逆正弦法則

大槻もも代

2023 年 10 月 31 日

目次

1　はじめに

コイントスなどの確率が 1/2 になる試行において勝ち越しと負け越しの回数の割合について興味深い現象がみられることを知ったので，それについて研究する．

2　ランダムウォーク

2.1　ランダムウォークとは

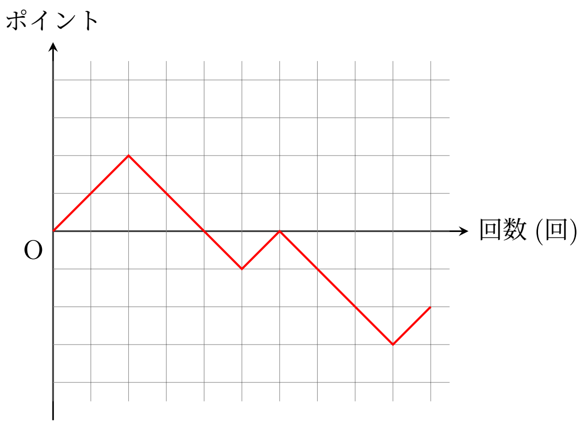
裏表がそれぞれ確率 1/2 ずつで出るコインを N 回投げる．表が出たら $+1$ ポイント，裏が出たら -1 ポイントを得るゲームを繰り返し行う．t 回目に投げた時の結果によって決まる確率変数 X_t を以下のように定義する．$(t = 1, 2, \cdots, N)$

$$X_t = \begin{cases} +1 & (i\text{ 回目に投げたコインが表のとき}) \\ -1 & (i\text{ 回目に投げたコインが裏のとき}) \end{cases}$$

よって，t 回後のポイントの総和 S_t は以下のように表せる．

$$S_t = X_1 + X_2 + \cdots + X_t = \sum_{i=1}^{t} X_i$$

ここでコインを 10 回投げた時の持っているポイントの動きをグラフにすると，一例として以下のようになる．



これは位置がランダムに決定されることから**ランダムウォーク**と呼ばれる．

2.2　ランダムウォークの例

- n 回後に x ポイントを持っている確率を求める．
  n 回のうち，$+1$ ポイントとなった回数 (つまり，表が出た回数) を k 回，-1 ポイントとなった回数 (つまり，裏が出た回数) を $n-k$ 回としたとき，n 回後のポイントの総和は $(+1) \times k + (-1) \times (n-k) = 2k - n$ であり，これが x と等しくなるとき，

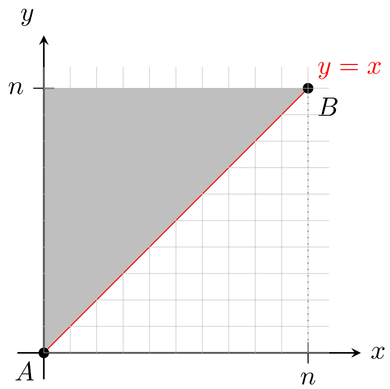
$$x = 2k - n \qquad \therefore k = \frac{n+x}{2}$$

このとき，k, n, x はいずれも自然数であるから k は偶数であり，$n+x$ が奇数のときそのような k は存在せず，確率は 0 となる．
$n+x$ が偶数のときを考える．反復試行の確率より，確率 $p = \dfrac{1}{2}$ で表が出る試行を独立に n 回行ったとき，n 回中 $k = \dfrac{n+x}{2}$ 回表が出る確率は ${}_nC_k \cdot p^k(1-p)^{n-k}$ と表せるので，

$$ {}_nC_k p^k (1-p)^{n-k} = {}_nC_{\frac{n+x}{2}} \left(\frac{1}{2}\right)^{\frac{n+x}{2}} \left(\frac{1}{2}\right)^{n-\frac{n+x}{2}} = \left(\frac{1}{2}\right)^{\frac{n+x}{2}} \left(\frac{1}{2}\right)^{\frac{n-x}{2}}_n C_{\frac{n+x}{2}} = \frac{1}{2^n} \cdot {}_nC_{\frac{n+x}{2}} $$

- $2n$ 回目で初めて原点に戻ってくる確率を求める．

Theorem 2.1. (x, y) 座標平面上で $(0,0)$ から (n,n) まで格子点をたどって進む最短経路のうち，$y \leq x$ の領域を通る経路の数はカタラン数 $c_n = \dfrac{{}_{2n}C_n}{n+1}$ である．

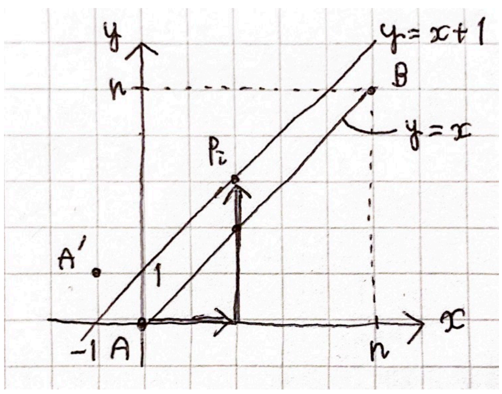


Proof. ($y \leq x$ の領域を通る経路の数)=(すべての最短経路の数 a_n)−(直線 $y = x$ を最低 1 回はまたぐ経路の数 b_n) であるから，余事象で考える．

(1) すべての最短経路の数 a_n は上を n 個，右を n 個並べる並べ方に等しいので

$$ a_n = \frac{(n+n)!}{n!n!} = \frac{2n \times (2n-1) \times \cdots \times n \times \cdots 1}{(n \times (n-1) \times \cdots \times 1) \times (n \times (n-1) \times \cdots \times 1)} = {}_{2n}C_n $$

(2) 直線 $y = x$ を最低 1 回はまたぐ経路の数 b_n は直線 $y = x+1$ を通る経路の数に等しい．
ここで点 $A(0,0)$，点 $B(n,n)$ とし，対角線 AB を点 (i,i) ではじめて踏み，$y = x+1$ を点 $P_i(i, i+1)$ ではじめて踏むとする．



図のように経路 $A \to P_i$ を $y = x+1$ に対して対称移動し，点 A の移動した先の点を A' とすると，$A'(-1,1)$ から $B(n,n)$ までの最短経路の数は b_n に等しく，上を $(n-1)$ 個，右を

$(n+1)$ 個並べる並べ方に等しいので

$$b_n = \frac{((n-1)+(n+1)!)}{(n-1)!(n+1)!} =_{2n} C_{n-1}$$

以上 (1),(2) より，

$$\begin{aligned} c_n = a_n - b_n &= \frac{(2n)!}{n!n!} - \frac{(2n)!}{(n-1)!(n+1)!} \\ &= \frac{(2n)!}{n!(n-1)!}\left(\frac{1}{n} - \frac{1}{n+1}\right) \\ &= \frac{(2n)!}{n!(n-1)!}\frac{1}{n(n+1)} \\ &= \frac{1}{n+1}\frac{(2n)!}{n!n!} = \frac{{}_{2n}C_n}{n+1} \end{aligned}$$

■

これを用いて，$2n$ 回目で初めて原点に戻ってくる確率を求める．
まず，$2n$ 回移動するすべての経路の数は 2^{2n} 通り．
1 回目が表のとき，1 回目の移動後から常に $y \geq 1$ の領域を通り，$(2n-1)$ 回目の移動後に $y = 1$ にいるような経路の数は c_{n-1} 通り．同様に，1 回目が裏のときも対称性から c_{n-1} 通りに等しい．
よって求める確率は

$$\frac{2c_{n-1}}{2^{2n}} = \frac{{}_{2n-2}C_{n-1}}{2^{2n-1}\cdot n}$$

となる．

3 逆正弦法則

3.1 最終的な累計ポイント数

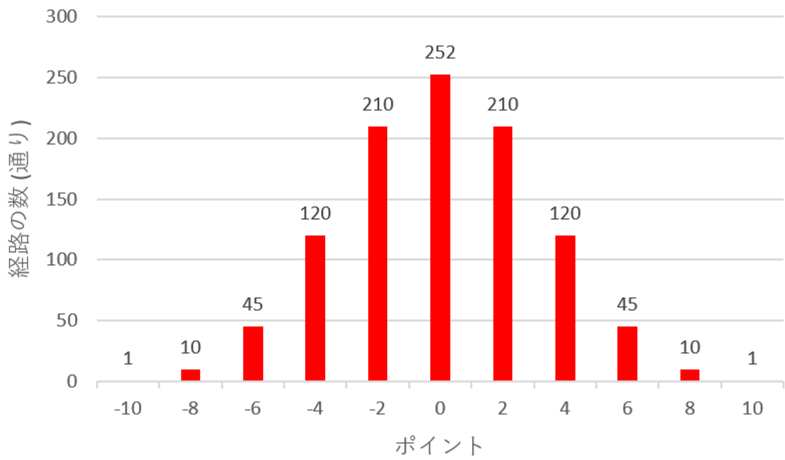
コインを 10 回投げて表が出たら +1 ポイント，裏が出たら −1 ポイントを得るゲームを繰り返し行ったときの 10 回目終了後のポイント総和 S_{10} を考える．

- $S_{10} = 0$ になる経路の数
  $S_{10} = 0$ になるのは +1 ポイント，−1 ポイントになる場合がそれぞれ 5 回ずつのときであるから，上を 5 個，下を 5 個並べる並べ方に等しいので $\dfrac{(5+5)!}{5!5!} =_{10} C_5 = 252$(通り)．
- $S_{10} = 1$ になる経路の数
  (2.2) より，$n + x = 11$ となり，奇数であるから，(2.2) における k は存在せず，0 通りである．
  このことから $S_{10} = 2m+1$ $(m \in \mathbb{Z})$ になる経路の数は 0 であることが分かる．
- $S_{10} = \pm 2$ になる経路の数
  $S_{10} = 2$ になるのは +1 ポイント，−1 ポイントになる場合がそれぞれ 6 回，4 回のときであるから，上を 6 個，下を 4 個並べる並べ方に等しいので $\dfrac{(6+4)!}{6!4!} =_{10} C_4 = 210$(通り)．
  一方，$S_{10} = -2$ になるのは +1 ポイント，−1 ポイントになる場合がそれぞれ 4 回，6 回のときであるから，上を 4 個，下を 6 個並べる並べ方に等しいので $\dfrac{(4+6)!}{4!6!} =_{10} C_4 = 210$(通り)．
  このことから，$S_{10} = l = -l$ となることがわかる．
- $S_{10} = \pm 4$ になる経路の数

$S_{10}=4$ になるのは $+1$ ポイント，-1 ポイントになる場合がそれぞれ 7 回，3 回のときであるから，上を 7 個，下を 3 個並べる並べ方に等しいので $\dfrac{(7+3)!}{7!3!} = {}_{10}C_3 = 120$(通り).
$S_{10}=-4$ になる経路の数も同様.

- $S_{10}=\pm 6$ になる経路の数
  $S_{10}=6$ になるのは $+1$ ポイント，-1 ポイントになる場合がそれぞれ 8 回，2 回のときであるから，上を 8 個，下を 2 個並べる並べ方に等しいので $\dfrac{(8+2)!}{8!2!} = {}_{10}C_2 = 45$(通り).
  $S_{10}=-6$ になる経路の数も同様.
- $S_{10}=\pm 8$ になる経路の数
  $S_{10}=8$ になるのは $+1$ ポイント，-1 ポイントになる場合がそれぞれ 9 回，1 回のときであるから，上を 9 個，下を 1 個並べる並べ方に等しいので $\dfrac{(9+1)!}{9!1!} = {}_{10}C_1 = 10$(通り).
  $S_{10}=-8$ になる経路の数も同様.
- $S_{10}=\pm 10$ になる経路の数
  $S_{10}=10$ になるのは $+1$ ポイント，-1 ポイントになる場合がそれぞれ 10 回，0 回のときであるから，上を 10 個，下を 0 個並べる並べ方に等しいので $\dfrac{(10+0)!}{10!0!} = 1$(通り).
  $S_{10}=-10$ になる経路の数も同様.

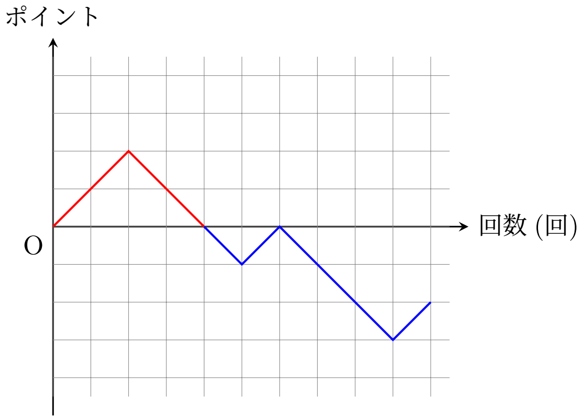
以上のことをまとめると，10 回目終了後のポイント総和 S_{10} のグラフは以下のようになる．



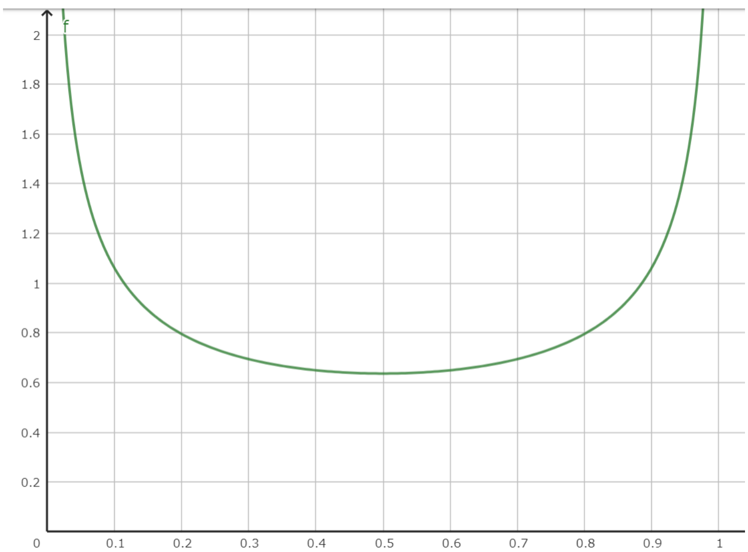
このグラフより，$S_{10}=\pm 10$ になる経路の数が一番少なく，$S_{10}=0$ になる経路の数は一番多くなっていることが分かる．

3.2 勝ち越しと負け越しの二極化

直感的に，勝ち越し (正の領域にいる回数の割合) と負け越し (負の領域にいる回数の割合) はほぼ等しいと考えられるが，これは間違いである．実際に，正の領域にいた回数の割合を計算する．例えば，下のグラフにおいて，赤線が正の領域にいた回数，青線が負の領域にいた回数であるから，$4/10=0.4$ となる．つまり負け越している回数の割合が大きいということが分かる．



コインを投げるゲームをできるだけ多く行った時の裏表のすべての出方について正の領域にいた回数割合 x $(0 \leq x \leq 1)$ を計算し，グラフにすると以下のようになる．



このグラフからずっと勝ち越している状態や負け越している状態の割合が一番大きく，逆に勝ち越しと負け越しの回数が同じとなる割合が一番小さいということが分かる．つまり，勝ち越しか負け越しのどちらか一方に偏ることが多いということが言える．

3.3　逆正弦法則

Theorem 3.1. コインを $2n$ 回投げた時，そのうち $2k$ $(k = 1, 2, \cdots, n)$ 回だけ正の領域にあり，$(2n-2k)$ 回だけ負の領域にある確率は,

$$P_{2n,2k} = u_{2k}\ u_{2n-2k}$$

である．ここで u_{2k} は $2k$ 回後においてポイントの総和が $S_{2k} = 0$ となっている確率であるから，一次元ランダムウォークの例より,

$$u_{2k} = \frac{{}_{2k}C_k}{2^{2k}}$$

である．ただし，$u_0 = 1$.

これを証明するにあたって，以下の式を用いる．

Lemma. (1) $2n$ 回目までに常に正の領域にいる確率は

$$u_2n = \frac{{}_{2k}C_k}{2^{2k}}, u_0 = 1$$

(2) $2n$ 回目に初めて原点に戻ってくる確率は

$$f_{2n} = \frac{{}_{2n-2}C_{n-1}}{2^{2n} \cdot n} = \frac{u_{2n-2}}{2n}, f_0 = 0$$

(3) $n \in \mathbb{N}$ ならば,

$$u_{2n} = \sum_{k=1}^{n} f_{2k} u_{2n-2k}$$

Proof of Theorem3.1.
$2n$ 回目までに常に正の領域にいる確率は式 (1) より u_{2n} であるから,

$$P_{2n,2n} = u_{2n} = u_{2n}u_0 = u_0u_{2n} = P_{2n,0}$$

よって，n についての数学的帰納法を用いて証明する.
まず，$n = 2$ のとき $P_{4,2k} = u_{2k}u_{4-2k}$ が成り立つと仮定して，$n = 3$ のとき $P_{6,2k} = u_{2k}u_{6-2k}$ が成り立つことを示す.
例えば，$P_{6,2}$ は正の領域にいた回数が 2 回，負の領域にいた回数が 4 回であるから，下図のような例が考えられる.

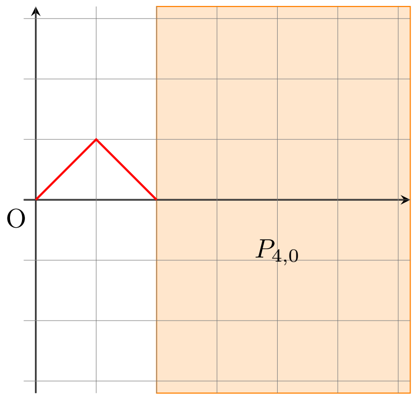


図 1　正の領域を通り，2 回目で原点回帰

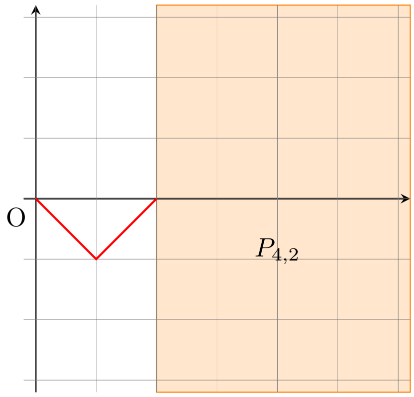


図 2　負の領域を通り，2 回目で原点回帰

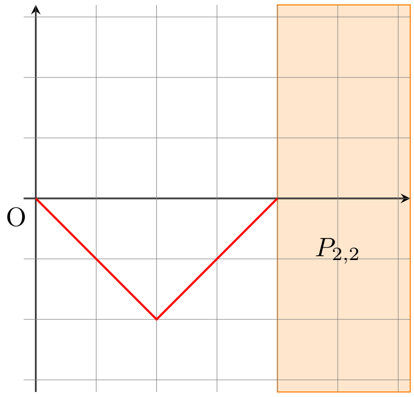


図 3　負の領域を通り，4 回目で原点回帰

よって式 (2),(3)，帰納法の仮定を用いると，

$$
\begin{aligned}
P_{6,2} &= \frac{1}{2}f_2P_{4,0} + \frac{1}{2}f_2P_{4,2} + \frac{1}{2}f_4P_{2,2} \\
&= \frac{1}{2}f_2u_0u_4 + \frac{1}{2}f_2u_2u_2 + \frac{1}{2}f_4u_2u_0 \\
&= \frac{1}{2}f_2u_0u_4 + \frac{1}{2}u_2(f_2u_2 + f_4u_0) \\
&= \frac{1}{2}u_2u_4 + \frac{1}{2}u_2u_4 \\
&= u_2u_4
\end{aligned}
$$

となり，$n = 2$ のときの成立を仮定すると，$n = 3$ のとき成り立つことが示された.

一般に，$N=1,2,\cdots,n-1$ のとき成り立つと仮定して，$N=n$ が成り立つことを示す．

$$\begin{aligned}
P_{2n,2k} &= \left(\frac{1}{2}f_2P_{2k-2,2n-2}+\cdots+\frac{1}{2}f_{2k}P_{0,2n-2k}\right)+\left(\frac{1}{2}f_2P_{2k,2n-2}+\cdots+\frac{1}{2}f_{2n-2k}P_{2k,2k}\right)\\
&= \sum_{r=1}^{k}\frac{1}{2}f_{2r}P_{2k-2r,2n-2r}+\sum_{r=1}^{n-k}\frac{1}{2}f_{2r}P_{2k,2n-2r}\\
&= \sum_{r=1}^{k}\frac{1}{2}f_{2r}u_{2k-2r}u_{2n-2k}+\sum_{r=1}^{n-k}\frac{1}{2}f_{2r}u_{2k}u_{2n-2k-2r}\\
&= \frac{1}{2}u_{2n-2k}\sum_{r=1}^{k}f_{2r}u_{2k-2r}+\frac{1}{2}u_{2k}\sum_{r=1}^{n-k}f_{2r}u_{2n-2k-2r}\\
&= \frac{1}{2}u_{2n-2k}u_{2k}+\frac{1}{2}u_{2k}u_{2n-2k}\\
&= u_{2n-2k}u_{2k}
\end{aligned}$$

よって $N=n$ のときも成り立つことが示された．ゆえに，

$$P_{2n,2k}=u_{2k}\ u_{2n-2k}$$

が成り立つ．

Example 3.1. $n=1,2,3$ のとき，Theorem3.1 を用いて勝ち越しの確率を考える．

- $n=1$ のとき

$$P_{2,2}=u_2\ u_0=\frac{{}_2C_1}{2^2}=\frac{1}{2}=P_{2,0}$$

- $n=2$ のとき
  $P_{4,4}$ は 4 回とも勝ち越している，つまり常に正の領域にいるので．グラフは図 1 のようになり，全部で 6 通りある．対称性より $P_{4,0}$ も 6 通りである．
  また，$P_{4,2}$ は 2 回勝ち越し，2 回負け越しているので 1 回目が表のときのとき，図 2 のようになり，2 通りである．さらに 1 回目が裏のときも対称性より 2 通りあるから，全部で 4 通りである．

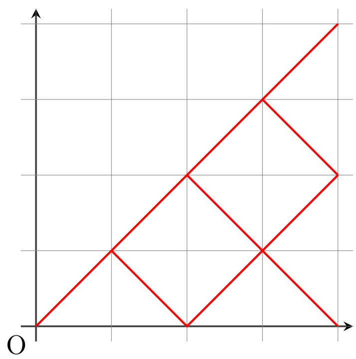


図 4　$P_{4,4}$ となるときの経路

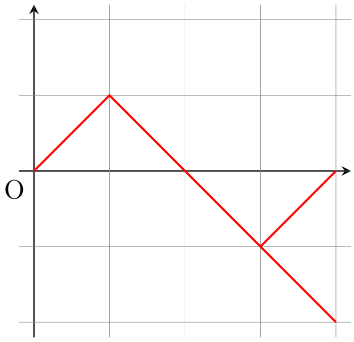


図 5　$P_{4,2}$ となるときの経路 (1 回目が表のとき)

実際に Theorem3.1 を用いて確率を計算すると，一致していることが分かる．

$$
\begin{aligned}
P_{4,0} &= u_0\ u_4 = \frac{{}_4C_2}{2^4} = \frac{6}{16} \\
P_{4,2} &= u_2\ u_2 = \frac{{}_2C_1}{2^2} \cdot \frac{{}_2C_1}{2^2} = \frac{4}{16} \\
P_{4,4} &= u_4\ u_0 = \frac{{}_4C_2}{2^4} = \frac{6}{16}
\end{aligned}
$$

- $n = 3$ のとき

$$
\begin{aligned}
P_{6,0} &= u_0\ u_6 = \frac{{}_6C_3}{2^6} = \frac{20}{2^6} = P_{6,6} \\
P_{6,2} &= u_2\ u_4 = \frac{{}_2C_1}{2^2} \cdot \frac{{}_4C_2}{2^4} \cdot \frac{{}_2C_1}{2^2} = \frac{12}{2^6} \\
P_{6,4} &= u_4\ u_2 = \frac{{}_4C_2}{2^4} \cdot \frac{{}_2C_1}{2^2} = \frac{12}{2^6}
\end{aligned}
$$

Example 3.2. ここで試行回数 n をもう少し大きくとってみる．$n = 10$ とすると

$$
\begin{aligned}
P_{20,0} &= P_{20,20} = 0.176197\cdots \\
P_{20,2} &= P_{20,18} = 0.046367\cdots \\
P_{20,4} &= P_{20,16} = 0.012273\cdots \\
&\vdots \\
P_{20,10} &= 0.000240\cdots
\end{aligned}
$$

となるのでグラフにすると下図のようになる．

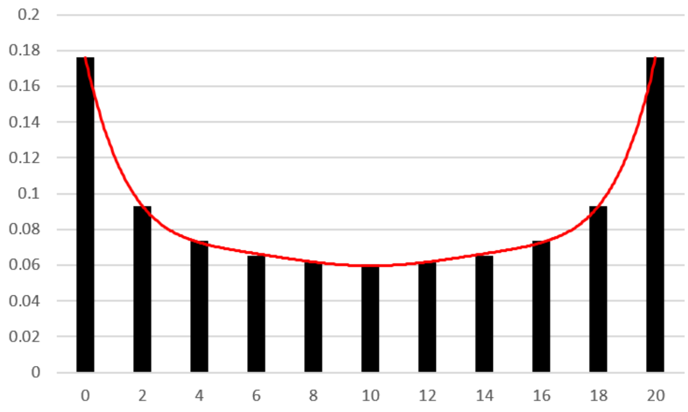


図 6　$P_{20,2k}$ における $2k$ の値

Example3.2 において，試行回数 n を限りなく大きくしたときに，近似的に，

$$f(x) = \frac{1}{\pi\sqrt{x(1-x)}}$$

と表される．これを**確率密度関数**という．

Proof. $2n$ 回後に原点にいる確率は，

$$u_{2n} = \frac{{}_{2n}C_n}{2^{2n}}$$

十分に大きい n に対し，スターリングの公式 $n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n$ を使うと，

$$u_{2n} \sim \frac{1}{2^{2n}} \cdot \frac{\sqrt{4\pi n}\left(\frac{2n}{e}\right)^{2n}}{2\pi n\left(\frac{n}{e}\right)^{2n}} = \frac{1}{\sqrt{\pi n}}$$

よって $P_{2n,2k} = u_{2k}u_{2n-2k}$ に代入すると，

$$P_{2n,2k} = u_{2k}u_{2n-2k} \sim \frac{1}{\sqrt{\pi n}} \cdot \frac{1}{\sqrt{\pi(n-k)}} = \frac{1}{n\pi} \cdot \frac{1}{\sqrt{\frac{k}{n}\left(1-\frac{k}{n}\right)}}$$

を得る．ここで，$x = \dfrac{k}{n}$ とおき，$a \leq x \leq b$ となる確率は，区分求積法を用いて，

$$\sum_{a \leq \frac{k}{n} \leq b} \sim \frac{1}{\pi} \sum_{a \leq \frac{k}{n} \leq b} \frac{1}{n\sqrt{\frac{k}{n}\left(1-\frac{k}{n}\right)}} = \frac{1}{\pi}\int_a^b \frac{1}{\sqrt{x(1-x)}}dx.$$

よって確率密度関数は

$$f(x) = \frac{1}{\pi\sqrt{x(1-x)}}$$

となる． ■

3.4　確率密度関数と累積分布関数

様々な試行において，確率分布は離散的または連続的に分けられる．例えば，さいころを投げた時の 1 が出る確率は $P(X=1) = \dfrac{1}{6}$ とかけるので確率分布は離散的である．一方，今回のような試行においい

て試行回数 n を限りなく大きくすることで確率分布は連続的なものになる．よって正の領域にいた回数の割合 $X = a$ $(0 \leq a \leq 1)$ と特定の値をとる確率はほぼ 0 である．ここで，$a \leq X \leq b$ となる確率は考えることができて，**確率密度関数** $f(x)$ を積分することで，その面積が求める確率となる．また，確率変数 X が $a \leq X \leq b$ となる確率を表す関数を**累積分布関数**という．

Definition 3.1. 確率密度関数
連続な確率変数 X において，$a \leq X \leq b$ となる確率 $P(a \leq X \leq b)$ が

$$P(a \leq X \leq b) = \int_a^b f(x)\ dx$$

と表せるときに被積分関数 $f(x)$ を**確率密度関数**という．

Definition 3.2. 累積分布関数
確率変数 X に対して，

$$F(x) = P(X \leq x) = P(x \in (-\infty, x])$$

を**累積分布関数**という．

Example 3.3. 確率変数が離散型であるとき
さいころを投げて出る目を確率変数 X とするとき，累積分布関数を計算すると以下のようになる．

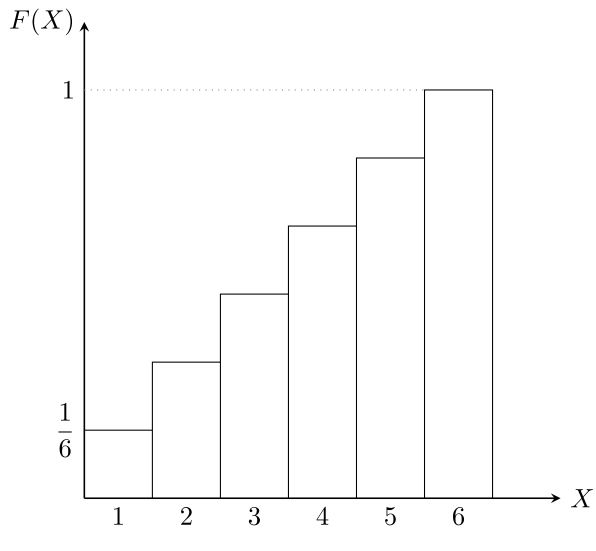
- 出る目が 3 以下になる確率

$$P(X \leq 3) = F(3) = \sum_{X \leq 3} P(x) = P(1) + P(2) + P(3) = \frac{1}{2}$$

- 出る目が 2 以上 5 以下になる確率
  言い換えると，(出る目が 5 以下になる確率)-(出る目が 1 以下になる確率) であるから，

$$P(2 \leq X \leq 5) = F(5) - P(1) = \sum_{X \leq 5} - \sum_{X \leq 1} = \frac{5}{6} - \frac{1}{6} = \frac{2}{3}$$

よって累積分布関数のグラフは下のようになる．



Example 3.4. 確率変数が連続型であるとき

今回取り扱っているコイントスの例を実際に用いて，$0 \leq x \leq \alpha$ になる確率を累積分布関数で計算する．確率密度関数は

$$f(x) = \frac{1}{\pi\sqrt{x(1-x)}}$$

であったから，求める確率は $f(x)$ の $0 \leq x \leq \alpha$ の面積である．

$$\begin{aligned}
P(0 \leq x \leq \alpha) &= \int_0^{\alpha} \frac{1}{\pi\sqrt{x(1-x)}} dx \\
&= \frac{1}{\pi} \int_0^{\alpha} \frac{1}{t\sqrt{1-t^2}} \cdot 2tdt \\
&= \frac{2}{\pi} \int_0^{\alpha} \frac{1}{\sqrt{1-t^2}} dt \\
&= \frac{2}{\pi} \left[\sin^{-1} t\right]^{\sqrt{\alpha}}{}_0 \\
&= \frac{2}{\pi} \sin^{-1} \sqrt{\alpha}.
\end{aligned}$$

(置換積分法と $(\sin^{-1} t)' = \dfrac{1}{\sqrt{1-t^2}}$ を用いた．)

ここで結果を見ると $\sin^{-1}$ の形が出てきている．このことから，勝ち越しか負け越しのどちらかに二極化する現象は逆正弦法則と呼ばれる．

4　まとめ

確率が同じくらいのゲームを何度も繰り返すと，勝ち越しか負け越しのどちらかに二極化し，その回数が同じになるということは少ない．これを逆正弦法則と呼ぶ．また，勝ち越している回数の割合は試行回数 n が大きいとき，確率密度関数

$$f(x) = \frac{1}{\pi\sqrt{x(1-x)}}$$

で表すことができる．

5　今後の課題

今回の研究を通して，賭けと確率は密接な関係にあると知った．今回は賭けの必勝法や攻略法について研究することができなかったので，これから研究していきたい．

参考文献

[1] 逆正弦法則の一般化に向けての一考察，鈴木将史，`https://aue.repo.nii.ac.jp/record/2888/files/epsilon39131140.pdf`.

[2] ランダムウォークの基礎（一次元，高校範囲），高校数学の美しい物語，`https://manabitimes.jp/math/1093`.

[3] 小針晛宏，確率・統計入門，岩波書店，1973.