More Than a Fifth of the Draws Needed to Collect Fifty Kinds Are Spent Waiting for the Last One

— Collecting all of **50** equally likely kinds takes **224.960267** draws on average, and the **50** spent waiting for the last kind are **0.222261** of them — the median of the whole is **0.955724** times the mean, but for the last kind alone it is **0.70** times — [Paper 609]

Yuuki YAMAGISHI (山岸 勇輝)
Yamagishi Complexity Torus Lab, Japan
September 2026

— From thinking to syncing. —

Abstract

When n kinds of prize come out one at a time with equal probability, how many draws are needed to collect them all (the coupon collector's problem)?Where those draws are spent is counted.No new theorem or law is claimed. First, collecting 50 kinds takes 224.960267 draws on average——holding j kinds, a new kind takes $50/(50-j)$ draws on average, and the sum is 50 times $1+1/2+\cdots+1/50$. Collecting 20000 times with random numbers gave a mean of 225.598950, within 1.5 standard errors of the formula (Section 2). **Second, and this is the core.More than a fifth of it is spent waiting for the last kind——once one kind remains, it takes 50 draws on average, 0.222261 of the whole. With 10 kinds the share is 0.341417, with 100 kinds 0.192776; it falls only slowly as the kinds increase** (Section 2). **Third, the whole and the last kind spread differently——the median of the whole is 215 draws, 0.955724 times the mean, close to it, but the median wait for the last kind alone is 35 draws, 0.70 times its mean of 50. For a single wait the median is about $\ln 2$ times the mean; adding many waits brings the ratio near 1** (Section 3). **Fourth, the separator is the least likely kind——with 49 kinds of weight 1 and one of weight 0.1, that kind has probability 0.002037, waiting for it alone takes 491 draws on average, and collecting everything takes 536.096370, 2.383071 times the equal case. The integral formula and random collection agreed for weights 1, 0.5, 0.2 and 0.1** (Section 4). **Keywords**: coupon collector's problem, harmonic numbers, expected value, last coupon, median.

Scope of this paper (scope note): No new theorem or law is claimed——that the mean number of draws is the number of kinds times the harmonic number, that its distribution approaches a fixed shape for many kinds (Erdős and Rényi, 1961), and that the mean for unequal probabilities is given by an integral (Flajolet, Gardy and Thimonier, 1992) are all known.The design or sale of lotteries is not discussed——the weight of the rare kind is an example for seeing the mechanism.Limits of the distribution are not treated——only the means for 10, 50 and 100 kinds and the random distribution for 50 kinds are looked at.Relation to earlier papers: Paper 554 showed that although a control chart runs 370 points on average before a false alarm, half alarm within 256 points (the median is close to $\ln 2$ times the mean) ——the last kind here is a wait of that type, and the whole is the sum of 50 of them, so the ratio of median to mean approaches 1.What is added is giving, for each number of kinds, the share of the draws taken by the last kind, setting the ratio of median to mean side by side for the whole and the last kind, and placing the separator on the least likely kind and checking it with the formula and random collection for several weights.

1. Notation and the Question

Setting: n kinds of prize come out one per draw, each with probability $1/n$. Let T be the number of draws until every kind has appeared at least once.Holding j kinds, a new kind appears with probability $(n-j)/n$, so the mean number of draws to the next new kind is $n/(n-j)$. Adding these gives

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

The bracket is called the harmonic number H_n.

Question: how much of $E[T]$ is spent waiting for the last kind, what is the shape of the spread, and what happens if one kind is less likely?

2. The Mean and the Last Kind

Computation 2.1: from the harmonic numbers the mean and standard deviation were found, together with the share of the whole taken by the n draws spent on the last kind.

kinds n	mean $E[T]$	standard deviation	share of the last kind
10	29.289683	11.211025	0.341417
50	**224.960267**	61.950557	**0.222261**
100	518.737752	125.821704	0.192776

Computation 2.2: **independently of the formula, 50 kinds were collected 20000 times with random numbers.The mean was 225.598950 draws; its difference 0.638683 from the formula is within 1.5 times the standard error 0.439149.**

> **Observation 2.3**
>
> **this is the core: more than a fifth of the draws are spent waiting for the last kind.The first kind comes in 1 draw, but with one kind left, only one draw in 50 brings something new——those 50 draws are 0.222261 of the whole 224.960267. The share is $1/H_n$, falling only from 0.341417 to 0.192776 when n grows tenfold.**

3. The Shape of the Spread

Computation 3.1: **the distribution of the 20000 random collections was examined.The median was 215 draws; 0.568800 finished at or below the mean and 0.005800 took more than twice the mean. The ratio of median to mean is 0.955724.**

Computation 3.2: **the wait for the last kind alone is the number of draws until an event of probability 1/50 occurs.Its median is 35 draws, 0.70 times the mean of 50, close to $\ln 2 = 0.693147$.**

> **Observation 3.3**
>
> **the whole and the last kind spread differently.A single wait stretches far to the right, pulling the median down to about $\ln 2$ times the mean. The whole is the sum of 50 waits, which evens out the stretch and brings the ratio of median to mean close to 1——still, the median 215 of the whole is below the mean 224.960267: more than half of collectors finish sooner than the mean.**

> **Observation 3.4**
>
> **beside Paper 554.Paper 554 showed that although a control chart runs 370 points on average before a false alarm, half alarm within 256——the last kind here is a wait of that type, with a median 0.70 times its mean. The whole number of draws is the sum of 50 waits of that type.**

4. The Separator

Computation 4.1: with 49 kinds of weight 1 and one kind of weight w, the mean number of draws was found numerically from the integral formula

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

and compared with random collection (4000 times each). The integral reproduced 224.960267 in the equal case.

weight w	probability of the rare kind	mean wait for it alone	mean to collect all	random
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

For every weight, the integral and random collection differed by less than 4 standard errors of the random mean. The mean for $w = 0.1$ is **2.383071** times the equal case.

> **Observation 4.2**
>
> the separator is the least likely kind.With weight 0.1, waiting for that kind alone takes 491 draws on average, most of the 536.096370 needed for everything——the other 49 kinds are usually complete while one waits for it. The number of draws comes to be decided less by the number of kinds than by the probability of the least likely one.

> **Observation 4.3**
>
> what is not said.The design or sale of lotteries is not discussed, and the limiting shape of the distribution is not entered——what is shown is the share of the mean taken by the last kind, the different spreads of the whole and the last kind, and what dominates when one kind is rarer.

5. Roll Call and Closing

1. **mean for 50 kinds 224.960267** (Computation 2.1): $50H_{50}$**.**
2. **random mean 225.598950** (Computation 2.2): **within 1.5 standard errors.**
3. **the last kind takes 0.222261** (Computation 2.1): **the core.**
4. **0.341417 for 10 kinds, 0.192776 for 100** (Computation 2.1): **falling only slowly.**
5. **median of the whole 215, ratio 0.955724** (Computation 3.1): **close to the mean.**
6. **the last kind alone: median 35, ratio 0.70** (Computation 3.2): **close to $\ln 2$.**
7. **a single wait stretches; a sum evens out** (Observation 3.3): **the type of Paper 554.**
8. **one kind at weight 0.1: mean 536.096370** (Computation 4.1): **2.383071 times.**
9. **the 491 draws for that kind are most of it** (Observation 4.2): **the separator.**
10. **integral and random agree for four weights** (Computation 4.1): **two routes.**

Closing: **Collecting all of 50 equally likely kinds takes 224.960267 draws on average.More than a fifth of that is spent waiting for the last kind——with one kind left, only one draw in 50 brings something new, and those 50 draws are 0.222261 of the whole. The share is the reciprocal of the harmonic number and is still 0.192776 for 100 kinds.The whole and the last kind spread differently——for the last kind alone the median is 0.70 times the mean, while for the sum of 50 waits it is 0.955724 times.The separator is the least likely kind——making one kind ten times less likely raises the mean to 536.096370, 2.383071-fold, and the 491 draws spent waiting for it take up most of it.Placed among the earlier papers——it looks at what is evened out, and what remains, when 50 waits of the type seen in Paper 554 are added.To be honest——the limiting shape of the distribution is not entered, and the design of lotteries is not discussed.**

References

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. Yamagishi, Y. (2026). *Half the Charts That "Last 370 Points on Average" Alarm by 256* [Paper 554]. Zenodo. 10.5281/zenodo.23014204

Yuuki YAMAGISHI (山岸 勇輝)

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