数据结构与算法 II 作业 (9.29)

中国人民大学　信息学院　崔冠宇　2018202147

P72, T5.3-4　Armstrong 教授建议用下面的过程来产生一个均匀随机排列：

PERMUTE-BY-CYCLIC(A)

```
n = A.length
let B[1..n] be a new array
offset = RANDOM(1, n)
for i = 1 to n
    dest = i + offset
    if dest > n
        dest = dest - n
    B[dest] = A[i]
return B
```

请说明每个元素 $A[i]$ 出现在 B 中任何特定位置的概率是 $1/n$。然后通过说明排列结果不是均匀随机排列，表明 Armstrong 教授错了。

解: 由于 `B[((i + offset - 1) MOD n) + 1] = A[i]`，一旦 `offset` 确定，整个序列就确定了，而 `offset` 是在循环开始前由 `RANDOM` 确定的 $1 \sim n$ 的均匀随机数，所以 $A[i]$ 出现在 B 中任何特定位置的概率都是 $1/n$。但是由于数组元素的相对顺序没变（将整个数组看成一个循环队列的话），只能产生 n 种排列，所以不是均匀随机排列。

P73, T5.3-5　证明：在过程 PERMUTE-BY-SORTING 的数组 P 中，所有元素都唯一的概率**至少**是 $1-1/n$。

证明:

$$
\begin{aligned}
P &= \frac{A_{n^3}^n}{(n^3)^n} \\
&= \frac{n^3}{n^3}\frac{n^3-1}{n^3}\cdots\frac{n^3-n+1}{n^3} \\
&\geq \frac{n^3-n}{n^3}\frac{n^3-n}{n^3}\cdots\frac{n^3-n}{n^3} \\
&> (1-\frac{n}{n^2}) \qquad (\text{因为}(1-x)(1-y) > 1-x-y, \text{当}a,b>0) \\
&= 1-\frac{1}{n}
\end{aligned}
$$

P79, T5.4-2　假设我们将球投入到 b 个箱子里，直到某个箱子中有两个球。每一次投掷都是独立的，并且每个球落入任何箱子的机会均等。请问投球次数次数期望是多少?

解: 将此问题转化为类似生日问题，即为“平均多少学生才能出现两人生日相同”。

设 k 人中（投球 k 次），指示器随机变量

$$
X_{ij} = I\{i,j\text{生日相同（在同一桶内）}\} = \begin{cases} 1, & i,j\text{生日相同（在同一桶内）} \\ 0, & \text{否则} \end{cases}
$$

则所有生日相同（在同一桶内）的对数 $X = \sum_{i=1}^{k} \sum_{j=i+1}^{k} X_{ij}$。

此时 $\mathbb{E}(X) = \sum_{i=1}^{k} \sum_{j=i+1}^{k} \mathbb{E}(X_{ij})$，而 $\mathbb{E}(X_{ij}) = b \times (1/b) \times (1/b) = 1/b$，所以

$$
\begin{aligned}
\mathbb{E}(X) &= \sum_{i=1}^{k} \sum_{j=i+1}^{k} \mathbb{E}(X_{ij}) \\
&= \sum_{i=1}^{k} \sum_{j=i+1}^{k} \frac{1}{b} \\
&= \frac{k(k-1)}{2b}
\end{aligned}
$$

令 $\mathbb{E}(X) = \frac{k(k-1)}{2b} = 1 (k \geq 1)$，解得 $k = \frac{1}{2}(1 + \sqrt{1 + 8b})$。