Randomized algorithm

Tutorial 3

Joyce

2008-11-13

Solution for assignment 2

[Question 1]:
Let S be a set of n numbers. The median-finding algorithm discussed in class finds the median of S with high probability, and its running time is $2n + o(n)$.
Can you generalize this algorithm so that it can find the kth largest item of S for any given value of k?
Prove that your resulting algorithm is correct, and bound its running time.

[Solution]:
We first consider the case where $k \leq n/2$.
For this case, we scan the n items and obtain the minimum value, say, t. Then, we add $n - 2k + 1$ items to the set of n items, each item having a value equal to t. It is easy to check that the kth largest item among the original n items will be the *median* of the new set of $2n - 2k + 1$ items.

[Cont.]: Therefore, we can apply the median-finding algorithm for the new set to obtain the desired kth largest item of the original set.
The running time for this case is $O(n)$. More precisely, we have spent n comparisons to find the minimum element, and at most $4n + o(n)$ further comparisons to find the median. The total time is thus at most $5n + o(n)$.
Could it be lower?
In fact, it is easy to see that we do not need to explicitly add the $n - 2k + 1$ items.

[Cont.]: Therefore, we can apply the median-finding algorithm for the new set to obtain the desired kth largest item of the original set.

The running time for this case is $O(n)$. More precisely, we have spent n comparisons to find the minimum element, and at most $4n + o(n)$ further comparisons to find the median. The total time is thus at most $5n + o(n)$.

Could it be lower?

In fact, it is easy to see that we do not need to explicitly add the $n - 2k + 1$ items.

[Cont.]: What we need to do is to modify the sampling algorithm a bit, so that with probability $p = n/(2n - 2k + 1)$, we are selecting an integer from the original n items, and with probability $1 - p$, we are selecting the minimum value t. Consequently, the time will be at most $3n + o(n)$.

For the case where $k > n/2$, we proceed by adding items with maximum value instead.

[Question 2]:
The weak law of large numbers state that, if $X_1, X_2, X_3, \ldots$ are independent and identically distributed random variables with finite mean μ and finite standard deviation σ,
then for any constant $\varepsilon > 0$ we have

$$\lim_{n\to\infty} \Pr\left(\left|\frac{X_1 + X_2 + X_3 + \ldots + X_n}{n} - \mu\right| > \varepsilon\right) = 0$$

Use Chebyshev's inequality to prove the weak law of large numbers.

[Solution]:
Firstly,

$$\Pr\left(\left|\frac{X_1 + X_2 + X_3 + ... + X_n}{n} - \mu\right| > \varepsilon\right)$$

$$= \Pr(|(X_1 + X_2 + ... + X_n) - n\mu| > n\varepsilon).$$

By Chebyshev's inequality and the independence of X_i's, we have:

$$\Pr(|(X_1 + X_2 + ... + X_n) - n\mu| > n\varepsilon)$$
$$\leq \frac{\mathrm{Var}[X_1 + X_2 + ... + X_n]}{(n\varepsilon)^2}$$

[Cont.]:

$$\frac{\mathrm{Var}[X_1 + X_2 + \ldots + X_n]}{(n\varepsilon)^2} = \frac{\sum_{i=1}^{n} \mathrm{Var}[X_i]}{n^2\varepsilon^2} = \frac{n\sigma^2}{n^2\varepsilon^2} = \frac{\sigma^2}{n\varepsilon^2}.$$

Combining, we have

$$\lim_{n\to\infty} \Pr\left(\left|\frac{X_1 + X_2 + X_3 + \ldots + X_n}{n}\right| - \mu > \varepsilon\right)$$

$$\leq \lim_{n\to\infty} \frac{\sigma^2}{n\varepsilon^2} = 0$$

which completes the proof of the weak law of large numbers.

[Question 3]:

1. Determine the moment generating function for the binomial random variable $Bin(n, p)$.
2. Let X be a $Bin(n, p)$ random variable and Y be a $Bin(m, p)$ random variable. Suppose that X and Y are independent. Use part (a) to determine the moment generating function of $X + Y$.
3. What can we conclude from the form of the moment generating function of $X + Y$?

[Solution]:

(a) Given $X \sim Bin(n, p)$. Then, the MGF of X, $M_X(t)$, can be calculated as follows:

$$
\begin{aligned}
M_X(t) &= \mathrm{E}[e^{tX}] \\
&= \Sigma_{k=0}^{n} \Pr(X = k) e^{tk} \\
&= \Sigma_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} e^{tk} \\
&= \Sigma_{k=0}^{n} \binom{n}{k} (pe^t)^k (1-p)^{n-k} \text{ (Binomial theorem)} \\
&= (pe^t + 1 - p)^n
\end{aligned}
$$

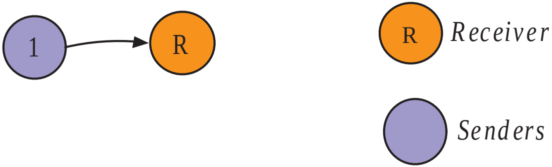
(Binomial theorem: $\Sigma_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} = (p + (1-p))^n$.)

[Cont.]:
(b) Since X and Y are independent, the MGF of $X + Y$ is thus:

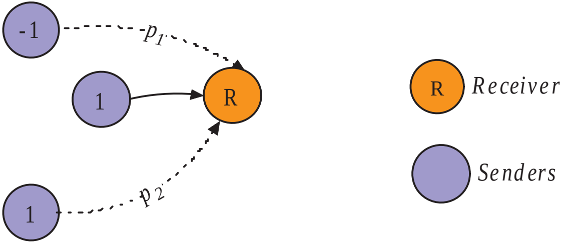
$$M_{X+Y}(t) = M_X(t) \times M_Y(t) = (pe^t + 1 - p)^{n+m}.$$

[Cont.]:
(c) The MGF in part (b) is the same as the MGF of a binomial random variable with parameters $n + m$ and p. Thus, $X + Y \sim \mathrm{Bin}(n + m, p)$.

[Question 4]:
In a wireless communication system, each receiver listens on a specific frequency. The bit $b(t)$ sent at time t is represented by a 1 or -1.



Unfortunately, noise from other nearby communications can affect the receiver's signal.



There are n senders and the ith has strength $p_i \leq 1$ The receiver obtains the signal $s(t)$ given by

$$s(t) = b(t) + \Sigma_{i=1}^{n} p_i b_i(t)$$

If $s(t)$ is closer to 1 than -1, the receiver assumes that the bit sent at time t was a 1; otherwise, it was a -1.
Assume that all the bit $b_i(t)$ can be considered independent, uniform random variables. Give a Chernoff bound to estimate the probability that the receiver makes an error in determining $b(t)$.

[Solution]:

Let X be the effect of combined noise. We may express X by p_i and b_i as:

$$X = p_1 b_1 + p_2 b_2 + \ldots + p_n b_n.$$

Also, $\mu = \mathrm{E}[X] = 0$.

From the question, we see that if the combined noise $|X| < 1$, the signal can be decoded properly. Thus,

$$\Pr(\text{receiver makes an error}) \leq \Pr(|X| \geq 1).$$

Below, we shall give a Chernoff bound for $\Pr(|X| \geq 1)$. First, we calculate the moment generating function of each $p_i b_i$:

$$\begin{aligned} M_{p_i b_i}(t) &= \mathrm{E}[e^{t p_i b_i}] = \frac{1}{2}\left(e^{t p_i} + e^{-t p_i}\right) \\ &= \sum_{k \geq 0} \frac{(t p_i)^{2k}}{(2k)!} \\ &\leq e^{t^2 p_i^2/2}. \end{aligned}$$

[Cont.]:
This implies

$$M_X(t) \le e^{t^2K/2}, \qquad \text{where } K = \sum_i p_i^2.$$

Thus, we have:

$$\begin{aligned}
\Pr(|X| \ge 1) &= 2\Pr(X \ge 1) \quad (X \text{ is symmetric around mean } \mu = 0) \\
&= 2\Pr(e^{tX} \ge e^t) \\
&\le \frac{2\,\mathrm{E}[e^{tX}]}{e^t} = \frac{2M_X(t)}{e^t} \\
&\le \frac{2e^{t^2K/2}}{e^t}.
\end{aligned}$$

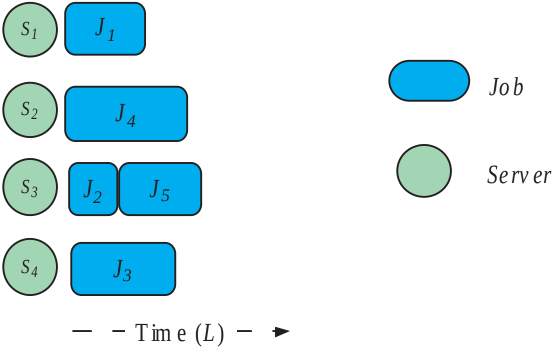
The last term is minimized when we set $t = 1/K$, so that we obtain:

$$\Pr(|X| \ge 1) \le 2e^{-1/(2K)}.$$

[Question 5]:

Let $J_1, J_2, \cdots, J_n$ be a set of jobs, where the i-th job J_i has an execution time of L_i seconds, $0 \leq L_i \leq 1$.

Suppose we have m servers, and we can assign the n jobs to run on them. The *load* of a server is the total execution time of all jobs assigned to it, and our goal is to find an assignment so that the load is as balanced as possible.


S_1
J_1
S_2
J_4
S_3
J_2
J_5
S_4
J_3
Time (L)
Job
Server

[Solution]:
(a) The expected load of each server is:

$$\bar{L} = \frac{\sum_{i=1}^{n} L_i}{m}.$$

[Cont.]:
(b) Let R_j denote the load assigned to server j. Then we have $\mathrm{E}[R_j] = \bar{L}$.
On the other hand, let $R_{j,i}$ be a random variable such that it has value L_i if job J_i is assigned to server j, and it has value 0 otherwise. Then, we see that:

$$R_j = R_{j,1} + R_{j,2} + \cdots + R_{j,n}.$$

[Cont.]:
(b) Since R_j is the sum of independent random variables whose range is between 0 and 1, we can apply Theorem 1 immediately, and obtain:

$$\begin{aligned}\Pr(R_j \geq c\bar{L}) &= \Pr(R_j \geq c\,\mathrm{E}[R_j]) \\ &\leq e^{-\alpha\,\mathrm{E}[R_j]} \qquad = e^{-\alpha\bar{L}},\end{aligned}$$

where $\alpha = c \ln c + 1 - c$.
Thus by union bound,

$$\Pr(\text{some server has load at least } c\bar{L}) \leq m e^{-\alpha\bar{L}}.$$

[Cont.]:
(c) Given $n = 100K$, $m = 10$, and the average job execution time is 0.25. Then $\Sigma_{i=1}^{n} L_i = 0.25n = 25K$ and $\bar{L} = 2500$.

$$\Pr(\text{all servers have load at most } 1.1\bar{L}) \geq 1 - e^{-9}.$$

means $c = 1.1$, thus α is about 0.004841.
By the result in part (b), we see that

$$\Pr(\text{some server has load at least } 1.1\bar{L}) \leq 10\, e^{-0.004841 \times 2500} \leq e^{-9}.$$

Thus, we get the desired bound that

$$\Pr(\text{all servers have load at most } 1.1\bar{L}) \geq 1 - e^{-9}.$$

[Question 6] (Bonus):
This question attempts to prove Theorem 1. Part (a) establishes a useful fact, while the remaining parts proceed to give the proof.
(a) Let c and λ be two positive real numbers, with $c > 1$ and $0 \leq \lambda \leq 1$.
Show that for any $z \in [-\lambda, 1-\lambda]$,

$$c^z \leq c^{-\lambda}(1 + \lambda(c-1)) + z(c^{1-\lambda} - c^{-\lambda}).$$

Furthermore, argue that

$$c^z \leq c^{-\lambda} e^{\lambda(c-1)} + z(c^{1-\lambda} - c^{-\lambda}).$$

(b) Next, we define $\Delta_i = X_i - \mathrm{E}[X_i]$. (Recall that each X_i is a random variable whose value is in between 0 and 1.) Thus, we have

$$-\mathrm{E}[X_i] \leq \Delta_i \leq 1 - \mathrm{E}[X_i] \quad \text{and} \quad \mathrm{E}[\Delta_i] = \mathrm{E}\left[X_i - \mathrm{E}[X_i]\right] = 0.$$

Let $\Delta = \sum_{i=1}^{n} \Delta_i$. Show that

$$\Pr(X \geq c\mathrm{E}[X]) \;=\; \Pr(\Delta \geq (c-1)\mathrm{E}[X]) \;\leq\; \frac{\mathrm{E}[c^{\Delta}]}{c^{(c-1)\mathrm{E}[X]}}.$$

(c) Using Part (a), show that if we set $\lambda = \mathrm{E}[X_i]$,

$$\mathrm{E}[c^{\Delta_i}] \le c^{-\lambda} e^{\lambda(c-1)} = e^{\lambda(c-1-\ln c)}.$$

(d) Argue that $\mathrm{E}[c^{\Delta}] \le e^{\mathrm{E}[X](c-1-\ln c)}$ and complete the proof of the theorem.

[Solution]:
(a)Let $f(z) = c^{-\lambda}(1 + \lambda(c-1)) + z(c^{1-\lambda} - c^{-\lambda}) - c^z$.
We may get $f'(z) = c^{1-\lambda} - c^{-\lambda} - zc^{z-1}$ and
$f''(z) = -z(z-1)c^{z-2}$.
Observe that $f''(z) < 0$ when $z \in [-\lambda, 0)$ and $f''(z) \geq 0$ when $z \in [0, 1-\lambda]$. This implies $f'(z)$ achieves the minimum value at $z = 0$. In other words:

$$f'(z) \geq f'(0) = c^{1-\lambda} - c^{-\lambda} \geq 0.$$

The above statement shows that $f(z)$ is monotonically increasing, so that

$$f(z) \geq f(-\lambda) = 0.$$

Thus, $c^z \leq c^{-\lambda}(1 + \lambda(c-1)) + z(c^{1-\lambda} - c^{-\lambda})$, so that

$$c^z \leq c^{-\lambda}e^{\lambda(c-1)} + z(c^{1-\lambda} - c^{-\lambda}).$$

[Cont.]:
(b)

$$\begin{aligned}
\Pr(X \geq c\,\mathrm{E}[X]) &= \Pr\left(\sum_i X_i \geq c\,\mathrm{E}[X]\right) \\
&= \Pr\left(\sum_i (\Delta_i + \mathrm{E}[X_i]) \geq c\,\mathrm{E}[X]\right) \\
&= \Pr(\Delta + \mathrm{E}[X] \geq c\,\mathrm{E}[X]) \\
&= \Pr(\Delta \geq (c-1)\mathrm{E}[X]) \\
&= \Pr(c^{\Delta} \geq c^{(c-1)\mathrm{E}[X]}) \\
&\leq \frac{\mathrm{E}[c^{\Delta}]}{c^{(c-1)\mathrm{E}[X]}}.
\end{aligned}$$

[Cont.]:

(c)By setting $\lambda = \mathrm{E}[X_i]$,

$$\begin{aligned}
\mathrm{E}[c^{\Delta_i}] &\le \mathrm{E}[c^{-\lambda} e^{\lambda(c-1)} + \Delta_i(c^{1-\lambda} - c^{-\lambda})] \\
&= \mathrm{E}[c^{-\lambda} e^{\lambda(c-1)}] \quad (\text{since } \mathrm{E}[\Delta_i] = \mathrm{E}[X_i - \mathrm{E}[X_i]] = 0) \\
&= c^{-\lambda} e^{\lambda(c-1)} \quad (\text{since } \lambda = \mathrm{E}[X_i] \text{ is a constant}) \\
&= e^{\lambda(c-1-\ln c)}.
\end{aligned}$$

[Cont.]:
(d)Since $\Delta_1, \Delta_2, \ldots, \Delta_n$ are independent, we have:

$$\mathrm{E}[c^{\Delta}] = \prod \mathrm{E}[c^{\Delta_i}] \leq \prod e^{\mathrm{E}[X_i](c-1-\ln c)} = e^{\mathrm{E}[X](c-1-\ln c)}.$$

Combining this with part (b), we have:

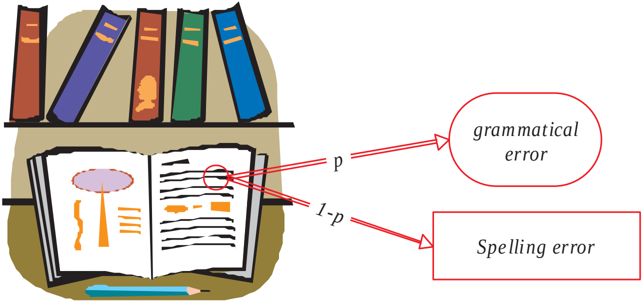
$$\begin{aligned}
\Pr(X \geq c\,\mathrm{E}[X]) &\leq \frac{e^{\mathrm{E}[X](c-1-\ln c)}}{c^{(c-1)\mathrm{E}[X]}} \\
&= \frac{e^{\mathrm{E}[X](c-1-\ln c)}}{e^{(c\ln c-\ln c)\mathrm{E}[X]}} \\
&= e^{\mathrm{E}[X](c-1-\ln c-c\ln c+\ln c)} \\
&= e^{-\alpha\,\mathrm{E}[X]},
\end{aligned}$$

where $\alpha = c\ln c + 1 - c$.

Hint for assignment 3

[Question 1]: Let X be a Poisson random variable with mean μ, representing the number of errors on a page of this book. Each error is independently a grammatical error with probability p and a spelling error with probability $1 - p$. If Y and Z be random variables representing the number of grammatical and spelling errors(respectively) on a page of this book, prove that Y and Z are Poisson random variables with means μp and $\mu(1 - p)$, respectively. Also prove that Y and Z are independent.

[Hint]:



By definition of Poisson random variable with some condition. Try to show $\Pr(Y = k) =$? and $\Pr(Z = k) =$?

[Question 2]: Let Z be a Poisson random variable of mean μ, where $\mu \geq 1$ is an integer.

1. Show that $\Pr(Z = \mu + h) \geq \Pr(Z = \mu - h - 1)$ for $0 \leq h \leq \mu - 1$.
2. Using part (a), argue that $\Pr(Z \geq \mu) \geq 1/2$.

[Hint]:
We may use the definition of Poisson distribution in page 14 of Lecture Note 13.

Definition:

A discrete Poisson random variable X with parameter μ is given by the following probability distribution for $r = 0, 1, 2, \ldots$:

$$Pr(X = r) = e^{-\mu}\mu^r/r!$$

[Question 3]:In Page 14 of Lecture Notes 14 we showed that, for any nonnegative functions f,

$$E[f(Y_1^{(m)}, ..., Y_n^{(m)})] \geq E[f(X_1^{(m)}, ..., X_n^{(m)})]\Pr(\Sigma Y_i^{(m)} = m)$$

1. Now suppose we further know that$E[f(X_1^{(m)}, ..., X_n^{(m)})]$ is monotonically increasing in m. Show that

   $$E[f(Y_1^{(m)}, ..., Y_n^{(m)})] \geq E[f(X_1^{(m)}, ..., X_n^{(m)})]Pr(\Sigma Y_i^{(m)} \geq m)$$

2. Combining part (a) and part (b) with the results in Question 2, prove the theorem in Page 20 of Lecture Notes 14.

[Hint]:
Re-examine the proof of Page 15 in Lecture Notes 14 (and show tighter bound when we know $f(x_1, x_2, ..., x_n)$ is monotonically increasing in m)

[Question 4]: 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 $X = \Sigma_{i=1}^{n} X_i$.
Let Y_i be independent Bernoulli random variable such that $Y_i = 1$ with probability $p = (1 - 1/n)^n$. Let $Y = \Sigma_{i=1}^{n} Y_i$.

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 $Pr(X \geq (1 + \delta)\mathrm{E}[X])$.

[Hint]:
Add oil.

Thank you