Randomized algorithm

Tutorial 2

Joyce

2009-10-27

Hint for Assignment 2
- Question 1
- Question 2
- Question 3
- Question 4
- Question 5
- Question 6

Solutions to Assignment 1
- Question 1
- Question 2
- Question 3
- Question 4
- Question 5
- Question 6
- Question 7

Interesting quiz
- Problem definition
- Solution

Hint for Assignment 2

[Question 1]: A fixed point of a permutation $\pi : [1, n] \rightarrow [1, n]$ is a value for which $\pi(x) = x$. Find the variance in the number of fixed points of a permutation chosen uniformly at random from all permutations.

[Hint]:
Let X_i be an indicator such that $X_i = 1$ if $\pi(i) = i$. Then, $\Sigma_{i=1}^{n} X_i$ is the number of fixed points.
You cannot use linearity to find $\mathrm{Var}[\Sigma_{i=1}^{n} X_i]$, but you can calculate it directly.

[Question 2]: Recall that the covariance of random variables X and Y is

$$\mathrm{Cov}[X, Y] = \mathrm{E}[(X - \mathrm{E}[X])(Y - \mathrm{E}[Y])]] = \mathrm{E}[XY] - \mathrm{E}[X]\mathrm{E}[Y]$$

Demonstrate an explicit example where $\mathrm{Cov}[X, Y] = 0$, yet X and Y are not independent.

[Question 3]:

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

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

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

[Hint]:
Chebyshev's inequality:

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

[Question 4]: Suppose you are given a biased coin that has $\Pr[heads] = p$. Also, suppose that we know $p \geq a$ for some fixed a. Now, consider flipping the coin n times and let n_H be the number of times a head came up.Naturally, we would estimate p by the value $\tilde{p} = n_H/n$.

1. Show that for any $\epsilon \in [0,1]$

$$\Pr(|p-\tilde{p}| > \epsilon p) < \exp\left(\frac{-na\epsilon^2}{2}\right) + \exp\left(\frac{-na\epsilon^2}{3}\right)$$

2. Show that for any $\epsilon \in [0,1]$, if

$$n > \frac{3\ln 2/\delta}{a\epsilon^2}$$

then $\Pr(|p-\tilde{p}| > \epsilon p) < \delta$

[Hint]:
Parameter estimation

[Question 5]:Let $X_1, X_2, ..., X_n$ be independent Poisson trials such that $\Pr(X_i) = p_i$. Let $X = \Sigma_{i=1}^{n} X_i$ and $\mu = [X]$. During the class, we have learnt that for any $\delta > 0$,

$$\Pr(X \geq (1+\delta)\mu) < \left(\frac{e^{\delta}}{(1+\delta)^{(1+\delta)}}\right)^{\mu}$$

In fact, the above inequality holds for the *weighted* sum of Poisson trials. Precisely, let $a_1, ..., a_n$ be real numbers in $[0, 1]$. Let $W = \Sigma_{i=1}^{n} a_i X_i$ and $\nu = \mathrm{E}[W]$. Then, for any $\delta > 0$,

$$\Pr(W \geq (1+\delta)\nu) < \left(\frac{e^{\delta}}{(1+\delta)^{(1+\delta)}}\right)^{\mu}$$

1. Show that the above bound is correct.
2. Prove a similar bound for the probability $\Pr(W \leq (1-\delta)\nu)$ for any $0 < \delta < 1$.

[Hint]:

- Moment Generating Function
- Markov Inequality

[Question 6]:Consider a collection $X_1, X_2, ..., X_n$ of n independent geometrically random variables with parameter $1/2$. Let $X = \Sigma_{i=1}^{n} X_i$ and $0 < \delta < 1$.

1. By applying the Chernoff bound to a sequence of $(1+\delta)(2n)$ fair coin tosses, show that

$$\Pr(X > (1+\delta)(2n)) < \exp\left(\frac{-n\delta^2}{2(1+\delta)}\right)$$

2. Derive a Chernoff bound on $\Pr(X > (1+\delta)(2n))$ using the moment generating function for geometric random variables

[Hint]:

- $(1 + \delta)(2n)$ is an integer.
- Sum of geometric random variable by binomial random variable.

Solutions to Assignment 1

[Question 1]:

$$
\begin{aligned}
& \Pr(\text{"Head appears on an odd toss"}) \\
= \; & \Pr(A_1 \cup A_3 \cup A_5 \cup \cdots) \\
= \; & \sum_{i=0}^{\infty} \Pr(A_{2i+1}) = p \sum_{i=0}^{\infty} q^{2i} \\
= \; & \frac{p}{1-q^2} = \frac{p}{(1+q)(1-q)} \\
= \; & \frac{1}{1+q} = \frac{1}{2-p}
\end{aligned}
$$

because $A_i \cap A_j = \emptyset$, $i \neq j$. Even for a fair coin, the probability of "Head first appears on an odd toss" is $2/3$.

[Question 2]:

$$\begin{aligned} W &= \{\text{"the transferred ball from Box 1 to Box 2 is white"}\}, \text{ and} \\ B &= \{\text{"the transferred ball from Box 1 to Box 2 is black"}\}. \end{aligned}$$

$\Pr(W) + \Pr(B) = 1$, and

$$\Pr(W) = \frac{a}{a+b} \text{ and } \Pr(B) = \frac{b}{a+b}.$$

Let A be the desired event that "the next ball drawn from Box 2 is white". Hence,

$$\begin{aligned}
\Pr(A) &= \Pr\{A \cap (W \cup B)\} \\
&= \Pr\{(A \cap W) \cup (A \cap B)\} \\
&= \Pr(A \cap W) + \Pr(A \cap B)
\end{aligned}$$

Since

$$\Pr(A|W) = \frac{c+1}{c+d+1} \text{ and } \Pr(A|B) = \frac{c}{c+d+1}.$$

we have

$$\Pr(A) = \frac{a(c+1)}{(a+b)(c+d+1)} + \frac{bc}{(a+b)(c+d+1)} = \frac{ac+bc+a}{(a+b)(c+d+1)}$$

[Question $3 - a$]:

$$\frac{a}{a+b} > \frac{a-1}{a+b-1} \text{ for } a, b \geq 1$$
$$\Rightarrow \quad (\frac{a}{a+b})^2 > (\frac{a}{a+b})(\frac{a-1}{a+b-1}) > (\frac{a-1}{a+b-1})^2.$$

Combining with

$$\frac{a}{a+b}\frac{a-1}{a+b-1} = \frac{1}{3}.$$

we get

$$(\frac{a}{a+b})^2 > \frac{1}{3} > (\frac{a-1}{a+b-1})^2$$

[Question $3 - b$]: From the left hand side of solution of (a)

$$\begin{aligned}
(\frac{a}{a+b})^2 &> \frac{1}{3} \\
\Rightarrow \quad (\frac{(\sqrt{3}+1)b}{2}) &< a.
\end{aligned}$$

Similarly, from the right inequality of part (a), that is

$$\begin{aligned}
\frac{1}{3} &> (\frac{a-1}{a+b-1})^2 \\
\Rightarrow \quad a &< 1 + \frac{(\sqrt{3}+1)b}{2}.
\end{aligned}$$

[Question $3 - c$]: Let W_i denote the event that the ith ball drawn is white. From the requirement, we know that there must be at least one black balls; otherwise,

$$\Pr(W_1 \bigcap W_2) = 1 \neq \frac{1}{3}.$$

If there are only one black ball (so that $b = 1$), we have $1.36 < a < 2.36$ from part (b). Thus there must be two white balls. By checking

$$\Pr(W_1 \bigcap W_2) = \frac{2}{3} \cdot \frac{1}{2} = \frac{1}{3},$$

we conclude that one black two whites can yield the desired probability. Thus, the smallest number of balls is 3.

[Question $3 - d$]: If $b = 2$ (so that b is set to minimum possible value), a would be 3, but

$$\Pr(W_1 \bigcap W_2) = \frac{3}{10}$$

which is not correct.
If $b = 4$ (so that b is set to the next minimum possible value), a would be 6, and

$$\Pr(W_1 \bigcap W_2) = \frac{1}{3}.$$

We conclude that 4 blacks and 6 whites can yield the desired probability, and the smallest number of balls (when b is even) is 10.

[Question 4]:
Let X denote the number of inversions in an array, and $X_{i,j}$ denote the indicator for the pair (i, j) being an inversion.

$$\mathrm{E}[X] = \mathrm{E}[\sum_{i<j} X_{i,j}] = \sum_{i<j} \mathrm{E}[X_{i,j}] = \sum_{i<j} \frac{1}{2} = \frac{n(n-1)}{4}.$$

[Question 5]:
Let X be the random variable which counts the number of pairs which are coupled. Let X_i be an indicator such that

$$\begin{aligned}\Pr(X_i) &= 1 \text{ if the } i\text{th pair are coupled}\\ \Pr(X_i) &= 0 \text{ otherwise}\end{aligned}$$

Then, $X = X_1 + X_2 + \cdots + X_20$, and $\mathrm{E}[X_i] = 1/20$. By linearity of expectation,

$$\mathrm{E}[X] = \mathrm{E}[\sum_{i=1} 20 X_i] = \sum_{i=1} 20\mathrm{E}[X_i] = \sum_{i=1} 20\frac{1}{20} = 1.$$

[Question $6 - a$]: $\min(X, Y)$ is equivalent to the number of times we need to perform two experiments together, each with success probability p and q, in order to obtain a success from at least one of these experiments. In other words, $\min(X, Y)$ is a geometric random variable with parameter $(1-(1-p)(1-q))$,i.e.,

$$\Pr(\min(X, Y) = k) = ((1-p)(1-q))^{k-1}(p+q-pq)$$

[Question $6 - b$]: At first, we define an indictor Z, such that

$$\begin{aligned} \mathrm{E}[X|X \leq Y] &= \sum x \Pr(X = k|X \leq Y) \\ &= \sum x \frac{\Pr(X = k \cap X \leq Y)}{\Pr(X \leq Y)} \\ &= \frac{1}{\Pr(X \leq Y)} \sum x \Pr(X = x) \Pr(Y \geq x) \end{aligned}$$

Now, we calculate $\Pr(X \leq Y)$ at first.

$$\begin{aligned}
\Pr(X \leq Y) &= \sum_i \Pr(X = i \cap Y \geq i) \\
&= \sum_i \Pr(X = i)\Pr(Y \geq i) \\
&= \sum_i (1-p)^{i-1} p (1-q)^{i-1} \\
&= \frac{p}{p+q-pq}
\end{aligned}$$

Next, we have:

$$\begin{aligned}
\sum x\Pr(X = x)\Pr(Y \geq x) &= \sum x(1-p)^{i-1}p \cdot \Pr(Y \geq x) \\
&= \sum x(1-p)^{i-1}p(1-q)^{i-1} \\
&= p\sum x(1-p)^{i-1}(1-q)^{i-1} \\
&= \frac{p}{(1-(1-p)(1-q))^2} = \frac{p}{(p+q-pq)^2}
\end{aligned}$$

Combining the above results, we have

$$\mathrm{E}[X|X \leq Y] = \frac{1}{(p+q-pq)}$$

[Question $6 - c$]:

$$\begin{aligned}
\Pr(X = Y) &= \Pr(X = Y = 1) + \Pr((X = Y) \cap ((X > 1) \cap (Y > 1))) \\
&= pq + \Pr(X = Y = 1) \\
&\quad + \Pr((X = Y) | ((X > 1) \cap (Y > 1))) \\
&\quad \cdot \Pr((X > 1) \cap (Y > 1)) \\
&= pq + \Pr(X = Y = 1) \\
&\quad + \Pr((X = Y) | ((X > 1) \cap (Y > 1)))(1 - p)(1 - q)) \\
&= pq + \Pr(X = Y)(1 - p)(1 - q),
\end{aligned}$$

where the third equality is by independence of X and Y, and the fourth equality is by memoryless property. Next, by rearranging terms, we get:

$$\Pr(X = Y) = \frac{pq}{1 - (1 - p)(1 - q)} = \frac{pq}{p + q - pq}.$$

[Question $6 - d$]:

$$
\begin{aligned}
\mathrm{E}[\max(X,Y)] &= \Pr(Y=1)\mathrm{E}[\max(X,Y)|(Y=1)] \\
&\quad +\Pr(Y>1)\mathrm{E}[\max(X,Y)|(Y=1)] \\
&= q \times \mathrm{E}[X] \\
&\quad +\Pr((X=1)\cap(Y>1)) + \mathrm{E}[\max(X,Y)|((X=1)\cap \\
&\quad +\Pr((X>1)\cap(Y>1)) + \mathrm{E}[\max(X,Y)|((X>1)\cap \\
&= q \times \mathrm{E}[X] \\
&\quad +p(1-q)\mathrm{E}[Y+1] \\
&\quad +(1-p)(1-q) \times \mathrm{E}[\max(X,Y)+1]
\end{aligned}
$$

where the third equality follows from memoryless properties. By rearranging terms,

$$\begin{aligned}(p+q-pq)\mathrm{E}[\max(X,Y)] &= \frac{q}{p}+\frac{p}{q}-pq+(1-p)(1-q)\\ &= \frac{q}{p}+\frac{p}{q}+1-p-q\\ &= \frac{q-pq}{p}+\frac{p-pq}{q}+1\\ &= \frac{p+q-pq}{p}+\frac{p+q-pq}{q}-1\end{aligned}$$

so that

$$\mathrm{E}[\max(X,Y)] = \frac{1}{p}+\frac{1}{q}-\frac{1}{p+q-pq}$$

[Question $7 - a$]:
To choose the ith candidate, we need

- $i > m$
- ith candidate is the best [Event B]
- the best of the first $i - 1$ candidates (say y) is among the first m candidates (otherwise, if y is not among the first m candidates, we will choose y by our interview strategy)

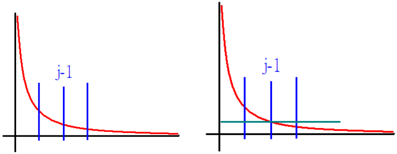
Then, we have

$$\Pr(E_i) = \begin{cases} 0 & \text{for } i \leq m \\ \Pr(B_i \cap Y_i) & \text{for} i > m \end{cases}$$

From our definition, we can see that $\Pr(E) = \sum_{i=1}^{n} \Pr(E_i)$. Thus, we have

$$\Pr(E) = \sum_{i=1}^{n} \Pr(E_i) = \sum_{i=m+1}^{n} \Pr(E_i) = \frac{m}{n} \sum_{i=m+1}^{n} \frac{1}{i-1}$$

[Question $7 - b$]:
Consider the curve $f(x) = 1/x$. The area under the curve from $x = j - 1$ to $x = j$ is less than $1/(j-1)$.



Thus,

$$\sum_{j=m+1}^{n} \frac{1}{j-1} \geq \int_{m}^{n} f(x)dx = \log_e n - \log_e m$$

Similarly, the area under the curve from $x = j - 2$ to $x = j - 1$ is greater than $1/(j-1)$. Thus,

$$\sum_{j=m+1}^{n} \frac{1}{j-1} \leq \int_{m-1}^{n-1} f(x)dx = \log_e(n-1) - \log_e(m-1)$$

Combining these two inequalities with part (a).

[Question $7-c$]: Let $g(m) = m\log_e n - \log_e m/n$. By differentiating $g(m)$, we get

$$g''(m) = \frac{-1}{mn} < 0$$

which indicates that $g(m)$ attains maximum when $m = n/e$. By substituting $m = n/e$ in the inequality of part(b), we get

$$\Pr(E) \geq \frac{m(\log_e n - \log_e m)}{n} = \frac{n(\log_e n - \log_e(n/e))}{ne} = \frac{n\log_e e}{ne} = \frac{1}{e}.$$

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, $\Pr(i$th 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