LECTURE 3: SAMPLE & MODIFY / ALTERATION.

Highlevel idea:

indirect 2-stage argument.

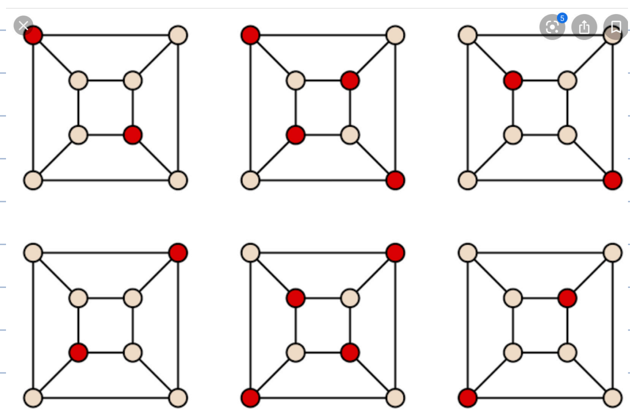
① sample: to construct random structure with some "blemishes".

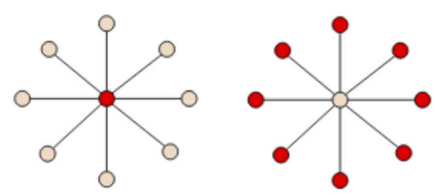
② modify: to satisfy required properties.

§ Application: Independent set.

(6.4.1 in M-U, 3.2 in A-S).

- Given a graph $G := (V, E)$, find the maximum sized independent set (also known as stable set, co-clique, anticlique), i.e. a set of vertices s.t. no two are adjacent.





Two independent sets for the star graph S_8 show how vastly different in size two maximal independent sets (the right being maximum) can be.

$\alpha(G)$ is the independence no. of graph G.

$\alpha(G) \geq t$ means $\exists\, t$ vertices with no edges between them

- NP hard!

• **Theorem**: Let $G=(V,E)$ be a connected graph on n vertices and m edges. Then $\alpha(G) \geq n^2/4m$.

$\to m \geq n-1$

$\to$ example : $m=n$.

Proof: Let $d = \frac{2m}{n} \geq 1$ be the average degree of vertices in G.

Consider following randomized algo:

1. Delete each vertex of G (together with its incident edges) w.p. $1 - 1/d$.
2. For each remaining edge, remove it and one of its adjacent vertices.

• Let X_i be the indicator random variable that vertex v_i survives the 'sample' step.

Then $\mathbb{E}[X_i] = 1/d$.

• Let X be #vertices survived after the 'sample' step.

Lin. of expectations

Then, $\mathbb{E}[X] = \mathbb{E}[\sum_i X_i] \overset{\uparrow}{=} \sum_{i=1}^{n} \mathbb{E}[X_i] = \frac{n}{d}$.

• Let Y_j be the indicator random variable that edge j survives 'sample' step.

Then $\mathbb{E}[Y_j] = \left(\frac{1}{d}\right)^2$ [As both its endpoints need to survive].

• Let Y be total number of edges survived after 'sample' step.

Then, $\mathbb{E}[Y] = \mathbb{E}[\sum_j Y_j] = \sum_{j=1}^{m} \mathbb{E}[Y_j] = \frac{nd}{2} \cdot \frac{1}{d^2}$

$= \frac{n}{2d}$.

In the second step, algorithm removes at most one vertex per edge.

Hence, it outputs an independent set of size $\geq X - Y$.

Now, $\mathbb{E}[X-Y] = \mathbb{E}[X] - \mathbb{E}[Y]$

$= \frac{n}{d} - \frac{n}{2d} = n/2d = n^2/4m$.

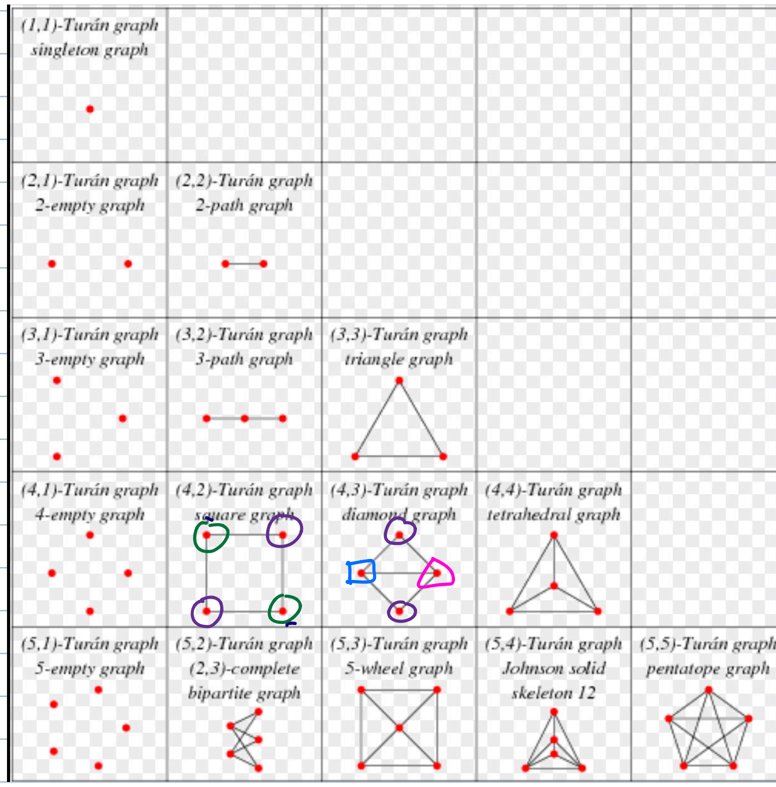
• This proves a weaker version of celebrated Turán's theorem.

• Turán's theorem : $\alpha(G) \geq \frac{n}{d+1}$. (A)

[Equivalent version : (B) Let G be a n-vertex K_{r+1}-free graph. Then it has at most $\frac{r-1}{r} \cdot \frac{n^2}{2}$ edges.]

This is infact tight:

Turán graph $T(n,r)$ is a complete multi-partite graph formed by partitioning a set of n vertices into r subsets, as equal as possible & connect two vertices iff they belong to different subsets.



Say, $r \mid n$, $\#\text{edges} = \frac{n}{2}\cdot\left(n - \frac{n}{r}\right) = n^2 \cdot \frac{(r-1)}{2r}$.

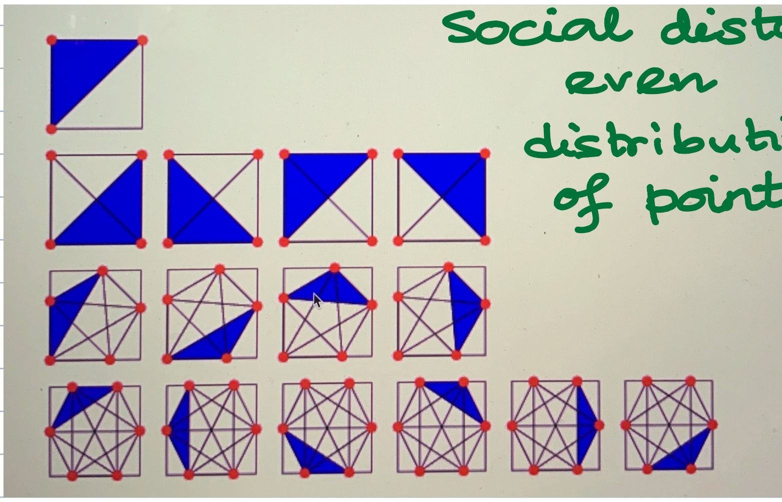
- Home work (not for submission): Show (A) & (B) are equivalent.

- Also check "proofs from the book" for many different proofs of Turán's theorem.

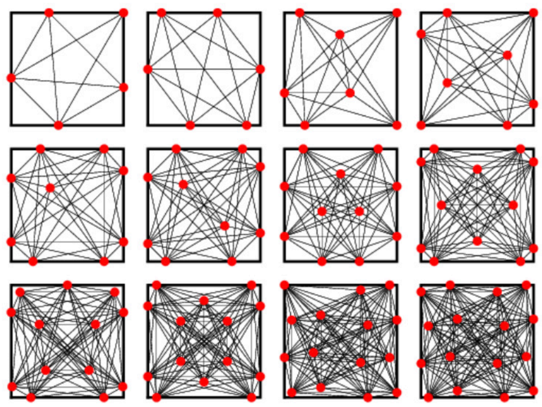
§ Application: Combinatorial Geometry.

- **Heilbronn's triangle problem:**

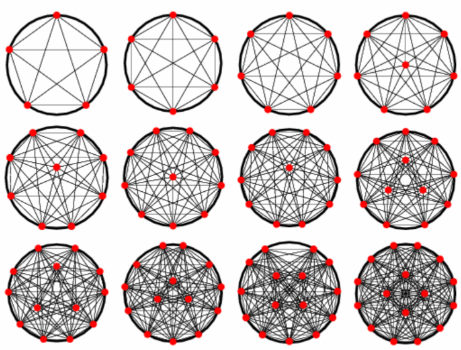
Place n points in a unit square, so as to maximize the area of the minimum $\triangle$ whose vertices are 3 of the n points.

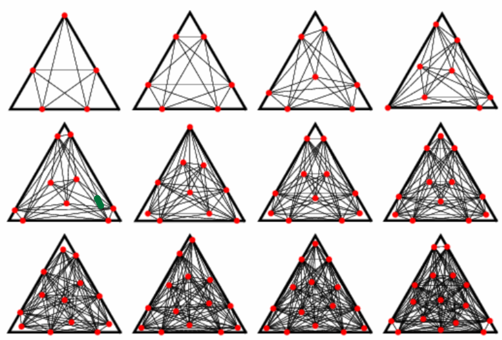


Social distancing: even distribution of points.



many variants are studied:





- Let S be a set of n points in $[0,1]\times[0,1]$.
- $T(S)$ be the min area of a triangle whose vertices are three distinct points of S.
- Let $T(n) = \max\limits_{S} T(S)$.
- Conjecture (Heilbronn): $O(1/n^2)$.
- Komlós, Pintz, Szemeredi (1982):

$T(n) = \Omega(\log n / n^2)$ using prob. methods.

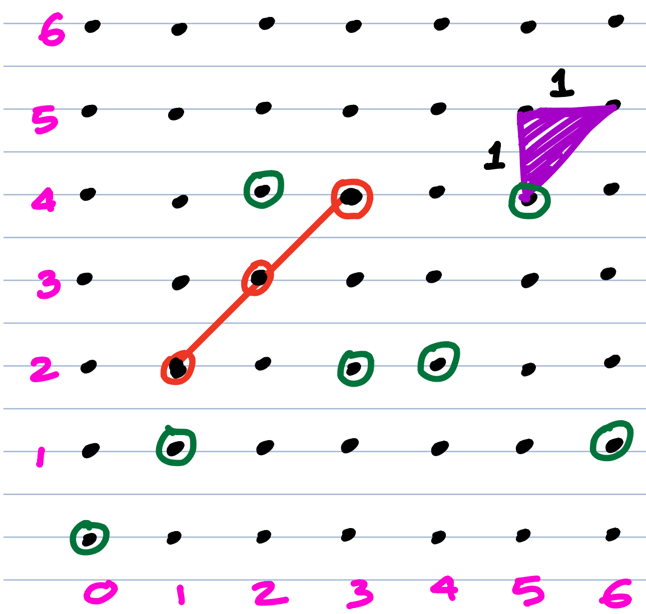
- Current best known upper bound:
  $T(n) \leq n^{-8/7 + o(1)}$. [KMS'81].

- We will see a simpler weaker result.

Theorem: (Thm 3.3.1 in A-S).
$\exists$ S of n points in unit square s.t.
$$T(S) \geq 1/(100n^2).$$

⊙ A nonprobabilistic proof:
[Due to Erdös, showing $1/2(n-1)^2$].

Let us start with a $n \times n$ uniform grid on $[0, n-1] \times [0, n-1]$



Problem: many points are collinear, giving Δ. area $= 0$.

If Δ vertices are not collinear, then minimum area triangle has area:
$$\frac{1}{2} \cdot 1 \cdot 1 = \frac{1}{2}. = \frac{1}{2(n-1)^2} \cdot (n-1)^2.$$

So we want to remove collinear points.
Let n be a prime.
Consider a set of n points from n^2 in grid. S:
(x, y), where $y \equiv x^2 \bmod n$, $0 \leq y < n$.
i.e. $S = \{(x, x^2) \in \mathbb{F}_n^2 : x \in \mathbb{F}_n\}$.

So, these points define a parabola.

A parabola meets a line $y = mx + b$ at ≤ 2 points.

[Otherwise, $x^2 - mx - b = 0$ has three distinct roots. A contradiction!]

Contracting the plane by a factor $(n-1)$ in both coordinates gives the desired set of n points with min Δ-area $\geq \frac{1}{2(n-1)^2}$.

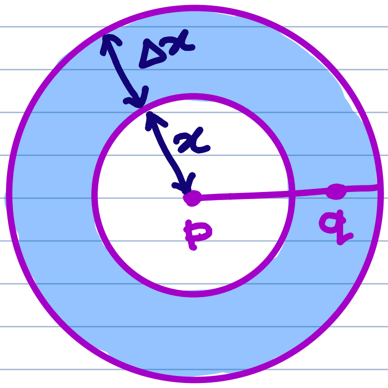
Often algebraic solutions are cute, but hard to modify/extend.

Combinatorial proofs might help us to use heavier hammers.

Proof:

- Let us sample 3 points p, q, r indeply. uniformly at random in unit square

What is the probability that the area of a Δpqr is at most ε?



Pick p first. Let $d_1 = dist(p, q)$.

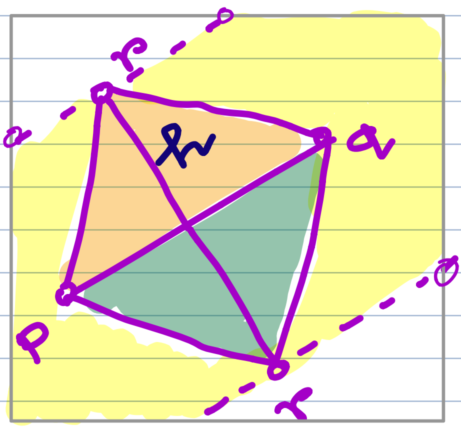
$\Pr[d_1 \in [x, x+\Delta x]]$

$\leq \pi(x+\Delta x)^2 - \pi x^2$

$\leq \pi(\Delta x . \Delta x + 2x\Delta x)$.

$\leq 7\, x . \Delta x$ for small Δx.

Now, fixing p and q at distance x, let h be altitude from r to line pq.



For area$(\Delta pqr) \leq \epsilon$,

we need $\frac{1}{2} . h x \leq \epsilon$

$\Rightarrow h \leq 2\epsilon / x$.

So, r must lie in a strip of width $4\epsilon/x$ and length $\sqrt{2}$ [As $\sqrt{2}$ is the maximum length of a line segment completely contained in a unit square].

The prob that r has above property

$$\le \frac{4\epsilon}{x} \cdot \sqrt{2}.$$

$$\Pr[\Delta pqr \text{ has area} \le \epsilon] \le 7x \Delta x \cdot \frac{4\epsilon}{x}\sqrt{2}$$

$$\le 40\, \epsilon \Delta x.$$

As, $0 \le x \le \sqrt{2}$, $\Pr[\Delta pqr \text{ has area} \le \epsilon]$

$$\le \int_0^{\sqrt{2}} 40\,\epsilon\, dx = 40\sqrt{2}\,\epsilon \le 60\,\epsilon.$$

① Sample:

- Choose $2n$ points, independently and uniformly at random in $[0,1] \times [0,1]$.

Let X denote the number of Δ's with area $\le 1/(100n^2)$.

For each triplet of points p, q, r

$$\Pr\left[\text{area}(\Delta pqr) \le \frac{1}{100n^2}\right] \le \frac{60}{100n^2}.$$

There are $\binom{2n}{3}$ such triplets.

Hence, $\mathbb{E}[X] \le \binom{2n}{3} \cdot \left(\frac{0.6}{n^2}\right)$

$$\le \frac{8n^3}{6} \cdot \left(\frac{0.6}{n^2}\right) < n.$$

② MODIFY / ALTER:

There exists a specific set of $2n$ vertices with fewer than n Δs of area $\leq 1/(100n^2)$.

Delete one vertex from the set from each such triangle.

This leaves at least n vertices, and now no triangle has area less than $1/100n^2$.

LECTURE 4: Second moment method.

Variance: of random variable X

$Var[X] = \mathbb{E}[X - \mathbb{E}[X]]^2 = \mathbb{E}[X^2] - \mathbb{E}[X]^2.$

• Chebyshev's inequality.

X be a random variable (RV) with $\mathbb{E}X < \infty$, variance $Var(X) < \infty$. Then for any $t > 0$,

$$\Pr[\,|X - \mathbb{E}[X]| \geq t\,] \leq \frac{Var[X]}{t^2}.$$

$$\Leftrightarrow \Pr[\,|X - \mathbb{E}[X]| \geq t\sigma\,] \leq \frac{1}{t^2}.$$

• Corollary 1.

$$\Pr(\,|X - \mathbb{E}[X]| \geq \varepsilon\, \mathbb{E}[X]) \leq \frac{Var[X]}{\varepsilon^2\, \mathbb{E}[X]^2}.$$

Comment: If X is an integer valued RV. Then

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

So, if $\mathbb{E}[X] \to 0$, then $\Pr[X > 0] \to 0$ i.e. $X \sim 0$ a.s.

However, if $\mathbb{E}[X] \to \infty$, that do not imply $X > 0$ a.s.

For that second moment is useful.

• **Corollary 2.** $\Pr[X=0] \le \mathrm{Var}[X] / (\mathbb{E}[X])^2$.

• **Corollary 3.**

If $\mathrm{Var}[X] = o(\mathbb{E}[X]^2)$, then w.h.p. $X > 0$.

Infact, we have a stronger property:

$X \sim \mathbb{E}[X]$ w.h.p

Summary:

• Use Markov/ first moment method, if you want to show that some nonnegative RV is 0 with high probability, (by showing $\mathbb{E}X \to 0$).

• Use Chebyshev/ second moment method if you want to show that it is nonzero with high probability by showing that variance / $(\text{mean})^2$ tends to 0.

- **Covariance**: $\mathrm{Cov}[X,Y]$

$= \mathbb{E}[(X-\mathbb{E}[X])(Y-\mathbb{E}[Y])] = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$.

- If X, Y are independent, then $\mathrm{Cov}[X,Y] = 0$.

- Let $X = X_1 + \ldots + X_m$, where X_i is the indicator RV for event A_i.

For indices $i, j (\neq i)$, if A_i, A_j are not independent, we write $i \sim j$

We set (the sum over ordered pairs)

$$\Delta = \sum_{i \sim j} \Pr[A_i \wedge A_j].$$

Now, $\mathrm{Cov}[X_i, X_j] = \mathbb{E}[X_i X_j] - \mathbb{E}[X_i]\mathbb{E}[X_j]$

$\leq \mathbb{E}[X_i X_j] = \Pr[A_i \wedge A_j]$, for $i \sim j$

and $\mathrm{Cov}[X_i, X_j] = 0$, for not $i \sim j$, $i \neq j$.

Lemma 1: $\mathrm{Var}[X] \leq \mathbb{E}[X] + \Delta$, [Lem 6.9 M-U]

Proof: $\mathrm{Var}[X] = \mathrm{Var}\left[\sum_{i=1}^{m} X_i\right]$

$$= \sum_{i=1}^{m} \mathrm{Var}[X_i] + \sum_{\substack{1\leq i,j\leq m \\ i\neq j}} \mathrm{Cov}(X_i, X_j).$$

Now, as X_i is 0/1 RV, $\mathbb{E}[X_i^2] = \mathbb{E}[X_i]$,

Thus, $\mathrm{Var}[X_i] = \mathbb{E}[X_i^2] - (\mathbb{E}(X_i))^2$

$$\leq \mathbb{E}[X_i^2] = \mathbb{E}[X_i].$$

Hence, $\sum_{i=1}^{m} \mathrm{Var}[X_i] = \sum_{i=1}^{m} \mathbb{E}[X_i] = \mathbb{E}\left[\sum_{i=1}^{m} X_i\right] = \mathbb{E}X.$

Also, $\sum_{\substack{1\leq i,j\leq m \\ i\neq j}} \mathrm{Cov}(X_i, X_j) = \sum_{i\sim j} \mathrm{Cov}(X_i, X_j) + \sum_{i\not\sim j} \mathrm{Cov}(X_i, X_j)$

$$= \sum_{i\sim j} \Pr[A_i \wedge A_j] = \Delta.$$

$\Rightarrow \mathrm{Var}[X] \leq \mathbb{E}[X] + \Delta.$

- **Corollary 4:** If $\mathbb{E}[X] \to \infty$, $\Delta = o(\mathbb{E}[X]^2)$, [i.e. $\mathrm{Var}[X] = o(\mathbb{E}[X]^2)$] then almost always $X > 0$ and $X \sim \mathbb{E}X$.

§ Application: Threshold behaviour in Random Graphs.

① Random Graphs:

#vertices   prob of edge

The random graph $G(n,p)$ is a probability space over the set of graphs on the vertex set $[n]$, determined by

$$\Pr[\{i,j\} \in E(G)] = p.$$

with these events mutually independent.

- Also called Erdős-Rényi random graph.

ON THE EVOLUTION OF RANDOM GRAPHS

by

P. ERDŐS and A. RÉNYI

Dedicated to Professor P. Turán at his 50th birthday.

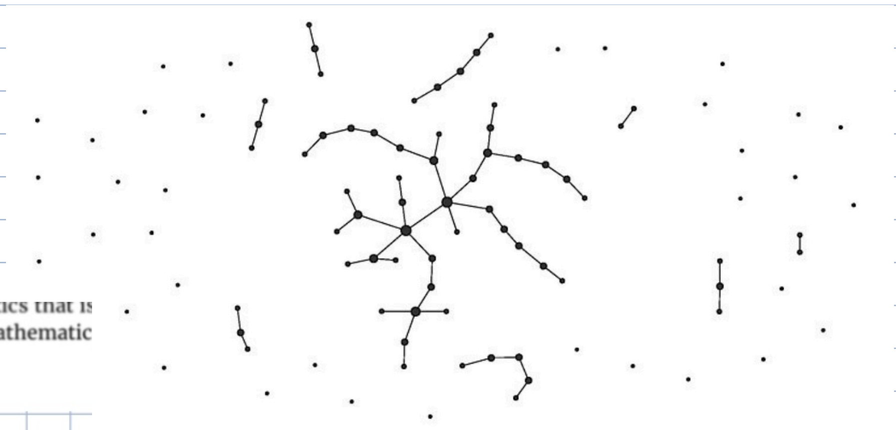
Random Graphs
books.google.co.in › books
This is a revised and updated version of the classic first edition.
Béla Bollobás, Bollobás Béla - 2001 - Preview - More editions

Random Graphs
books.google.co.in › books
Poised to stimulate research for years to come, this new work cove
providing a much-needed, modern overview of this fast-growing ar
Svante Janson, Tomasz Luczak, Andrzej Rucinski - 2011 - Preview

Random Graphs
books.google.co.in › books
Results of research on classical combinatorial structures such as r
systems of random linear equations in finite fields.
V. F. Kolchin, Valentin Fedorovich Kolchin, G. C. Rota - 1999 - Prev

Introduction to Random Graphs
books.google.co.in › books
The text covers random graphs from the basic to the advanced, in
recommendations for further reading.
Alan Frieze, Michał Karoński - 2016 - Preview - More editions

$G(n,p)$ with $p = 0.01$



atics that is
mathematic

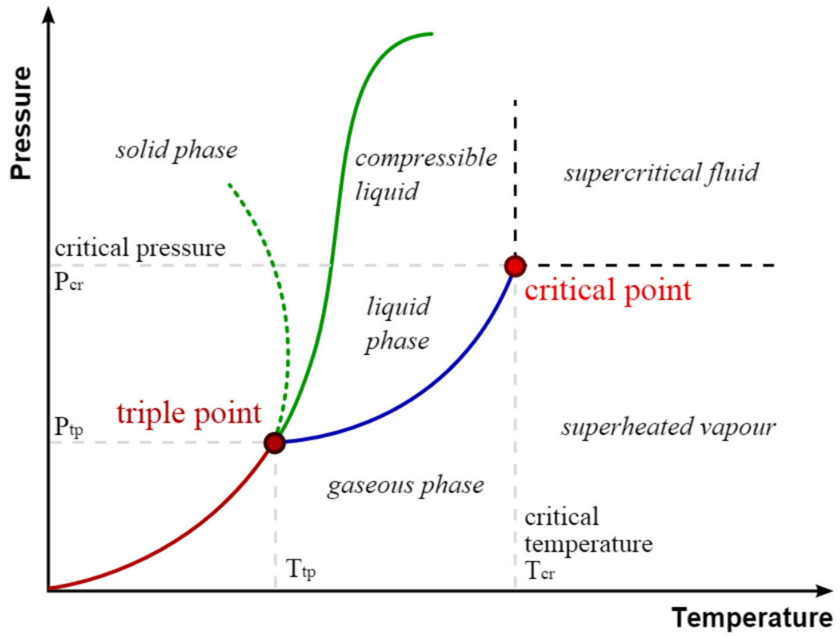
$p = 0 \Rightarrow$ all isolated vertices.

As p increases, it becomes denser.

$p = 1 \Rightarrow G(n,1) = K_n$.

② Threshold behaviour:

phase transition



Curie point

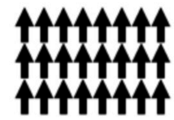


Figure 1. Below the Curie temperature, neighbouring magnetic spins align parallel to each other in ferromagnet in the absence of an applied magnetic field

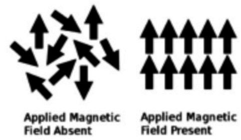


Figure 2. Above the Curie temperature, the magnetic spins are randomly aligned in a paramagnet unless a magnetic field is applied

- The second moment method is used to prove certain properties of random graphs.

- In $G(n,p)$ model often there is a threshold function f such that:

(a) $p < f(n) \Rightarrow$ almost no graph has the desired property.

(b) $p > f(n) \Rightarrow$ almost every graph has the desired property.

Theorem: (Thm 4.4.1 in A-S, 6.8 in M-U).

Let $\omega(G)$ be the number of vertices in the maximum clique of graph G.

Let Π be the property that $\omega(G) \geq 4$.

Then Π has threshold fn $n^{-2/3}$.

[**Reformulation**:

for any $\varepsilon > 0$, & sufficiently large n.

① if $p = o(n^{-2/3})$, then

$\Pr[\omega(G) \geq 4 \text{ for } G \sim G(n,p)] \leq \varepsilon.$

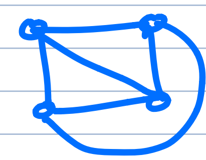
② if $p = \omega(n^{-2/3})$,

$\Pr[\omega(G) \geq 4 \text{ for } G \sim G(n,p)] \geq 1 - \varepsilon.$]

Proof:

Let S be a set of four vertices in $G(n,p)$ and A_S be the event: "S is a clique", and X_S its indicator RV.

Then $\mathbb{E}[X_S] = \Pr[A_S] = p^6.$



K_4 has six edges.

Let X be the total number of 4-cliques in G, i.e.

$$X = \sum_{\substack{|S|=4,\\ S \subset V(G)}} X_S$$

Hence, $\omega(G) \geq 4 \iff X > 0$.

Linearity of expectations gives

$$\mathbb{E}[X] = \sum_{|S|=4} \mathbb{E}[X_S] = \binom{n}{4} p^6 \leq \frac{n^4 p^6}{24}. \quad \cdots (*)$$

If $p = o(n^{-2/3})$, $\frac{n^4 p^6}{24} = o(n^{-4} \cdot n^4) = o(1)$.

So, $\mathbb{E}[X] \leq \varepsilon$ for sufficiently large n.

$\therefore \Pr[X \geq 1] \leq \mathbb{E}[X] \leq \varepsilon$.

Thus, $\Pr[\omega(G) \geq 4 \text{ for } G \sim G(n,p)] \leq \varepsilon$.

Now suppose $p = \omega(n^{-2/3})$.

Then, $\mathbb{E}[X] = \frac{n^4 p^6}{24} = \omega(1) \to \infty$ as $n \to \infty$.

However, this is not sufficient to say that w.h.p. a graph chosen from $G(n,p)$ will have $\omega(G) \geq 4$. [e.g. may not hold true if the $\mathrm{Var}[X]$ is high].

Now, $\Pr[X = 0] \leq \mathrm{Var}[X] / (\mathbb{E}[X])^2$. [Cor. 2]

Thus if $\mathrm{Var}[X] = o(\mathbb{E}[X]^2)$, we get $\Pr[X=0] = o(1)$. $\Rightarrow \Pr[X=0] \leq \varepsilon$.

There $m := \binom{n}{4}$ possible 4-tuples.

Let $C_1, C_2, \ldots, C_m$ be an enumeration of all subsets of four vertices.

Let $X_1, X_2, \ldots, X_m$ are RVs corrs. to $C_1, \ldots, C_m$.

Now from Lemma 1,

$$\mathrm{Var}[X] = \mathrm{Var}\left[\sum_{i=1}^{m} X_i\right]$$
$$\leq \mathbb{E}\left[\sum_{i=1}^{m} X_i\right] - \sum_{\substack{1 \leq i,j \leq m \\ i \neq j}} \mathrm{Cov}(X_i, X_j)$$

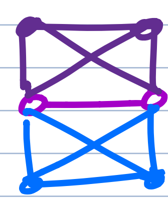
Now, $\mathbb{E}\left[\sum_{i=1}^{m} X_i\right] = \binom{n}{4} p^6$. [From (*)].

So, we compute the covariance term

<u>Case 1</u>. $|C_i \cap C_j| \leq 1$.

Then C_i and C_j are edge-disjoint.
Thus X_i and X_j are independent
and $\mathrm{Cov}(X_i, X_j) = 0$.

<u>Case 2</u>. $|C_i \cap C_j| = 2$.



They share a common edge.
For $X_i = X_j = 1$, all eleven edges must appear.

$\therefore \mathrm{Cov}[X_i, X_j] \leq \mathbb{E}[X_i X_j] \leq p^{11}$.

There are $\binom{n}{6}$ ways to choose 6 vertices.

There are $\binom{6}{2;2;2}$ ways to split them into C_i and C_j.

[2 for $C_i \cap C_j$,
2 for $C_i \setminus C_j$ and
2 for $C_j \setminus C_i$].

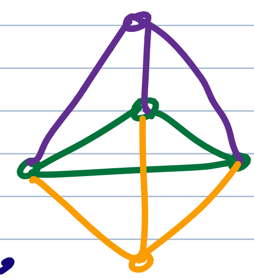
multinomial coefficients

$$\binom{n}{k_1; k_2; \ldots; k_r} = \frac{n!}{k_1! k_2! \cdots k_r!}$$

where $\sum_{i=1}^{r} k_i = n$

Case 3: $|C_i \cap C_j| = 3$.

They share 3 (green) edges.



$\mathrm{Cov}[X_i, X_j] \leq \mathbb{E}[X_i X_j] \leq p^9$,
as all 9 edges must appear.

There are $\binom{n}{5}$ ways to choose 5 vertices.
$\binom{5}{3;1;1}$ ways to split into C_i, C_j.

Hence, $\mathrm{Var}[X]$

$$\leq \binom{n}{4}p^6 + \binom{n}{6}\binom{6}{2;2;2}p^{11} + \binom{n}{5}\binom{5}{3;1;1}p^9$$

$$= O(n^4)p^6 + O(n^6)p^{11} + O(n^5)p^9$$

$$= o(n^8 p^{12}) = o((n^4 p^6)^2) = o((\mathbb{E}[X])^2).$$

as $(\mathbb{E}[X])^2 = \left(\binom{n}{4} p^6\right)^2 = \Theta(n^8 p^{12})$.

which completes the proof $\blacksquare$

• Another alternative proof using conditional expectation inequality. [6.6 in M-U].

• Theorem A: Let $X = \sum_{i=1}^{m} X_i$, where each X_i is a 0-1 RV. Then $\boxed{\Pr(X > 0) \geq \sum_{i=1}^{m} \frac{\Pr(X_i = 1)}{\mathbb{E}[X \mid X_i = 1]}}$.

Proof:

Define $Y = 1/X$ i.e. $XY = 1$, if $X > 0$.

$Y = 0$ i.e. $XY = 0$, if $X = 0$.

$\therefore \mathbb{E}[XY] = 1 \cdot \Pr(XY = 1) + 0 \cdot \Pr(XY = 0)$

$= \Pr(X > 0)$.

Hence, $\Pr(X) > 0 = \mathbb{E}[XY] = \mathbb{E}\left[\sum_{i=1}^{m} X_i Y\right]$

$= \sum_{i=1}^{m} \mathbb{E}[X_i Y]$ [lin of exp.]

$= \sum_{i=1}^{m} \left(\mathbb{E}[X_i Y \mid X_i = 1] \Pr(X_i = 1) + \mathbb{E}[X_i Y \mid X_i = 0] \Pr(X_i = 0)\right)$

$= \sum_{i=1}^{m} \mathbb{E}[1 \cdot Y \mid X_i = 1] \Pr(X_i = 1)$

$= \sum_{i=1}^{m} \mathbb{E}\left[\frac{1}{X} \mid X_i = 1\right] \Pr(X_i = 1)$

$$\geq \sum_{i=1}^{m} \frac{\Pr(X_i = 1)}{\mathbb{E}[X \mid X_i = 1]}$$

[From Jensen's inequality:
$\mathbb{E}[\phi(x)] \geq \phi(\mathbb{E}[X])$,
for convex fn ϕ.
Here, we take $\phi(x) = \frac{1}{x}$.]

Not using variance

• **Theorem**: If $p = \omega(n^{-2/3})$, w.h.p. a graph chosen from $G(n,p)$ will have $\omega(G) \geq 4$.

As before $m = \binom{n}{4}$; and $X = \sum_{i=1}^{m} X_i$.

$\Pr(X_j = 1) = p^6$.

$\therefore \mathbb{E}[X \mid X_j = 1] = \sum_{i=1}^{m} \mathbb{E}[X_i \mid X_j = 1]$ [linearity of exp.]

As X_i is 0/1 RV, $\mathbb{E}[X_i \mid X_j = 1] = \Pr[X_i = 1 \mid X_j = 1]$

Case 1. $|C_i \cap C_j| = 0$. [edge-disjoint]

$\exists \binom{n-4}{4}$ sets C_i; Each $X_i = 1$ w.p. p^6.

Case 2. $|C_i \cap C_j| = 1$. [edge-disjoint]

$\exists \binom{4}{1}\binom{n-4}{3}$ sets C_i; Each $X_i = 1$ w.p. p^6.

Case 3. $|C_i \cap C_j| = 2$. [one edge common]

$\exists \binom{4}{2}\binom{n-4}{2}$ sets C_i; Each $X_i = 1$ w.p. p^5.

Case 4. $|C_i \cap C_j| = 3$. [three edges common]

$\exists \binom{4}{3}\binom{n-4}{1}$ sets C_i; Each $X_i = 1$ w.p. p^3.

Hence, $\mathbb{E}[X \mid X_j = 1] = \sum_{i=1}^{m} \mathbb{E}[X_i \mid X_j = 1]$

$$= 1 + \binom{n-4}{4}p^6 + 4\binom{n-4}{3}p^6 + 6\binom{n-4}{2}p^5 + 4\binom{n-4}{1}p^3.$$

$\uparrow$
$|G_i \cap G_j| = 4$
i.e. $i = j$.

Using Thm A, $\Pr(X > 0) \geq$

$$\left(\frac{\binom{n}{4}p^6}{1 + \binom{n-4}{4}p^6 + 4\binom{n-4}{3}p^6 + 6\binom{n-4}{2}p^5 + 4\binom{n-4}{1}p^3} \right)$$

$\rightarrow 1$ as $n \rightarrow \infty$, due to $p = \omega(n^{-2/3})$.

$\blacksquare$

[Intuitively,

① $\binom{n-4}{4}p^6 \approx \binom{n}{4}p^6$, for $n \rightarrow \infty$.

② $\binom{n-4}{4} \approx \Theta(n^4) \gg \Theta(n^3) \approx \binom{n-4}{3}$.

$\Rightarrow \binom{n-4}{4}p^6 \gg 4\binom{n-4}{3}p^6$.

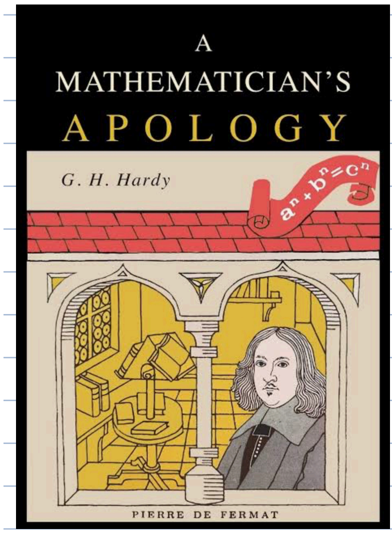
③ $\binom{n-4}{4}p = \omega(n^{4-2/3}) \gg 6\binom{n-4}{2}$.

$\Rightarrow \binom{n-4}{4}p^6 \gg 6\binom{n-4}{2}p^5$.

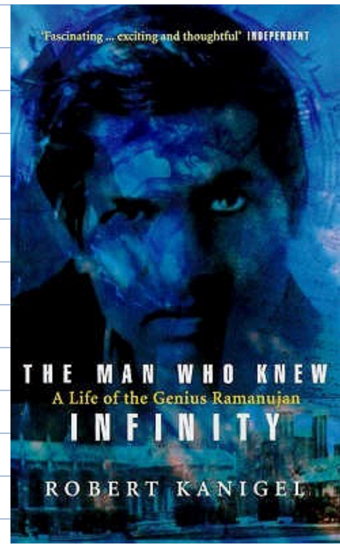
④ $\binom{n-4}{4}p^3 = \omega(n^{4-2}) = \omega(n^2) \gg 4\binom{n-4}{1}$

$\Rightarrow \binom{n-4}{4}p^6 \gg 4\binom{n-4}{1}p^3$.]

§ Application in Number Theory



G.H. Hardy (1877-1947) and
Srinivasa Ramanujan (1887-1920)



Let $\nu(n)$ = number of prime divisors of n.

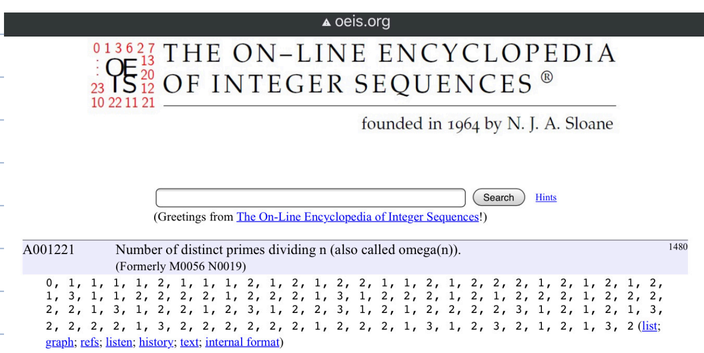
• **Theorem [Hardy & Ramanujan 1920]**
For all $\varepsilon > 0$, there exists a constant c such that all but ε fraction of numbers $x \in \{1, 2, .., n\}$ satisfy

$$|\nu(x) - \ln\ln n| \le c\sqrt{\ln\ln n}.$$

[Thm 4.2.1 in A.S, Proof below is by Turán]

Intuitively, "almost all" n have $\log\log n\,(1+o(1))$ prime factors.

prime omega function



▲ oeis.org

0 1 3 6 2 7 13 OEIS 20 12 23 10 22 11 21 THE ON-LINE ENCYCLOPEDIA OF INTEGER SEQUENCES ®

founded in 1964 by N. J. A. Sloane

Search Hints

(Greetings from The On-Line Encyclopedia of Integer Sequences!)

A001221 Number of distinct primes dividing n (also called omega(n)). 1480
(Formerly M0056 N0019)

0, 1, 1, 1, 1, 2, 1, 1, 1, 2, 1, 2, 1, 2, 2, 1, 1, 2, 1, 2, 2, 2, 1, 2, 1, 2, 1, 2, 1, 3, 1, 1, 2, 2, 2, 2, 1, 2, 2, 2, 1, 3, 1, 2, 2, 2, 1, 2, 1, 2, 2, 2, 1, 2, 2, 2, 2, 2, 1, 3, 1, 2, 2, 1, 2, 3, 1, 2, 2, 3, 1, 2, 1, 2, 2, 2, 2, 3, 1, 2, 1, 2, 1, 3, 2, 2, 2, 2, 1, 3, 2, 2, 2, 2, 2, 2, 1, 2, 2, 2, 1, 3, 1, 2, 3, 2, 1, 2, 1, 3, 2 (list; graph; refs; listen; history; text; internal format)

Later Erdős-Kac showed $\nu(x)$ behave like normal distr. with mean & variance $\ln\ln n$.

• Insight: Statistically primes have many properties that make them seem random, even if the primes themselves are not.

Proof:

• We will also use the following basic results from analytic number theory:

Merten's theorem: Adding over all primes upto N, $\sum_{p \le N} \frac{1}{p} = \ln \ln N + O(1).$

• Let x be chosen uniformly at random from $[n]$ and p be a prime.

Define $X_p = \begin{cases} 1 & \text{if } p \mid x \\ 0 & \text{otherwise} \end{cases}$

Then the number of prime divisors of x that are $\le M$ is:

$$X = \sum_{p \le M} X_p$$

Pick $M = n^{1/10}$ [works for any large constant instead of 10]

As no $x \le n$ can have ≥ 10 prime factors larger than M, we have

$$v(x) - 10 \le X(x) \le v(x).$$

so deviation bounds for $X \sim$ d.b. for v.

$\therefore \mathbb{E}[X_p] = \dfrac{\lfloor n/p \rfloor}{n} = \dfrac{1}{p} + O(\frac{1}{n})$. $\left[\because y-1 \leq \lfloor y \rfloor \leq y\right]$

By linearity of expectations,

$$\mathbb{E}[X] = \mathbb{E}\Big[\sum_{p \leq M} X_p\Big] = \sum_{p \leq M}\Big(\frac{1}{p} + O(\frac{1}{n})\Big)$$

$$\leq \ln\ln n + O(1). \quad \left[\text{Merten's theorem}\right]$$

Now, we want to compute $Var[X]$.

Note, $Var[X_p] \leq (\frac{1}{p} + \frac{c}{n})(1 - \frac{1}{p} - \frac{c}{n})$ for some constant c

$$\leq \frac{1}{p} + \frac{c}{n} = \frac{1}{p} + O(\frac{1}{n})$$

$\therefore \sum_{p \leq M} Var[X_p] \leq \ln\ln n + O(1)$. $\left[\text{same as } \mathbb{E}X \text{ computation}\right]$

Now, we focus on covariance.

For distinct primes p, q; $X_p X_q = 1$ iff $(p \mid x$ and $q \mid x) \iff (pq \mid x)$.

Hence, $Cov[X_p, X_q] = \mathbb{E}[X_p X_q] - \mathbb{E}[X_p]\mathbb{E}[X_q]$

$$= \frac{\lfloor n/pq \rfloor}{n} - \frac{\lfloor n/p \rfloor}{n}\frac{\lfloor n/q \rfloor}{n}$$

$$\leq \Big(\frac{1}{pq}\Big) - \Big(\frac{1}{p} - \frac{1}{n}\Big)\Big(\frac{1}{q} - \frac{1}{n}\Big)$$

$$\leq \frac{1}{n}\Big(\frac{1}{p} + \frac{1}{q}\Big).$$

$$\therefore \sum_{p\neq q} \mathrm{Cov}[X_p, X_q] \leq \frac{1}{n} \sum_{p \neq q} \left(\frac{1}{p} + \frac{1}{q}\right)$$

$$\leq \frac{O(M)}{n} \sum \frac{1}{p}$$

$$\leq O(n^{-9/10} \cdot \underbrace{\ln\ln n}_{o(1)})$$

Hence, $\mathrm{Var}[X] = \sum_{p \leq M} \mathrm{var}[X_p] + \sum_{p\neq q} \mathrm{Cov}[X_p, X_q]$.

$$\leq \ln\ln n + o(1) \approx \mathbb{E}X.$$

Thus, Chebyshev's inequality imply:

$$\Pr[|X - \ln\ln n| > \lambda\sqrt{\ln\ln n}] \leq \frac{(\mathrm{Var}\,X)^2}{\lambda^2 (\ln\ln n)}$$

$< \lambda^{-2} + o(1)$ for any constant $\lambda > 0$.

As $|X - \nu| \leq 10$, this finally imply, w.h.p.

$$|\nu(x) - \ln\ln n| \leq c\sqrt{\ln\ln n}. \quad \blacksquare$$

§ An application to analysis.

- Recommended (Optional) read.

- **Theorem:** (Thm 4.32 in MIT)

Weierstrass Approximation Theorem

Let $f : [0,1] \to \mathbb{R}$ be a continuous function on a bounded interval.
Given $\varepsilon > 0$, it is possible to approximate f by a polynomial $p(x)$ such that

$$|p(x) - f(x)| \leq \varepsilon, \quad \forall x \in [0,1].$$

– Very important result in numerical analysis (polynomial interpolation).
used in ML & convex optimization

Karl Weierstraß
(1815 - 1897)

← Father of 'analysis.'
Had no formal college degree. Received honorary doctorate due to his contributions.