Notes by Manuel Blum/Douglas Young/Alistair Sinclair.

Lecture Note 2

Random permutations

A permutation on n elements is a 1-1 function from the elements to themselves.

Generate a random permutation on n elements, all equally likely. (Recall that any permutation may be uniquely represented as a collection of cycles.)

Q1: What is the expected number of cycles of length 1?

Q2: What is the expected total number of cycles?

Q3: What is the probability that the permutation is a single cycle?

Sample space: all $n!$ possible permutations, each with probability $\frac{1}{n!}$.

Answer to Q3

Let π be a random permutation (sample point). Then

$$\Pr[\pi \text{ is a single cycle}] = \#\text{ cycles on } n \text{ elements} \times \tfrac{1}{n!} = \tfrac{(n-1)!}{n!} = \tfrac{1}{n}.$$

Answer to Q1

First attempt: define a random variable

$$X = \#\text{ cycles of length 1 in } \pi,$$

(More formally, X is a function that maps sample point π to the number of cycles of length 1 in π.)
Then we could try to compute $\mathrm{E}(X)$ by figuring out $\Pr[X = k]$ for each k.

Ex: Use this approach to compute $\mathrm{E}(X)$ in the cases $n = 2$ and $n = 3$. $\square$

Unfortunately, the probabilities $\Pr[X = k]$ are not so simple to write down. (Have a go at doing this!) But we can get over this problem by using a different r.v. In fact, we use a family of n r.v.'s:

$$X_i = \begin{cases} 1 & \text{if } \pi \text{ maps element } i \text{ to itself;} \\ 0 & \text{otherwise.} \end{cases}$$

Then $X = \sum_{i=1}^{n} X_i$.

Now $\mathrm{E}(X_i) = \Pr[X_i = 1] = \frac{1}{n}$. (Do you believe this? Why?)
And now

$$\mathrm{E}(X) = \mathrm{E}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathrm{E}(X_i) = \sum_{i=1}^{n} \frac{1}{n} = 1.$$

Surprised? You should be: this says that $\mathrm{E}(X)$ is *independent of* n!!

Ex: After any such calculation, and particularly when the answer is surprising, you should always do a sanity check by verifying the answer for some values of n. Is the above answer in line with your solutions to the previous exercise? $\square$

Answer to Q2

First attempt: define r.v. Y = total number of cycles in π.

Ex: Compute $\mathrm{E}(Y)$ in the cases $n = 2$ and $n = 3$. $\square$

Once again, this approach is not good because it is messy to figure out the distribution of Y. Here is a cleaner approach. For $1 \le i \le n$, define the r.v.

$$Y_i = 1/(\text{length of cycle containing } i).$$

Then $\sum_{i=1}^{n} Y_i = \#$ cycles in π. (Why?) Therefore $\mathrm{E}(Y) = \mathrm{E}(\sum_i Y_i) = \sum_i \mathrm{E}(Y_i)$.
And what is $\mathrm{E}(Y_i)$? Well,

$$\begin{aligned}
\mathrm{E}(Y_i) &= \sum_{k=1}^{n} \tfrac{1}{k} \cdot \Pr\left[Y_i = \tfrac{1}{k}\right] \\
&= 1 \cdot \Pr[Y_i = 1] + \tfrac{1}{2} \cdot \Pr\left[Y_i = \tfrac{1}{2}\right] + \tfrac{1}{3} \cdot \Pr\left[Y_i = \tfrac{1}{3}\right] + \cdots \\
&= 1 \cdot \tfrac{1}{n} + \tfrac{1}{2} \cdot \tfrac{1}{n} + \tfrac{1}{3} \cdot \tfrac{1}{n} + \cdots \\
&= \tfrac{1}{n}\left(1 + \tfrac{1}{2} + \tfrac{1}{3} + \cdots + \tfrac{1}{n}\right) \\
&\sim \tfrac{1}{n}(\ln n + \gamma),
\end{aligned}$$

where $\gamma = 0.5772\ldots$ is *Euler's constant.*

You should think *carefully* about the third line in the above derivation. Here is an explanation:

$$\Pr[Y_i = \tfrac{1}{2}] = \Pr[\text{for some } j \neq i,\ \pi \text{ maps } i \text{ to } j \text{ and } j \text{ to } i] = (n-1)(n-2)! \times \tfrac{1}{n!} = \tfrac{1}{n},$$

and similarly for other values of k.
So finally we have

$$\mathrm{E}(Y) = \sum_{i=1}^{n} \mathrm{E}(Y_i) = \ln n + \gamma.$$

I.e., *the expected number of cycles in a random permutation on n elements as $n \to \infty$ is* $\ln n +$ constant.

How do we generate a random permutation?

Choose a destination $\pi(1)$ for 1 uniformly at random (u.a.r.) from $\{1, \ldots, n\}$, then a destination $\pi(2)$ for 2 u.a.r. from $\{1, \ldots, n\} - \{\pi(1)\}$, then a destination $\pi(3)$ for 3 u.a.r. from $\{1, \ldots, n\} - \{\pi(1), \pi(2)\}$, and so on.
So we have a sequence of n experiments (or *trials*). Are they independent?
Apparently not: e.g., $\Pr[\pi(1) = i \wedge \pi(2) = i] = 0$!!??

Correct view: sample space is really of the form $\mathcal{S} = \mathcal{S}_1 \times \mathcal{S}_2 \times \cdots \times \mathcal{S}_n$, where $\mathcal{S}_i = \{1, \ldots, n-i+1\}$ (outcomes of ith trial). And for each sample point $(x_1, \ldots, x_n)$ we have $\Pr[(x_1, \ldots, x_n)] = \Pr_{\mathcal{S}_1}[x_1] \times \Pr_{\mathcal{S}_2}[x_2] \times \cdots \times \Pr_{\mathcal{S}_n}[x_n] = \frac{1}{n!}$.

Ex: Check that $\mathcal{S}$ is the correct sample space by showing precisely how each sample point $(x_1, \ldots, x_n)$ corresponds to a *distinct* permutation. $\square$

In fact, we can use the same sample space $\mathcal{S}$ together with *any* method that associates each point with a distinct permutation. E.g.:

Ex: Let i be an arbitrary element. Choose $\pi(i)$ u.a.r. from $\{1, \ldots, n\}$, then $\pi(\pi(i))$ u.a.r. from the remaining possibilities, then $\pi(\pi(\pi(i)))$, and so on until the cycle is closed. Then take some arbitrary remaining j and continue until all n elements have been mapped. Justify this method by showing (carefully) how each point in $\mathcal{S}$ corresponds to a distinct permutation. $\square$

Independent trials give a different (often easier) way of calculating probabilities. E.g.:

$$\Pr[j \text{ is a fixed point}] = \tfrac{1}{n};$$
$$\Pr[\text{both } i \text{ and } j \text{ are fixed points}] = \tfrac{1}{n} \times \tfrac{1}{n-1} = \tfrac{1}{n(n-1)};$$
$$\Pr[i \text{ is in a cycle of length } k] = \tfrac{n-1}{n} \times \tfrac{n-2}{n-1} \times \ldots \times \tfrac{n-k+1}{n-k+2} \times \tfrac{1}{n-k+1} = \tfrac{1}{n}.$$

We can also view the balls-and-bins example of Note 1 as a sequence of independent trials: $\mathcal{S} = \mathcal{S}_1 \times \cdots \times \mathcal{S}_m$, where each $\mathcal{S}_i = \{1, \ldots, n\}$, the possible choices of bin for ball i. Then

$$\Pr[\text{bin } i \text{ is empty}] = \textstyle\prod_{j=1}^m \Pr[\text{ball } j \text{ misses bin } i] = (1 - \tfrac{1}{n})^m;$$
$$\Pr[\text{balls 1 and 2 land in first bin}] = \tfrac{1}{n} \times \tfrac{1}{n} = \tfrac{1}{n^2};$$
$$\Pr[\text{balls 1 and 2 land in same bin}] = \tfrac{1}{n}.$$

Aside: Two completely different methods for generating a random permutation

1. Select n random numbers $\{a_i\}_{i=1}^n$ uniformly from the interval $[0, 1]$. Sort the numbers. The sorted indices form a random permutation.

   Note: In practice, we would use a random number generator with limited precision, so instead of selecting numbers from the interval $[0, 1]$ we would be selecting them from the set $\{\frac{i}{2^k} : i = 0, 1, \ldots, 2^k - 1\}$, where k is the number of bits of precision. For the method to work, we require that all selected numbers be different. How large does k have to be (as a function of n) to ensure that this happens with high probability? (Hint: recall the birthday problem.)

2. Starting with the order $1, \ldots, n$, do the following "many times": pick two elements at random and interchange them. Output the final ordering. (Is it random? How many times is enough?)

Variance

The expectation gives only very partial information about a r.v. X. More information can be obtained from the *variance*, which measures how much X is expected to deviate from $\mathrm{E}(X)$.

Definition: For a r.v. X with expectation $\mu = \mathrm{E}(X)$, the <u>variance</u> of X is $\mathrm{Var}(X) = \mathrm{E}\left((X-\mu)^2\right)$. The <u>standard deviation</u> of X is $\sqrt{\mathrm{Var}(X)}$. $\square$

Example

Roll a single die; let r.v. X = number of pips.

$\mu = \mathrm{E}(X) = \sum_k \Pr[X=k]\cdot k = \frac{1}{6}\{1+2+\cdots+6\} = 3.5$

$\sigma^2 = \mathrm{Var}(X) = \sum_k \Pr[X=k]\cdot(k-\mu)^2 = \frac{1}{6}\left\{(1-3.5)^2+(2-3.5)^2+\cdots+(6-3.5)^2\right\} = \frac{35}{12} \approx 2.92$

$\sigma = \sqrt{2.92} \approx 1.71$ (standard deviation)

Simple Theorem: $\mathrm{Var}(X) = \mathrm{E}\left(X^2\right) - \mu^2$.

Proof:

$$\mathrm{Var}(X) = \mathrm{E}\left((X-\mu)^2\right) = \mathrm{E}\left((X^2-2\mu X+\mu^2)\right) = \mathrm{E}\left(X^2\right) - 2\mu\mathrm{E}(X) + \mu^2 = \mathrm{E}\left(X^2\right) - \mu^2. \qquad \square$$

Another example

Let's look again at the sample space of random permutations on n elements, and the r.v. X = # cycles of length 1. We have seen that $\mathrm{E}(X) = 1$. What is $\mathrm{Var}(X)$?
By the Simple Theorem, $\mathrm{Var}(X) = \mathrm{E}\left(X^2\right) - 1$. To compute $\mathrm{E}\left(X^2\right)$, recall that $X = \sum_i X_i$. Then

$$\mathrm{E}\left(X^2\right) = \mathrm{E}\left(\left(\sum_{i=1}^n X_i\right)^2\right) = \mathrm{E}\left(\sum_i X_i^2 + \sum_{i\neq j} X_i X_j\right) = \sum_i \mathrm{E}\left(X_i^2\right) + \sum_{i\neq j} \mathrm{E}\left(X_i X_j\right).$$

Since X_i is a 0/1 r.v., $\mathrm{E}\left(X_i^2\right) = \mathrm{E}(X_i)$, so the first sum is just $\sum_i \mathrm{E}(X_i) = \mu = 1$.

What about $\mathrm{E}(X_i X_j)$? Since X_i, X_j are both 0/1 r.v.'s, $\mathrm{E}(X_i X_j) = \Pr[X_i = 1 \wedge X_j = 1] = \frac{1}{n(n-1)}$. (Why?) So the second sum above is $\sum_{i\neq j} \mathrm{E}(X_i X_j) = n(n-1)\cdot\frac{1}{n(n-1)} = 1$.

Putting these together, we get $\mathrm{E}\left(X^2\right) = 1+1 = 2$, and hence $\mathrm{Var}(X) = 2-1 = 1$.

So, for this r.v. X, we have $\mu = 1$ and $\sigma^2 = 1$.

Ex: What does this mean? What extra constraints does this put on the distribution of X? $\square$

Ex: Recall Example 1 from Note 1: let the r.v. X be the number of empty bins when m balls are tossed at random into n bins. We have seen that $\mathrm{E}(X) = n\left(1-\frac{1}{n}\right)^m$. What is $\mathrm{Var}(X)$? $\square$

Conditional Probability

The homeworks of n students are randomly shuffled and returned. Pr[I get my own hw] $= \frac{1}{n}$.
Does this probability change if you tell me that you got your own homework?
Yes: looking at the reduced sample space of outcomes where you get your own hw, the prob is

$$\frac{\#\text{ perms in which we both get our own hw}}{\#\text{ perms in which you get your own hw}} = \frac{(n-2)!}{(n-1)!} = \frac{1}{n-1}.$$

Working instead in the *original* sample space, we arrive at the notion of *conditional probability*:

Definition: For events E, F on the same sample space, the conditional probability of E given F is defined as

$$\Pr[E|F] = \tfrac{\Pr[E \wedge F]}{\Pr[F]}. \qquad (*)$$

So $\Pr[\text{I get my own hw} \,|\, \text{You get your own hw}] = \frac{1/n(n-1)}{1/n} = \frac{1}{n-1}$.

Some useful equalities

1. $\Pr[E|F]\Pr[F] = \Pr[F|E]\Pr[E] = \Pr[E \wedge F]$.
2. $\Pr[E_1 \wedge E_2 \wedge \ldots \wedge E_k] = \Pr[E_1] \times \Pr[E_2|E_1] \times \Pr[E_3|E_1 \wedge E_2] \times \cdots \times \Pr[E_k| \wedge_{i<k} E_i]$.
3. $\Pr[E] = \Pr[E|F]\Pr[F] + \Pr[E|\overline{F}]\Pr[\overline{F}]$, where $\overline{F}$ is the complement of F.

Ex: Verify the above formulae using the definition $(*)$. $\square$

(Alternative) Definition: Events E, F are independent if $\Pr[E|F] = \Pr[E]$.

Ex: Convince yourself that this is equivalent to the definition of Note 1, page 4. $\square$

Ex: In the example above, show that the events "I get my own homework" and "You get your own homework" are *not* independent. Why does this not contradict our view of a random permutation as a sequence of *independent* trials? $\square$

Ex: Consider the balls-and-bins experiment with $n = m = 4$ (i.e., 4 labeled balls are thrown into 4 labeled bins). Define the events E_i = first i balls land in different bins. Show that $\Pr[E_4] = \frac{3}{32}$, $\Pr[E_4|E_2] = \frac{1}{8}$, $\Pr[E_4|E_3] = \frac{1}{4}$. What is $\Pr[E_3|E_4]$? $\square$