Probability & Statistics for EECS:
TA Lecture #03 - Random Variables

Problem 1 (EXTRA BH CH2 # 62: Difference equation)

There are n types of toys, which you are collecting one by one. Each time you buy a toy, it is randomly determined which type it has, with equal probabilities. Let $p_{i,j}$ be the probability that just after you have bought your i^{th} toy, you have exactly j toy types in your collection, for $i \geq 1$ and $0 \leq j \leq n$. (This problem is in the setting of the coupon collector problem, a famous problem which we study in Example 4.3.11.)

(a) Find a recursive equation expressing p_{ij} in terms of $p_{i-1,j}$ and $p_{i-1,j-1}$, for $i \geq 2$ and $1 \leq j \leq n$.

(b) Describe how the recursion from (a) can be used to calculate $p_{i,j}$.

Solution

(a) Given that you have bought your ith toy, you have exactly j toy types in your collection, the possible cases in previous buying are as follows:

- After buying the $(i-1)$th toy, you have exactly j toy types, which suggests that the toy you bought in ith round is same as some of toys you've already got.
- After buying the $(i-1)$th toy, you have exactly $(j-1)$ toy types, which suggests in the ith round, you bought a toy different from any of those you've got.

As a result,

$$p_{i,j} = p_{i-1,j} \cdot \frac{j}{n} + p_{i-1,j-1} \cdot \frac{n-j+1}{n}$$

for $i \geq 2$ and $1 \leq j \leq n$, with $j \leq i$ and

$$p_{i,j} = 0, \forall j > i.$$

(b) Firstly, we know $p_{i,0}$ is the probability that you've still got nothing after i ($i \geq 1$) rounds, which is impossible and therefore $p_{i,0} = 0$. We just need to focus on calculating $p_{i,j}$ from basic cases for $i \geq j$, since $p_{i,j} = 0$ for $j > i$. Starting from $p_{1,1} = 1$, we could calculate

$$p_{2,1} = p_{1,1}\frac{1}{n}, p_{2,2} = p_{1,1}\frac{n-1}{n}.$$

With $p_{2,1}$ and $p_{2,2}$, we could calculate

$$\begin{aligned} p_{3,1} &= p_{2,1}\frac{1}{n} \\ p_{3,2} &= p_{2,2}\frac{2}{n} + p_{2,1}\frac{n-1}{n} \\ p_{3,3} &= p_{2,2}\frac{n-2}{n}. \end{aligned}$$

Likewise, given $p_{k,m}$ for $m = 1, 2, \ldots, k$, we could calculate the values of $p_{k+1,m'}$ for $m' = 1, 2, \ldots, k+1$, using the recursive equations.