

MAT220 第5回講義：確率入門

Contents

1	確率とは何か？	2
1.1	不確実性	2
1.2	確率を定義する	3
1.3	古典的確率	3
1.4	数え上げによって確率を割り当てる	4
2	確率の規則	7
2.1	単純事象、標本空間、事象	7
2.2	確率は密かに集合論である	8
3	複合確率	9
3.1	複合事象を作る	9
3.2	補集合： $\text{not } A$	10
3.3	共通部分： $A \text{ and } B$	11
3.4	条件付き確率	11
3.5	独立事象と従属事象	14
3.6	独立事象の乗法公式	14
3.7	和集合： $A \text{ or } B$	16
3.8	互いに排反な事象	17

第4回講義では、確率論に必要となる集合論と数え上げの方法を導入した。今日は、それらの考え方を使って、確率を数学的に記述する。中心となる考え方は、事象は可能な結果の部分集合として表され、その事象の確率は 0 から 1 までの数である、ということである。古典的確率論を扱うとき、不確実性に関する問題は、数え上げの問題に変えることができる。この講義では、この考え方を詳しく見ていく。

1 確率とは何か？

1.1 不確実性

このコースの最初の数回の講義では、データと統計の基本的な言葉を導入した。データは、たいてい複雑な現実世界の状況から集められ、それから数学的に分析できるものへと変換される。今日は、別の種類の複雑さを扱う。それは不確実性である。

次の2つの文を考えよう。

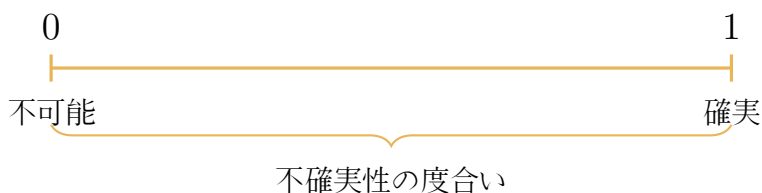
$$2 + 2 = 4 \quad \text{and} \quad 2 + 2 = 5.$$

知っての通り、最初の文は真であり、2つ目の文は偽である。これは明らかであるだけでなく、一義的である。つまり、これらの文が別のあり方をする余地はない。これに対して、次の2つの文を考えよう。

「Dr. O'Connell のミドルネームは John である」と「Dr. O'Connell のミドルネームは Jennifer である」。

分かるように、これらの非数学的な文は、もはやそれほど明確ではない。もし Dr. O'Connell のことを知らなければ、彼のミドルネームについて不確実である。しかし、私のミドルネームを知らなくても、イングランド人のミドルネームが Jennifer であるのは珍しく、John である方がもっともらしい、と想像することはできる。したがって、これらの文の正確な真偽をはっきり指し示すことはできなくても、それらのもっともらしさを比較することはできる。

確率論は、この不確実性を数によって捉えようとする。現実には、世界には純粋に論理的ではない不確実なものがたくさんある。そして確率論は、そのような信念を数値化し、比較できるようにする方法を与えてくれる。数 0 を「絶対に不可能」、数 1 を「絶対に確実」と解釈するなら、私たちの不確実な文はすべて、その間のどこかに位置することになる。



基本的な考え方

確率とは、不確実性の数学である。確率は、事象に数を割り当て、それらの事象がどれくらい起こりやすいかを比較する方法を与えてくれる。

注意すべきこととして、確率は真理のスペクトルではない。¹むしろ、それは確実性のスペクトルである。したがって、確率の値が 0.5 であるということは、何かが「半分真である」という意味ではない。そうではなく、その文が真である可能性を 50% と見積もっている、という意味である。

¹興味のある人へ：実はこれを扱う論理体系があり、「ファジー論理」と呼ばれる。しかし、これは確率とは違う。

1.2 確率を定義する

数学的には、確率とは事象に数を割り当てる関数である。文字 A が、起こるかもしれないし起こらないかもしれない何かを表すとき、 A が起こる確率を $P(A)$ と書く。この意味で、文字 P は、現実世界の事象を食べて、その確実性を表す数値を吐き出す小さな機械だと考えることができる。



一般に、このような「起こりうるもの」を事象と呼び、 A, B, C などの大文字で表す。形式的には、事象に割り当てられる確率は、次のようにまとめられる。

確率の値

任意の事象 A に対して、その確率は

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

を満たす。ここで $P(A) = 0$ は A が不可能であることを意味し、 $P(A) = 1$ は A が確実であることを意味する。

確率は 0 から 1 の間の線分上のどこかに対応するので、異なる事象の不確実性の度合いを直接比較できる。2つの事象 A と B があり、それらの確率 $P(A)$ と $P(B)$ が分かっているとしよう。もし $P(A) < P(B)$ なら、事象 B は事象 A よりも起こりやすい。例えば、

$$P(A) = 0.2 \quad \text{and} \quad P(B) = 0.7$$

なら、 B の方が A よりも起こりやすい。なぜなら、 $P(B)$ は $P(A)$ よりも絶対的な確実性である 1 に近いからである。

1.2.1 確率の表し方

任意の確率は 0 から 1 までの数なので、いくつかの同値な形で書き表すことができる。例えば、ある事象 A の確率が 50% であるなら、次のどの形でも書くことができる。

$$P(A) = \frac{1}{2} = 0.5 = 50\% = \frac{250}{500}.$$

ここで $\frac{1}{2}$ は約分された分数、0.5 は小数、50% は百分率、 $\frac{250}{500}$ は約分されていない分数である。最後の場合、例えば可能性が 500 個あるシステムを扱っていて、そのうち 250 個の結果が有利である、という状況かもしれない。約分されていない分数 $\frac{250}{500}$ はこの情報を伝えてくれるが、約分された分数 $\frac{1}{2}$ はその情報を伝えない。

1.3 古典的確率

この講義では、主に古典的確率に焦点を当てる。これは、可能な結果が有限個であり、それぞれが同じように起こりやすい設定である。

定義：古典的確率

可能な結果が有限個であり、それぞれの結果が同じように起こりやすいとき、その確率問題は古典的確率の問題である。

例としては、次のようなものがある。

- 公平なサイコロを振ること、
- 公平なコインを投げること、
- よくシャッフルされたデッキからカードを配ること、
- 真偽問題の答えを選ぶこと。

この設定では、確率は通常、数え上げによって計算できる。

1.3.1 ランダム性についての注意

真のランダム性という問題は、非常に哲学的である。世界は完全に決定論的であり、すべては過去の出来事と物理法則によって決まっている、と考える人もいる。一方で、例えば量子力学における観測のように、本当にランダムなものが存在すると考える人もいる。しかし、そこでも議論は続いている。いずれにせよ、私たちの目的にとって、真のランダム性が存在するかどうかは重要ではない。このコースでは、不確実性を数量化するために確率を使う。したがって、現実には物事が真にランダムでないとしても、私たちにはランダムに見えることがある。

このコースにおけるランダム性

このコースでは、実験が行われる前に確実に予測できない結果を、ランダムな結果とみなす。

この実用的な解釈で、このコースで行う種類の確率計算には十分である。

1.4 数え上げによって確率を割り当てる

古典的確率を扱うとき、個々の事象の確率は、「有利な場合」の数を、すべての可能性の集まりと比較する比として計算できる。任意の事象に対して、次の公式で確率を割り当てることができる。

$$\text{事象の確率} = \frac{\text{その事象に有利な結果の数}}{\text{結果の総数}}.$$

言い換えると、その事象を真にする結果の数を数え、それを可能な結果の総数で割る。この分数では、分子は常に分母以下なので、比を取ると常に 0 から 1 までの値が得られる。

古典的確率の黄金公式

すべての結果が同じように起こりやすいなら、

$$P(A) = \frac{A \text{ に有利な結果の数}}{\text{結果の総数}}.$$

ここから、この公式を使って基本的な確率を計算する方法を示す。

1.4.1 個々の結果の確率

任意の個々の結果の確率は、この公式で求めることができる。例えば：

- 公平なコインでは、コインが表になる確率は $\frac{1}{2}$ である。なぜなら、可能性は2つしかなく、そのうち1つが表だからである。
- 公平なサイコロを振るとき、5 が出る確率は $\frac{1}{6}$ である。なぜなら、サイコロには 5 が1つだけ書かれており、数字は全部で 6 個あるからである。
- シャッフルされたデッキからランダムにカードを1枚引くとき、スペードのエースを引く確率は $\frac{1}{52}$ である。なぜなら、デッキには 52 枚のカードがあり、そのうちスペードのエースは1枚だけだからである。

一般に、古典的確率における任意のシステムでは、上の公式から次が分かる。

$$\text{任意の個々の結果の確率} = \frac{1}{\text{結果の総数}}.$$

1.4.2 より複雑な事象の確率

黄金公式を使うと、より複雑な確率も計算できる。先ほど述べたように、これは可能性の空間を調べ、その事象を真にする結果をすべて数え上げることによって行える。例として、公平なサイコロを振っているとしよう。事象 A を

$$A = \text{「偶数が出る」},$$

とし、事象 B を

$$B = \text{「5 より小さい数が出る」}$$

とする。

これらの確率を求めるには、それぞれの「有利な結果」の数を計算し、それを 6 で割ればよい。なぜなら、サイコロには 6 つの面があるからである。事象 A については、

$$P(A) = \frac{\text{偶数はいくつあるか}}{\text{全部でいくつの数があるか}} = \frac{3}{6} = \frac{1}{2},$$

となり、同様に B の確率は

$$P(B) = \frac{\text{5 より小さい数はいくつあるか}}{\text{全部でいくつの数があるか}} = \frac{4}{6} = \frac{2}{3}.$$

である。

1.4.3 例：おもちゃの棚

熱心すぎるいところが、棚の中に 58 個の異なるおもちゃを持っていて、そのうち1つをランダムに選んであなたに見せるとしよう。彼がそれをする前には、あなたはどれを選ぶのか分からないので、あなたの視点から見ると、その選択は本質的にランダムである。下のおもちゃの棚を見てみよう。



事象 A を「選ばれたおもちゃが青い車である」とする。上の画像を調べると、棚には青い車が2つあることが分かる。したがって、黄金公式より

$$P(A) = \frac{2}{58}.$$

事象 B を「選ばれたおもちゃが4段目から来る」とする。上の画像では、4段目には8個のおもちゃがあると数えられるので、事象 B は8個の結果を含む。したがって、

$$P(B) = \frac{8}{58}.$$

どちらの場合も計算は同じである。有利なおもちゃを数え、それをおもちゃの総数で割る。

演習：真偽問題の解答列

3問の宿題が真偽問題であるとする。正しい答えをランダムに選ぶために、Dr. O'Connell は可能なすべての真偽列を書き出し、そのうち1つを帽子から選ぶ。3つの T/F が並ぶ可能な列は8個ある。

TTT, TTF, TFT, FTT, TFF, FTF, FFT, FFF

- (a) 3つの答えがすべて False である確率は何か。
- (b) ちょうど2つの答えが False である確率は何か。
- (c) 学生が3つの答えをすべて当てずっぽうで選ぶ場合、3問すべて正解する確率は何か。

解答

これら3つの問題はすべて、先ほどの黄金公式を使って答えられる。まず、可能な真偽列は8個あることに注意する。

(a) 有利な結果は FFF の1つだけである。したがって、

$$P(3つすべて False) = \frac{1}{8}.$$

(b) 有利な結果は TFF, FTF, FFT である。したがって、

$$P(\text{ちょうど2つが False}) = \frac{3}{8}.$$

(c) 正しい列はちょうど1つだけなので、

$$P(3問すべて正解) = \frac{1}{8}.$$

2 確率の規則

2.1 単純事象、標本空間、事象

ここで、確率の語彙をもう少し丁寧に整理する。最も重要な対象は、単純事象、標本空間、事象である。

重要な定義

- (1) **単純事象**とは、実験の結果をちょうど1つだけ含む事象である。
- (2) **標本空間**は、通常 S と書かれ、可能なすべての結果の集合である。
- (3) **事象**とは、標本空間の任意の部分集合であり、1つの結果を含むことも、多くの結果を含むことも、まったく含まないこともある。
- (4) 事象 A の確率は $P(A)$ と書く。これは 0 から 1 までの数である。
- (5) すべての単純事象の確率の和は常に 1 である。

これらの定義を見るために、おもちゃの棚の例に戻ろう。この場合、標本空間は棚の中の 58 個すべてのおもちゃの集合である。

$$S = \{\text{棚の中のすべてのおもちゃ}\}.$$

この例では、単純事象とは1つのおもちゃを選ぶことであり、例えば「この特定のおもちゃを選ぶ」というようなものである。数学的には、これは S の部分集合で、大きさが1である任意のものである。一般の事象は S の任意の部分集合であり、これらの部分集合は「4段目からおもちゃを選ぶ」や「青い車を選ぶ」といった、より一般的な文を表す。

強調しておく、事象とは標本空間の部分集合である。例えば、上のおもちゃの棚の例では、事象 A を「青い車を選ぶ」と考え、その確率を $\frac{2}{58}$ と求めた。事象 A の数学的な定義は、棚から取られた2つの青い車を含む集合である。

2.2 確率は密かに集合論である

もう1つの例を考えよう。Dr. O'Connell のクラスに 11 人の学生がいるとする。彼は学生を2列に並べる。前列に 6 人、後列に 5 人である。



議論のために、Dr. O'Connell が前列からランダムに1人を選ぶ確率は何だろうか。各選択が同じように起こりやすいと仮定すると、黄金公式より、

$$P(A) = \frac{\text{前列の学生の数}}{\text{学生の総数}} = \frac{6}{11}.$$

先ほどの用語で言えば、標本空間 S はすべての学生の集合に対応し、事象 A は前列の学生からなる部分集合である。

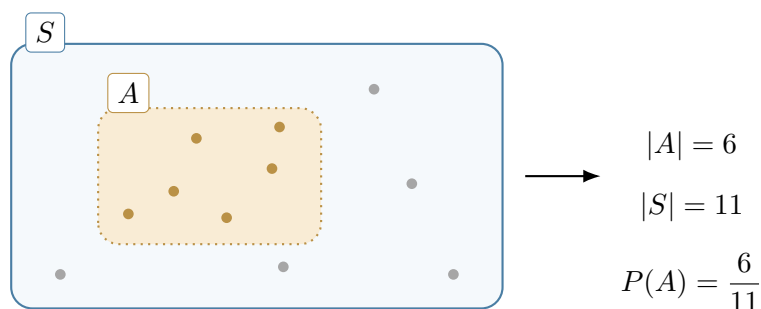
$$A = \{ \text{前列の6人} \}$$

$$S = \{ \text{前列の6人, 後列の5人} \}$$

すぐに数えると、 $|A| = 6$ かつ $|S| = 11$ である。この比を取ると確率 $P(A)$ が得られる。したがって、集合論の言葉では、黄金公式は次のように書ける。

$$P(A) = \frac{|A|}{|S|}.$$

これは、事象 A の確率を計算するには、部分集合 A の大きさを数え、それを標本空間 S の大きさを割ればよい、ということである。さらに抽象的に言えば、下の図のような数え上げを行っている。



今後の参照のために、この公式をもう一度述べておく。

古典的確率の黄金公式の言い換え

すべての結果が同じように起こりやすいなら、事象 A の確率は

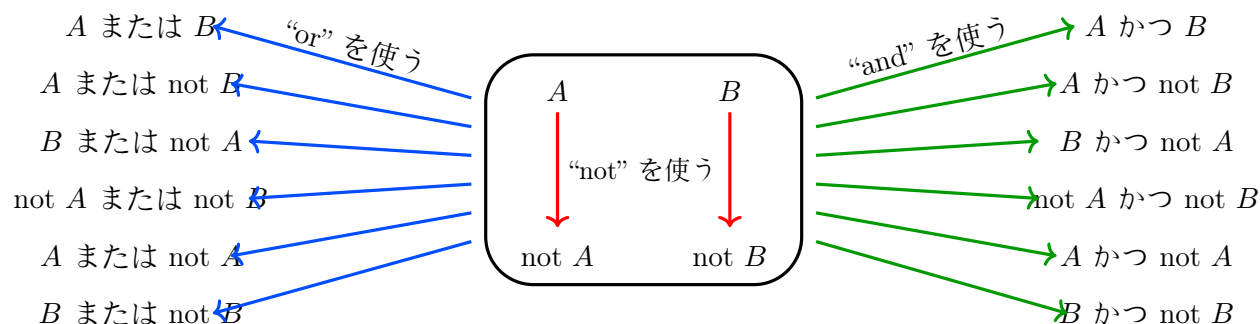
$$P(A) = \frac{|A|}{|S|},$$

すなわち、 A の濃度を標本空間 S の濃度で割ることによって求められる。

3 複合確率

3.1 複合事象を作る

2つの事象 A と B があるとしよう。ここまでの議論から、 $P(A)$ と $P(B)$ の求め方は分かっている。 A または B に対応する結果の数を数え、それを標本空間の大きさに割ればよい。実は、 A と B から、*not*、*and*、*or* という3つの論理的な部品を使って、多くの別の事象を作ることができる。例えば、これら3つの言葉を A と B に適用すると、次のような多くの文を作ることができる。それらはすべてそれ自体として事象である。



この図は、第4回講義で集合論を議論したときに見た図とかなり似ている。実際、それは同じ図である。*not*、*and*、*or* という3つの操作は、次のように集合論へ直接翻訳できる。

言葉	集合の操作	意味
<i>not A</i>	A^c	A の補集合
A and B	$A \cap B$	A と B の共通部分
A or B	$A \cup B$	A と B の和集合

前の節によれば、任意の事象の確率は、その部分集合の大きさを数え、それを標本空間の大きさに割ることで計算できる。上の3つの文の場合、これによって次の3つの新しい公式が得られる。

$$P(\text{not } A) = \frac{|A^c|}{|S|},$$

$$P(A \text{ and } B) = \frac{|A \cap B|}{|S|},$$

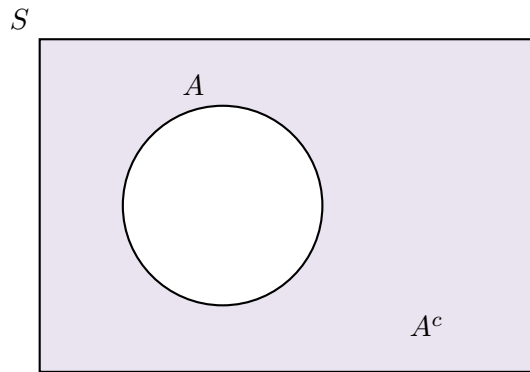
そして

$$P(A \text{ or } B) = \frac{|A \cup B|}{|S|}.$$

したがって、これらの集合の大きさを求めることができれば、対応する複合確率を計算する公式を作ることができる。ここから、これらを1つずつ説明する。

3.2 補集合：not A

補集合 A^c とは、標本空間の中で、部分集合 A に入っていないすべてのものである。下の図では、補集合は塗られた領域である。



古典的確率の公式によれば、not- A の確率を計算するには、補集合 A^c の大きさを数えればよい。

$$P(\text{not } A) = \frac{|A^c|}{|S|}.$$

A と A^c を合わせると標本空間全体を覆うので、

$$|A| + |A^c| = |S|.$$

したがって、

$$|A^c| = |S| - |A|.$$

これを確率の公式に代入すると、

$$P(A^c) = \frac{|A^c|}{|S|} = \frac{|S| - |A|}{|S|} = \frac{|S|}{|S|} - \frac{|A|}{|S|} = 1 - P(A).$$

補集合の規則

任意の事象 A に対して、

$$P(\text{not } A) = 1 - P(A).$$

例。真偽列の例で、 E を「FFF が選ばれる」という事象とする。8 個の列がすべて同じように起こりやすいので、

$$P(E) = \frac{1}{8}.$$

補集合 E^c は「FFF が選ばれない」という事象である。有利な結果は、FFF 以外のすべて、つまり 7 個である。したがって、

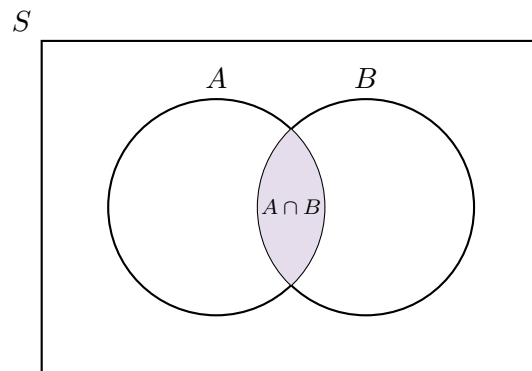
$$P(E^c) = \frac{7}{8}.$$

同値に、補集合の規則を使うと、

$$P(E^c) = 1 - P(E) = 1 - \frac{1}{8} = \frac{7}{8}.$$

3.3 共通部分：A and B

A と B の共通部分は $A \cap B$ と書かれ、 A と B の両方に同時に入っている結果の集合である。視覚的には、2つの集合の重なりである。



黄金公式によれば、複合事象 A and B の確率は、それらの共通部分の大きさを含むはずである。

$$P(A \text{ and } B) = \frac{|A \cap B|}{|S|}.$$

残念ながら、 $|A|$ と $|B|$ だけから $A \cap B$ の大きさを求める一般的な公式は存在しない。2つの大きな集合は、大きく重なることもあれば、少しだけ重なることもあり、まったく重ならないこともある。したがって、 $P(A \text{ and } B)$ を計算するには、2つの事象がどのように関係しているかについての情報が必要である。

3.4 条件付き確率

$P(A \text{ and } B)$ の公式を取り、少し代数計算をしてみよう。うまく工夫して、 $\frac{|B|}{|B|}$ という形に変装した 1 を人工的に挿入する。基本的な代数計算を行うと、次が得られる。

$$P(A \text{ and } B) = \frac{|A \cap B|}{|S|} = \frac{|A \cap B|}{|S|} \cdot \frac{|B|}{|B|} = \frac{|A \cap B|}{|B|} \cdot \frac{|B|}{|S|}.$$

これにより、私たちはこの確率を2つの別の比の積として表すことができた。2つ目の分数は慣れたもので、 $P(B)$ の定義である。一方、1つ目の比は新しいものであり、条件付き確率と呼ばれる特別な確率である。

定義：条件付き確率

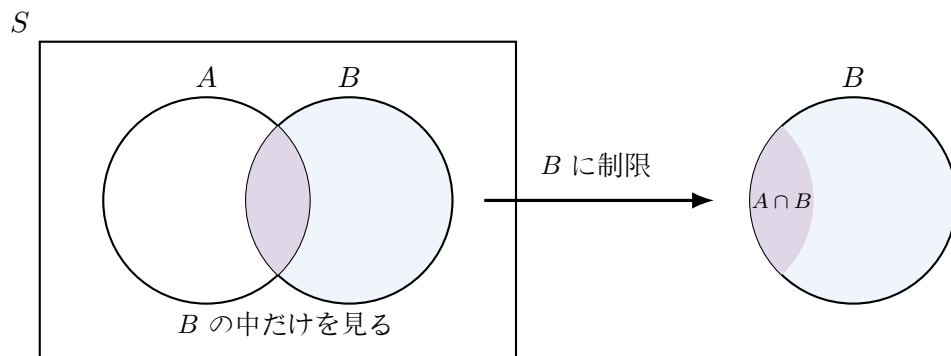
条件付き確率 $P(A | B)$ は、 B が起こったと仮定したときに A が起こる確率である。古典的確率では、

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

ただし $|B| \neq 0$ とする。

条件付き確率という概念は、最初は分かりにくく見える。そこで、集合論を使って何が起きているかを説明しよう。

先ほど述べたように、共通部分 $A \cap B$ の大きさを計算する完全な公式は存在しない。この大きさを求める1つの方法は、集合 B の中を見て、そこに見つかる A の事例をすべて数えることである。これは、標本空間全体を取り、すでに B に入っていないものを選択的に無視するようなものである。



この新しい設定に基づいて確率を計算しようとする、標本空間はもはや S ではなく、部分集合 B であることに気づく。この意味で、標本空間が今や B であるため、この制限された状況で A が起こる確率を数えようとするなら、黄金公式の分母には B を使う必要がある。

$$P(A, B \text{ の中で}) = \frac{A \text{ の元でもあり } B \text{ の元でもあるものの数}}{B \text{ の大きさ}}.$$

A の元であり、同時に B の元でもあるものは、ちょうど共通部分 $A \cap B$ に入る。したがって、上の確率から、先ほどの条件付き確率の公式を取り戻すことができる。

$$P(A | B) = P(A, B \text{ の中で}) = \frac{A \text{ の元でもあり } B \text{ の元でもあるものの数}}{B \text{ の大きさ}} = \frac{|A \cap B|}{|B|}.$$

例として、公平なサイコロを振るとしよう。

$$A = \{\text{偶数が出る}\} \quad \text{and} \quad B = \{4 \text{ より小さい数が出る}\}.$$

部分集合として書き出すと、事象 A と B は次のようになる。

$$A = \{2, 4, 6\}, \quad \text{and} \quad B = \{1, 2, 3\}.$$

そこで、あなたがサイコロを空中に投げたとする。サイコロが着地する前に、神が天から降りてきて、あなたの耳元で「結果は 4 より小さくなる」とささやいたと想像しよう。これは、その瞬間、4、5、6 が出ることはあり得ないと確実に知っている、ということである。この状況で、事象 A が起こる可能性を評価し、新しい確率を計算しようとするなら、可能性の空間を

$$\{1, 2, 3, 4, 5, 6\}$$

から、より小さな集合

$$\{1, 2, 3\}$$

へと切り詰めることになる。この小さな集合の中で、偶数は 2 だけである。したがって、

$$P(A | B) = \frac{1}{3}.$$

集合記法を使うと、

$$A \cap B = \{2\},$$

なので、

$$P(A | B) = \frac{|A \cap B|}{|B|} = \frac{1}{3}.$$

これらをまとめると、

$$P(B) = \frac{3}{6} = \frac{1}{2}, \quad P(A \text{ and } B) = \frac{1}{6}, \quad P(A | B) = \frac{1}{3}.$$

面白いことに、たまたま

$$\frac{1}{6} = \frac{1}{2} \cdot \frac{1}{3}$$

が成り立つ。しかし、これは偶然ではない。これは先ほど述べた一般公式の現れである。

$$P(A \text{ and } B) = \frac{|A \cap B|}{|S|} = \frac{|A \cap B|}{|B|} \cdot \frac{|B|}{|S|} = P(A | B)P(B).$$

一般の乗法公式

関連する条件付き確率が定義されているとき、任意の事象 A と B に対して、

$$P(A \text{ and } B) = P(B)P(A | B).$$

同値に、

$$P(A \text{ and } B) = P(A)P(B | A).$$

3.5 独立事象と従属事象

場合によっては、 $P(A \text{ and } B)$ の公式を大きく簡単にできる。その方法をきちんと説明するためには、まず独立事象と従属事象の区別をする必要がある。

独立事象と従属事象

A と B を事象とする。

- A と B は、一方が起こったと知っても他方の確率が変わらないとき、**独立**である。
- A と B は、一方が起こったと知ることによって他方の確率が変わるとき、**従属**である。

思い出すと、確率は不確実性の科学である。したがって、従属事象の背後にある考え方は、事象 B が起こったことを知ると、 A の確率に影響するかもしれない、ということである。これは、 B についての情報が何らかの形で事象 A に関係しているときに起こりうる。

先ほどのサイコロの例では、 A を「偶数が出る」、 B を「4 より小さい数が出る」とした。この2つの事象は従属である。なぜなら、 B が起こると知ることによって、私たちの確率が $\frac{1}{2}$ から $\frac{1}{3}$ に実際に変わったからである。言い換えると、一方の事象についての情報が、他方の可能性についての私たちの知識に影響したのである。

しかし、ある情報を知っても確率の評価を変える理由がない状況も考えられる。例えば、事象 C を「2 より大きい数が出る」としよう。このとき、部分集合としては

$$C = \{3, 4, 5, 6\}.$$

もし何らかの理由で、サイコロの結果が 2 より大きくなると絶対に分かっているなら、標本空間を

$$\{1, 2, 3, 4, 5, 6\}$$

から部分集合

$$\{3, 4, 5, 6\}$$

へと制限できる。この制限された集合は奇数2つと偶数2つから成るので、偶数が出る確率は依然として $\frac{1}{2}$ である。これは $P(A)$ と同じなので、 C が起こると知っても、このシステムについて私たちが持つ確実性の度合いは変わらない。したがって、事象 A と C は独立である。

記号で書くと、 A と B は

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

のとき独立であり、同様に、

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

である。

3.6 独立事象の乗法公式

A と B が独立なら、 $P(A | B) = P(A)$ を使って、 $P(A \text{ and } B)$ の公式を大幅に簡単にできる。代入を行うと、次の乗法公式が得られる。

独立事象の乗法公式

A と B が独立事象なら、

$$P(A \text{ and } B) = P(A)P(B).$$

例：2回のコイン投げ。 公平なコインを2回投げるとする。

A を「1回目のコイン投げが裏である」という事象、 B を「2回目のコイン投げが裏である」という事象とする。これらの事象は独立である。なぜなら、1回目の結果は2回目の結果に影響しないからである。

したがって、

$$P(A \text{ and } B) = P(A)P(B) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}.$$

同値に、標本空間は

$$S = \{HH, HT, TH, TT\},$$

であり、有利なのは TT だけである。

例：2つの別々のデッキ。 2つの別々のカードデッキをシャッフルするとする。

A を「デッキ1の一番上のカードがエースである」という事象、 B を「デッキ2の一番上のカードがエースである」という事象とする。これらの事象は独立であり、それぞれの確率は

$$\frac{4}{52} = \frac{1}{13}.$$

したがって、

$$P(A \text{ and } B) = \left(\frac{1}{13}\right) \left(\frac{1}{13}\right) = \frac{1}{169}.$$

同値に、一番上のカードの順序付き組は $52 \times 52 = 2704$ 通りあり、そのうち2枚ともエースである組は $4 \times 4 = 16$ 通りである。したがって、

$$\frac{16}{2704} = \frac{1}{169}.$$

次は従属事象の例である。

例：1つのデッキから2枚のエース。 52 枚のカードからなる1つのデッキをシャッフルするとする。

$$A = \{1\text{枚目のカードがエースである}\},$$

かつ

$$B = \{2\text{枚目のカードがエースである}\}.$$

これらの事象は従属である。なぜなら、1枚目がエースなら、2枚目を取る時には残り 51 枚の中にエースが 3 枚しか残っていないからである。計算すると、

$$P(A) = \frac{4}{52} = \frac{1}{13},$$

また

$$P(B | A) = \frac{3}{51}.$$

したがって、

$$P(A \text{ and } B) = P(A)P(B | A) = \frac{1}{13} \cdot \frac{3}{51} = \frac{1}{221}.$$

$P(A \text{ and } B)$ の計算方法

Step 1: 事象が独立かどうかを判断する。もし $P(A) = P(A | B)$ なら、 A と B は独立である。

Step 2: A と B が独立なら、

$$P(A \text{ and } B) = P(A)P(B).$$

を使う。

Step 3: 独立であっても従属であっても、任意の事象に対して、

$$P(A \text{ and } B) = P(A)P(B | A),$$

または同値に、

$$P(A \text{ and } B) = P(B)P(A | B).$$

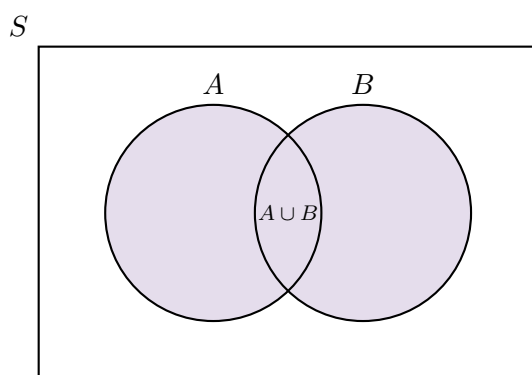
を使うことができる。

3.7 和集合： $A \text{ or } B$

Section 3.1 によれば、複合事象「 $A \text{ or } B$ 」は集合論的な和集合 $A \cup B$ に対応する。したがって、確率 $P(A \text{ or } B)$ を計算するには、この部分集合の大きさを数え、それを標本空間の大きさに割ればよい。

$$P(A \text{ or } B) = \frac{|A \cup B|}{|S|}.$$

A と B の和集合は $A \cup B$ と書かれ、 A に入っている、 B に入っている、または両方に入っている結果の集合である。



先ほど議論した共通部分の場合と同様に、一般には、 A と B の大きさだけから和集合 $A \cup B$ の大きさを求めることはできない。実際、前回の講義で見たように、素朴に

$$|A \cup B| = |A| + |B|$$

と書きたくなる。しかし、これは一般には間違っている。なぜなら、重なり $A \cap B$ がうっかり2回数えられてしまうからである。例えば、

$$A = \{1, 2, 3\}, \quad \text{and} \quad B = \{2, 3, 4\},$$

なら、それらの和集合は

$$A \cup B = \{1, 2, 3, 4\},$$

であり、その濃度は 4 である。しかし、

$$|A| + |B| = 3 + 3 = 6 \neq 4.$$

問題は、共通部分 $\{2, 3\}$ の2つ分が、和 $3 + 3 = 6$ の中に隠れていることである。

一般に、この問題を直すためには、共通部分を1つ分引けばよい。

$$|A \cup B| = |A| + |B| - |A \cap B|.$$

この観察を使うと、 $P(A \text{ or } B)$ を $P(A)$ 、 $P(B)$ 、 $P(A \text{ and } B)$ によって表す一般公式を導ける。

$$\begin{aligned} P(A \text{ or } B) &= \frac{|A \cup B|}{|S|} \\ &= \frac{|A| + |B| - |A \cap B|}{|S|} \\ &= \frac{|A|}{|S|} + \frac{|B|}{|S|} - \frac{|A \cap B|}{|S|} \\ &= P(A) + P(B) - P(A \text{ and } B). \end{aligned}$$

これは加法公式として知られており、専用の箱に入れる価値がある。

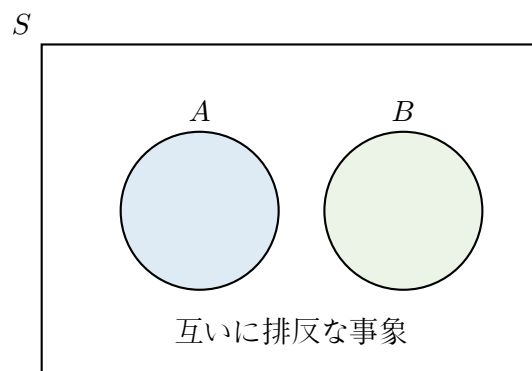
加法公式

任意の事象 A と B に対して、

$$P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B).$$

3.8 互いに排反な事象

ときには、2つの事象が同時には起こりえないことがある。例えば、1回のコイン投げで、同時に表と裏が出ることはない。このような事象を互いに排反であるという。視覚的には、互いに排反な事象には重なりがない。



定義：互いに排反な事象

2つの事象 A と B は、同時に起こりえないとき、互いに排反である。同値に、 $A \cap B = \emptyset$ であり、これは

$$P(A \text{ and } B) = \frac{|A \cap B|}{|S|} = \frac{0}{|S|} = 0.$$

を意味する。

この観察を使うと、 A or B の確率の公式において、互いに排反な事象では第3項が消えることが分かる。

$$\begin{aligned} P(A \text{ or } B) &= \frac{|A \cup B|}{|S|} \\ &= \frac{|A| + |B| - \cancel{|A \cap B|}}{|S|} \\ &= \frac{|A|}{|S|} + \frac{|B|}{|S|} - \cancel{\frac{|A \cap B|}{|S|}} \\ &= P(A) + P(B) - \cancel{P(A \text{ and } B)} \\ &= P(A) + P(B). \end{aligned}$$

これは、公式が A と B の確率の単純な足し算へと簡約されることを意味する。これは、互いに排反な事象に対する加法公式として知られている。

互いに排反な事象に対する加法公式

A と B が互いに排反なら、

$$P(A \text{ or } B) = P(A) + P(B).$$

例として、標準的な 52 枚のカードデッキをシャッフルするとしよう。次の事象を定義する。

$$A = \{\text{一番上のカードがエースである}\},$$

$$B = \{\text{一番上のカードがスペードである}\},$$

そして

$$C = \{\text{一番上のカードがキングである}\}.$$

事象 A と B は互いに排反ではない。なぜなら、スペードのエースはエースでもあり、スペードでもあるからである。しかし、 A と C は互いに排反である。なぜなら、一番上のカードが同時にエースでありキングであることはできないからである。

演習：エース、スペード、キング

次を計算しなさい。

(a) $P(A \text{ or } B)$.

(b) $P(A \text{ or } C)$.

解答

(a) 次が成り立つ。

$$P(A) = \frac{4}{52}, \quad P(B) = \frac{13}{52}, \quad P(A \text{ and } B) = \frac{1}{52}.$$

したがって、

$$P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B) = \frac{4}{52} + \frac{13}{52} - \frac{1}{52} = \frac{16}{52} = \frac{4}{13}.$$

(b) A と C は互いに排反なので、

$$P(A \text{ or } C) = P(A) + P(C).$$

したがって、

$$P(A \text{ or } C) = \frac{4}{52} + \frac{4}{52} = \frac{8}{52} = \frac{2}{13}.$$

まとめ：複合確率

この講義における主な複合事象は、*not*、*and*、*or* という言葉から作られる。集合記法では、これらは次のように対応する。

$$\text{not } A \longleftrightarrow A^c, \quad A \text{ and } B \longleftrightarrow A \cap B, \quad A \text{ or } B \longleftrightarrow A \cup B.$$

補集合については、

$$P(\text{not } A) = 1 - P(A)$$

を使う。

共通部分については、乗法公式

$$P(A \text{ and } B) = P(A)P(B | A),$$

または同値に、

$$P(A \text{ and } B) = P(B)P(A | B)$$

を使う。事象が独立なら、これは

$$P(A \text{ and } B) = P(A)P(B)$$

へと簡約される。

和集合については、加法公式

$$P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$$

を使う。事象が互いに排反なら、 $P(A \text{ and } B) = 0$ なので、これは

$$P(A \text{ or } B) = P(A) + P(B)$$

へと簡約される。