Mathematics Internal Assessment:

How to Collect All Stickers with Minimum Purchasement of Pokemon Breads Using the Coupon Collector's Problem

Introduction

“Pokemon Bread craze” was the hottest issue at the start of 2022. Pokemon Bread, which was re-released after 16 years, gained unbelievable popularity among maniacs and enthusiasts who wanted to collect the stickers inside rather than the taste. As the popularity of stickers soared through social media on the internet, the younger generation - who generally have the strong desire to keep up with trends - became more enthusiastic about Pokemon bread even though they don’t have a strong inclination to literally consume the bread. Despite many people trying to buy the bread from the convenience store, Pokemon Breads weren’t able to be seen easily in many stores since the demand was unrealistically higher compared to the demand. The price of each sticker online skyrocketed, and some even bought the sticker separately to collect the stickers. One main goal of the sticker collectors, which included myself, was to assemble all 160 types of Pokemon stickers.

There are three aims to this investigation: first, I will calculate the minimum number of Pokemon bread that needs to be purchased in order to collect all 160 stickers and calculate the price (money needed); then, I will derive the most economic way to collect all stickers by comparing three methods - purchasing Pokemon breads only, purchasing Pokemon bread for some stickers but purchase some stickers independently online, or purchasing all stickers online; lastly, I will obtain a more realistic scope of collecting stickers by assuming that the probability of each sticker showing up may not be uniform.

This topic was chosen out of personal interest in the area of advanced probability and economics. I wanted to investigate a real life situation where the circumstances are often very different from ideal mathematical situations. In many contextual situations in mathematics textbooks, too many considerations are omitted. By adding 3 more hypothetical variables in the investigation, I am willing to provide a mathematical calculation or model that better reflects probability in real life.

Background Information

Bernoulli Experiment and Geometric Distribution

Bernoulli experiment is a random experiment that has two possible outcomes: success and failure. It counts the number of failures X until it meets the first success and since the experiment is random, X is a random variable. Here, if each repetition trial is independent of each other, the distribution of X is referred to as geometric distribution.

Geometric distribution will be used in order to find the minimum number of Pokemon breads I need to buy. Geometric distribution is a type of discrete probability distribution of a random variable X which fulfils certain conditions. In order to use a geometric distribution in a situation, the situation must have a series of trials, each trial must have only 2 possible outcomes of success and failure, and the probability of success should be equal for all trials. In other words, the geometric distribution shows the probability of success after N number of trials. Each trial has the success probability - the probability of desired outcome - of p and probability of failure of $q = 1 - p$. In a hypothetical situation, the first trial will have the success probability of p and the second will have $q \times p$ since it needs to fail in the first trial to succeed in the second trial. The random discrete variable X’s geometric distribution can be written as

$$X \sim Geo(p)$$

and the probability of success in the *r*th trial can be formulated as:

$$P(X = r) = q^{r-1} \times p$$

And the probability that X is greater than the number of trials *r* will be written as:

$$P(X > r) = 1 - P(X \leq r)$$

The equation can be re-written as:

$$= 1 - [P(X = 1) + P(X = 2) + P(X = 3) + ... + P(X = r)$$

Since $P(X = r) = q^{r-1} \times p$, this upper equation can be written as:

$$= 1 - (p + qp + q^2p + ... + q^{r-1}p)$$

According to the geometric sum, $S_n = a + ar + ar^2 + ... + ar^{n-1}$ which is $\frac{a(1-r^n)}{1-r}$. Thereby, the upper equation can be written as:

$$= 1 - p[\frac{1(1-q^r)}{1-q}]$$

$$\therefore P(X > r) = 1 - (1 - q^r) = q^r$$

In conclusion, the probability that random discrete variable X is larger than the number of trials *r* is equal to the probability of failure *q* multiplied *r* times.

Here, the expected value or mean of the distribution of X is shown as:

$$E(X) = \Sigma xP(X = x)$$

$$E(X) = p + 2qp + 3q^2p + 4q^3p + ... (a)$$

$$qE(X) = qp + 2q^2p + 3q^3p + ... (b)$$

When we subtract (b) from (a):

$$= E(X) - qE(X) = p + qp + q^2p + q^3p + ...$$

$$= (1 - q)E(X) = p(1 + q + q^2 + q^3 + ...)$$

$$= (1 - q)E(X) = p(\frac{1}{1-q})$$

$$= pE(X) = 1$$

$$\therefore E(X) = \frac{1}{p}$$

To brief the equations, in a geometric distribution $X \sim Geo(p)$,

<u>The Coupon Collector's Problem</u>

The coupon collector's problem uses geometric distribution in order to find the minimum number of coupons to collect all types of coupons. Each coupon is numbered 1, 2, …, N and each successive item is an independent random variable that assume the value *k* with the probability p_k. Here, the main inquiry is the expected number of coupons we need to collect in order to complete the collection of all N coupons.

If we assume that there are N coupons that are equally likely to be chosen, the probability that each coupon would appear at any time will be written as $\frac{1}{N}$. There are geometric distribution approaches that I have explored previously that can be used to explain the coupon collector's problem.

Just like I have explored in the Bernoulli Experiment and geometric distribution, I will let *X* stand for the random amount of coupons we need to buy to complete the collection and $X = X_1 + X_2 + ... + X_N$. For every $i \in \{1, 2,..., N\}$, X_i represents the additional number of coupons that we must buy in order to move from $i - 1$ to *i* distinct sorts of coupons in our collection. Since I am discussing the situation of a uniform distribution and trivially $X_1 = 1$, it follows that after collecting *i* different types of coupons, a new coupon purchased will be of a different type with a probability of $\frac{N-i+1}{N}$. The random variable X_i, for $i \in \{2, 3, ..., N\}$, is independent of the other variables and has a geometric rule with parametre of $\frac{N-i+1}{N}$. Hence, the expected number of coupons needed to purchase to complete the collection can be derived as:

$$E[X] = E[X_1] + ... + E[X_N]$$

$$= 1 + \frac{N}{N-1} + \frac{N}{N-2} + ... + \frac{N}{2} + N$$

$$= N\sum_{i=1}^{N}\frac{1}{i}$$

When the expected value E[X] is plotted as the y axis and number of coupons *N* as the axis, the following graph can be drawn:

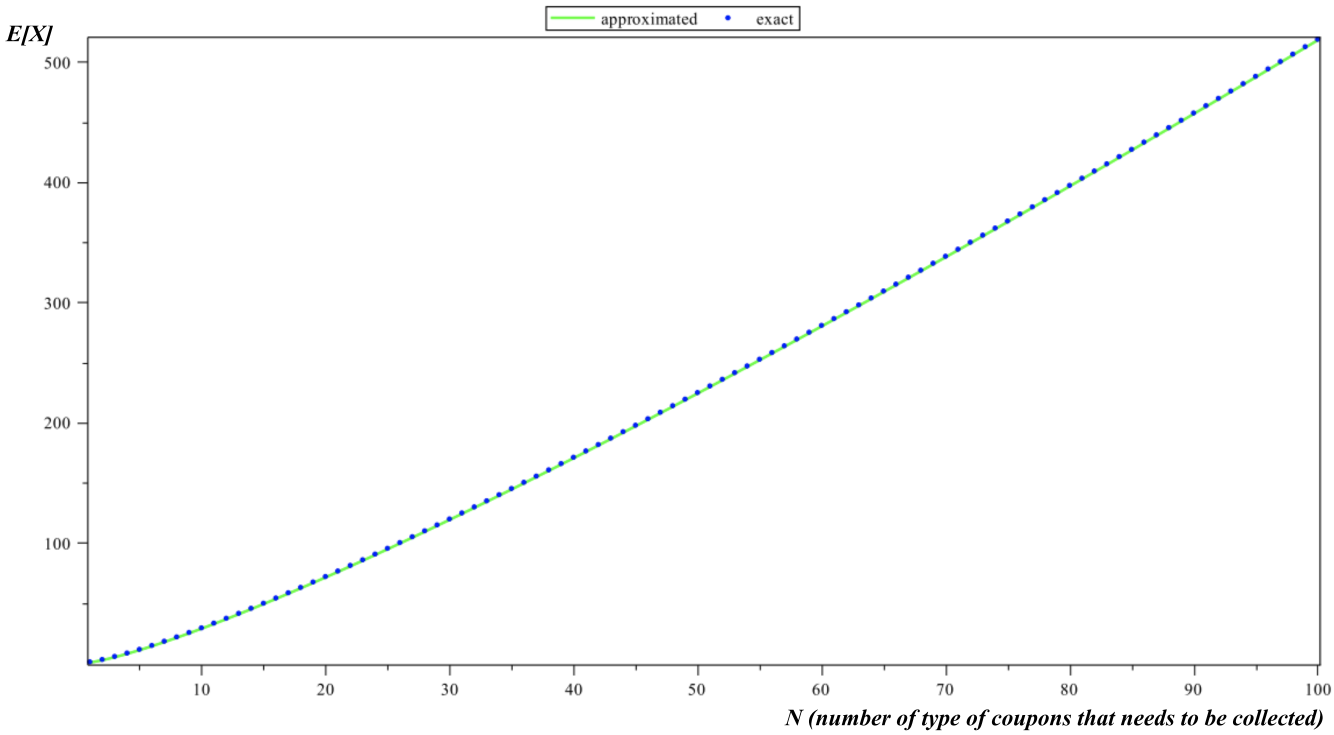

Figure 1. Exact and approximated values of E[X] depending on *N*

When the Probability of Obtaining Each Sticker is Uniform

My investigation is primarily rooted in the coupon collector's problem, where instead of coupons, stickers are collected. This means that I will employ the geometric distribution and expected value formula from the coupon collector's problem.

When Purchasing the Stickers Separately Online

https://taeyeobv.tistory.com/entry/%EC%83%88%EB%A1%9C%EB%82%98%EC%98%A8-%ED%8F%AC%EC%BC%93%EB%AA%AC-%EB%9D%A0%EB%B6%80%EB%9D%A0%EB%B6%80%EC%94%B0-%EC%8B%9C%EC%84%B8%EB%8A%94-%EC%96%BC%EB%A7%88%EC%9D%BC%EA%B9%8C%EC%98%A4%EB%A5%98%EC%94%B0-%ED%8F%AC%ED%95%A8

This website shows the current price of each sticker. Although the price of stickers are different, I will calculate the average price.

Average price=sum of all price total number of stickers (160)

By combining with approach 1, I will find the most economical method.
→ how many pieces of bread should I buy and how many stickers should I buy separately to spend minimal money in collecting all stickers?

When Probability of Pokemon Bread not being In the Convenience Store

Conclusion

Bibliography

https://www.statlect.com/fundamentals-of-probability/random-variables
https://brilliant.org/wiki/euler-mascheroni-constant/
https://mat.uab.cat/matmat_antiga/PDFv2014/v2014n02.pdf