

Probability and Computing – Probabilistic Method

Stefan Walzer | WS 2024/2025



The Probabilistic Method (pioneered by Paul Erdős)

Show that something exists by proving that it has a positive probability of arising from a random process.

- Used to prove statements that don't involve randomness at all.
- Probabilistic arguments replace combinatorial arguments.

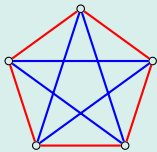
First Example: Ramsey Numbers

Definition: Ramsey Number

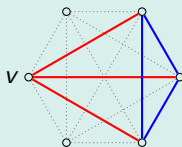
$R(k, k) := \min\{n \in \mathbb{N} \mid \text{any red-blue colouring of the edges of } K_n \text{ contains a monochromatic } k\text{-clique}\}.$ ^a

^aThe general definition of $R(r, b)$ asks for red r -clique or blue b -clique.

$R(3, 3) > 5$



$R(3, 3) \leq 6$



- v has 3 incident edges of the same colour
- wlog that colour is red
- if there is no red triangle then w_1, w_2, w_3 form a blue triangle.

Hence: $R(3, 3) = 6$.

First Example: Ramsey Numbers

Definition: Ramsey Number

$R(k, k) := \min\{n \in \mathbb{N} \mid \text{any red-blue colouring of the edges of } K_n \text{ contains a monochromatic } k\text{-clique}\}.$

Theorem: $R(k, k) > 2^{k/2}$ for $k \geq 6$. // actually $k \geq 3$ suffices

Proof.

- To show: Edges of K_n with $n \leq 2^{k/2}$ can be coloured while avoiding a monochromatic k -clique.
- Plan: Show that *uniformly random colouring* avoids monochromatic k -clique with positive probability.
- There are $\binom{n}{k}$ k -cliques. Each is monochromatic with probability $2^{-\binom{k}{2}+1}$.
- The number M of monochromatic k -cliques satisfies:

$$\mathbb{E}[M] = \binom{n}{k} \cdot 2^{-\binom{k}{2}+1} \leq \frac{n^k}{k!} \cdot 2^{-k^2/2+k/2+1} \leq \frac{(2^{k/2})^k}{(k/2)^{k/2}} \cdot 2^{-k^2/2} 2^{k/2} 2 = 2 \left(\frac{4}{k}\right)^{k/2} < 1.$$

Since $\mathbb{E}[M] < 1$ it is possible that $M = 0$. In particular a colouring with no monochromatic k -cliques exists. \square

We have implicitly used:

$$\Pr[X \leq \mathbb{E}[X]] > 0 \text{ and } \Pr[X \geq \mathbb{E}[X]] > 0.$$

Probabilistic Method with Expectation Argument

Show that an object x with $f(x) \leq b$ exists by proving that a random object X satisfies $\mathbb{E}[f(X)] \leq b$.

Simple Use Case

Any graph $G = (V, E)$ admits a cut of weight at least $|E|/2$.

Proof.

- Assign each $v \in V$ to V_1 or V_2 uniformly at random.
- Each edge crosses the cut (V_1, V_2) with probability $1/2$.
- $\mathbb{E}[\text{weight of } (V_1, V_2)] = \mathbb{E}\left[\sum_{e \in E} [e \text{ crosses } (V_1, V_2)]\right] = \sum_{e \in E} \Pr[e \text{ crosses } (V_1, V_2)] = |E| \cdot \frac{1}{2}. \quad \square$

Example: Independent Sets

Theorem

Let $G = (V, E)$ with $n = |V|$, $m = |E|$ and $m \geq \frac{n}{2}$.

Then there exists an independent set of size $\frac{n^2}{4m} \cdot \Theta\left(\frac{n}{\text{average degree}}\right)$

Proof.

- `sampleAndReject` computes an independent set $V^+ \setminus V^-$.

- $\mathbb{E}[|V^+|] = n \cdot \frac{n}{2m} = \frac{n^2}{2m}$.

- $\mathbb{E}[|V^-|] \leq \sum_{\{u,v\} \in E} \Pr[u \in V^+, v \in V^+] = \sum_{\{u,v\} \in E} \left(\frac{n}{2m}\right)^2 = \frac{n^2}{4m}$.

- $\mathbb{E}[|V^+ \setminus V^-|] = \mathbb{E}[|V^+|] - \mathbb{E}[|V^-|] \geq \frac{n^2}{2m} - \frac{n^2}{4m} = \frac{n^2}{4m}$. \square

Algorithm `sampleAndReject`:

```
V+ ← ∅
for v ∈ V do
  with probability  $\frac{n}{2m}$  do
    V+ ← V+ ∪ {v}
V- ← ∅
for {u, v} ∈ E do
  if u ∈ V+ and v ∈ V+ then
    with probability  $\frac{1}{2}$  do
      V- ← V- ∪ {u}
    otherwise
      V- ← V- ∪ {v}
return V+ \ V-
```

Remark: `sampleAndReject` seems suitable for a parallel / distributed setting.

Context

Given: Family $\mathcal{E} = \{E_1, \dots, E_n\}$ of “bad” events with $\Pr[E_i] \leq p < 1$.

Want: Show $\Pr[\bar{E}_1 \cap \dots \cap \bar{E}_n] = \Pr[\text{none of } \mathcal{E}] > 0$.

Observation: Easy if \mathcal{E} is independent

If \mathcal{E} is an independent family then $\Pr[\text{none of } \mathcal{E}] = \prod_{i=1}^n \Pr[\bar{E}_i] \geq (1 - p)^{|\mathcal{E}|} > 0$.

Observation: Expectation arguments only gets us so far

If $np < 1$ then $\mathbb{E}[\#\text{events from } \mathcal{E} \text{ occurring}] \leq np < 1$, hence $\Pr[\text{none of } \mathcal{E}] > 0$.

If $np = 1$ then $\Pr[\text{none of } \mathcal{E}] = 0$ is possible, e.g. $X \sim \mathcal{U}([n])$ and $E_i := \{X = i\}$.

Lovász Local Lemma (László Lovász and Paul Erdős, 1973)

If each $E \in \mathcal{E}$ has $\Pr[E] < p$ and depends on at most d events^a from \mathcal{E} and $4pd \leq 1$ then $\Pr[\text{none of } \mathcal{E}] > 0$.

^aLittle challenge: State what this means formally.

Example for Lovász Local Lemma (by Wikipedia User *Kevinatilusa*)

Lovász Local Lemma (László Lovász and Paul Erdős, 1973)

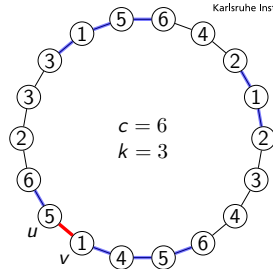
If each $E \in \mathcal{E}$ has $\Pr[E] < p$ and depends on at most d events from \mathcal{E} and $4pd \leq 1$ then $\Pr[\text{none of } \mathcal{E}] > 0$.

Setting

Consider a necklace of ck beads with k beads of each of c colours.

An *independent rainbow* is a set of beads

- containing one bead of each colour // rainbow
- and not containing a pair of adjacent beads. // independent



Claim: If $k \geq 16$ then an independent rainbow always exists. // $k \geq 11$ also suffices

Consider any necklace. Let R contain a random bead of each color. // Goal: $\Pr[R \text{ independent}] > 0$.

One bad event per pair of adjacent beads:

$$E_{\{u,v\}} := \{u \in R \wedge v \in R\}, \quad \Pr[E] \leq \frac{1}{k^2} =: p.$$

$E_{\{u,v\}}$ depends on $E_{\{u',v'\}}$ only if u' or v' share the colour of u or v .

$2k$ relevant beads, hence $4k - 2$ relevant pairs.

$$\Rightarrow d = 4k - 2, \quad 4pd \leq 4 \frac{1}{k^2} (4k - 2) < \frac{16}{k} \leq 1.$$

$$\Pr[R \text{ independent}] = \Pr[\text{none of } (E_{\{u,v\}})_{u,v}] \stackrel{\text{LLL}}{>} 0.$$

Proof of Lovász Local Lemma

Lovász Local Lemma (László Lovász and Paul Erdős, 1973)

If each $E \in \mathcal{E}$ has $\Pr[E] < p$ and depends on at most d events from \mathcal{E} and $4pd \leq 1$ then $\Pr[\text{none of } \mathcal{E}] > 0$.

Claim: $\forall S \subseteq \mathcal{E} : \forall E^* \in \mathcal{E} \setminus S : \Pr[E^* \mid \text{none of } S] \leq 2p$.

Proof of LLL using the Claim.

$$\Pr[\text{none of } \mathcal{E}] = \prod_{i=1}^n \Pr[\bar{E}_i \mid \text{none of } \{E_1, \dots, E_{i-1}\}] \geq (1 - 2p)^n \stackrel{4pd \leq 1}{>} 2^{-n} > 0. \quad \square$$

Proof of Lovász Local Lemma

Lovász Local Lemma (László Lovász and Paul Erdős, 1973)

If each $E \in \mathcal{E}$ has $\Pr[E] < p$ and depends on at most d events from \mathcal{E} and $4pd \leq 1$ then $\Pr[\text{none of } \mathcal{E}] > 0$.

Claim: $\forall S \subseteq \mathcal{E} : \forall E^* \in \mathcal{E} \setminus S : \Pr[E^* \mid \text{none of } S] \leq 2p$.

Proof of the Claim by Induction on $|S|$.

- Base case: If $|S| = 0$ then $\Pr[E^* \mid \text{none of } \emptyset] = \Pr[E^*] \leq p \leq 2p$. \checkmark Let now $|S| > 0$.
- Partition $S = I \dot{\cup} D$ such that E^* is independent of I and $1 \leq |D| \leq d$. $\parallel \leq d$ possible by assumption, > 0 is our choice.
- $\Pr[\text{none of } D \mid \text{none of } I] = 1 - \Pr[\bigcup_{E \in D} E \mid \text{none of } I] \stackrel{\text{UB}}{\geq} 1 - \sum_{E \in D} \underbrace{\Pr[E \mid \text{none of } I]}_{\leq 2p \text{ (Induction, using } |I| < |S|)} \stackrel{4pd \leq 1}{\geq} 1 - 2dp \geq \frac{1}{2}$. (\star).

$$\begin{aligned} \Pr[E^* \mid \text{none of } S] &= \frac{\Pr[E^* \wedge \text{none of } S]}{\Pr[\text{none of } S]} \leq \frac{\Pr[E^* \wedge \text{none of } I]}{\Pr[\text{none of } D \mid \text{none of } I] \Pr[\text{none of } I]} \\ &= \frac{\Pr[E^*] \Pr[\text{none of } I]}{\Pr[\text{none of } D \mid \text{none of } I] \Pr[\text{none of } I]} = \frac{p}{\Pr[\text{none of } D \mid \text{none of } I]} \stackrel{(\star)}{\leq} \frac{p}{1/2} = 2p. \quad \square \end{aligned}$$

What the Probabilistic Method is all About

- Goal: Prove the existence of objects with certain properties.
- Use probabilistic language as a tool.

Vanilla Variant:

Goal: Show that $P \subseteq \Omega$ is not empty.

- 1 Define a random object $X \in \Omega$.
- 2 Show: $\Pr[X \in P] > 0$.
- 3 Conclude: $\exists x \in \Omega : x \in P$.

Variant with Expectation Argument

Goal: Show that $f : \Omega \rightarrow \mathbb{R}$ has maximum at least q .

- 1 Define a random object $X \in \Omega$.
- 2 Show: $\mathbb{E}[f(X)] \geq q$.
- 3 Conclude: $\exists x \in \Omega : f(x) \geq q$.

Variant with Lovász Local Lemma

Goal: Show that $P \subseteq \Omega$ is not empty.

- 1 Define random object X .
- 2 Define family \mathcal{E} of bad events such that $\bigcap_{E \in \mathcal{E}} \bar{E} \Rightarrow X \in P$.
- 4 Show that $E \in \mathcal{E}$ satisfies $\Pr[E] \leq p$.
- 5 Show $E \in \mathcal{E}$ depends on at most d other events from \mathcal{E} .
- 6 Show $4dp \leq 1$.
- 7 Conclude with LLL: $\exists x : x \in P$.

Anhang: Mögliche Prüfungsfragen I

- Was ist das Ziel der probabilistischen Methode?
- Bezüglich der grundlegenden Methode:
 - Welche “Kreativleistung” muss man erbringen und was muss man dann ausrechnen?
 - Verdeutliche die Methode an einem Beispiel.
- Bezüglich der Variante mit Erwartungswertargument:
 - Welche “Kreativleistung” muss man erbringen und was muss man dann ausrechnen?
 - Verdeutliche die Methode an einem Beispiel.
 - Wir haben gezeigt, dass jeder Graph einen Schnitt von Gewicht $|E|/2$ besitzt. Wie?
 - Wir haben gezeigt, dass jeder Graph eine unabhängige Menge der Größe $\frac{n^2}{4m}$ besitzt. Wie?
- Bezüglich Lovász Local Lemma:
 - Formuliere die Aussage des Lemmas.
 - Was ist der Bezug zur probabilistischen Methode?
 - Wir haben gezeigt, dass gefärbte Graphen unabhängige Regenbogenmengen gewisser Größe besitzen. Wie sind wir vorgegangen?