A GENERALIZATION OF BANACH'S MATCHBOX PROBLEM (with a use in computer science)

SREEJIT ROY, NILADRI KAL, DEBARSHI CHAKRABORTY
(Department of Statistics, 2nd Year)

1. INTRODUCTION.

While we went through Banach's matchbox problem for the first time, during our probability theory 1 course, it was nothing more than all the other rather harmless problems. Later, on a random day, while discussing some other problems, a thought on generalization of that problem emerged. As anybody else would do, we, after failing to generalize so in a few attempts, asked the internet. To our surprise, we could not find any sufficient answer to our certain problem. We kept on trying.

In the following sections, a discussion on the generalization, along with one use of that in the field of computer science, is discussed in detail. Reader's familiarity with rudimental concepts of basic probability theory and distribution theory is assumed.

2. A MEMORY ALLOCATION PROBLEM.

Suppose, k works are being performed by a memory unit, with the i^{th} one occupying n_i bytes initially. The probability of one byte being emptied from the i^{th} task is p_i. Then, what is the probability of having x bytes of the memory occupied, when one task is found to be complete?

3. IS IT BANACH'S?

An obvious question that may arise is that what is the connection of the aforementioned problem with a generalization of Banach's matchbox problem. It would be clearer at the onset of section 5.

4. SIMPLE BANACH'S PROBLEM.

Setup 1. A certain mathematician carries two matchboxes, one on his right pocket and one on his left, containing n_1 and n_2 matches initially. When in need of a match, he may pick the right one, with probability p_1, or the left one, with probability p_2.

Problem 1. Let
X := Number of matches remaining in one box when the other is found to be empty.
Then, for setup 1, evaluate Pr(X=x).
Solution. It is clear that he can either find the box on the right pocket empty, or that on the left. Hence, we can write,

Pr(X=x) = Pr(right has x | left empty) + Pr(left has x | right empty)

$$= \binom{n_2+(n_1-x)}{n_2} p_2^{n_2} p_1^{(n_1-x)} p_2 + \binom{n_1+(n_2-x)}{n_1} p_1^{n_1} p_2^{(n_2-x)} p_1$$

$$= p_1^{n_1} p_2^{n_2} \left\{ \binom{n_2+(n_1-x)}{n_2} \left(\frac{p_1}{p_2^x}\right) + \binom{n_1+(n_2-x)}{n_1} \left(\frac{p_2}{p_1^x}\right) \right\}$$

Special Case. For $n_i = n \ \forall i$ and $p_i = 0.5 \ \forall i$, we have Pr(X=x) = $\binom{2n-x}{n}(0.5)^{2n-x}$.

5. GENERALIZED BANACH'S PROBLEM.

Setup 2. A certain mathematician carries k matchboxes, with the i^{th} one containing n_i matches initially. When in need of a match, he may pick the i^{th} box, with probability p_i.

With setup 2 in hand, the fact that the problem stated in section 2 is indeed a generalization of simple Banach's problem, should be clear to the reader.

Problem 2. Let
X := A random vector of order k-1 denoting the number of matches remaining in the boxes left, when one is found to be empty.
Then, for setup 2, evaluate Pr(**X**=**x**).
Solution. Due to the lack of adequate mathematical groundwork, we consider different values of k, n_i, and p_i, in an attempt to reach to a conclusion about the general case.

Case A. $(k = 3, n_i = n \ \forall i, p_i = \frac{1}{k} \ \forall i)$

Pr(**X**=**x**) = Pr(X_1=x_1, X_2=x_2)

$$= 3 \binom{n+(n-x_1)+(n-x_2)}{n-x_1, n-x_2} \left(\frac{1}{3}\right)^n \left(\frac{1}{3}\right)^{n-x_1} \left(\frac{1}{3}\right)^{n-x_2} \left(\frac{1}{3}\right)$$

$$= \left(\frac{1}{3}\right)^{3n} \binom{3n-(x_1+x_2)}{n-x_1, n-x_2} \left(\frac{1}{3}\right)^{-(x_1+x_2)}$$

Case B. $(k, n_i = n \ \forall i, p_i = \frac{1}{k} \ \forall i)$

$$\Pr(\mathbf{X}=\mathbf{x}) = k \binom{n+(n-x_1)+(n-x_2)+\cdots+(n-x_{k-1})}{n-x_1, n-x_2, \ldots, n-x_{k-1}} \left(\frac{1}{k}\right)^n \left(\frac{1}{k}\right)^{n-x_1} \left(\frac{1}{k}\right)^{n-x_2} \ldots \left(\frac{1}{k}\right)^{n-x_{k-1}} \left(\frac{1}{k}\right)$$

$$= \left(\frac{1}{k}\right)^{kn} \binom{kn-(x_1+x_2+\cdots+x_{k-1})}{n-x_1, n-x_2, \ldots, n-x_{k-1}} \left(\frac{1}{k}\right)^{-\sum_{i=1}^{k-1} x_i}$$

Case C. $(k, n_i, p_i = \frac{1}{k} \ \forall i)$

$$\Pr(\mathbf{X}=\mathbf{x}) = k \binom{n_k+(n_1-x_1)+(n_2-x_2)+\cdots+(n_{k-1}-x_{k-1})}{n_1-x_1, n_2-x_2, \ldots, n_{k-1}-x_{k-1}} \left(\frac{1}{k}\right)^{n_k} \left(\frac{1}{k}\right)^{n_1-x_1} \left(\frac{1}{k}\right)^{n_2-x_2} \ldots \left(\frac{1}{k}\right)^{n_{k-1}-x_{k-1}} \left(\frac{1}{k}\right)$$

$$= \left(\frac{1}{k}\right)^{\sum_1^k n_i} \binom{\sum_1^k n_i - \sum_1^{k-1} x_i}{n_1-x_1, n_2-x_2, \ldots, n_{k-1}-x_{k-1}} \left(\frac{1}{k}\right)^{-\sum_1^{k-1} x_i}$$

Case D. (k, n_i, p_i)

Pr(**X**=**x**)

$$= \sum_k \binom{n_k+(n_1-x_1)+(n_2-x_2)+\cdots+(n_{k-1}-x_{k-1})}{n_1-x_1, n_2-x_2, \ldots, n_{k-1}-x_{k-1}} (p_k)^{n_k} (p_1)^{n_1-x_1} (p_2)^{n_2-x_2} \ldots (p_{k-1})^{n_{k-1}-x_{k-1}} (p_k)$$

$$= \left(\prod_{i=1}^k p_i^{n_i}\right) \sum_k \binom{\sum_1^k n_i - \sum_1^{k-1} x_i}{n_1-x_1, n_2-x_2, \ldots, n_{k-1}-x_{k-1}} \prod_{i=1}^{k-1} p_i^{-x_i}$$ ●