Randomized algorithm

Tutorial 5

Joyce

2009-12-08

Solution to assignment 3

[Question 1]: Let X be a Poisson random variable with mean λ.

1. What is the most likely value of X
   1.1 when λ is an integer?
   1.2 when λ is not an integer?

   Hint: Compare $\Pr(X = k + 1)$ with $\Pr(X = k)$.
2. We define the median of X to be the least number m such that $\Pr(X \leq m) \geq 1/2$. What is the median of X when $\lambda = 3.9$?

[Solution]: X is a Poisson random variable, therefore we know that

$$Pr(X = k) = \frac{e^{-\lambda}\lambda^k}{k!}$$

$$Pr(X = k+1) = \frac{e^{-\lambda}\lambda^{k+1}}{(k+1)!}$$

Compare $Pr(X = k)$ and $Pr(X = k+1)$

$$\frac{Pr(X = k+1)}{Pr(X = k)} = \frac{\frac{e^{-\lambda}\lambda^{k+1}}{(k+1)!}}{\frac{e^{-\lambda}\lambda^k}{k!}} = \frac{\lambda}{k+1}$$

When λ is an integer, both λ and $\lambda - 1$ are the most likely values.
When λ is not an integer, $X = \lfloor \lambda \rfloor$ is the most likely value.

[Solution]: By knowing that $Pr(X = k) = \frac{e^{-\lambda}\lambda^k}{k!}$ and $\frac{Pr(X=k+1)}{Pr(X=k)} = \frac{\lambda}{k+1}$, we may calculate $Pr(X = i)$ directly.

$$\begin{aligned}
\Pr(X = 0) &= \frac{e^{-3.9} * -3.9^0}{0!} = 0.0202419114 \\
\Pr(X = 1) &= \Pr(X = 0)\frac{3.9}{1} = 0.0789434548 \\
\Pr(X = 2) &= \Pr(X = 1)\frac{3.9}{2} = 0.153939737 \\
\Pr(X = 3) &= \Pr(X = 2)\frac{3.9}{3} = 0.200121658 \\
\Pr(X = 4) &= \Pr(X = 3)\frac{3.9}{4} = 0.195118617
\end{aligned}$$

$$\Sigma_{i=0}^{4} Pr(X = i) = 0.6483653782 > 0.5$$

[Question 2]: Let X be a Poisson random variable with mean μ, representing the number of criminals in a city. There are two types of criminals: For the first type, they are not too bad and are reformable. For the second type, they are flagrant. Suppose each criminal is independently reformable with probability p (so that flagrant with probability $1 - p$). Let Y and Z be random variables denoting the number of reformable criminals and flagrant criminals (respectively) in the city. Show that Y and Z are independent Poisson random variables.

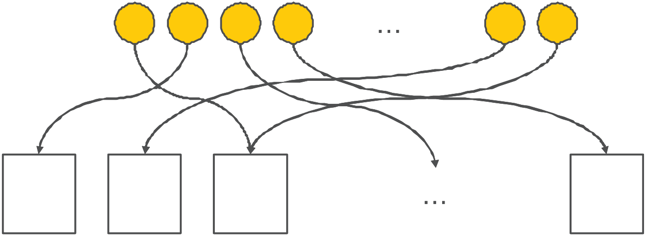
[Solution]: Suppose we have $Y = k$.

$$\begin{aligned}
\Pr(Y=k) &= \sum_{m=k}^{\infty} Pr(Y=k|X=m)Pr(X=m) \\
&= \sum_{m=k}^{\infty} \binom{m}{k} p^k (1-p)^{m-k} \frac{e^{-\lambda}\lambda^m}{m!} \\
&= \sum_{m=k}^{\infty} \frac{m!}{k!(m-k)!} p^k (1-p)^{m-k} \frac{e^{-\lambda}\lambda^m}{m!} \\
&= \frac{(\lambda p)^k e^{-\lambda}}{k!} \sum_{m=k}^{\infty} \frac{(1-p)^{m-k}\lambda^{m-k}}{(m-k)!} \\
&= \frac{(\lambda p)^k e^{-\lambda}}{k!} e^{\lambda(1-p)} \\
&= \frac{(\lambda p)^k e^{-p\lambda}}{k!}
\end{aligned}$$

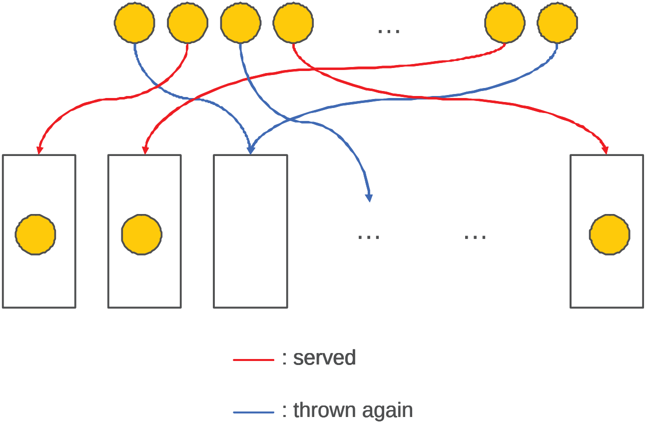
[Solution]: We may also get $Pr(Z = k) = \lambda(1-p)^k e^{-(1-p)\lambda}/k!$ by the same way. Now we are going to show that Y and Z are independent.

$$
\begin{aligned}
\Pr(Y = k \bigcap Z = j) &= \Pr(X = k + j)\Pr(Y = k | X = k + j) \\
&= \frac{\lambda^{k+j} e^{-\lambda}}{(k+j)!} \binom{k+j}{k} p^k (1-p)^j \\
&= \frac{(\lambda p)^k e^{-p\lambda}}{k!} \frac{(\lambda(1-p))^j e^{-(1-p)\lambda}}{j!} \\
&= \Pr(Y = k)\Pr(Z = j)
\end{aligned}
$$

[Question 3]: Consider assigning some balls to n bins as follows: In the first round, each ball chooses a bin independently and uniformly at random.



After that, if a ball lands at a bin by itself, the ball is *served* immediately, and will be removed from consideration. For the number of bins, it remains unchanged.



In the subsequent rounds, we repeat the process to assign the remaining balls to the bins. We finish when every ball is served.

1. Suppose at the start of some round b balls are still remaining. Let $f(b)$ denote the expected number of balls that will remain after this round. Given an explicit formula for $f(b)$.
2. Show that $f(b) \le b^2/n$.
   Hint: You may use *Bernoulli's inequality*:
   $$\forall r \in \mathbb{N} \text{ and } x \ge -1, \qquad (1+x)^r \ge 1 + rx.$$
3. Suppose we have $\frac{n}{k}$ balls initially, for some fixed constant $k > 1$. Every round the number of balls served was exactly the expected number of balls to be served. Show that all the balls would be served in $O(\log \log n)$ rounds.

[Solution]: (1)

$$Pr(\text{the } i\text{-th bin has exactly 1 ball}) = b\frac{1}{n}(1-\frac{1}{n})^{b-1}$$

$$\mathrm{E}[\text{number of bins have 1 ball}] = b(1-\frac{1}{n})^{b-1}$$

$$f(b) = b - \mathrm{E}[\text{number of bins have 1 ball}] = b(1-(1-\frac{1}{n})^{b-1})$$

[Solution]: (2) By Bernoulli's inequality: $(1-\frac{1}{n})^{b-1} \geq 1-\frac{b-1}{n}$.

$$
\begin{aligned}
f(b) &= b(1-(1-\frac{1}{n})^{b-1})) \\
&\leq b(1-(1-\frac{b-1}{n})) \\
&= \frac{b(b-1)}{n} \\
&\leq \frac{b^2}{n}
\end{aligned}
$$

[Solution]: (3) Consider we have n balls initially.

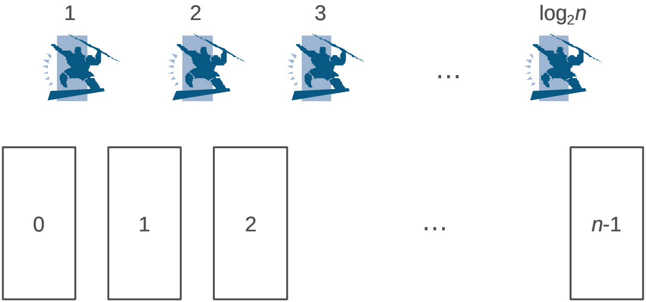
$$\lim_{n\to\infty} n(1-(1-\frac{1}{n})^{n-1}) = n(1-(1-\frac{1}{n-1})^{n}) = n(1-\frac{1}{e})$$

It is like the value of n/k while k is a small constant.

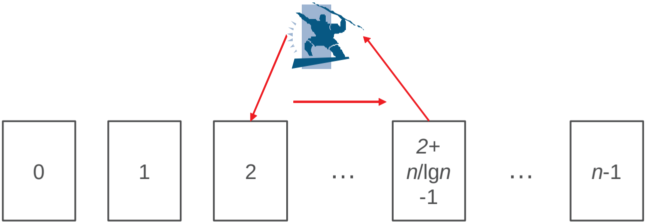
$$f(n/k) = n/k^2 \Rightarrow f^r(n/k) = n/k^{2^r}$$

Let $r = log_k log_2 n = O(loglogn)$, then $f^r(n/k) = 1$.

[Question 4]: Suppose that we vary the balls-and-bins process as follows. For convenience let the bins be numbered from 0 to $n-1$. There are $\log_2 n$ players.



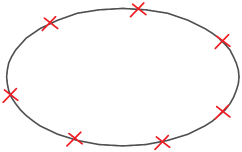
Each player chooses a starting location ℓ uniformly at random from $[0, n-1]$ and then places one ball in each of the bins numbered $\ell \bmod n, \ell+1 \bmod n, \ldots, \ell + n/\log_2 n - 1 \bmod n$.(Assume that n is a multiple of $\log_2 n$.)



Show that the maximum load in this case is only $O(\log\log n/\log\log\log n)$ with probability that approaches 1 as $n \longrightarrow \infty$.

[Solution]:

Since there are $log_2 n$ players to choose a starting location, we choose $\log n$ bins as representatives.



$$
\begin{aligned}
& \Pr(\text{some bins get } M \text{ balls}) \\
\leq\ & \Pr(\text{there exist 2 representatives with } M/2 \text{ balls inside}) \\
\leq\ & \log n\left(\frac{1}{(M/2)!}\right) \\
\leq\ & \log n\left(\frac{2e}{M}\right)^{M/2}
\end{aligned}
$$

Now we let $M = 6\log\log n/\log\log\log n$,

$$\begin{aligned}
& \log n(\frac{2e}{M})^{M/2} \\
\leq \quad & \log n(\frac{\log\log\log n}{\log\log n})^{3\log\log n/\log\log\log n} \\
\leq \quad & e^{\log\log n}(e^{\log\log\log\log n - \log\log\log n})^{3\log\log n/\log\log\log n} \\
\leq \quad & e^{-2\log\log n}e^{3\log\log n\log\log\log\log n/\log\log\log n} \\
\leq \quad & \frac{1}{\log n}
\end{aligned}$$

[Question 5]: We consider another way to obtain Chernoff-like bound in the balls-and-bins setting without using the theorem in Page 13 of Lecture 14.
Consider n balls thrown randomly into n bins. Let $X_i = 1$ if the i-th bin is empty and 0 otherwise. Let $X = \sum_{i=1}^{n} X_i$.
Let Y_i be independent Bernoulli random variable such that $Y_i = 1$ with probability $p = (1 - 1/n)^n$. Let $Y = \sum_{i=1}^{n} Y_i$.

1. Show that $\mathrm{E}[X_1 X_2 \cdots X_k] \leq \mathrm{E}[Y_1 Y_2 \cdots Y_k]$ for any $k \geq 1$.
2. Show that ${X_1}^{j_1} {X_2}^{j_2} \cdots {X_k}^{j_k} = X_1 X_2 \cdots X_k$ for any $j_1, j_2, \ldots, j_k \in \mathbb{N}$.
3. Show that $\mathrm{E}[e^{tX}] \leq \mathrm{E}[e^{tY}]$ for all $t \geq 0$.
   Hint: Use the expansion for e^x and compare $\mathrm{E}[e^{tX}]$ to $\mathrm{E}[e^{tY}]$.
4. Derive a Chernoff bound for $Pr(X \geq (1 + \delta)\mathrm{E}[X])$.

[Solution]:

(1) We may calculate the expectation directly.

$$\mathrm{E}[X_1X_2\cdots X_k] = Pr(X_1 = 1 \cap X_2 = 1 \cap \ldots \cap X_k = 1) = (\frac{n-k}{n})^n$$

$$\mathrm{E}[Y_1Y_2\cdots Y_k] = Pr(Y_1 = 1 \cap Y_2 = 1 \cap \ldots \cap Y_k = 1) = (1-\frac{1}{n})^{kn}$$

$$(\frac{n-k}{n})^n \leq (1-\frac{1}{n})^{kn}$$

The alternative solution(By Induction):

$$
\begin{aligned}
& \mathrm{E}[X_1 X_2 \cdots X_k X_{k+1}] \\
= \; & \mathrm{E}[X_1 X_2 \cdots X_k | X_{k+1} = 1] Pr(X_{k+1} = 1) \\
\leq \; & \mathrm{E}[X_1 X_2 \cdots X_k] Pr(X_{k+1} = 1) \\
\leq \; & (1 - \frac{1}{n})^{(k+1)n} \\
= \; & \mathrm{E}[Y_1 Y_2 \cdots Y_k Y_{k+1}]
\end{aligned}
$$

[Solution]:
(2) X_i is an indicator and $X_i = 1$ when X_i is empty.
If ${X_1}^{j_1}{X_2}^{j_2}\cdots{X_k}^{j_k} = 1$, $X_i = 1$ for all i.
Then $X_1X_2\cdots X_k = 1$
If ${X_1}^{j_1}{X_2}^{j_2}\cdots{X_k}^{j_k} = 0$, there exists $X_i = 0$ for at least one i.
Then $X_1X_2\cdots X_k = 0$

[Solution]:
(3)

$$\mathrm{E}[e^{tX}] = \mathrm{E}[1 + tX + \frac{(tX)^2}{2!} + ...] = \mathrm{E}[1] + t\mathrm{E}[X] + \frac{t^2}{2!}\mathrm{E}[X^2] + ...$$

By (a), we get $\mathrm{E}[X] \leq \mathrm{E}[Y]$.

$$\begin{aligned} E[X^2] &= \mathrm{E}[(X_1X_2 \cdots X_n)^2] \\ &= \mathrm{E}[(X_1^2 + 2(X_1X_2 + X_1X_3 + \cdots + X_1X_n) + X_2^2 + \cdots + X_n^2] \\ &\leq \mathrm{E}[(Y_1^2 + 2(Y_1Y_2 + Y_1Y_3 + \cdots + Y_1Y_n) + Y_2^2 + \cdots + Y_n^2] \\ &= E[Y^2] \end{aligned}$$

The remaining inequalities $\mathrm{E}[X^i] \leq \mathrm{E}[X^i]$ can be proved by the same way.
Thus, $\mathrm{E}[e^{tX}] \leq \mathrm{E}[e^{tY}]$.

[Solution]:

(4) First, we set $\mathrm{E}[X]$ as μ.

$$Pr(X \geq (1+\delta)\mathrm{E}[X]) = Pr(e^{tX} \geq e^{t(1+\delta)\mathrm{E}[X]}) \leq \frac{\mathrm{E}[e^{tX}]}{e^{t(1+\delta)\mu}}$$

$$\mathrm{E}[e^{tY}] = \prod \mathrm{E}[e^{tY_i}] = [(1-p)+pe^t]^n = [1+p(e^t-1)] \leq e^{np(e^t-1)}$$

Combine these two inequalities, we get

$$Pr(X \geq (1+\delta)\mathrm{E}[X]) \leq \frac{\mathrm{E}[e^{tX}]}{e^{t(1+\delta)\mu}} = \frac{e^{\mu(e^t-1)}}{e^{t\mu(1+\delta)}}$$

Now we set $t = \ln(1+\delta)$,

$$Pr(X \geq (1+\delta)\mathrm{E}[X]) \leq \frac{e^{\mu(e^t-1)}}{e^{t\mu(1+\delta)}} \leq \frac{e^{\mu\delta}}{(1+\delta)^{(1+\delta)\mu}}$$

Hint for assignment 4

[Question 1]: Let G be a random graph drawn from the $G_{n,1/2}$ model.

1. (10%) What is the expected number of 5-clique in G?
2. (10%) What is the expected number of 5-cycle in G?

1. No hint.
2. A 5-cycle is a subset of vertices with size 5.

[Question 2]: Suppose we have a set of n vectors, $v_1, v_2, \ldots, v_n$, in R^m. Each vector is of unit-length, i.e., $\|v_i\| = 1$ for all i. In this question, we want to show that, there exists a set of values, $\rho_1, \rho_2, \ldots, \rho_n$, each $\rho_i \in \{-1, +1\}$, such that

$$\|\rho_1 v_1 + \rho_2 v_2 + \cdots + \rho_n v_n\| \leq \sqrt{n}.$$

Intuitively, if we are allowed to "reflect" each v_i as we wish (i.e., by replacing v_i by $-v_i$), then it is possible that the vector formed by the sum of the n vectors is at most $\sqrt{n}$ long.

1. Let $V = \rho_1 v_1 + \rho_2 v_2 + \cdots + \rho_n v_n$, and recall that

$$\|V\|^2 = V \cdot V = \sum_{i,j} \rho_i \rho_j v_i \cdot v_j.$$

Suppose that each ρ_i is chosen uniformly at random to be -1 or $+1$. Show that

$$\mathrm{E}[\|V\|^2] = n.$$

2. (5%) Argue that there exists a choice of $\rho_1, \rho_2, \ldots, \rho_n$ such that $\|V\| \leq \sqrt{n}$.
3. (5%) Your friend, Peter, is more ambitious, and asks if it is possible to to choose $\rho_1, \rho_2, \ldots, \rho_n$ such that

$$\|V\| < \sqrt{n}$$

instead of $\|V\| \leq \sqrt{n}$ we have just shown. Give a counter-example why this may not be possible.

[Hint]:

- Can you show that $\mathrm{E}[\rho_i \rho_j] = 0$ when $i \neq j$?
- What is the value of $\mathrm{E}[\rho_i \rho_i]$?
- What is the value of $v_i \cdot v_i$?

[Question 3]: Let F be a family of subsets of $N = \{1, 2, ..., n\}$. F is called an antichain if there are no $A, B \in F$ satisfying $A \subset B$.
Ex:
$N = \{1, 2, 3\}$
All subsets of N :$\{\}, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}$
$\{\{1, 3\}, \{1, 2\}, \{2, 3\}\}$ is an antichain.
$\{\{1, 2, 3\}, \{1, 2\}\}$ is not an antichain. ($\{1, 2\} \subset \{1, 2, 3\}$)

We want to prove that $|F| \leq \binom{n}{\lfloor n/2 \rfloor}$. How?
Let $\sigma \in S_n$ be a random permutation of the elements of N and consider the random variable

$$X = |\{i : \{\sigma(1), \sigma(2), ...\sigma(i)\} \in F\}|$$

What is X?

Let us set $F = \{\{2\}, \{1, 3\}\}$.
For $\sigma = \{2, 3, 1\}$, $X = 1$. $\because \{2\} \in F$.
For $\sigma = \{3, 2, 1\}$, $X = 0$. $\because$ None of $\{3\}$, $\{3, 2\}$,$\{3, 2, 1\}$ are in F.
For $\sigma = \{3, 1, 2\}$, $X = 1$. $\because \{1, 3\} \in F$.
Try to find the relation between X and F.

[Hint]: You may also try this. F can be partitioned into n parts according to the size of subsets, like k_i denotes a size-i set. Therefore, $|F| = k_1 + k_2 + ... + k_n$.

[Question 4]: Consider a graph in $G_{n,p}$, with $p = 1/n$. Let X be the number of triangles in the graph, where a triangle is a clique with three edges. Show that

$$Pr(X \geq 1) \leq 1/6$$

and that

$$\lim_{n \to \infty} Pr(X \geq 1) \geq 1/7$$

[Hint]: Conditional expectation inequality.

Thank you