Lecture 2: 31/10/2017

Reminder:

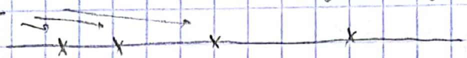
the wigner model: $X_N = \frac{1}{\sqrt{n}} \overbrace{\begin{pmatrix} Y_1 & Z_{1,2} & Z_{1,3} & \cdots \\ & Y_2 & Z_{2,3} & \cdots \\ & \text{Sym} & \ddots & \\ & & & Y_n \end{pmatrix}}^{N} \Big\} N$

$(Z_{i,j})_{j>i}$ IID $\nwarrow$ indep.
(Y_i) IID $\swarrow$

we want to study the eigenvalues of X_N: $\lambda_1^N \leq \lambda_2^N \leq \ldots \leq \lambda_N^N$

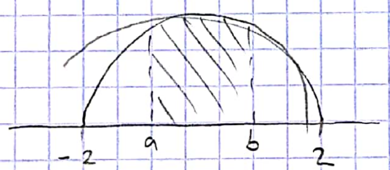
empirical ~~distrinbe~~ distribution: $\mathcal{L}_N := \frac{1}{N} \sum_{j=1}^{N} \delta_{\lambda_j^N}$ ← this is a measure.

atoms of size $\frac{1}{N}$



Wigners semicircle distribution:

$\sigma(x)dx$ where $\sigma(x) = \frac{1}{2\pi} \sqrt{4-x^2}\, \mathbb{1}_{[-2,2]}(x)$



Wigner thm.: $\mathcal{L}_N \longrightarrow \sigma(x)dx$ in distribution (weakly), in probability, that is:

$\forall f \in C_b(\mathbb{R})$ (f bdd. cont. on $\mathbb{R}$):

$$P\left(\left| \int f(x)\, d\mathcal{L}_N(x) - \int f(x)\, \sigma(x)\, dx \right| > \varepsilon \right) \xrightarrow[N \to \infty]{} 0$$

under the assumptions:

1) $\mathbb{E} Y_i = \mathbb{E} Z_{i,j} = 0$

2) $\mathbb{E} Z_{i,j}^2 = 1$

3) $\forall k \geq 2$, $\max\left(\mathbb{E}[|Y_i|^k], \mathbb{E}[|Z_{i,j}|^k] \right) < \infty$

Idea of method of moments:

(Notation: $\int f(x)\, d\mu(x) =: \langle \mu, f \rangle$)

$$\langle \mathcal{L}_N, x^k \rangle = \frac{1}{N} \sum_{j=1}^{N} (\lambda_i^N)^k = \frac{1}{N} \mathrm{Tr}(X_N^k)$$

$\Downarrow$

moments of σ:

prop.:

$$\forall k \geq 0: \quad \langle \sigma, x^k \rangle = \begin{cases} 0 & k \text{ odd} \\ C_{\frac{k}{2}} & k \text{ even} \end{cases}$$

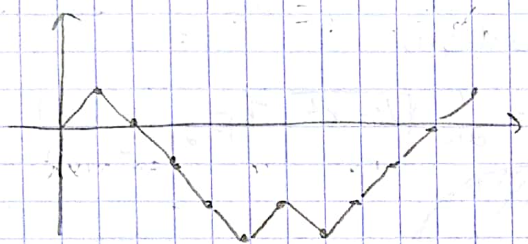
where: (Catalan):

$$C_m = \frac{1}{m+1} \cdot \binom{2m}{m} = \frac{(2m)!}{(m+1)!\, m!} \qquad m \geq 0$$

Short detour on Catalan numbers:

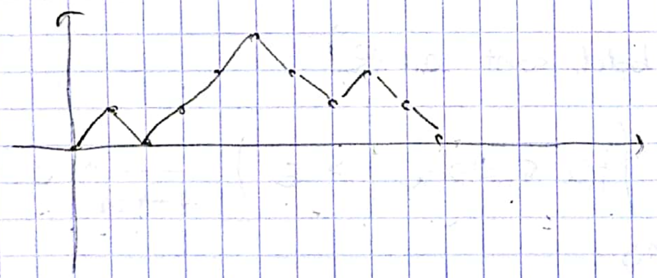
Bernoulli walk of length $2k$:

$x = (x_0, x_1, \dots, x_{2k})$ where $x_0 = 0$, $|x_{i+1} - x_i| = 1$ $\forall i \geq 0$.



Dyck path:

$x_p \geq 0$, $\forall p$. & $x_{2k} = 0$

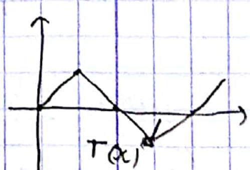


Lemma:

C_k equals the number of Dyck paths of length $2k$.

Proof:

There are $\binom{2k}{k}$ walks of length $2k$ with $x_{2k} = 0$ (exactly k steps go up & k steps go down)



For each such walk x that is not a Dyck path, let $T(x)$ be the first index when $x_{T(x)} = -1$. there is a mapping (the reflection mapping) to a walk ending at -2. This mapping is a bijection. We conclude that the number of Dyck paths of len. $2k$ is:

$$\binom{2k}{k} - \binom{2k}{k-1} = \frac{2k!}{k!\,k!} - \frac{2k!}{(k-1)!\,(k-1)!} = C_k$$

$\boxtimes$

<u>Lemma:</u>

מה זה פונקציה יוצרת?

The generating function of C_k is:

$$f(z) := \sum_{m=0}^{\infty} C_m z^m$$

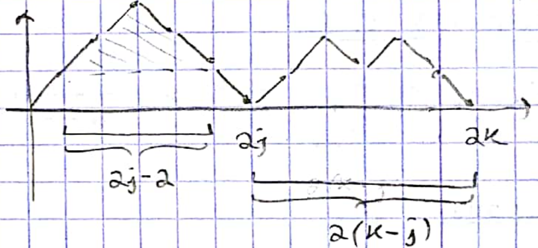
satisfies: $f(z) = \frac{1-\sqrt{1-4z}}{2z}$, $|z| < \frac{1}{4}$

<u>Proof:</u>

First, $C_m \leq 4^m$ (the dyck paths of length $2k$ are a subset of the bernoulli paths of length $2k$. there are 2^{2k} of these)

for each Dyck path, by considering the first time $2j$ where it equals 0, get:

$$(*) \quad C_k = \sum_{j=1}^{k} C_{j-1} C_{k-j}$$



we calculate that:

$$f(z) - 1 = \sum_{k=1}^{\infty} C_k z^k \overset{(*)}{=} \sum_{k=1}^{\infty} \sum_{j=1}^{k} C_{j-1} C_{k-j} z^k$$

$$\& \quad f(z)^2 = \left(\sum_{m=0}^{\infty} C_m z^m\right)^2 = \sum_{k,k'=0}^{\infty} C_k C_{k'} z^{k+k'} = \sum_{q=0}^{\infty} z^q \sum_{j=0}^{q} C_j C_{q-j}$$

Then: $f(z) - 1 = z(f(z))^2$

This means:

$$f(z) \in \left\{ \frac{1 \pm \sqrt{1-4z}}{2z} \right\}$$ (by solving the quadratic equation)

since $f(0) = 1$, $f(z) = \frac{1-\sqrt{1-4z}}{2z}$

$\boxtimes$

Proof of Wigner's thm:

Def.:

$\bar{L}_N$ is the probability measure $\mathbb{E}(L_N)$. That is, for any measurable $f:\mathbb{R}\to[0,\infty)$, $\langle \bar{L}_N, f\rangle := \mathbb{E}[\langle L_N, f\rangle]$

(in particular, for $f = \mathbb{1}_A$, A Borel set).

One checks that $\bar{L}_N$ is a probability measure.

E.g. σ additivity is the claim that if $A_1, A_2, \ldots$ are disjoint Borel sets, then

$$\langle \bar{L}_N, \mathbb{1}_{A_1} + \mathbb{1}_{A_2} + \ldots\rangle = \sum_{j=1}^{\infty} \underbrace{\langle \bar{L}_N, \mathbb{1}_{A_j}\rangle}_{\mathbb{E}\langle L_N, \mathbb{1}_{A_j}\rangle}$$

$$\mathbb{E}\langle L_N, \mathbb{1}_{A_1} + \ldots\rangle \overset{\text{by monotone convergance theorem}}{=} \mathbb{E}\left[\sum_{j=1}^{\infty}\langle L_N, \mathbb{1}_{A_j}\rangle\right]$$

Rely on two lemmas for thm.:

Lemma 1:

$$\forall k\geq 0.\ \langle \bar{L}_N, x^k\rangle \xrightarrow[N\to\infty]{} \langle \sigma, x^k\rangle$$

Lemma 2:

for any $k\geq 0$ & $\varepsilon>0$. $P\left(|\langle L_N, x^k\rangle - \langle \bar{L}_N, x^k\rangle| > \varepsilon\right) \xrightarrow[N\to\infty]{} 0$

Now the thm. follows from the lemmas:

Fix $f\in C_b(\mathbb{R})$. Fix $\delta>0$. Let Q_δ be a poly. satisfying

$$|f(x) - Q_\delta(x)| \leq \delta \quad \text{for } -5\leq x\leq 5$$

$$|\langle L_N, f\rangle - \langle \sigma, f\rangle| = |\langle L_N, \underbrace{(f-Q_\delta)\mathbb{1}_{|x|\leq 5}}_{|\cdot|\leq\delta} + (f-Q_\delta)\mathbb{1}_{|x|>5} + Q_\delta\rangle -$$

$$- \langle \sigma, \underbrace{(f-Q_\delta)\mathbb{1}_{|x|\leq 5}}_{|\cdot|\leq\delta} + \underbrace{(f-Q_\delta)\mathbb{1}_{|x|>5}}_{\langle\sigma,\cdot\rangle = 0} + Q_\delta\rangle| \leq$$

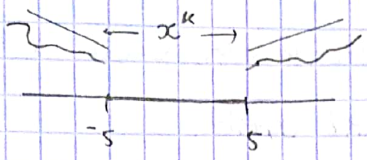
$$\leq 2\delta + \underbrace{\langle L_N, |f-Q_\delta|\mathbb{1}_{|x|\geq 5}\rangle}_{(*)} + \underbrace{|\langle L_N, Q_\delta\rangle - \langle\sigma, Q_\delta\rangle|}_{(**)}$$

$$(**) \leq \underbrace{|\langle \underbrace{L_N, Q_\delta}_{\text{random}}\rangle - \langle \bar{L}_N, Q_\delta\rangle|}_{\forall\varepsilon,\ P(\cdot>\varepsilon)\xrightarrow[N\to\infty]{} 0 \text{ (lemma 2)}} + |\langle \bar{L}_N, Q_\delta\rangle - \langle\sigma, Q_\delta\rangle|$$

deterministic

lemma 1 $\xrightarrow[N\to\infty]{} 0$

To control $(*)$ note that $\exists k_0$ s.t. $|f-Q_\delta|\mathbb{1}_{|x|>5} \le |x|^k \mathbb{1}_{|x|>5}$, $k>k_0$



$|f-Q_\delta|\mathbb{1}_{|x|>5}$

(this is possible because f is bounded)

So that: $(*) = \langle \mathcal{Z}_N, |f-Q_\delta|\mathbb{1}_{|x|>5} \rangle \le \langle \mathcal{Z}_N, |x|^k \mathbb{1}_{|x|>5} \rangle$

$$\le \frac{\langle \mathcal{Z}_N, x^{2k} \rangle}{5^k}$$

$\forall x\ |x|^k \mathbb{1}_{|x|>5} \le \dfrac{x^{2k}}{5^k} \rightarrow$

Now we take expectation:

$$\mathbb{E}\langle \mathcal{Z}_N, |f-Q_\delta|\mathbb{1}_{|x|>5} \rangle \le \frac{1}{5^k}\langle \bar{\mathcal{Z}}_N, x^{2k} \rangle \xrightarrow[N\to\infty]{} \frac{C_k}{5^k} \le \left(\frac{4}{5}\right)^k$$

$\forall k \ge k_0$

$$\Rightarrow \limsup_{N\to\infty} \mathbb{E}\langle \mathcal{Z}_N, |f-Q_\delta|\mathbb{1}_{|x|>5} \rangle = 0$$

Conclude with Markov's inequality.

<u>Proof of lemma 1:</u>

For $k \ge 1$ (for $k=0$ it is obvious)

$$\langle \bar{\mathcal{Z}}_N, x^k \rangle = \mathbb{E}\langle \mathcal{Z}_N, x^k \rangle = \frac{1}{N}\mathbb{E}[\mathrm{tr}(X_N^k)]$$

$$\left(\text{notice: } (A\cdot B\cdot C)_{ij} = \sum_{m_1,m_2=1}^{N} A_{im_1} B_{m_1m_2} C_{m_2j} \right.$$
$$\left. A,B,C - N\times N \text{ matrices.} \right)$$

Have $\mathrm{Tr}(X_N^k) = \displaystyle\sum_{i_1,i_2,\dots,i_k=1}^{N} X_N(i_1,i_2)X_N(i_2,i_3)\cdots X_N(i_{k-1},i_k)X_N(i_k,i_1)$

Denoting: $i=(i_1,\dots,i_k)$ we write

$$T_i^N = X_N(i_1,i_2)\cdots X_N(i_k,i_1)$$

Thus, $\langle \bar{\mathcal{Z}}_N, x^k \rangle = \dfrac{1}{N}\displaystyle\sum_{i\in\{1,\dots,N\}^k} \mathbb{E}(T_i^N)$

Notation:

Given an alphabet A (in our case $A=\{1,\dots,N\}$),

a letter is an element of A.

a word ω is a finite sequence of letters.

a word ω is closed if first letter equals last letter

& ω_1 is equiv. to ω_2 (written $\omega_1 \sim \omega_2$) if there is a bijection from A to itself taking ω_1 to ω_2.

Example: 1424371 - closed, $supp(\omega)=\{1,2,3,4,7\}$

$\sim$ $wt(\omega) = 5$

2414372 $l(\omega) = 7$

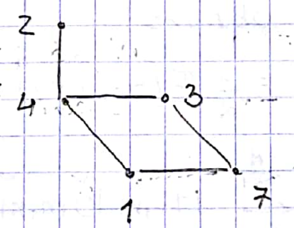
the support of ω, $supp(\omega)$, is the set of letters in ω.

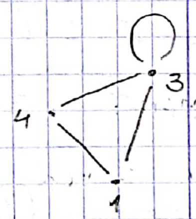
the weight of ω is $wt(\omega) = |supp(\omega)|$

the length of ω, denoted $l(\omega)$, is the length of ω as a sequence.

Associated to each word $\omega = \omega_1 \dots \omega_l$ is a graph:

$G_\omega(V_\omega, E_\omega)$ with $V_\omega = Supp(\omega)$, $E_\omega = \{\{\omega_i, \omega_{i+1}\} \mid 1 \le i \le l-1\}$

Examples: 1424371: 

, 1433431: 

for each edge e, N_e^ω is the number of times e was traversed in ω.

We also partition E_ω into E_ω^s, E_ω^c

self edges $\{v,v\}$ connecting edges $\{v,u\}$, $v \neq u$

With this notation, for any $i \in \{1,\dots,N\}^k$, $\omega_i = i_1 i_2 \dots i_k i_1$

$$\mathbb{E}(T_i^N) = \frac{1}{N^{k/2}} \prod_{e \in E^c_{\omega_i}} \mathbb{E}(Z_{1,2}^{N_e^{\omega_i}}) \prod_{e \in E^s_{\omega_i}} \mathbb{E}(Y_1^{N_e^{\omega_i}})$$

Note, if $\omega_{i_1} \sim \omega_{i_2}$ then $\mathbb{E}\, T_{i_1}^N = \mathbb{E}\, T_{i_2}^N$

Write $\omega_{k,t}$ for a set of representatives for equiv. classes of closed words of length $k+1$ & weight t.

Each equiv class of a word in $\omega_{k,t}$ has size:

$C_{N,t} = N(N-1)(N-2)\cdot\ldots\cdot(N-t+1)$

Thus, $\langle \bar{L}_N, x^k \rangle = \frac{1}{N} \sum'_{i \in \{1,\ldots,N\}^k} \mathbb{E}\, T_i^N = \sum_{t=1}^{k} \frac{C_{N,t}}{N^{k/2+1}} \sum'_{\omega \in \omega_{k,t}} \prod_{e \in E_\omega^c} \mathbb{E}(Z_{1,2}^{N_e^\omega}) \prod_{e \in E_\omega^s} \mathbb{E}(Y_1^{N_e^\omega})$

There are k steps in path of ω. To get non-zero expectation, have to pass each edge at least twice. So G_ω has at most $\lfloor \frac{k}{2} \rfloor$ edges.

A connected graph with e edges has $\leq e+1$ vertices, with equality iff graph is a tree.

So G_ω has $\leq \lfloor \frac{k}{2} \rfloor + 1$ vertices. That is $wt(\omega) \leq \lfloor \frac{k}{2} \rfloor + 1$.

Thus,

$$\langle \bar{L}_N, x^k \rangle = \sum_{t=1}^{\lfloor \frac{k}{2} \rfloor + 1} \frac{C_{N,t}}{N^{k/2+1}} \sum'_{\omega \in \omega_{k,t}} (\square) \leq \sum_{t=1}^{\lfloor \frac{k}{2} \rfloor + 1} \frac{C_{N,t}}{N^{\frac{k}{2}+1}} C_k'''$$

$(\square) \leq C_k'$

$|\omega_{k,t}| \leq C_k''$

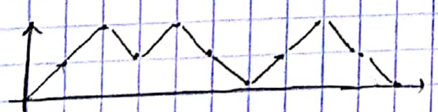
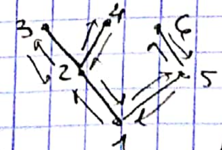
the only non-negligible contribution is from $t = \frac{k}{2} + 1$.

Conclude: $\langle \bar{L}_N, x^k \rangle \xrightarrow[N \to \infty]{} 0$, k odd

for k even, get $\langle \bar{L}_N, x^k \rangle \xrightarrow[N \to \infty]{} \sum'_{\omega \in \omega_{k,\frac{k}{2}+1}} (\square) = |\omega_{k,\frac{k}{2}+1}|$

the graph corresponding to $\omega \in \omega_{k,\frac{k}{2}+1}$ is necessarily a tree with edge traversed exactly twice in path of ω.

Next, $|\omega_{k,\frac{k}{2}+1}|$ equals the number of rooted trees where the children of each vertex are ordered.

 $\Leftrightarrow$  $\Leftrightarrow$ 12324215651

bijection

bijection

More formally, given ω of length $k+1$ & weight $\frac{k}{2}+1$, define a Dyck path of length k by $x_0 = 0$, $x_i =$ height above root in corresponding tree.

This is a bijection.

from this we conclude that $|\omega_{k,\frac{k}{2}+1}| = C_k$.

Altogether, $\langle L_N, x^k \rangle \longrightarrow C_k$, k even

∎

Proof of lemma 2:

Suffices that $Var(\langle L_N, x^k \rangle) \xrightarrow[N\to\infty]{} 0$ (from chebyshev)

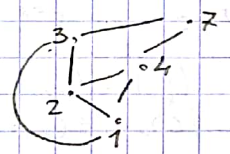
Recall, $\langle L_N, x^k \rangle = \frac{1}{N} \sum_{i \in \{1,\dots,N\}^k} T_i^N$

where $T_i^N = X_N(i_1,i_2) \cdot \ldots \cdot X_N(i_k,i_1)$

Hence, $Var(\langle L_N, x^k \rangle) = \frac{1}{N^2} \sum_{i,i' \in \{1,\dots,k\}^N} Cov(T_i^N, T_{i'}^N) =$

$$= \frac{1}{N^2} \cdot \sum_{i,i' \in \{1,\dots,N\}^k} E(T_i^N T_{i'}^N) - E(T_i^N) E(T_{i'}^N)$$

$G_{\omega_i} G_{\omega_{i'}}$ - the idea is to draw the two graphs on the same vertices:



$i = 123241$ $\qquad$ $i' = 14731$

Inorder to get Expectance $\neq 0$ we need Each edge traversed ≥ 2 in union of graphs, at least one edge in common to both graphs.