高级算法作业三

1190200122 袁野

2022 年 5 月 10 日

3.1

向长度为 $2r$ 的正方形内随机投掷 n 个点，设其中有 x 个点落进圆内，令 $q = \frac{x}{n} \approx \frac{\pi r^2}{4r^2} = \frac{\pi}{4}$，我们用 $4q$ 来估计 π，由 $E[x] = \frac{n\pi}{4}$ 有

$$Pr[4q > \pi + \delta] = Pr[\frac{4x}{n} > \pi + \delta] = Pr[x > E[x](1 + \frac{\pi}{\delta})] \leq e^{-\frac{n\delta^2}{12\pi}}$$

$$Pr[4q < \pi - \delta] = Pr[\frac{4x}{n} < \pi - \delta] = Pr[x < E[x](1 - \frac{\pi}{\delta})] \leq e^{-\frac{n\delta^2}{8\pi}}$$

则有

$$\begin{aligned} Pr[4q \in [\pi - \delta, \pi + \delta]] &= Pr[x \in [E(x)(1 - \frac{\delta}{\pi})], E(x)(1 + \frac{\delta}{\pi})] \\ &\geq 1 - e^{-\frac{n\delta^2}{12\pi}} - e^{-\frac{n\delta^2}{8\pi}} \end{aligned}$$

所以置信区间为 $[\pi - \delta, \pi + \delta]$，置信度为 $1 - e^{-\frac{n\delta^2}{12\pi}} - e^{-\frac{n\delta^2}{8\pi}}$

3.2

由 $Pr[\frac{kn}{m} \in [(1-\lambda)|A|, (1+\lambda)|A|]] > 1 - \gamma$ 可以得到

$$Pr[\frac{kn}{m} < (1-\lambda)|A|] + Pr[\frac{kn}{m} > (1+\lambda)|A|] < \gamma$$

令 $X = kn$，则 $E[X] = |A| * m$ 上式即为

$$Pr[X < (1-\lambda)E[X]] + Pr[X > (1+\lambda)E[X]] < \gamma$$

根据切尔诺夫界可以得到

$$Pr[X < (1-\lambda)E[X]] \leq e^{-\frac{|A|m\lambda^2}{2}}$$

$$Pr[X > (1-\lambda)E[X]] \leq e^{-\frac{|A|m\lambda^2}{3}}$$

因此 m 需要满足

$$e^{-\frac{|A|m\lambda^2}{2}} + e^{-\frac{|A|m\lambda^2}{3}} \leq \gamma$$

3.3

(a) 要证任意一个由根到叶结点的路径，好结点的数量不超过 $c_1 log_2 n$，等价于证明好结点最多的路径，好结点的数量不超过 $c_1 log_2 n$。好结点数量最多时，所有的好结点划分都恰好为 $\frac{s}{3}$ 和 $\frac{2s}{3}$。设好结点的数量为 k，因此有

$$n * \frac{2}{3}^k \geq 1$$

由上式可得到

$$k \leq \frac{log_2 n}{log_2 \frac{3}{2}}$$

因此 $c_1 = \frac{1}{log_2 \frac{3}{2}}$

(b) 设 $X_i = 1$ 表示 i 是坏结点，$X_i = 0$ 表示 i 是好结点。

$$Pr[X_i = 0] = \frac{1}{3}$$

$$Pr[X_i = 1] = \frac{2}{3}$$

设当前路径上有 N 个结点，记 $X = \sum X_i$，则有 $E[X] = \frac{2}{3}N$，好结点的个数即为 $N - X$ 个。

由 (a) 知，

$$N - X \leq \frac{log_2 n}{log_2 \frac{3}{2}}$$

由此可得

$$E[N - X] = E[N] - E[X] = \frac{1}{3}E[N] le \frac{log_2 n}{log_2 \frac{3}{2}}$$

可得

$$E[N] \leq \frac{3\, log_2 n}{log_2 \frac{3}{2}}$$

则

$$Pr(N \geq \frac{3\, log_2 n}{log_2 \frac{3}{2}} * (1 + \delta)) \leq Pr(N \geq E[N] * (1 + \delta)) < e^{-\frac{\mu\delta^2}{3}}$$

令 $\delta = \sqrt{2\, ln\, \frac{3}{2}}$，由切尔诺夫界则有

$$Pr(N \geq \frac{3(1 + \sqrt{2\, ln\, \frac{3}{2}})}{log_2 \frac{3}{2}} * log_2 n) < e^{-\frac{2\mu\, ln\, \frac{3}{2}}{3}} = \frac{1}{n^2}$$

令 $c_2 = \frac{3(1+\sqrt{2\, ln\, \frac{3}{2}})}{log_2 \frac{3}{2}}$ 则

$$Pr(N \geq c_2\, log_2 n) < \frac{1}{n^2}$$

即 $Pr(N \leq c_2\, log_2 n) \geq 1 - \frac{1}{n^2}$，其中 $c_2 = \frac{3(1+\sqrt{2\, ln\, \frac{3}{2}})}{log_2 \frac{3}{2}}$。

(c) 设所有的叶结点到根结点的长度由小到大分别是 $N_1, N_2, \cdots, N_n$，那么有

$$\begin{aligned} Pr(N_n \le c_2\, log + 2\, n) &= 1 - Pr(N_n > c_2\, log + 2\, n) \\ &= 1 - Pr(\bigcup_{i=1}^{n} N_i > c_2\, log + 2\, n) \\ &\ge 1 - \sum_{i=1}^{n} Pr(N_i > c_2\, log + 2\, n) \\ &> 1 - \sum_{i=1}^{n} \frac{1}{n^2} \\ &= 1 - \frac{1}{n} \end{aligned}$$

(d) 设对 n 个元素进行快排的时间为 $T(n)$，那么对 $n > 1$ 我们有

$$T(n) = T(x) + T(n-x) + O(n)$$

由 (c) 可知树深度不超过 $c_2\, log + 2\, n$ 的概率为 $1 - \frac{1}{n}$，则 $Pr(T(n) = O(n\, log_2\, n)) > 1 - \frac{1}{n}$。

3.4

由题意知 $E[X_i] = \frac{X_{i-1}+1}{2}$

$$\begin{aligned} E[Y_i | Y_0, Y_1, Y_2, \cdots, Y_{i-1}] &= E[2^i(1 - X_i) | Y_0, Y_1, Y_2, \cdots, Y_{i-1}] \\ &= 2^i E[1 - X_i | Y_0, Y_1, Y_2, \cdots, Y_{i-1}] \\ &= 2^i (1 - \frac{X_{i-1}+1}{2}) \\ &= 2^{i-1}(1 - X_{i-1}) \\ &= Y_{i-1} \end{aligned}$$

所以 $Y_0, Y_1 \cdots$ 是一个鞅。

3.5

设 Y_i 为第 i 次进行步骤二的 $|x_1, \cdots, x_n / y_1, \cdots, y_{k-1}|$,$T$ 进行步骤一的次数,那么总复杂度为 $E[T] * E[Y_i]$。

$$E[T] = \sum_{i=1}^{\infty} i * \frac{1}{n!} * (\frac{n! - 1}{n!})^{i-1} = n!$$

$$E[Y_i] = \frac{1}{n!} \sum_{i=1}^{n} C_n^i (n-i)! < e$$

总复杂度为 $O(n!)$

3.6

设 S_i 为 w_i 的前缀和，即 $S_i = \sum_{j=1}^{i} w_j$。我们可以生成一个 $[1, S_n]$ 的随机数 x，通过二分法寻找 i 满足 $S_i < x \le S_{i+1}$，显然 i 是必然存在的，那么我们选取 o_i 作为选取对象，这样的话概率即为 $\frac{w_i}{\sum w_i}$。二分的时间复杂度为 $O(log\, n)$，那么选取的过程的复杂度即为 $O(log\, n)$。