1 Events and Probabilities

1. Compare the following events:

    - We flip a fair coin 26 times, and consider the event that we see 13 heads and 13 tails.
    - We select the top 26 cards from a well-shuffled pack, and consider the event that we end up selecting 13 red and 13 black cards.

    Without explicit calculation, which event do you expect is more likely?

2. You are hoping to make a uniform random choice from $\{1, 2, \ldots, k\}$. Regrettably, all you have is an m-sided dice, for $m \neq k$. How would you proceed (a) if $m > k$? (b) if $m < k$?

3. Original problem: A fictional college committee contains a proportion p of dons, who never change their minds about anything, and a proportion $1 - p$ of student reps who change their minds completely at random (with probability r) between successive votes on the same issue.

    Now, imagine that p is not known, but $r = 1/2$. A *Varsity* investigative article claims that $p = 0.01$. You are unconvinced. You pick a member uniformly at random and observe their voting. They vote the same for 20 consecutive votes. Discuss whether you might obtain useful further information to assess the claim by watching them for the next 20 votes.

    What about if *The Cambridge Student* had responded, claiming that $p = 0$?

4. A *lattice path* $\mathcal{S} = (S_0, S_1, \ldots, S_n)$ satisfies $S_0 = 0$ and $S_k - S_{k-1} = \pm 1$ for each $k = 1, \ldots, n$. We will study how many lattice paths $\mathcal{S}$ satisfy $S_n = 0$ and $S_1, \ldots, S_{n-1} \geq 0$. There are many approaches to this enumeration problem.

    (a) What can you say when n is odd?

    (b) When $n = 2m$ is even, here is one approach. Instead, study lattice paths of length $2m+1$ satisfying $S_{2m+1} = -1$, and $S_0, S_1, \ldots, S_{2m} \geq 0$, so that $2m + 1$ is the '*hitting time*' of -1. Show that the number of such lattice paths is $\frac{1}{2m+1}\binom{2m+1}{m}$.

    (c) Explain the connection to the original problem, and express the answer to the original problem in the most natural form.

    (d*) Using notation from later in the course, suppose now that the lattice path is generated randomly by assuming that

    $$\mathbb{P}(S_k - S_{k-1} = +1) = \frac{1}{2}, \quad \mathbb{P}(S_k - S_{k-1} = -1) = \frac{1}{2},$$

    for every $k \geq 1$ independently. This is the simplest example of a *random walk*. Part (b) makes a connection between

    $$\mathbb{P}(S_{2m+1} = -1) \quad \text{and} \quad \mathbb{P}(S_{2m+1} = -1,\, S_1, \ldots, S_{2m} \geq 0).$$

    How far can this connection be generalised?

    (e*) If you are re-reading this problem set later in the course, you might like to revisit this problem using generating functions.

Dominic Yeo djy21
dominic.yeo@kcl.ac.uk

2 Random variables, expectations, notable distributions

1. Let X be a geometric random variable on $\{0,1,2,\dots\}$ with parameter $p \in (0,1)$, and Y a Poisson random variable with parameter $\lambda > 0$. Compare the distributions of i) $X-n$, conditional on $X \geq n$; and ii) $Y-n$, conditional on $Y \geq n$, when n is large.

2. Can you construct non-negative integer valued random variables X and $(X_n)_{n\geq 1}$, all taking non-negative integer values, such that $\mathbb{E}[X], \mathbb{E}[X_n]$ are all finite, and such that

$$\mathbb{P}(X_n = k) \to \mathbb{P}(X = k) \text{ as } n \to \infty,$$

holds for every $k \in \{0,1,\dots\}$, but for which $\mathbb{E}[X_n] \not\to \mathbb{E}[X]$?

3. Let σ_n be a uniformly chosen permutation from Σ_n, and consider the cycle decomposition of σ_n. Let $\alpha_{n,k}$ be the number of cycles in σ_n of length k. Find $\mathbb{E}[\alpha_{n,k}]$. Let α_n be the total number of cycles in σ_n. Show that $\mathbb{E}[\alpha_n] \to \infty$ as $n \to \infty$.

   Now, let ℓ_n be the length of the cycle of σ_n which includes the element 1. Find the distribution of ℓ_n, and also $\mathbb{E}[\ell_n]$.

   Note that $\mathbb{E}[\ell_n]\mathbb{E}[\alpha_n] \gg n$. Explain why this is not a contradiction.

4. In a community of N people, birthdays are independent, and uniformly chosen from the 365 days of the non-leap year. How would you try to find an expression for the probability that at least k people share a birthday? (Ie, at least one day d such that at least k people were born on day d.)

   Would this analysis be easier if instead you assumed the number of people in the community was random, with Poisson(N) distribution?

   What if the birthdays were independent but not distributed uniformly through the year?

3 Random walks, branching processes

1. Let q_λ be the extinction probability of a branching process with Poisson(λ) offspring distribution. Consider $\zeta_\lambda = 1 - q_\lambda$ the corresponding *survival probability*. Show that ζ_λ satisfies $\zeta_\lambda = 1 - e^{-\lambda\zeta_\lambda}$. Describe the behaviour of $\zeta_{1+\epsilon}$ as $\epsilon \downarrow 0$, in the form $\zeta_{1+\epsilon} \sim C\epsilon^\alpha$, for constants C, α to be determined.

   [*Hint: start by showing $\zeta_\lambda \to 0$ as $\lambda \downarrow 1$, then expand.*]

2. Let G be a *graph*, consisting of a finite collection of vertices $V(G)$, some pairs of which are connected by an edge. We declare the *neighbours* of a vertex v to be those vertices w such that v, w are directly connected by an edge. We assume the graph is *connected*, so that there is a path of edges joining any pair of vertices. Let A be a non-empty subset of the vertices, and $B = V(G) \setminus A$. For each vertex $v \in A$, we declare a value $a_V \in \mathbb{R}$.

   Use a probabilistic argument to show that there exists a function $f : V(G) \to \mathbb{R}$ such that $f(v) = a_v$ for all $v \in A$, and for all $v \in B$, $f(v)$ is the average of the values taken by f on the neighbours of v.

   Can you justify that f is unique?

   [*Note: such a function f is called the (discrete) harmonic extension of (a_v).*]

3. Let $\mathcal{T}$ be a branching process tree with offspring distribution X satisfying $\mu = \mathbb{E}[X] \leq 1$. Order the individuals in a *breadth-first* manner, so the root is x_1, and the the children of the root are $x_2, \ldots, x_{1+Z_1}$, and the individuals in the $(n+1)$th generation are

   $$x_{Z_0+Z_1+\ldots+Z_n+1},\ x_{Z_0+Z_1+\ldots+Z_n+2} \cdots\ , x_{Z_0+Z_1+\ldots+Z_n+Z_{n+1}}.$$

   Let $c(x_i)$ be the number of children of individual x_i.

   (a) Consider the random process given by $S_0 = 0$ and $S_m = c(x_1) + \ldots + c(x_m) - m$. Explain briefly why $(S_0, S_1, \ldots)$ is a random walk whose increments have distribution $X - 1$.

   (b) Show that $|\mathcal{T}|$, the total number of individuals in the population, has the same distribution as $\tau := \inf\{m \geq 0 : S_m = -1\}$.

   (c) Prove that $\mathbb{P}(|\mathcal{T}| = m) = \frac{1}{m}\mathbb{P}(S_m = -1)$.

   (d) Suppose that X takes the values 2 and 0 each with probability 1/2. Explain why $\mathbb{E}[|\mathcal{T}|] = \infty$, and find constants C, α such that $\mathbb{P}(|\mathcal{T}| = m) \sim Cm^{-\alpha}$.

4. Let $\mathcal{T}$ be a branching process tree with supercritical offspring distribution X satisfying $\mu = \mathbb{E}[X] > 1$. Denote by Ψ the *extinction* event $\{|\mathcal{T}| < \infty\}$, and assume $q = \mathbb{P}(\Psi) \in (0, 1)$.

   (a) Explain briefly why the conditional branching process tree $(\mathcal{T} \mid \Psi)$ is itself a branching process, and describe its offspring distribution $\hat{X}$.

   (b) Show that if $X \sim \text{Po}(\mu)$, then $\hat{X} \sim \text{Po}(\nu)$, where $\nu \neq \mu$ and satisfies $\mu e^{-\mu} = \nu e^{-\nu}$. Show that ν is monotone as a function of μ.

   (c) Now return to $\mathcal{T}$, with $X \sim \text{Po}(\mu)$. Colour blue all individuals with infinitely many descendents, and colour red all others, so that $\Psi^c = \{\text{root is blue}\}$. State the distribution of the number of blue children of the root. Now characterise the distribution of the number of blue children of the root, conditional on Ψ^c, and check that the conditional mass function sums to 1.

Give as complete as description as you can manage for the structure of the blue and red individuals in $\mathcal{T}$, conditional on Ψ^c.

4 Continuous random variables, and limits of random variables

1. Let us define discrete random variables X and Y by a joint probability mass function as shown:

Y \ X	-1	0	1
-1	a	b	c
0	d	e	f
1	g	h	i

Write down the marginal mass functions $p_X(\cdot)$ and $p_Y(\cdot)$ of X and Y.

A lecturer is trying to choose $\{a, b, \ldots, i\}$ to illustrate the general principle that $\mathrm{Cov}(X, Y) = 0$ does *not* imply X, Y independent. Describe geometrically the sets of $\{a, b, \ldots, i\}$s which (i) induce X, Y independent, and (ii) induce $\mathrm{Cov}(X, Y) = 0$, and conclude that if the lecturer chooses *uniformly at random from the set in (ii)* (*), this will illustrate their point with probability one.

[You may reflect on how to make (*) formal, but this is not part of the question.]

2. (a) The Polya urn model for contagion is as follows. We start with an urn which contains one white marble and one black marble. At each second we choose a marble at random from the urn and replace it together with one more marble of the same colour. Calculate the probability that when n marble are in the urn, i of them are white, and conclude a limit result for convergence in distribution of the proportion of white marbles in the urn.

   (b) Let $U_1, U_2, \ldots$ be IID Unif$[0, 1]$ random variables in probability space $(\Omega, \mathcal{F}, \mathbb{P})$. For each n, we define $\sigma^{(n)} = (\sigma_1^{(n)}, \ldots, \sigma_n^{(n)})$ a random permutation of $\{1, 2, \ldots, n\}$ to be the ordering of $(U_1, \ldots, U_n)$. That is, $\sigma^{(n)}$ is the unique permutation for which

$$U_{\sigma_1^{(n)}} < U_{\sigma_2^{(n)}} < \ldots < U_{\sigma_n^{(n)}}.$$

   [Note, in this construction, we are ignoring the possibility that any two of the U_ns are equal. Fortunately, this event has probability zero.]

   i) Can you state a simpler way to construct $\sigma^{(n+1)}$ from $\sigma^{(n)}$ directly, without using the auxiliary $U_1, U_2, \ldots$ random variables?

   ii) Prove that $(\frac{1}{n}\sigma_k^{(n)})_{n \geq k}$ converges almost surely for each $k \geq 1$.

   (c) Can you use relate part (b) to part (a) to prove almost sure convergence for Polya's urn?

3. We extend a problem from Sheet 2. Let σ_n be a uniformly chosen permutation from Σ_n, and α_n the number of cycles in σ_n. Denote by $\mathcal{C} = (C_1, C_2, \ldots, C_{\alpha_n})$ the lengths of these cycles, in decreasing order. Note that conditional on $\mathcal{C}$, the elements $\{1, 2, \ldots, n\}$ are assigned to the cycles uniformly.

   Let ℓ_n be the length of the cycle containing the element 1. Let β_n be the length of a cycle chosen uniformly from the α_n possible cycles. Find expressions for $\mathbb{E}[\ell_n \,|\, \mathcal{C}]$ and $\mathbb{E}[\beta_n \,|\, \mathcal{C}]$ in terms of $\mathcal{C}$, and prove that $\mathbb{E}[\ell_n] \geq \mathbb{E}[\beta_n]$. Finally, prove that $\mathbb{E}[\beta_n]\mathbb{E}[\alpha_n] \geq n$.

Dominic Yeo djy21
dominic.yeo@kcl.ac.uk

4. Daniel is playing *Snakes and Ladders*, which we model by adapting a random walk.

   (a) To avoid tears, initially there are no snakes. Let $X_1, X_2, \ldots$ be IID uniform choices from $\{1, 2, \ldots, 6\}$, denoting dice rolls, with $S_n = X_1 + \ldots + X_n$ denoting position after n moves.

   Daniel plays until time $T_N := \inf\{n \geq 0 : S_n \geq N\}$, where N is large for the sake of his parents' productivity. Derive a CLT for T_N, ie a limit in distribution for $\frac{T_N - a_N}{b_N}$, for some a_N, b_N.

   (b) Let $\alpha_N = \lfloor N/3 \rfloor$ and $\beta_N = \lfloor 2N/3 \rfloor$. We introduce a snake from $\beta_N \mapsto \alpha_N$, so that now

$$S_{n+1} := \begin{cases} S_n + X_n & \text{if } S_n + X_n \neq \beta_N \\ \alpha_N & \text{if } S_n + X_n = \beta_N. \end{cases}$$

   When N is large, state the limiting probability that the second case (ie the snake move) occurs at least once. [You do not need to prove the validity of this limit.]

   For large N, describe approximately the distribution of $T_N := \inf\{n \geq 0 : S_n \geq N\}$, and explain why there is no choice of a_N, b_N such that $\frac{T_N - a_N}{b_N}$ has a continuous limit in distribution.