

「確率とその応用」ノート、その式

逆瀬川浩孝

2 確率の定義、確率変数

サイコロ振りのように、ランダムな結果の起こりうる試行を考えたとき、試行の可能な結果（1の目が出る、2の目が出る、...）を**標本点**といい、それらをすべて集めたものを**標本空間**といいます（サイコロ振りの場合 $S = \{1, 2, 3, 4, 5, 6\}$ と書くことにしよう）。コイン投げ試行ならば、標本点は「おもて」「うら」の2通り、標本空間は $S = \{\text{おもて}, \text{うら}\}$ です。サイコロを二つ振る場合は、標本点は $(i, j) (i, j = 1, 2, \dots, 6)$ の36通りです。

多くの場合、ランダム試行で関心があるのは、たとえば、サイコロを振って偶数の目が出る（2か4か6の目が出る）かどうかという場合のように、標本空間の中のある部分集合に含まれる標本点が観測されるか否か（2か4か6の目が出るかどうか）、観測されるとすればどれくらいの可能性があるかということです。標本空間 S の部分集合を**事象**といい、事象 A の起きやすさを**確率**と定義し、 $P(A)$ と書くことにします（たとえば、偶数かどうかということに関心があるのであれば、 $A = \{2, 4, 6\}$ です。そして、その確率は $P(A) = P(\{2, 4, 6\})$ と書きます）。

2.1 確率の決め方

すべての事象 A の確率 $P(A)$ を決めることを**確率規則を決める**と言うことにします。標本空間が有限集合（ $S = \{s_1, s_2, \dots, s_N\}$ ）の場合、事象の確率は、その事象を構成する標本点一つだけからなる事象（**根元事象**という）の確率（標本点の確率ともいう）を決めてから、それらの確率を合計したものとして事象の確率を求める、というやり方が分かりやすいでしょう。根元事象の確率を決める方法としては次のようなものがあります。

根元事象（標本点）の確率

1. 等可能性の原理を適用する（テキスト 37 ページ）
2. 実験に基づく相対頻度を用いる
3. 過去の経験と理屈を使って決める
4. 主観的な判断で決める

見たところ正常なサイコロを振って出る目に関心がある場合、6つの目のどれもが他のどれよりも出やすいということは考えにくいので、どの目が出る確率も6分の1とするのが正当でしょう。このように、どの標本点の出方も等しいと考える考え方を**等可能性の原理**といいます。何も予備知識が無い場合には、皆がこの原理に従わざるを得ない、と納得するに違いありません。標本点の個数を N としたとき、どの標本点の確率も等しく N 分の1とします。

しかし、その決め方が絶対に正しいかというと、そういうわけでもありません。実際、意図的に1が出やすいサイコロを作ることができますから、実際に振ってみたらやけに1が出やすいサイコロだ、というように目の出方が均等でないということが分かったら、等可能性の原理にこだわるのは適切ではありません。その場合、個々の標本点の出方の違いを調べるには、もし可能ならば実際に実験をしてみて、実験結果に合うように決めればよいでしょう。たとえば、サイコロの場合、100回くらい振ってみて、その頻度を求め、それらを相対度数に直したものを各目の出

る確率とするのが良いかもしれません。実験を何回くらい繰り返せばよいのか決まった規則はありませんし、相対度数はいつも同じとは限りませんから（100 回振ったときの結果と 101 回振ったときの結果は確実に違います）、相対度数をそのまま確率とするのではなく、それを参考にしきりのよい数に決めるというのが普通です。

実験することが不可能、あるいは信頼できる相対度数を計算できるほど多数回の実験を繰り返すことはできない、というような場合は、いろいろな情報を集め検討し、最終的に自分はどう思うという数値を採択します。そのようにして決めた確率は主観確率といいます。

例 2.1 赤と白の二つのサイコロを振って目を調べるというランダム試行では、 $(1,1), (1,2), \dots, (6,6)$ という 36 通りの目の組み合わせが結果のすべて。 $(2,3)$ と $(3,2)$ は異なる標本点とみなします。何も事前の情報が無い場合、一つ一つの組み合わせに対して $\frac{1}{36}$ の確率を割り当てるのが普通です。これが等可能性の原理。二つの区別できないサイコロを振ると、目の組み合わせは $(i,j) (1 \leq i \leq j \leq 6)$ という 21 通りになりますが、標本点は 21 通り、とするのが間違いというわけではありません。しかし、等可能性の原理を当てはめることが出来なくなり、扱いがやっかいです。なんとか区別を付けて 36 通りとするほうが、後の議論が簡単になります。

例題 2.1 二つの正しい n 面体さいころを振って、大きい方の目が k である確率を求めなさい。

二つのサイコロを振ったときに出る目の組み合わせというと紛らわしいので、2つのサイコロは色が違うと考えると、目 (i,j) の組み合わせは n^2 通りある。正しいさいころなので、そのどれもが同じ確率で実現すると考えてよい。そうすると、該当する場合の数を数えて n^2 で割ったものがもとめたい確率になる。そろ目の場合を除くと、 $i < k$ となる (i,k) の組み合わせは $k-1$ 通り、同じく (k,i) の組み合わせも $k-1$ 通り、 (k,k) の一通りとを合わせて全部で $2k-1$ 通り、従って、もとめたい確率は $(2k-1)/n^2$

これでもよいのだが、これからも頻繁に利用することになるであろうスマートな考え方があるので、慣れておくとよい。大きい方の目が k ということは、両方の目が k 以下、しかし、両方とも $k-1$ 以下ではない、と考える。サイコロの目が k 以下となるのは k 通り、両方ともそうなるのは k^2 通り、したがって、ちょうど k になる確率は以下のように与えられる。

$$\frac{k^2 - (k-1)^2}{n^2} = \frac{2k-1}{n^2}, \quad k = 1, 2, \dots, n$$

例 2.2 工場で生産される製品の検査工程を考えたとき、得られる結果は不良か不良でないかのいずれかです。結果は二通りですから、等可能性の原理を適用して不良品を生産する確率は 0.5、というのは乱暴すぎます。この場合は、過去の実績データから不良率を推定するのが受け入れられやすいでしょう。

例 2.3 さいころを 100 回振って、各目の出た度数がそれぞれ 25, 14, 16, 12, 13, 20 回だったとき、相対頻度の立場に立てば、各目の出る確率は 0.25, 0.14, 0.16, 0.12, 0.13, 0.2 とするのが良さそうだが、あと 10 回実験を追加して相対度数を計算し直すと、この「確率」とは違う数値が得られるので、この決め方はちょっと安直。正しいサイコロでもこれくらいの度数の違いは起こりうるはずとして、等可能性の原理を適用し、 $\{\frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}\}$ とするのも一つの考え方。あるいは 1, 6 の目が出やすい形になっているのかもしれないと考えて $\{0.2, 0.15, 0.15, 0.15, 0.15, 0.2\}$ とするのも一つの考え方。どちらが正しいかは、今は判断できない。このさいころを使って賭けをするのであれば、賭けの参加者が集まって決めればよい。

標本点が無限個ある場合でも、 $p_1 + p_2 + \dots = 1$ となるような正数の組が見つければ、（無限

個の) 根元事象の確率を使って事象の確率を決めることができます。例えば、コインをおもてが出るまで投げ続けた場合の投げる回数は $1, 2, 3, \dots$ という可能性があり、限りがありません。 i 回目に初めておもてが出る確率 p_i は $p(1-p)^{i-1}$ で与えられ、それらを $i = 1, 2, 3, \dots$ について合計すると 1 になります。従って、根元事象の個数は無限ですが、 $\{p_i = p(1-p)^i, i = 1, 2, \dots\}$ を使った上の方法で確率モデルを構成することができます。これは**幾何分布**と呼ばれます。

あるいは、

$$p_i = \frac{a^i}{i!} e^{-a}, \quad i = 0, 1, 2, \dots \quad (2.1)$$

もその例です。

$$e^a = 1 + a + \frac{a^2}{2!} + \frac{a^3}{3!} + \frac{a^4}{4!} + \dots \quad (2.2)$$

という公式 (指数関数のマクローリン展開) を使うと、 $p_0 + p_1 + p_2 + \dots = 1$ が導かれるからです。これは**ポワソン分布** (あるいは、ポアソン分布) と呼ばれます。

事象の確率

標本点の確率が決まると、あとは、任意の事象 $A \subset S$ の確率は、 A に含まれるすべての標本点の確率を足したものとする、と約束すれば、すべての事象の確率が決まります。

例題 2.2 起こりうる結果が n 通りあるようなランダム試行を説明する確率モデルとして、標本空間を $S = \{s_1, s_2, \dots, s_n\}$ とし、 $p_1 + p_2 + \dots + p_n = 1$ となるような n 個の正数を決めて、 S の任意の部分集合 A に対して

$$P(A) = \sum_{s_i \in A} p_i$$

と定義すると、 $P(\cdot)$ は確率の公理を満たすことを説明しなさい。

こたえ $P(A) \geq 0, P(S) = 1$ は定義から明らか。3 番目の公理は、 A と B が互いに排反ならば

$$\sum_{s_i \in A} p_i + \sum_{s_i \in B} p_i = \sum_{s_i \in A \cup B} p_i$$

となることから分かります。

$A = \{s_i\}$ とすると、 $P(s_i) = p_i$ ですから、 p_i は根元事象 $\{s_i\}$ の確率になっていることに注意して下さい。標本空間が離散 (で標本点の個数が n 個) の場合は、このように、 $p_1 + p_2 + \dots + p_n = 1$ となるような n 個の正数を決めて、各根元事象の確率とし、事象の確率は、その事象に含まれる根元事象の確率を合計したもの、と設定すれば、矛盾なく、確率を定義することができます。 $n = 6, p_i = \frac{1}{6}$ とした場合が正しいサイコロの確率モデルです。一般の場合の確率モデルは、サイコロをイメージして、一般に n **面体の歪んだサイコロ** を振った時に出る目の確率規則をモデル化したもの、と考えると良いでしょう。

もし、標本点の確率が等可能性の原理に基づくものだとすると、 $p_i = \frac{1}{n}$ ですから、任意の事象 $A \subset S$ の確率は、 A の標本点の数を $|A|$ として、 $P(A) = \frac{|A|}{N}$ とすることができるので、あとは A に含まれる標本点の個数を計算する問題になります。高校の確率で主に学んだのはこの標本点の個数を計算するための 2 項係数を駆使した計算でした (「袋の中に赤い玉と黒い玉が 4 個ずつ入っています。4 個取り出したときに赤い玉と黒い玉が半々となる確率はいくつですか?」)。

結局、どの決め方が良いのか？

さあ、どれがよいでしょうね。正解はありません。可能ならば試行を繰り返して頻度確率を求め、さもなければ、過去の経験、類似の例などを参考に「適当に」定め、関係者の間で了解を取り付ける、という方法しかないのですね。世の中には真実の一つ、と思い込んでいる人には「目から鱗」(って知っていますか？ ⇒「言葉のページ」参照) かもしれません。

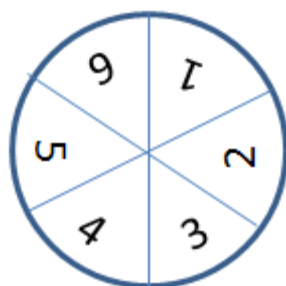
結局、「確率」の絶対的(真実の)定義は存在するわけではなく

1. ある事象の確率は「**矛盾が起きない限りは**」つまりは「**確率の第3公理をみたす限りは**」皆が納得するように、どう決めても構わない。
2. その結果として不都合が起きたとしても、皆で納得して決めたことなのだから誰も文句を言うことはできない。
3. 皆が不都合と思うのだったら、新たに決め直せばよい。

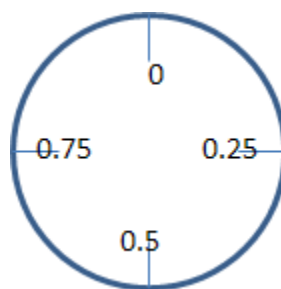
というのが現実的な対処法です。

2.1.1 標本点が連続無限個ある場合の確率の決め方

標本点が無数にある(連続無限個ある)場合は、標本点の確率を決めてから事象の確率を決めるというやり方がうまくいきません。人生ゲームについているコマの形をしたサイコロを思い浮かべてください。人生ゲームのマスを進めるために、コマはサイコロのように6等分されていますが、その区切りを取り払い、円周上に0から1までの目盛りを書いたコマを使うとしたらどうでしょう。ゲーム盤のマスも取っ払って一本の道にして、コマの目盛り分だけ進むというようにするのです。



コマ型のサイコロ



連続目のサイコロ

普通のコマの場合、3マス進める確率は6分の1ですが、連続目のコマの場合、ちょうど、たとえば0.426進む確率はゼロです。なぜならば、「目」は無数にあり、どの目の出る可能性も等しいので、その一つ一つに正の確率を与えていたら、確率の合計が1を超えてしまうからです。しかし、コマを回せばどこかで停止するのですから、そのときの「目」の確率について何かを決めておかなければ議論ができません。そこで、考え出されたのが、標本点ではなく、それらを集めた集合としての事象に確率を与える方法です。たとえば、0.15と0.34の間の「目」が出る確率は、区間の長さに比例して0.19であると決める方法です。標本点は0以上1未満の実数全体、従って標本空間は $S = [0, 1)$ という半开区間とします。事象は S の任意の部分集合ですが、たとえば $[a, b) \subset S$ に対しては $P([a, b)) = b - a$ と約束するのです。確率とは面積のようなものと言いましたが、この場合は確率とは長さのようなもの、といってもよいでしょう。この決め方は、

確率の公理をすべて満たすことを確認してください（重ならない区間の合計の長さは、個々の区間の長さの合計に等しい、など）。

確率の第3公理の意味（テキスト 35 ページ）

標本空間が有限（可算無限）の場合は自然にそうなるように決めるのが普通で、意識することはないでしょう。本当に必要なのは、標本空間が連続無限の場合です。このとき、離散の場合のように、事象に含まれるすべての標本点を数え上げることが出来ませんから、標本点一つ一つに確率を与える方法では事象の確率は計算できません。そこで、一つ一つの事象に確率を「矛盾なく」定義する必要があるのです。まあしかし、あまり奇抜なことを考えなければ、**面積のように**、「部分集合の大きさ」を確率としておけば、（公理を）意識する必要はありません。

例 2.4 ダーツボードにダーツを投げるという試行では、標本点は「ある点に刺さること」で、その確率はゼロ。「現に刺さったのだから、確率がゼロのわけがない」と考える人もいるようですが、ここは数学の話。「ある点」は二つの実数値の組 (x, y) で確定します。実数値は無限桁を持ち、ランダムに選んだ一組の実数値が完全に一致することはありません。だから、確率ゼロ。ダーツボードに描かれた円盤に刺さるという事象の確率を決めるために、特定の点に刺さるという標本点の確率から求めることはできないので、その場合は、例えば円盤のある微少な部分 Δ に刺さるという事象の確率として、円盤の面積を 1 としたときの Δ の相対面積としよう、というように、標本点を介さずに、事象の確率を与える必要があります。

ランダムに投げるとすれば、それでも良いのですが、高い点数に当てようという目的があって投げているのであれば、点数の高い場所ほど刺さる可能性は大きくなりますから、単純な面積を確率にしたのでは不都合でしょう。その場合は、中心部のある範囲に刺さる確率と周辺部の同じ大きさの範囲に刺さる確率に差を付けるようにします。この決め方は、歪んだサイコロの目の確率の決め方と同じことです。いずれにしても「矛盾がない」ように、公理 3 を満たすように決める必要があります。

2.2 いくつかの例題

例題 2.3 壺に赤、白、黄色の玉が、それぞれ 5 個、3 個、2 個はいっている。そこから一度に二つをランダムに選び出すという実験をするとき、標本空間としてどのようなものを考えればよいですか、確率規則を決めなさい。

こたえ 答えは一通りではありません。

決め方その 1： 三つの色の組み合わせ「赤赤」「赤白」「赤黄」「白白」「白黄」「黄黄」の 6 通りを標本点とする。

「赤赤」の確率は 10 個の中から 2 個を取る取り出し方の数 $\binom{10}{2} = 45$ を分母、5 個の中から 2 個を取る取り出し方の数 $\binom{5}{2} = 10$ を分子にしたものです。以下同様に

$$P(\text{赤白}) = \frac{\binom{5}{1}\binom{3}{1}}{45} = \frac{15}{45}, \quad P(\text{赤黄}) = \frac{\binom{5}{1}\binom{2}{1}}{45} = \frac{10}{45},$$

$$P(\text{白白}) = \frac{\binom{3}{2}}{45} = \frac{3}{45}, \quad P(\text{白黄}) = \frac{\binom{3}{1}\binom{2}{1}}{45} = \frac{6}{45}, \quad P(\text{黄黄}) = \frac{\binom{2}{2}}{45} = \frac{1}{45}$$

となります。最初は誰でも考える標本点の選び方で、間違いではありませんが、確率計算をすることを考えると、標本点はなるべく個別の結果（根元事象）にしておいた方が便利です。「二つ

のさいころを振ったときの結果をリストアップしてください」と言われたら、(1,1) から (6,6) までの 36 通りと答えるでしょう、(1,6) と (6,1) を区別せず 21 通りと考えるよりは、表現がシンプルになるからです。その説明に納得できるなら、次の決め方その 2 もすんなり、納得できるのではありませんか。

決め方その 2: 10 個の玉に通し番号を付けて、10 個の玉の中からランダムに選ぶ選び方 $\binom{10}{2} = 45$ 通りの結果を標本点とする。

こうすると、どの標本点も等しく選ばれることになりますので、等可能性の原理が適用できて、

$$P(\{i, j\}) = \frac{1}{45}, \quad 1 \leq i < j \leq 10$$

とすることができます。事象と対応させるには、 $i = 1, 2, 3, 4, 5$ ならば「赤」、 $i = 6, 7, 8$ ならば「白」、 $i = 9, 10$ ならば「黄」とすれば、(1, 4) が選ばれたら「赤赤」という事象が起きた、(3, 9) が選ばれたら「赤黄」という事象が起きた、というように見なすことができます。

例題 2.4 トランプカード 52 枚の中からランダムに 1 枚取り出して元に戻すということを 5 回繰り返すという試行において、取り出された 5 枚の中に含まれるハートの枚数に興味があるという場合、標本空間をどのようにすればよいですか。また、標本点の確率はどのように定義しますか。

こたえ これも答えは一通りではありません。

決め方その 1: 標本点は「 $\spadesuit A \heartsuit 4 \diamondsuit 3 \clubsuit J \clubsuit J$ 」などのように、取り出された 5 枚のカードを並べたもので、全部で $52^5 \approx 3.8 \times 10^8$ 通りある。標本点の確率は等可能性の原理で決めて良い。

決め方その 2: しかし、ハートかそれ以外かということだけに興味があるというのであれば、 $A, 2, 3, \dots, Q, K$ の区別は必要ないので、エース以外のカードを破棄して、4 枚のエースから選んでも良い。というわけで、「 $\spadesuit \heartsuit \diamondsuit \clubsuit$ 」を使って出来る長さ 5 の記号列 (単語という) $4^5 = 1024$ 個を標本点とする (だいぶ減りました)。この場合も、標本点の確率は等可能性の原理で決めて良い。

決め方その 3: もっと簡略化する場合は表に $\heartsuit$ を刻印した歪んだコインで、 $\heartsuit$ の出現する確率が $\frac{1}{4}$ となるようなものを用意し、それを 5 回振った結果を標本点とする、という考え方がある。そうすると、 $\heartsuit$ を H (Head、おもて)、そうでない場合を T (Tail、うら) として、標本点は H, T の 2 文字を使って出来る長さ 5 の「単語」 $2^5 = 32$ 個になる。この場合は「等可能性の原理」が適用できず、「数学的確率」すなわち、 H を $p = 0.25$ 、 T を $1 - p = 0.75$ で置き換えた 5 個の確率の積を一つの標本点の確率とする。たとえば $P(HHTHT) = 0.25^3 \times 0.75^2 = 0.0088$ 、など。

決め方その 4: さらに簡略化して、 $\heartsuit$ の枚数、0, 1, 2, 3, 4, 5 の六つだけを標本点にするという考え方も、当然ありうる。その場合、標本点の確率は 2 項分布を使って計算しなければいけないのでややこしくなる。

このような組合せの問題では等可能性の原理を利用したほうがいろいろな問題を扱うときに便利で、過度の簡略化はあまり得にならない。

例題 2.5 二つの事象 A, B に対して、次の「確率の和の公式」が成り立つことを証明しなさい。

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \quad (2.3)$$

こたえ A と A^c は排反なので A と $A^c \cap B$ ($\subset A^c$) は排反。 A と $A^c \cap B$ の和集合は $A \cup B$ なので、

$$P(A \cup B) = P(A) + P(A^c \cap B) \quad (2.4)$$

同じ理由で $A \cap B$ と $A^c \cap B$ は排反。 $A \cap B$ と $A^c \cap B$ の和事象は B なので、

$$P(B) = P(A \cap B) + P(A^c \cap B) \quad (2.5)$$

$P(A^c \cap B)$ を二つの式から削除すると与式が得られる。

良くある間違い、その1：「(重なり合う○を二つ描いて) ベン図式から明らか」という「証明」。ベン図は思考の整理には役立ちますが、証明にはなりません。たとえば、さいころを二つ振ったときの標本空間の中で、そろ目が出るという事象はどういう○になりますか、そんなもの決まった書き方はありません。あなたがそう描けば○になりますが、□で描いても間違いではありません。飛び飛びに描いてもかまいません。そうなるような図を自分で描いて「明らか」というのは、「明らかになるように描いただけ」ということなのです。

良くある間違い、その2：「 $P(A) = P(A \cap B) \cup P(A \cap B^c)$ と分解出来る」という表現。「 $\cup$ 」は集合演算の記号、 $P(\cdot)$ は数ですから、 $P(\cdot) \cup P(\cdot)$ という表現はありません。

例題 2.6 (確率問題) 久しぶりのクラス会に 50 人が集まった。この中に同じ誕生日を持つ人がいる可能性はどれくらいあると思いますか。半分か？ 30 人だったらどうか。5 分の 1 くらい？ 確率が半々になる人数は何人必要でしょうか。

こたえ 同じ誕生日を持つ人がいない、という確率を計算してそれを 1 から引く方が考えやすい。1 年 365 日で、どの日に生まれる確率も等しいと仮定すると、組み合わせを数える問題になる。最初の人の誕生日は 365 通り、2 番目の人は 364 通りの中から選び、3 番目の人は 363 通りの中から選び、というように、前の人の誕生日を避けるように選択する。一方、50 人の誕生日の組み合わせは全部で 365^{50} 通り、したがって同じ誕生日を持つペアが存在しない確率は

$$\frac{365 \times 364 \times 363 \times \cdots \times 316}{365^{50}} = \frac{364}{365} \frac{363}{365} \cdots \frac{316}{365} \approx 0.0296$$

となり、目的の確率は 0.97 となる。ほぼ確実に誰かと誰かは同じ誕生日を持つ。同じように 30 人の場合を計算してみると

$$\frac{365!}{335!} 365^{-30} \approx 0.2937$$

したがって、約 71% の確率で同じ誕生日を持つペアがいる。出席者をどんどん減らしていくと、

$$\begin{aligned} \frac{365!}{342!} 365^{-23} &\approx 0.4927 \\ \frac{365!}{343!} 365^{-22} &\approx 0.5243 \end{aligned}$$

となり、結局、23 人いれば、同じ誕生日を持つペアが見つかる確率が 0.5 を超えることが分かる。これは想像したより少ない。

例題 2.7 部屋を借りるために、不動産を見て歩く。見るのは一つの物件に付き 1 回だけ。見たときに決めないと、チャンスは消える。全部で n の物件を見ることが出来るとしたとき、なるべく良い物件と契約できるようにするためにどのような方針で臨めば良いか、という問題について考えてみましょう。最初の m 個については価格水準を見るために、契約しない。その m 件の中の最良の物件を覚えておいて、次にその物件より条件の良いものに会ったら、それを借りることにする、というのが現実的な戦略でしょう。 m が多すぎると、良い物件を見逃してしまう可能性が大きく、 m が小さすぎると良い物件を見ないまま契約をしてしまう危険性が高い。 n に

対して m をどれくらいにしたら、最良の物件を借りる可能性を最大にすることが出来るでしょうか。

こたえ 契約をするまでに見て回る物件の数を k とすると、 $k = m + 1, m + 2, \dots, n$ のいずれかであるから、それについて条件を付けて全確率の公式を適用する。 A を最良の物件と契約するという事象、 B_k を k 軒目の物件と契約するという事象とすると、

$$P(A) = \sum_{k=m+1}^n P(A \cap B_k)$$

と表される。

カードの問題に置き換えて考える。1 から n までの数を記した n 枚のカードを用意してランダムに並べる。数は、物件の相対順位を表すと考え、最初から k 枚目のカードの数は、 k 番目に見る物件の相対順位を示している。最初の m 枚の中の最小値が、契約基準を表す。 $m + 1$ 枚目以降で、その最小値よりも小さい数が出てきたら、そこで物件巡りを終えて、契約する。そのカードに書かれている数が 1 である確率はいくつか、というのが問題。

事象 $A \cap B_k$ は、 k 枚目の数が 1 で、 $k - 1$ 枚目までのカードの最小値が最小 m 枚目までにある、という事象を表す。そのようなカードの並び方が何通りあるか分かれば、それを n 枚から k 枚取り出して一列に並べる並べ方の数で割れば、求めたい確率が計算できる。

n 枚から k 枚を取り出して一列に並べる並べ方は $(n)_k = n(n-1) \cdots (n-k+1)$ 通り。 n 枚から 1 を取り出す取り出し方は、もちろん一通り。 $k - 1$ 枚を取り出す取り出し方は $\binom{n-1}{k-1}$ 通り、その中の最小数を除いた $k - 2$ 枚を一列に並べる並べ方は $(k-2)!$ 通り、最小数を最初の m 番目までに並べる並べ方は m 通り。したがって、目的の確率は

$$\frac{\binom{n-1}{k-1} (k-2)! \times m}{n(n-1) \cdots (n-k+1)} = \frac{1}{n} \frac{m}{k-1}$$

となる。

したがって、目的の確率は

$$P(A) = \frac{m}{n} \sum_{k=m+1}^n \frac{1}{k-1}$$

これだけでは見通しが悪いので、その特徴を調べるために n を大きくしてみる。 m は当然 n とともに増えていくと思われるので、 $m/n (= a)$ を一定として、 n を大きくする。 m も n も大きければ $\sum_{k=m+1}^n \frac{1}{k-1}$ は積分で近似的に計算できる。

$$P(A) = \frac{m}{n} \sum_{k=m+1}^n \frac{1}{k-1} \approx a \int_m^n \frac{dx}{x} = a \log \frac{n}{m} = -a \log a$$

これを最大にする a はいくつかという問題を解く。

通常の微分で、 $0 < a < 1$ の範囲の最大値を求めることが出来る

$$-\log a - 1 = 0 \Leftrightarrow a = e^{-1} \approx 0.368$$

したがって、見て回ることの出来る物件の数が決まったら、そのうちの 37% くらいをチェックして、その中の最高水準を記憶し、その後、その最高水準を上回る良い物件に出会ったら契約をする、というのが賢い方法である。

その値を $P(A)$ の式に代入すると、 a に等しくなり、最良の物件が得られない可能性が 3 分の 2 くらいあるというのが、結論である。

練習問題 2.1 事象 A が起きないという事象 A^c に対して、 $P(A^c) = 1 - P(A)$ が成り立つことを「証明」しなさい。

注意：「 $P(A)$ と $P(A^c)$ は互いに排反」とはいわない。

練習問題 2.2 確率の公理（教科書 36 ページ）で、「事象」を「土地の区画」、「確率」を「面積」、「全事象」を「日本全土」、「和事象」を「合わせた区画」というように、読み替えたとき、三つの公理が成り立つことを説明しなさい。ただし、「面積」は日本全土の面積を 1 としたときの相対面積と考えます。

練習問題 2.3 「(たとえば、「このさいころを振ったとき 1 の目が出る確率は 6 分の 1」というように) ある事象が起きる確率はいくつであるというのは、それが正しいからそう決めたわけではなく、単なる約束事である。納得が行かなければ皆で相談して決め直せばよい」という命題について、もし、それを確率を知らない人に話をしたときに予想される反論・疑問を書き、それらに対するわかりやすい説明を考えなさい。

練習問題 2.4 円の弦をでたらめに引いたとき、その弦の長さが内接正四辺形の一辺の長さより長い可能性はどれくらいありますか。

2.3 確率変数と確率分布

例題 2.8 表の出る確率が p であるようなコインを繰り返し投げて、初めて表が出たらおしまいにします。このとき、コインを投げた回数を X とします。(1) X の取りうる値を列挙しなさい。(2) $X = 3$ の確率はいくつですか。(3) $X > 5$ の確率はいくつですか。(4) $X \leq 100$ の確率は「おおよそ」いくつですか。

こたえ (1) 正の整数「 $X = 1, 2, 3, \dots$ 」が正解。「無限」、あるいは「 $X \geq 1$ 」としてはいけません。 X は整数値だけです。

(2) 1 回目、2 回目は裏、3 回目が表、なので $p(1-p)^2$

(3) 少なくとも最初の 5 回は全部裏、なので、 $(1-p)^5$

(4) 相当に歪んだコインでも、100 回連続で裏ということは考えにくいので、「おおよそ 1」。「おおよそ」というのは「だいたい」とか、「おおざっぱに」とか、「アバウトに言う」とか、「いい加減に見積もると」というような意味です。

$$P(X \leq 100) = 1 - (1-p)^{100}$$

は「厳密解」で、「おおよそ」の答えとは言えません。

どれくらい歪んでいたら「おおよそ 1」とは言えなくなるのか、せっかく Excel が使えるのだから、試してみましょう。 $p = 0.1$ とすると、 $0.9^{100} = 2.6561 \times 10^{-5}$ となりますから、「おおよそ 1」で良いでしょう。

もっと歪んだコインにすると、 $p = 0.01$ で $0.99^{100} = 0.36603$ ですから、3 分の 2 くらいになり、「おおよそ 1」とは言えなくなります。

練習問題 2.5 確率変数 X は、1 と 6 の目がほかの目よりも 2 倍出やすいという（直方体の）さいころを振って出た目を 3 で割った「あまり」を表すものとします。 X の確率関数を求めなさい。

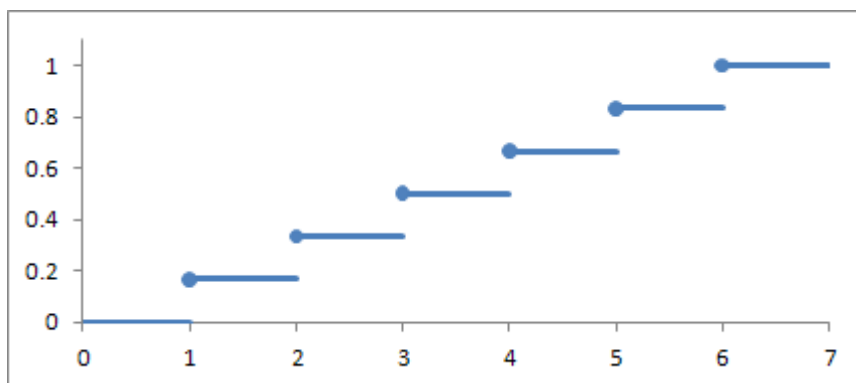
例題 2.9 正しいさいころを振った時に出る目の確率分布の確率関数を求めなさい。累積分布関

数のグラフを描きなさい。

こたえ 確率関数は以下の通り。

$$P(X = i) = \frac{1}{6}, \quad i = 1, 2, 3, 4, 5, 6 \quad (2.6)$$

累積分布関数は 1, 2, 3, 4, 5, 6 で $\frac{1}{6}$ ずつジャンプして 6 で 1 になるような階段関数。



さいころの目の累積分布関数

エクセルヒント：グラフの描き方

Excel で階段関数を描く方法

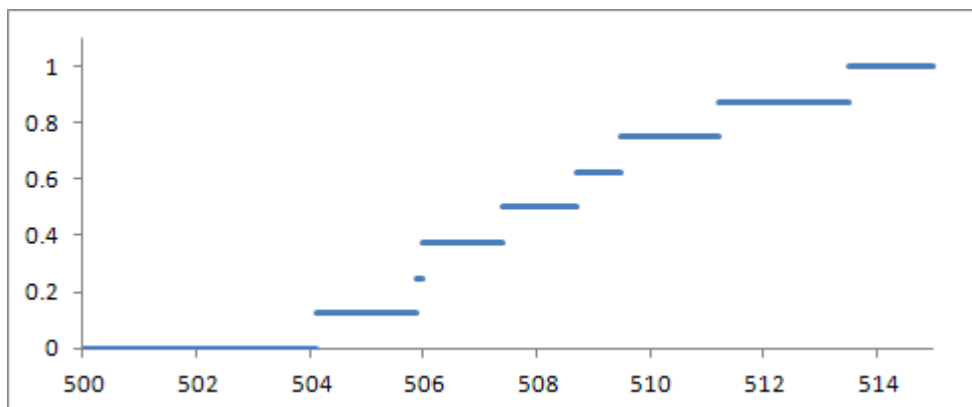
折れ曲がる点を全部入力し、それらの点をデータと思って（各点を結ぶ）散布図を描けばよい。縦棒を表示させないようにするには、間に空行を入れればよい。

おわり

例題 2.10 次のデータが与えられたとき、経験分布関数を描きなさい。

507.4, 509.5, 506.0, 513.5, 511.2, 508.7, 504.1, 505.9

こたえ 経験分布関数は、データの個数が n 個あったとき、各データの値を $1/n$ で取る離散一様分布する確率変数 X の累積分布関数と同じことです。言い換えれば、 n 面体のさいころの各面に 1, 2, 3, ... ではなくてデータを書いたようなものを振った時に出る目を X としたとき、 X の累積分布関数ですから、前の例題のようなグラフになります。ただし、ジャンプする場所が不等間隔で、1 回のジャンプ幅が $1/n$ になります。

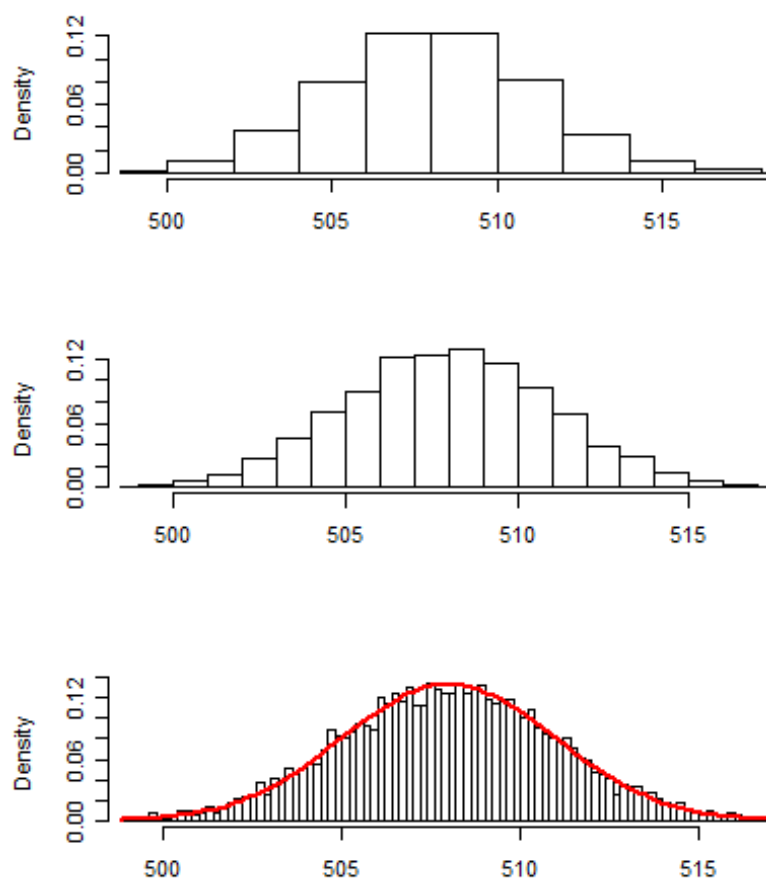


練習問題 2.6 あるパソコンショップの1日の売り上げ台数（せいぜい5台程度しか売れない）を20日間記録したものを仮想的に作成し、1日の売り上げ代数の経験分布関数を描きなさい。

2.4 密度関数と累積分布関数

密度関数は確率ではない、ということについて補足説明します。データを集めてヒストグラムを作るところから話を進めましょう。たとえばペットボトルの内容量、測定データをたくさん集めて、区間幅を適当に決めて度数分布を作り、ヒストグラムを描くのですが、このとき、棒の面積を全部足したものが1になるように高さを調整してやります。そうすると、各区間の代表値を $a_1, a_2, \dots$ 、クラスの幅を $2h$ としたとき、各棒の面積はペットボトルの内容量がたとえば $a_i - h$ 以上 $a_i + h$ 未満の確率を与えるといっていよいでしょう。棒の高さを $f(a_i)$ とすれば、その確率は $f(a_i) \times 2h$ で与えられます。

区間の幅を半分にすると、全部の棒の合計面積を1としたままだと、各棒が半分に分割されたようなヒストグラムになり、棒の高さ $f(a_i)$ はほとんど変わらず、面積（確率）が半分になります。全部の棒の合計面積を1としたまま、区間の幅をどんどん狭くしていくと、棒の面積（確率）はますますゼロに近づきますが、（データが十分にあれば）ヒストグラムの上辺はなめらかな曲線 $f(x)$ に近づくでしょう。それが真のペットボトルの内容量という連続型確率変数 X の密度関数です。



したがって、 X の密度関数 $f(x)$ はその作り方から、

1. 非負: $f(x) \geq 0$
2. 全域で積分すると 1
3. 区間 $[a, b]$ で積分すると、 X が a 以上 b 以下となる確率が得られる

特に $a = b$ とすると、 $X = a$ となる確率はゼロ、つまり、密度関数 $f(a)$ は $X = a$ となる確率ではありません。たとえば、上の図で 505 から 510 までの間の面積（積分）は、ペットボトルの内容量が 505 以上、510 以下の確率を表しています。つまり、確率を計算するには定積分の計算が必要ですが、もし、密度関数 $f(x)$ に不定積分 $F(x)$ があれば、いちいち定積分を計算しなくても、 $F(510) - F(505)$ で求められることは知っています。この不定積分で、 $F(\infty) = 1$ としたものを累積分布関数と言います。

数学の知識（微積分の基本定理！）から、累積分布関数 $F(x)$ を微分すると密度関数 $f(X)$ が得られ、密度関数を $-\infty$ から x まで積分したものが $F(x)$ になります。このことから、確率変数 X の累積分布関数 $F(x)$ の性質として、次が導かれます。

1. $F(x)$ は $X \leq x$ の確率を表す。
2. $F(x)$ は単調非減少（非負の密度関数を積分したものだから。あるいは、導関数が非負だから）
3. $F(-\infty) = 0$ （ $X \leq -\infty$ は空事象を表し、その確率はゼロだから）
4. $F(\infty) = 1$ （ $X < \infty$ は全事象を表し、その確率は 1 だから）

例題 2.11 次の二つの関数はいずれも確率変数 X の密度関数を表すものとします。(1) 密度関数の性質を使って c の値を決めなさい。(2) 密度関数の略図を描きなさい。(3) 累積分布関数を計算しなさい。(4) 累積分布関数の略図を描きなさい。

$$f_1(x) = \begin{cases} c & (0.5 \leq x \leq 1) \\ 0 & \text{それ以外} \end{cases}, \quad f_2(x) = \begin{cases} c(2 - |x|), & |x| \leq 2 \\ 0, & |x| > 2 \end{cases}$$

こたえ 密度関数は非負で、全域で積分すると 1 になるということを使います。

$f_1(x)$: (1) 全域で積分すると 1 になるのだから、 $c = 2$ 。(2) 省略。(3)

$$F_1(x) = \begin{cases} 0, & x < 0.5 \\ 2(x - 0.5), & 0.5 \leq x < 1 \\ 1 & x \geq 1 \end{cases}$$

(4) 省略

$f_2(x)$: (1) 同様に考えると $c = 0.25$ 。(2) 底辺が $[0, 2]$ 、高さ 0.5 の二等辺三角形、省略。(3) $x < 0$ のとき

$$F_2(x) = \int_{-2}^x \frac{1}{4}(2+x)dx = \left[\frac{x}{2} + \frac{x^2}{8} \right]_{-2}^x = \frac{x}{2} + \frac{x^2}{8} + \frac{1}{2}$$

$x > 0$ のときは、0 までの積分の値を足すことを忘れないように。

$$\begin{aligned} F_2(x) &= \int_{-2}^0 \frac{1}{4}(2+x)dx + \int_0^x \frac{1}{4}(2-x)dx = \frac{1}{2} + \left[\frac{x}{2} - \frac{x^2}{8} \right]_0^x \\ &= \frac{x}{2} - \frac{x^2}{8} + \frac{1}{2} = 1 - \frac{1}{8}(x-2)^2 \end{aligned}$$

まとめると、

$$F_2(x) = \begin{cases} 0, & x < -2 \\ \frac{1}{8}(x+2)^2, & -2 \leq x < 0 \\ 1 - \frac{1}{8}(x-2)^2, & 0 \leq x < 2 \\ 1, & x \geq 2 \end{cases}$$

(4) 二つの放物線を $(0, 0.5)$ でつなげたもの、省略

練習問題 2.7 確率変数 X の累積分布関数が次のように与えられているものとします。(1) 累積分布関数の満たす条件を書き、 $F_2(x)$ がその条件を満たしていることを確かめなさい。(2) それぞれ対応する密度関数を計算しなさい。(3) 密度関数のグラフを描きなさい。

$$F_1(x) = \begin{cases} 0, & x \leq -1 \\ \frac{1}{2}x + \frac{1}{2}, & -1 < x < 1 \\ 1, & x \geq 1 \end{cases}, \quad F_2(x) = \begin{cases} 0, & x \leq 0 \\ -x^2 + 2x, & 0 < x < 1 \\ 1, & x \geq 1 \end{cases}$$

練習問題 2.8 (正規分布) 平均 8、標準偏差 2 の正規分布の概形を描きなさい (1 回微分して極値を求め、2 回微分して変曲点を求め、関数の増減表を書き、いくつかの通る点を押さえること)。Excel を使って確認しなさい (正規分布の密度関数を計算する関数は `NORMDIST()`)。最初から Excel を使ってはいけません。

練習問題 2.9 (指数分布) $x \geq 0$ で $F(x) = 1 - e^{-2x}$ 、 $x < 0$ で $F(x) = 0$ という関数のグラフを描き、それが累積分布関数の性質を備えていることを示しなさい。また、その導関数を計算してそのグラフを描きなさい (この累積分布関数を持つ確率変数は**指数分布**に従う、と言われます)。

例題 2.12 確率変数 X は、区間 $\{1, 2, 3, 4, 5, 6\}$ 上で一様分布するものとします。また、 $Z = 2X + 2$ と定義します。このとき、 X の確率関数 $f_X(x)$ と、 Z の確率関数 $f_Z(x)$ を書きなさい。

こたえ $f_X(x) = \frac{1}{6} (x = 1, 2, 3, 4, 5, 6)$ 。 $Z = 2X + 2$ の取りうる値は 4, 6, 8, 10, 12, 14、一様分布することは同じなので、 $f_Z(x) = \frac{1}{6} (x = 4, 6, 8, 10, 12, 14)$

例題 2.13 X の密度関数を $f_X(x) = \frac{1}{c} (0 < x < c)$ とします (それ以外は 0)。確率変数 Y を $Y = aX + b (a > 0)$ と定義したとき、 Y の密度関数 $f_Y(x)$ を求めるために、以下の問いに答えなさい。(1) 事象 $\{Y \leq x\}$ を X を使って表しなさい。(2) Y の累積分布関数 $F_Y(x)$ を、 X の累積分布関数 $F_X(x)$ を使って表しなさい。(3) 累積分布関数を微分することにより、 Y の密度関数を $f_X(x)$ を使って表しなさい。

こたえ (1) $\{Y \leq x\} \Leftrightarrow \{aX + b \leq x\} \Leftrightarrow \{X \leq \frac{x-b}{a}\}$ 。(2) 事象が等しければ確率も等しい、したがって

$$P(Y \leq x) = P\left(X \leq \frac{x-b}{a}\right)$$

これを累積分布関数を使って表現すると

$$F_Y(x) = P(Y \leq x) = P\left(X \leq \frac{x-b}{a}\right) = F_X\left(\frac{x-b}{a}\right)$$

(3)

$$f_Y(x) = \frac{d}{dx} F_Y(x) = \frac{d}{dx} F_X\left(\frac{x-b}{a}\right) = \frac{1}{a} f_X\left(\frac{x-b}{a}\right)$$

よくある間違い $Y = aX + b$ ならば $F_Y(x) = aF_X(x) + b$ 、なんていう「ルール」はありません。確実に分かる知識だけを使って、地道に計算しなさい。

練習問題 2.10 (標準正規分布) X はパラメータ μ, σ の正規分布に従う確率変数とし、 $Z = (X - \mu)/\sigma$ ($\sigma > 0$) とします。このとき、 Z の密度関数を計算しなさい。

例題 2.14 (参考：正規分布の密度関数) 参考 (正規分布) 次の関数が密度関数になっていることを証明しなさい。

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \quad (2.7)$$

ただし μ と $\sigma (\neq 0)$ は任意の定数、 π は円周率、 $\exp(x)$ は e^x の別記法で、 e は自然対数の底の数でネイピア数と呼ばれる定数 ($= 2.718281828\dots$) である。

こたえ $f(x)$ の $(-\infty, \infty)$ における定積分が 1 になることを言えばよい。 $z = (x - \mu)/\sigma$ と変数変換をすると、

$$I = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz \quad (2.8)$$

と表される。 $e^{-z^2/2}$ の定積分の 2 乗

$$J = \left(\int_{-\infty}^{\infty} e^{-z^2/2} dz\right)^2 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-(u^2+v^2)/2} du dv \quad (2.9)$$

を考える。 $u = r \sin \theta, v = r \cos \theta$ という変数変換をすると、

$$J = \int_0^{\infty} \int_0^{2\pi} r e^{-r^2/2} dr d\theta = 2\pi \int_0^{\infty} r e^{-r^2/2} dr = 2\pi \left[-e^{-r^2/2}\right]_0^{\infty} = 2\pi \quad (2.10)$$

したがって、

$$\int_{-\infty}^{\infty} e^{-z^2/2} dz (= \sqrt{2\pi} I) = \sqrt{2\pi} \quad (2.11)$$

が成り立つ。したがって、 $I = 1$ となる。