



確率論と社会

ギャンブルからネットワークへ

～ コイントスで見えるもの ～

東北大学大学院情報科学研究科
尾畑 伸 明

1. はじめに

ちょっと回りを見渡せばチャンス（偶然）の愛好者が大勢見当たります。おなじみのものから、かなり怪しく意味不明なものまでいろいろです。

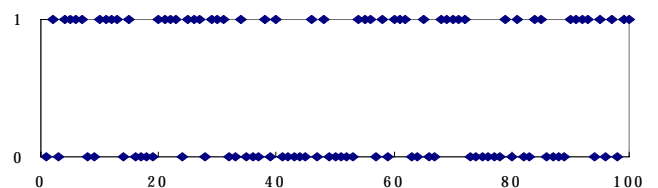
- ・ 降水確率
- ・ 宝くじの当選確率
- ・ おもちゃのカンヅメ当選確率を調査
- ・ お菓子のおまけが全て揃う確率
- ・ 横浜市民が犯罪にあう確率（横浜市統計ポータルサイト）
- ・ 宮城県沖地震の発生確率（仙台市消防局 17.1.31）
- ・ 地震による原発事故発生の「確率」について
- ・ 浜岡原子力発電所のそれぞれの原子炉施設への航空機落下確率（中部電力 14.9.30）
- ・ 小惑星 2004MN4 が 2029 年 4 月 13 日に地球に衝突する確率は 1/300（NASA ジェット推進研究所 2004.12.24）
- ・ 財務データを評価分析し、企業の倒産確率を算出
- ・ あなたの 10 年後の死亡確率
- ・ 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（Reconstruction） などなど

今回は、コイントスをキーワードとして確率の楽しみをお話ししながら、最近大いに注目されている「数理ファイナンス」と「スモール・ワールド・ネットワーク」を垣間見たいと思います。

2. コイントス

最も単純な偶然現象、そして最も基本的な確率モデルです。コイントスの結果は表（Heads）または裏（Tails）の 2 通りです。とりあえず、公正なコインだけ考えることにします。数式なら、 $P(H) = P(T) = 1/2$ と簡潔になります。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
```



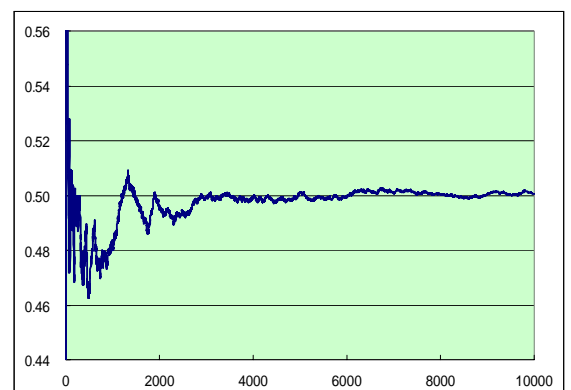
(1) 表の出る相対頻度を調べよう

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。

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

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

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。5000 回ころから 0.50 ± 0.002 くらいの揺らぎに見えるので、

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

が読み取れるでしょう。これは、今の一例についてだけではなく、一般にいつでも成り立つ事実（数学的証明があるのです）で、「大数の法則」と呼ばれています。

(2) ギャンブル

本質的に違いはないのですが、先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、

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

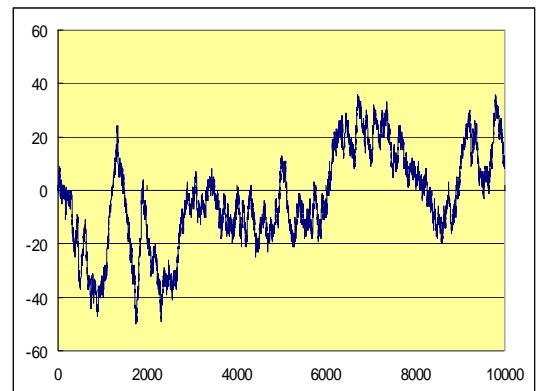
なので、より実感が沸くかもしれません。大数の法則は、

$$\frac{S_n}{n} \rightarrow 0 \quad \text{as } n \rightarrow \infty$$

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。公平なゼロサムゲームの証と言えます。

(3) 原点回帰

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。



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

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（平均回帰時間）は無量大です！

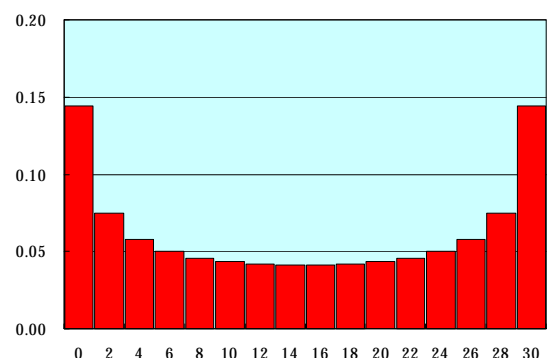
(4) 幸福な時間・不幸な時間

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。公正なコインは表裏が均等に出るはずで、しかも 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。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「公正さ」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 $S_k \geq 0$ かつ $S_{k+1} < 0$ のとき幸福な時間、 $S_k < 0$ かつ $S_{k+1} \geq 0$ 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、

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

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、ヒストグラムは、逆正弦則と呼ばれる曲線に近づくことが知られています。

教え：公平なゼロサムゲームにおいては、「儲かる人は長時間にわたり儲け続け、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。総じて、資金力のある相手には勝てません。

3. 確率論小史

貴族から賭博に関する質問を受けた若い Pascal (1623-62) が当代随一の数学者 Fermat (1601-65) と始めた文通から「組合せ論的確率論」が誕生したことに異論はないでしょう。しかしながら、何事にも先駆者はあるもので、Cardano (1501-76)、Galileo (1564-1642) の名を忘れると、イタリアの友人からは一言。Cardano は Pascal と Fermat が文通を始める 100 年も前に、「偶然のゲームに関する本」を著したものの、出版されたのは 1663 年だったのです。

その後、Newton (1642-1727) による微積分学が融合し、Laplace (1749-1827)、Gauss (1777-1855)、Poisson (1781-1840) らによる「解析的確率論」の全盛期を迎えます。Gauss の誤差論、Poisson 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、統計的推論に数理的基礎が与えられました。

20 世紀初頭には、Kolmogorov (1903-87) の「確率論の基礎概念 (1933)」に象徴される公理主義が起り、「現代確率論」となって今日に至っています。特に、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)において、Wiener (1894-1964) と Lévy (1886-1971) の功績が輝いています。ただし、ブラウン運動の数学的理論(問題意識は物理学ですが)は 1905 年(奇跡の年)に Einstein が発表しました。1950 年頃になると、確率過程論に確率微分方程式論が付け加わって著しい発展を遂げました。その第一人者は伊藤清(1915-)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。なお、伊藤清はウォールストリートで最も有名な日本人にもあげられています。

4. 金融工学あるいは数理ファイナンス

金融派生商品(デリバティブ)開発において、合理的な理論なしにリスク管理は不可能です。Black と Scholes 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。この公式は「ブラック・ショールズ式」と呼ばれ、今日では、実務において必ず参照されるようになってきました。1977 年に Scholes と Merton が「デリバティブ価格決定の新手法」によってノーベル経済学賞を受賞(Black 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。

ブラック・ショールズ式は、実は、コイントスからも導かれます。行使期限 T のコールオプションの価格を決めるために、現在の株価 $S(0)$ から将来の株価 $S(T)$ を予測します。時間区間 $[0, T]$ を長さ Δt の小区間に等分して、第 k 時間小区間における株価の変化をコイントスでモデル化します。つまり、第 k 時間小区間の始値を $S(k)$ として、終値 $S(k+1)$ は、確率 p で $uS(k)$ に値上がりし、確率 $q = 1-p$ で $dS(k)$ に値下がりするものとします。ただし、 $u = 1 + \mu\Delta t + o(\Delta t)^{1/2}$ 、 $d = 1 + \mu\Delta t - \sigma(\Delta t)^{1/2}$ です。 p, q はリスク中立という考えに基づいて、

$$p = \frac{1}{2} + \frac{\rho - \mu}{2\sigma} \sqrt{\Delta t} \quad q = \frac{1}{2} - \frac{\rho - \mu}{2\sigma} \sqrt{\Delta t}$$

のように求まります。小区間における変化を $t=0$ から $t=T$ まで総合し、時間幅 $\Delta t \rightarrow 0$ として、連続量である株価の予想値

$$S(T) = S(0) \exp \left\{ \sigma \sqrt{T} Z + \left(\rho - \frac{\sigma^2}{2} \right) T \right\}$$

が得られます。ここで、 ρ は安全利子率（金利）、 σ は株価の変動率であり、市場から知ることができます。また、 Z は標準正規分布と呼ばれ、確率論で基本的な分布です。上で求めた $S(T)$ を用いて、ある種の平均値を計算すると、ブラック・ショールズ式に到達します。

$$C = S(0)\Phi(d_1) - e^{-\rho T} K\Phi(d_2)$$

$$d_1 = \frac{1}{\sigma\sqrt{T}} \left\{ \log \frac{S(0)}{K} + \left(\rho + \frac{\sigma^2}{2} \right) T \right\}, \quad d_2 = \frac{1}{\sigma\sqrt{T}} \left\{ \log \frac{S(0)}{K} + \left(\rho - \frac{\sigma^2}{2} \right) T \right\}$$

K は権利行使価格、 Φ は標準正分布の分布関数です。上の $S(T)$ は、「株価の時間変化が幾何ブラウン運動に従う」ことを意味しますが、その背後にはコイントスが潜っていたのです。

5. スモールワールド・ネットワーク

1967年に Milgram（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お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。

ハンガリーの数学者 Erdős（1913-96）は、自分の家を持たず、世界中の数学者の家に泊めてもらいながら数学を議論し、生涯に 1500 編もの論文を発表しました。Erdős と共著の論文を書いたことのある人のエルデユス数を 1 とし、その誰かと共著論文がある人のエルデユス数は 2、といった具合に Erdős との「親密度」を計る遊びがあります。多くの数学者のエルデユス数は 4~5 のようで、私自身大変驚いたものです。似たような遊びでは「Bacon 数」がとても有名です。

友人関係はグラフ（頂点と辺からなる図形）で表すのが便利です。現実世界に観察される興味深いグラフは一般に巨大ですから、まずは、グラフの細部の構造より、全体の特徴を大まかにつかむことが重要です。その指標として基本的なものは次の 3 つです。全頂点数を n とします。

(1) 度数分布

$$p(k) = \frac{\text{度数 } k \text{ の頂点数}}{n}, \quad k = 0, 1, 2, \dots \quad (\text{枝が } k \text{ 本出ている頂点の次数が } k)$$

(2) 平均頂点間距離

$$L = \frac{2}{n(n-1)} \sum \text{頂点对間の距離}, \quad \text{なお、頂点对の総数} = \frac{1}{2}n(n-1)$$

(3) クラスタ係数

$$C = \frac{1}{n} \sum C(i), \quad C(i) = \frac{\text{頂点 } i \text{ に隣接している頂点の対で } i \text{ とともに 3 角形をなすものの個数}}{\text{頂点 } i \text{ に隣接している頂点の対の総数}}$$

現実世界のネットワークでは、

(1) ベキ則 ($p(k) \propto k^{-\gamma}$) など (2) $L = O(\log n)$ 小さい (3) $C = O(1)$ 大きい

といった特徴が見られます。このような性質をもったネットワークが形成されるメカニズムについては、新しい問題でもありまだ解明されていません。ランダムグラフ（おもとは Erdős）、スモールワールド・モデル（Watts & Strogatz）、優先的選択成長モデル（Barabasi & Albert）などの興味深い研究は、いずれもコイントスによって現実世界を模倣するアイデアが背後にあります。このような新しい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に期待が寄せられています。

ご意見などはメールでお寄せください。よろしければ、ウェブページにもお越しください。

E-mail は obata@math.is.tohoku.ac.jp Web Page は <http://www.math.is.tohoku.ac.jp/~obata>