③ Turan's Theorem '1941 [AS]

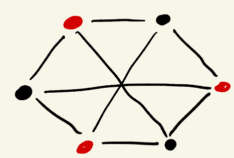
"If in a group of 100 people each person knows 3 other people, then there are 25 people who don't know each other".

More generally:

every vertex has exactly d neighbors

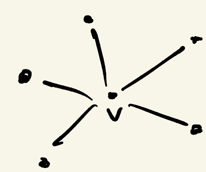
Thm $\forall$ d-regular graph on n vertices $G(V,E)$
$\exists$ an independent set of cardinality $\geq \frac{n}{d+1}$

a subset of vertices no two of which are neighbors



Independent set

Proof Choose a random ordering of V (uniformly)



- $I := \{v \in V : v$ is the least element among v and its neighbors$\}$

$$\mathbb{E}|I| = \mathbb{E}\sum_{v\in V} \mathbb{1}_{\{v\in I\}} = \sum_{v\in V} \underbrace{P\{v\in I\}}_{= \frac{1}{d+1}} = \frac{n}{d+1}.$$

(prob. that v is the least among $d+1$ elements in random order)

- I is an independent set:

if $u, v \in I$ and $(u,v) \in E$ then $u < v$ and $v < u \Rightarrow \lightning$ $\square$.

Remark: Turan's bound is optimal for the union of $\frac{n}{d+1}$ cliques of size $d+1$.

④ k-SAT problem

- A Boolean formula on n literals is a function $\phi : \{\underset{\text{TRUE}}{T}, \underset{\text{FALSE}}{F}\}^n \to \{T, F\}$
- ϕ is satisfiable if $\exists$ an assignment x such that $\Phi(x) = T$
- $\forall$ Boolean formula can be expressed in the conjunctive normal form (CNF): AND of ORs (of x_i's and $\bar{x}_i$'s), e.g.

$\bar{x}_i$ ← negation

$$\phi(x) = (x_1 \overset{\text{OR}}{\vee} \bar{x}_2) \overset{\text{AND}}{\wedge} (x_1 \vee \bar{x}_2 \vee x_3) \wedge \bar{x}_1 .$$

Even 3-sat $\in$ NP

Q Consider a random "k-SAT" formula with rn clauses each of length k, where each literal is chosen independently uniformly from $\{x_1, \dots, x_n, \bar{x}_1, \dots, \bar{x}_n\}$. What is the probability that ϕ is satisfiable?

$$\phi(x) = \underbrace{\underbrace{(? \vee ? \vee \dots \vee ?)}_{k} \wedge \underbrace{(? \vee ? \vee \dots \vee ?)}_{k} \wedge \dots \wedge \underbrace{(? \vee ? \vee \dots \vee ?)}_{k}}_{rn \text{ clauses}}$$

each "?" can be a literal x_i or its negation $\bar{x}_i$

Thm If $r > 2^k \ln 2$ and $n \to \infty$ (while r, k are fixed), then $P\{\phi \text{ is satisfiable}\} \to 0$

Proof $P\{\phi \text{ is satisfiable}\} = P\{\exists x \in \{T,F\}^n : \phi(x) = T\}$

$$\le \sum_{x \in \{T,F\}^n} P\{\phi(x) = T\} \;\; \text{(=)} \quad \text{(union bound)}$$

$$\phi(x) = \phi_1(x) \wedge \dots \wedge \phi_{rn}(x) \implies$$

$$P\{\phi(x) = T\} = P\{\phi_i(x) = 1 \;\; \forall i = 1, \dots, rn\} \overset{\text{indep}}{=} \prod_{i=1}^{rn} \underbrace{P\{\phi_i(x) = 1\}}_{\substack{\parallel \\ 1 - \left(\frac{1}{2}\right)^k}} = \left(1 - \frac{1}{2^k}\right)^{rn}$$

(since $\phi_k = k$ or's)

$$\text{(=)} \;\; 2^n \left(1 - \frac{1}{2^k}\right)^{rn} \overset{1-x \le e^{-x}}{\le} \Big(\underbrace{2e^{-r/2^k}}_{\substack{\wedge \\ 1 \text{ if } r > 2^k \ln 2}}\Big)^n \to 0 \text{ as } n \to \infty. \;\square$$

Remark If $k \ge 3$, $r < 2^k \ln 2 - k$, $P\{\phi \text{ is satisfiable}\} \to 1$ $(n \to \infty)$ [Alon-Naor-Peres 05]

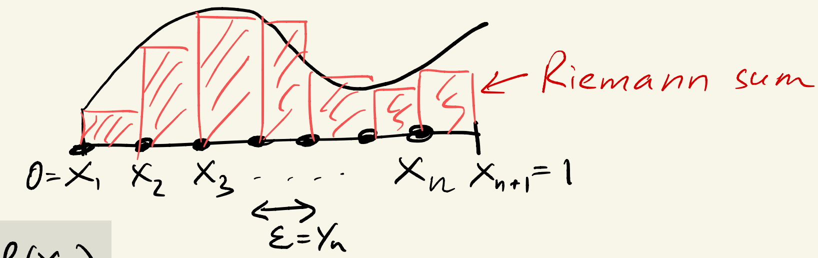
5 Monte-Carlo Integration

Problem For a given function $f: [0,1]^d \to \mathbb{R}$, numerically compute

$$\int_0^1 \cdots \int_0^1 f(x_1, \dots, x_d)\, dx_1 \cdots dx_d = \int_{[0,1]^d} f(x)\,dx$$

METHOD 1

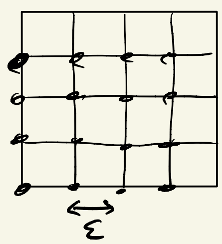
- If $d = 1$: use the grid



$$\int_0^1 f(x)\,dx \approx \sum_{i=1}^{n} \underbrace{(x_{i+1} - x_i)}_{\overset{!!}{\varepsilon = 1/n}} f(x_i) = \boxed{\frac{1}{n}\sum_{i=1}^{n} f(x_i)}$$

"Resolution"

- If $d = 2$, do similarly but use the grid



$\Rightarrow$ $n = (1/\varepsilon)^2$ pts.

- For general dimension d, we need $n = (1/\varepsilon)^d$ pts.

Exponential in d. Too large.

"THE CURSE OF DIMENSIONALITY" ☹

Method 2 : Monte-Carlo

- Instead of choosing X_i on the grid, choose them at random independently: $X_i \sim \mathrm{Unif}([0,1]^d)$

$$\frac{1}{n}\sum_{i=1}^{n} f(X_i) \overset{?}{\approx} \int_{[0,1]^d} f(x)\,dx$$

$\Rightarrow$ $f(X_i)$ are i.i.d. r.v.'s.

- $$\mathbb{E}\left[\frac{1}{n}\sum_{i=1}^{n} f(X_i)\right] = \frac{1}{n}\sum_{i=1}^{n} \underbrace{\mathbb{E}f(X_i)}_{\mathbb{E}f(X) \text{ by identical distribution}} = \mathbb{E}f(X)$$

$$= \int_{\mathbb{R}^d} f(x)\,p(x)\,dx, \quad \text{where density is } p(x) = \begin{cases} 1, & x \in [0,1]^d \\ 0, & \text{———} \end{cases}$$

$$= \int_{[0,1]^d} f(x)\,dx. \qquad ☺$$

$\Rightarrow$ we have an unbiased estimator of the integral

- Rate of convergence? L^2 error ("MSE"):

$$\mathbb{E}\Big(\underbrace{\frac{1}{n}\sum_{i=1}^{n} f(X_i)}_{=:Z} - \underbrace{\int_{[0,1]^d} f(x)\,dx}_{=\mathbb{E}Z}\Big)^2 = \mathrm{Var}\left(\frac{1}{n}\sum_{i=1}^{n} f(X_i)\right)$$

$$= \frac{1}{n^2}\sum_{i=1}^{n} \underbrace{\mathrm{Var}(f(X_i))}_{\mathrm{Var}(f(X)) \text{ by identical distribution}} = \frac{\mathrm{Var}(f(X))}{n}$$

$$\le \frac{1}{n} \quad \text{e.g. if } |f(x)| \le 1 \ \forall x.$$

- Chebyshev's inequality $\Rightarrow$ with probability ≥ 0.99,

$$\left|\frac{1}{n}\sum_{i=1}^{n} f(x_i) - \int_{[0,1]^d} f(x)\,dx\right| \le \frac{10}{\sqrt{n}} \qquad ☺$$

Regardless of dimension d!