Randomized algorithm

Tutorial 3

Joyce

2009-11-10

Solution to Assignment 2

[Question 1]:(10%) 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.

[Solution]: Let X_i be an indicator such that $X_i = 1$ if $\pi(i) = i$. Then, $\sum_{i=1}^{n} X_i$ is the number of fixed points.

$$\begin{aligned} \mathrm{Var}[X] &= \mathrm{E}[X^2] - (\mathrm{E}[X])^2 \\ &= \sum_{i=1}^{n} \mathrm{E}[X_i^2] + \sum_{i \neq j} \mathrm{E}[X_i X_j] - (\mathrm{E}[X])^2 \end{aligned}$$

Since $\mathrm{E}[X_i^2] = \mathrm{Pr}(X_i^2 = 1) = \frac{1}{n}$ and

$$\begin{aligned} \mathrm{E}[X_i X_j] &= \mathrm{Pr}(X_i X_j = 1) \\ &= \mathrm{Pr}(X_i = 1)\mathrm{Pr}(X_j = 1 \mid X_i = 1) \\ &= \frac{1}{n(n-1)}, \end{aligned}$$

we get

$$\mathrm{Var}[X] = n \times \frac{1}{n} + n(n-1) \times \frac{1}{n(n-1)} - 1 = 1.$$

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

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

We have seen that if X and Y are independent, then the covariance is 0. Interestingly, if X and Y are not independent, the covariance may still be 0.
(15%) Construct an example where X and Y are not independent, yet $\mathrm{Cov}[X, Y] = 0$.

[Solution]: Let the sample space be $\{-1, 0, 1\}$, with each outcome having the same probability to occur. Let X denote the outcome and let $Y = X^2$. We know that $\Pr(X = 0 \cap Y = 0) = 1/3$ and $\Pr(X = 0)\Pr(Y = 0) = 1/9$ so that X and Y are not independent. However,

$$\begin{aligned} \mathrm{Cov}[X, Y] &= \mathrm{E}[(X - \mathrm{E}[X])(Y - \mathrm{E}[Y])] \\ &= \frac{1}{3}((-1)(1 - \mathrm{E}[Y]) + (0)(0 - \mathrm{E}[Y]) + (1)(1 - \mathrm{E}[Y])) \\ &= 0. \end{aligned}$$

[Question 3]:
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 $\epsilon > 0$ we have

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

(15%) Use Chebyshev's inequality to prove the weak law of large numbers.

[Solution]:

$$\begin{aligned} & \Pr\left(\left|\frac{X_1 + X_2 + X_3 + \ldots + X_n}{n} - \mu\right| > \epsilon\right) \\ = \; & \Pr\left(|X_1 + X_2 + X_3 + \ldots + X_n - n\mu| > n\epsilon\right) \end{aligned}$$

By Chebyshev's inequality:

$$\begin{aligned} & \Pr\left(|X_1 + X_2 + X_3 + \ldots + X_n - n\mu| > n\epsilon\right) \\ \leq \; & \frac{\mathrm{Var}[X_1 + X_2 + X_3 + \ldots + X_n]}{(n\epsilon)^2} \\ = \; & \frac{\Sigma \mathrm{Var}[X_i]}{n^2\epsilon^2} = \frac{n\sigma^2}{n^2\epsilon^2} = \frac{\sigma^2}{n\epsilon^2} \end{aligned}$$

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

[Question 4]: Suppose you are given a biased coin that has $\Pr(\text{head}) = p$. 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 comes up. Naturally, we would estimate p by the value $\tilde{p} = n_H/n$.

1. Show that for any $\epsilon \in (0,1)$,

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

2. Show that for any $\epsilon \in (0,1)$, if

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

then

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

[Solution]: (1)

$$
\begin{aligned}
& \Pr(|p - \tilde{p}| > \epsilon p) \\
= & \Pr(n\tilde{p} < np - \epsilon pn) + \Pr(n\tilde{p} > np + \epsilon pn) \\
= & \Pr(X < \mathsf{E}[X](1-\epsilon)) + \Pr(X > \mathsf{E}[X](1+\epsilon)) \\
< & e^{-np\epsilon^2/2} + e^{-np\epsilon^2/3} \\
< & e^{-na\epsilon^2/2} + e^{-na\epsilon^2/3}
\end{aligned}
$$

[Solution]: (2)

$$
\begin{aligned}
& n > \frac{3\ln(2/\delta)}{a\epsilon^2} \\
\Rightarrow\quad & na\epsilon^2/3 > \ln(2/\delta) \\
\Rightarrow\quad & \ln(\delta/2) > -na\epsilon^2/3 \\
\Rightarrow\quad & \delta/2 > e^{-na\epsilon^2/3} \\
\Rightarrow\quad & \delta > e^{-na\epsilon^2/3} + e^{-na\epsilon^2/2} \\
\Rightarrow\quad & \delta > \Pr(|p - \tilde{p}| > \epsilon p)
\end{aligned}
$$

[Question 5]:(20%) Let $X_1, X_2, ..., X_n$ be independent Poisson trials such that $\Pr(X_i) = p_i$. Let $X = \sum_{i=1}^{n} X_i$ and $\mu = \mathrm{E}[X]$. 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 $W = \sum_{i=1}^{n} a_i X_i$ and $\nu = \mathrm{E}[W]$. Then, for any $\delta > 0$,

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

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

[Solution]: (1) Since $W = \Sigma_{i=1}^{n} a_i X_i$, we have

$$\nu = \mathrm{E}[W] = \Sigma_{i=1}^{n} a_i \mathrm{E}[X_i] = \Sigma_{i=1}^{n} a_i p_i$$

For any i,

$$\mathrm{E}[e^{ta_i X_i}] = p_i e^{ta_i} + (1 - p_i) = 1 + p_i(e^{ta_i} - 1) \le e^{p_i(e^{ta_i} - 1)}$$

Claim 1. For any $x \in [0, 1]$, $e^{tx} - 1 \le x(e^t - 1)$

$$\begin{aligned} & f(x) = x(e^t - 1) - e^{tx} + 1 \\ \Rightarrow\ & f'(x) = (e^t - 1) - te^{tx} \\ \Rightarrow\ & f'(x) = 0 \text{ (when } x = x^* = (\ln(e^t - 1) - \ln t)/t) \\ \Rightarrow\ & f''(x) = -t^2 e^{tx} \le 0 \end{aligned}$$

In other words, for $x \in [0, 1]$, $f(x)$ achieves minimum value either at $f(0)$ or $f(1)$. So $f(x) \ge min\{f(0), f(1)\} = 0$ for all $x \in [0, 1]$.

[Solution]: Hence,

$$\mathsf{E}[e^{eta_i X_i}] \le e^{p_i(e^{ta_i}-1)} \le e^{a_i p_i(e^t-1)},$$

By the independence of X_i's and property of MGF,

$$\mathrm{E}[e^{tW}] = \prod_{i=1}^{n} \mathrm{E}[e^{ta_i X_i}] \le \prod_{i=1}^{n} \mathrm{E}[e^{a_i p_i(e^t-1)}] = e^{\nu(e^t-1)}$$

For any $t > 0$, we have

$$\Pr(W \ge (1+\delta)\nu) = \Pr(e^{tW} \ge e^{t(1+\delta)\nu}) \le \frac{\mathrm{E}[e^{tW}]}{e^{t(1+\delta)\nu}} \le \frac{e^{\nu(e^t-1)}}{e^{t(1+\delta)\nu}}$$

Then, for any $\delta > 0$, we can set $t = \ln(1+\delta) > 0$ and obtain:

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

[Solution]: (2) For any $t < 0$, we have

$$\begin{aligned} & \Pr(W \leq (1-\delta)\nu \\ = & \Pr(e^{tW} \geq e^{t(1-\delta)\nu}) \\ \leq & \frac{\mathrm{E}[e^{tW}]}{e^{t(1-\delta)\nu}} \\ \leq & \frac{e^{\nu(e^t-1)}}{e^{t(1-\delta)\nu}} \end{aligned}$$

Then, for any $0 < \delta < 1$, we can set $t = \ln(1-\delta) < 0$ and obtain:

$$\Pr(W \leq (1-\delta)\nu) < \left(\frac{e^{\delta}}{(1-\delta)^{(1-\delta)}}\right)^{\nu}$$

[Question 6]:(30%) Consider a collection $X_1, X_2, ..., X_n$ of n independent geometric random variables with parameter $1/2$. Let $X = \sum_{i=1}^{n} X_i$ and $0 < \delta < 1$.

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

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

[Solution]:(1)

X_i : a sequence of coin flips until the first heads comes.

ΣX_i : a sequence of coin flips until we see the n-th head.

$X > (1+\delta)2n$: the n-th head does not occur among the first $(1+\delta)2n$ coin flips.

Let Y be the random variable giving the number of heads among the first $(1+\delta)2n$ coin flips. Then we have

$\Pr(X > (1+\delta)2n) = \Pr(Y < n)$

Noting that $\mathrm{E}[Y] = (1+\delta)n$, we have

$$\begin{aligned}
\Pr(X > (1+\delta)2n) &= \Pr(Y < n) \\
&= \Pr\left(Y < (1 - \frac{\delta}{1+\delta})(1+\delta)n\right) \\
&\leq \exp\left(-(1+\delta)n \cdot \frac{\delta^2}{2(1+\delta)^2}\right) \\
&= \exp\left(\frac{n\delta^2}{2(1+\delta)}\right)
\end{aligned}$$

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

(i) Show that for $e^t < 2$,

$$\mathrm{E}\left[e^{tX_i}\right] = \frac{e^t}{2-e^t}.$$

(ii) Show that for $t \in (0, \ln 2)$,

$$\left|\frac{1}{(2-e^t)e^{t(1+2\delta)}}\right| \quad \text{is minimized when } t = \ln\left(1 + \frac{\delta}{(1+\delta)}\right).$$

(iii) Show that

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

[Solution]:(2-i)

Recall that $e^t < 2$. Then we have:

$$\begin{aligned}
\mathrm{E}[e^{tX_i}] &= \frac{1}{2}e^t + \frac{1}{4}e^2t + \frac{1}{8}e^3t + ... \\
&= \Sigma_{k=1}\infty\left(\frac{e^t}{2}\right)^k \\
&= \frac{e^t/2}{1-e^t/2} \\
&= \frac{e^t}{2-e^t}
\end{aligned}$$

(2-ii)

Use differentiation with respect to t.

[Solution]:(2-iii) Apply Markov inequality,

$$\begin{aligned}
\Pr(X > (1+\delta)2n) &= \Pr(\exp(tX) > \exp(t(1+\delta)2n)) \\
&\leq \frac{\mathrm{E}[\exp(tX)]}{\exp(t(1+\delta)2n)} \\
&= \frac{\prod \mathrm{E}[\exp(tX_i)]}{\exp(t(1+\delta)2n)}
\end{aligned}$$

Combine former results and apply $t = \ln(1 + \frac{\delta}{1+\delta})$,

$$\begin{aligned}
\Pr(X > (1+\delta)2n) &= \frac{\prod \mathrm{E}[\exp(tX_i)]}{\exp(t(1+\delta)2n)} \\
&\leq \frac{e^{tn}}{(2-e^t)e^{t(1+\delta)2n}} \\
&= \left(\left(1 - \frac{\delta}{1+\delta}\right)\left(1 + \frac{\delta}{1+\delta}\right)^{1+2\delta}\right)^{-n}
\end{aligned}$$

3. It is known that when δ is small, there exists $\epsilon > 0$ such that

$$1 - \frac{\delta}{1+\delta} > e^{-\epsilon}, \qquad \left(1 + \frac{\delta}{1+\delta}\right)^{(1+\delta)/\delta} > e^{1-\epsilon},$$

$$\text{and} \qquad \frac{(1+2\delta)\delta}{1+\delta} > \delta^2.$$

Show that in this case, the bound in 6(b)-(iii) becomes

$$\Pr(X > (1+\delta)(2n)) < \exp\left(-n(1-\epsilon)\delta^2 - \epsilon\right).$$

Conclude that when δ is small enough such that ϵ is arbitrarily close to 0, the above bound is tighter than the bound obtained in 6(a).

[Solution]:(3) By substitution

$$
\begin{aligned}
& \left(\left(1-\frac{\delta}{1+\delta}\right)\left(1+\frac{\delta}{1+\delta}\right)^{1+2\delta}\right)^{-n} \\
\leq\ & \left(\exp\left(-\epsilon+(1-\epsilon)(\frac{\delta^2}{1+\delta}+\delta)\right)\right)^{-n} \\
\leq\ & \exp\left(-n((1-\epsilon)\delta^2-\epsilon)\right)
\end{aligned}
$$