

確率

第 1 回講義で、「くじびきが当たる確率 50%」などと、はじめから簡単に「確率」という言葉を使いました。ここで、「確率」の意味をもう一度よく考えてみましょう。

「可能性」の集合

いま、くじをひくと、当たりが出たとします。現実世界では、くじは確かに当たったのであって、それ以外の結果は現れていません。

しかし、われわれは、くじびきとはいつも当たるものではなく、いま現れている「当たり」は偶然による結果だということを知っています。「偶然による」というのは、他の可能性もあった、つまり偶然によって他の結果になるかもしれない、ということの意味をしています。この例の場合ならば、「はずれ」が出るという可能性もあった、ということになります。このような「結果が偶然によって決まる現象」を**ランダム現象**といいます。

統計学の世界では、つねに、この「可能性の集合」を念頭において、考えを進めます。この例の場合ならば、「今は『当たり』という結果が現れたが、『はずれ』が現れる可能性もあった」と考えている、ということです。

そして、さらに「どの結果が、どのくらい現れやすいか」を考えます。これを数字で表したのが**確率**です。「現れやすさ」などというものを、どのように数字で表せばよいのでしょうか。ひとつの考え方は、下のようなものです。

ある結果が現れる確率とは、

**これからその結果が現れる可能性のある 十分多くの回数の機会があるとき、
そのうち本当にその結果が現れる回数の割合である。**

次にその結果が現れる確率とは、

**遠い将来までの十分多くの回数の機会を考えて初めて言える「結果の回数の割合」を、
次の 1 回の機会にあてはめて述べたものにすぎない。**

例えば、くじ引きを**十分多くの回数**行なうとき、10 回に 3 回の割合で当たりが出るとすれば、「あたりが出る確率」は $1/3$ であると考えます。

ただし、次にくじを 1 回ひくとき、当たりが出るかどうかは何とも言えません。ただ、「これからもくじをひきつづけると、長い目で見れば 10 回に 3 回の割合で当たりが出るだろう」という数値で、次の 1 回の機会での当たりくじの「出やすさ」を表現しようというのが、確率の考え方です。

ギャンブルの例で言えば、プロのギャンブラーは日常的に多くの賭けをし、長い目で見た利益を考えていますから、常に確率が大きい方に賭けるほうが有利です。しかし、1 回しか賭けをしない人にとっては、「確率が大きい」ことと「次の賭けで勝てる」こととは直接は結びつかないことになります。

ここでいう「あたりが出る」などの「結果」を、確率論の言葉では**事象**といいます。また、事象が起きる機会、この例ならば「くじを引くこと」を**試行**といいます。また、このような確率の考え方を、**頻度による確率の定義**（統計的確率）といいます。つまり、確率とは「特定の結果がおきる回数の割合」ですから、その値は 0 から 1 (0%~100%) の範囲になります。

さいころの各目が出る確率はどれも $1/6$ か？

高校までの教科書で確率を学ぶ時には、「さいころの各目が出る確率は、いずれも $1/6$ である」ということを前提にしていたと思います。

しかし、頻度による確率の定義から考えれば、次にさいころをふったときにある目が出る確率は、十分に多くの回数さいころを振ってみななければわからないことになります。しかも、「十分に多くの回数」振らなければなりません、何回なら十分なのでしょう？ 実は、数学でいう「十分に多く」というのは、「誰も文句を言わないくらい多く」という意味であって、何回振っても十分ではないのです。

また、さいころを1万回ふって、そのうち1の目が $1/6$ の割合で出たとしても、それはあくまで「過去の実績」であって、その次に1万回さいころをふっても、1の目は1回も出ないかもしれません。つまり、頻度による定義では、現実には確率を定めることはできないことになります。

では、なぜ「さいころの各目が出る確率は、いずれも $1/6$ である」と言われているのでしょうか？ それは、

1. 各目が同じ確率で出る
2. 各目が出る確率は、いつさいころを振っても同じである

ということをお互いが認めているからです。そこで「さいころには全部で6種類の目があって、いずれの目も常に同じ確率で出るから、各目が出る確率は $1/6$ 」ということになります。

高校までに習った確率の問題は、このような仮定を認めたうえで、確率すなわち「特定の結果が現れる回数の割合」の問題を、「(さいころの目の種類などの) 可能性のある結果の種類」の問題に置き換えたものです。このような確率の考え方を**ラプラスの定義**（数学的確率）といいます。

しかし、このラプラスの定義も、よく考えるとおかしいところがあります。上で「このような仮定を認めれば」と書きましたが、これが認められるかどうかは、さいころを十分な回数振ってみないとわかりません。これでは堂々めぐりです。

つまり、確率の定義にはどのように考えてもあやしいところがあります。確率は、**遠い将来までを長い目で見てはじめて言える「特定の結果が現れる回数の割合」を、次の1回の機会にあてはめて述べたものにすぎません。** また、確率の定義には「十分多くの回数さいころを投げる」という現実には実行不可能な操作や、「各目が同じ確率で出る」という真偽を確かめられない仮定が含まれています。ですから、確率は**測定するものではなく、何らかの仮定をおいて「定義する」もの**なのです。この講義で扱う統計学では、概ね常識的に確率を理解しておけば十分ですが、ここまで述べた確率の「あやしき」は承知しておいてもらいたいと思います¹。

条件付き確率と「独立」

統計学では、「独立」という言葉がよく出てきます。これは、簡単にいえば、2つのランダム現象があるとき、一方の結果がもう一方の結果に影響しない、という意味です。例えば、2つのくじ引きがあるとき、一方に当たるともう一方にも当たりやすくなる、というときは、2つのくじ引きは独立ではありません。

¹現代の数学では、確率は現実の問題から離れて、集合を測る尺度（測度）のひとつとしてとらえられています。

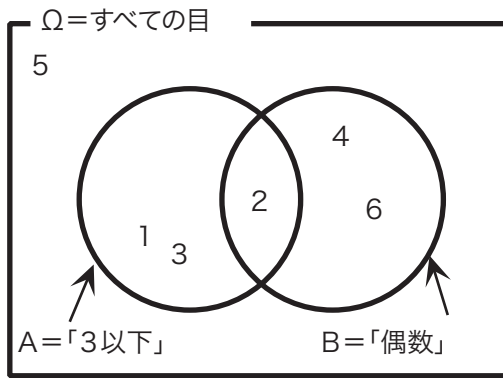


図 1: 2つの事象とベン図

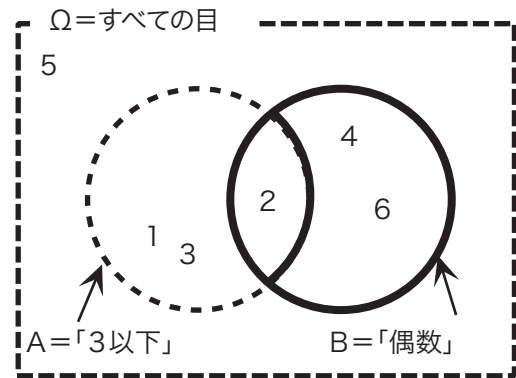


図 2: 条件付き確率

独立の概念は、正確には**条件付き確率**を使って定義されます。単に「明日雨がふる確率」よりも、「明日雨が降るといふ予報が出たときに、本当に雨がふる確率」のほうが大きい、というのは、日常感することです。後者のような確率が条件付き確率とよばれるものです。以下では、その意味を、さいころの各目が出る確率を例にとって説明します。

さいころで、「3以下の目が出る確率」を図に表すことを考えます。さいころで、「可能なすべての目」は1, 2, 3, 4, 5, 6の6通りで、これを集合 Ω で表します。一方、「3以下の目」は1, 2, 3の3通りで、これを Ω の内部にある集合 A で表します。

このとき、「3以下の目が出る確率」は、集合 A の要素がおきる確率なので、「事象 A がおきる確率」で、 $P(A)$ で表します。 $P(A)$ は、「集合 A の要素の数」を $|A|$ で表すと、

$$P(A) = |A|/|\Omega| = 3/6 = 1/2 \quad (1)$$

となります。

さらにもうひとつ、「偶数の目が出る確率」を考えます。同様にして、「偶数の目」は2, 4, 6の3通りで、これを集合 B で表すと、「偶数の目が出る確率」 $P(B)$ は

$$P(B) = |B|/|\Omega| = 3/6 = 1/2 \quad (2)$$

となります。これらを目に見えるように表したのが「ベン図」で、図1となります。

では、「3以下かつ偶数の目が出る」確率を考えましょう。この事象は集合 $A \cap B$ で表されますから、その確率 $P(A \cap B)$ は

$$P(A \cap B) = |A \cap B|/|\Omega| = 1/6 \quad (3)$$

となります。

ここで、 $|A \cap B|/|B|$ という確率を考えてみましょう。図2の太線の部分です。分母が $|\Omega|$ から $|B|$ に変わっていますから、ここでは、「偶数の目」が、ここでの「可能なすべての目」になっています。一方、 $A \cap B$ は「3以下かつ偶数の目が出る」という事象ですが、今は「偶数の目が出る」という事象の中でしか考えていませんから、この事象は単に「3以下の目が出る」という事象ということができます。したがって、

$|A \cap B|/|B|$ = 偶数の目が出るとうわかつている時（偶数の目が出るのが確実な時）、それが3以下である確率

になります。

これを、「**Bを条件とするAの条件付き確率**」といい、 $P(A|B)$ で表します。 $P(A|B) = |A \cap B|/|B| = 1/3$ ですから、「偶数の目が出た」という情報が得られている時は、そうでないときよりも「3以下の目が出る」確率は小さくなることがわかります。

ところで、

$$P(A|B) = \frac{|A \cap B|}{|B|} = \frac{|A \cap B|/|\Omega|}{|B|/|\Omega|} = \frac{P(A \cap B)}{P(B)} \quad (4)$$

と表され、これを条件付き確率の定義としている本もあります。ただし、この場合、分母分子それぞれの確率は、いずれも同じ $|\Omega|$ を分母とする確率でなければならないことに、注意する必要があります。また、(4)式から

$$P(A \cap B) = P(A|B)P(B) \quad (5)$$

となります。(5)式は、簡単に言えば

「**AとBの両方が起きる確率**」 = 「**Bが起きたとしたときにAが起きる確率**」
× 「**本当にBが起きる確率**」

ということです。 $P(A|B)$ と $P(A \cap B)$ の違いも、これでわかると思います。

では、上の例の事象Aが、「3以下の目」ではなく「2以下の目」だったらどうでしょう。このときは、「2以下の目が出る確率」 $P(A) = 1/3$ です。一方、 $P(A \cap B) = 1/6$ や $P(B) = 1/2$ は変わりませんから、 $P(A|B) = |A \cap B|/|B| = 1/3$ もかわりません。

したがって、このときは $P(A|B) = P(A)$ となります。このときは、「事象Aが起きる確率」と「事象Bが起きるとわかっているときに、事象Aが起きる確率」が同じですから、事象Bが起きるかどうかには関係がないことを意味しています。このとき、事象Aと事象Bは**独立**であるといいます。

事象Aと事象Bが独立のとき、(4)式から

$$P(A \cap B) = P(A)P(B) \quad (6)$$

となります。事象Aと事象Bが独立のときこうなるのであって、いつもこうなるのではないことに注意してください。

確率のパラドックス

次の問題を考えてみましょう。

囚人が3人いて、「A, B, Cの3人のうち、2人は明日処刑される。あとの一人は釈放される。誰が処刑されるかは明日朝発表される」と言い渡されています。以下は、処刑前夜の囚人Aと看守の会話です。

囚人A「看守さんは、明日誰が処刑されるかを知っているそうですね。私が明日処刑されるかどうかを、教えてくれとは言いません。ただ、私以外のB, Cのうち、どちらが処刑されるかを教えてもらえませんか？」

看守「本当は言うてはいけないのだが... 実は、Cは処刑されるのだ。」

囚人A「よく教えてくださいました。これで少し安心しました。」

看守「なぜだ？」

囚人A「看守さんの話を聞く前は、私が処刑される確率は2/3でした。ところが、看守さんにCが処刑されると教えていただきました。ということは、あとは私とBの2人にひとりだけが処刑されるので、私が処刑される確率は1/2に減りました。」

さて、なにかおかしいと思いませんか？

A, Cが処刑される確率をそれぞれ $P(A)$, $P(C)$ とすると、求める確率は $P(A|C)$ で、 $P(C) = 2/3$ で $P(A \cap C) = 1/3$ だから、 $P(A|C) = P(A \cap C)/P(C) = 1/2$ でいいんじゃないですか？

この問題に答えるポイントは、「『Cが処刑される』という情報は、Aが処刑されるかどうかについての手がかりになるか？」ということです。

実は、看守が同じように「Cは処刑される」と答えても、その答えに至る過程によって、「Aが処刑される確率」は変わってきます。それはなぜか、以下の例で考えてみましょう。

看守が「Cは処刑される」と答えるに至る、以下の3通りの場合を考えます。看守は判決を知っており、その判決は「Bは釈放され、A, Cが処刑される」とであると仮定しましょう。

1. 看守は、「A, B, Cのどれについて答えるか」を確率1/3ずつで決めるくじをひき、Cについて答えることになったので、「Cは処刑される」と答えた。
2. 看守は、Aについては何も答えないことに決めていたので、BとCのどちらについて答えるかを確率1/2ずつで決めるコイン投げを行ない、Cについて答えることになったので、「Cは処刑される」と答えた。
3. (問題文と同じ例) 看守は、Bが釈放されることを答えてしまうと、Aが処刑されることを答えるのと同じことになってしまうので、あえてBについては触れず、「Cは処刑される」と答えた。

これらの3通りの例は、いったいどう違うのでしょうか。

1. の場合、看守のくじの結果は「Cについて答える」となりましたが、他に「Aについて答える」「Bについて答える」という結果になる可能性もありました。つまり、

- Cについて答える → Cは処刑 → Aの運命は未定
- Bについて答える → Bは釈放 → Aの処刑が確定
- Aについて答える → Aは処刑 → Aの処刑が確定

の3通りの可能性のうち、実際に起きたのはいちばん上のできごとですが、下の2通りが起きるかもしれないなかったわけです。

ここで、「ある事象が起きる確率」とは、「その事象以外にすべての可能な事象を考えて、それらの事象が起きうる試行を十分に多くの回数行なうとき、その特定の事象が起きるような試行の回数の割合」であることを思い出してください。

1. の場合にも、「看守がくじをひく」ことを、仮に十分に多くの回数行なうとしましょう。このときの「起きうる事象」には、上の3通りの可能性があります。ところが、1. の文では、3通りの可能な事象のうち、いちばん上の「くじの結果、Cについて答えると決まった」場合しか考えていません。

ということは、「十分に多くの試行」の回数から、下の2通りの事象が起きた試行を除いたうえで、確率を計算する必要があります。全体の試行の回数が変わりますから、Aが処刑される確率も、看守がくじをひく前に比べて変化します。この「全体の試行の回数を変える」計算とは、すなわち、前節で述べた「条件付き確率」の計算ということになります。

また、2. の場合、「Aについては何も答えないことに決めていた」のは確かですが、この場合でも

- Cについて答える → Cは処刑 → Aの運命は未定
- Bについて答える → Bは釈放 → Aの処刑が確定

の2つの可能性があり、そのうち上の場合しか考えていません。したがって、上の例と同様に、やはりこのコイン投げはAが処刑される確率に影響しています。

ところが、3. の場合、つまり本文の問題文の例では、「Cについて答える」ことは初めから決まっておき、他の可能性はありません。したがって、看守が「Cについて答えた」ことがわかっても、上の説明でいう「十分に多くの試行」の回数は変わりません。したがって、Aが処刑される確率も変わらないことになります。つまり、この問題の答は「Aが処刑される確率は、看守の言葉を聞いたとしても、依然 $2/3$ 」ということになります。

このように、実際に看守が「Cが処刑される」と答えたのは同じでも、「他にどのような可能性があったのか」が、確率の計算に影響するのです。このことは、看守の答は同じでも、看守の「癖」や「心の中」が問題になる、ということの意味しています。しかし、心の中を客観的に調べることは、通常はできません。本文の問題文では、看守は、Aの運命を答えてしまうことにならないように「Cが処刑される」と答えたと思われそうですが、それが本当かどうかは囚人Aにはわかりません。ですから、この問題の答えは、「看守の言うことを信じれば」確率はこうなる、という意味でしかありません。「確率は、測るものではなく、定義するもの」なのです。

上の「看守と囚人」の問題は、「モンティ・ホールのパラドックス(逆説)」として知られているのと同じものです。インターネットで「モンティ・ホール」で検索すると、いろいろな解説が出てきます。講義のウェブサイトの「統計データ・ツールへのリンク」には、この問題をはじめ、ルーレットやさいころにおける確率の問題についての解説記事へのリンクを掲載しています。