

数学 A 1 章「場合の数と確率」 2 節「確率」 カンペ

2024 年 5 月 5 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 \emptyset 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 <https://github.com/sodesudesu/suA.git> から見られる。

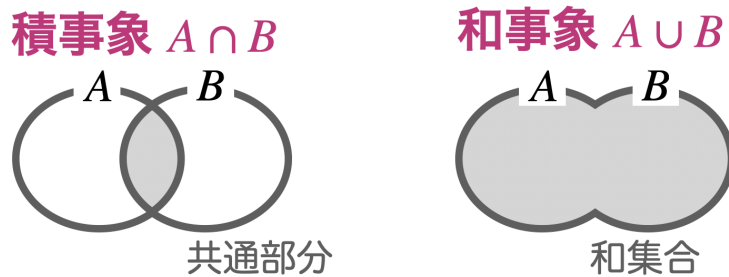


図 1: 積事象 $A \cap B$ のベン図 (左) と和事象 $A \cup B$ のベン図 (右)。

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。

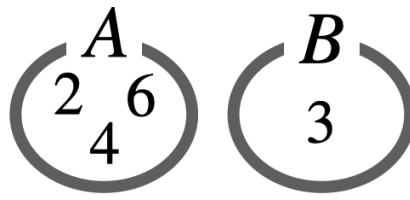


図 2: 事象 $A = \{2, 4, 6\}$ と事象 $B = \{3\}$ のベン図。 $A \cap B = \emptyset$ つまり、ベン図でこれらの集合は交わらない。3 つ以上の事象についても、ベン図を描いて交わらないならば、それらは互いに排反である。

■**排反事象** サイコロを 1 回投げる試行について（再び）考える。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)	(0, 2)	(1, 2)	(2, 2)
		(1, 0)	(1, 1)	(2, 1)	
			(2, 0)		

のように0~4の各和に対し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。上の例で言えば、

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！

となる。

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 枚とも表になる確率を計算せよ。

（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。根元事象は、(硬貨 A の表裏, 硬貨 B の表裏) と表すと、

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

のように選べば、これらは同様に確からしい ($\{2 \text{ 枚とも表}\}, \{1 \text{ 枚は表, } 1 \text{ 枚は裏}\}, \{2 \text{ 枚とも裏}\}$, 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 前節の復習だが、事象 A の確率は、

$$P(A) = \frac{\text{事象 } A \text{ の要素の個数 } n(A)}{\text{全事象 } U \text{ の要素の個数 } n(U)} \quad (15)$$

で計算された（この式は (8) と同じもの）。

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、

$$n(\emptyset) \leq n(A) \leq n(U) \quad (16)$$

が成り立つ。各辺を $n(U)$ で割ると、

$$\frac{n(\emptyset)}{n(U)} \leq \frac{n(A)}{n(U)} \leq \frac{n(U)}{n(U)} \quad (17)$$

すなわち、

$$0 \leq P(A) \leq 1 \quad (18)$$

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求め方 (15) を使った。(17) と (18) 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、

$$n(A \cup B) = n(A) + n(B) \quad (19)$$

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 (15) と (19) を用いることで、和事象 $A \cup B$ の確率 $P(A \cup B)$ は、

$$\begin{aligned} P(A \cup B) &= \frac{n(A \cup B)}{n(U)} \\ &= \frac{n(A) + n(B)}{n(U)} \\ &= \frac{n(A)}{n(U)} + \frac{n(B)}{n(U)} \end{aligned} \quad (20)$$

となるので、

$$P(A \cup B) = P(A) + P(B) \quad (21)$$

という表式が得られる。(21) は**確率の加法定理** と呼ばれる。

(21) 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。

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 (21) 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

$$n(A \cup B) = n(A) + n(B) - n(A \cap B) \quad (22)$$

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 (15) とあわせて、

$$\begin{aligned} P(A \cup B) &= \frac{n(A \cup B)}{n(U)} \\ &= \frac{n(A) + n(B) - n(A \cap B)}{n(U)} \\ &= \frac{n(A)}{n(U)} + \frac{n(B)}{n(U)} - \frac{n(A \cap B)}{n(U)} \end{aligned} \quad (23)$$

であるから、和事象の確率の公式

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

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、(24) の右辺第 3 項が消えて、

$$P(A \cup B) = P(A) + P(B) \quad (25)$$

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である。つまり、確率の加法定理 (25) 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 (24) に含まれている。

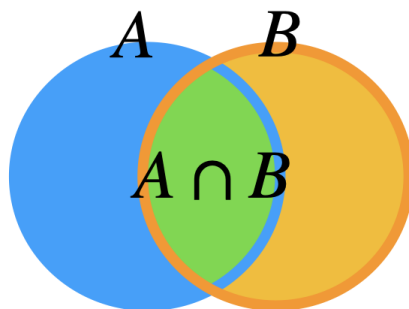


図 3: 互いに排反でない事象 A と事象 B のベン図。和事象 $A \cup B$ の要素の個数はどう表されるか。各集合の要素の個数を $n(A) + n(B)$ のように足し算してしまうと、共通部分 $A \cap B$ (図の緑色の部分) を余分に数え上げることになる。したがって、2つの集合の要素数の和 $n(A) + n(B)$ から、重複して数えてしまった $n(A \cap B)$ を引いたものが、和事象 $A \cup B$ の要素の個数となる： $n(A \cup B) = n(A) + n(B) - n(A \cap B)$ 。

■余事象の確率

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 問だった。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（多いですね〜）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、

$$P(A) = \frac{65}{100} = \frac{13}{20} \quad (26)$$

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

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

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、

$$P(A) = \frac{10}{25} = \frac{2}{5} \quad (27)$$

となる。

(26) も (27) 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。(27) 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、(26) 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)。(27) 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 (27) の表式を見ると、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、

$$P_B(A) = \frac{n(A \cap B)}{n(B)} \quad (28)$$

と書ける。この式の右辺の分母分子を $n(U)$ で割ると、

$$P_B(A) = \frac{P(A \cap B)}{P(A)} \quad (29)$$

という表式も得られる。我々は、(28) や (29) 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 4 (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) (28) より、

$$P_{\bar{A}}(B) = \frac{n(\bar{A} \cap B)}{n(B)} = \frac{3}{7} \quad (30)$$

となる。

3.2 確率の乗法定理

(29) の両辺を $P(A)$ 倍すると、

$$P(A \cap B) = P(A)P_B(A) \quad (31)$$

が得られる。これは、**確率の乗法定理**と呼ばれる。

例 5 (確率の乗法定理 (31) 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 (31) より、1 回目も 2 回目も赤玉を取り出す確率は、

$$\begin{aligned} P(1 \text{ 回目赤} \cap 2 \text{ 回目赤}) &= P(1 \text{ 回目赤}) \times P_{1 \text{ 回目赤}}(2 \text{ 回目赤}) \\ &= \frac{3}{5} \times \frac{1}{2} \\ &= \frac{3}{10} \end{aligned} \tag{32}$$

と計算される。

3.3 原因の確率

■本節でやりたいこと

例 6 (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 メールといひ、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	1000

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。