五十種類を揃える回数の二割強は、最後の一種類を待つ時間である

— 50 種類が等しく出るくじを全部揃えるのに要る回数の期待値は 224.960267 回で、最後の一種類を待つ 50 回がその 0.222261 を占める — 全体の中央値は期待値の 0.955724 倍だが、最後の一種類だけなら 0.70 倍 — [論文609]

山岸 勇輝 (Yuuki YAMAGISHI)
Yamagishi Complexity Torus Lab, Japan｜山岸勇輝ラボ
2026年9月

— From thinking to syncing. —

要旨

n 種類の景品が等しい確率で一つずつ出るとき、全部を揃えるには何回引けばよいか（クーポン収集の問題）。その回数が、どこで費やされているかを数える。新しい定理も法則も主張しない。 第一に、50 種類を揃える回数の期待値は 224.960267 回である——すでに j 種類持っているとき、新しい種類が出るまでの期待値は $50/(50-j)$ 回で、それを足すと 50 に $1+1/2+\cdots+1/50$ を掛けた値になる。乱数で 20000 回揃えさせると平均は 225.598950 回で、式との差は標準誤差の 1.5 倍以内だった（第2節）。 第二に、これが本稿の芯である。そのうち二割強は、最後の一種類を待つ時間である——残り一種類になってから出るまでの期待値は 50 回で、全体の 0.222261 を占める。10 種類なら 0.341417、100 種類なら 0.192776 で、種類が増えても割合はゆっくりしか減らない（第2節）。 第三に、全体と最後の一種類では、ばらつきの形が違う——全体の中央値は 215 回で期待値の 0.955724 倍と近いが、最後の一種類だけを待つ時間の中央値は 35 回で、期待値 50 回の 0.70 倍である。一つの待ち時間では中央値が期待値の $\ln 2$ 倍ほどになり、多くの待ち時間を足すと 1 倍に近づく（第3節）。 第四に、分離子は、いちばん出にくい一種類である——49 種類は重み 1、一種類だけ重みを 0.1 にすると、その一種類の確率は 0.002037、それだけを待つ期待値は 491 回になり、全部揃う期待値は 536.096370 回、等しいときの 2.383071 倍になる。積分の式と乱数が、重み 1・0.5・0.2・0.1 のどれでも合った（第4節）。 キーワード：クーポン収集問題、調和数、期待値、最後の一種類、中央値。

> 本稿の射程（射程注記）：新しい定理も法則も主張しない——揃える回数の期待値が種類の数と調和数の積であること、その分布が大きい種類数で決まった形に近づくこと（エルデシュとレーニ、1961）、出る確率が等しくないときの期待値を積分で表すこと（フラジョレ、ガルディ、ティモニエ、1992）は、いずれも既知である。くじの設計や売り方を論じない——出にくい一種類の重みは、仕組みを見るための例である。分布の極限は扱わない——10・50・100 種類の期待値と、50 種類の乱数の分布だけを見る。既刊との関係：論文554 は、管理図が誤報を出すまでの点数が平均 370 点でも、半分は 256 点までに鳴ることを示した（中央値は平均の $\ln 2$ 倍に近い）——本稿の最後の一種類はその型の待ち時間で、全体はそれを 50 本足したものなので、中央値と期待値の比が 1 に近づく。加えたのは、揃える回数のうち最後の一種類が占める割合を種類数ごとに出したこと、全体と最後の一種類で中央値と期待値の比を並べたこと、分離子を「いちばん出にくい一種類」に置き、重みを変えて式と乱数で確かめたことである。

1. 記法と問い

設定：n 種類の景品が、一回引くごとに等しい確率 $1/n$ で一つ出る。全部の種類が一つ以上揃うまでに引く回数を T とする。すでに j 種類持っているとき、新しい種類が出る確率は $(n-j)/n$ なので、次の新しい種類までの回数の期待値は $n/(n-j)$ である。足し合わせると

$$E[T] = n\left(1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n}\right)$$

になる。括弧の中を調和数 H_n と呼ぶ。

問い：$E[T]$ のうち、どれだけが最後の一種類を待つ時間か。ばらつきはどんな形か。一種類だけ出にくいと、何が起きるか。

2. 期待値と最後の一種類

計算 2.1：調和数から期待値と標準偏差を出し、最後の一種類を待つ期待値 n 回が全体に占める割合を求めた。

種類 n	期待値 $E[T]$	標準偏差	最後の一種類の割合
10	29.289683	11.211025	0.341417
50	**224.960267**	61.950557	**0.222261**
100	518.737752	125.821704	0.192776

計算 2.2：式とは独立に、乱数で 50 種類を 20000 回揃えさせた。平均は 225.598950 回で、式との差 0.638683 は標準誤差 0.439149 の 1.5 倍以内だった。

> **観察 2.3**
>
> これが本稿の芯である。揃える回数の二割強は、最後の一種類を待つ時間である。最初の一種類は 1 回で出るが、残り一種類になると 50 回に一回しか新しいものが出ない——最後の一種類の 50 回は、全体の 224.960267 回の 0.222261 である。割合は $1/H_n$ で、n が 10 倍になっても 0.341417 から 0.192776 にしか減らない。

3. ばらつきの形

計算 3.1：乱数の 20000 回の分布を見た。中央値は 215 回で、期待値以下で揃ったのは 0.568800、期待値の 2 倍を超えたのは 0.005800 だった。中央値と期待値の比は **0.955724** である。

計算 3.2：最後の一種類だけを待つ時間は、毎回 1/50 の確率で出るまでの回数である。その中央値は 35 回で、期待値 50 回の 0.70 倍、$\ln 2 = 0.693147$ に近い。

> **観察 3.3**
>
> 全体と最後の一種類では、ばらつきの形が違う。一つの待ち時間は右に長く伸びるので、中央値が期待値の $\ln 2$ 倍ほどに下がる。全体は 50 本の待ち時間の和なので、伸びが均されて、中央値と期待値の比は 1 に近づく——それでも全体の中央値 215 回は期待値 224.960267 回より小さく、半分以上の人は期待値より早く揃う。

> **観察 3.4**
>
> 論文554 と並べる。554 は、管理図が誤報を出すまでの点数が平均 370 点でも、半分は 256 点までに鳴ることを示した——本稿の最後の一種類はその型の待ち時間で、中央値は期待値の 0.70 倍だった。揃える回数の全体は、その型の待ち時間を 50 本足したものである。

4. 分離子

計算 4.1：49 種類は重み 1、一種類だけ重み w で出るとして、全部揃う期待値を積分の式

$$E[T] = \int_0^\infty \left(1 - \prod_i \left(1 - e^{-p_i t}\right)\right) dt$$

で数値的に求め、乱数（4000 回ずつ）と比べた。この積分は、等しいときに 224.960267 を再現した。

重み w	出にくい一種類の確率	それだけを待つ期待値	全部揃う期待値	乱数
1	0.020000	50	224.960267	223.740500
0.5	0.010101	99	234.159532	232.313250
0.2	0.004065	246	323.833899	320.603000
0.1	0.002037	491	**536.096370**	530.662250

どの重みでも、積分の式と乱数の差は乱数の標準誤差の 4 倍以内だった。$w = 0.1$ の期待値は、等しいときの **2.383071** 倍である。

> 観察 4.2
>
> 分離子は、いちばん出にくい一種類である。重みを 0.1 にすると、その一種類だけを待つ期待値が 491 回になり、全部揃う期待値 536.096370 回のほとんどを占める——ほかの 49 種類は、その一種類を待つあいだにたいてい揃ってしまう。揃える回数は、種類の数よりも、いちばん出にくい一種類の確率で決まるようになる。

> 観察 4.3
>
> 言わないこと。くじの設計や売り方は論じない。分布の極限の形には立ち入らない——示したのは、期待値のうち最後の一種類が占める割合、全体と最後の一種類のばらつきの違い、そして一種類だけ出にくいときに何が支配するかである。

5. 点呼と結び

1. 50 種類の期待値 **224.960267** 回（計算 2.1）：$50 H_{50}$。
2. 乱数の平均 225.598950 回（計算 2.2）：標準誤差の 1.5 倍以内。
3. 最後の一種類が **0.222261**（計算 2.1）：本稿の芯。
4. 10 種類で 0.341417、100 種類で 0.192776（計算 2.1）：ゆっくりしか減らない。

5. 全体の中央値 **215** 回、比 **0.955724**（計算 3.1）：期待値に近い。
6. 最後の一種類だけなら中央値 **35** 回、比 **0.70**（計算 3.2）：$\ln 2$ に近い。
7. 一つの待ち時間は伸び、足すと均される（観察 3.3）：論文**554** の型。
8. 一種類だけ重み **0.1** で期待値 **536.096370** 回（計算 4.1）：**2.383071** 倍。
9. その一種類だけを待つ **491** 回がほとんど（観察 4.2）：分離子。
10. 積分の式と乱数が四つの重みで合う（計算 4.1）：二つの道筋。

結び：**50** 種類が等しく出るくじを全部揃えるには、期待値で **224.960267** 回引く。そのうち二割強は、最後の一種類を待つ時間である——残り一種類になると **50** 回に一回しか新しいものが出ず、その **50** 回が全体の **0.222261** を占める。割合は種類の数の調和数の逆数で、**100** 種類でも **0.192776** にしか減らない。全体と最後の一種類では、ばらつきの形が違う——最後の一種類だけなら中央値は期待値の **0.70** 倍だが、**50** 本の待ち時間を足した全体では **0.955724** 倍になる。分離子は、いちばん出にくい一種類である——一種類だけ重みを十分の一にすると期待値は **536.096370** 回、**2.383071** 倍になり、その一種類を待つ **491** 回がほとんどを占める。既刊との位置——論文**554** が見た一つの待ち時間の型を、**50** 本足すと何が均され、何が残るかを見た。正直に言えば——分布の極限の形には入らず、くじの設計も論じていない。

参考文献

1. Erdős, P., & Rényi, A. (1961). On a classical problem of probability theory. *Magyar Tudományos Akadémia Matematikai Kutató Intézetének Közleményei*.
2. Flajolet, P., Gardy, D., & Thimonier, L. (1992). Birthday paradox, coupon collectors, caching algorithms and self-organizing search. *Discrete Applied Mathematics*.
3. 山岸勇輝 (2026). 『「平均 **370** 点までもつ」の半分は、**256** 点までに鳴る』[論文554]. Zenodo. 10.5281/zenodo.23014204

山岸 勇輝 (Yuuki YAMAGISHI)

© 2026 Yamagishi Complexity Torus Lab, Japan · Licensed under Creative Commons Attribution 4.0 International (CC BY 4.0)