CS174 Lecture Note 4

Based on notes by Alistair Sinclair, September 1998; based on earlier notes by Manuel Blum/Douglas Young.

More on random permutations

We might ask more detailed questions, such as:

Q3: What is the probability that π contains at least one 1-cycle (cycle of length 1)?

Q4: What is the distribution of the number of 1-cycles?

Before we can answer these questions, we need to recall the inclusion exclusion principle. The version we use is adapted to probabilities. Suppose we start with n properties (events) $E_1, \ldots, E_n$. First define $p_i = \Pr[E_i]$ and $p_{ij} = \Pr[E_i \wedge E_j]$ and $p_{ijk} = \Pr[E_i \wedge E_j \wedge E_k]$ and so on. (The indices i, j, k here are assumed to be distinct.) Now we define sums S_i as

$$S_1 = \sum_{i=1}^{n} p_i \qquad S_2 = \sum_{1 \le i < j \le n} p_{ij} \qquad S_3 = \sum_{1 \le i < j < k \le n} p_{ijk} \cdots$$

The following theorem, known as the Principle of Inclusion/Exclusion, expresses $\Pr[E_1 \vee \ldots \vee E_n]$ in terms of the easier-to-compute S_k.

Theorem 1: $\Pr[E_1 \vee E_2 \vee \ldots \vee E_n] = S_1 - S_2 + S_3 - S_4 + \cdots \pm S_n$.

Proof: Let s be any sample point in $E_1 \vee \ldots \vee E_n$. How often is it counted on the right-hand-side? Suppose s occurs in exactly r of the E_i. Then it appears r times in S_1, $\binom{r}{2}$ times in S_2, $\binom{r}{3}$ times in S_3, and so on. (Why?) So the contribution of $\Pr[s]$ to the r.h.s. is

$$\Pr[s] \left\{ \binom{r}{1} - \binom{r}{2} + \binom{r}{3} - \cdots \pm \binom{r}{r} \right\}. \qquad (**)$$

But now if we look at the binomial expansion of $(1-x)^r$ we see

$$0 = (1-1)^r = 1 - \binom{r}{1} + \binom{r}{2} - \binom{r}{3} + \cdots \pm \binom{r}{r},$$

so the term in braces in $(**)$ is exactly 1. Thus s contributes exactly $\Pr[s]$ to the r.h.s., which proves the theorem. $\square$

Now we return to Q3. Let E_i be the event that π maps i to itself. Q3 asks for $\Pr[E_1 \vee E_2 \vee \ldots \vee E_n]$. This seems hard to compute ...

What probabilities *can* we compute easily? We have

$$p_i = \tfrac{(n-1)!}{n!} = \tfrac{1}{n}; \qquad p_{ij} = \tfrac{(n-2)!}{n!} = \tfrac{1}{n(n-1)}; \qquad p_{ijk} = \tfrac{(n-3)!}{n!};$$

and so on. (Check this!) So we get $S_1 = n \cdot \frac{1}{n} = 1$; $S_2 = \binom{n}{2} \cdot \frac{1}{n(n-1)} = \frac{1}{2}$; and generally

$$S_k = \binom{n}{k} \cdot \tfrac{(n-k)!}{n!} = \tfrac{n!}{k!(n-k)!} \cdot \tfrac{(n-k)!}{n!} = \tfrac{1}{k!}. \qquad (*)$$

We can now answer our Q3 about random permutations. From Theorem 1, and using the values $S_k = \frac{1}{k!}$ from $(*)$, we get:

$$\Pr[\pi \text{ contains at least one 1-cycle}] = 1 - \tfrac{1}{2!} + \tfrac{1}{3!} - \tfrac{1}{4!} + \cdots \pm \tfrac{1}{n!} \sim 1 - \mathrm{e}^{-1} = 0.632\ldots$$

Ex: How good is this last approximation for $n = 6$? $\square$

Now let's think about Q4. For a family of events $\{E_i\}$, define

$$q_k = \Pr[\text{exactly } k \text{ of the } E_i \text{ occur}].$$

To compute this, we first need a generalization of Theorem 1:

Theorem 1': $\Pr[\text{at least } k \text{ of the } E_i \text{ occur}] = S_k - \binom{k}{k-1}S_{k+1} + \binom{k+1}{k-1}S_{k+2} - \binom{k+2}{k-1}S_{k+3} + \cdots \pm \binom{n-1}{k-1}S_n$. $\square$

Ex: verify that Theorem 1 is a special case of Theorem 1', and (harder!) prove Theorem 1'. $\square$

From Theorem 1', we can easily deduce:

Theorem 2: $q_k = S_k - \binom{k+1}{k}S_{k+1} + \binom{k+2}{k}S_{k+2} - \binom{k+3}{k}S_{k+3} + \cdots \pm \binom{n}{k}S_n$.

Proof: From the definition of q_k, we have

$$q_k = \Pr[\text{at least } k \text{ of the } E_i \text{ occur}] - \Pr[\text{at least } k+1 \text{ of the } E_i \text{ occur}].$$

From Theorem 1', the coefficient of S_{k+i} in the difference of these two series (neglecting the sign) is

$$\binom{k+i-1}{k-1} + \binom{k+i-1}{k} = \tfrac{(k+i-1)!}{(k-1)!i!} + \tfrac{(k+i-1)!}{k!(i-1)!} = \tfrac{(k+i-1)!(k+i)}{k!i!} = \binom{k+i}{k}.$$

Since the signs alternate, this gives us exactly the series claimed. $\square$

Going back to the special case of random permutations, recall from $(*)$ that $S_k = \frac{1}{k!}$, so Theorem 2 gives us:

$$\begin{aligned}
q_0 &= 1 - 1 + \tfrac{1}{2!} - \tfrac{1}{3!} + \cdots \pm \tfrac{1}{n!} \\
q_1 &= 1 - 1 + \tfrac{1}{2!} - \tfrac{1}{3!} + \cdots \mp \tfrac{1}{(n-1)!} \\
q_2 &= \tfrac{1}{2!}\left\{1 - 1 + \tfrac{1}{2!} - \tfrac{1}{3!} + \cdots \pm \tfrac{1}{(n-2)!}\right\} \\
q_3 &= \tfrac{1}{3!}\left\{1 - 1 + \tfrac{1}{2!} - \tfrac{1}{3!} + \cdots \mp \tfrac{1}{(n-3)!}\right\} \\
&\vdots \\
q_{n-2} &= \tfrac{1}{(n-2)!}\left\{1 - 1 + \tfrac{1}{2!}\right\} \\
q_{n-1} &= \tfrac{1}{(n-1)!}\left\{1 - 1\right\} = 0 \\
q_n &= \tfrac{1}{n!}.
\end{aligned}$$

Ex: Give simple arguments to explain why $q_{n-1} = 0$ and $q_n = \frac{1}{n!}$. $\square$

Thus we see that, for every fixed k, $q_k \sim \frac{1}{k!}\mathrm{e}^{-1}$.

The probabilities $\{\frac{1}{k!}\mathrm{e}^{-1}\}$ play a special role: they define the *Poisson distribution* (with parameter 1).

Definition: A r.v. X has the Poisson distribution with parameter λ if

$$\Pr[X = k] = \mathrm{e}^{-\lambda}\tfrac{\lambda^k}{k!} \qquad \text{for all integers } k \geq 0$$

(and $\Pr[X = x] = 0$ for all other values of x). $\square$

Ex: Check that this *is* always a probability distribution, i.e., that $\sum\limits_{k=0}^{\infty} \mathrm{e}^{-\lambda}\frac{\lambda^k}{k!} = 1$. $\square$

So we see that, as $n \to \infty$, the distribution of the number of 1-cycles in a random permutation on n elements behaves like the Poisson distribution with $\lambda = 1$.

Ex: For $n = 10$, compute the q_k exactly and compare them with the approximate values $\frac{1}{k!}\mathrm{e}^{-1}$. How good is the approximation? $\square$

Mean and Variance for a Poisson R.V. For a Poisson R.V. X, the expected value is

$$\mathrm{E}\left(X\right) = \sum_{k=0}^{\infty} k\frac{e^{-\lambda}\lambda^k}{k!} = \lambda\sum_{k=1}^{\infty}\frac{e^{-\lambda}\lambda^{k-1}}{(k-1)!}$$

and substituting $l = k - 1$ gives

$$\mathrm{E}\left(X\right) = \lambda\sum_{l=0}^{\infty}\frac{e^{-\lambda}\lambda^l}{l!} = \lambda e^{-\lambda}e^{\lambda} = \lambda$$

So a random Poisson variable X always has $\mathrm{E}\left(X\right) = \lambda$. The variance of a random variable is defined as

$$\mathrm{Var}\left(X\right) = \mathrm{E}\left(\left(X - \mathrm{E}\left(X\right)\right)^2\right)$$

and its not hard to show that this simplifies to $\mathrm{Var}\left(X\right) = \mathrm{E}\left(X^2\right) - \mathrm{E}\left(X\right)^2$. We know that $\mathrm{E}\left(X\right) = \lambda$, so lets compute $\mathrm{E}\left(X^2\right)$:

$$\mathrm{E}\left(X^2\right) = \sum_{k=0}^{\infty} k^2\frac{e^{-\lambda}\lambda^k}{k!} = \sum_{k=1}^{\infty} k\frac{e^{-\lambda}\lambda^k}{(k-1)!} = \sum_{k=1}^{\infty}(k-1)\frac{e^{-\lambda}\lambda^k}{(k-1)!} + \sum_{k=1}^{\infty}\frac{e^{-\lambda}\lambda^k}{(k-1)!}$$

After cancelling and substituting $i = k - 2$, $j = k - 1$, the last two sums become

$$\mathrm{E}\left(X^2\right) = \lambda^2\sum_{i=0}^{\infty}\frac{e^{-\lambda}\lambda^i}{i!} + \lambda\sum_{j=0}^{\infty}\frac{e^{-\lambda}\lambda^j}{j!} = \lambda^2 + \lambda$$

and finally

$$\mathrm{Var}\left(X\right) = \mathrm{E}\left(X^2\right) - \mathrm{E}\left(X\right)^2 = (\lambda + \lambda^2) - (\lambda)^2 = \lambda$$

so we have the surprising result that the mean *and* variance for a Poisson distribution is λ:

$$\mathrm{E}\left(X\right) = \mathrm{Var}\left(X\right) = \lambda$$

The Poisson distribution shows up naturally in many contexts. Here is another example, which also introduces another important distribution, the *binomial distribution*.

Bernoulli trials

A coin comes up heads with probability p, tails with probability $1 - p$.

- Suppose it is tossed n times. What is $\Pr[\text{exactly } k \text{ heads}]$?

This question arises very frequently in applications in Computer Science. In place of coin flips, we can think of a sequence of n identical independent trials, each of which succeeds (heads) with probability p. It is also a special case of Theorem 2 above, where E_i is the event "the ith toss is heads": the difference here is that *the events E_i are now independent*, so things are now much simpler.

Define the r.v. X = # heads in above experiment.

Ex: By writing $X = \sum_i X_i$ for suitable indicator r.v.'s X_i, show that $\mathrm{E}(X) = np$ and $\mathrm{Var}(X) = np(1-p)$. $\square$

What does the distribution of X look like? Well, consider any outcome of the experiment in which $X = k$, i.e., in which there are exactly k heads. We can view this as a string $s \in \{\mathrm{H}, \mathrm{T}\}^n$ containing k H's and $n-k$ T's. Now since all coin tosses are independent, we must have $\Pr[s] = p^k(1-p)^{n-k}$. The number of such strings s is $\binom{n}{k}$. Summing over sample points in the event "$X = k$" gives

$$\Pr[X = k] = \binom{n}{k} p^k (1-p)^{n-k}.$$

Definition: The above distribution is known as the binomial distribution with parameters n and p.

Examples

1. The probability of exactly k heads in n tosses of a fair coin is $\binom{n}{k} 2^{-n}$.
2. When we toss m balls into n bins, the probability that any given bin (say, bin i) contains exactly k balls is $\binom{m}{k}(\frac{1}{n})^k(1-\frac{1}{n})^{m-k}$.

We'll have a lot more to say about the binomial distribution later. Here, we just consider a special case in which $p = \lambda/n$ for some constant λ. Note that this means that $\mathrm{E}(X) = np = \lambda$ remains constant as $n \to \infty$.

Writing $q_k = \Pr[X = k]$, we have

$$q_0 = (1-p)^n = (1 - \tfrac{\lambda}{n})^n \sim \mathrm{e}^{-\lambda} \qquad \text{as } n \to \infty.$$

Also,

$$\frac{q_k}{q_{k-1}} = \frac{\binom{n}{k} p^k (1-p)^{n-k}}{\binom{n}{k-1} p^{k-1} (1-p)^{n-k+1}} = \frac{n-k+1}{k} \cdot \frac{p}{1-p} = \frac{n-k+1}{k} \cdot \frac{\lambda}{n-\lambda}.$$

For any fixed k, we therefore have $\frac{q_k}{q_{k-1}} \sim \frac{\lambda}{k}$ as $n \to \infty$. So we get

$$\begin{aligned} q_1 &\sim \lambda q_0 \sim \lambda \mathrm{e}^{-\lambda} \\ q_2 &\sim \tfrac{\lambda}{2} q_1 \sim \tfrac{\lambda^2}{2!} \mathrm{e}^{-\lambda} \\ &\vdots \\ q_k &\sim \tfrac{\lambda}{k} q_{k-1} \sim \tfrac{\lambda^k}{k!} \mathrm{e}^{-\lambda}. \end{aligned}$$

Once again, we get the Poisson distribution, this time with parameter $\lambda = np$.

Example: Suppose we toss $m = cn$ balls into n bins, where c is a constant. Then for any fixed k,

$$\Pr[\text{bin } i \text{ contains exactly } k \text{ balls}] \sim \tfrac{c^k}{k!} \mathrm{e}^{-c}. \qquad \square$$