Collector's Problem

A cereal company includes a hockey card in each cereal box that it sells. A complete collection consists of 10 cards. Suppose that all cards are equally likely to appear in any cereal box. On average, how many cereal boxes does someone have to buy to get a complete collection?

To solve this problem, we will use the following Theorem.

Theorem. *Let X be the number of trials until the first success in a sequence of Bernoulli trials (e.g.,* FFFS *gives $X = 4$). Then,*

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

where p is the probability of a success.

Proof. If $q = 1 - p$ is the probability of a failure, we have

$$P(X = n) = q^{n-1}p, \quad \text{for all } n = 1, 2, 3, \ldots$$

since $X = n$ means that the first $n - 1$ trials were failures and the n-th trial was a success. We then have that

$$\begin{aligned} E(X) &= \sum_{n=1}^{\infty} nP(X = n) = \sum_{n=1}^{\infty} \left(nq^{n-1}p\right) = p \cdot \sum_{n=1}^{\infty} nq^{n-1} \\ &= p \cdot \frac{d}{dq}\left(\sum_{n=0}^{\infty} q^n\right) = p \cdot \frac{d}{dq}\left(\frac{1}{1-q}\right) = \frac{p}{(1-q)^2} \\ &= \frac{1}{p}. \end{aligned}$$

□

Let's now solve the collector's problem if there are n different cards to collect. Let T be the total number of cereal boxes needed until one gets the complete collection. We have that

$$T = X_1 + X_2 + X_3 + \cdots + X_n$$

where X_i is the number of trials needed to get a different card if we already have $i - 1$ of the cards. For the random variable X_i, the probability of a success is

$$p_i = \frac{n - i + 1}{n}, \quad \text{for all } i = 1, 2, 3, \ldots, n.$$

Using the above Theorem, we deduce that

$$\begin{aligned} E(T) &= E(X_1) + E(X_2) + E(X_3) + \cdots + E(X_n) \\ &= \frac{1}{p_1} + \frac{1}{p_2} + \frac{1}{p_3} + \cdots + \frac{1}{p_n} \\ &= \frac{n}{n} + \frac{n}{n-1} + \frac{n}{n-2} + \cdots + \frac{n}{1} \\ &= n\left(1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n}\right). \end{aligned}$$

If there are 10 different hockey cards to collect, we have

$$E(T) = 10\left(1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{10}\right) \approx 29.3$$

which means that on average, someone has to buy about 29 cereal boxes to get a complete collection.