Randomness and Computation 2018/19
notes for Week 4 tutorial (Tues 5th, Wed 6th February, 2019)

1. **(a), (b)** Imagine taking our biased coin and flipping it *twice*. After doing this we have the possibility of four outcomes: two "heads", "heads"-then-"tails", "tails"-then-"heads", two "tails".

   Now notice that because flips are independent and identical (with the unknown probability p), that the probabilities of these four outcomes are $p^2, p(1-p), (1-p)p$ and $(1-p)^2$ respectively. In particular, "heads"-then-"tails" and "tails"-then-"heads" have identical probability of being generated. We will use this fact to identify "heads"-then-"tails" with the overall outcome "heads" and "tails"-then-"heads" with the overall outcome "tails", these each having identical probability. If the pair of flips generates two "heads" or two "tails", we re-run the experiment with two new flips of the coin.

   Algorithm BIASNOMORE(p)

   (a) flip1 =0 , flip2 =0;
   (b) **while** $\mathsf{flip1} = \mathsf{flip2}$ **do**
   (c) $\quad\mathsf{flip1} \leftarrow \mathsf{B}(1, p)$;
   (d) $\quad\mathsf{flip2} \leftarrow \mathsf{B}(1, p)$;
   (e) **od**
   (f) **if** $\mathsf{flip1} = 1$
   (g) $\quad$**return** "heads"
   (h) **else**
   (i) $\quad$**return** "tails"

   I've already argued that *on any particular two flips*, "heads" and "flips" are equally likely to be returned ($p(1-p)$ each). These is true regardless of whether we take $2, 4, 6, 8, \ldots, 2i$ flips to return a value - the final pair of flips determines what is returned, and "heads" and "tails" are equally likely at that point. Hence the probability, over all possible sequences of flips that end with a returned value, is equal for "heads" and "tails".

   (c) For this bit, we can write the *expected number of coin flips* used to be

$$2\sum_{j=1}^{\infty}(1-2p(1-p))^{j-1}2p(1-p).$$

   The value inside the sum is the geometric distribution with parameter $2p(1-p)$, and therefore its expectation is the inverse of this parameter, which is $\frac{1}{2p(1-p)}$.

   Multiplying-in the 2 from outside the sum, overall the expected number of flips is $\frac{2}{2p(1-p)}$, ie $p^{-1}(1-p)^{-1}$.

2. We start with a bin containing one black ball and one white ball, and repeatedly do the following: choose one ball from the bin uniformly at random, and then put the ball back in the bin with another ball of the same colour. We repeat until there are n balls in the bin.

   Claim: by the time that we have n balls (after $n-2$ steps), the number of white balls is equally likely to be any number between 1 and $n-1$.

   We will prove this by induction on n.

   We should note that no matter what choices are made, we will always have at least one white ball and at least one black ball in the bin.

   base case: $n = 2$. In this case we definitely have 1 white ball in the bin. The range $1, \ldots, n-1$ is just 1, so the hypothesis is trivially correct.

   Induction step: Suppose we have shown the claim for $n = k$ *(Induction Hypothesis (IH))*. We now show it must also hold for $n = k+1$.

   If the claim holds for $n = k$, then when it comes to carry out the following step, we know that we have k balls in the bin, and that $\Pr_{k \text{ balls}}[j \text{ white balls}] = \frac{1}{k-1}$ for every $j = 1 \ldots, k-1$. This is by our (IH).

   Suppose the bin has j white balls before the final step (out of k total balls). If we draw a black ball (with probability $1 - \frac{j}{k}$), then there are still j white balls after this extra step. If we draw a white ball (with probability $\frac{j}{k}$) there are instead $j+1$ white balls after this final step. So for every $j = 2, \ldots, k-1$ (so that both $j-1$ and j are possible white ball counts for the prior step with k balls)

$$\begin{aligned}
\Pr_{k \text{ balls}}[j \text{ white balls}] &= (1 - \tfrac{j}{k}) \Pr_{k \text{ balls}}[j \text{ white balls}] + \tfrac{j-1}{k} \Pr_{k \text{ balls}}[j-1 \text{ white balls}]. \\
&= (1 - \tfrac{j}{k})\tfrac{1}{k-1} + \tfrac{j-1}{k}\tfrac{1}{k-1} \qquad \textit{by the (IH) for } k-1 \\
&= \frac{1}{k-1} + \frac{1}{k(k-1)}(j-1-j) \\
&= \frac{1}{k-1} - \frac{1}{k(k-1)} \\
&= \frac{1}{k-1}\frac{k-1}{k} \\
&= \frac{1}{k},
\end{aligned}$$

   as required.

   For the case of $j = 1$ for $k+1$ balls, the only way we can achieve this is if we previously had a single white ball in the (k balls) bin, and we drew a black ball on the final step; the probability of this happening is

$$\frac{k-1}{k} \Pr_{k \text{ balls}}[1 \text{ white ball}],$$

   which is $\frac{k-1}{k}\frac{1}{k-1} = \frac{1}{k}$.

For the case of $j = k$ for $k = 1$ balls, a similar argument to the $j = 1$ case will work, or alternatively we can just note that the remaining probability will be

$$1 - \sum_{j=1}^{k-1} \Pr_{k+1 \text{ balls}}[j \text{ white balls}],$$

which is $1 - (k-1)\frac{1}{k}$, which is $\frac{1}{k}$.

3. (a) One run of the algorithm performs $n-2$ edge contractions. Therefore, we have $2(n-2)$ edge contractions in two runs.

   (b) We have $n-k$ edge contractions in the first phase. Subsequently, in each of the l runs we have $k-2$ edge contractions. Hence, we have a total of $n-k+l(k-2)$ edge contractions.

   (c) Let C be a minimum cutset of G. We will give a lower bound on the probability that our modified Karger's-min-cut algorithm returns C. Clearly, no edge in C can be contracted in the first $n-k$ edge contractions if the algorithm is to have any chance to output C at the end. Therefore, we have that

$$\begin{aligned}
\Pr[\text{no edge in C contracted in 2nd phase}] &\geq (1-\frac{2}{n})(1-\frac{2}{n-1})\dots(1-\frac{2}{n-(n-(k+1))}) \\
&= \frac{n-2}{n}\cdot\frac{n-3}{n-1}\cdots\frac{k-1}{k+1} \\
&= \frac{k(k-1)}{n(n-1)}
\end{aligned}$$

   Given that no edge in C is contracted in the first phase, the algorithm will output C if at least one of the l runs in the second phase produces C. We have

$$\begin{aligned}
&\Pr[\text{2nd phase outputs C} \mid \text{no C edge contracted in 1st phase}] \\
&= 1 - \Pr[\text{no run produces C} \mid \text{no C edge contracted in 1st phase}] \\
&\geq 1-(1-\tfrac{2}{k(k-1)})^{l}
\end{aligned}$$

   Therefore, putting it all together we have

$$\begin{aligned}
\Pr[\text{algorithm outputs C}] &= \Pr[\text{2nd phase outputs C} \mid \text{no C edge contracted in 1st phase}] \\
&\quad * \Pr[\text{no C edge contracted in 1st phase}] \\
&\geq \left(1-\left(1-\frac{2}{k(k-1)}\right)^{l}\right)\frac{k(k-1)}{n(n-1)}
\end{aligned}$$

   (d) If we want to do the same number of edge contractions with our modified algorithm as we would do with just two runs of original algorithm, then we must have

$$n-k+l(k-2) = 2(n-2)$$

Solving for l we get

$$l = \frac{n+k-4}{k-2}$$

We want to maximise

$$\left(1-(1-\frac{2}{k(k-1)})^{l}\right)\frac{k(k-1)}{n(n-1)} = \left(1-(1-\frac{2}{k(k-1)})^{\frac{n+k-4}{k-2}}\right)\frac{k(k-1)}{n(n-1)}$$

with respect to k such that $2 \leq k \leq n$.

Then things become messy: we can do some transformations using $1-x \leq e^{-x}$. The idea is to use $1-x \leq e^{-x}$ where $x = \ 1-\frac{2}{k(k-1)}$, and also probably use the fact that $\frac{k(k-1)}{n(n-1)} \geq \frac{k^2}{n^2}$ for $k \leq n$, so that we have

$$\text{[the probability expression to maximise]} \geq (1-e^{-\frac{2(n+k-4)}{k(k-1)(k-2)}}) \cdot \frac{k^2}{n^2}.$$

And then maybe take derivatives, etc, to "finish off" (where the the finishing of is probably a few pages.....) by maximising the new expression on the left hand side of the last inequality. In any case, without doing all that, at this stage it already seems that:
1) k should not be a constant, because otherwise the factor k^2/n^2 already gives probabability only $O(1/n^2)$, which is the same as the original Karger's, and

2) we don't want $k >> n^{1/3}$, because then the factor $(1-e^{-\frac{2(n+k-4)}{k(k-1)(k-2)}})$ would be going to 0 fast enough to dwarf the 'gains' in $\frac{k^2}{n^2}$.

3) However, if we take $k \sim (2n)^{1/3}$, then the exponent in $e^{-\frac{2(n+k-4)}{k(k-1)(k-2)}}$ will become approximately -1, which will make $(1-e^{-\frac{2(n+k-4)}{k(k-1)(k-2)}})$ approximately $1-e^{-1}$, with the $\frac{k^2}{n^2}$ term becoming about $\frac{2^{2/3}}{n^{4/3}}$, a considerable improvement over the standard "2 runs of Karger" result.

4. Let Y be a nonnegative integer-valued random variable with (strictly) positive expectation. Prove that

$$\frac{(\mathrm{E}[Y])^2}{\mathrm{E}[Y^2]} \ \leq \ \Pr[Y \neq 0] \ \leq \mathrm{E}[Y].$$

Proof: First let's do the **right-hand side**. For this, notice that by Y's range being non-negative and integer, we know

$$\begin{aligned}
E[Y] &= \sum_{j=0}^{\infty} j \cdot \Pr[Y = j] \\
&= \sum_{j=1}^{\infty} j \cdot \Pr[Y = j] \\
&\geq 1 \cdot \sum_{j=1}^{\infty} \Pr[Y = j], \\
&= \Pr[Y \geq 1] \;=\; \Pr[Y \neq 0],
\end{aligned}$$

where the first step (expansion of $E[Y]$) and final step (equality of $\Pr[Y \geq 1]$ and $\Pr[Y \neq 0]$) both follow from the fact that Y only takes on non-negative integers.

For the **left-hand side**, we have two ways of proving it:

method 1: Consider the conditional expectations $E[Y \mid Y \neq 0]$ and $E[Y^2 \mid Y \neq 0]$. Note that the function $f(x) = x^2$ is convex. Therefore, by Jensen's inequality we know that

$$(E[Y \mid Y \neq 0])^2 \leq E[Y^2 \mid Y \neq 0]$$

We have that

$$\begin{aligned}
E[Y \mid Y \neq 0] &= \sum_{j=0}^{\infty} j \cdot \frac{\Pr[Y = j, Y \neq 0]}{\Pr[Y \neq 0]} \\
&= \frac{1}{\Pr[Y \neq 0]} \sum_{j=0}^{\infty} j \cdot \Pr[Y = j, Y \neq 0] \\
&= \frac{1}{\Pr[Y \neq 0]} \sum_{j=1}^{\infty} j \cdot \Pr[Y = j] \\
&= \frac{1}{\Pr[Y \neq 0]} E[Y]
\end{aligned}$$

Similarly, we have

$$\begin{aligned}
E[Y^2 \mid Y \neq 0] &= \sum_{j=0}^{\infty} j^2 \cdot \frac{\Pr[Y = j, Y \neq 0]}{\Pr[Y \neq 0]} \\
&= \frac{1}{\Pr[Y \neq 0]} \sum_{j=0}^{\infty} j^2 \cdot \Pr[Y = j, Y \neq 0] \\
&= \frac{1}{\Pr[Y \neq 0]} \sum_{j=1}^{\infty} j^2 \cdot \Pr[Y = j] \\
&= \frac{1}{\Pr[Y \neq 0]} E[Y^2]
\end{aligned}$$

Therefore, we know that

$$\left(\frac{1}{\Pr[Y \neq 0]} \mathrm{E}[Y]\right)^2 \leq \frac{1}{\Pr[Y \neq 0]} \mathrm{E}[Y^2]$$

Multiplying both sides by $\frac{(\Pr[Y \neq 0])^2}{\mathrm{E}[Y^2]}$ we get

$$\frac{(\mathrm{E}[Y])^2}{\mathrm{E}[Y^2]} \leq \Pr[Y \neq 0]$$

as claimed.

method 2: The alternative way of proving the **left-hand side** is to show it directly, by first expanding out the terms $\Pr[Y \neq 0], \mathrm{E}[Y^2]$ and $\mathrm{E}[Y]^2$. We have

$$\begin{aligned}
\mathrm{E}[Y^2] &= \sum_{j=1}^{\infty} j^2 \Pr[Y = j], \quad \text{and} \\
\mathrm{E}[Y]^2 &= \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} ij \Pr[Y = i] \Pr[Y = j], \quad \text{and} \\
\Pr[Y \neq 0] &= \sum_{j=1}^{\infty} \Pr[Y = j].
\end{aligned}$$

We will calculate $\mathrm{E}[Y^2] \Pr[Y \neq 0]$ as:

$$\begin{aligned}
&\left(\sum_{i=1}^{\infty} i^2 \Pr[Y = i]\right) \cdot \left(\sum_{j=1}^{\infty} \Pr[Y = j]\right) \\
= &\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} i^2 \cdot \Pr[Y = i] \Pr[Y = j]
\end{aligned}$$

Now, proving the left-hand side is equivalent to proving $\mathrm{E}[Y]^2 \leq \mathrm{E}[Y^2] \cdot \Pr[Y \neq 0]$, and (with the expanded form) this happens if and only if we have

$$\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} ij \Pr[Y = i] \Pr[Y = j] \quad \leq \quad \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} i^2 \cdot \Pr[Y = i] \Pr[Y = j] \tag{1}$$

Let's pair up the terms as follows:

- First, for any $h \in \mathbb{N}$, if we take h in *both* sums on the lhs of(1) ($i \leftarrow h, j \leftarrow h$) we get $h^2 \Pr[Y = h] \Pr[Y = h]$. Taking $i \leftarrow h, j \leftarrow h$ on the rhs, we get $h^2 \Pr[Y = h] \Pr[Y = h]$ also. So they match, and don't affect the $\leq$.

- Second, for any h, k with $h \neq k$, we will see $2hk \Pr[Y = h] \Pr[Y = k]$ on the lhs of the inequality (1) (because we may take $i \leftarrow h$ in the first sum and $j \leftarrow k$ in the second, but may also take $i \leftarrow k, j \leftarrow h$).

  On the rhs of (1), if we take $i \leftarrow h, j \leftarrow k$ we get the value $h^2 \Pr[Y = h] \Pr[Y = k]$ and if we take $i \leftarrow k, j \leftarrow h$ we get the value $k^2 \Pr[Y = h] \Pr[Y = k]$.

  Now, recall that for any positive values h, k, we always have

$$hk \leq \frac{h^2 + k^2}{2},$$

  hence $2hk \leq h^2 + k^2$, and multiplying across by the non-negative value $\Pr[Y = h] \Pr[Y = k]$, we see

$$2hk \cdot \Pr[Y = h] \Pr[Y = k] \leq (h^2 + k^2) \cdot \Pr[Y = h] \Pr[Y = k].$$

  Therefore (1)'s lhs-value for $h, k, h \neq k$ is always $\leq$ than (1)'s rhs-value for $h, k, h \neq k$.

Putting all these together for the various h, k pairs, we see that we must have

$$\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} ij \Pr[Y = i] \Pr[Y = j] \quad \leq \quad \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} i^2 \cdot \Pr[Y = i] \Pr[Y = j]$$

Mary Cryan