Randomized algorithm

Tutorial 2

Joyce

2008-10-28

Hints 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.

[Hint]:
When the kth number is the median, we may use the randomized algorithm describe in Lecture 9.
How about the other values?

[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.

[Hint]:
Chebyshev's inequality:

$$\Pr(|X - \mathrm{E}[X]| \geq a) \leq \frac{\mathrm{Var}[X]}{a^2}$$

[Question 3]:

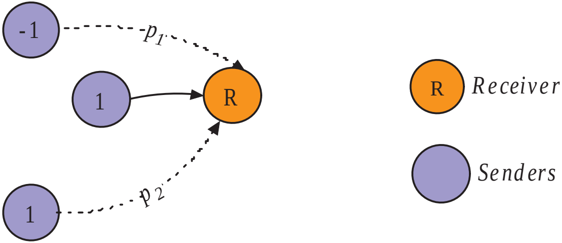
1. Determine the moment generating function for the binomial random variable $Bin(n, p)$.
2. Let X be a $Bin(n, p)$ random variable and Y be a $Bin(m, p)$ random variable. Suppose that X and Y are independent. Use part (a) to determine the moment generating function of $X + Y$.
3. What can we conclude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

[Hint]:

1. Binomial theorem.
   $\sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} = ?$
2. If X and Y are independent random variables, $M_{X+Y}(t) = ?$

[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

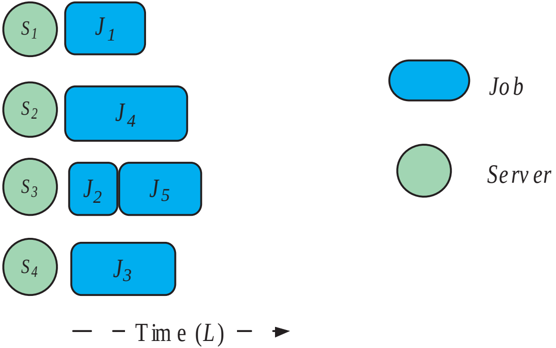
$$s(t) = b(t) + \Sigma_{i=1}^{n} p_i b_i(t)$$

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 $b(t)$.

[Hint]:

1. Let X be noise. Can you express X in terms of p_i and x_i?
2. What is the condition that an error occurs?

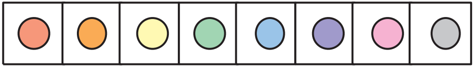
[Question 5]:



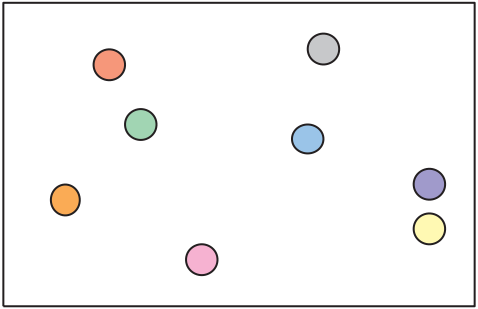
[Hint]:
This question can be done by following the steps.

Michael's algorithm

Determine the shortest distance between a pair of points in the array. (The points are in 2-d)



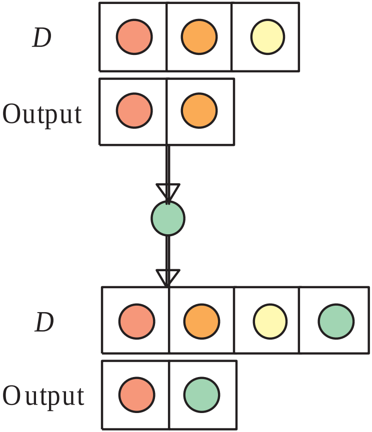
Information in the array represents the position of each node in 2D plane.



Now we have a storage data structure D.

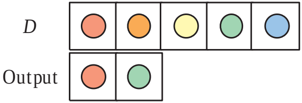
Each time we insert a point.

When we give a new point to D, it stores the point and answers the shortest distance of all points in D.

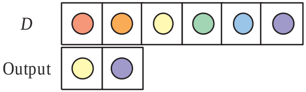


The time D takes depends on whether the answer change or not.

1. If output is the same: D takes 1 clock tick.



2. If output is the same: D takes $|D|$ clock tick.



Here comes the question.
Show that the expected total number of clock ticks used by D is $O(n)$.

Let X_i be the clock ticks when inserting the ith point and X be the total clock ticks.

1. What is the probability that the ith point causes answer to change?
2. What is $\mathrm{E}[X_i]$?

1. Pr(ith point causes answer to change)
   $=$ Pr(ith point is one of the shortest pair among the first i points)
   $= 2/i$ –(why?)
2. $E[X_i]$
   $= i * 2/i + 1 - 2/i$
   < 3
3. $E[X] = 1 + \Sigma E[X_i] = O(n)$

One of three

A company is going to develop a predict system by using machine learning.
For a given user, the algorithm runs
$\Pr(\mathit{success}) = p_1$
$\Pr(\mathit{failure}) = p_2$
$\Pr(\mathit{notsure}) = p_3$

The company runs their algorithm for n different items. (Assume the results are independent.)

Let

X_1: the total number of correct prediction.

X_2: the total number of failure prediction.

X_3: the total number of not sure prediction.

The question is to compute $\mathrm{E}[X_1 | X_3 = m]$.

1. $X_3 = m \rightarrow X_1 + X_2 = n - m$
2. Therefore, Pr(ith prediction is correct | not not sure) $= p_1/(p_1 + p_2)$.
3. Now we let X_1 be binomial random variable (n', p'),
   $\mathrm{E}[X_1|X_3 = m]$
   $= n'p'$
   $= (n - m)p_1/(p_1 + p_2)$

Thank you