高级算法设计与分析 Lecture 4

授课时间: 2020 年 3 月 9 日　授课教师: 孙晓明　记录人: 张硕

1 Balls and Bins

问题设定：将 m 个同样的球, 依次独立随机地放入 n 个编号为 $1, 2, \ldots, n$ 的盒子中。设随机变量 $X_1, \ldots, X_n$ 表示对应盒子中的球数，$X_i (1 \leq i \leq n)$ 同分布，且 $\sum_{i=1}^{n} X_i = m$. 考虑随着 m 的变化，$\{X_i\}$ 以及相关的函数 $f(X_1, X_2, \ldots, X_n)$ 的分布情况（例如：$Y = \max\{X_i\}$）。

1.1 Birthday Paradox

问题设定：当人数达到多少时，以概率 $1 - \varepsilon$，会出现两个人生日相同。

生日悖论问题可以看作是球盒模型，此时同学看作球，日期看作盒子，即 m 取到多大时，以概率 $1 - \varepsilon$，存在某一个盒子中至少有两个球。在第一节课我们已经证明，当 $m \sim \Theta(\sqrt{n})$ 时，$\exists X_i \geq 2$，以概率 $1 - \varepsilon$. 记随机变量 $Y = \max\{X_1, \ldots, X_n\}$，即球数最多的盒子中的球数，有 $\Pr(Y \geq 2) > 1 - \varepsilon$。

1.2 Load Balancing

问题设定：每个任务，被随机的分配到一个服务器上，当任务数与服务器数相等时，最大负荷服务器的承载量。即球盒问题中 $m = n$ 时，记 $Y = \max\{X_1, \ldots, X_n\}$，$Y$ 的分布。

定理 1. *当 $m = n$ 时，$Y \sim \Theta(\frac{\ln n}{\ln \ln n})$，w.h.p., 即以下两个式子同时成立*

$$\Pr(Y = O(\frac{\ln n}{\ln \ln n})) = 1 - o(1)$$
$$\Pr(Y = \Omega(\frac{\ln n}{\ln \ln n})) = 1 - o(1)$$

证明　下证

$$\Pr(Y \leq \frac{4 \ln n}{\ln \ln n}) = 1 - o(1)$$
$$\Pr(Y \geq \frac{\ln n}{4 \ln \ln n}) = 1 - o(1)$$

不妨令 $\frac{\ln n}{\ln \ln n} = t$ ，

1) 有 $\Pr(Y \leq 4t) = 1 - \Pr(Y > 4t) = 1 - \Pr(\max X_i > 4t)$ 。

先考虑对于第一个盒子，$X_1 > 4t$ 的概率。

$$
\begin{aligned}
\Pr(X_1 > 4t) &= \Pr(\bigcup_{1\le j_1<\ldots<j_{4t+1}\le n} \#j_1,...,j_n \text{ balls } \to \#1 \text{ bin}) \\
&\le \sum_{1\le j_1<\ldots<j_{4t+1}\le n} \Pr(\#j_1,...,j_{4t+1} \text{ balls} \to \#1 \text{ bin}) \qquad (1.1)\\
&= \binom{n}{4t+1}\cdot(\frac{1}{n})^{4t+1} \\
&\le (\frac{ne}{4t+1})^{4t+1}\cdot(\frac{1}{n})^{4t+1} \qquad (1.2)\\
&< (\frac{1}{t})^{4t+1} = (\frac{\ln\ln n}{\ln n})^{4t+1} \\
&< (\frac{\sqrt{\ln n}}{\ln n})^{4t+1} = (\frac{1}{\ln n})^{\frac{4t+1}{2}} \\
&< (\ln n)^{-2\frac{\ln n}{\ln\ln n}} = (e^{\ln\ln n})^{-2\frac{\ln n}{\ln\ln n}} = \frac{1}{n^2}
\end{aligned}
$$

其中，不等式 (1.1) 是由于 Union Bound：$\Pr(A\bigcup B) = \Pr(A) + \Pr(B) - \Pr(A\bigcap B) \le \Pr(A) + \Pr(B)$. 不等式 (1.2) 是由组合数的估计，$(\frac{n}{m})^m \le \binom{n}{m} = \frac{n(n-1)\ldots(n-m+1)}{m(m-1)\ldots1} \le (\frac{ne}{m})^m$

再考虑对于球数最多的盒子，$\max X_1 > 4t$ 的概率。

$$
\begin{aligned}
\Pr(\max X_i > 4t) &= \Pr((X_1 > 4t)\cup(X_2 > 4t)\cup\ldots\cup(X_n > 4t)) \\
&\le \sum_{i=1}^{n}\Pr(X_i > 4t) < n\cdot\frac{1}{n^2} = \frac{1}{n}
\end{aligned}
$$

因此 $\Pr(Y \le 4(\frac{\ln n}{\ln\ln n})) = 1 - \Pr(Y > 4t) > 1 - \frac{1}{n} = 1 - o(1)$.

2) 对于 $\Pr(Y \ge \frac{1}{4}t) = 1 - o(1)$ 也先考虑对于第一个盒子，$X_1 \ge \frac{1}{4}t$ 的概率。

$$
\begin{aligned}
\Pr(X_1 \ge \frac{1}{4}t) &= \sum_{k\ge\frac{1}{4}t}\Pr(X_1 = k) \\
&\ge \Pr(X_1 = \frac{1}{4}t) = \binom{n}{\frac{t}{4}}\cdot(\frac{1}{n})^{\frac{t}{4}}\cdot(1-\frac{1}{n})^{n-\frac{t}{4}} \\
&\ge (\frac{n}{\frac{t}{4}})^{\frac{t}{4}}\cdot(\frac{1}{n})^{\frac{t}{4}}\cdot e^{-1} = e^{-1}\cdot(\frac{4\ln\ln n}{\ln n})^{\frac{\ln n}{4\ln\ln n}} \qquad (1.3)\\
&\ge e^{-1}\cdot(\frac{1}{\ln n})^{\frac{\ln n}{4\ln\ln n}} = e^{-1}\cdot(e^{\ln\ln n})^{-\frac{\ln n}{4\ln\ln n}} = e^{-1}n^{-1/4} \qquad (1.4)\\
&\ge n^{-1/3} \qquad (1.5)
\end{aligned}
$$

其中，(1.3) 第一个不等号，由 $(1-\frac{1}{n})^{n-\frac{t}{4}} \ge (1-\frac{1}{n})^n \sim e^{-1}$. 当 n 足够大时，$4\ln\ln n \ge 1$ ，且 $n^{\frac{1}{4}} \ll n^{\frac{1}{3}}$ ，式 (1.4) 和 (1.5) 中不等号成立。

引入 0-1 随机变量 Z_i，

$$
Z_i = \begin{cases} 1 & \text{if } X_i \ge \frac{t}{4} \\ 0 & \text{otherwise} \end{cases}
$$

等价的，

$$
\Pr(Z_i = 1) = \Pr(X_i \ge \frac{1}{4}t) \ge n^{-\frac{1}{3}}
$$

代入 $\Pr(Y \geq \frac{1}{4}t)$.

$$
\begin{aligned}
\Pr(Y \geq \tfrac{1}{4}t) = \Pr(\max X_i \geq \tfrac{1}{4}t) &= \Pr\left(\bigcup_{i=1}^{n}(X_i \geq \tfrac{1}{4}t)\right) \\
&= \Pr\left(\bigcup_{i=1}^{n}(Z_i = 1)\right) \\
&= 1 - \Pr\left(\bigcap_{i=1}^{n}(Z_i = 0)\right) \\
&= 1 - \Pr\left(\sum_{i=1}^{n} Z_i = 0\right)
\end{aligned}
$$

定义 $Z = Z_1 + Z_2 + \ldots + Z_n$，由 Chebyshev 不等式

$$
\begin{aligned}
\Pr(Z=0) &= \Pr(Z - \mathbb{E}(Z) = -\mathbb{E}(Z)) \\
&\leq \Pr(|Z - \mathbb{E}(Z)| = |\mathbb{E}(Z)|) \\
&\leq \frac{Var(Z)}{\mathbb{E}(Z)^2} \\
&\leq \frac{n \cdot Var(Z_1)}{n^{4/3}} \qquad (1.6) \\
&\leq \frac{n \cdot \frac{1}{4}}{n^{4/3}} = \frac{1}{4n^{1/3}} = o(1) \qquad (1.7)
\end{aligned}
$$

其中，不等式 (1.6)，由期望的线性性，$\mathbb{E}(Z) = n\mathbb{E}(Z_1) \geq n \cdot n^{-\frac{1}{3}} = n^{\frac{2}{3}}$ ；以及由 Z_i 之间负相关，$Var(X+Y) = Var(X) + Var(Y) + 2\operatorname{cov}(X,Y) \leq Var(X) + Var(Y)$ 。不等式 (1.7)，由 0-1 随机变量的性质，$Var(Z_i) \leq \frac{1}{4}$。

证得，$\Pr(Y = \Omega(\frac{\ln n}{\ln \ln n})) = 1 - o(1)$.

结合 1)、2) ，原命题得证。 □

Two-Choice Load Balancing：如果不是直接随机投放，而是随机选择 2 个（常数个同理）盒子，询问当前盒子内的球数,然后投放到比较少的一个盒子内,那么当 $m = \Theta(n)$ 时,$Y \sim \Theta(\ln \ln n)$,w.h.p.

1.3 Coupon Collector

问题设定：当 m 继续增大，达到什么量级时，以高概率，每个盒子中都有球，即 $\min X_i > 0$

记随机变量 Y_k 为占据 k 个盒子，所需最少的球数，$k = 1, 2, ..., n$。显然有，$Y_1 = 1$，Y_n 即为所求，记 $Y = Y_n$，不妨设 $Y_0 = 0$。

记随机变量 $Z_k = Y_k - Y_{k-1}$，表示在 $k-1$ 个盒子非空的基础上，还需要多少次投球，才能投入一个空盒子，$k = 1, 2, ..., n$。此时投入空盒子的概率为 $p_k = 1 - \frac{k-1}{n}$ ，Z_k 服从几何分布，Z_k 的期望

$\mathbb{E}(Z_k) = \frac{1}{p_k}$ ，方差 $Var(Z_k) = \frac{1-p_k}{{p_k}^2}$。$Z_k$ 之间相互独立，取值只与 p_k 有关。

$$
\begin{aligned}
\mathbb{E}(Y) &= \mathbb{E}((Y_n - Y_{n-1}) + (Y_n - Y_{n-1}) + ... + (Y_2 - Y_1) + Y_1) \\
&= \mathbb{E}(Y_n - Y_{n-1}) + \mathbb{E}(Y_{n-1} - Y_{n-2}) + ... + \mathbb{E}(Y_2 - Y_1) + \mathbb{E}(Y_1 - Y_0) \\
&= \sum_{k=1}^{n} \frac{1}{p_k} = \sum_{k=1}^{n} \frac{1}{1 - \frac{k-1}{n}} = \sum_{k=1}^{n} \frac{n}{n-k+1} = n \cdot \sum_{k=1}^{n} \frac{1}{k} \\
&\sim n \cdot (\ln n + \mathrm{O}(1)) \sim n \cdot \ln n + \mathrm{O}(n)
\end{aligned}
$$

Y 的期望 $\mathbb{E}(Y) \sim n \cdot \ln n + \mathrm{O}(n)$ ，但由 Load balancing 知，只有期望不足以描述 Y 的分布，我们需要讨论 $Var(Y)$ ，如果 $\sqrt{Var(Y)}$ 在量级上比 $n \ln n$ 小，则我们可以说 Y 的主项是 $n \ln n$。

$$
\begin{aligned}
Var(Y) &= Var((Y_n - Y_{n-1}) + (Y_{n-1} - Y_{n-2}) + ... + (Y_1 - Y_0)) \\
&= Var(Z_n + Z_{n-1} + ... + Z_1) \\
&= \sum_{i=1}^{n} Var(Z_i) = \sum_{i=1}^{n} \frac{1-p_i}{{p_i}^2} = \sum_{i=1}^{n} (\frac{1}{{p_i}^2} - \frac{p_i}{{p_i}^2}) \qquad (\text{第一个} = \text{是因为}Z_i\text{独立}) \\
&= n^2 \sum_{i=1}^{n} \frac{1}{i^2} - n \sum_{i=1}^{n} \frac{1}{i} \\
&\sim \frac{\pi}{6} \cdot n^2 - n \ln n
\end{aligned}
$$

得到 $\mathbb{E}(Y) \sim n \cdot \ln n + \mathrm{O}(n)$，$Var(Y) = \frac{\pi}{6} \cdot n^2 + o(n^2)$，代入 Chebyshev 不等式 $\Pr(|X - \mathbb{E}(X)| \geq c) \leq \frac{Var(X)}{c^2}$。取 $c = \Theta(n)$，有 $\Pr(|Y - n \ln n| \geq K \cdot n) \leq \varepsilon$，其中 K 是一个常数，也就是说，以概率 1-ε 有，$Y \sim (n \ln n - K \cdot n, n \ln n + K \cdot n)$ 。通过调整 c 的取值，可以使概率达到 $1 - o(1)$ ，例如取 $c = \Theta(n \ln \ln n)$。

2 Chernoff's Bound

在 Coupon Collector 问题中，要想得到 $\min X_i > 0$，需要 $m \sim \Theta(n \ln n)$。事实上，当我们取一个更大的常数，仍然在 $m \sim \Theta(n \ln n)$，w.h.p，$\min X_i, \max X_i \sim \Theta(\ln n)$ 。换言之，球数开始趋于平均数。

为了证明这个结论，我们引入新的工具，Chernoff's Bound。

定理 2. *(Chernoff's Bound)* $X_1, \ldots, X_n$ 是独立的 *0-1* 随机变量，有 $\Pr(X_i = 1) = p_i, \Pr(X_i = 0) = 1 - p_i$ 。记 $X = X_1 + \cdots + X_n$ ，$\mathbb{E}(X) = p_1 + \cdots + p_n = \mu$ ，则有

$$
\Pr(X \geq (1+\delta)\mu) \leq \left[\frac{e^{\delta}}{(1+\delta)^{1+\delta}} \right]^{\mu}, \forall\ \delta > 0 \tag{2.1}
$$

$$
\Pr(X \leq (1-\delta)\mu) \leq \left[\frac{e^{-\delta}}{(1-\delta)^{1-\delta}} \right]^{\mu}, 0 < \delta < 1 \tag{2.2}
$$

证明 由对称性，只证 $\Pr(X \geq (1+\delta)\mu) \leq [\frac{e^{\delta}}{(1+\delta)^{1+\delta}}]^{\mu}$.

$$
\begin{aligned}
\Pr(X \geq (1+\delta)\mu) &= \Pr(e^{\lambda X} \geq e^{\lambda(1+\delta)\mu}) \\
&\leq \frac{\mathbb{E}(e^{\lambda X})}{e^{\lambda(1+\delta)\mu}} = \frac{\mathbb{E}(e^{\lambda(X_1+\ldots+X_n)})}{e^{\lambda(1+\delta)\mu}} \quad (2.3)\\
&= \frac{\mathbb{E}(e^{\lambda X_1}) \cdot \mathbb{E}(e^{\lambda X_2}) \cdot \ldots \cdot \mathbb{E}(e^{\lambda X_n})}{e^{\lambda(1+\delta)\mu}} \quad (2.4)\\
&= \frac{[(1-p_1)+p_1e^{\lambda}] \cdot [(1-p_2)+p_2e^{\lambda}] \cdot \ldots \cdot [(1-p_n)+p_ne^{\lambda}]}{e^{\lambda(1+\delta)\mu}} \\
&= \frac{[1+p_1(e^{\lambda}-1)] \cdot [1+p_2(e^{\lambda}-1)] \cdot \ldots \cdot [1+p_n(e^{\lambda}-1)]}{e^{\lambda(1+\delta)\mu}} \\
&\leq \frac{e^{p_1(e^{\lambda}-1)} \cdot e^{p_2(e^{\lambda}-1)} \cdot \ldots \cdot e^{p_n(e^{\lambda}-1)}}{e^{\lambda(1+\delta)\mu}} \quad (2.5)\\
&= \frac{e^{(e^{\lambda}-1)(p_1+\ldots+p_n)}}{e^{\lambda(1+\delta)\mu}} = \left[\frac{e^{e^{\lambda}-1}}{e^{\lambda(1+\delta)}}\right]^{\mu}
\end{aligned}
$$

其中，式 (2.3) 中不等号由 Markov 不等式得到；式 (2.4) 由 X_i 相互独立得到；不等式 (2.5) 通过放缩 $1+x \leq e^x$. 可以通过调整 λ 的取值，来控制不等式的上界，求导得 $\lambda = \ln(1+\delta)$ 时，取到最小值。此时有 $\Pr(X \geq (1+\delta)\mu) \leq [\frac{e^{1+\delta-1}}{(1+\delta)^{1+\delta}}]^{\mu} = [\frac{e^{\delta}}{(1+\delta)^{1+\delta}}]^{\mu}$.

式 (2.2) 证明方法类似，此处略去。 □

进一步地，当 $0 < \delta < 1$ 时，我们来估计指数的底，对指数式取对数，

$$
\begin{aligned}
\ln \frac{e^{\delta}}{(1+\delta)^{1+\delta}} &= \delta - (1+\delta)\ln(1+\delta) \\
&= \delta - (1+\delta)\left(\delta - \frac{1}{2}\delta^2 + \frac{1}{3}\delta^3 \cdots\right) \\
&= \delta - \left(\delta + \frac{1}{2}\delta^2 - \frac{1}{6}\delta^3 + \frac{1}{12}\delta^4 \cdots\right) \\
&= -\frac{1}{2}\delta^2 + \frac{1}{6}\delta^3 - \frac{1}{12}\delta^4 \cdots \\
&\leq -\frac{1}{2}\delta^2 + \frac{1}{6}\delta^3 \\
&\leq -\frac{1}{2}\delta^2 + \frac{1}{6}\delta^2 \\
&= -\frac{1}{3}\delta^2
\end{aligned}
$$

式 (2.2) 类似可证，故当 $0 < \delta < 1$ 时，有

$$
\begin{aligned}
\Pr(X \geq (1+\delta)\mu) &\leq \left[\frac{e^{\delta}}{(1+\delta)^{1+\delta}}\right]^{\mu} \leq e^{-\frac{\delta^2\mu}{3}} \\
\Pr(X \leq (1-\delta)\mu) &\leq \left[\frac{e^{-\delta}}{(1-\delta)^{1-\delta}}\right]^{\mu} \leq e^{-\frac{\delta^2\mu}{2}}
\end{aligned}
$$

补充：对于 (2.1) 式，下面给出一个统一的估计，需要更细致的放缩，

$$
\begin{aligned}
\ln \frac{e^\delta}{(1+\delta)^{1+\delta}} &= \delta - (1+\delta)\ln(1+\delta) \\
&= \delta - (1+\delta)\left(\delta - \frac{1}{2}\delta^2 + \frac{1}{3}\delta^3 - \cdots\right) \\
&= \delta - \left(\delta + \frac{1}{2}\delta^2 - \frac{1}{6}\delta^3 + \frac{1}{12}\delta^4 + \cdots\right) \\
&= -\frac{1}{2}\delta^2 + \frac{1}{6}\delta^3 - \frac{1}{12}\delta^4 - \cdots \\
&= -\frac{\delta^2}{1\times 2} + \frac{\delta^3}{2\times 3} - \frac{\delta^4}{3\times 4} + \frac{\delta^5}{4\times 5} - \frac{\delta^6}{5\times 6} + \cdots \\
&= -\frac{\delta^2}{4} - \left(\frac{1}{2}\cdot\frac{\delta^2}{1\times 2} - \frac{\delta^3}{2\times 3} + \frac{1}{2}\cdot\frac{\delta^4}{3\times 4}\right) - \left(\frac{1}{2}\cdot\frac{\delta^4}{3\times 4} - \frac{\delta^5}{4\times 5} + \frac{1}{2}\cdot\frac{\delta^6}{5\times 6}\right) - \cdots \\
&\le -\frac{\delta^2}{4}. \qquad (2.6)
\end{aligned}
$$

其中，(2.6) 是由均值不等式，

$$
\begin{aligned}
&\frac{1}{2}\cdot\frac{\delta^{2k}}{(2k-1)2k} - \frac{\delta^{2k+1}}{2k(2k+1)} + \frac{1}{2}\cdot\frac{\delta^{2k+2}}{(2k+1)(2k+2)} \\
&\ge \frac{1}{2}\cdot 2\sqrt{\frac{\delta^{2k}}{(2k-1)2k}\cdot\frac{\delta^{2k+2}}{(2k+1)(2k+2)}} - \frac{\delta^{2k+1}}{2k(2k+1)} \\
&= \delta^{2k+1}\left(\sqrt{\frac{1}{(2k-1)2k(2k+1)(2k+2)}} - \frac{1}{2k(2k+1)}\right) \\
&\ge 0.
\end{aligned}
$$

于是得到，

$$
\Pr(X \ge (1+\delta)\mu) \le \left[\frac{e^\delta}{(1+\delta)^{1+\delta}}\right]^\mu \le e^{-\frac{\delta^2\mu}{4}}.
$$

下面我们利用 Chernoff's Bound 来证明，当 $m \sim \Theta(n\ln n)$ 时，X_i 的分布。

定理 3. 当 $m > 8n\ln n$，有 $\min X_i, \max X_i \sim \Theta(\ln n)$，*w.h.p.*

证明 下证 $\Pr(\frac{m}{2n} \le \min X_i, \max X_i \le \frac{2m}{n}) = 1 - o(1)$. 由 Union Bound 得到，

$$
\begin{aligned}
\Pr\left(\frac{m}{2n} \le \min X_i, \max X_i \le \frac{2m}{n}\right) &= 1 - \Pr((\min X_i < \frac{m}{2n}) \cup (\max X_i > \frac{2m}{n})) \\
&\ge 1 - \Pr(\min X_i < \frac{m}{2n}) - \Pr(\max X_i > \frac{2m}{n})
\end{aligned}
$$

1) 先估计 $\Pr(\max X_i > \frac{2m}{n})$

$$
\Pr(\max X_i > \frac{2m}{n}) = \Pr((X_1 > \frac{2m}{n}) \cup ... \cup (X_n > \frac{2m}{n})) \le n\cdot\Pr(X_1 > \frac{2m}{n}) \qquad (2.7)
$$

其中，不等号由 Union Bound 得到。现在将 X_1 改写为 0-1 随机变量，定义 $Z_1, \ldots, Z_m$ 表示第 i 个球是否投放进第一个盒子中，

$$
Z_i = \begin{cases} 1 & \text{if Ball\#i} \to \text{Bin\#1} \\ 0 & \text{otherwise} \end{cases}
$$

有 $X_1 = Z_1 + \ldots + Z_m$，$\mathbb{E}(X_1) = \mathbb{E}(Z_1 + \ldots + Z_m) = \sum_{i=1}^{m} \mathbb{E}(Z_i) = \frac{m}{n}$。
由 Chernoff's Bound 得，

$$\Pr(X_1 > \frac{2m}{n}) = \Pr(X_1 > (1+1) \cdot \mathbb{E}(X_1)) \leq [\frac{e^1}{(1+1)^2}]^{\frac{m}{n}} = (\frac{e}{4})^{8\ln n} \leq (\frac{1}{e^2})^{\ln n} = n^{-2}$$

代入式 (2.7)，

$$\Pr(\max X_i > \frac{2m}{n}) \leq n \cdot \Pr(X_1 > \frac{2m}{n}) \leq \frac{1}{n} = o(1)$$

2) 再估计 $\Pr(\min X_i < \frac{m}{2n})$

$$\begin{aligned}
\Pr(\min X_i < \frac{m}{2n}) &= \Pr((X_1 < \frac{m}{2n}) \cup (X_2 < \frac{m}{2n}) \cup \cdots \cup (X_n < \frac{m}{2n})) \\
&\leq n \cdot \Pr(X_1 < \frac{m}{2n}) \\
&= n \cdot \Pr(X_1 < (1 - \frac{1}{2}) \cdot \mathbb{E}(X_1)) \\
&\leq n \cdot \left[\frac{e^{-\frac{1}{2}}}{(1-\frac{1}{2})^{\frac{1}{2}}} \right]^{\frac{m}{n}} = n \cdot (\frac{2}{e})^{\frac{m}{2n}} \\
&\leq n \cdot (\frac{2}{e})^{4\ln n} \leq n \cdot 0.3^{\ln n} \leq n^{1+\ln 0.3} < n^{-0.2} \\
&= o(1)
\end{aligned}$$

结合 1)、2)，有 $\Pr(\frac{m}{2n} \leq \min X_i, \max X_i \leq \frac{2m}{n}) = 1 - o(1)$. □

综上，随着 m 的增加，球盒模型中球的分布呈以下规律：

- (Birthday Paradox) $m \sim \Theta(\sqrt{n})$，$\Pr(\max X_i > 1) \geq 1 - \varepsilon$.
- (Load Balancing) $m = n$，$\max X_i \sim \Theta(\frac{\ln n}{\ln \ln n})$，w.h.p.
- (Coupon Collector) $m \in (n \ln n - C \cdot n, n \ln n + C \cdot n)$，$\Pr(\min X_i > 0) \geq 1 - \varepsilon$.
- $m \sim Cn \ln n$ for large C，$\min X_i, \max X_i \sim \Theta(\ln n)$，w.h.p.