コイントスから広がる確率論の世界

東北大学大学院情報科学研究科
尾　畑　伸　明

１．はじめに

おみくじや宝くじ、トランプゲーム、為替や株価、明日の天気など「偶然」に左右される事柄を確実に予測することはできません。おみくじや宝くじなら何等賞かとか、為替や株価の変動ならいくら上がるか下がるかのように、起こりうる結果はあらかじめ知ってはいるのですが、何通りもあるので、そのどれが起こるか分からないのです。しかし、どの結果がどの程度の割合で起こるかを知ることができれば、個人レベルで得をするだけでなく、危険を避けたりすることで日々の生活も豊かになるに違いありません。物事の起こりやすさ、起こりにくさを数値化したものが「確率」です。インターネットをちょっと眺めるだけでも、確率の愛好者たちが次々に現れます。

- 宝くじの当選確率
- 降水確率
- 試験に合格する確率
- 宮城県沖地震の発生確率（仙台市消防局 17.1.31）
- 横浜市民が犯罪にあう確率（横浜市統計ポータルサイト）
- 地震による原発事故発生の「確率」について
- 浜岡原子力発電所のそれぞれの原子炉施設への航空機落下確率（中部電力 14.9.30）
- 小惑星 2004MN4 が 2029 年 4 月 13 日に地球に衝突する確率は 1/300（NASA ジェット推進研究所 2004.12.24）
- 成功確率 50％のプロジェクトに GO を出すべきか（エイチアイエス社長インタビュー）
- 阪神星野仙一 SD(58)が今季の優勝確率を「100％以上」と断言 （2005.7.11）
- 恋に落ちる確率（映画です、原題 Reconstruction） などなど

確率を知ることは、偶然の結果をチャンスに変える玉手箱を手に入れたようなものです。

でも、確率を考えるための前提があります。結果が不確かだからと言って、とにかく確率を持ちこめば何とかなるというものではありません。確率が役に立つのは、その事柄の実験や観察を繰り返し行うことが可能な時です。たとえば、コイントスは何回でも繰り返し実験ができますから、確率を考える対象になります。しかし、明日熱を出して学校を休む確率、交差点で友人に会う確率、恋に落ちる確率とかは慎重にならざるをえません。

２．組合せ確率論の問題

問題１．3 個のサイコロを同時に投げて 1 の目が出れば勝ち。このゲームを 3 回行うときに 3 連勝する確率は？

世は大航海時代。悪天候で海に出られない船乗りたちのすることは。。。３次方程式の解の公式を発見したことで有名なカルダーノ（1501-76）はギャンブラー。生きているうちには出版されなかった（秘密だったから？）「偶然のゲームに関する本」（1663 年出版）で議論されています。本気で研究したことでしょう。

問題 2．3 つのサイコロを投げるとき、その和が 9 になる場合と 10 になる場合の数は等しいので、どちらに賭けても同等であると思っていたが、実際には 10 になる方が少し多く感じられるのは、どうした わけであろうか？

中世の貴族はばくちがお好き。相当やりこまなければ、この違いには気付かないと思われます。ガリレオ（1564-1642）は、もちろん正しく答えています。

問題 3．公平なコイン投げのゲームで先に 4 勝した方が賞金 1 万円を受け取る。やむを得ない事

情で、2 勝 1 敗の状況でゲームを中断せざるを得なくなった。賞金を公平に分配するにはどうすればよいだろうか?

「人は考える葦である」パスカル（1623-62）は貴族の質問をきっかけに、当代随一の数学者フェルマ（1601-65）と文通を始めました。相当数の往復書簡のなかで様々な確率の問題が論じられ、「組合せ確率論」はここに始まったとされています。

問題 4. 100 発 100 中の大砲 1 門は、100 発 1 中の大砲 100 門に勝る?

日本海海戦を勝利に導いた東郷平八郎（1847-1934）の訓示として有名な一言です。興味があったら「坂の上の雲」（司馬遼太郎）を読んでみてください。１つ１つを大事にすることは尊いけれど、100 発 1 中の大砲 100 門に囲まれたくはないなあ。

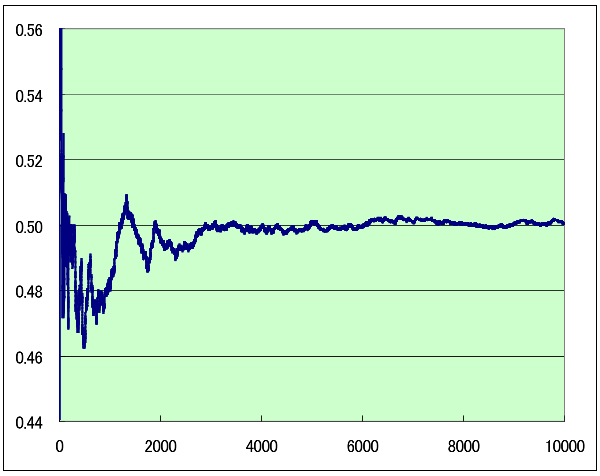
暇があったら考えてみてね（答えは文末）。

３. コイントスの法則を探ろう

コイントスの結果は表（Heads）と裏（Tails）の 2 通りです。したがって、コイントスは最も単純な偶然現象であり、最も基本的な「確率モデル」になります。公平なコインでは、$P(\mathrm{H}) = P(\mathrm{T}) = 1/2$ と表わされます。実際のコイントスの結果をお見せしましょう。

01011 11001 11101 00001 11101 11011 10010 00101 00000 10100
00011 10101 11001 00111 11000 00010 10011 00001 11101 01011

（１）表の出る相対頻度（大数の法則）



k 回目のコイントスの値 X_k を、表が出たら 1、裏が出たら 0 と定めると便利です。

$$S_n = \sum_{k=1}^{n} X_k : 1 \sim n \text{ 回目までに表の出た回数}$$

$$f_n = \frac{S_n}{n} \quad \text{：その相対頻度}$$

となるからです。右図は、10000 回のコイントスの結果の一例です。横軸に n をとり、縦軸に相対頻度をとってあります。

$$f_n \to \frac{1}{2} \quad as \quad n \to \infty$$

が読み取れるでしょう。これが「大数の法則」です（数学的に証明できます）。この 1/2 というのがコイントスで表の出る確率（もちろん、裏の出る確率も同じ）になるのです。

（２）ギャンブル

本質的に違いはないのですが、（１）のコイントスの値を変更します。k 回目のコイントスの値 X_k を 1（表が出たとき）、-1（裏が出たとき）と定めます。これは、勝てば 1（円）を獲得し、負ければ 1（円）を失うコイントスによる 2 人ゲームと考えられます。このとき、

$$S_n = \sum_{k=1}^{n} X_k : \text{ゲームが } n \text{ 回終わった時点での獲得金の総額}$$

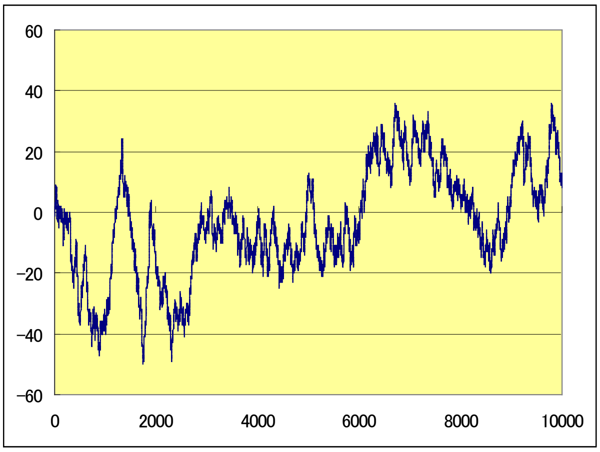
なので、より実感が沸くかもしれません。大数の法則によって、

$$\frac{S_n}{n} \to 0 \quad as \quad n \to \infty$$

となりますが、これは、ゲームを多数回続ければ、1 回あたりの平均損得はほぼ 0 であるということです。公平なゼロサムゲームの証と言えます。

（３）再帰性

もっと興味があるのは、平均値ではなく 1 回ごとの損得かも知れません。右のグラフは 10000 回のコイントスによる獲得金額の推移を表しています。興味深いことが読み取れます。それは、出発点の損得 ±0（原点と呼びましょう）からどんなにはずれても必ず将来のある時点で原点に戻ることです。



$$P\text{（ある時刻 } n \text{ で } S_n = 0\text{）} = 1$$

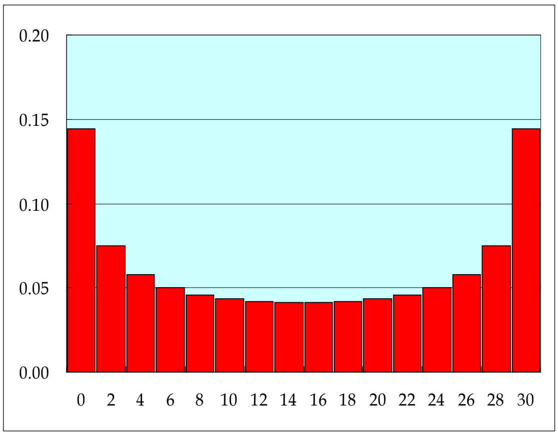
が証明されます。しかしながら、$S_n = 0$ となるのに要する回数 n の平均値（平均再帰時間）は無限大です！

（４）幸福な時間・不幸な時間

もう一度、グラフを観察しましょう。ずいぶん長時間にわたって獲得賞金がプラスまたはマイナスの側に偏りすぎているとは思いませんか？ たとえば、6000 回から 8000 回にかけての部分を見てください。公平なコインは表裏が均等に出るはずですね。2000 回も振っていますから、表裏はほぼ半々の割合で出ているはずです（大数の法則)。したがって、獲得賞金がプラスになっている時間（幸福な時間）とマイナスになっている時間（不幸な時間）も半々になるのが「公平さ」の証であるように思うのも自然です。しかし、この直感は間違いなのです。

時間区間 $[k,k+1]$ において $S_k \geqq 0$ かつ $S_{k+1} \geqq 0$ のとき幸福な時間, $S_k \leqq 0$ かつ $S_{k+1} \leqq 0$ のとき不幸な時間と呼ぶことにします。$2n$ 回のコイントスのうち幸福の時間があわせて $2k$ となる(よって不幸の時間は $2n-2k$ です）確率 $P_{2k/2n}$ は、多少の技巧で計算できて、

$$P_{2k/2n} = \binom{2k}{k}\binom{2n-2k}{n-k}\frac{1}{2^{2n}} \qquad (k = 0,1,2,\cdots,n)$$



右図は、30 回のコイントスの場合（n =15）です。横軸が（30 回のコイントスのうちの）幸福な時間数、縦軸がその確率です。なお、n が大きくなるとき、ヒストグラムは、逆正弦則と呼ばれる曲線

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

に近づくことが知られています。

（１）－（５）を「ギャンブルの数理」としてまとめてみましょう。公平なゼロサムゲームにでは、「儲かる人は長時間にわたり儲け続け、逆に、借金地獄からは短時間では這い上がれない」のです。しかしながら、「ゲームを続行できるなら、いつかは必ず損得±０に戻ってやり直しがきく」わけです。でも、実際には、資金力という壁が立ちはだかります。

問題 5. 資産 10 円をもつ大富豪 A と資産 1 円をもつ大貧民 B がコイン投げによる 1 円やりとりのゲーム（上で説明したもの）で対戦します。一方が破産するまでゲームを続ける時、それぞれの勝つ確率はどうなるか。（資金力のある相手には勝てません。）

4. 研究の最前線へ

高等学校で習う「確率論」は、場合の数の比によって確率を与える「組合せ確率論」です。きちんと場合分けして数え上げるという、隙のない論理展開を学ぶ上でも好都合な題材です。パスカルやフェルマの時代の確率論は、まさに「組合せ確率論」でした。その後、ニュートン（1642-1727）による微積分学が融合し、ベルヌイ（1654-1705）、ラプラス（1749-1827）、ガウス（1777-1855）、ポアソン（1781-1840）らによる「解析的確率論」の全盛期を迎えます。ガウスの誤差論（中心極限定理）、ポアソンの小数の法則（「確率の計算の一般規則に基づく刑事および民事裁判の確率に関する研究（1837）」で発表）などが生まれ、統計的推論の基礎ができてきました。

「現代確率論」はコルモゴロフ（1903-87）の有名な著書「確率論の基礎概念（1933）」にさかのぼります。特に、ランダムウォークやブラウン運動に代表される「偶然現象の時間変化」の研究（確率過程論）が始まりました。1950 年頃になると、確率過程に微積分が高度に融合した「確率解析」が起こりました。その第一人者は伊藤清（1915−2008）です。今日では「伊藤解析（Ito Calculus）」と呼ばれ、自然科学から社会科学への広い範囲に多大な影響を与えています。ところで、伊藤清はウォールストリートで最も有名な日本人としても知られています。金融派生商品（デリバティブ）の価格は、未来の株価を基にして決めるので、株価予想の合理的な理論なしには不可能です。ブラックとショールズは 1973 年に（ヨーロッパ型コール）オプションの理論価格を与える画期的な公式を発表しました。これは「ブラック・ショールズ式」と呼ばれ、銀行や証券の実務において必ず参照される重要な公式です。この功績によって、ショールズとマートンは 1997 年ノーベル経済学賞を受賞（ブラックは 1995 年に死去）しましたが、伊藤解析が本質的な役割を演じていたので、伊藤こそノーベル賞にふさわしいという声が多くあがりました。

現代はインターネットに代表されるネット社会であるといわれます。複雑なネットワークは、どのようにして形成され、どのように壊れてゆくのでしょうか。あるいは、どのような性質が隠されているのでしょうか。1967 年にミルグラム（社会学者）は、全米から無作為に選んだ 300 人に対し、彼らが直接知らない受取人（おおまかな居場所や職業は知らされている）に、友人を仲介にして手紙を転送してもらうという実験をしました。結果は、平均 6 回の転送で最終受取人に手紙が届くというもので、大きな反響を呼び、映画にもなりました。全米の人口を考えると驚くほど早いように思えます。現実世界の様々なネットワークに関連して、その発達の仕方やネットワーク上の流れをコイントスによって模倣することで、確率論の応用がどんどん広がっています。このようなネットワーク研究が進めば、2003 年の SARS があれほど急速に世界中に広まった理由が解明され、伝染病の蔓延の阻止に貢献できるでしょう。また、堅牢な通信ネットワークの構築や社会の合意形成の仕組みの解明なども期待できます。もちろん、ネットワークへの応用は確率論の数ある応用の一つに過ぎません。皆さんも確率の世界を楽しんでください。

初出：神奈川県立神奈川総合高等学校講演会（2008.12.10）改訂（2011.1.19）

問題の答え (1) 約 7.5% (2) 和が 9 になる確率は 25/216, 和が 10 になる確率は 27/216 (3) 11:5 に配分する (4) 1 対 1 の対戦で一方が全滅するまで続けるのなら、概ね 0.634:0.366 で 100 門の大砲を所有しているほうが有利 (5) 10:1 で大富豪の勝ち。一般に資金 a,b をもつ A,B の 2 人が対戦する時、それぞれの勝つ確率は $b/(a+b)$ と $a/(a+b)$ となる。