期待値

確率の研究は、賭け事の疑問を解消することから始まったと言われている。1 個のさいころを投げることを 4 回続けて、少なくとも 1 回は 6 の目を出す賭けでは、賭けを持ちかけた張本人が実際に 6 の目を出すことができて、賭けに勝ち続けた。次に、2 個のさいころを同時に投げることを 24 回続けて、少なくとも 1 回は目の和が 12 となる賭けでは、張本人は思い通りに目の和を 12 にできず、賭けに負け続けた。たしか、こんな話だったと思う。

賭けを仕掛けた本人は、1 個のさいころで 6 の目が出る確率が $\dfrac{1}{6}$ で、2 個のさいころで目の和が 12 になる確率が $\dfrac{1}{36}$ であることは知っていた。だから 1 個のさいころなら、4 回投げれば $\dfrac{1}{6}\times 4=\dfrac{2}{3}$ の割合で有利と考え、同様に 2 個のさいころなら、24 回投げれば $\dfrac{1}{36}\times 24=\dfrac{2}{3}$ の割合で同様に有利と考えたようである。

思い通りの結果にならなかった彼が頼ったのがパスカル[1]であった。パスカルはフェルマー[2]と手紙のやり取りをしながら研究したことが、後の確率論や統計論に発展していったということである。

いま例に挙げたさいころの賭けは、反復試行の確率の計算をすることで、どちらがどれだけ有利または不利なのか求められるのだが、この章では賭け事などに欠かせない知識を得ることを優先しよう。

X	1 等（1,000 円）	2 等（500 円）	3 等（100 円）	はずれ（0 円）	計
本数	1（本）	10（本）	25（本）	64（本）	100（本）
確率 (p)	$\frac{1}{100}$	$\frac{10}{100}$	$\frac{25}{100}$	$\frac{64}{100}$	1
$X\times p$	10（円）	50（円）	25（円）	0（円）	85（円）

いま、表のような宝くじがあるとしよう。各賞が当たる確率は本数に占める割合であるから、それが表中に示された確率である。このとき、1,000 円の値とそれが当たる確率 $\dfrac{1}{100}$ の積を求めると 10 円になる。これは他の賞に対しても行えて、それぞれの積が $X\times p$ の欄に書かれている。このとき $X\times p$ の総和—表では計の欄の 85 円—を**期待値**と呼ぶ。期待値の呼び方から連想できるように、この値が宝くじ 1 本の価値を表していることに気づくかもしれない。

一般には期待値とは、確率によって与えられる変数 X が $x_1, x_2, \ldots, x_n$ のいずれか 1 つの値をとり、各々の確率が $p_1, p_2, \ldots, p_n$（ただし、$p_1+p_2+\cdots+p_n=1$）で与えられるとき、

$$\boxed{E(X)=x_1p_1+x_2p_2+\cdots+x_np_n}$$

[1] ブレーズ・パスカル（1623–1662）: フランスの哲学者・数学者。
[2] ピエール・ド・フェルマー（1607?–1665）: フランスの数学者。

を、**確率変数** X **の期待値**と呼ぶ。確率変数とは、いろいろな状況に応じて、いろいろな値をとる変数のことで、この場合は、くじを引いて起こるいろいろな等級に応じて支払われる、いろいろな金額をとる値のことである。

*　*　*

1 個のさいころを投げることを 4 回続けて、少なくとも 1 回は 6 の目を出す賭けについて触れておこう。この賭けに負ける場合は、さいころを 4 回投げて 4 回とも 6 以外の目が出たときである。それは

$$\left(\frac{5}{6}\right)^4 = \frac{5^4}{6^4} \approx 0.4822$$

である。これ以外の場合は、4 回のうち少なくとも 1 回はどこかで 6 の目が出ているので、逆に勝てる確率は $1 - 0.4822 = 0.5178$ で 5 割を超えている。1 回 1 万円の賭けとすると、期待値は

$$10000\text{ 円} \times 0.5178 + (-10000\text{ 円}) \times 0.4822 = 356\text{ 円}$$

である。1 万円が 356 円の利益を生むのだから、3.56%の利回りということになる。

2 個のさいころを同時に投げることを 24 回続けて、少なくとも 1 回は目の和が 12 になる賭けの場合に、賭けに負ける場合の計算をすると

$$\left(\frac{35}{36}\right)^{24} = \frac{35^{24}}{36^{24}} \approx 0.5085$$

であるから、すでに負ける確率が 5 割を超えている。同様に期待値を計算すると

$$10000\text{ 円} \times 0.4915 + (-10000\text{ 円}) \times 0.5085 = -170\text{ 円}$$

である。▮

期待値の意味すること

期待値を定義したところで、期待値が何を意味するか述べておこう。確率変数 X の各々に対する確率が p_1、 p_2、$\ldots$、p_n（ただし、$p_1 + p_2 + \cdots + p_n = 1$）で与えられるということは、確率変数が全部で n 通りあり、それぞれが前述の宝くじの本数のように、a_1 通り、a_2 通り、$\ldots$、a_n 通り（ただし、$a_1 + a_2 + \cdots + a_n = m$）と数えられるとしても同じことである。この場合、たとえば $p_1 = \dfrac{a_1}{m}$ であることに注意されたい。すると期待値の計算は

$$\begin{aligned} E(X) &= x_1p_1 + x_2p_2 + \cdots + x_np_n \\ &= x_1\frac{a_1}{m} + x_2\frac{a_2}{m} + \cdots + x_n\frac{a_n}{m} \\ &= \frac{x_1a_1 + x_2a_2 + \cdots + x_na_n}{m} \end{aligned}$$

としていることと同じことである。この最後の計算式は、加重平均を求める計算でもある。すなわち期待値とは、平均値を求めていることと同じなのである。

すると先の宝くじの例は、

くじ 1 本当たりの平均的価値が 85 円である

ことに等しいことになる。もし、くじが 1 本 100 円で売られているならば、“損” なくじである。では、いったい何が損なのだろうか。むしろ、1,000 円や 500 円が当たれば “得” であり、その確率は 11%強にもなっているのである。

さて、何かが見えてきたかもしれない。期待値が平均値を表すというとき、それは大数の法則のように母数が非常に多いときの平均値であり、また、同じことを大量に繰り返したときの平均値なのである。だから、例に挙げた宝くじを買って 500 円を得たら、それは期待値より「高い値」を得たのである。くじに当たらなかったら、それは期待値より「低い値」を得たのである。このように考えると期待値は単に基準値であり、とりわけ損得勘定があるわけではないのである。

しかし 1 人の人間には、基準値としての期待値の値は意味をなさない。1 人で宝くじのほとんどを購入するならば、1 本当たりの利益は基準値に近づくだろうが、少量の宝くじを買った 1 人の人間ということなら、大きく当たることもあれば、かすりもしないこともある、というに過ぎないのである。

*　*　*

宝くじは買った分の何割かしか戻ってこないとか、ギャンブルは決して儲からない、などと言うとき、そこには期待値の確固たる計算結果が根拠になっている。たしかに宝くじやあらゆるギャンブルは、参加料に対する返戻金の割合は低い。言い換えれば、胴元が必ず利益を得られるようになっている。たとえば、何万枚も用意した宝くじの売れ行きが極端に悪く、ほんの数百枚しか売れず、しかも売れたくじがほとんど高額当選であったために胴元は大赤字だった、などということは考えられないだろう。宝くじは何万枚も売れるから、そこに大数の法則が働き、結果として胴元が利益を得るのである。

ギャンブルや株の売買で、個人が利益を得るのはなかなか難しい。たとえば競馬において自分なりの予想法で買い続けた場合、長い間には大抵損をするはずである。それは、自分なりの予想と言いながらも、長く続けることで結果的に無作為な予想と大差なくなってしまうからであろう。そうなると大数の法則が働きやすいのである。もし利益を得られるとしたら、少なくとも、長く続けても無作為にならない予想法でなくてはならないだろう。▌

期待値の計算例

期待値は事象に対して考えるだけでなく、試行に対しても考えられる。以前、日本シリーズの勝敗について調べたときは、それぞれの事象が起こる割合を計算した。そこで求めた確率を試合数に乗ずると、試合数の期待値、すなわち平均して何試合行われるかが求められる。

試合数 (X)	4 試合	5 試合	6 試合	7 試合	計
確率 (p)	$\frac{1}{2^4} \times 2$	$\frac{4}{2^5} \times 2$	$\frac{10}{2^6} \times 2$	$\frac{20}{2^7} \times 2$	1
$X \times p$	$\frac{8}{2^4}$	$\frac{40}{2^5}$	$\frac{120}{2^6}$	$\frac{280}{2^7}$	$\frac{744}{2^7} = 5.8125$

表は、先の日本シリーズの勝敗に対する割合を、合計試合数に対する確率として計算し直したものである。この結果、日本シリーズにおける試合数は平均すると約 5.8 試合行われているはずである。実際、2023 年シリーズまでの加重平均は—引き分けを無視して—約 5.81 試合であるから、実によいところをついている。

ただ、平均値というのは分布を正確に教えてくれるものではない。平均が 5.8 試合と聞くと、多くが 6 試合で決着し、5 試合や 7 試合で決着する場合は若干 5 試合の方が多いのではないかと思うかもしれない。しかし実際は、6 試合と 7 試合で決着することが大半で、4 試合で決着することも少なからずあるのである。大半が 6 試合か 7 試合で決着しながら、加重平均が 6.5 あたりの値になっていないのは、4 試合で決着する試合数の影響が大きいためである。期待値は注意して見る必要があるということである。

さいころの賭けの期待値

さいころを投げることを 4 回続けて、少なくとも 1 回は 6 の目を出す賭けにおいては、期待値はどうなっているだろうか。

6 の目 (X)	0 回	1 回	2 回	3 回	4 回	計
確率 (p)	$(\frac{1}{6})^0(\frac{5}{6})^4$	${}_4C_1(\frac{1}{6})^1(\frac{5}{6})^3$	${}_4C_2(\frac{1}{6})^2(\frac{5}{6})^2$	${}_4C_3(\frac{1}{6})^3(\frac{5}{6})^1$	$(\frac{1}{6})^4(\frac{5}{6})^0$	1
$X \times p$	0	$\frac{500}{6^4}$	$\frac{300}{6^4}$	$\frac{60}{6^4}$	$\frac{4}{6^4}$	$\frac{864}{6^4} = \frac{2}{3}$

$\dfrac{864}{6^4} = \dfrac{2}{3}$ であるから、6 の目が出る期待値は $\dfrac{2}{3}$ 回となる。つまり、1 回賭けをするたびに、6 の目が $\dfrac{2}{3}$ 回の割合で出るわけである。ということは、3 回に 2 回の割合で 6 の目が出ることになって、賭ける方に有利であるような気がする。

2 個のさいころを同時に投げることを 24 回続けて、少なくとも 1 回は目の和が 12 になる賭けの期待値も調べてみよう。これは、目の和が 12 になることが 0 回から 24 回まで可能性があるので、Microsoft Excel の力を借りる方がよいだろう。

◇	A	B	C	D	...	Z	AA
1	目の和 12(X)	0	1	2	...	24	計
2	24Cr	1	24	276	...	1	
3	(1/36)^r	(※ **B3**)	→右へコピーする	→	→	→	
4	(35/36)^(24-r)	(※ **B4**)	→右へコピーする	→	→	→	
5	確率 (p)	(※ **B5**)	→右へコピーする	→	→	→	(※ **AA5**)
6	X × p	(※ **B6**)	→右へコピーする	→	→	→	(※ **AA6**)
7							

※ セルの式
(**B3**) **=(1/36)^B1**
(**B4**) **=(35/36)^(24-B1)**
(**B5**) **=B2*B3*B4**
(**AA5**) **=SUM(B5:Z5)**
(**B6**) **=B1*B5**
(**AA6**) **=SUM(B6:Z6)**

シートは、さいころを投げることを 4 回続けて、少なくとも 1 回は 6 の目を出す賭けの表と同じ形式である。ただ、24 回まで計算する必要があるため、表は AA 列まで広がっていることと、計算の都合上、2〜4 行目までを計算補助に利用している。いままで示してきた期待値計算のための表は、ワークシートの 1、5、6 行目だけを見てきたことになる。

2 行目の 24Cr の行は、パスカルの三角形の 24 行目にあたる数を列挙してある。ここは、以前パスカルの三角形を作った章を参考に入力してもらいたい。

B 列は目の和が 12 になるのが 0 回である計算、C 列は目の和が 12 になるのが 1 回である計算、. . .、Z 列は目の和が 12 になるのが 24 回である計算である。すなわち、B5 セルは目の和が 12 になるのが 0 回である確率、C5 セルは目の和が 12 になるのが 1 回である確率、. . .、Z5 セルは目の和が 12 になるのが 24 回である確率を求めているのである。確率の計算が正しいことは、AA5 セルの値が 1 になっていることで確かめられる。そうでなければ、おそらく 2 行目のパスカルの三角形の値を入力間違いしている可能性が高い。

さて、この場合の期待値は AA6 セルに表示されている。それはおそらく 0.666666667 だろう。さっきのさいころ 4 回の例の期待値と同じであることに気づいただろうか。そして、このことは偶然ではないのである。

2 つの例をもう一度要約しておこう。はじめの例は 1 回の試行で 6 の目が出る確率は $\dfrac{1}{6}$ であり、これを独立に 4 回繰り返したのであった。次の例は 1 回の試行で目の和が 12 になる確率は $\dfrac{1}{36}$ であり、これを独立に 24 回繰り返したのであった。いずれも単に

$$\frac{1}{6}\times 4=\frac{2}{3}\quad \text{または}\quad \frac{1}{36}\times 24=\frac{2}{3}$$

を計算したことになっている。

実はこのことは一般に成り立つことで、

確率 p で起こる独立試行 A を n 回繰り返すと、A の起こる期待値は np

であることがいえるのである。証明は別の機会に譲ることになるが、このような確率分布は**2 項分布**に従い、その平均値が X の期待値のことだからである。