PROBABILISTIC COMBINATORICS

Michael Drmota

Institut für Diskrete Mathematik und Geometrie
TU Wien
michael.drmota@tuwien.ac.at
www.dmg.tuwien.ac.at/drmota/

Contents

Introduction

G.C. Rota:

,,Probability is just combinatorics divided by n."

The Probabilistic Method

The **Probabilistic Method** has been initiated by Paul Erdős (1947) in order to prove the existence of certain combinatorial objects. The principle idea is to define a proper probability distribution on a class of (discrete) objects and to show that the probability of a certain property is positive. Of course this also proves that there exists such an object with this property. We will apply this approach to various problems on **random graphs**.

The Probabilistic Method

Ramsey Numbers

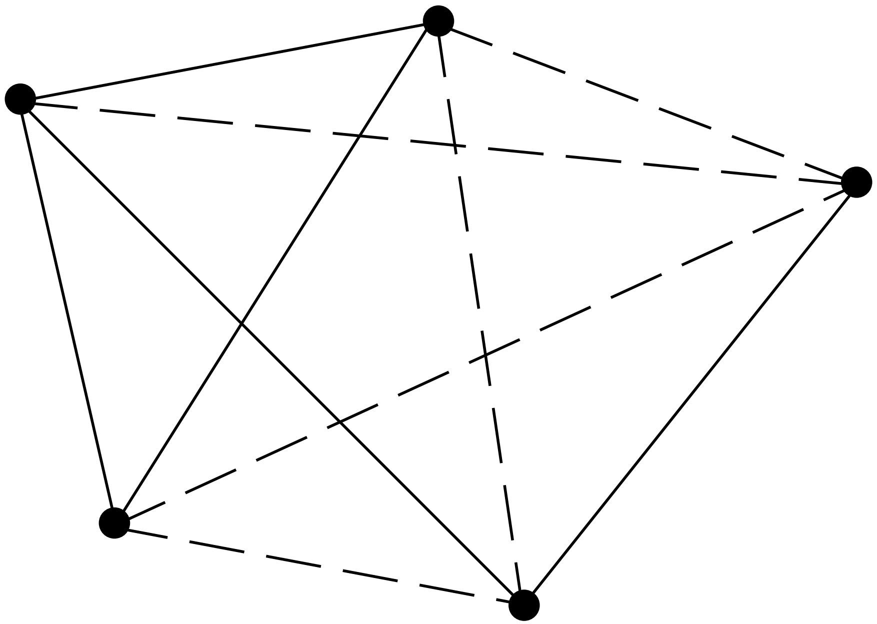
Definition. *The* **Ramsey number** $R(k,l)$ *is the smallest number* n *such that any 2-coloring of the edges on the conplete graph* K_n *on* n *vertices contains either a monochromatic* K_k *(in* K_n*) of the first color or a monochromatic* K_l *(in* K_n*) of the second color.*

Ramsey's theorem: $R(k,l)$ exists for all positive integers k and l.

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

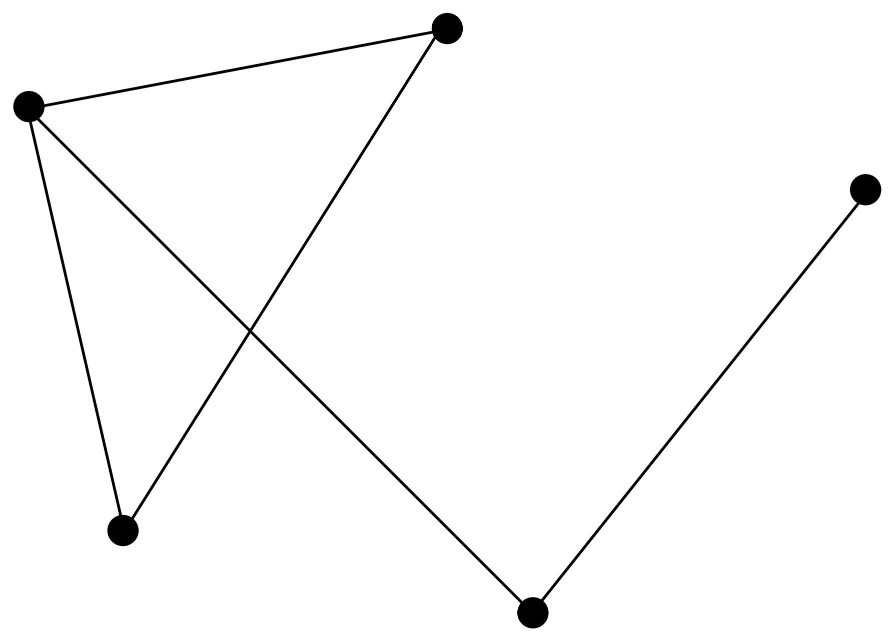
The Probabilistic Method

Ramsey Numbers



The Probabilistic Method

Ramsey Numbers



The Probabilistic Method

Upper bounds for the Ramsey Number

Lemma

$$\boxed{R(k,l) \leq R(k-1,l) + R(k,l-1).}$$

$$R(k,1) = R(1,k) = 1.$$

Corollary.

$$\boxed{R(k,l) \leq \binom{k+l-2}{k-1} \leq 2^{k+l-2}}$$

$$\Longrightarrow \quad \boxed{R(k,k) \leq 4^{k-1}}$$

The Probabilistic Method

Proof (color 1 = blue, color 2 = red)

Consider a complete graph K_n with $n = R(k-1,l) + R(k,l-1)$ nodes.

Pick a vertex v and let M be the blue neighbors and N the red neighbors.

$$\implies \quad |M| \geq R(k-1,l)| \quad \text{or} |N| \geq R(k,l-1).$$

$|M| \geq R(k-1,l) \implies K_n$ contains a blue K_k or a red K_l.

$|N| \geq R(k,l-1) \implies K_n$ contains a blue K_k or a red K_l.

The Probabilistic Method

Lower Bound for the Ramsey Number

Theorem

$$\boxed{R(k,k) > 2^{k/2}}$$

for all $k \geq 3$.

Proof

K_n ... complete graph with vertex set $\{1, 2, \ldots\}$

Take a random 2-coloring of the $\binom{n}{2}$ edges
(Each edge is colored independently and with equal probability $\frac{1}{2}$.)

The Probabilistic Method

Lower Bound for the Ramsey Number

$R \subseteq \{1, 2, \ldots\}$, $|R| = k$

$$A_R := \{\text{the induced subgraph of } R \text{ is monochromatic}\}$$

$$\Longrightarrow \quad \mathbb{P}(A_R) = 2\,\frac{1}{2^{\binom{k}{2}}} = 2^{1-\binom{k}{2}}$$

$$\Longrightarrow \quad \mathbb{P}\{\exists\, R \subseteq \{1, 2, \ldots\} : |R| = k,\ A_R \text{ occurs}\} \leq \binom{n}{k} 2^{1-\binom{k}{2}}.$$

The Probabilistic Method

Lower Bound for the Ramsey Number

$n = \lfloor 2^{k/2} \rfloor$ (and $k \geq 3$)

$$\implies \quad \binom{n}{k} 2^{1-\binom{k}{2}} < 2\frac{n^k}{k!}\frac{1}{2^{k^2/2-k/2}} \leq 2\frac{2^{k/2}}{k!} < 1$$

$$\implies \quad \mathbb{P}\{\forall\, R \subseteq \{1,2,\ldots\} : |R| = k,\ R \text{ is not monochromatic}\} > 0$$

$$\implies \quad \boxed{R(k,k) > n}.$$

First Moment Method

Theorem

X ... discrete random variable on **non-negative integers**.

$$\Longrightarrow \quad \boxed{\mathbb{P}\{X > 0\} \le \mathbb{E}\, X}.$$

Proof

$$\mathbb{E}\, X = \sum_{k \ge 0} k\, \mathbb{P}\{X = k\} \ge \sum_{k \ge 1} \mathbb{P}\{X = k\} = \mathbb{P}\{X > 0\}.$$

First Moment Method

As an **first application** we prove $R(k,k) > 2^{k/2}$ a second time:

K_n ... complete graph with vertex set $\{1, 2, \ldots\}$

Take a random 2-coloring of the $\binom{n}{2}$ edges

$\mathcal{S}_{n,k}$ ... set of all subgraphs of K_n with k nodes

$$\Longrightarrow \quad \boxed{X_n := \sum_{R \in \mathcal{S}_{n,k}} \mathbb{I}_{[R \text{ is monochromatic}]}}$$

is the **(random) number of monochromatic subgraphs of K_n that are isomorphic to K_k.**

First Moment Method

$$X_n = \sum_{R \in \mathcal{S}_{n,k}} \mathbb{I}_{[R \text{ is monochromatic}]}$$

$$\implies \quad \mathbb{E}\, X_n = \sum_{R \in \mathcal{S}_{n,k}} \mathbb{P}\{R \text{ is monochromatic}\} = \binom{n}{k} 2 \; 2^{-\binom{k}{2}}$$

$$\implies \quad \mathbb{P}\{X_n > 0\} \le \binom{n}{k} 2^{1-\binom{k}{2}} < 2 \frac{2^{k/2}}{k!} < 1$$

$$\implies \quad \boxed{\mathbb{P}\{X_n = 0\} > 0}.$$

Random Graphs

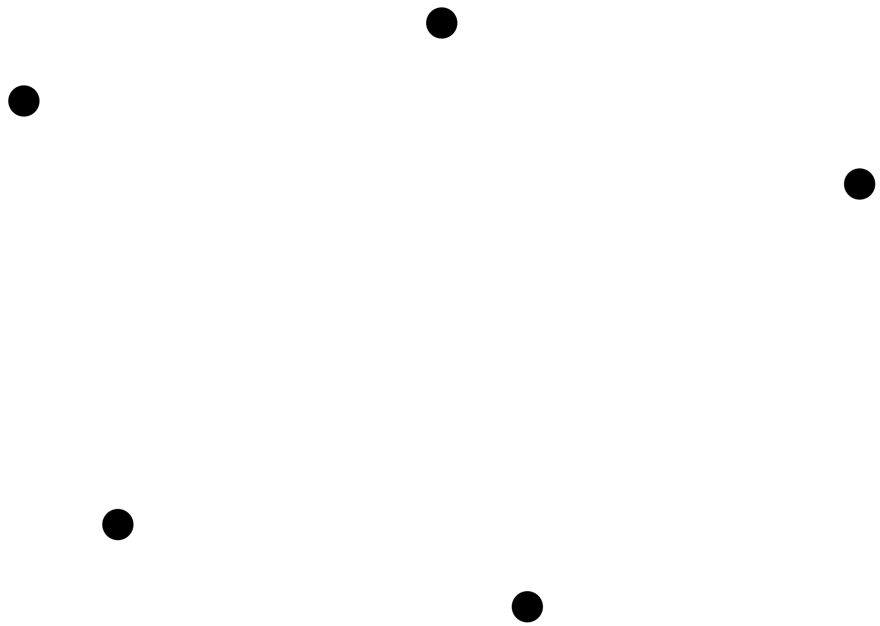
Definition *Let n be a positive integer and p a real number with $0 \leq p \leq 1$. The* **random graph** *$G(n,p)$ is a probability space over the set of graphs on the vertex set $\{1, 2, \ldots, n\}$ determined by*

$$\mathbb{P}\{(i,j) \in G\} = p$$

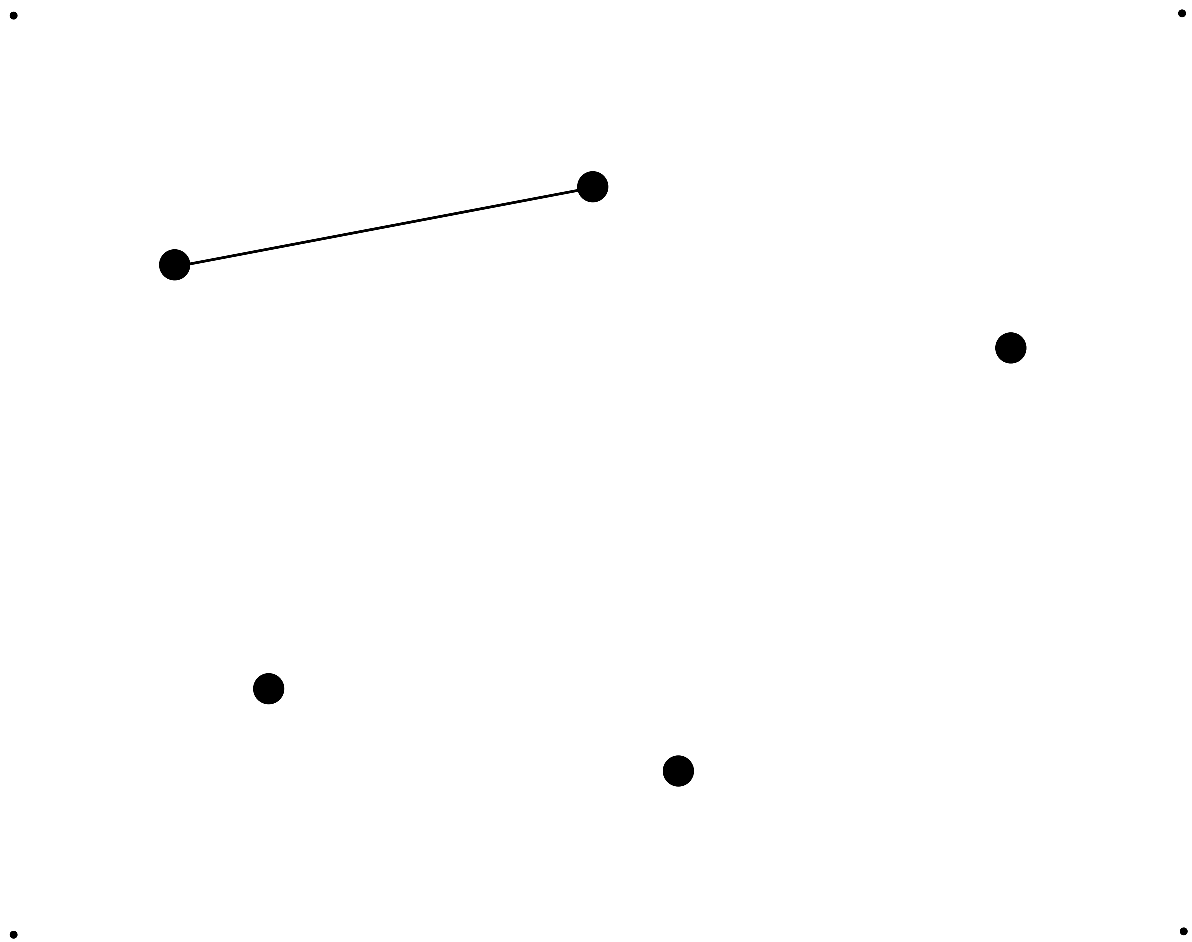
for all possible (undirected) edges (i,j) with $1 \leq i,j \leq n$ and $i \neq j$ with these events mutually independent.

Similarly one also considers random graphs $G(n,m)$, where m is also a given integer with $0 \leq m \leq \binom{n}{2}$. Here one considers the set of all graphs on the set of vertices $\{1, 2, \ldots, n\}$ with exactly m (undirected) edges where each of these graphs is equally likely. Due to the law of large numbers $G(n,m)$ will have very similar properties as $G(n,p)$ with $p = m/\binom{n}{2}$.

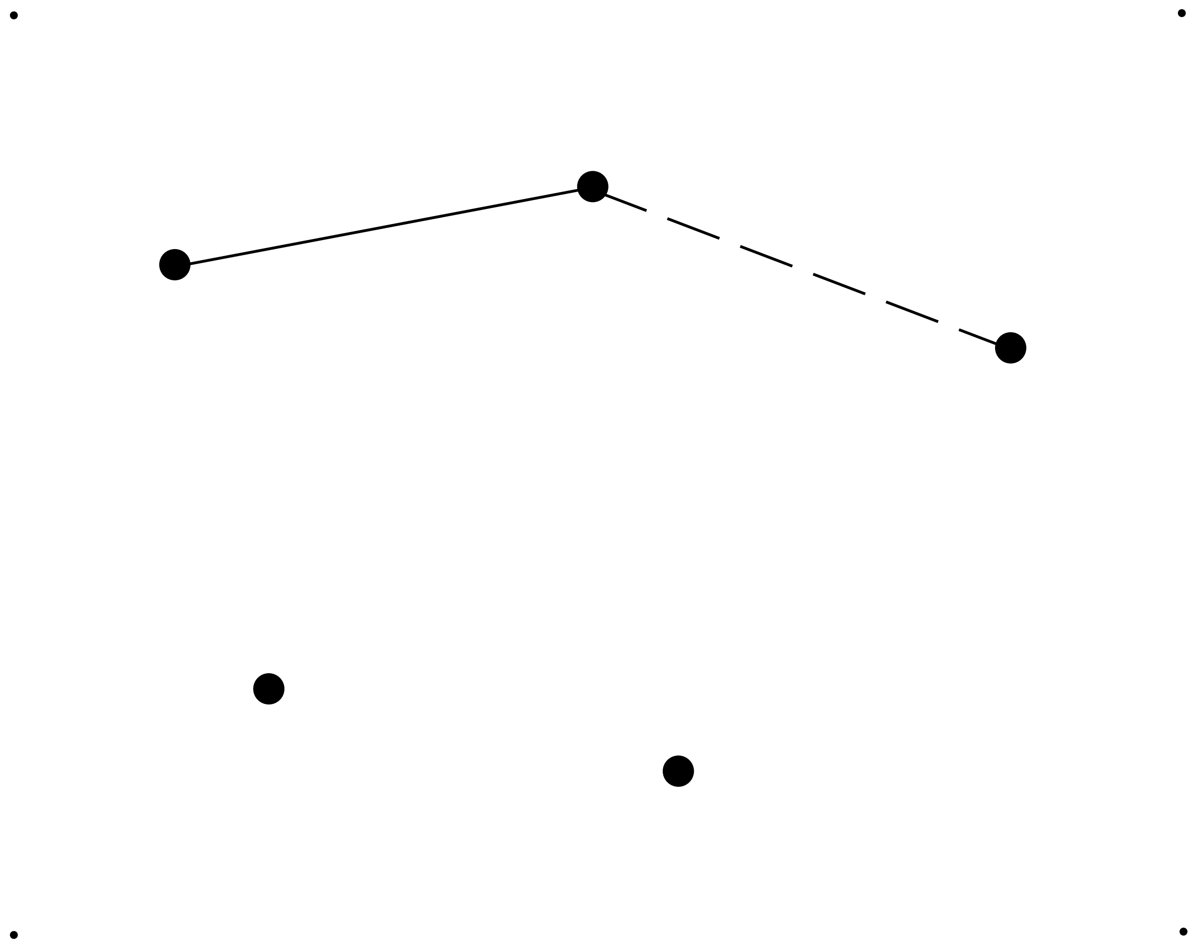
Random Graphs



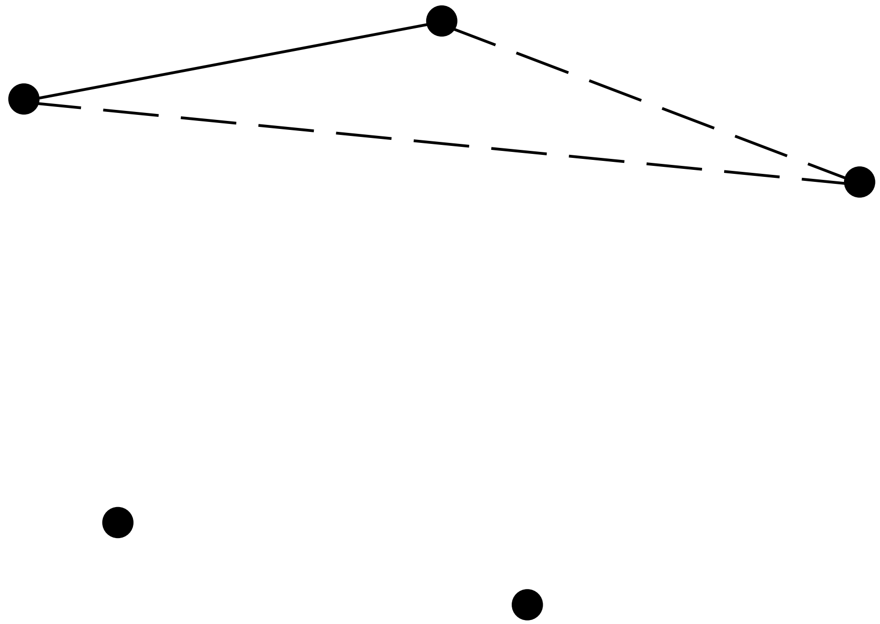
Random Graphs



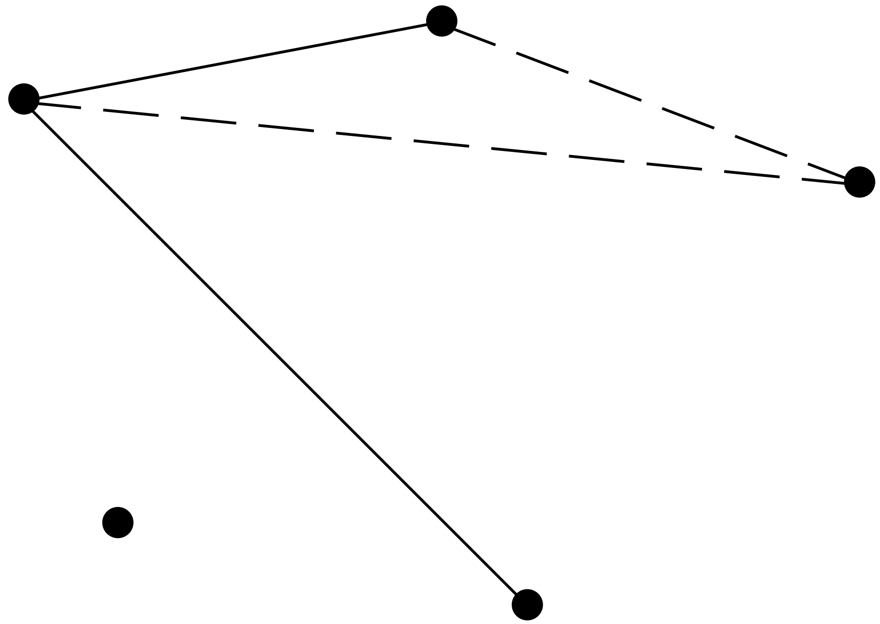
Random Graphs



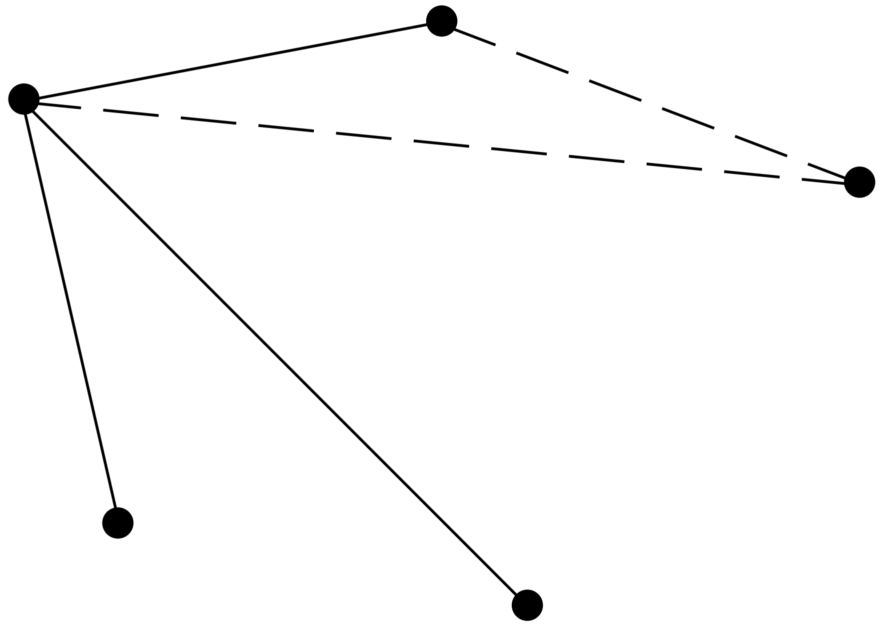
Random Graphs



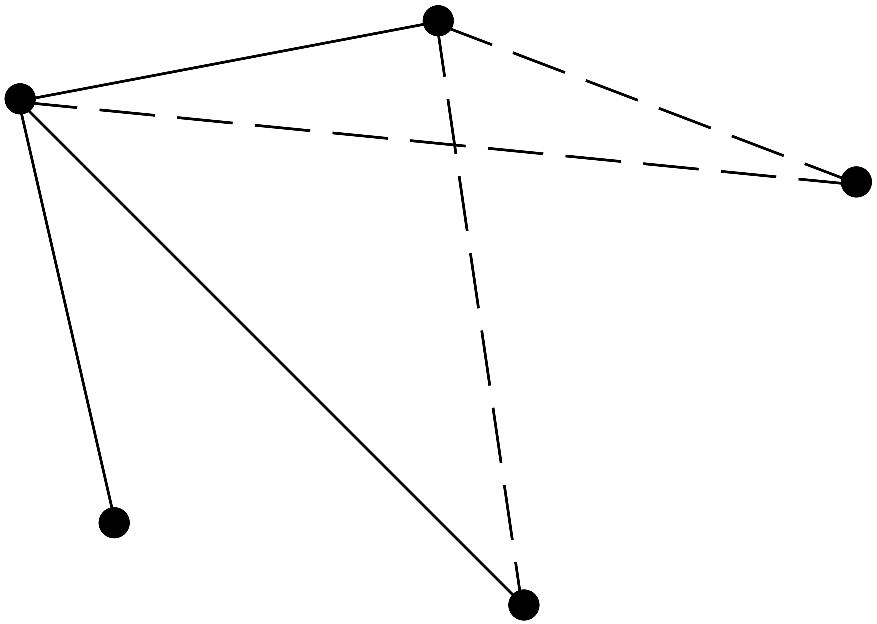
Random Graphs



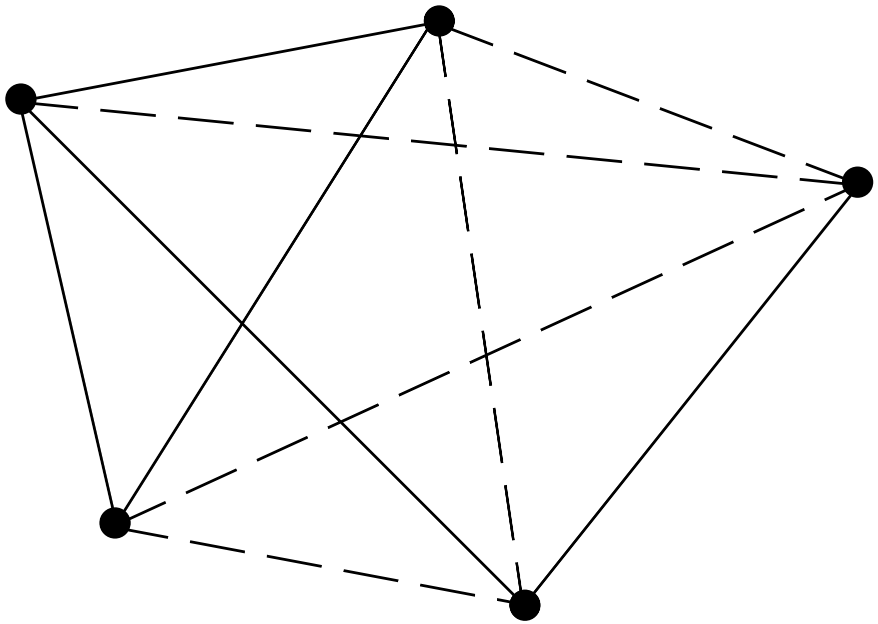
Random Graphs



Random Graphs



Random Graphs



Random Graphs

Definition *The* **girth** girth(G) *of a graph* G *is the size of the shortest cycle.*

The **chromatic number** $\chi(G)$ *of a graph* G *is the smallest number* k *such that there exists a regular* k*-coloring of the vertices of* G*, that is, a coloring of at* k *colors of the vertices such that adjacent vertices have different colors.*

Theorem [Erdős 1959]

For all (positive integers) k and ℓ there exists a graph G with

$$\boxed{\text{girth}(G) > \ell} \quad \text{and} \quad \boxed{\chi(G) > k}.$$

Random Graphs

Proof

$p = n^{\theta-1}$ for some $0 < \theta < 1/\ell$ (n be chosen sufficiently large)

$V = \{1, 2, \ldots, n\}$ ... vertex set of a random graph:

$$\mathbb{P}\{e \in E(G)\} = p \qquad \text{(independently)}$$

X ...number of cycles of size $\leq \ell$.

$\theta\ell < 1$

$$\Longrightarrow \quad \mathbb{E}\, X = \sum_{i=3}^{\ell} \frac{(n)_i}{2i} p^i \leq \sum_{i=3}^{\ell} \frac{n^i}{2i} n^{(\theta-1)i} = \sum_{i=3}^{\ell} \frac{n^{\theta i}}{2i} = o(n).$$

Random Graphs

$$\mathbb{E}\,X \geq \mathbb{E}\left(X \cdot \mathbb{I}_{[X \geq n/2]}\right) \geq \frac{n}{2}\,\mathbb{P}\{X \geq n/2\}$$

$$\mathbb{E}\,X = o(n)$$

$$\Longrightarrow \quad \boxed{\mathbb{P}\{X \geq n/2\} = o(1)}.$$

Random Graphs

$$\begin{aligned}
\mathbb{P}\{\alpha(G) \geq m\} &= \mathbb{P}\{\exists\, S \subseteq \{1, 2, \ldots, n\} : |S| = m,\ S \text{ is independent}\} \\
&\leq \mathbb{E}\left(\sum_{|S|=m} \mathbb{I}_{[S \text{ is independent}]} \right) \\
&= \sum_{|S|=m} \mathbb{P}\{S \text{ is independent}\} \\
&= \binom{n}{m} (1-p)^{\binom{m}{2}} \\
&\leq \frac{n^m}{m!} e^{-p\binom{m}{2}} \\
&\leq \left(n e^{-p(m-1)/2}\right)^m
\end{aligned}$$

Random Graphs

$$m = m(n) = \lceil \tfrac{3}{p} \log n \rceil \sim 3n^{1-\theta} \log n$$

$$\Longrightarrow \quad n e^{-p(m-1)/2} \to 0 \qquad (n \to \infty)$$

$$\Longrightarrow \quad \boxed{\mathbb{P}\{\alpha(G) \geq m(n)\} \to 0} \qquad (n \to \infty)$$

Random Graphs

n sufficiently large that $\mathbb{P}\{X \geq n/2\} < \frac{1}{2}$ and $\mathbb{P}\{\alpha(G) \geq m(n)\} < \frac{1}{2}$.

- Take G with $X < n/2$ (less than $n/2$ cycles of length at most ℓ) and $\alpha(G) < m(n) \sim 3n^{1-\theta} \log n$.

- Remove from G a vertex from each cycle of length at most ℓ.

- New graph G^* has at least $n/2$ vertices, $\boxed{\text{girth}(G^*) > \ell}$

- $\alpha(G^*) \leq \alpha(G)$

$$\implies \quad \chi(G^*) \geq \frac{|G^*|}{\alpha(G)} \geq \frac{n/2}{3n^{1-\theta} \log n} = \frac{n^\theta}{6 \log n}.$$

- n sufficiently large that $n^\theta/(6 \log n) > k \implies \boxed{\chi(G) > k}$.

Random Graphs

Theorem [Bollobas] We have, almost always in $G(n, \frac{1}{2})$,

$$\boxed{\chi(G) \sim \frac{n}{2 \log_2 n}}.$$

Proof (Lower bound)

Almost always there exists no complete subgraph $K_{\lfloor 2 \log_2 n \rfloor}$ in $G(n, \frac{1}{2})$.

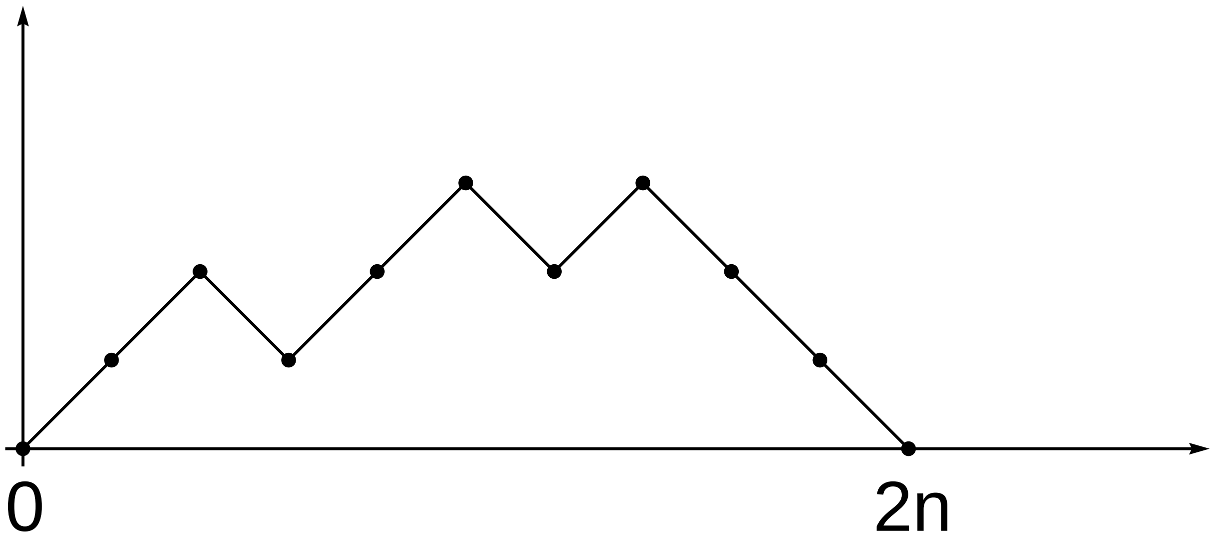
The same holds for the complement. Consequently almost always there is no independent set of size $\lfloor 2 \log_2 n \rfloor$.

$$\Longrightarrow \quad \chi(G) \geq \frac{n}{\alpha(G)} \geq \frac{n}{2 \log_2 n}.$$

($\alpha(G)$ ... independece number of G.)

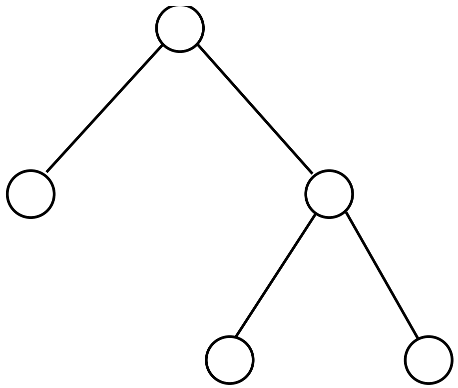
Approximation by Continuous Processes

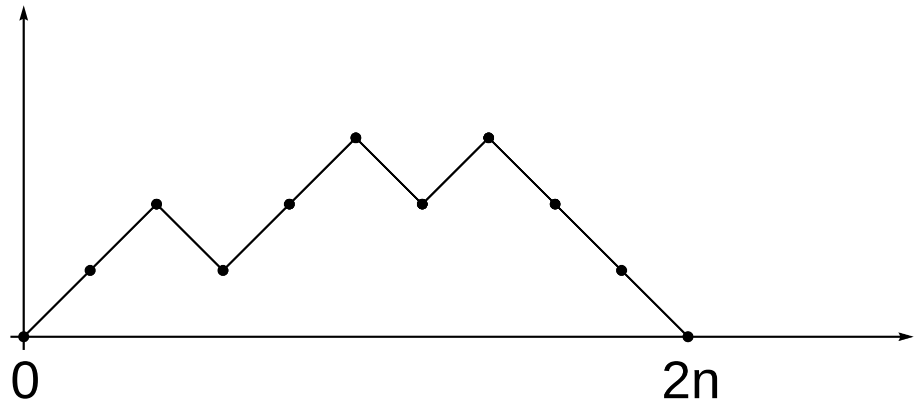
Dyck paths



Approximation by Continuous Processes

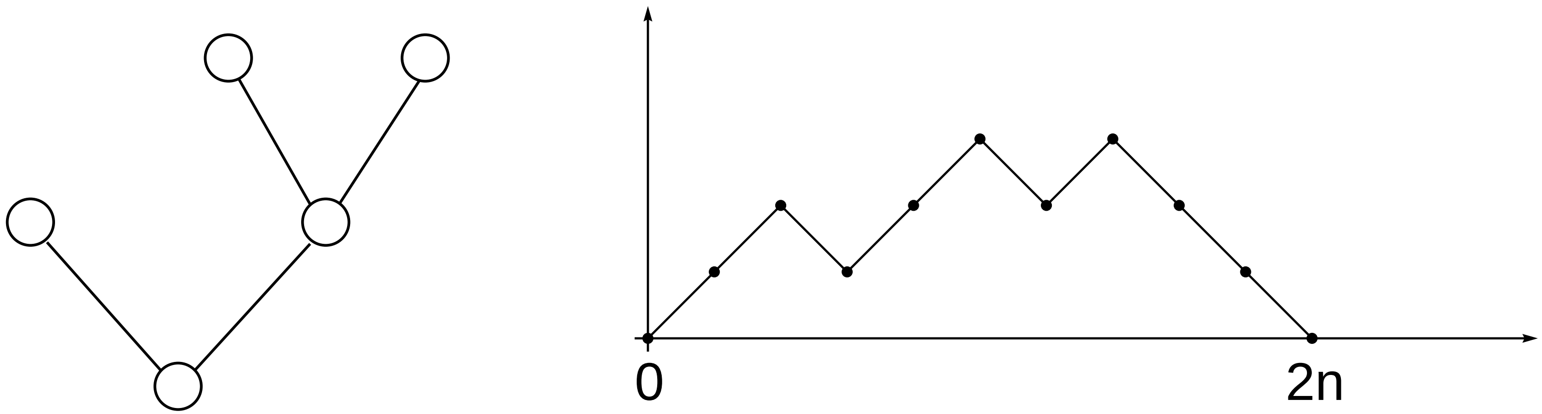
Dyck paths and Trees





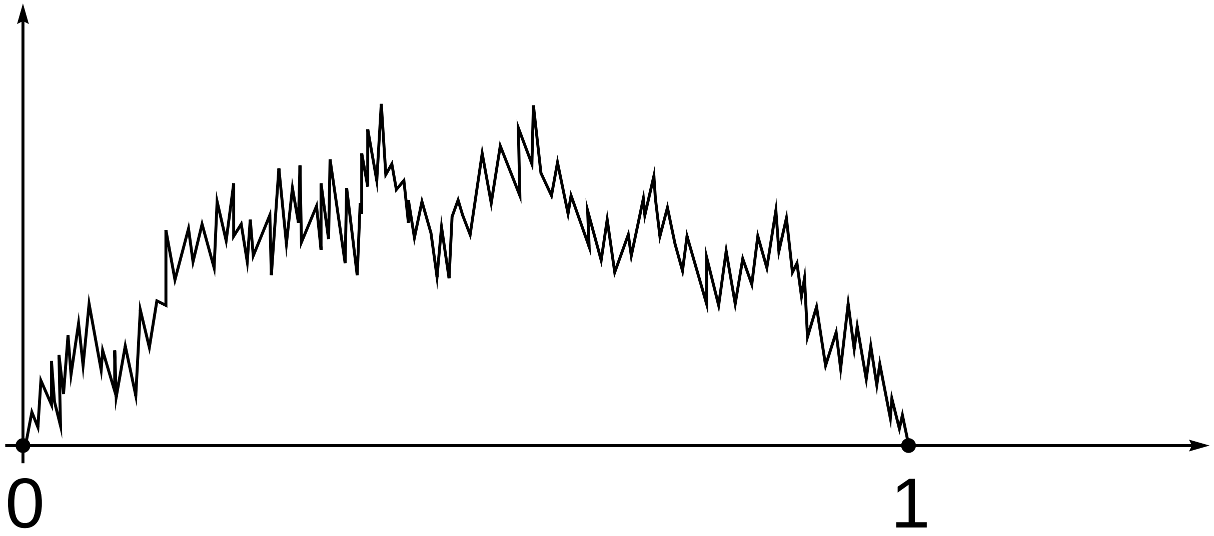
Approximation by Continuous Processes

Dyck paths and Trees



Approximation by Continuous Processes

Brownian Excursion



$(e(t), 0 \leq t \leq 1)$ ... Brownian motion between two zeros and scaled.

Approximation by Continuous Processes

Theorem. [Donsker's Theorem]

$(X_n(t), 0 \le t \le 2n)$ ... process of Dyck paths (every Dyck path is equally likely)

$(e(t), 0 \le t \le 1)$ ... Brownian excursion

$\Longrightarrow$

$$\boxed{\left(\frac{1}{\sqrt{n}} X_n(2nt),\ 0 \le t \le 1\right) \to (e(t), 0 \le t \le 1)}$$

Longest Increasing Subsequence in R. P.

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 1 & 5 & 6 & 2 & 4 \end{pmatrix}$$

Longest Increasing Subsequence in R. P.

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 1 & 5 & 6 & 2 & 4 \end{pmatrix}$$

Longest Increasing Subsequence in R. P.

S_n ... the set of permutations of the numbers $\{1, 2, \ldots, n\}$
(We assume that every permutation in S_n is equally likely.)

For $\pi \in S_n$ we say that $\pi(i_1), \pi(i_2), \ldots, \pi(i_k)$ is an increasing subsequence in π if $i_1 < i_2 < \cdots < i_k$ and $\pi(i_1) < \pi(i_2) < \cdots < \pi(i_k)$.

$L_n = L_n(\pi)$ ... length of the longest increasing subsequence.

Ulam's Problem: $\mathbb{E}\, L_n \sim$?

Ulam's conjecture: $\boxed{\mathbb{E}\, L_n \sim c\sqrt{n}}$ for some constant $c > 0$.

Longest Increasing Subsequence in R. P.

Erdős Szekeres 1935: $c \geq \frac{1}{2}$

Logan and Shepp 1977: $c \geq 2$

Vershik and Kerov 1977: $\boxed{c = 2}$.

(Alternate proofs are due to Aldous and Diaconis, Seppäläinen, and Johansson).

Frieze 1991: L_n is concentrated

Bollobás and Brightwell 1992, Talagrand 1995: $\mathbb{V}\, L_n = O(\sqrt{n})$.

Odlyzko and Rains 2000: order of $\mathbb{V}\, L_n$ should be $n^{1/3}$.

Baik, Deift, and Johansson 1999: complete solution.

Longest Increasing Subsequence in R. P.

Theorem (Baik, Deift, and Johansson 1999)

Let S_n be the group of permutations of n numbers with uniform distribution and L_n the longest increasing subsequence. Then there exists a random variable Y such that

$$\boxed{\frac{L_n - 2\sqrt{n}}{n^{1/6}} \xrightarrow{\mathrm{d}} Y.}$$

Furthermore, we have convergence of all moments.

Remark.

The limiting distribution Y is exactly the same at the limiting distribution of the largest eigenvalue in random Hermitian matrices. However, it seems that there is no direct connection between these two problems.

Longest Increasing Subsequence in R. P.

Tracy-Widom distribution: $F(t) = \mathbb{P}\{Y \leq t\}$

$u(x)$ ... solution of the Painlevé II equation

$$u'' = 2u^3 + xu, \quad u(x) \sim Ai(x) \quad (\text{as } x \to \infty),$$

where $Ai(x)$ denotes the Airy function.

Then

$$\boxed{F(t) = \exp\left(\int_t^{\infty} (x-t)^2 u(x)^2 \, dx\right)}$$

Longest Increasing Subsequence in R. P.

Proof Method

Basically one determines the asymptotic behaviour of the Poisson transform

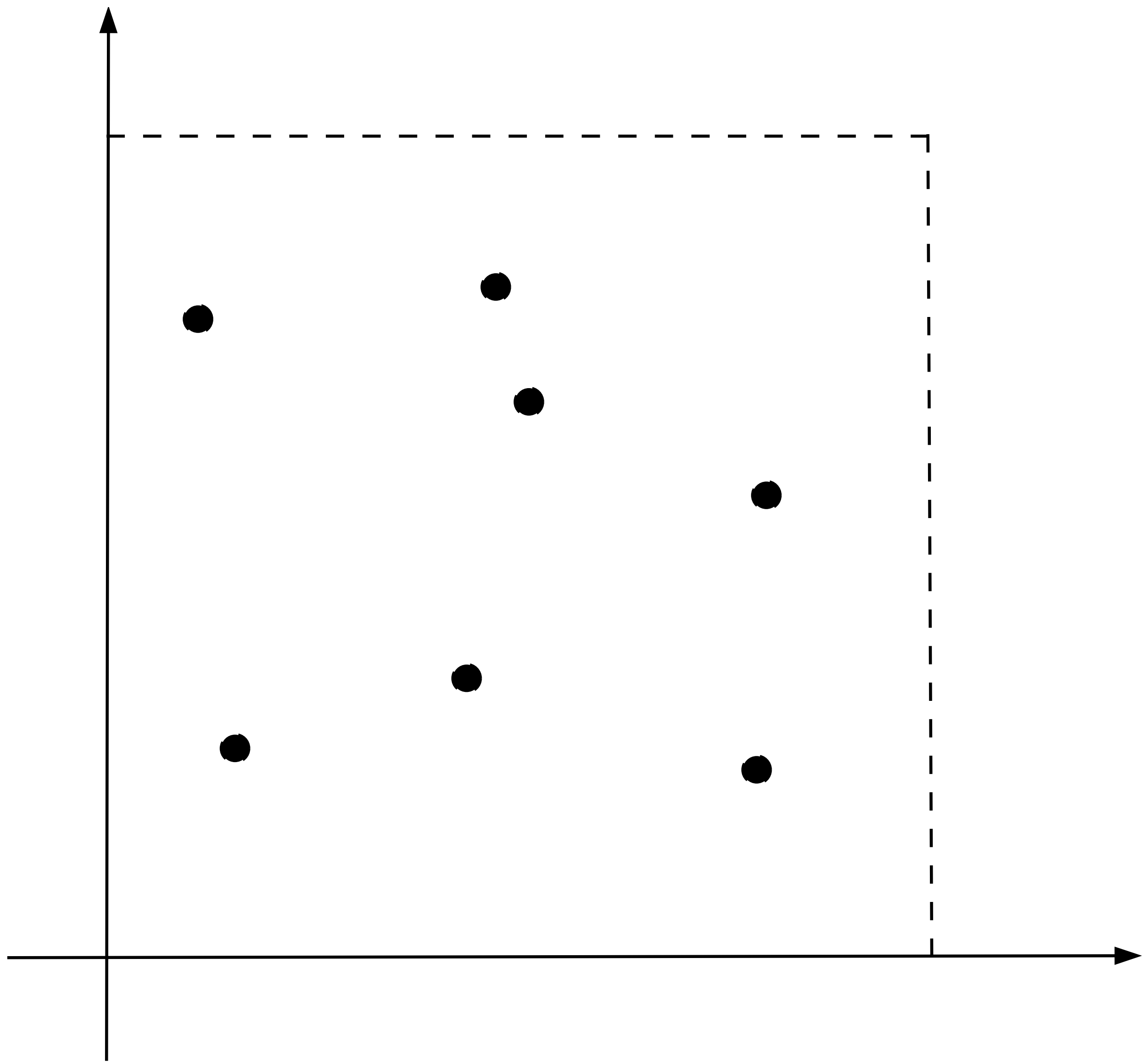
$$\phi_k(\lambda) = \sum_{n=0}^{\infty} \frac{e^{-\lambda}\lambda^n}{n!} \mathbb{P}\{L_n \leq k\}$$

that can be represented as

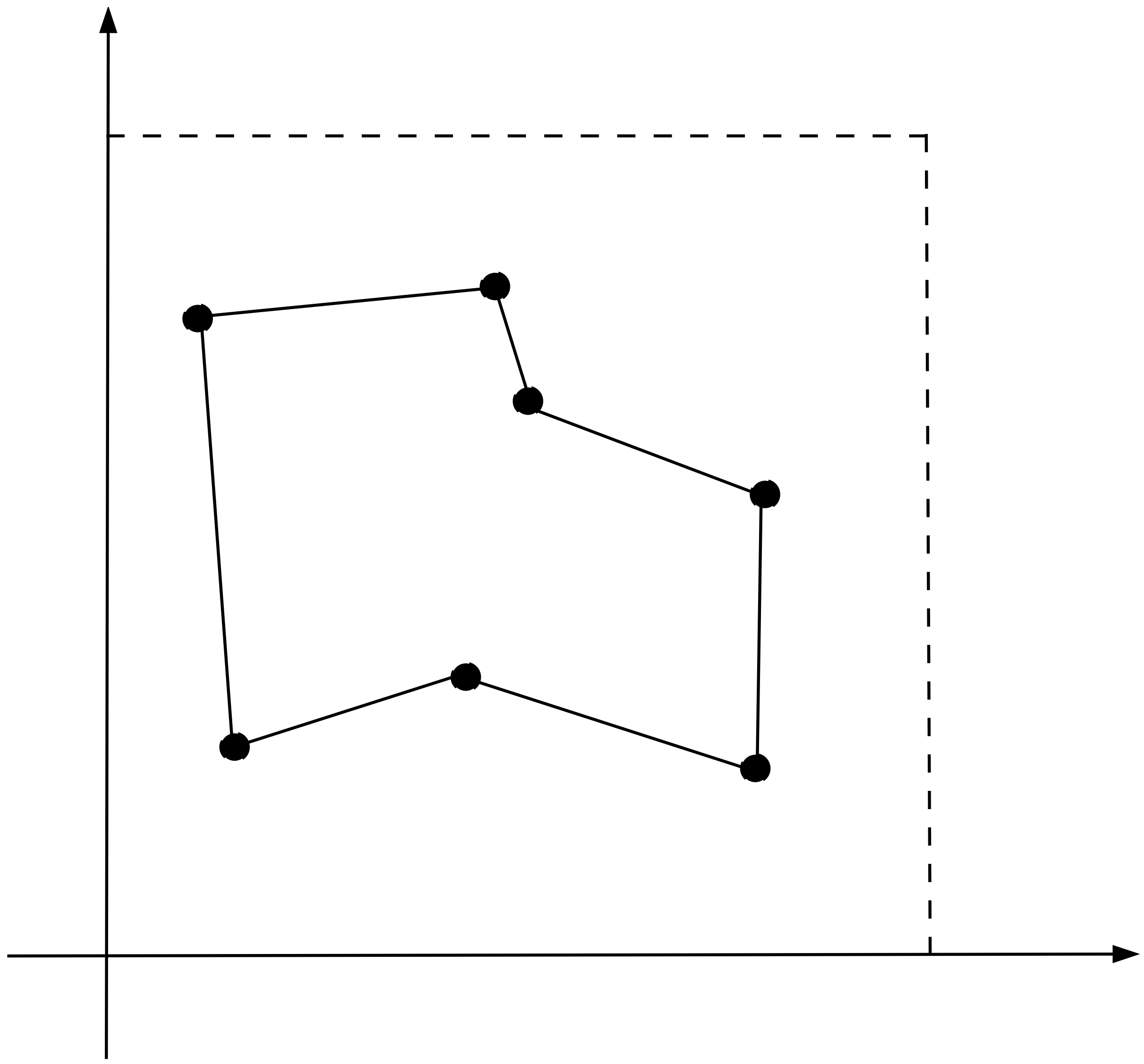
$$\phi_k(\lambda) = \frac{e^{-\lambda}}{(2\pi)^k k!} \int_{[-\pi,\pi]^k} \exp\left(2\sqrt{\lambda}\sum_{j=1}^{k} \theta_j\right) \prod_{1\leq j<\ell\leq k} \left|e^{i\theta_j} - e^{i\theta_\ell}\right| \, d\theta_1 \cdots d\theta_k.$$

One has to use the theory of orthogonal polynomials on the unit circle, sophisticated Riemann-Hilbert problem techniques and certain properties on eigenvalues of random matrices.

Travelling Salesman Problem



Travelling Salesman Problem



Travelling Salesman Problem

$\mathbf{X} = (X_1, X_2, \ldots, X_n)$ ... n-tuple of random point selected uniformly and independently in the unit square $[0,1]^2$

Length of the minimum (travelling salesman) tour:

$$\boxed{\mathrm{TSP}(\mathbf{X}) = \min_{\pi \in S_n} \sum_{j=1}^{n} \left| X_{\pi(j)} - X_{\pi(j+1)} \right|}$$

Theorem (Beardwood, Halton and Hammersley 1959)

$$\boxed{\frac{\mathrm{TSP}(\mathbf{X})}{\sqrt{n}} \to \beta_2} \quad \text{in prob.}$$

for some $\beta_2 > 0$.

Remark: Up to now there is no known analytic expression for β_2.

Travelling Salesman Problem

Notation: M(Y) ... median of r.v. Y

Theorem (Rhee and Talagrand)

$$\boxed{\mathbb{P}\{|\text{TSP}(\mathbf{X}) - \text{M}(\text{TSP}(\mathbf{X}))| \geq t\} < 4e^{-t^2/c}.}$$

for some constant $c > 0$.

Corollary. All central moments of TSP($\mathbf{X}$) are bounded.

(However, the exact location of the mean is unknown.)

Travelling Salesman Problem

Talagrand's Inequality

$\Omega_1, \Omega_2, \ldots, \Omega_n$... probability spaces, $\Omega = \Omega_1 \times \cdots \times \Omega_n$

$\mathbf{X} = (X_1, X_2, \ldots, X_n)$ ... independent random variables, X_k taking values in Ω_k.

Weighted Hamming distance related to $\alpha = (\alpha_1, \ldots, \alpha_n)$ with $\alpha_k \geq 0$:

$$d_\alpha(\mathbf{x}, \mathbf{y}) = \sum_{x_i \neq y_i} \alpha_i$$

Talagrand's convex distance

$$\boxed{d_T(\mathbf{x}, A) = \sup_{\alpha \geq 0, \|\alpha\|=1} \inf_{\mathbf{y} \in A} d_\alpha(\mathbf{x}, \mathbf{y})}$$

between $\mathbf{x} \in \Omega$ and $A \subset \Omega$.

Travelling Salesman Problem

Talagrand's Inequality

$$\boxed{\mathbb{P}\{\mathbf{X} \in A\} \cdot \mathbb{P}\{d_T(\mathbf{X}, A) \geq t\} < e^{-t^2/4}.}$$

Travelling Salesman Problem

Theorem

f … real valued function on $\Omega = \Omega_1 \times \cdots \times \Omega_n$

For every $\mathbf{x} \in \Omega$ there exists a non-negative unit n-vector α and a constant $c > 0$ such that for all $\mathbf{y} \in \Omega$

$$\boxed{f(\mathbf{x}) \leq f(\mathbf{y}) + c\, d_\alpha(\mathbf{x}, \mathbf{y}).}$$

Then, for every random n-tuple $\mathbf{X} = (X_1, \ldots, X_n)$ of independent random variables X_k taking values in Ω_k we have

$$\boxed{\mathbb{P}\{|f(\mathbf{X}) - \mathsf{M}(f(\mathbf{X}))| \geq t\} \leq 4\, e^{-t^2/(4c^2)}.}$$

Travelling Salesman Problem

Lemma

For every $\mathbf{x} \in ([0,1]^2)^n$ there exists non-negative unit vector α and a constant $c > 0$ such that for all $\mathbf{y} \in ([0,1]^2)^n$

$$\boxed{\mathsf{TSP}(\mathbf{x}) \leq \mathsf{TSP}(\mathbf{y}) + c\, d_\alpha(\mathbf{x},\mathbf{y}).}$$

(Elementary proof that uses an approximate minumum tour to construct α.)

Remark

This method can be applied to several other problem, for example to the minimal Steiner tree problem etc.

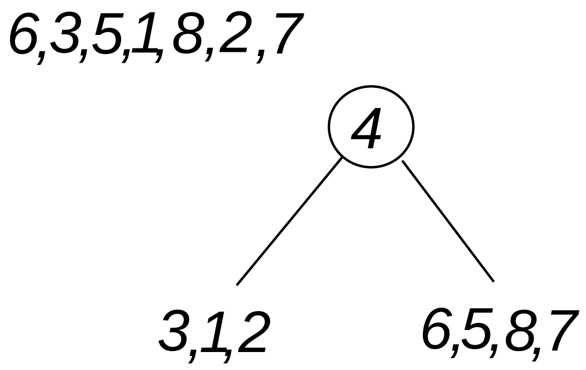
Quicksort

Sorting Data

4,6,3,5,1,8,2,7

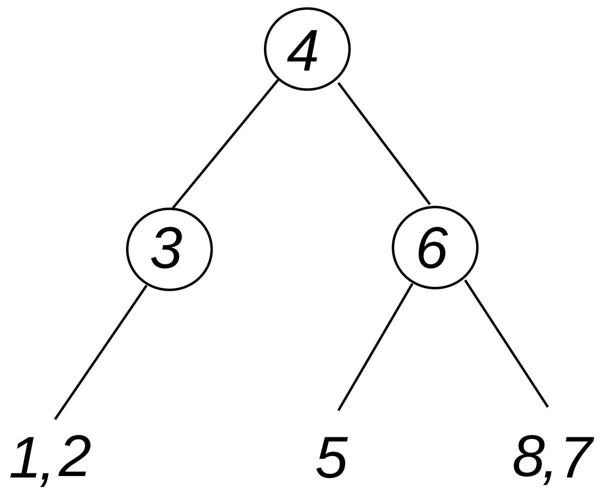
Quicksort

Sorting Data



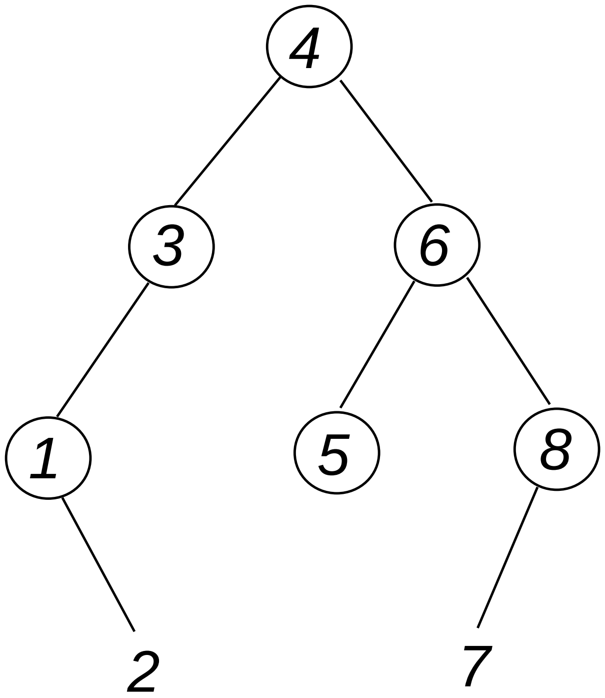
Quicksort

Sorting Data



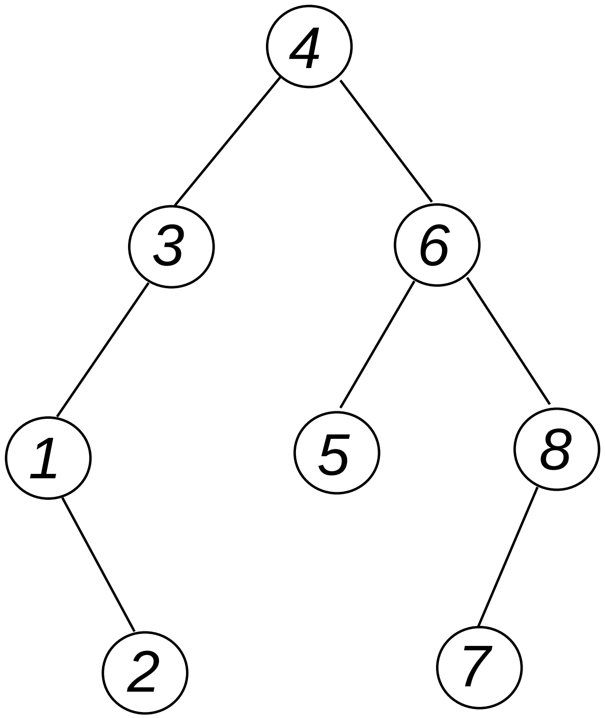
Quicksort

Sorting Data



Quicksort

Sorting Data



Quicksort

Probabilistic Model

Every permutation of the input data $\{1, 2, \ldots, n\}$ is equally likely.

Number of Comparisions

L_n ... number of comparisions that are needed to sort a random permutation of $\{1, 2, \ldots, n\}$ with Quicksort.

$$\boxed{\mathcal{L}(L_n) = \mathcal{L}\left(L_{Z_n - 1} + \overline{L}_{n - Z_n} + n - 1\right)}, \qquad n \geq 2,$$

where $L_0 = L_1 = 0$, $L_2 = 1$, Z_n is uniformly distributed on $\{1, 2, \ldots, n\}$, $\mathcal{L}(L_j) = \mathcal{L}(\overline{L}_j)$, and Z_n, L_j, $\overline{L}_j$ $(1 \leq j \leq n)$ are independent.

$\mathcal{L}(X)$ ... distribution of X.

Quicksort

Expected Number of Comparisions

$$\mathbb{E}\, L_n = n - 1 + \frac{1}{n} \sum_{j=1}^{n} \Big(\mathbb{E}\, L_{j-1} + \mathbb{E}\, L_{n-j} \Big)$$

$\Longrightarrow$

$$\begin{aligned} \mathbb{E}\, L_n &= 2(n+1) \sum_{h=1}^{n+1} \frac{1}{h} - 4(n+1) + 2 \\ &= 2n \log n + n(2\gamma - 4) + 2 \log n + 2\gamma + 1 + \mathcal{O}\left((\log n)/n\right) \end{aligned}$$

with $\gamma = 0.57721...$ being Euler's constant.

Quicksort

Theorem. [Régnier, Rösler]

The normalized number of comparisions

$$\boxed{Y_n = \frac{L_n - \mathbb{E}\, L_n}{n}}$$

converges weakly to a random variable Y:

$$\boxed{Y_n \to Y},$$

which is defined by the fixed point equation

$$\boxed{\mathcal{L}(Y) = \mathcal{L}\Big(UY + (1-U)\overline{Y} + c(U)\Big)},$$

where U is uniformly distributed on $[0,1]$, $\mathcal{L}(\overline{Y}) = \mathcal{L}(Y)$, $U, \overline{Y}, Y$ are independent, and

$$c(x) = 2x \log x + 2(1-x)\log(1-x) + 1.$$

Quicksort

Wasserstein metric d_2

D ... space of distribution functions with finite second moment and zero first moment.

$$\boxed{d_2(F,G) = \inf_{\mathcal{L}(X)=F,\mathcal{L}(Y)=G} \|X - Y\|_2},$$

(D, d_2) constitutes a Polish space.

A sequence F_n converges to F in D if and only if F_n converges weakly to F and if the second moments of F_n converge to the second moment of F.

Quicksort

$Y_n = (L_n - \mathbb{E}\, L_n)/n \Longrightarrow$

$$\boxed{\mathcal{L}(Y_n) = \mathcal{L}\left(Y_{Z_n-1}\frac{Z_n-1}{n} + \overline{Y}_{n-Z_n}\frac{n-Z_n}{n} + c_n(Z_n)\right)}, \qquad n \geq 2,$$

where $Y_0 = Y_1 = 0$, Z_n is uniformly distributed on $\{1, 2, \ldots, n\}$, and $\mathcal{L}(Y_j) = \mathcal{L}(\overline{Y}_j)$, and Z_n, Y_j, $\overline{Y}_j$ $(1 \leq j \leq n)$ are independent, and

$$c_n(j) = \frac{n-1}{n} + \frac{1}{n}\Big(\mathbb{E}\, L_{j-1} + \mathbb{E}\, L_{n-j} - \mathbb{E}\, L_n\Big).$$

If Y_n has a limiting distribution Y then

$$\boxed{\mathcal{L}(Y) = \mathcal{L}\Big(UY + (1-U)\overline{Y} + c(U)\Big)}.$$

Quicksort

Lemma

Let $S : D \to D$ be a map defined by

$$\boxed{S(F) := \mathcal{L}(UX + (1-U)\overline{X} + c(U))},$$

where $X, \overline{X}, U$ are independent, $\mathcal{L}(\overline{X}) = \mathcal{L}(X) = F$, and U is uniformly distributed on $[0, 1]$.

Then S is a **contraction** with respect to the Wasserstein metric d_2 and, thus, there is a unique fixed point $F \in D$ with $S(F) = F$.

Quicksort

Proof

$F, G \in D$, $\mathcal{L}(\overline{X}) = \mathcal{L}(X) = F$, $\mathcal{L}(\overline{Y}) = \mathcal{L}(Y) = G$, U u.d. on $[0,1]$, $U, \overline{X}, X$ and $U, \overline{Y}, Y$ are independent.

$S(F) = \mathcal{L}(UX + (1-U)\overline{X} + c(U))$,
$S(G) = \mathcal{L}(UY + (1-U)\overline{Y} + c(U))$
$\Longrightarrow$

$$\begin{aligned} d_2^2(S(F), S(G)) &\leq \|UX + (1-U)\overline{X} - UY - (1-U)\overline{Y}\|_2^2 \\ &= \|U(X-Y) + (1-U)(\overline{X} - \overline{Y})\|_2^2 \\ &= \mathbb{E}\,(X-Y)^2 \cdot \mathbb{E}\,U^2 + \mathbb{E}\,(\overline{X} - \overline{Y})^2 \cdot \mathbb{E}\,(1-U)^2 \\ &= \frac{2}{3}\mathbb{E}\,(X-Y)^2. \end{aligned}$$

$\Longrightarrow$

$$d_2(S(F), S(G)) \leq \sqrt{\frac{2}{3}} \cdot d_2(F, G),$$

References

Noga Alon and Joel H. Spencer. *The probabilistic method.* Second edition. Wiley-Interscience, New York, 2000

J. Baik, P. Deift and K. Johansson, On the distribution of the length of the longest increasing subsequence of random permutations, *J. Amer. Math. Soc.* 12 (1999), 1119–1178.

Béla Bollobás, *Random graphs*. Second edition. Cambridge Studies in Advanced Mathematics, 73. Cambridge University Press, Cambridge, 2001.

M. Drmota, Stochastic analysis of tree-like data structures, *Proc. R. Soc. Lond. A 460* (2004), 271–307.

References

M. Drmota, Concentration Properties of Extremal Parameters in Random Discrete Structures, *DMTCS*, proc. AG 1–30, 2006.

Svante Janson, Tomasz Łuczak, and Andrzej Rucinski, *Random graphs*. Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley-Interscience, New York, 2000

C. McDiarmid, *Concentration*, Probabilistic methods for algorithmic discrete mathematics, 195–248, Algorithms Combin., 16, Springer, Berlin, 1998.