Chapter 11 Probabilistic Method

Definition A probability space is a triple $(\Omega, \Sigma, \mathrm{P})$, where Ω is a set, $\Sigma \subseteq 2^\Omega$ is a σ-algebra on Ω (a collection of subsets containing Ω and closed on complements, countable unions and countable intersections), and P is a countably additive measure on Σ with $\mathrm{P}[\Omega] = 1$. The elements of Σ are called events and the elements of Ω are called elementary events. For an event A, $\mathrm{P}[A]$ is called the probability of A.

We will consider Ω finite and $\Sigma = 2^\Omega$ in our examples later.

1: Give an example of $(\Omega, \Sigma, \mathrm{P})$.

2: Why is P on Σ and not on Ω?

3: Show that for any collection of events $A_1, \ldots, A_n$,

$$\mathrm{P}\left[\bigcup_{i=1}^{n} A_i\right] \leq \sum_{i=1}^{n} \mathrm{P}(A_i).$$

Events A, B are independent if $\mathrm{P}[A \cap B] = \mathrm{P}[A]\mathrm{P}[B]$. More generally, events $A_1, \ldots, A_n$ are independent if for any subset of indices $I \subseteq [n]$

$$\mathrm{P}\left[\bigcap_{i \in I} A_i\right] = \prod_{i \in I} \mathrm{P}[A_i].$$

4: Find three events A_1, A_2 and A_3 that are pairwise independent but not mutually independent.
(You need to say what is $(\Omega, \Sigma, \mathrm{P})$ as well.)
Hint: $\Omega = \{a, b, c, d\}$ *and* $\mathrm{P}[x] = \frac{1}{4}$ *for each* $x \in \Omega$ *could work.*

For events A and B with $\mathrm{P}[B] > 0$, we define the conditional probability of A, given that B occurs, as

$$\mathrm{P}[A|B] = \frac{\mathrm{P}[A \cap B]}{\mathrm{P}(B)}.$$

5: Simplify the formula for independent events A and B.

A real random variable on a probability space $(\Omega, \Sigma, \mathrm{P})$ is a function $X : \Omega \to \mathbb{R}$ that is P-measurable. (That is, for any $a \in \mathbb{B}$, $\{\omega \in \Omega : X(\omega) \leq a\} \in \Sigma$.)

We use Ω discrete, so no trouble with measurable in our case.

Expectation for finite Ω can be expressed as $E[X] = \sum_{\omega \in \Omega} \mathrm{P}[\omega] X(\omega)$.

Real random variables X, Y are independent if for every two measurable sets $A, B \subseteq \mathbb{R}$,

$$\mathrm{P}[X \in A \text{ and } Y \in B] = \mathrm{P}[X \in A] \cdot \mathrm{P}[Y \in B].$$

For verification, it is enough to check

$$\mathrm{P}[X \leq a \text{ and } Y \leq b] = \mathrm{P}[X \leq a] \cdot \mathrm{P}[Y \leq b].$$

6: What is $\mathrm{P}[X \in A]$?

7: Show the following for a finite probability space. If X and Y are independent random variables, then $E[XY] = E[X] \cdot E[Y]$.

2-coloring hypergraphs - Construct something random.

A k-uniform hypergraph (V,E) has V as a set of vertices and edges $E \subseteq \binom{V}{k}$. That is, edges are k-subsets.

A hypergraph is c-colorable if its vertices can be colored with c colors so that no edge is monochromatic i.e., at least two different colors appear in every edge.

Let $m(k)$ denote the smallest number of edges in a k-uniform hypergraph that is not 2-colorable.

8: What is $m(2)$?

9: Use probabilistic method to show that for any $k \geq 2$,

$$m(k) \geq 2^{k-1}.$$

Hint: Union bound.

$E[X] = \sum_{x \in \Omega} x \cdot p(x)$

Linearity of Expectation

Linearity of Expectation Let $X_1, \cdots, X_n$ be random variables, $X = c_1X_1 + \cdots, +c_nX_n$, then

$$\mathbb{E}[X] = c_1\mathbb{E}[X_1] + \cdots + c_n\mathbb{E}[X_n].$$

Definition For an event A, the indicator random variable I_A has value 1 if event A occurs and has value 0 otherwise. $E[I_A] = Pr(A)$

10: Calculate the expected number of fixed points of random permuation σ on $\{1,\ldots,n\}$, i.e., the number of i such that $\sigma(i) = i$.

Let X_i be the event that σ fixes i.

We want to know $\mathbb{E}[X]$ where

$X = X_1 + X_2 + \cdots + X_n$

$\mathbb{E}(X] = \underline{\mathbb{E}[X_1]} + \cdots + \mathbb{E}[X_n] = \underbrace{\frac{1}{n} + \cdots + \frac{1}{n}}_{n} = 1$

$\exists \left\lceil \frac{n!}{e} \right\rceil$ derangements

11: Show that there is a tournament on n vertices that has at least $\frac{n!}{2^{n-1}}$ Hamiltonian paths.
Remark: Alon(1990) proved that the maximum number of Hamiltonian paths is at most $cn^{3/2}\frac{n!}{2^{n-1}}$.

A tournament on n vertices is an orientation of K_n.

A Hamiltonian path is a path that sees all the vertices.
Select a random tournament. Consider a permutation σ. and the corresponding path $\sigma(1), \sigma(2), \dots, \sigma(n)$
The probability this is Hamiltonian is $(1/2)^{n-1}$. Let $X = \sum_\sigma X_\sigma$
$\mathbb{E}[X] = \sum_\sigma \mathbb{E}[X_\sigma] = \frac{n!}{2^{n-1}}$. $\exists$ an instance w/ $\geq \frac{n!}{2^{n-1}}$ H.P.s

12: Show that any graph G with e edges contains a bipartite subgraph with at least $e/2$ edges.
Hint: randomly partition vertices into two parts.

Construct a partition $V = A \sqcup B$, where $\Pr(v \in A) = \frac{1}{2}$ $\forall v \in V$.
Let X_{uv} be the event uv is a crossing edge
$\mathbb{E}[X_{uv}] = \frac{1}{2}$. Let $X = \sum X_{uv}$
$\mathbb{E}[X] = \sum_{uv \in E} \mathbb{E}[X_{uv}] = \frac{1}{2}|E(G)|$. So $\exists$ a partition w/ $\geq \frac{e}{2}$ crossing edges.

The above result can be improved:

13: Show that if G has $2n$ vertices and e edges, then it contains a bipartite subgraph with at least $\frac{n}{2n-1}e$ edges.
If G has $2n+1$ vertices and e edges, then it contains a bipartite subgraph with at least $\frac{n+1}{2n+1}e$ edges

PIN

14: Given vectors $v_1, \dots, v_n \in R^n$ with $|v_i| = 1$. Show that there exist $\varepsilon_1, \dots, \varepsilon_n = \pm 1$ such that

$$|\varepsilon_1 v_1 + \cdots + \varepsilon_n v_n| \leq \sqrt{n},$$

Discrepancy

and also there exist $\varepsilon_1, \dots, \varepsilon_n = \pm 1$ such that

$$|\varepsilon_1 v_1 + \cdots + \varepsilon_n v_n| \geq \sqrt{n}.$$

Hint: pick ε_i randomly And show $\mathbb{E}[|\varepsilon_1 v_1 + \cdots + \varepsilon_n v_n|^2] = n$

Let $\varepsilon_i = \begin{cases} -1 & w/ \text{ prob } 1/2 \\ 1 & 1/2 \end{cases}$ Let $X = \left|\sum \varepsilon_i v_i\right|^2$

$= \sum_{i,j} \varepsilon_i \varepsilon_j (v_i \cdot v_j)$

$\mathbb{E}[X] = \sum_{i,j} \mathbb{E}[\varepsilon_i \varepsilon_j (v_i \cdot v_j)] = \sum_{i,j} \mathbb{E}[\varepsilon_i \varepsilon_j](v_i \cdot v_j)$

$= \sum_i \underbrace{\mathbb{E}[\varepsilon_i^2]}_{1}\underbrace{(v_i \cdot v_i)}_{1} + \underbrace{\sum_{i \neq j} \mathbb{E}[\varepsilon_i \varepsilon_j](v_i \cdot v_j)}_{0}$

$= n(1)(1)$

15: Given vectors $v_1, \dots, v_n \in R^n$ with $|v_i| \le 1$. Let $p_1, \dots, p_n \in [0,1]$ be arbitrary, and set $w = p_1 v_1 + \cdots + p_n v_n$. Then there exist $\varepsilon_1, \dots, \varepsilon_n \in \{0,1\}$ so that set $v = \varepsilon_1 v_1 + \cdots + \varepsilon_n v_n$, we have

$$|w - v| \le \frac{\sqrt{n}}{2}.$$

Choose ε_i indep. Bernoulli with prob p_i Ber(p_i)

$\varepsilon_i \sim \text{Ber}(p_i) \iff \begin{array}{l} Pr(\varepsilon_i = 1) = p_i \\ Pr(\varepsilon_i = 0) = 1 - p_i \end{array}$

Let $X = |w - v|^2$

$\mathbb{E}(X) = \mathbb{E}[|w-v|^2] = \sum_{i,j} \underbrace{\mathbb{E}[(p_i - \varepsilon_i)(p_j - \varepsilon_j)]}_{*} v_i \cdot v_j$

If $i = j$, then $* = p_i(1-p_i) \le 1/4$

If $i \neq j$, then $* = 0$, by indep.

Therefore $\mathbb{E}[X] \le \frac{n}{4}$

16: Let F be a family of subsets of $[n] = \{1, \dots, n\}$ such that there are no $A, B \in F$ satisfying $A \subset B$. Let σ be a random permutation of $[n]$. Consider the random variable $X = |\{i : \{\sigma(1), \sigma(2), \dots, \sigma(i)\} \in F\}|$. Prove $|F| \le \binom{n}{\lfloor n/2 \rfloor}$ by considering the expectation of X.

Sperner's Theorem for antichains in posets.

$X \le 1$ $\mathbb{E}[X] \le 1$

$Pr(A \text{ is chosen}) = \frac{|A|!(n-|A|)!}{n!} = \binom{n}{|A|}^{-1} \ge \binom{n}{\lfloor n/2 \rfloor}^{-1}$

$|F| \binom{n}{\lfloor n/2 \rfloor}^{-1} \le \mathbb{E}[X] \le 1$

$$\mathbb{E}\left[(p_i-\varepsilon_i)^2\right] = \underbrace{p_i}_{Pr}(p_i-1)^2 + \underbrace{(1-p_i)}_{Pr}(p_i-0)^2$$

$$= p_i(1-p_i)\left[(1-p_i)+p_i\right] = p_i(1-p_i)$$

Some estimates:

$1-x \geq e^{-x-x^2}$

if $x \geq -\frac{1}{2}$

$$n! \le n^n \qquad n! \approx \sqrt{2\pi n}\left(\frac{n}{e}\right)^n$$

$$\left(\frac{n}{e}\right)^n \le n! \le en\left(\frac{n}{e}\right)^n \qquad \left(\frac{n}{k}\right)^k \le \binom{n}{k} \le \left(\frac{en}{k}\right)^k$$

$$\frac{2^{2m}}{\sqrt{2m}} \le \binom{2m}{m} \le \frac{2^{2m}}{2\sqrt{m}}$$

$$(1-p)^m \le e^{-pm} \qquad (1-p) \ge e^{-2p} \text{ for } 0 \le p \le \frac{1}{2}$$

$1-x \leq e^{-x}$

17: *(Bonus)* Let $(\Omega, 2^\Omega, P)$ be a finite probability space, where all elementary events have the same probability. Show that if $|\Omega|$ is a prime, then there does not exist a pair of non-trivial independent events. Trivial events are $\emptyset$ and Ω.