# Antrittsvorlesung
# Random combinatorial structures: graphs, permutations and representations

Valentin Féray
Assistenzprofessor für reine Mathematik, Universität Zürich

This presentation is recorded and will be made publicy available at
`www.uzh.ch/outreach/events/av.html`

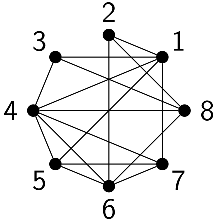
October 27th, 2014

Universität Zürich[UZH]

# First example of random combinatorial structures: random graphs

Erdös-Rényi model of random graphs $G(n, p)$:

- $G$ has $n$ vertices labelled $1, \ldots, n$;
- each edge $\{i, j\}$ is taken independently with probability $p$;

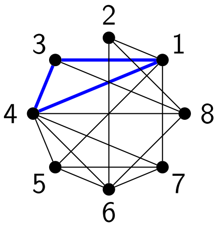



*Example* : $n = 8, p = 1/2$

# First example of random combinatorial structures: random graphs

Erdös-Rényi model of random graphs $G(n, p)$:

- $G$ has $n$ vertices labelled $1, \dots, n$;
- each edge $\{i, j\}$ is taken independently with probability $p$;




*Example* : $n = 8, p = 1/2$

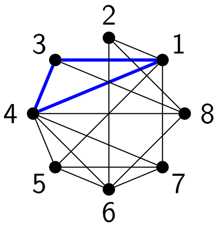
**Question**

Does $G$ contains a triangle? If yes, how many?

# First example of random combinatorial structures: random graphs

Erdös-Rényi model of random graphs $G(n,p)$:

- $G$ has $n$ vertices labelled $1,\dots,n$;
- each edge $\{i,j\}$ is taken independently with probability $p$;




*Example* : $n=8, p=1/2$

**Question**

Does $G$ contains a triangle? If yes, how many?

$\rightarrow$ we look for an asymptotic answer.

# Second example of random combinatorial structure: random permutation

A uniform random permutation of size 20

$$[9, 10, 5, 19, 7, 16, 18, 2, 14, 20, 17, 1, 6, 12, 8, 15, 11, 13, 4, 3]$$

# Second example of random combinatorial structure: random permutation

A uniform random permutation of size 20

$$[9, 10, 5, 19, 7, 16, 18, 2, 14, 20, 17, 1, 6, 12, 8, 15, 11, 13, 4, 3]$$

**Question**

Does $G$ contains an adjacency. If yes, how many?

adjacency = consecutive values in consecutive places.

# Some general motivations

## Some general motivations

- Biological observation: from one person to another, order of the *genes* on a chromosome are the same.
  From one species to another: this order changes, one can encode that by a permutation.
  Number of adjacencies measures how close the two species are.
  $\rightarrow$ One has to compare with a random permutation !

# Some general motivations

- Biological observation: from one person to another, order of the *genes* on a chromosome are the same.
  From one species to another: this order changes, one can encode that by a permutation.
  Number of adjacencies measures how close the two species are.
  $\rightarrow$ One has to compare with a random permutation !
- Spectrum of random (GOE) matrices describe accurately the spectra of heavy atoms (Wigner, 50').

# Some general motivations

- Biological observation: from one person to another, order of the *genes* on a chromosome are the same.
  From one species to another: this order changes, one can encode that by a permutation.
  Number of adjacencies measures how close the two species are.
  $\rightarrow$ One has to compare with a random permutation !
- Spectrum of random (GOE) matrices describe accurately the spectra of heavy atoms (Wigner, 50').
- Graphs can represent networks, *e.g.* internet (vertices are web pages, edges hyperlinks).
  But Erdős-Rényi random graph is not a good model for internet (not the good degree distribution!).
  $\rightarrow$ A lot of other models of random graphs have been introduced and studied.

# Outline of the talk

1. Presentation of the moment method

2. Two facets of my work related to moment methods
   - Random permutations and small cumulants
   - Random representations

# First moment: description of the method

**Lemma**

Let $X$ be a random variable with non-negative integer values (for example $X$ is counting something). Then

$$P(X=0) > 1 - E(X).$$

Proof: $E(X) = \sum_k kP(X=k) \geq \sum_{k \geq 1} P(X=k) = 1 - P(X=0)$. $\square$

# First moment: description of the method

**Lemma**

Let $X$ be a random variable with non-negative integer values (for example $X$ is counting something). Then

$$P(X = 0) > 1 - E(X).$$

Proof: $E(X) = \sum_k kP(X = k) \geq \sum_{k \geq 1} P(X = k) = 1 - P(X = 0)$. $\square$

**Corollary (first moment method)**

If $E(X_n)$ tends to 0, then $X_n = 0$ with high probability (that is, $P(X_n = 0)$ tends to 1).

# First moment: application to triangles in random graphs

Let $T_n$ be the number of triangles in $G(n, p_n)$. Then

$$T_n = \sum_{\{i,j,k\} \subset [n]} \Delta_{\{i,j,k\}},$$

where

$$\Delta_{\{i,j,k\}} = \begin{cases} 1 & \text{if } G \text{ contains the triangle } i, j, k; \\ 0 & \text{otherwise.} \end{cases}$$

# First moment: application to triangles in random graphs

Let $T_n$ be the number of triangles in $G(n, p_n)$. Then

$$T_n = \sum_{\{i,j,k\} \subset [n]} \Delta_{\{i,j,k\}},$$

where

$$\Delta_{\{i,j,k\}} = \begin{cases} 1 & \text{if } G \text{ contains the triangle } i, j, k; \\ 0 & \text{otherwise.} \end{cases}$$

But $E(\Delta_{\{i,j,k\}}) = p_n^3$ for all $\{i, j, k\}$ and thus

$$E(T_n) = \sum_{\{i,j,k\}} E(\Delta_{\{i,j,k\}}) = \binom{n}{3} p_n^3.$$

# First moment: application to triangles in random graphs

Let $T_n$ be the number of triangles in $G(n, p_n)$. Then

$$T_n = \sum_{\{i,j,k\} \subset [n]} \Delta_{\{i,j,k\}},$$

where

$$\Delta_{\{i,j,k\}} = \begin{cases} 1 & \text{if } G \text{ contains the triangle } i, j, k; \\ 0 & \text{otherwise.} \end{cases}$$

But $E(\Delta_{\{i,j,k\}}) = p_n^3$ for all $\{i, j, k\}$ and thus

$$E(T_n) = \sum_{\{i,j,k\}} E(\Delta_{\{i,j,k\}}) = \binom{n}{3} p_n^3.$$

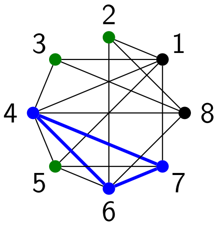
**Theorem**

*If $n\, p_n \to 0$, then $G(n, p_n)$ has no triangles with high probability.*

# Variant: probabilistic method applied to Ramsay number.

Fix $p = 1/2$ and $k \geq 1$. Then consider $X_n$ the number of sets $W$ of $k$ vertices in $G(n, 1/2)$ which are:

- either a clique, all pairs of vertices of $W$ is linked by an edge;
- or an independent set, *i.e.* there is no edge between two vertices of $W$

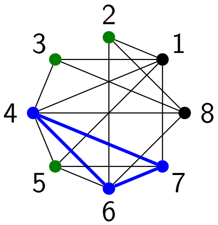



*Example* : $n = 8, k = 3$

# Variant: probabilistic method applied to Ramsay number.

Fix $p = 1/2$ and $k \geq 1$. Then consider $X_n$ the number of sets $W$ of $k$ vertices in $G(n, 1/2)$ which are:

- either a clique, all pairs of vertices of $W$ is linked by an edge;
- or an independent set, *i.e.* there is no edge between two vertices of $W$




*Example* : $n = 8, k = 3$

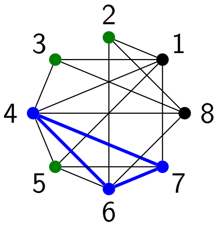
Then

$$P(X_n = 0) > 1 - E(X_n) = 1 - 2^{1-\binom{k}{2}} \binom{n}{k}$$

# Variant: probabilistic method applied to Ramsay number.

Fix $p = 1/2$ and $k \geq 1$. Then consider $X_n$ the number of sets $W$ of $k$ vertices in $G(n, 1/2)$ which are:

- either a clique, all pairs of vertices of $W$ is linked by an edge;
- or an independent set, *i.e.* there is no edge between two vertices of $W$




*Example* : $n = 8, k = 3$

Then

$$P(X_n = 0) > 1 - E(X_n) = 1 - 2^{1-\binom{k}{2}} \binom{n}{k}$$

**Theorem (Erdős, 1947)**

*If* $1 - 2^{1-\binom{k}{2}} \binom{n}{k} > 0$, *then* **there exists** *a graph* $G$ *with* $n$ *vertices and neither cliques nor independent sets of size* $k$.

# Second moment: description of the method

⚠ $\lim_{n\to\infty} E(X_n) = \infty$ does NOT imply $\lim_n P(X_n = 0) = 0$.

# Second moment: description of the method

⚠ $\lim_{n\to\infty} E(X_n) = \infty$ does NOT imply $\lim_n P(X_n = 0) = 0$.

How to prove that "$\lim_n P(X_n = 0) = 0$" ?

Idea: use variance $\mathrm{Var}(X) := E(X^2) - E(X)^2 = E\big[(X - E(X))^2\big]$.

# Second moment: description of the method

⚠ $\lim_{n\to\infty} E(X_n) = \infty$ does NOT imply $\lim_n P(X_n = 0) = 0$.

How to prove that "$\lim_n P(X_n = 0) = 0$" ?

Idea: use variance $\mathrm{Var}(X) := E(X^2) - E(X)^2 = E\big[(X - E(X))^2\big]$.

**Lemma (Chebyshev's inequality)**

$$P\left(|X - E(X)| \geq \lambda\sqrt{\mathrm{Var}(X)}\right) \leq 1/\lambda^2.$$

In particular,

$$P(X = 0) \leq \left(\frac{E(X)}{\sqrt{\mathrm{Var}(X)}}\right)^{-2}.$$

Proof: If $Y$ is a non-negative r.v., then $P(Y \geq a) \leq E(Y)/a$.

Apply this to $Y = (X - E(X))^2$ and $a = \lambda^2 \mathrm{Var}(X) = \lambda^2 E(Y)$. □

# Second moment: description of the method

⚠ $\lim_{n\to\infty} E(X_n) = \infty$ does NOT imply $\lim_n P(X_n = 0) = 0$.

How to prove that "$\lim_n P(X_n = 0) = 0$" ?

Idea: use variance $\mathrm{Var}(X) := E(X^2) - E(X)^2 = E\left[(X - E(X))^2\right]$.

## Lemma (Chebyshev's inequality)

$$P\left(|X - E(X)| \geq \lambda\sqrt{\mathrm{Var}(X)}\right) \leq 1/\lambda^2.$$

In particular,

$$P(X = 0) \leq \left(\frac{E(X)}{\sqrt{\mathrm{Var}(X)}}\right)^{-2}.$$

## Second moment method

If $\sqrt{\mathrm{Var}(X_n)}/E(X_n) \to 0$, then $X_n > 0$ with high probability.

# Second moment: application to triangles in random graphs

Recall that $E(T_n) \sim c n^3 p_n^3$. But

$$\mathrm{Var}(T_n) = E\left[(T_n - E(T_n))^2\right] = E\left[\left(\sum_{\{i,j,k\}} \Delta_{\{i,j,k\}} - E(\Delta_{\{i,j,k\}})\right)^2\right]$$

$$= \sum_{\substack{\{i,j,k\}\\ \{i',j',k'\}}} E\left[\left(\Delta_{\{i,j,k\}} - E(\Delta_{\{i,j,k\}})\right) \cdot \left(\Delta_{\{i',j',k'\}} - E(\Delta_{\{i',j',k'\}})\right)\right]$$

# Second moment: application to triangles in random graphs

Recall that $E(T_n) \sim c n^3 p_n^3$. But

$$\mathrm{Var}(T_n) = E\big[(T_n - E(T_n))^2\big] = E\left[\left(\sum_{\{i,j,k\}} \Delta_{\{i,j,k\}} - E(\Delta_{\{i,j,k\}})\right)^2\right]$$

$$= \sum_{\substack{\{i,j,k\} \\ \{i',j',k'\}}} E\left[\left(\Delta_{\{i,j,k\}} - E(\Delta_{\{i,j,k\}})\right) \cdot \left(\Delta_{\{i',j',k'\}} - E(\Delta_{\{i',j',k'\}})\right)\right]$$

- if $\{i,j,k\}$ and $\{i',j',k'\}$ do not share an edge, then $\Delta_{\{i,j,k\}}$ and $\Delta_{\{i',j',k'\}}$ are independent and $E[\dots]$ is zero.

# Second moment: application to triangles in random graphs

Recall that $E(T_n) \sim c n^3 p_n^3$. But

$$\mathrm{Var}(T_n) = E\big[(T_n - E(T_n))^2\big] = E\left[\left(\sum_{\{i,j,k\}} \Delta_{\{i,j,k\}} - E(\Delta_{\{i,j,k\}})\right)^2\right]$$

$$= \sum_{\substack{\{i,j,k\} \\ \{i',j',k'\}}} E\left[\left(\Delta_{\{i,j,k\}} - E(\Delta_{\{i,j,k\}})\right) \cdot \left(\Delta_{\{i',j',k'\}} - E(\Delta_{\{i',j',k'\}})\right)\right]$$

- if $\{i,j,k\}$ and $\{i',j',k'\}$ do not share an edge, then $\Delta_{\{i,j,k\}}$ and $\Delta_{\{i',j',k'\}}$ are independent and $E[\dots]$ is zero.

But this is true for most of the terms in the sum! One can show:

**Theorem**

If $n\, p_n \to \infty$, then $\sqrt{\mathrm{Var}(T_n)}/E(T_n)$ tends to 0 and thus $G(n, p_n)$ contains a triangle with high probability.

# Transition

- To prove that our random structure does not contain a given type of substructure:
  set $X_n$ the number of these substructures $X_n = \sum_{\alpha} I_\alpha$.
  
  Compute $E(X_n)$ by linearity and show that it tends to 0.

# Transition

- To prove that our random structure does not contain a given type of substructure:
  set $X_n$ the number of these substructures $X_n = \sum_{\alpha} I_\alpha$.
  Compute $E(X_n)$ by linearity and show that it tends to 0.
- To prove that our random structure does contain a given type of substructure:
  consider $E(X_n^2)$ (or $\mathrm{Var}(X_n)$), expand, use linearity of expectation and independence and show that

$$\lim_{n\to\infty} \sqrt{\mathrm{Var}(T_n)}/E(T_n) = 0.$$

# Transition

- To prove that our random structure does not contain a given type of substructure:
  set $X_n$ the number of these substructures $X_n = \sum_{\alpha} I_\alpha$.
  Compute $E(X_n)$ by linearity and show that it tends to 0.
- To prove that our random structure does contain a given type of substructure:
  consider $E(X_n^2)$ (or $\mathrm{Var}(X_n)$), expand, use linearity of expectation and independence and show that

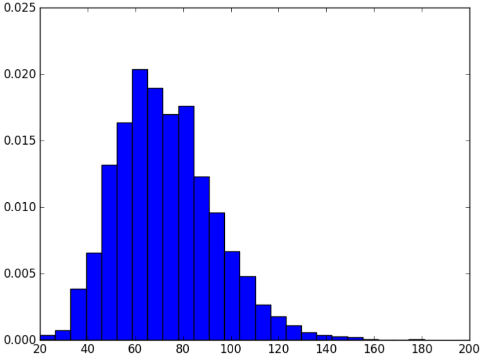
$$\lim_{n\to\infty} \sqrt{\mathrm{Var}(T_n)}/E(T_n) = 0.$$

**Question**

In the case where there is a triangle with high probability, how many are there ?

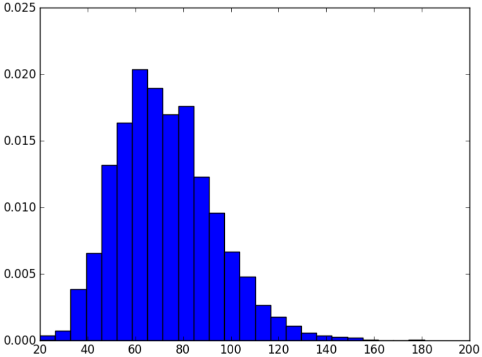
# Number of triangles in random graphs: a simulation

Here is a histogram number of triangles in 5000 independent random graphs $G(20, .4)$.



# Number of triangles in random graphs: a simulation

Here is a histogram number of triangles in 5000 independent random graphs $G(20, .4)$.




**Theorem (Ruciński, 1988)**

If $np_n \to \infty$ and $n^2(1-p_n) \to \infty$, then $T_n$ is asymptotically Gaussian.

# The moment method

### Theorem (the moment method)

Let $X_n$ be a sequence of random variable and $N$ a standard normal random variable. If for each $\ell \geq 1$,

$$\lim_{n \to \infty} E(X_n^\ell) = E(N^\ell),$$

then $X_n$ converges to $N$ in distribution.

Remark:

$$E(N^\ell) = \begin{cases} 1 \cdot 3 \cdots (2m-1) & \text{if } \ell = 2m \text{ is even;} \\ 0 & \text{if } \ell \text{ is odd.} \end{cases}$$

# Sketch of proof of Ruciński theorem

Let $T_n$ be the number of triangles in $G(n, p_n)$ and $X_n = \frac{T_n - E(T_n)}{\sqrt{\mathrm{Var}(T_n)}}$. Then

$$E(X_n^\ell) = \frac{1}{\mathrm{Var}(T_n)^{\ell/2}} E\left( (T_n - E(T_n))^\ell \right)$$

But $T_n - E(T_n) = \sum_{\{\{i,j,k\}\}} \Delta_{\{i,j,k\}} - E(\Delta_{\{i,j,k\}})$. Thus

$$\mathrm{Var}(T_n)^{\ell/2} E(X_n^\ell) = \sum_{\substack{\{i_1,j_1,k_1\} \\ \vdots \\ \{i_\ell,j_\ell,k_\ell\}}} E\Big[ (\Delta_{\{i_1,j_1,k_1\}} - E(\Delta_{\{i_1,j_1,k_1\}})) \cdots (\Delta_{\{i_\ell,j_\ell,k_\ell\}} - E(\Delta_{\{i_\ell,j_\ell,k_\ell\}})) \Big]$$

# Sketch of proof of Ruciński theorem

Let $T_n$ be the number of triangles in $G(n, p_n)$ and $X_n = \frac{T_n - E(T_n)}{\sqrt{\mathrm{Var}(T_n)}}$. Then

$$E(X_n^\ell) = \frac{1}{\mathrm{Var}(T_n)^{\ell/2}} E\left( (T_n - E(T_n))^\ell \right)$$

But $T_n - E(T_n) = \sum_{\{\{i,j,k\}\}} \Delta_{\{i,j,k\}} - E(\Delta_{\{i,j,k\}})$. Thus

$$\mathrm{Var}(T_n)^{\ell/2} E(X_n^\ell) = \sum_{\substack{\{i_1,j_1,k_1\} \\ \vdots \\ \{i_\ell,j_\ell,k_\ell\}}} E\Big[ (\Delta_{\{i_1,j_1,k_1\}} - E(\Delta_{\{i_1,j_1,k_1\}})) \cdots (\Delta_{\{i_\ell,j_\ell,k_\ell\}} - E(\Delta_{\{i_\ell,j_\ell,k_\ell\}})) \Big]$$

Because of independence, most of the terms vanish... One has to study asymptotics carefully and find $\lim E(X_n^\ell) = E(N^\ell)$.

# A TCL for the number of prime divisors

Each integer $X \geq 1$ writes uniquely as a product of primes:

$$18984 = 2^3 * 3 * 7 * 113.$$

**Theorem (Erdős, Kac, 1940)**

Let $D_n$ be the number of prime divisors of a uniform random integer between 1 and $n$. Then

$$X_n := \frac{D_n - E(D_n)}{\sqrt{Var(D_n)}} \longrightarrow_d N.$$

$$D_n = \sum_{\substack{p \text{ prime} \\ p \le n}} I_p(n) \text{ where } I_p(n) = \begin{cases} 1 & \text{if } p \text{ divides } n; \\ 0 & \text{otherwise.} \end{cases}$$

As usual, we want to compute $E(D_n^\ell)$, we expand and use linearity...

$$D_n = \sum_{\substack{p \text{ prime} \\ p \le n}} I_p(n) \text{ where } I_p(n) = \begin{cases} 1 & \text{if } p \text{ divides } n; \\ 0 & \text{otherwise.} \end{cases}$$

As usual, we want to compute $E(D_n^\ell)$, we expand and use linearity...

No independence, but, for distinct primes,

$$E(I_{p_1} \dots I_{p_k}) = \frac{1}{n} \left\lfloor \frac{n}{p_1 \dots p_k} \right\rfloor \simeq \underbrace{\frac{1}{p_1 \dots p_k}}_{\substack{\text{what we would have} \\ \text{if } I_{p_1} \dots I_{p_k} \text{ were independent}}} .$$

$$D_n = \sum_{\substack{p \text{ prime} \\ p \leq n}} I_p(n) \text{ where } I_p(n) = \begin{cases} 1 & \text{if } p \text{ divides } n; \\ 0 & \text{otherwise.} \end{cases}$$

As usual, we want to compute $E(D_n^\ell)$, we expand and use linearity...

No independence, but, for distinct primes,

$$E(I_{p_1} \dots I_{p_k}) = \frac{1}{n} \left\lfloor \frac{n}{p_1 \dots p_k} \right\rfloor \simeq \underbrace{\frac{1}{p_1 \dots p_k}}_{\substack{\text{what we would have} \\ \text{if } I_{p_1} \dots I_{p_k} \text{ were independent}}} .$$

So

$$E(D_n^\ell) \simeq E(Y_n^\ell),$$

where $Y_n$ is a sum of independent copies of the $I_p$. It converge after normalization to a normal law.

$$D_n = \sum_{\substack{p \text{ prime} \\ p \leq n}} I_p(n) \text{ where } I_p(n) = \begin{cases} 1 & \text{if } p \text{ divides } n; \\ 0 & \text{otherwise.} \end{cases}$$

As usual, we want to compute $E(D_n^\ell)$, we expand and use linearity...

No independence, but, for distinct primes,

$$E(I_{p_1} \dots I_{p_k}) = \frac{1}{n} \left\lfloor \frac{n}{p_1 \dots p_k} \right\rfloor \simeq \underbrace{\frac{1}{p_1 \dots p_k}}_{\substack{\text{what we would have} \\ \text{if } I_{p_1} \dots I_{p_k} \text{ were independent}}} .$$

So

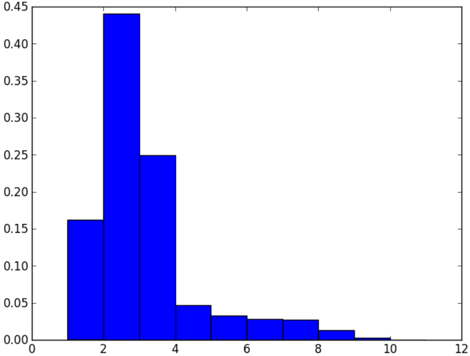
$$E(D_n^\ell) \simeq E(Y_n^\ell),$$

where $Y_n$ is a sum of independent copies of the $I_p$. It converge after normalization to a normal law.

$\Rightarrow D_n$ also converges after normalization to a normal law.

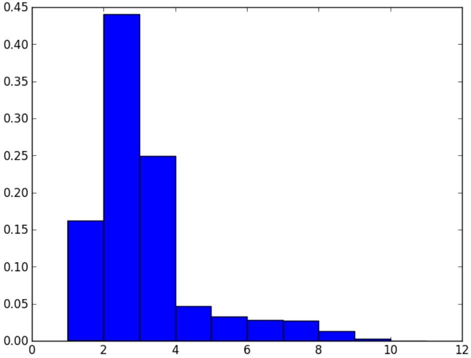
# Number of prime divisors: a simulation

$n = 28511235679461510605581038657982805983853648817939444953417128836$



# Number of prime divisors: a simulation

$n = 2851123567946151060558103865798280598385364881793944495341712883 6$




$\ln(\ln(n)) \sim 5$, but the empirical mean is 2.5.

Erdős-Kac theorem is an asymptotic result!

# Random permutations and non independence

**Definition**

A permutation of $n$ is a word with letters from 1 to $n$, which contains exactly one each letter.

A uniform random permutation of size 20:

$$[9, 10, 5, 19, 7, 16, 18, 2, 14, 20, 17, 1, 6, 12, 8, 15, 11, 13, 4, 3]$$

The first element $\sigma_1$ is uniform between 1 and $n$. Also the second $\sigma_2$, ...
But they are not independent.

# Measuring their dependence

Set $\sigma'_1 = \sigma_1/n$, $\sigma'_2 = \sigma_2/n$. Their *covariance* is

$$\mathrm{Cov}(\sigma'_1, \sigma'_2) := E(\sigma'_1 \sigma'_2) - E(\sigma'_1)E(\sigma'_2) = \cdots = O(1/n).$$

Small correlation!

# Measuring their dependence

Set $\sigma'_1 = \sigma_1/n$, $\sigma'_2 = \sigma_2/n$. Their *covariance* is

$$\mathrm{Cov}(\sigma'_1, \sigma'_2) := E(\sigma'_1\,\sigma'_2) - E(\sigma'_1)E(\sigma'_2) = \cdots = O(1/n).$$

Small correlation! Let us compute the 3rd cumulant:

$$\begin{aligned}\kappa_3(\sigma'_1, \sigma'_2, \sigma'_3) = E(\sigma'_1\,\sigma'_2\,\sigma'_3) - E(\sigma'_1\,\sigma'_2)E(\sigma'_3) - \cdots + 2E(\sigma'_1)E(\sigma'_2)E(\sigma'_3)\\ = \cdots = O(1/n^2).\end{aligned}$$

# Measuring their dependence

Set $\sigma'_1 = \sigma_1/n$, $\sigma'_2 = \sigma_2/n$. Their *covariance* is

$$\mathrm{Cov}(\sigma'_1, \sigma'_2) := E(\sigma'_1\, \sigma'_2) - E(\sigma'_1)E(\sigma'_2) = \cdots = O(1/n).$$

Small correlation! Let us compute the 3rd cumulant:

$$\begin{aligned}\kappa_3(\sigma'_1, \sigma'_2, \sigma'_3) = E(\sigma'_1\, \sigma'_2\, \sigma'_3) - E(\sigma'_1\, \sigma'_2)E(\sigma'_3) - \cdots + 2E(\sigma'_1)E(\sigma'_2)E(\sigma'_3) \\ = \cdots = O(1/n^2).\end{aligned}$$

In general, $\kappa_\ell(\sigma'_1, \ldots, \sigma'_\ell) = O(n^{-\ell+1})$ (F., 2013)

# Measuring their dependence

Set $\sigma'_1 = \sigma_1/n$, $\sigma'_2 = \sigma_2/n$. Their *covariance* is

$$\mathrm{Cov}(\sigma'_1, \sigma'_2) := E(\sigma'_1\,\sigma'_2) - E(\sigma'_1)E(\sigma'_2) = \cdots = O(1/n).$$

Small correlation! Let us compute the 3rd cumulant:

$$\kappa_3(\sigma'_1, \sigma'_2, \sigma'_3) = E(\sigma'_1\,\sigma'_2\,\sigma'_3) - E(\sigma'_1\,\sigma'_2)E(\sigma'_3) - \cdots + 2E(\sigma'_1)E(\sigma'_2)E(\sigma'_3)$$
$$= \cdots = O(1/n^2).$$

In general, $\kappa_\ell(\sigma'_1, \ldots, \sigma'_\ell) = O(n^{-\ell+1})$ (F., 2013)

For "local" statistics, allows to use moment method as if $\sigma_1, \sigma_2, \ldots$ were independent.

Example: the number of adjacencies is asymptotically Poisson distributed!

# Measuring their dependence

Set $\sigma'_1 = \sigma_1/n$, $\sigma'_2 = \sigma_2/n$. Their *covariance* is

$$\mathrm{Cov}(\sigma'_1, \sigma'_2) := E(\sigma'_1\, \sigma'_2) - E(\sigma'_1)E(\sigma'_2) = \cdots = O(1/n).$$

Small correlation! Let us compute the 3rd cumulant:

$$\begin{aligned}\kappa_3(\sigma'_1, \sigma'_2, \sigma'_3) = E(\sigma'_1\, \sigma'_2\, \sigma'_3) - E(\sigma'_1\, \sigma'_2)E(\sigma'_3) - \cdots + 2E(\sigma'_1)E(\sigma'_2)E(\sigma'_3) \\ = \cdots = O(1/n^2).\end{aligned}$$

In general, $\kappa_\ell(\sigma'_1, \dots, \sigma'_\ell) = O(n^{-\ell+1})$ (F., 2013)

For "local" statistics, allows to use moment method as if $\sigma_1, \sigma_2, \dots$ were independent.

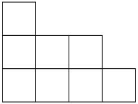
Small cumulants also appear in random graphs with fixed number of edges, random orthogonal/unitary matrices, ...

# Partitions

**Definition**

A partition (of $n$) is a non-increasing list of integer (of sum $n$).

Example : $(4, 3, 1)$ is a partition of 8.

Representation as Young diagram : 

# Plancherel measure

Representation theory of symmetric group associates to each partition $\lambda$ of $n$ a vector space $V_\lambda$, called representation, such that

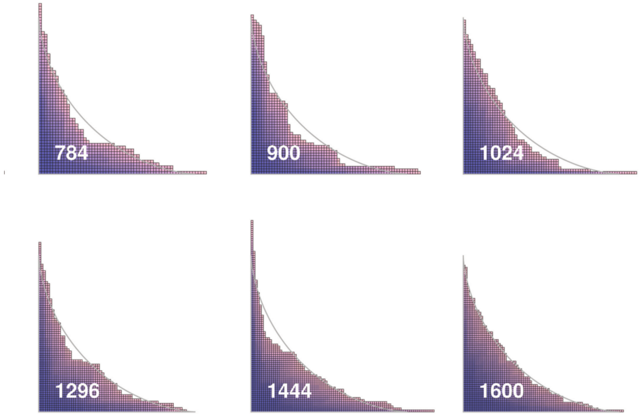
$$\sum_{\substack{\lambda \text{partition} \\ \text{of } n}} (\dim V_\lambda)^2 = n!$$

We consider Plancherel measure on partitions of size $n$:

$$P(\lambda) = \frac{(\dim V_\lambda)^2}{n!}.$$

This defines a model for a random partition of size $n$, but not uniform.

# Random partitions under Plancherel measure: simulation




©Notices of the AMS, Feb. 2011, front cover.

Nice limit shape!

# Moment method on characters

**Lemma (easy algebraic statement rewritten in probabilistic terms)**

$$E(\hat{\chi}^\bullet(\sigma)) = \begin{cases} 1 & \text{if } \sigma = \mathrm{id}; \\ 0 & \text{otherwise.} \end{cases},$$

where $\hat{\chi}$ are irreducible characters.

We would like to use moment method and compute $E(\hat{\chi}^\bullet(\sigma)^m)$.

# Moment method on characters

**Lemma (easy algebraic statement rewritten in probabilistic terms)**

$$E(\hat{\chi}^\bullet(\sigma)) = \begin{cases} 1 & \text{if } \sigma = \mathrm{id}; \\ 0 & \text{otherwise.} \end{cases},$$

where $\hat{\chi}$ are irreducible characters.

We would like to use moment method and compute $E(\hat{\chi}^\bullet(\sigma)^m)$.

"Solution": express $\hat{\chi}^\bullet(\sigma)^m$ as a linear combination of irreducible characters.

This is a classical kind of question in algebraic combinatorics!

# Moment method on characters

**Lemma (easy algebraic statement rewritten in probabilistic terms)**

$$E(\hat{\chi}^\bullet(\sigma)) = \begin{cases} 1 & \text{if } \sigma = \mathrm{id}; \\ 0 & \text{otherwise.} \end{cases},$$

where $\hat{\chi}$ are irreducible characters.

We would like to use moment method and compute $E(\hat{\chi}^\bullet(\sigma)^m)$.

"Solution": express $\hat{\chi}^\bullet(\sigma)^m$ as a linear combination of irreducible characters.

This is a classical kind of question in algebraic combinatorics!

Algebraic combinatorics provides a lot of non-trivial probabilistic models that can be analysed by moment method $\rightarrow$ integrable probability.