Warm-up

Problem 1. Consider a deck of $4n$ cards, with n ♠, n ♡, n ♢, and n ♣. After it is shuffled uniformly at random, what is the expected number of consecutive pairs of the same suit?

Solution 1. This is $n-1$: Let $X_i \in \{♠, ♡, ♢, ♣\}$ denote the suit of the i-th card in the permuted deck. We want to compute the expectation of

$$X = \sum_{i=1}^{4n-1} \mathbb{1}_{\{X_i = X_{i+1}\}}$$

For every $1 \leq i \leq 4n-1$,

$$\Pr[\, X_i = X_{i+1} \,] = \frac{n-1}{4n-1}$$

since once X_i has a given suit, then there remain $n-1$ cards of that particular suit, out of $4n-1$ cards left in total. The result then follows from linearity of expectation. As a sanity check: for $n = 13$ (standard deck of 52 cards), we get $n-1 = 12$, retrieving the result mentioned in class.

Problem 2. A computer randomly generates a 2024-bit long binary string. What is the expected number of consecutive runs of 3 ones? (For instance, the 4-bit binary string 1111 has 2 such consecutive runs, while 0111 only has 1.)

Solution 2. This is again by linearity of expectation, looking at the $2024 - 2 = 2022$ indicator random variables $Y_1, \dots, Y_{2022}$ defined by

$$Y_i = \mathbb{1}_{\{X_i = X_{i+1} = X_{i+2} = 1\}}$$

where $X = (X_1, \dots, X_{2024}) \in \{0,1\}^{2024}$ is the binary string. Now,

$$\mathbb{E}[Y_i] = \Pr[\, X_i = X_{i+1} = X_{i+2} = 1 \,] = \Pr[\, X_i = 1 \,] \cdot \Pr[\, X_{i+1} = 1 \,] \cdot \Pr[\, X_{i+2} = 1 \,] = \frac{1}{2^3}$$

(the second inequality as the bits are independent) and so the answer is $\frac{2024-2}{2^3} = \frac{2022}{8} = 252.75$.

Problem 3. An integer $1 \leq i \leq n$ is called a *fixed point* of a given permutation $\pi\colon \{1, 2, \dots, n\} \to \{1, 2, \dots, n\}$ if $\pi(i) = i$. Show that the expected number of fixed points of a uniformly randomly chosen permutation π is 1. What is the variance?

Solution 3. Linearity of expectation. Specifically, let $X_i \in \{0,1\}$ be the indicator random variable of whether $\pi(i) = i$ (that is, it is 1 if, and only if, i is a fixed point of the random permutation π). Of course, the X_i's are not independent (but we don't care). Since a given element i (only looking at this element) has value

$\pi(i)$ equal to any fixed element of $\{1, 2, \ldots, n\}$ with the same probability, we get $\Pr[\, X_i = 1\,] = 1/n$. The expected number $\mathbb{E}[X]$ of fixed points is then

$$\mathbb{E}[X] = \mathbb{E}\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} \mathbb{E}[X_i] = \sum_{i=1}^{n} \Pr[\, X_i = 1\,] = \sum_{i=1}^{n} \frac{1}{n} = 1$$

The answer is 1 for the variance as well, apply linearity of expectation after expanding the square of the sum and dividing it into $\sum_i$ and $\sum_{i \neq j}$. In more detail, since $\operatorname{Var} X = \mathbb{E}\left[X^2\right] - \mathbb{E}[X]^2$ and we just got that $\mathbb{E}[X]^2 = 1^2 = 1$, it suffices to show that $\mathbb{E}\left[X^2\right] = 2$. We have:

$$\begin{aligned}
\mathbb{E}\left[X^2\right] &= \mathbb{E}\left[\left(\sum_{i=1}^{n} X_i\right)^2\right] = \mathbb{E}\left[\sum_{i=1}^{n}\sum_{j=1}^{n} X_i X_j\right] \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} \mathbb{E}\left[X_i X_j\right] \\
&= \sum_{i=1}^{n} \mathbb{E}\left[X_i^2\right] + \sum_{1 \leq i \neq j \leq n}^{n} \mathbb{E}\left[X_i X_j\right] \\
&= \sum_{i=1}^{n} \mathbb{E}[X_i] + \sum_{1 \leq i \neq j \leq n} \Pr[\, \pi(i) = i \text{ and } \pi(j) = j\,] && (\text{since } X_i^2 = X_i) \\
&= 1 + \sum_{1 \leq i \neq j \leq n} \frac{1}{n} \cdot \frac{1}{n-1} && (\text{see below}) \\
&= 1 + 1 = 2 && (\text{The second sum has } n(n-1) \text{ terms})
\end{aligned}$$

concluding the proof. Now, why do we have $\Pr[\, \pi(i) = i \text{ and } \pi(j) = j\,] = \frac{1}{n} \cdot \frac{1}{n-1}$? We pick uniformly at random two distinct values in $\{1, 2, \ldots, n\}$ for $(\pi(i), \pi(j))$: there are $n(n-1)$ possibilities; out of these, only one is good.

Problem 4. (1) Give a random variable X over $[0, \infty)$ such that $\mathbb{E}[X] = \infty$. (2) Give a random variable X over $\mathbb{N}$ such that $\mathbb{E}[X] = \infty$.

Solution 4. (1) Random variable with probability density function $f(x) = \frac{2}{\pi}\frac{1}{x^2+1}$.
(2) Random variable with probability mass function $p(n) = \frac{6}{\pi^2}\frac{1}{(n+1)^2}$.

Problem 5. Prove the fact from the lecture: if X has a finite variance, then $\operatorname{Var} X = \mathbb{E}\left[X^2\right] - \mathbb{E}[X]^2$.

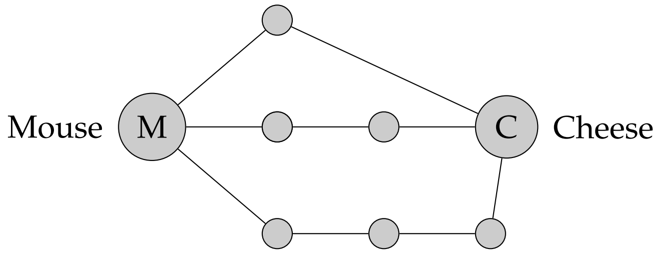
Solution 5. Expand inside the expectation, hope for the best.

Problem solving

Problem 6. Prove the fact from the lecture: if X takes values in $\mathbb{N} = \{0, 1, 2, \dots,\}$ and $\mathbb{E}[X]$ is finite, then $\mathbb{E}[X] = \sum_{n=1}^{\infty} \Pr[\, X \geq n \,]$.

Solution 6. Swap sums, hope for the best.

Problem 7. Consider the following map: each edge represents a path (of length one) between two different locations. To reach the cheese, the mouse needs to take a path connecting locations M and C.



Unfortunately, cats have heard of this plan, and will try to intercept the mouse. These cats are not the brightest, thankfully, and behave randomly: namely, each edge will be occupied by a cat, independently of all other edges, with some fixed probability $p \in (0, 1)$. The mouse cannot go on any edge that has a cat, of course. (Once the cats have randomly decided their position at the beginning, they stay there once and for all, effectively "killing" that edge as far as the mouse is concerned.)

a) Give the probability that the mouse still has a path leading to the cheese.
b) Give the probability that the mouse still has a path of length at most 3 leading to the cheese.
c) Give the expected numbers of cats on the map.

Solution 7.

a) The idea here is to simplify the problem by asking questions of the form "is there *no* cat" because while there are a large number of ways to have "*a* cat," there is only one way to have none of something.

$$
\begin{aligned}
\Pr[\text{There is a path}] &= 1 - \Pr[\text{There is no path}] \\
&= 1 - \Pr[\text{Top path has a cat}] \cdot \Pr[\text{Middle path has a cat}] \\
&\qquad \cdot \Pr[\text{Bottom path has a cat}] \\
&= 1 - (1 - \Pr[\text{Top path has no cats}]) \cdot (1 - \Pr[\text{Middle path has no cats}]) \\
&\qquad \cdot (1 - \Pr[\text{Bottom path has no cats}]) \\
&= 1 - (1 - (1-p)^2)(1 - (1-p)^3)(1 - (1-p)^4)
\end{aligned}
$$

Small "sanity check": for $p = 0$, we get a probability $1 - 0 \cdot 0 \cdot 0 = 1$, which makes sense (there is no cat anywhere). For $p = 1$, we get $1 - 1 \cdot 1 \cdot 1 = 0$, which also makes sense (there are cats everywhere).

b) We just remove the bottom path from our working above, giving

$$1 - (1 - (1-p)^2)(1 - (1-p)^3).$$

c) If we define an indicator variable X_e such that $X_e = 1$ if edge e (in the set of edges E) has a cat, and count by taking $X = \sum_{e \in E} X_e$, we get

$$\mathbb{E}[X] = \mathbb{E}\left[\sum_{e \in E} X_e\right] = \sum_{e \in E} \mathbb{E}[X_e] = 9p$$

Problem 8. Let A be an array of n distinct numbers. We say that an index $1 \leq i \leq n$ is "prefix-maximum" if $A[i]$ is the biggest number so far, that is, if $A[j] < A[i]$ for all $j < i$. Let $\mathrm{pf}(A)$ denote the number of prefix-maximum indices of A.

a) What is $\mathrm{pf}(A)$ if A is sorted (increasing)?

b) Suppose that we permute the elements of A uniformly at random to get an array B. Show that

$$\mathbb{E}[\mathrm{pf}(B)] = H_n = O(\log n)\,,$$

where $H_n = 1 + 1/2 + 1/3 + \cdots + 1/n$ is the n-th Harmonic number.

Solution 8.

a) Every element is bigger than the ones before it (because the array is sorted) and we say that the first element is a prefix maximum (as it is bigger than nothing) so $\mathrm{pf}(A) = n$.

b) Define an indicator variable X_i which is 1 if B_i is a prefix maximum. Then we have $\mathrm{pf}(B) = \sum_{i=1}^{n} X_i$,

$$\mathbb{E}[\mathrm{pf}(B)] = \sum_{i=1}^{n} \mathbb{E}[X_i]$$

Where $\mathbb{E}[X_1]$ is 1 (it is always the largest so far), $\mathbb{E}[X_2]$ is the probability that given two distinct elements, we randomly choose the largest, which is 1/2, and so on. This gives the series

$$\mathbb{E}[\mathrm{pf}(B)] = \sum_{i=1}^{n} \frac{1}{i} = O(\log n).$$

Advanced

Problem 9. Given two values $x, y \in \{0,1\}$, their XOR $x \oplus y$ is equal to their sum modulo 2, or equivalently, is 1 if $x + y$ is odd, and 0 otherwise. This generalises to n bits as follows: for $x_1, \dots, x_n \in \{0,1\}$,

$$x_1 \oplus x_2 \oplus \cdots \oplus x_n = \begin{cases} 0 \text{ if } \sum_{i=1}^n x_i \text{ is even} \\ 1 \text{ if } \sum_{i=1}^n x_i \text{ is odd} \end{cases}$$

Suppose that $X_1, \dots, X_n, \dots$ are independent Bernoulli random variables with parameter $p \in [0,1]$, and, for any $n \geq 1$, let $Y_n = X_1 \oplus X_2 \oplus \cdots \oplus X_n$. This is itself a Bernoulli random variable: let's call its parameter p_n.

a) Compute the first few values of p_n when $p = 1/2$, $p = 0$, and $p = 1$. Establish the expression of p_n (as a function of n) for these particular cases. Interpret the result.

b) In general, as a function of p, what is p_0? p_1? p_2?

c) Give a recurrence relation for p_n.

d) Solve the recurrence to obtain the expression for p_n. Show that it always converge to $1/2$. How fast?

Solution 9.

a) For $p = 1/2$, we get $p_0 = 0$, and $p_n = 1/2$ for every $n \geq 1$. (XOR-ing independent fair random bits still gives a fair random bit). For $p = 0$, then $Y_n = 0$ for all n, and so $p_n = 0$ for $n \geq 0$. For $p = 1$, then $Y_n = 0$ (with probability 1) for n even and 1 for n odd, and so $p_n = 1 - (-1)^n$ for $n \geq 0$.

b) $p_0 = 0$ (Y_0 is the sum of…nothing, so is always equal to 0); $p_1 = p$ by definition, while $p_2 = 2p(1-p)$ (the probability that $X_1 \neq X_2$: exactly one of the two is must be equal to 1, the other 0).

c) We have $Y_{n+1} = Y_n \oplus X_{n+1}$. So for Y_{n+1} to be equal to 1, we need either (1) $Y_n = 0$ and $X_{n+1} = 1$ or (2) $Y_n = 1$ and $X_{n+1} = 0$. These are disjoint events, so the probability $p_{n+1} = \Pr[Y_{n+1} = 1]$ is the sum of the probabilities of these two events; by recurrence, we have that the first has probability $(1 - p_n) \cdot p$, and the second has probability $p_n \cdot (1 - p)$.
We then get the recurrence

$$p_{n+1} = (1-p)p_n + p(1-p_n)$$

or, massaging the right-hand-side,

$$p_{n+1} = (1-2p)p_n + p$$

d) Now, solving this recurrence will give the solution:

$$p_n = \frac{1}{2}(1 - (1-2p)^n), \qquad n \geq 0$$

Before proving it, note that this converges exponentially fast to $1/2$ for $p \in (0,1)$, is stationary at 0 for $p = 0$, and does not converge for $p = 1$. This is consistent with questions a) and b) (important to check!)

How do we get there?

- **Nice way:** $1/2$ seems special (as it is a fixed point of the recurrence relation, as shown in a)), so let's "center" on $1/2$: that is, let $q_n = p_n - 1/2$, and rewrite:

  $$q_{n+1} + 1/2 = (1-2p)(q_n + 1/2) + p$$

  which gives, expanding and simplifying:

  $$q_{n+1} = (1-2p)q_n$$

  from which we can easily get $q_{n+1} = (1-2p)^{n+1}q_0$. But that means

  $$p_{n+1} = (1-2p)^{n+1}q_0 + \frac{1}{2}$$

  and since $q_0 = p_0 - 1/2 = -1/2$, we get

  $$p_{n+1} = \frac{1}{2}\left(1 - (1-2p)^{n+1}\right)$$

  as claimed.

- **Painful, general way:** first rewrite it as a linear system:

  $$\begin{pmatrix} p_{n+1} \\ 1 \end{pmatrix} = \begin{pmatrix} 1-2p & p \\ 0 & 1 \end{pmatrix} \begin{pmatrix} p_n \\ 1 \end{pmatrix}$$

  so that

  $$\begin{pmatrix} p_{n+1} \\ 1 \end{pmatrix} = \begin{pmatrix} 1-2p & p \\ 0 & 1 \end{pmatrix}^n \begin{pmatrix} p \\ 1 \end{pmatrix}$$

  We can then try to diagonalise the matrix to compute its n-power more easily: if

  $$\begin{pmatrix} 1-2p & p \\ 0 & 1 \end{pmatrix} = P^{-1}\Delta P$$

  with Δ diagonal, then

  $$\begin{pmatrix} 1-2p & p \\ 0 & 1 \end{pmatrix}^n = P^{-1}\Delta^n P$$

  and Δ^n is easy to compute (as Δ is diagonal).

Diagonalising gives

$$\Delta = \begin{pmatrix} 1 & 0 \\ 0 & 1-2p \end{pmatrix}, \qquad P = \frac{1}{2}\begin{pmatrix} 0 & 2 \\ 2 & -1 \end{pmatrix}, \qquad P^{-1} = \frac{1}{2}\begin{pmatrix} 1 & 2 \\ 2 & 0 \end{pmatrix}$$

and so

$$\begin{aligned}
\begin{pmatrix} p_{n+1} \\ 1 \end{pmatrix} &= \frac{1}{2}\begin{pmatrix} 1 & 2 \\ 2 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & (1-2p)^n \end{pmatrix} \cdot \frac{1}{2}\begin{pmatrix} 0 & 2 \\ 2 & -1 \end{pmatrix}\begin{pmatrix} p \\ 1 \end{pmatrix} \\
&= \frac{1}{4}\begin{pmatrix} 1 & 2 \\ 2 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & (1-2p)^n \end{pmatrix}\begin{pmatrix} 2 \\ 2p-1 \end{pmatrix} \\
&= \frac{1}{4}\begin{pmatrix} 1 & 2 \\ 2 & 0 \end{pmatrix}\begin{pmatrix} 2 \\ -(1-2p)^{n+1} \end{pmatrix} \\
&= \begin{pmatrix} \frac{1}{2} - \frac{1}{2}(1-2p)^{n+1} \\ 1 \end{pmatrix}
\end{aligned}$$

and we again get $p_{n+1} = \frac{1}{2}\left(1 - (1-2p)^{n+1}\right)$, for all $n \geq 0$.