UC Berkeley, CS 174: Combinatorics and Discrete Probability (Fall 2010)

Solutions to Problem Set 4

1. **(MU 5.4)** In a lecture hall containing 100 people, you consider whether or not there are three people in the room who share the same birthday. Explain how to calculate this probability exactly, using the same assumptions as in our previous analysis.

   There are a number of ways to count this value, so several solutions were accepted. One such solution is given here.

   Let A_0 be the event that everyone in the room has a distinct birthday. Let A_i be the event that i pairs of people share a birthday. Then the probability we are interested in is

$$\mathbb{P}[\text{at least 3 people share a birthday}] = 1 - \mathbb{P}[A_0] - \sum_{i=1}^{n/2} \mathbb{P}[A_i] \tag{1}$$

   where there are n people in the room.

   First, we note that $\mathbb{P}[A_0]$ is in the book and is $\mathbb{P}[A_0] = \frac{\binom{365}{n} n!}{365^n}$. Second, we wish to compute $\mathbb{P}[A_i]$. For $i = 1$, we have

$$\mathbb{P}[A_1] = 365 \frac{\binom{n}{2}\binom{365-1}{n-2}(n-2)!}{365^n}.$$

   For $i = 2$ we have

$$\mathbb{P}[A_2] = \frac{\binom{365}{2}\binom{n}{2}\binom{n-2}{2}\binom{365-2}{n-4}(n-4)!}{365^n}.$$

   This generalizes for i pairs of people sharing i distinct birthdays

$$\mathbb{P}[A_i] = \frac{\binom{365}{n}\left(\prod_{j=1}^{i}\binom{n-2j+2}{2}\right)\binom{365-i}{n-2i}(n-2i)!}{365^n}.$$

   Plugging everything into Equation 1 and putting in $n = 100$, we can compute a value for this probability. We get $\mathbb{P}[\text{at least 3 people share a birthday}] \simeq 0.6459$.

2. **(MU 5.8)** Our analysis of Bucket sort in Section 5.2.2 assumed that n elements were chosen independently and uniformly at random from the range $[0, 2^k)$. Suppose instead that n elements are chosen independently from the range $[0, 2^k)$ according to a distribution with the property that any number $x \in [0, 2^k)$ is chosen with probability at most $a/2^k$ for some fixed constant $a > 0$. Show that, under these conditions, Bucket sort still requires linear expected time.

   Recall from the proof in section 5.2.2 that the expected sorting time for all the buckets is at most

$$\mathbb{E}\left[\sum_{j=1}^{n} c(X_j)^2\right] = c\sum_{j=1}^{n} \mathbb{E}[X_j^2].$$

   We will now use the probability that any bin is chosen with probability at most $a/2^k$. The number of elements any bin are upper bounded by a random variable $Y \sim Bin(n, a/n)$. To upper bound the expected time for sorting, we see that

$$
\begin{aligned}
c\sum_{j=1}^{n}\mathbb{E}[X_j^2] &\le cn\mathbb{E}[Y^2] \\
&= cn\left[n(n-1)\left(\frac{a}{n}\right)^2 + n\left(\frac{a}{n}\right)\right] \qquad \text{(by p48 of book)} \\
&= cn\left[\left(1-\frac{1}{n}\right)a^2 + a\right] \\
&\le cn(a^2+a).
\end{aligned}
$$

This is linear-time sorting, since c and $a^2 + a$ are both constants.

3. **(MU 5.9, a and b)**

   Consider the probability that every bin receives exactly one ball when n balls are thrown uniformly at random into n bins. Let X_i be the event that the i'th bin has one ball, and let X be the event that all the bins have a single ball: $X = \cup_{i=1}^{n} X_i$.

   (a) Using a Poisson approximation, we find an upper bound. Let Y_i be a Poisson random variable for bin i, and recall that $m = n$

   $$
   \begin{aligned}
   \mathbb{P}[Y_i = j] &= \frac{e^{-n/n}(\frac{n}{n})^j}{j!} \\
   \mathbb{P}[Y_i = 1] &= \frac{1}{e}
   \end{aligned}
   $$

   Applying the constant factor for the Poisson approximation (Cor. 5.9), we can bound the probability of X using $Y = \cup_{i=1}^{n}\{Y_i = 1\}$.

   $$
   \begin{aligned}
   \mathbb{P}[X] &\le e\sqrt{n}\mathbb{P}[Y] \\
   &= e\sqrt{n}\left(\frac{1}{e}\right)^n.
   \end{aligned}
   $$

   (b) Now, we determine the exact probability of this event. There are n^n ways to throw n balls into n bins. There is only one way, given a specific ordering of the balls, to put a single ball into each bin. And there are $n!$ possible orderings of the balls. Thus the exact probability of n balls in n bins is:

   $$\mathbb{P}[X] = \frac{n!}{n^n}.$$

4. **(MU 5.11)** The following problem models a simple distributed system wherein agents contend for resources but "back off" in the face of contention. Balls represent agents, bins represent resources.

   The system evolves over rounds. Every round, balls are thrown independently and uniformly at random into n bins. Any ball that lands in a bin by itself is served and removed from consideration. The remaining balls are thrown again in the next round. We begin with n balls in the first round, and we finish when every ball is served.

(a) If there are b balls in a stage, the probability that a particular ball lands in a bin by itself is $(1-1/n)^{b-1}$. By linearity of expectation, the expected number of balls that get served is $b(1-1/n)^{b-1}$, and the expected number of those that remain is $b - b(1-1/n)^{b-1}$.

(b) Suppose that 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 $O(loglogn)$ rounds.

We use the previous part together with the assumption that in each round the expected number of balls are served. To obtain the second line below, notice also that $(1-x)^k \geq 1-kx$ for $k \geq 0$.

$$\begin{aligned} x_{j+1} &= x_j(1-(1-1/n)^{x_j-1}) \\ &\leq x_j(1-(1-(x_j-1)/n)) \\ &= x_j(x_j-1)/n \\ &\leq x_j^2/n. \end{aligned}$$

Notice from the exact formula in part (a) that a constant number of rounds i suffices to reduce n balls to $n/2$ balls. By recursively expanding the above relation, we see that

$$\begin{aligned} x_{i+t} &\leq x_i^{2^t}/n^{2^t-1} \\ &\leq (n/2)^{2^t}/n^{2^t-1} \\ &\leq n/2^{2^t}. \end{aligned}$$

This value is 1 when $t = O(\log\log n)$. Since i is constant, the total number of rounds is $i + t = O(\log\log n)$.

5. **(MU 5.13, a and b)**

Let Z be a Poisson random variable with mean μ where $\mu \geq 1$ is an integer.

(a) Show that $\mathbb{P}[Z = \mu + h] \geq \mathbb{P}[Z = \mu - h - 1]$ for $0 \leq h \leq \mu - 1$.

$$\begin{aligned} \mathbb{P}[Z = \mu + h] &\geq \mathbb{P}[Z = \mu - h - 1] \\ \frac{e^{-\mu}\mu^{\mu+h}}{(\mu+h)!} &\geq \frac{e^{-\mu}\mu^{\mu-h-1}}{(\mu-h-1)!} \\ \frac{\mu^{\mu+h}}{(\mu+h)!} &\geq \frac{\mu^{\mu-h-1}}{(\mu-h-1)!} \\ \mu^{2h+1} &\geq \frac{(\mu+h)!}{(\mu-h-1)!} \\ \mu^{2h+1} &\geq (\mu-h)(\mu-h+1)...(\mu-1)\mu(\mu+1)...(\mu+h-1)(\mu+h) \end{aligned}$$

Now multiply pairs of the coefficients in the product on the right hand side: $(\mu-h)(\mu+h)$,..., $(\mu-1)(\mu+1)$. We see that there are h of these new coefficients $(\mu-(h-i))(\mu+(h-i)) = (\mu^2-(h-i)^2)$, and certainly we have the inequality $(\mu^2-(h-i)^2) \leq \mu^2$. So we see that the inequality does hold for $h \geq 0$ since

$$\begin{aligned} \mu^{2h+1} &\geq (\mu^2-h^2)(\mu^2-(h-1)^2)...(\mu^2-1)\mu \\ \mu^{2h} &\geq (\mu^2-h^2)(\mu^2-(h-1)^2)...(\mu^2-1). \end{aligned}$$

(b) Using part (a), we can see that $\mathbb{P}[Z \geq \mu] \geq 1/2$.

$$\begin{aligned}
\mathbb{P}[Z \geq \mu] &= \sum_{h=\mu}^{\infty} \mathbb{P}[Z = h] \\
&= \sum_{h=0}^{\infty} \mathbb{P}[Z = \mu + h] \\
&\geq \sum_{h=0}^{\mu-1} \mathbb{P}[Z = \mu + h] \\
&\geq \sum_{h=0}^{\mu-1} \mathbb{P}[Z = \mu - h - 1] \qquad \text{by part (a)} \\
&= \mathbb{P}[Z < \mu]
\end{aligned}$$

Since $\mathbb{P}[Z < \mu] + \mathbb{P}[Z \geq \mu] = 1$ and $\mathbb{P}[Z \geq \mu] \geq \mathbb{P}[Z < \mu]$ from above, we have $\mathbb{P}[Z \geq \mu] \geq 1/2$.

6. **(MU 5.15)**

We consider another way to obtain Chernoff-like bounds in the setting of balls and bins without using Theorem 5.7. Consider n balls thrown randomly into n bins. Let $X_i = 1$ if the ith bin is empty and 0 otherwise. Let $X = \sum_{i=1}^{n} X_i$. Let Y_i, $i = 1, ..., n$, be independent Bernoulli random variables that are 1 with probability $p = (1 - 1/n)^n$. Let $Y = \sum_{i=1}^{n} Y_i$.

(a) Show that $\mathbb{E}[X_1 X_2 \cdots X_k] \leq \mathbb{E}[Y_1 Y_2 \cdots Y_k]$ for any $k \geq 1$.

For the X_i's, the product is 1 only if the first k bins are empty. So there are $(n-k)^n$ ways to throw n balls into $n-k$ bins. And there are n^n ways to throw n balls into n bins. So, we have

$$\mathbb{E}[X_1 X_2 \cdots X_k] = \frac{(n-k)^n}{n^n} = (1 - k/n)^n .$$

For the Y_i's, since they are independent, we simply have the product of k binomial random variables with $p = (1 - 1/n)^n$. This is

$$\mathbb{E}[Y_1 Y_2 \cdots Y_k] = (1 - 1/n)^{kn}.$$

From the book, we know that $1 - k/n \leq (1 - 1/n)^k$ for all positive integers n, k. So we have $\mathbb{E}[X_1 X_2 \cdots X_k] \leq \mathbb{E}[Y_1 Y_2 \cdots Y_k]$.

(b) Show that $\mathbb{E}[e^{tX}] \leq \mathbb{E}[e^{tY}]$ for all $t \geq 0$. (Hint: use the expansion for e^x and compare $\mathbb{E}[X^k]$ to $\mathbb{E}[Y^k]$.

Using the Taylor series for e^x and linearity of expectation, we can write the MGF of X and Y as:

$$\mathbb{E}\left[e^{tX}\right] = \sum_{i=0}^{\infty} \frac{t^i}{i!} \mathbb{E}[X^i]$$

$$\mathbb{E}\left[e^{tY}\right] = \sum_{i=0}^{\infty} \frac{t^i}{i!} \mathbb{E}[Y^i]$$

Now, we want to compare term-wise, the MGF of X and the MGF of Y. This means that we want to show $\mathbb{E}[X^i] \leq \mathbb{E}[Y^i]$.

Let $Z = \sum_{i=1}^{n} Z_i$ be an indicator random variable which is the sum of independent and identically distributed indicators Z_i. For example, both X and Y satisfy the properties of Z. Then

$$\mathbb{E}[Z^j] = \mathbb{E}\left[\left(\sum_{i=1}^{n} Z_i\right)^j\right].$$

After expanding out the jth power, we get a sum of terms where each term is a product of various Z_i's to powers C_i, i.e. $Z_1^{c_i} Z_2^{c_2} \cdots Z_n^{c_n}$. Since the Z_i are indicators, each term simplifies to

$$Z_1^{c_i} Z_2^{c_2} \cdots Z_n^{c_n} = \prod_{i=1}^{n} Z_i^{I\{c_i \neq 0\}} = \prod_{i=1}^{k} Z_i,$$

where $I\{c_i \neq 0\}$ indicates the event that c_i is positive, and $k = \sum_{i=1}^{n} I\{c_i\ ne0\}$. Notice that the last equality follows from the Z_i being independent and identically distributed.

Notice from this and by application of part (a) to each term, we can write

$$\mathbb{E}[X^j] = \mathbb{E}\left[\left(\sum_{i=1}^{n} X_i\right)^j\right] = \sum_{c=(c_1,\ldots,c_n)} \mathbb{E}\left[\prod_{i=1}^{n} X_i^{c_i}\right] = \sum \mathbb{E}\left[\prod_{i=1}^{k} X_i\right] \leq \sum \mathbb{E}\left[\prod_{i=1}^{k} Y_i\right] = \mathbb{E}[Y^j].$$

This proves that for a particular term, i, in the MGF of X and the MGF of Y, the inequality holds. Which means we have shown that $\mathbb{E}[e^{tX}] \leq \mathbb{E}[e^{tY}]$.

(c) Derive a Chernoff bound for $\mathbb{P}\left[X \geq (1+\delta)\mathbb{E}[X]\right]$.

Using Markov's inequality and part (b), we get

$$\mathbb{P}[X \geq x] = \mathbb{P}[e^{tX} \geq e^{tx}] \leq \frac{\mathbb{E}[e^{tX}]}{e^{tx}} \leq \frac{\mathbb{E}[e^tY]}{e^{tx}}.$$

The problem ask for us to set $x = (1+\delta)\mathbb{E}[X]$, and we know that the MGF of Y is upper-bounded by $e^{\mathbb{E}[Y](e^t-1)}$ (page 64). Now, we have

$$\mathbb{P}[X \geq (1+\delta)\mathbb{E}[X]] \leq \frac{e^{\mathbb{E}[Y](e^t-1)}}{e^{t(1+\delta)\mathbb{E}[X]}}$$

By the problem set-up, we know that $\mathbb{E}[X] = \mathbb{E}[Y]$, since $\mathbb{E}[X] = \sum \mathbb{E}[X_i] = \sum(1-1/n)^n = \mathbb{E}[Y]$. Because of this, we write

$$\mathbb{P}[X \geq (1+\delta)\mathbb{E}[X]] \leq \left(\frac{e^{(e^t-1)}}{e^{t(1+\delta)}}\right)^{\mathbb{E}[Y]}$$

Setting $t = ln(1+\delta)$ we get

$$\mathbb{P}\left[X \geq (1+\delta)\mathbb{E}[X]\right] \leq \left(\frac{e^{\delta}}{(1+\delta)^{(1+\delta)}}\right)^{\mathbb{E}[Y]}.$$