1 回でも 6 の目が出たら勝ち。
- サイコロ 2 つを同時に 24 回投げて、1 回でもゾロ目 (6, 6) が出たら勝ち。

(答え)

- 求めたいのは、事象 A 「4 回中少なくとも 1 回は 6 の目が出る」の確率 $P(A)$ である。 A の余事象 \bar{A} は「4 回全て 1 以外の目が出る」であり、この確率 $P(\bar{A})$ は

$$P(\bar{A}) = \frac{n(\bar{A})}{n(U)} = \frac{5 \times 5 \times 5 \times 5}{6 \times 6 \times 6 \times 6} = \left(\frac{5}{6}\right)^4 \quad (34)$$

と求まる。余事象の確率の公式 (33) より、

$$P(A) = 1 - \left(\frac{5}{6}\right)^4 \quad (35)$$

とわかる。

- 求めたいのは、事象 B 「24 回中少なくとも 1 回は (6, 6) が出る」の確率 $P(B)$ である。 B の余事象 \bar{B} は「24 回全て (6, 6) 以外の目が出る」であり、この確率 $P(\bar{B})$ は

$$P(\bar{B}) = \frac{n(\bar{B})}{n(U)} = \frac{35^{24}}{36^{24}} = \left(\frac{35}{36}\right)^{24} \quad (36)$$

となる。余事象の確率の公式 (33) を用いれば、

$$P(B) = 1 - \left(\frac{35}{36}\right)^{24} \quad (37)$$

となる。

計算機の助けを借りると、 $P(A) = 0.518$, $P(B) = 0.491$ が得られる。よって、1 つめ賭けの方が勝率が高い。

2.2 独立な試行の確率

■独立な試行とは 2 つの試行について、互いに他方の結果に影響を及ぼさない時、これらの試行は**独立**であるという。硬貨を何回か続けて投げる時、表のあとは裏が出やすいということはない。つまり、毎回のコイントスは独立である。

例 7 (引いたカードを戻すか戻さないかで独立かどうか変わる) 袋に赤玉が 3 つと白玉が 2 つは入っている。目隠しをして袋から玉を一つ取り出す。玉の色を確認したのちに、

(a) 玉を戻さずにもう一度玉を取り出す。1 回目と 2 回目で袋の中の玉の状態が変わってしまうので、これらは独立な試行ではない。

(b) 玉を戻してからもう一度玉を取り出す。1 回目と 2 回目で袋の中の玉の状態は同じ。よって、1 回目で何色を取り出そうが 2 回目の結果には影響を及ぼさない。これらは独立な試行となる。

■独立な試行の計算 次の例題の考察から、独立な試行の確率の公式を導こう (公式と呼ぶは大袈裟な気もするが)。

例 8 (独立な試行の確率の公式を導く) a, b, c の 3 種類のアルファベットが印字されたカードと、1, 2, 3, 4 がそれぞれ印字された数字カードがある。3 枚のアルファベットカードから 1 枚、4 枚の数字カードの中から 1 枚を (目隠しをして) 選ぶ。これらは独立な試行である。この時、事象 A 「アルファベットカードは b を引き、数字カードは 1 か 3 を引く」の確率 $P(A)$ を求めよ。

(答え) $P(A) = n(A)/n(U)$ のように事象 A の確率を求められるので、図 5 を見ながら $n(A)$ と $n(U)$ を数えれば、

$$n(A) = 1 \times 2, \quad n(U) = 3 \times 4 \quad (38)$$

となる。よって、

$$P(A) = \frac{n(A)}{n(U)} = \frac{1 \times 2}{3 \times 4}$$

| | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| a | | | | |
| b | | | | |
| c | | | | |

図 5: 3 枚のアルファベットカードから 1 枚選ぶ試行と 4 枚の数字カードから 1 枚ずつ選ぶ試行において、起こりうる結果をまとめたもの。選ばれるカードの組合せの総数（つまり全事象 U の要素の個数）は表のマス目の個数であるから $n(U) = 3 \times 4$ となる。アルファベットカードから b を引き、数字カードから 1 または 3 を引くという事象 A は、表の濃く塗りつぶしてあるマスに対応しており、 $n(A) = 1 \times 2$ である。

であるが、最右辺をよく見ると、

$$\begin{aligned}
 &= \frac{1}{3} \times \frac{2}{4} \\
 &= P(\text{アルファベットカードの b を引く}) \times P(\text{数字カードの 2 または 4 を引く}) \quad (39)
 \end{aligned}$$

という形をしていることに気づく。つまり、アルファベットカードから 1 枚選ぶという試行において b を選ぶ確率と数字カードから 1 枚選ぶという試行において 1 または 3 を選ぶ確率を別個計算しておき、それらをかけ合わせたものが $P(A)$ となる。

この例でわかった事実をほんのちょっとだけ一般的に書き表そう：独立な試行 T_1, T_2 について、試行 T_1 では事象 A が起こり、試行 T_2 では事象 B が起こるという事象が起こる確率 P は、

$$P = P(A)P(B) \quad (40)$$

で計算される。

例 9（俺には KISUKE しかない） KISUKE には 4 つのステージがある：

1. 5 杯のジュースのうち、1 杯は超酸っぱい！ ラッピージュース。
2. 4 つのジャンプ台のうち、1 つは着地地点に足つぽマット。

3. 3つのシュークリームのうち、1つは激辛からし入り。
4. 2脚の椅子のうち、1脚はビリビリ椅子。

ダメージをくらったらそこで失格。KISUKE をクリアする確率を計算せよ。

(答え) 1st ステージから 4th ステージまで、それぞれの挑戦は独立である。だから、各ステージでの成功確率を計算し、最後にそれらをかけ算すればよい：

$$P = \frac{4}{5} \times \frac{3}{4} \times \frac{2}{3} \times \frac{1}{2} = \frac{1}{5}. \quad (41)$$

2.3 反復試行の確率

独立な試行を繰り返すという試行を**反復試行**という。

例 10 (サイコロを使った反復試行) サイコロを 4 回投げる。4 回の試行はお互いの結果に影響を及ぼさない。例えば、1 が出た後は 1 が出づらいということはない。つまりこれらは独立な試行であり、独立な試行を繰り返しておこなっているので反復試行と言える。4 回中ちょうど 2 回、1 の目が出る確率を計算せよ。

(答え) まず、4 回中 2 回、1 の目が出るシチュエーションを全て列挙しよう：

| | 1 回目 | 2 回目 | 3 回目 | 4 回目 |
|----------|------|------|------|------|
| <i>A</i> | 1 | 1 | 1 以外 | 1 以外 |
| <i>B</i> | 1 | 1 以外 | 1 | 1 以外 |
| <i>C</i> | 1 | 1 以外 | 1 以外 | 1 |
| <i>D</i> | 1 以外 | 1 | 1 | 1 以外 |
| <i>E</i> | 1 以外 | 1 | 1 以外 | 1 |
| <i>F</i> | 1 以外 | 1 以外 | 1 | 1 |

表から全 6 パターンあることがわかる。総数の“6”は、4 回の試行に対し 2 つの 1 を割り当てる場合の数の計算であるから、 ${}_4C_2$ で求められる。説明の都合で 6 つそれぞれに *A* ~ *F* と名前をつけた。*A* が起こる確率は、独立な試行の確率の公式 (40) より、

$$P(A) = \frac{1}{6} \times \frac{1}{6} \times \frac{5}{6} \times \frac{5}{6} = \left(\frac{1}{6}\right)^2 \times \left(\frac{5}{6}\right)^2 \quad (42)$$

である。 B, C, \dots, F が起こる確率も同様に、

$$P(B) = \frac{1}{6} \times \frac{5}{6} \times \frac{1}{6} \times \frac{5}{6} = \left(\frac{1}{6}\right)^2 \times \left(\frac{5}{6}\right)^2, \quad (43)$$

$$P(C) = \frac{1}{6} \times \frac{5}{6} \times \frac{5}{6} \times \frac{1}{6} = \left(\frac{1}{6}\right)^2 \times \left(\frac{5}{6}\right)^2, \quad (44)$$

$$P(D) = \frac{5}{6} \times \frac{1}{6} \times \frac{1}{6} \times \frac{5}{6} = \left(\frac{1}{6}\right)^2 \times \left(\frac{5}{6}\right)^2, \quad (45)$$

$$P(E) = \frac{5}{6} \times \frac{1}{6} \times \frac{5}{6} \times \frac{1}{6} = \left(\frac{1}{6}\right)^2 \times \left(\frac{5}{6}\right)^2, \quad (46)$$

$$P(F) = \frac{5}{6} \times \frac{5}{6} \times \frac{1}{6} \times \frac{1}{6} = \left(\frac{1}{6}\right)^2 \times \left(\frac{5}{6}\right)^2 \quad (47)$$

と計算される。結局、 $(1/6)^2 \times (5/6)^2$ が、 ${}_4C_2 = 6$ 組できるとわかる。

A, \dots, F は互いに排反であるから（排反な事象については 1.1 節を復習せよ）、確率の加法定理 (20) が利用できる。ゆえに、4 回中ちょうど 2 回、1 の目が出る確率 P は、

$$\begin{aligned} P &= P(A) + P(B) + P(C) + P(D) + P(E) + P(F) \\ &= \left(\frac{1}{6}\right)^2 \times \left(\frac{5}{6}\right)^2 + \left(\frac{1}{6}\right)^2 \times \left(\frac{5}{6}\right)^2 + \left(\frac{1}{6}\right)^2 \times \left(\frac{5}{6}\right)^2 \\ &\quad + \left(\frac{1}{6}\right)^2 \times \left(\frac{5}{6}\right)^2 + \left(\frac{1}{6}\right)^2 \times \left(\frac{5}{6}\right)^2 + \left(\frac{1}{6}\right)^2 \times \left(\frac{5}{6}\right)^2 \\ &= {}_4C_2 \times \left(\frac{1}{6}\right)^2 \times \left(\frac{5}{6}\right)^2 \end{aligned} \quad (48)$$

で求められる。

一般に、ある試行を n 回繰り返す時、事象 A がちょうど r 回起こる確率 P は、

$$P = {}_nC_r \times (P(A))^r \times (1 - P(A))^{n-r} \quad (49)$$

である。上のサイコロの例を思いだせば、すぐに (49) は復元できるだろう。だから無理に公式として覚えるまでもない。もちろん、(49) において、 $n = 4, r = 2, P(A) = 1/6$ と置き換えれば (48) が導かれる。

表 1: 生徒 100 人に「文系か理系か」、「爬虫類を飼っているか」という質問をして得られたデータ。100 人からランダムに 1 人選んだとき、その生徒が文系であるという事象 A 、理系であるという事象 \bar{A} 、爬虫類を飼っているという事象 B 、爬虫類を飼っていないという事象 \bar{B} の確率を考える。単に選ばれた生徒が文系である確率を計算したければ、表の 4 行目に着目することで、 $P(A) = 65/100$ が求まる。選ばれた生徒が爬虫類を飼っていることをあらかじめ知らされている場合、表の 2 行目に着目し、条件付き確率 $P_B(A) = 10/25$ が求まる。同じ「文系である確率」の計算でも、爬虫類を飼っているかどうかの情報の有無によって確率の計算過程（ここでは表のどこに注目して計算を進めるか）が変わる。

| | 文系 A | 理系 \bar{A} | 計 |
|----------------------|--------|--------------|-----|
| 爬虫類を飼っている B | 10 | 15 | 25 |
| 爬虫類を飼っていない \bar{B} | 55 | 20 | 75 |
| 計 | 65 | 35 | 100 |

3 条件付き確率

3.1 条件付き確率

■条件付き確率ってなんなのさ 生徒会が全校生徒 100 人にアンケートをとった。アンケートは「あなたは文系ですか、理系ですか？」と「あなたは爬虫類を飼っていますか？」の 2 問。結果は表 1 のようになった。

アンケートに答えた 100 人の中からランダムに 1 人選び出したとき、その生徒が文系であるという事象を A 、その生徒が爬虫類を飼っているという事象を B とする。すると、その生徒が理系である事象と爬虫類の飼い主ではないという事象はそれぞれ A , B の余事象 \bar{A} , \bar{B} で表される。

表 1 から、回答者 100 人のうち文系は 65 人いるとわかった。この事実をもとに、事象 A の確率（選ばれたその生徒が文系である確率）は、

$$P(A) = \frac{65}{100} = \frac{13}{20} \quad (50)$$

と計算できる。再び 100 人の中から 1 人をランダムに選ぶ。今回はあらかじめその生徒が爬虫類を飼っていることがわかっていたとする。表 1 から、爬虫類を飼っている生徒は

25 人いて、そのうち文系は 10 人いると読み取れるので、この生徒が文系である確率は、

$$P(A) = \frac{10}{25} = \frac{2}{5} \quad (51)$$

となる。

(50) も (51) も文系である確率 $P(A)$ を計算するものだが、値が異なっている。なぜだろうか。(51) を計算する際、事前に「その生徒が爬虫類を飼っている (事象 B)」という情報を得ていた。言い換えると、事象 B が起こったことがわかった上で事象 A の確率を計算していた。一方で、(50) のときは爬虫類を飼っているかの情報はなかった。事前情報の有無によって確率の計算結果が変わってしまうのだ (表 1 の説明も参照)。(51) について、事象 A の確率を計算する際に「事象 B が起こったことを事前に知っていましたよ」ということを強調するために、確率の記号として $P(A)$ に下付き文字を加えた $P_B(A)$ を用いる。 $P_B(A)$ を事象 B が起こったときに事象 A が起こる**条件付き確率**と呼ぶ。

■条件付き確率の計算 条件付き確率 (51) を見ると、文系かつ爬虫類を飼っている生徒の人数 $n(A \cap B) = 10$ 人を、爬虫類を飼っている生徒の人数 $n(B) = 25$ 人で割っているとわかる。つまり、

$$P_B(A) = \frac{n(A \cap B)}{n(B)} \quad (52)$$

と書ける。この式の右辺の分母分子を $n(U)$ で割ると、

$$P_B(A) = \frac{P(A \cap B)}{P(B)} \quad (53)$$

という表式も得られる。我々は、(52) や (53) を用いて、事象 B が起こったときに事象 A が起こる条件付き確率 $P_B(A)$ を求める。

例 11 (条件付き確率の計算問題) 表 1 の 100 人から選び出された生徒が理系であることがわかっているとき、その生徒が爬虫類を飼っている確率 $P_{\bar{A}}(B)$ を計算せよ。

(答え) (52) より、

$$P_{\bar{A}}(B) = \frac{n(\bar{A} \cap B)}{n(\bar{A})} = \frac{3}{7} \quad (54)$$

となる。

3.2 確率の乗法定理

(53) の両辺を $P(A)$ 倍すると、

$$P(A \cap B) = P(B)P_B(A) \quad (55)$$

が得られる。これは、**確率の乗法定理**と呼ばれる。

例 12 (確率の乗法定理 (55) の使いどころ) 箱に赤玉 3 個と白玉 2 個が入っている。目をつぶって箱から玉をひとつ取り出し、その玉を箱に戻さずにもう一度目をつぶって箱から玉を取り出した。2 つとも赤玉である確率を計算せよ。

(答え) 1 回目赤玉である確率は $P(1 \text{ 回目赤}) = 3/5$ で、1 回目に赤玉を引いた時に 2 回目に赤玉を引く条件付き確率は $P_{1 \text{ 回目赤}}(2 \text{ 回目赤}) = 1/2$ となる。確率の乗法定理 (55) より、1 回目も 2 回目も赤玉を取り出す確率は、

$$\begin{aligned} P(1 \text{ 回目赤} \cap 2 \text{ 回目赤}) &= P(1 \text{ 回目赤}) \times P_{1 \text{ 回目赤}}(2 \text{ 回目赤}) \\ &= \frac{3}{5} \times \frac{1}{2} \\ &= \frac{3}{10} \end{aligned} \quad (56)$$

と計算される。

3.3 原因の確率

■本節でやりたいこと 前節までに述べたことの繰り返しになるが、ある事象 B が起こった時に事象 A が起こる条件付き確率は $P_B(A)$ と表され、(52) や (53) で計算され

る。あらかじめ知っている「 A が起こったという事実」をもとに、 B が起こる確率を計算している。

本節では、『 A が起こったという事実』を確認していないが、結果的に事象 B が起こった。じゃあ A が起こっていたのだろうか？」ということを確認率を用いて考えたい。

例 13 (spam メールの判別) 受信者の意図を無視した一方的なメールを spam メールといい、spam ではないものを ham と呼ぶ。

手元に 10000 件のメールがあり、そのうち 3000 件が spam で残り 7000 件が ham である。spam メールのうち 80 %には本文に”claim”という単語が含まれている。一方、ham の 5 %には本文に”claim”という単語が含まれている。一応表にしておく：

| | spam (A) | ham (\bar{A}) | 計 |
|----------------------------|--------------|-------------------|-------|
| claim が含まれる (B) | 2400 | 350 | 2750 |
| claim が含まれない (\bar{B}) | 600 | 6650 | 7250 |
| 計 | 3000 | 7000 | 10000 |

この 10000 件のメールの中から、無作為に 1 件メールを選んだところ本文に”claim”という単語が含まれていた。このメールが spam である確率を計算せよ。

(答え) 選んだメールが spam であるという事象を A 、選んだメールに claim という単語が含まれている事象を B とする。条件付き確率の公式 (52) より、

$$P_B(A) = \frac{n(A \cap B)}{n(B)} = \frac{48}{55} \quad (57)$$

である。 $n(\cdot)$ には表を睨みながら数字を代入するだけ。終わり！

以上のように、表を書けば計算は一発だが、ここではあえて愚直に計算を進めよう。問題文に「10000 件のメールがあり、そのうち 3000 件が spam で残り 7000 件が ham」とあるので、

$$P(A) = \frac{3000}{10000}, P(\bar{A}) = \frac{7000}{10000} \quad (58)$$

とわかる。同じく、「spam メールのうち 80 %には本文に”claim”という単語が含まれている。一方、ham の 5 %には本文に”claim”という単語が含まれている。」という文から、

$$P_A(B) = \frac{80}{100}, P_{\bar{A}}(B) = \frac{5}{100} \quad (59)$$

もわかる。問題文から直ちに得られる (58), (59) から $P_B(A)$ の表式を導くことが、ここでの目標である。

我々の目的地である $P_B(A)$ は、条件付き確率の公式 (53) より、

$$P_B(A) = \frac{P(A \cap B)}{P(B)} \quad (60)$$

と表された。この式の右辺の分子は確率の乗法定理 (55) より、

$$P(A \cap B) = P(A)P_A(B) \quad (61)$$

と書ける。(60) の右辺分母は、

$$\begin{aligned} P(A) &= P(A \cap B) + P(A \cap \bar{B}) \\ &= P(A)P_A(B) + P(\bar{A})P_{\bar{A}}(B) \end{aligned} \quad (62)$$

と (かなりテクニカルに) 変形できる。1 行目では、 $A = A \cap U = A \cap (B \cup \bar{B}) = (A \cap B) \cup (A \cap \bar{B})$ という (一見複雑そうに見える) 変形をおこなったが、ベン図を描くのと直感的に意味を理解できる (図 6)。2 行目を導くために確率の乗法定理 (55) を利用した。(61) と (62) を (60) に代入すれば、

$$P_B(A) = \frac{P(A)P_A(B)}{P(A)P_A(B) + P(\bar{A})P_{\bar{A}}(B)} \quad (63)$$

が得られる。右辺は、問題文からすぐに計算できる (58), (59) だけで構成されている。先ほど提示した我々の目標が達成された! (63) を「事象 B が起こったことを知って、それが原因 A から起こったと考えられる『原因の確率』」の公式だと認めしまい、(58) や (59) の計算結果を代入すれば、機械的に原因の確率は計算できる (実際にやってみよ)。ただし、今回の問題では表を書けば秒殺であった。

答えが長くなってしまったので、ここまでのストーリーをまとめておく。迷子になってしまった人はこれをロードマップに読み返してほしい。

1. 我々がはじめに与えられていた情報は $P(A)$, $P(\bar{A})$, $P_A(B)$, $P_{\bar{A}}(B)$ の 4 つの確率である ((58), (59))。
2. これらの確率を利用して $P_B(A)$ を計算したい。
3. 条件付き確率の公式 (53) や確率の乗法定理 (55) を利用して $P_B(A) = \frac{P(A)P_A(B)}{P(A)P_A(B) + P(\bar{A})P_{\bar{A}}(B)}$ が導けた!

例 14 (おまけの問題) ジョーカーを除いたトランプ 52 枚をよくシャッフルした後、伏せて机に置く。その束から無作為に 1 枚選んで表を開けずに適当な箱に封印する。残りの束から無作為に 1 枚引いて表を見たところ、ハートの A であった。このとき、はじめに引いて封印していたカードが赤か黒かで賭けをしたい。

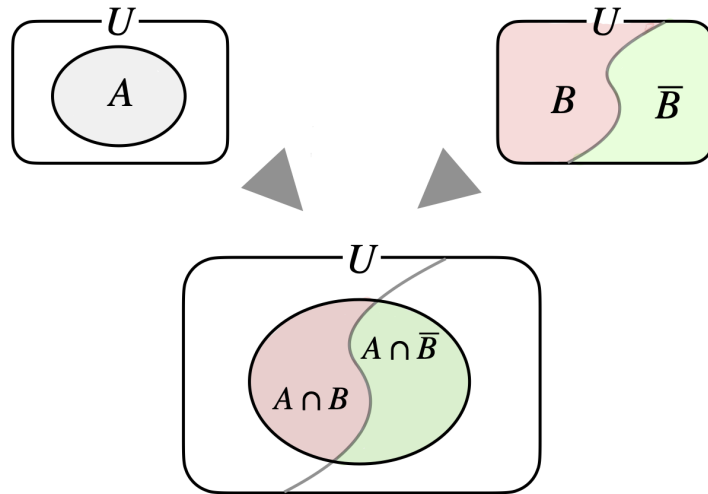


図 6: 左上の図は全事象 U と事象 A のベン図。右上の図は全事象 U と事象 B とその余事象 \bar{B} のベン図。これらを重ね合わせたものが下の図である。ここに示されているように、 A は $A \cap B$ と $A \cap \bar{B}$ の 2 つに分断できる。したがって、 $n(A) = n(A \cap B) + n(A \cap \bar{B})$ であり、(62) の 1 行目の変形: $P(A) = \frac{n(A)}{n(U)} = \frac{n(A \cap B) + n(A \cap \bar{B})}{n(U)} = P(A \cap B) + P(A \cap \bar{B})$ が導かれる。

和さん 「2 枚目のカードが赤だったから、1 枚目にはきっと黒のカードを引いただろう。」

健さん 「1 枚目を引く段階では、黒と赤が半々なのでどちらを引く確率は $1/2$ である。それを封印した後はその箱に手を加えていないのだから、確率は $1/2$ のまま変わるわけがない。したがって、赤と黒のどちらを選んでも確率は同じである。」

2 人の意見のうち賛成できるのはどちらか。なぜそう言えるのか。また、反対意見の相手を説得せよ。

(答え) 結論を言うと、2 枚目が赤いカードであれば 1 枚目に黒いカードが選ばれていた確率の方が高い。

実際に計算してみよう。1 枚目に赤を引くという事象を X_R 、1 枚目に黒を引くという事象を X_B 、2 枚目に赤を引くという事象を Y_R 、2 枚目に黒を引くという事象を Y_B とする。2 枚目が赤カードであったときに、1 枚目に封印していたのが黒カードであった確率

は、例 13 と同様の方法を用いて、

$$\begin{aligned}
 P_{Y_R}(X_B) &= \frac{P(Y_R \cap X_B)}{P(Y_R)} \\
 &= \frac{P(X_B)P_{X_B}(Y_R)}{P(X_R \cap Y_R) + P(X_B \cap Y_R)} \\
 &= \frac{P(X_B)P_{X_B}(Y_R)}{P(X_B)P_{X_B}(Y_R) + P(X_R)P_{X_R}(Y_R)} \\
 &= \frac{(1/2) \times (26/51)}{(1/2) \times (26/51) + (1/2) \times (25/51)} \\
 &= \frac{26}{51} \geq \frac{1}{2}
 \end{aligned} \tag{64}$$

とわかる。確かに確率は $1/2$ より大きく、1 枚目に黒いカードを引いた確率の方が大きいと言える。

健さんは、1 枚目のカードを封印したあとは何をしてもはじめの確率は揺るがないなどと息巻いていたが、彼にはぜひ 3.1 節の説明を読んでもらいたい。選んだ生徒が文系である確率を計算する際、「爬虫類を飼っているかいなか」という情報を知っているかいなかで、得られる確率の値が異なっていたことを思い出そう。ここでおこなったトランプの議論でも同様に、2 枚目に引いたカードが何色か、という情報を得たことにより、はじめに引いたカードが赤である確率（黒である確率も同様）の値は $1/2$ から変わることがある。