数学クイズ
東北大学オープンキャンパス 2017

担当:田中亮吉 [*]

1　すべてが P になる.

では早速, 問題です.

問題 1.1 (★). 例えば, あなたの所持金は 1000 円だとします. 当たれば 1000 円もらえる宝くじが 500 円だとします. 宝くじの当たる確率は 50 %で, 外れれば, 1 円ももらえません. あなたはこの宝くじを買うべきでしょうか? 買わないべきでしょうか?

「当たっても 1000 円しかもらえないのか」と思われるかもしれませんが, では次の問題はどうでしょうか?

問題 1.2 (★). 当たれば 1 億円もらえる宝くじが 500 円だとします. ただしこの場合, 宝くじの当たる確率は 1 億分の 1 とします. あなたはこの宝くじを買うでしょうか?

確率論は, このように起こる可能性が割合で与えられた現象を扱います. 元々は, 17 世紀に上記のような賭け事に関わる問題が確率論の起源と言われていますが, 現在では, より複雑な保険の掛け金に関わる問題や, 統計学, また物理現象や生命科学における現象を正確に理解するのになくてはならない分野になっています. 数学における確率論では, 自然現象から単純なモデルを考えて, その「数学的構造」を理解することを考えます. ここでは, そのような例として, 確率論におけるいくつかの「壺」たちを紹介します. ちなみに, P は確率 (probability) のことです. なお, 星★の数は問題の難易度を表し, 星★の数が多いほど内容が高級になるのは, レストランの質と同じです.

1.1　ポリヤの壺

今 1 つの壺に, 黒の球が 1 つ, 白の球が 1 つあるとします. 壺の中に手を入れ, どの球も同等の確率で取り出します. もしそれが黒の球ならば, その球を壺に戻し, もう 1 つ黒の球を追加します. またもしそれが白の球ならば, その球を壺に戻し, 白の球を 1 つ追加します. 今度は 3 つの球が壺にあるわけですが, さらに壺に手を入れ, 球を 1 つ同等の確率で取り出します. 先ほどと同じように, 取り出した球を壺に戻すとき

[*]東北大学大学院理学研究科 rtanaka@m.tohoku.ac.jp

に, それと同じ色の球を追加することにします. これを繰り返していくと, 壺の中の球の数は増えていくわけですが, 黒の球, 白の球の割合はどのように変わっていくでしょうか?

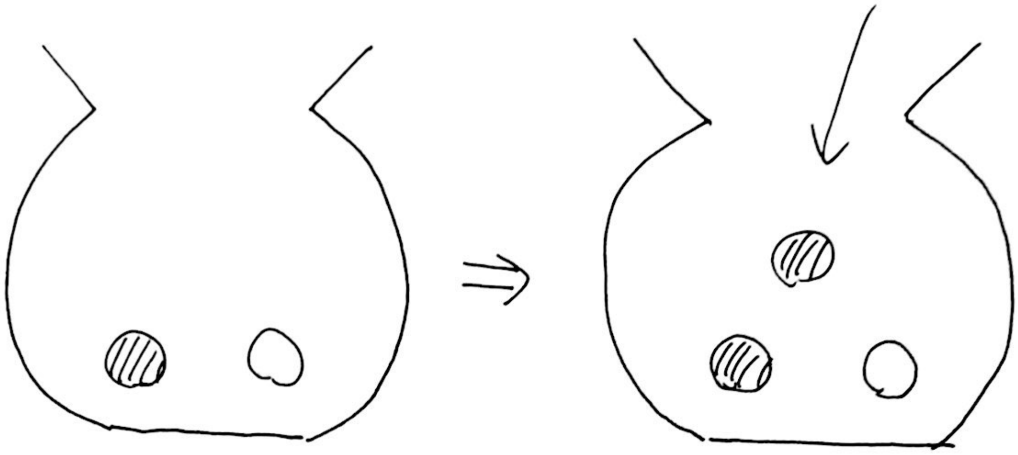

Figure 1: ポリヤの壺

問題 1.3 (★). 2 回球を追加した後, 黒色の球が 1 つである確率はいくつでしょうか? また 2 回球を追加した後, 黒色の球が 2 つである確率はいくつでしょうか?

問題 1.4 (★★). 3 回球を追加した後, 黒色の球が 1 つである確率はいくつでしょうか? また 3 回球を追加した後, 黒色の球が 2 つである確率はいくつでしょうか?

問題 1.5 (★★★). 100 回球を追加した後, 黒の球が 46 個である確率はいくつでしょうか？

1.2　エーレンフェストの壺

今 2 つの壺 A と壺 B があり, 片方の壺 A に球が 10 個あり, もう片方の壺 B には球が 0 個あるとします (つまりは球がないということ!). どの球も入っている壺に関係なく同等の確率で選び出し, その球が入っていなかった方の壺に入れ替えるとします. つまり壺 A から球を取り出した場合, 壺 B に移します.

問題 1.6 (★). 2 回球を移し替えたとき, 壺 B に球が 0 個である確率はいくつでしょうか?

問題 1.7 (★★). 3 回球を移し替えたとき, 壺 B に球が 0 個である確率はいくつでしょうか?

問題 1.8 (★★★). 5 回球を移し替えたとき, 壺 A と壺 B に球が 5 個ずつある確率は $\frac{1}{2}$ より大きいでしょうか?

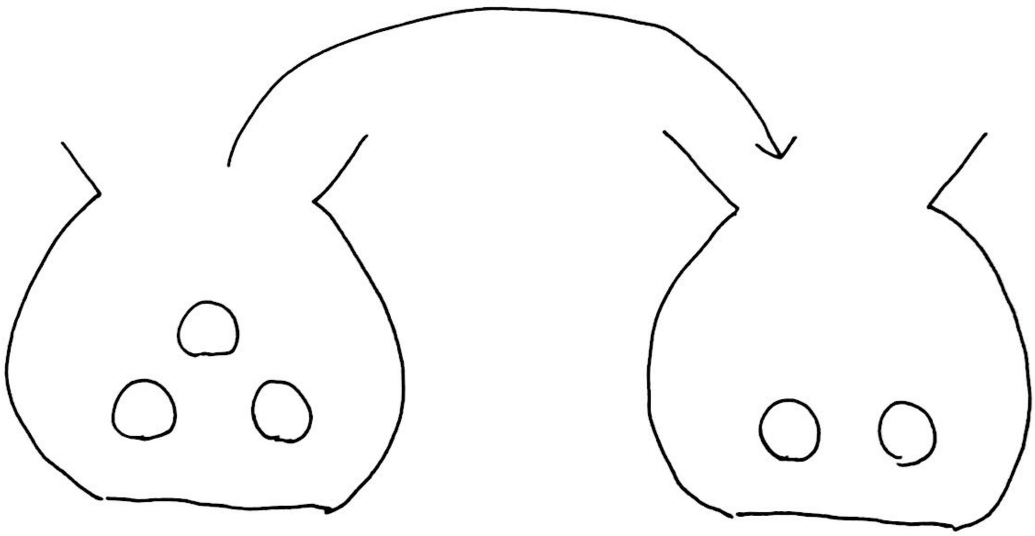

Figure 2: エーレンフェストの壺

1.3　クーポンコレクター問題

これは壺ではありませんが, やはり確率モデルが実際の現状を理解するのに役立つ例として紹介します. 今 10 種類のクーポンをすべてそろえたいとします. 毎回 (例えばスーパーのレジで) クーポンを受け取るのに, どの種類のクーポンをもらうのもすべて同じ確率だとします. 同じクーポンを何枚ももらう可能性がありますが, すべての種類を集めきるのに, 何回クーポンを引く必要があるでしょうか?

問題 1.9 (★). 2 種類のクーポンを集めるとき, 初めて 2 種類のクーポンが集まる回数の平均はいくつでしょうか?

問題 1.10 (★★). 10 種類のクーポンを集めるとき, 手持ちですでに 9 種類のクーポンがそろっているとします. 最後の 1 種類のクーポンが手に入るのにかかる回数の平均はいくつでしょうか?

問題 1.11 (★★★). 10 種類のクーポンを集めるとき, 初めて 2 種類のクーポンが集まる回数の平均はいくつでしょうか?

2　証明はディナーの後で.

ポリヤ[1]の壺は, もし最初に黒の球が 100 個, 白の球が 1 個しかなければ, 黒の球の数がどんどん増えていく確率が高くなるだろうということは, すぐに予想がつくと思

[1]George Pólya: 1887-1985. ハンガリー生まれの数学者.

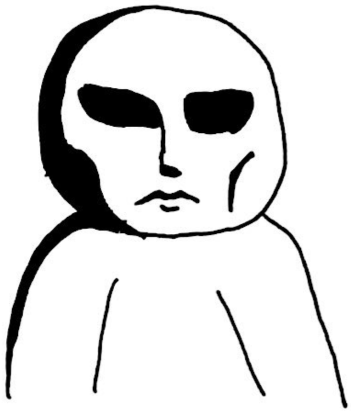

Figure 3: クーポンコレクター

います. また途中で, 黒の球が白の球より, 多くなれば, 黒の方が「有利」になり白が巻き返すのが難しくなっていくように想像されます. この壺のモデルは,「富める者はより富む[2]」という現象を理解するモデルになっています.

さて, ここで次のような変種を考えてみます. 黒の球を取り出した場合, その球を壺に戻し, さらに 2 つの黒の球を追加します. 白の球を取り出した場合, 再びその球を壺に戻し, 今度は 1 つの黒の球と 1 つの白の球を追加します. このモデルでは, 回数ごとに 2 つずつ球の総数は増えていきます. さて, n 回操作を行った後, 白の球の平均の個数はいくつでしょうか？ 実は, この場合だいたい $\sqrt{n}$ であることが証明できます. n に対して $\sqrt{n}$ はずっと小さいですが, それでも 1/2 の指数で増えていくことがわかります. これはランダムなネットワークのモデルで出てくる現象を理解するのに使われます. 例えば, 人間を点と思って, お友達である 2 つの点を辺で結びます. (このような点と辺からなる構造をグラフあるいはネットワークと呼びます. 関数の "グラフ" とは違うものです.) ここで, 毎回新しい人間がこのネットワークに加わっていく状況を考えます. 例えば, その新しい人間は, すでにネットワークにいる人たちから, 友達の多さに比例した確率で, 人を選び出し, その人 (またその人だけ) と友達になることにします. では, ある A さんを指定して, A さんのお友達の輪はどのように広がっていくでしょうか? 上のルールから A さんの友達の数は, 次のように増えていくことが分かります. まず各辺を半分に分け, それらを片側辺と呼ぶことにします (各点から千手観音のようにたくさんの手が伸びていて, その腕を片側辺と呼んでいます). A さんから出ている片側辺を白色に塗ります. それ以外の人の片側辺を黒色に塗ります. こうして, ネットワークの辺に半分白色, 半分黒色の辺と, 両方とも黒色の辺が現れます. 白色の片側辺の個数が A さんの友達の数です. 新しい人がネットワークに参入した場合, A さんの友達になる確率は白色の片側辺の個数に比例します. 友達にならない確率は黒色の片側辺の個数に比例します. 新しい人が A さんの友達になった場合, 白色の片辺側が 1 つ増え, 黒色の片側辺が 1 つ増えます. その人が A さんの友達にならなかった場合, 白色の片側辺は増えず, 黒色の片側辺が 2 つ増えます. 片側辺を球に置き換えれば, これはまさに上で考えたポリヤの壺

[2]「それ誰にても, 持てる人は与えられていよいよ豊かならん. 」(マタイ伝福音書 13 章 12 節)

の変種になっています.

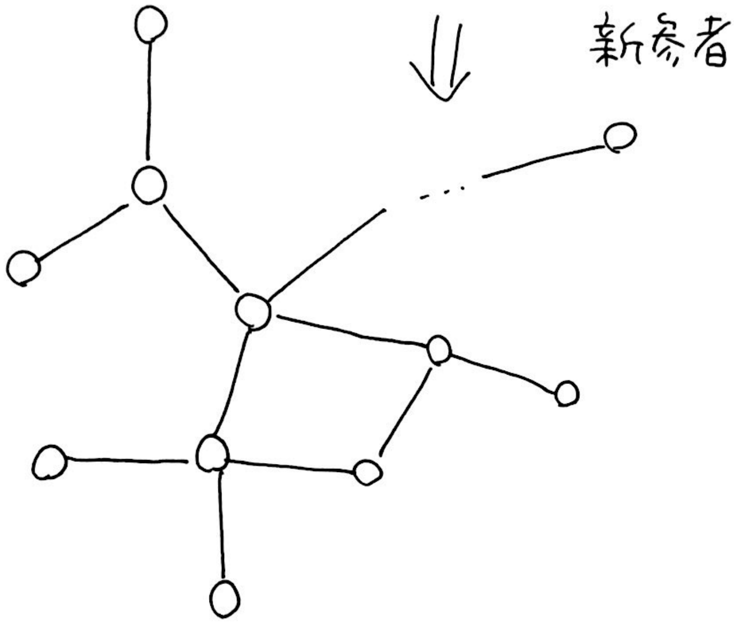

Figure 4: 増大するネットワーク

問題 2.1 (★★★★). 上のモデルで n 回操作を行った後, 白の球の平均の個数はだいたい $\sqrt{n}$ であることを示してください.

問題 2.2 (★★★★). 他にどのような変種が考えられるでしょうか? またその性質はどのようなものでしょうか? たとえば, 黒, 白, 赤の三色ではどのようなモデルが考えられるでしょうか?

エーレンフェスト[3] の壺は, 気体の温度がどのように一定になっていくかをモデル化したものです. 気体の温度を気体分子の運動から説明しようというのが統計力学です. このモデルは, そうした試みから考案された "由緒正しい"確率モデルです. これはある容器に仕切りがあり, 2 つに分けられた状況で, 片側の気体分子をもう片側に移し替えることを繰り返すことを考えていると思うことができます. 何度も移し替えて行くと, あるところから, 2 つの壺でだいたい同じ個数の球があるという状況から変化しにくくなっていく様子が分かります. このようなあまり変わらない状態を平衡状態と呼びます. 平衡状態を正確に記述するのが, 多くの問題で重要になります. 例えば, エーレンフェストの壺の場合, n 個の球 (粒子) がある場合, 片側の壺 A に k 個の球があり, もう片側の壺 B に $n-k$ この球があるような配置は, $\binom{n}{k} = \frac{n!}{k!(n-k)!}$ 通りあります. すべての可能な配置全体は 2^n 通りあります. このモデルでは, 壺 A

[3]Paul Ehrenfest: 1880-1933. オーストリア-オランダの物理学者, Tatiana Ehrenfest: 1876-1964. ロシア-オランダの数学者・物理学者.

に k 個の球があるという確率が $\frac{1}{2^n}\binom{n}{k}$ である分布が平衡分布 (平衡状態) になっています. 例えば $n=100$ のとき, 壺 A に球が 40 から 60 個ある確率は $0.96\ldots$, 0 から 40 個ある確率は $0.02\ldots$ になります. これは統計力学的には, 氷を常温で放置しておくと水になるが, 水を常温で放置しておいても氷になることは (現実的に) あり得ない, ということを説明するモデルになっています[4].

問題 2.3 (★★★★). このモデルで $n=100$ のとき, 壺 A に球が 40 から 60 個ある確率を求めてください.

クーポンコレクター問題では, 10 種類のクーポンの完全セットを得るのに, 少なくとも 10 回はクーポンを引かなければならないことは, 当たり前ですが, 10 回よりどれくらい多く引く必要があるだろうか, という問題です. 最初のうちは引くたびに新しいクーポンが得られるかもしれませんが, もし 9 種類集まっていてあと 1 種類のクーポンを引いたら終わり, という段階になって, 引くたびに, 持っていたクーポンと同じものを引いてしまって, 最後の 1 種類がなかなか集まらないという状況が起きえます. この効果が無視出来ずに残ることから, 結論から言うと, 10 種類のクーポンを集めるのに, 平均して約 23 回必要だということが分かります (2 種類では平均して 3 回です). より一般的に考えて n 種類のクーポンを初めて集めきる時間の平均はどうなるか考えてみます. 中間段階も考えて, 異なる k 種類のクーポンが初めて揃う時間の平均を T_k とします. T_n が求めたいものです. $T_k - T_{k-1}$ はいくつか考えてみることにします. これは今 $k-1$ 種類のクーポンを持っていて新しい種類のクーポンが手に入るまでの時間の平均です. ここで確率 $\frac{n-k+1}{n}$ で新しいクーポンが手に入り, 確率 $\frac{k-1}{n}$ ですでに手持ちのクーポンを引きます. ここから (少し計算してみると) 平均して $\frac{n}{n-k+1}$ で新しいクーポンが手に入ることが分かります. T_n はこれらの差 $T_k - T_{k-1}$ を足し上げれば良いので,

$$T_n = \sum_{k=1}^{n}(T_k - T_{k-1}) = \sum_{k=1}^{n}\frac{n}{n-k+1} = n\sum_{k=1}^{n}\frac{1}{k}$$

となることが分かります. $\sum_{k=1}^{n}\frac{1}{k}$ はだいたい $\log n$ なので $T_n \approx n\log n$ が近似値を与えます.

問題 2.4 (★★★★). 上の説明で計算を省略した部分を補ってください. つまり, $k-1$ 種類のクーポンを持っていて新しい種類のクーポンが手に入るまでの時間の平均は $\frac{n}{n-k+1}$ であることを示してください.

3 失われた記憶を求めて.

最後に確率論で重要な概念である**独立性**と**マルコフ性**について簡単にコメントをしておきます. ここで紹介したモデルたちでは, 同様の操作を何度も繰り返すということを行います. これは現在ある状況だけから次の取り得る状況になる確率が決まっていて, それ以前のその状況にどのように至ったか, ということを全く問いません.

[4]実際アボガドロ数のオーダー $n=1.0\times 10^{23}$ で考えると配置全体は途方もない場合の数になります.

そういう意味で, これらの操作は今ある状況以前のことは完全に忘れる, ということが仮定されています. この性質をマルコフ性と呼びます[5]. 一方で独立性は, ある試行の情報が別の試行の情報に全く影響しない状況を厳密に定式化する概念です. こうした数少ない仮定から多くの豊かな構造が説明されます. しかし, これらをきちんと定式化し説明するにはさらなる準備が必要なようです. ここでは, これらの概念が本質的に関わって, モデルの解析を可能にしている, ということを指摘するのに留めておくことにします.

問題 3.1 (★★★★★). 身の周りで簡単な確率モデルで説明できる現象を探してください.

[5] Andrei Markov: 1856-1922. ロシアの数学者.