Homework 4

吴昊 2016K8009929015（本科生）

邮件：wuhao164@mails.ucas.ac.cn

1 compute the variance of Z_i

Z_i denotes the steps that it takes from $i-1$ boxes unempty to i boxes unempty. We can treat it as a **geometric distribution** with successful probability $p=\frac{n-i+1}{n}$.

$$\begin{aligned} Var(Z_i) &= \sum_{i=1}^{\infty}(i-\frac{1}{p})^2 p(1-p)^{i-1} \\ &= \sum_{i=1}^{\infty}(i^2-\frac{2i}{p}+\frac{1}{p^2})p(1-p)^{i-1} \end{aligned} \tag{1}$$

We already knew that

$$\sum_{i=1}^{\infty} ip(1-p)^{i-1}=\frac{1}{p} \tag{2}$$

So we only need to compute $\sum_{i=1}^{\infty} i^2p(1-p)^{i-1}$, denote this by S, we have

$$(1-p)S=\sum_{i=1}^{\infty}(i-1)^2p(1-p)^{i-1} \tag{3}$$

Let equation (2) minus (3)

$$S-(1-p)S=pS=\sum_{i=1}^{\infty}2ip(1-p)^{i-1}-\sum_{i=1}^{\infty}p(1-p)^{i-1}=\frac{2}{p}-1$$

Bring equation (3) to (1), (1) can be simplified as

$$Var(Z_i)=\frac{2}{p^2}-\frac{1}{p}-\frac{2}{p^2}+\frac{1}{p^2}=\frac{1-p}{p^2}=\frac{ni-n}{(n-i+1)^2}$$

2 E(number of empty boxes)

let $X_i = 1$ represent that the i-th box is empty, so $X = \sum_{i=1}^{n} X_i$. It is obvious that $X_1 = X_2 = ... = X_n$, we have $X = nX_i$

$$X_i = (\frac{n}{n-1})^n$$

so

$$X = n(\frac{n}{n-1})^n$$

3 deviation between X and E(X)

Applying the Chebyshev ineq, we have

$$Pr(|X - E(X)| > c) \leq \frac{Var(X)}{c^2}$$

Now we try to give an upper bound for Var(X)

$$\begin{aligned} Var(X) &= Var(\sum_{i=1}^{n} X_i) \\ &= \sum_{i=1}^{n} Var(X_i) + \sum_{i,j} Cor(X_i, X_j) \end{aligned}$$

$Var(X_i)$ can be computed by

$$Var(X_i) = E(X_i^2) - E(X_i)^2$$

As for $Cor(X_i, X_j)$, we try to prove it is less than 0.

$$Cor(X_i, X_j) = E(X_i, X_j) - E(X_i)E(X_j) = E(X_j)[E(X_i|X_j) - E(X_i)]$$

Obvious that $E(X_i|X_j) - E(X_j) < 0$, so $Cor(X_i, X_j) < 0$
We have

$$Var(X) < \sum_{i=1}^{n} Var(X_i) = E(X_i^2) - E(X_i)$$

Give the result that

$$Pr(|X - E(X)| > c) < \frac{n(\frac{n-1}{n})^n - n^2(\frac{n-1}{n})^{2n}}{c^2}$$