

MAT120：数量的推論 第10講ハンドアウト

確率入門

前のパートでは、モデリングを学びました。つまり、方程式を用いて変数どうしの関係を記述してきました。この講義では、新しい話題である確率に入ります。確率とは、不確実性の数学です。出来事がどれくらい起こりやすいかを語るための方法を与えてくれ、次に続く統計の単元の土台にもなります。

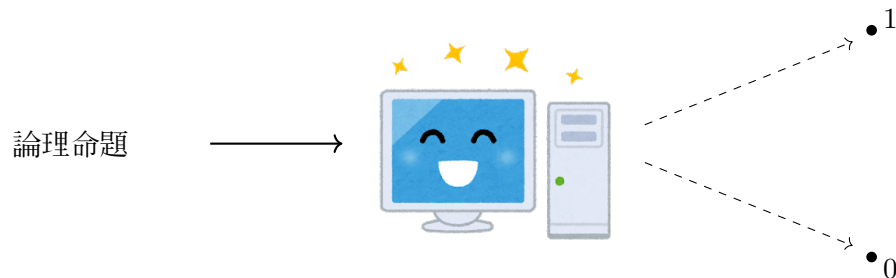
1 確率とは何か？

1.1 論理と不確実性

この授業の最初の数回では、論理を導入しました。論理とは、議論の構造を表現するために使える「形式言語」でした。ある議論は *valid* である、つまり正しい論理構造に従っていること、またある議論は *sound* である、つまりその前提が実際に現実世界で真であることを見ました。よい議論とは、*valid* でもあり *sound* でもあるものです。すなわち、真の前提から出発し、正しい論理の手順に従って、やはり真である結論に到達する議論のことです。

この考え方の根本にあるのは、あらゆる命題に「真理値」を割り当てられるという考えです。つまり、あらゆる論理命題は真か偽かのどちらかです。実際、第5講では、コンピュータもこれとよく似た仕方で動作していることを見ました。コンピュータは電気信号を「オン」か「オフ」（すなわち1と0）として解釈し、そこで行われる計算は、そうした入力信号の論理的組合せから作られる演算にすぎません。簡単に言えば、コンピュータは論理で考え、2進法で数えるのです。

コンピュータは一種の「論理機械」と考えることができます。どんな論理命題でも受け取って、それが真か偽かの評価を返してくれるからです。もし T を1に、 F を0に置き換えると、コンピュータによる命題の評価は次のように視覚化できます。



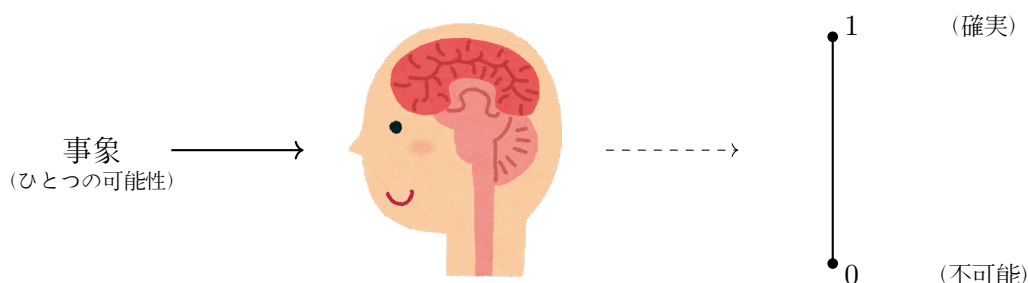
コンピュータは、数学や電気信号のような、明確で精密な対象を含む論理命題を扱うには非常に便利です。しかし、人間はこれとまったく同じ仕方で考えているわけではない、という意味もあります。現実世界では、ある命題が真か偽かを十分に知らないまま推論しなければならないことがしばしばあるからです。たとえば、「明日の朝は雨が降るだろう」という命題を考えてみましょう。もし私たちが気象学者なら答えを知っているかもしれませんが、そうではないので、正直に言えるのは「わからない、たぶん？」くらいです。この意味で、私たちの限界はその命題に真理値がないことではなく、どちらの真理値が正しいのかを私たちが知らないということです。言い換えれば、世界は私たちにとって不確かでありうるのに、それでも私たちはその不確実性のもとで推論しているのです。

1.2 確率論

確率論とは、不確実性を数学的に研究するものです。論理と異なるのは次の点です。論理は、与えられた前提から確実にどんな結論が従うかを教えてくれますが、確率は、情報が完全では

ないときに、さまざまな結果がどれくらい起こりやすいかを測る方法を与えてくれます。ある意味で、確率論は信念を数量化する学問だと考えることもできます。

この授業では、実験（サイコロを振るとか、コインを投げるといったもの）がいくつかの可能な結果をもつと考えます。しかし、実験が行われる前には、どの結果が実際に起こるのかはわかりません。確率は、それぞれの可能な結果に対して0 から1 の間の数を割り当て、その数をその結果がどれくらいもってもらいやすいかの尺度として解釈します。確率論はこの考えを出発点として、私たちがすでにもっている情報や、考えている実験の設定に基づいて、それらの数を決めるための巧妙な規則を発展させていきます。先ほどのコンピュータとは対照的に、不確かな人間の頭脳は、予測できない結果をむしろ連続的なスペクトラムとして捉えるかもしれません。



ここで注意してほしいのは、確率は0 と1 のあいだの「真理のスペクトラム」ではないということです。¹むしろ、確率は確かさのスペクトラムです。各事象 A （結果の集まり）に対して、0 から1 のあいだの数 $P(A)$ を割り当てることができ、 $P(A) = 0$ はその結果が不可能であること、 $P(A) = 1$ はそれが必ず起こることを意味し、その中間の値は、どの結果が起こるかについての不確実性の程度を表します。

1.3 ランダム性

この授業では、私たちの扱う系で実際に起こる結果はランダムであると仮定します。つまり、どの結果が起こるかを定める規則が（少なくとも私たちに利用可能な形では）存在しないということです。そこで次のように問いたくなります。そもそもランダム性とは何でしょうか。現実世界は本当にランダムなのでしょうか。たとえば、私が1 から10 のあいだの数をランダムに選んで、たまたま3 を選んだとします。なぜ私はその数を選んだのでしょうか。本当にランダムに選んだのか、それとも過去の経験の総和によって3 を選ぶように導かれただけなのでしょうか。

同じことを別の言い方で言えば、コインを投げたとします。その結果は本当にランダムなのでしょうか。もちろん何らかの意味では物理法則によって決まっているはずですが。もし私がコインを空中にはじき上げるときに加えた力についてすべてを知っていたら、その結果を確実に計算できるのでしょうか。

真のランダム性という問いは、深く哲学的な問題です。世界は完全に決定論的であり、あらゆるものが過去の出来事と物理法則によって決まると考える人もいます。他方で、本当にランダムなことがあると考える人もいます（たとえば量子力学における観測など）。しかし、そこで

¹興味のある人のために言えば、実際にそのようなことを行う論理理論は存在します。それは「ファジィ論理」と呼ばれます。しかし、これは確率ではありません。

もなお議論はあります。いずれにせよ、この授業の目的にとって、真のランダム性が存在するかどうかは重要ではありません。この授業では、不確実性を数量化するために確率を用います。したがって、現実には本当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のです。だからこそ、確率論は依然として有用です。不確実性のなかで意思決定をする助けになるのです。

2 確率の基本的性質

2.1 概念としての確率

数学的には、確率とはあるランダムな過程の可能な結果に数を割り当てる関数です。確率論では、結果の集まりを事象と呼び、通常これを大文字 A, B, C, \dots で表します。そして $P(A)$ を、事象 A の確率という意味で書きます。結果とはひとつの具体的な帰結（たとえば「4を出す」）のことです。すべての結果を集めた集合を標本空間と呼び、 S と書きます。そして事象 A は S の部分集合（結果の集まり）です。確率では次のような重要な基本ルールがあります。

- (1) 任意の事象 A について、常に $0 \leq P(A) \leq 1$ である。
- (2) $P(A) = 0$ は「不可能」を意味し、 $P(A) = 1$ は「確実」を意味する。
- (3) $P(A) < P(B)$ なら、事象 B のほうが事象 A より起こりやすい。

この授業では主に古典的確率を扱います。つまり、可能な結果が有限個あり、それらをすべて等しく起こりやすいとみなす場合です。例としては、サイコロを振ること、コインを投げること、カードを配ること、あるいは正誤問題を当てずっぽうで答えることなどがあります。

2.2 確率を割り当てる

ある実験が有限個の等確率な結果をもつとします（つまり、各結果が起こる確率がすべて同じであるということです）。「事象」とは結果の集まりにすぎないことを思い出してください。任意の事象について、次の公式で確率を割り当てることができます。

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

言い換えれば、有利な結果の数を数えて、全体の可能な結果の数で割るのです。この分数では、分子は常に分母以下ですから、その比は常に0から1のあいだの値になります。

また、上の公式は、各結果が等確率であるという仮定とも整合しています。ひとつの個別の結果は、大きさ1の集まりだと考えられるので、上の公式は次の特別な場合に縮約されます。

$$\text{任意の個別結果の確率} = \frac{1}{\text{結果の総数}}.$$

2.2.1 確率の表し方

確率は常に0から1のあいだの数ですが、同じ値をいくつかの等価な形で書くことができます。確率をどう書くかによって、系に関する追加情報が得られることもあります。たとえば、事象 A の確率が50%なら、次のように等価に書けます。

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

ここで $\frac{1}{2}$ は既約分数, 0.5は小数, 50%は百分率, $\frac{250}{500}$ は既約でない分数です。これらはすべて同じ確率を表しています。最後の形では, もしかすると500通りの可能性をもつ系を扱っていて, そのうち250通りが望ましい結果なのかもしれません。既約でない分数 $\frac{250}{500}$ はその情報を伝えますが, 既約分数 $\frac{1}{2}$ からはそれはわかりません。

2.3 例：おもちゃの戸棚

あなたのやたら熱心ないところは, 58個の異なるおもちゃが入った戸棚を持っていて, その中からひとつをランダムに選んで見せてくれることにしました。選ぶ前の時点では, どれを取るかはわかりませんし, どのおもちゃも同じ確率で選ばれると仮定します。



事象 A を「選ばれたおもちゃが青い車である」とします。上の図を見ると, 戸棚の中には青い車が2台あります。したがって,

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

事象 B を「選ばれたおもちゃが第4列から来る」とします。上の図を数えると, 第4列には8個のおもちゃがあります。したがって, 事象 B には8個の結果が含まれます。よって,

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

事象 C を「選ばれたおもちゃが緑色である」とします。戸棚の緑色のおもちゃを調べて数えると, 緑のおもちゃは10個あります。したがって,

$$P(C) = \frac{10}{58}.$$

2.4 練習：真偽の組合せ

Dr. O'Connellは, 3問のtrue/false問題からなる宿題を作成しました。彼はすべての可能な解答パターンを並べ, その中からひとつをランダムに選びます。

標本空間には, 等確率な8個の結果があります。

$$\{TTT, TTF, TFT, FTT, TFF, FTF, FFT, FFF\}.$$

練習1

次に答えなさい。

- (1) 3つの答えがすべてFalseである確率は何か。
- (2) ちょうど2つの答えがTrueである確率は何か。
- (3) 学生が完全に当てずっぽうで答えるとき、3問すべて正解する確率は何か。

解答

- (1) 有利な結果は1つだけ (FFF) なので、

$$P(3つともFalse) = \frac{1}{8}.$$

- (2) 有利な結果はTTF, TFT, FTT の3つなので、

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

- (3) 正しいパターンはちょうど1つだけなので、

$$P(3問すべて正解する) = \frac{1}{8}.$$

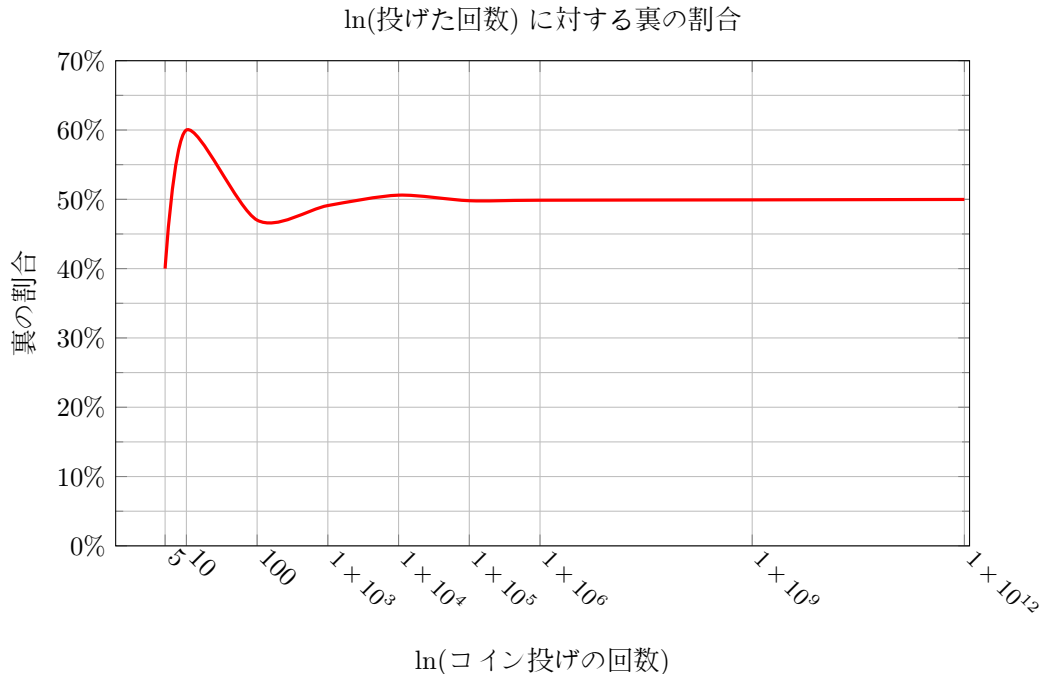
2.5 大数の法則

大数の法則とは、ある実験を何度も繰り返すと、その平均的な結果が理論上の平均に近づいていくというものです。例として、公平なコインを考えましょう。コインを投げて裏が出る確率は50%、すなわち $P(T) = \frac{1}{2}$ だと期待します。確率の解釈によれば、長い目で見れば裏はだいたい50% 出るはずだと期待するわけです。

これを実際に見てみましょう。コインを何度も何度も投げるとします。乱数生成器を使って得たコイン投げの標本における裏の割合の例を下に示します。

- 5回投げる：HHTTH。 $T = 2, H = 3$ なので、裏 = 40%。
- 10回投げる：HHTTHTHTTT。 $T = 6, H = 4$ なので、裏 = 60%。
- 100回投げる： $T = 47, H = 53$ なので、裏 = 47%。
- 1000回投げる： $T = 491, H = 509$ なので、裏 = 49.1%。
- 10000回投げる： $T = 5062, H = 4938$ なので、裏 = 50.6%。
- 100000回投げる： $T = 49872, H = 50128$ なので、裏 = 49.8%。
- 100万回投げる：裏 = 49.87%。
- 10億回投げる：裏 = 49.92%。
- 1兆回投げる：裏 = 49.982%。

これをグラフにすると次のようになります（ここではコイン投げの回数を同じ軸に収めるために対数尺度 $\ln(\cdot)$ をとっています）。



見てわかるように、最初のうちは結果にかなり大きなランダムな揺れがあります。しかし時間がたつにつれて、そうしたランダムな変動はおおむね打ち消し合い、裏の割合は理論平均である50%に収束していきます。これによって、確率のひとつの解釈が得られます。

確率の解釈

$P(A) = x\%$ と言うとき、それは「もしその実験を非常にたくさん繰り返したなら、事象Aに対応する結果の割合は $x\%$ にますます近づいていく」という意味です。

2.6 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 とは、起こりうるすべての結果（単純事象）の集合である。
- (3) 一般の事象A とは標本空間の任意の部分集合であり、1つの結果だけを含んでもよいし、多くの結果を含んでもよい。
- (4) 標本空間に含まれるすべての単純事象の確率の和は常に1である。
- (5) 結果が等確率である場合、確率は有利な結果の数を数えて、全結果数で割ることで求める。

これらの定義がどう働くかを見るために、例2.3のおもちゃの戸棚に戻りましょう。この場合、標本空間は

$$S = \{\text{戸棚の中のすべてのおもちゃ}\},$$

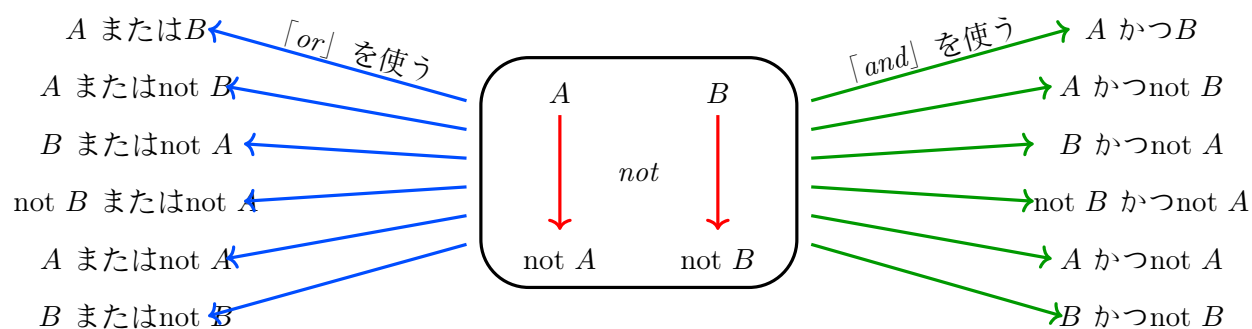
ですから、集合Sの大きさは58です。この例では、ひとつの結果とは個々のおもちゃが選ばれること、つまり「この特定のおもちゃが選ばれる」というようなものです。数学的には、これは大きさ1のSの部分集合にあたります。一般の事象とはSの任意の部分集合であり、「第4列からおもちゃを選ぶ」とか「青い車を選ぶ」といった、より一般的な主張を表します。

ここで強調しておきたいのは、事象そのものが標本空間の部分集合だということです。たとえば、Section 2.3 では「緑のおもちゃを選ぶ」という意味の事象 C を考え、その確率は $\frac{10}{58}$ だと求めました。事象 C の数学的定義は、戸棚の中のすべての緑のおもちゃからなる集合です。

3 複合確率

3.1 複合事象

2つの事象 A と B があるとします。これまでの議論から、 $P(A)$ と $P(B)$ の求め方はわかっています。単に、 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（つまり、起こりうるすべての結果の個数）で割ればよいのです。実は、**and**, **or**, **not** という3つの論理的な基本操作を使えば、さらに多くの事象を作ることができます。



これらは、前の講義で見た命題論理の結合子によく似ています。しかし違いもあります。ここでは、1と0の厳密な世界を扱っているのではなく、不確実性に割り当てる数について推論しようとしているのです。したがって、上の図は命題論理の論理結合子によく似ていますが、確率論ではこれらの結合子はまったく同じようには振る舞いません。実際、確率論では、これらの論理操作を集合論的操作へ翻訳するほうが有用です。

3.2 集合論的操作

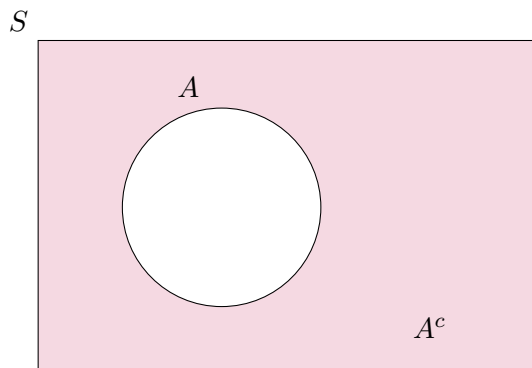
not, *and*, *or* という3つの操作は、集合論では次のように翻訳されます。

- (1) 任意の事象 A に対して、「not A 」とは、 A が起こらない事象を意味する。集合の言葉では、「not A 」は A の補集合であり、 A^c と書く。
- (2) 2つの事象 A と B に対して、「 A and B 」とは、 A と B の両方が起こる事象を意味する。集合の言葉では、「 A and B 」は A と B の共通部分であり、 $A \cap B$ と書く。
- (3) 2つの事象 A と B に対して、「 A or B 」とは、 A が起こるか、 B が起こるか、あるいはその両方が起こる事象を意味する。集合の言葉では、「 A or B 」は A と B の和集合であり、 $A \cup B$ と書く。

以下では、これらの集合論的操作をひとつずつ説明します。

3.2.1 集合の補集合

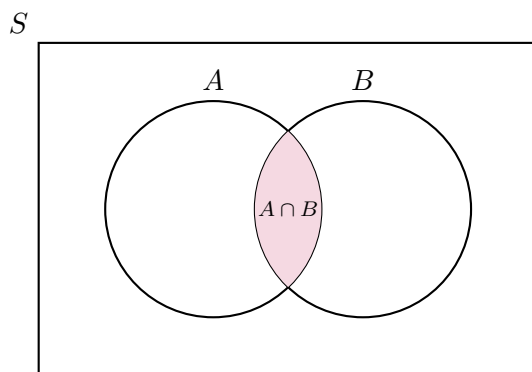
補集合 A^c とは、標本空間のうち部分集合 A に入っていないすべてのものを意味します。下の図では、補集合が塗られている部分です。



例として、 $S = \{0, 1, 2, 3, \dots\} = \mathbb{N}$ (自然数全体) とします。部分集合 $A = \{\text{偶数}\}$ をとると、 A の補集合は S の中で A に入っていない数すべてになります。言い換えれば、補集合 A^c は A から抜けている S の元全体です。 A が偶数全体なので、補集合は $A^c = \{\text{奇数}\}$ となります。

3.2.2 2つの集合の共通部分

2つの部分集合 A と B に対して、その共通部分 $A \cap B$ は、 A と B の両方に同時に属するものからなる別の部分集合です。「intersection」という言葉は、道路についてもよく使われます。交差点とは、2本の道路が交わる場所のことです。その意味で、もしちょうど交差点の上に立っていれば、同時に2本の道路の上に立っていることになります。集合論における共通部分も同じです。部分集合 $A \cap B$ は2つの集合の重なりに対応し、下の図ではその部分が塗られています。

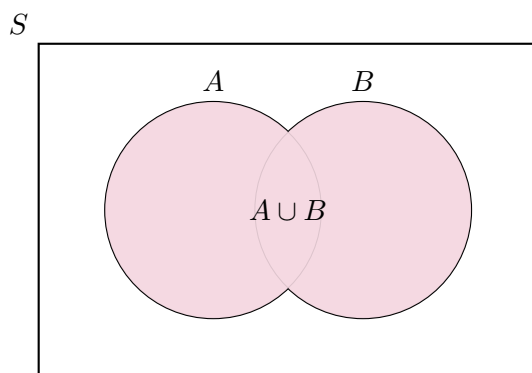


例として、再び $S = \mathbb{N}$ および $A = \{\text{偶数}\}$ を考えましょう。さらに S の別の部分集合として、 $B = \{10 \text{ 未満の数}\}$ を考えます。すると、 A と B の共通部分は、 A と B の両方に同時に入っている数の集合になります。 A は偶数全体、 B は10未満の数全体ですから、共通部分 $A \cap B$ は10未満である偶数の集合です。書き下すと、

$$A \cap B = \{0, 2, 4, 6, 8\}.$$

3.2.3 2つの集合の和集合

「union」という言葉はしばしば「いくつかのものが結び合わされたもの」を意味します。たとえばEuropean Unionは、いくつかのヨーロッパ諸国が結び合わさってできた政治的組織です。集合論でも考え方は似ています。2つの部分集合 A と B に対して、和集合 $A \cup B$ は、 A に入っているか、 B に入っているか、あるいはその両方に入っているものすべてからなる部分集合です。図で表すと、 A と B の両方を覆うものとしてイメージできます。



簡単な例として、 $A = \{1, 2, 3\}$ 、 $B = \{3, 4, 5, 6\}$ としましょう。このとき和集合 $A \cup B$ は、 A と B のメンバーをすべてまとめてできる集合なので、

$$A \cup B = \{1, 2, 3, 4, 5, 6\}.$$

となります。ここで3は2回書きません。1回だけ数えます。

3.2.4 練習：集合の計算

$S = \mathbb{N} = \{0, 1, 2, 3, 4, \dots\}$ を自然数全体とします。2つの部分集合 $A = \{\text{偶数}\}$ および $B = \{10 \text{ 未満の数}\}$ を考えます。

練習2

次を求めなさい。

- (1) B^c (B の補集合)
- (2) $A \cap B$
- (3) $A^c \cap B$

解答

(1)

$$B^c = \{10 \text{ 以上の数}\}.$$

(2)

$$A \cap B = \{0, 2, 4, 6, 8\}.$$

(3) A^c は奇数全体なので、

$$A^c \cap B = \{1, 3, 5, 7, 9\}.$$

3.3 確率 $P(\text{not } A)$

任意の事象 A について、事象 A と事象 A^c は標本空間全体を覆い、しかも重なりません。集合の言葉で言えば、 $A \cup A^c = S$ かつ $A \cap A^c = \emptyset$ です。したがって、それらの確率の和は1になります。

$$P(A) + P(A^c) = 1.$$

よって、

$$P(A^c) = 1 - P(A).$$

3.3.1 練習：FFFの補集合

先ほどのtrue/false実験では、 $P(\text{FFF}) = \frac{1}{8}$ でした。

E を「FFFが起こる」という事象とします。すると E^c は「FFFではない」という意味で、ほかの7つの結果のいずれかを表します。

練習3

$P(E^c)$ を求めなさい。

解答

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

3.4 確率 $P(A \text{ and } B)$

私たちの扱っている「古典的確率」の設定では、確率 $P(A \text{ and } B)$ は、共通部分の大きさを数えて、それを標本空間の大きさを割ることで決まるはずですが、すなわち、

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

もっとも単純な状況では、確率 $P(A \text{ and } B)$ は単に数えることで計算できます。たとえば公平なサイコロがあり、事象 A を「偶数が出る」、事象 B を「4未満の数が出る」としましょう。このとき $P(A \text{ and } B)$ は、有利な結果を具体的に数えるだけで求められます。1から6のあいだで、4未満かつ偶数である数はただ1つ、それは2です。したがって、部分集合 $A \cap B = \{2\}$ の大きさは1であり、対応する確率は $P(A \text{ and } B) = \frac{1}{6}$ です。

この例がうまくいくのは、標本空間が大きさ6の非常に小さな集合だったからです。しかし、もっと複雑な状況では、 $A \cap B$ を数えるためにすべての可能性を書き出して整理するのはたいへん難しくなります。実際には、確率 $P(A \text{ and } B)$ を $P(A)$ や $P(B)$ と結びつけるいくつかの公式を使うと便利なのがよくあります。 A と B の大きさがわかっているなら、前と同じような数え上げの公式で $P(A)$ と $P(B)$ はいつでも計算できます。しかし、部分集合 A と B の大きさだけでは、その共通部分の大きさは決まりません。したがって、確率 $P(A \text{ and } B)$ の計算は少し微妙であり、2つの事象がどのように関係しているかに大きく依存します。

3.4.1 条件付き確率

共通部分 $A \cap B$ の大きさを調べるひとつの方法は、集合 B の中を見て、その中にある A の例を数えることです。確率の言葉で言えば、これは標本空間全体 S ではなく、それが部分集合 B に制限されたと仮定して、そのもとで A が起こる確率を考えることに対応します。つまり、 B が起こるとわかったときに、 A が起こる可能性を評価しようとしているのです。実際、この考えは条件付き確率と呼ばれ、 $P(A|B)$ と書きます。これは「 B が起こると仮定したときの A の確率」と読みます。

サイコロの例に戻りましょう。事象 A は偶数が出ること、事象 B は4未満の数が出ることでした。サイコロを振る前に、神様があなたの耳元で「結果は4未満になる」とささやいたと想像してみてください。するとその瞬間、4, 5, 6は出ないとわかります。この状況で事象 A の起こる可能性を評価しようとする、可能性の空間は $\{1, 2, 3, 4, 5, 6\}$ から、それより小さい集合 $\{1, 2, 3\}$ に縮まります。この集合はちょうど B です。したがって、可能性を評価するときには

$$P(A|B) = \frac{A \cap B \text{ の大きさ}}{B \text{ の大きさ}}$$

という数え方をすることになります。 A と B に対応する部分集合はそれぞれ $\{2, 4, 6\}$ と $\{1, 2, 3\}$ です。したがって、共通部分 $A \cap B = \{2\}$ の大きさは1です。 B の大きさは3なので、条件付き確率は $P(A|B) = \frac{1}{3}$ になります。

ここで有用な観察をします。この例では、 $P(B) = \frac{3}{6} = \frac{1}{2}$ 、 $P(A \text{ and } B) = \frac{1}{6}$ 、そして $P(A|B) = \frac{1}{3}$ でした。するとたしかに $\frac{1}{6} = \frac{1}{2} \cdot \frac{1}{3}$ です。記号で書けば、これは一般公式のひとつの例になっています。

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

したがって、複合事象「 A and B 」の確率を計算するには、条件付き確率を含む積をとればよいのです。（同値な形として、 $P(A \text{ and } B) = P(A) \cdot P(B|A)$ でもあります。）

3.4.2 確率的独立

場合によっては、 $P(A \text{ and } B)$ の公式を大きく簡単にすることができます。そのためには、独立な事象と従属な事象を区別する必要があります。

独立事象と従属事象

2つの事象 A と B に対して、次のように言います。

- A と B が**独立**であるとは、 A について知っても、 B が起こるかどうかにして何の情報も得られない場合である。
- A と B が**従属**であるとは、 A が起こったと知ることによって、 B が起こるかどうかにして何らかの情報が得られる場合である。

思い出してください。確率とは不確実性の科学です。したがって、従属事象という考え方の背景には、「事象 A が起こったと知ることによって、 B の確率が変わるかもしれない」という考えがあります。これは、 A に関する情報が何らかの形で B に関係しているときに起こりえます。

前のサイコロの例では、 A が「偶数が出る」、 B が「4未満の数が出る」でした。この2つは従属事象です。なぜなら、一方についての情報が、他方の可能性についての知識に影響を与えるからです。しかし、一方の情報を知っても、確率の見積もりを変える理由が何もないような状況も考えられます。たとえば、事象 C を「3の倍数が出る」としましょう。このとき集合としては $C = \{3, 6\}$ です。もし今度は神様が耳元で「結果は3の倍数になる」とささやいたとすると、その瞬間、標本空間は $\{1, 2, 3, 4, 5, 6\}$ から $\{3, 6\}$ に制限されます。この新しい制限された集合には奇数が1つ、偶数が1つ含まれているので、偶数が出る確率は依然として $\frac{1}{2}$ です。これは $P(A)$ と同じです。したがって、 C が起こるとわかって、系についての確かさの度合いは変わりません。この意味で、事象 A と C は独立です。

この考え方によれば、追加情報を知っても不確実性が実際には変わらないとき、その2つの事象は独立だということになります。言い換えれば、条件付き確率が無条件の確率と一致するとき、2つの事象は独立です。すなわち、 $P(A|B) = P(A)$ かつ $P(B|A) = P(B)$ のとき、事象 A と B は独立です。すると、 $P(A \text{ and } B)$ の公式は次のように簡単になります。

独立事象の乗法則

A と B が独立事象ならば、

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

例：公平なコインを2回投げることを考えましょう。事象 A を「1回目か裏」、事象 B を「2回目か裏」とします。これらは独立事象です。1回目のコインを投げて、次のコイン投げの結果については何の情報も得られません。したがって、独立事象の公式を使って

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

と結論できます。同じことを標本空間 $\{HH, HT, TH, TT\}$ で見れば、 TT だけが有利な結果なので、確率は $\frac{1}{4}$ であり、上の計算と一致します。

3.4.3 例：独立事象 (2組のカード)

2組の別々のトランプの山を取り、それぞれをよくシャッフルするとします。 $A =$ 「デッキ1の一番上のカードがエースである」 $B =$ 「デッキ2の一番上のカードがエースである」とします。

これらの事象は独立であり、それぞれの確率は $\frac{4}{52} = \frac{1}{13}$ です。したがって、

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

3.4.4 例：従属事象 (1組のカードから2枚のエース)

52枚のトランプ1組をシャッフルするとします。 $A =$ 「1枚目がエースである」 $B =$ 「2枚目がエースである」とします。

これらは従属事象です。なぜなら、1枚目がエースだと、山札に残っているエースの枚数が減るからです。

計算すると、

$$P(A) = \frac{4}{52} = \frac{1}{13}, \quad P(B|A) = \frac{3}{51}.$$

したがって、

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

3.4.5 乗法法則のまとめ

Step 1: 事象が独立かどうかを判断する。

(B が不可能事象でないとは仮定して。) もし $P(A) = P(A | B)$ なら独立である。

Step 2: 正しい乗法法則を使う。

独立なら $P(A \text{ and } B) = P(A) \cdot P(B)$ 。

独立でないなら $P(A \text{ and } B) = P(A) \cdot P(B | A)$ 。

3.5 確率 $P(A \text{ or } B)$

Section 3.2 によれば、複合事象「 $A \text{ or } B$ 」は集合論的には和集合 $A \cup B$ に対応します。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て、それを標本空間の大きさを割ればよいことになります。

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

先ほど議論した共通部分と同様に、一般には A と B の大きさだけから、和集合 $A \cup B$ の大きさを決めることはできません。2つの集合がどのように重なっているかについての情報も必要です。幸い、私たちはすでにその重なりを数える方法を導いているので、 $A \cup B$ の大きさを A , B , そして共通部分 $A \cap B$ の大きさと関係づけることができます。その公式は

$$A \cup B \text{ の大きさ} = (A \text{ の大きさ}) + (B \text{ の大きさ}) - (A \cap B \text{ の大きさ}).$$

です。この公式が成り立つのは、 A と B の両方に入っているものをすべて数えたいからです。すると、 A の大きさと B の大きさを単に足して、それが和集合 $A \cup B$ の大きさだと言いたくなります。しかしそうすると、共通部分 $A \cap B$ を二重に数えてしまう危険があります。自分で確かめるために、2つの集合 $A = \{1, 2, 3\}$ と $B = \{2, 3, 4\}$ を考えてください。どちらも大きさ3で、その和は $3 + 3 = 6$ です。しかし、両方のメンバーをすべて集めた和集合 $A \cup B$ は $\{1, 2, 3, 4\}$ であり、大きさは4です。6ではありません。 $3 + 3 = 6$ と計算するとき、実際には元2と3を誤って2回数えてしまっています。ところが、余分な $\{2, 3\}$ を取り除けば、 $6 - 2 = 4$ となり、これが正しい和集合の大きさです。この観察を使うと、 $P(A \text{ or } B)$ を $P(A)$, $P(B)$, そして $P(A \text{ and } B)$ で表す一般公式が導けます。

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

場合によっては、事象 A と B に対応する部分集合がまったく重ならないこともあります。このとき、共通部分 $A \cap B$ は空集合 (つまり大きさ0) なので、二重に数える危険はありません。

このようなとき、 A と B は互いに排反であると言います。 A と B の部分集合は、それぞれ事象 A と B を実現する結果の集合に対応しているので、互いに排反な事象は同時には起こりえません。すなわち、 $P(A \text{ and } B) = 0$ です。

A と B が互いに排反であれば、 $P(A \text{ or } B)$ の公式は

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

に簡約されます。

3.5.1 例：エース、スペード、キング（一番上のカード）

トランプを1組シャッフルするとします。 $A =$ 「一番上のカードがエースである」 $B =$ 「一番上のカードがスペードである」 $C =$ 「一番上のカードがキングである」とします。

A と B は互いに排反ではありません。スペードのエースが A と B の両方に含まれているからです。 A と C は互いに排反です。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す。

練習4

次を求めなさい。

- (1) $P(A \text{ or } B)$
- (2) $P(A \text{ or } C)$

解答

(1)

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

したがって

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

(2) A と C は互いに排反なので、

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

4 練習問題

4.1 サイコロを振る

公平な6面体のサイコロを振るとします。 $A =$ 「偶数が出る」 $B =$ 「 ≥ 4 の数が出る」とします。

練習5

次を求めなさい。

- (1) $P(A)$
- (2) $P(B)$
- (3) $P(A \text{ and } B)$
- (4) $P(A \text{ or } B)$

解答

(1) $A = \{2, 4, 6\}$ なので,

$$P(A) = \frac{3}{6} = \frac{1}{2}.$$

(2) $B = \{4, 5, 6\}$ なので,

$$P(B) = \frac{3}{6} = \frac{1}{2}.$$

(3) $A \text{ and } B$ は「偶数かつ ≥ 4 」を意味するので $\{4, 6\}$, したがって

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

(4) $A \text{ or } B$ は $\{2, 4, 5, 6\}$ を意味するので,

$$P(A \text{ or } B) = \frac{4}{6} = \frac{2}{3}.$$