Fluctuations from the Semicircle Law
Lecture 2

Ioana Dumitriu
University of Washington

Women and Math, IAS 2014

May 21, 2014
Calculation of the Variance

Convergence in Probability and Almost Surely
We are examining \( \sum_{I,J \in \mathcal{I}} (\mathbb{E}(w_I w_J) - \mathbb{E}(w_I)\mathbb{E}(w_J)) \), where \( \mathcal{I} \) is the set of all \( k \)-tuples.

To examine this we need the graphs \( G_I, G_J, \) and their union, as well as the walks given by \( I \) and \( J \) on this graph.

Every edge appears \( \geq 2 \) times, \( v - 1 \leq e \leq k \).

We determined that terms with \( v = k + 1 \) have 0 contribution.

We want \( v \) as large as possible.
Why must $v$ be large?

\[
\sum_{I,J \in \mathcal{I}} \left( \mathbb{E}(w_I w_J) - \mathbb{E}(w_I) \mathbb{E}(w_J) \right) = \frac{1}{n^k} \sum_{G_I, G_J, G_I \cup G_J} n(n-1)(n-2) \ldots (n-v+1) \cdot Q(k, v, e),
\]

\[
= \frac{1}{n^k} \sum_{G_I, G_J, G_I \cup G_J} n^v (1 + o(1)) \cdot Q(k, v, e).
\]

The sum is like a polynomial in $n$, whose asymptotics is determined by the highest power. Thus, we want the coefficient of that highest power. Established yesterday that $v < k + 1$. 
Second attempt: \( v = k, k \text{ even} \)

Note that we must have some edge overlap between the graph induced by \( I \) and the one induced by \( J \). Otherwise term contributes 0 to covariance.

As \( v = k \), one of these must be true:

- \( e = k = v \); graph has a cycle or loop;
- \( e = k - 1 = v - 1 \); graph is a tree; one edge has multiplicity 4;
- \( e = k - 1 = v - 1 \); graph is a tree; two edges have multiplicity 3.
Second attempt: \( v = k, \ k \text{ even} \)

Note that we must have *some* edge overlap between the graph induced by \( I \) and the one induced by \( J \). Otherwise term contributes 0 to covariance.

As \( v = k \), one of these must be true:
- \( e = k = v \); graph has a cycle; or loop;
- \( e = k - 1 = v - 1 \); graph is a tree; one edge has multiplicity 4;
- \( e = k - 1 = v - 1 \); graph is a tree; two edges have multiplicity 3.
Second attempt: $v = k$, $k$ even

Note that we must have some edge overlap between the graph induced by $I$ and the one induced by $J$. Otherwise term contributes 0 to covariance.

As $v = k$, one of these must be true:

- $e = k = v$; graph has a cycle; or loop;
- $e = k - 1 = v - 1$; graph is a tree; one edge has multiplicity 4;
- $e = k - 1 = v - 1$; graph is a tree; two edges have multiplicity 3.
Count for $v = k$, $k$ is even, graph has a cycle

Let $3 \leq r \leq k$ be the length of the cycle. The shape of the graph is a cycle with dangling trees.

The trees divide in two subsets with same number of edges; half, i.e., $(k - r)/2$, belong to the graph induced by $I$, the others to the graph induced by $J$. 
Count for $v = k$, $k$ is even, graph has a cycle

Let $3 \leq r \leq k$ be the length of the cycle. The shape of the graph is a cycle with dangling trees.

No overlap on the trees! We can think of each vertex of the cycle as having its own pair $(T_{i,I}, T_{i,J})$ of trees, possibly empty, one from $I$, the other from $J$. Call their edge sizes $k_{i,I}$ and $k_{i,J}$. 
Count for $\nu = k$, $k$ is even, graph has a cycle

Must have

\[
\sum_{i=1}^{r} k_{i,I} = \sum_{i=1}^{r} k_{i,J} = \frac{k - r}{2}.
\]

For each pair of partitions of $(k - r)/2$, one has a total number of possible combinations of trees which is

\[
\prod_{i=1}^{r} C_{k_{i,I}} \prod_{i=1}^{r} C_{k_{i,J}}.
\]
Count for $v = k$, $k$ is even, graph has a cycle

Assume that $k$ vertices have been chosen in
\[(n)_k = n(n - 1) \ldots (n - k + 1) = n^k(1 + o(1))\] ways.

Given a cycle-and-dangling-trees graph, properly partitioned, there are
- a total number $k^2$ of ways of choosing starting vertices;
- 2 choices on direction (same or opposite);
- a rotational invariance overcount (must divide by $r$).
Count for \( v = k \), \( k \) is even, graph has a cycle

Summing over all possible cycle lengths \( r \):

\[
n^k(1 + o(1)) \sum_{r=3}^{k} \frac{2k^2}{r} \left( \sum_{\kappa \vdash (k-r)/2; \text{length}(\kappa) = r} \prod_{i=1}^{r} C_i \right)^2.
\]
Count for $v = k$, $k$ is even, graph has a cycle

Sharing the cycle implies that both $\mathbb{E}(w_I) = \mathbb{E}(w_J) = 0$, and since each edge appears twice, there are $k$ of them, and $\mathbb{E}(w_{ij}^2) = 1/n$, the contribution is asymptotically

$$\sum_{r=3}^{k} \frac{2k^2}{r} \left( \sum_{\kappa \vdash (k - r)/2 \atop \text{length}(\kappa) = r} \prod_{i=1}^{r} C_i \right)^2.$$
Count for $v = k$, $k$ is even, graph is a tree

Recall that one edge is overlapped and appears 4 times.

Given a set of $k$ labels, have $C_{k/2}$ possible trees and for each pair, $(k/2)^2$ choices of pair of overlapped edges, and 2 choices of orientation. Total choices:

$$n^k (1 + o(1)) 2(k/2)^2 C_{k/2}^2 ;$$

each choice comes with weight

$$\left( \mathbb{E}(Z^4) - 1 \right) \left(1/n \right)^k ,$$

the contribution is, asymptotically,

$$2(k/2)^2 C_{k/2}^2 \left( \mathbb{E}(Z^4) - 1 \right) .$$
Asymptotical variance for $k$ even

Putting it all together for $k$ even:

$$\sigma_k^2 := \text{Var} \left( W_n^k \right) = \frac{k^2}{2} C_{k/2}^2 \left( \mathbb{E}(Z^4) - 1 \right) + \sum_{r=3}^{k} \frac{2k^2}{r} \left( \sum_{\kappa \vdash (k-r)/2 \atop \text{length}(\kappa) = r} \prod_{i=1}^{r} C_i \right)^2.$$
Second attempt: \( v = k, \, k \text{ odd} \)

As \( v = k \), one of these must be true:

- \( e = k = v \); graph has a cycle or loop.
- \( e = k - 1 = v - 1 \); graph is a tree; one edge has multiplicity 4;
- \( e = k - 1 = v - 1 \); graph is a tree; two edges have multiplicities 3.
Second attempt: \( v = k, \, k \text{ odd} \)

As \( v = k \), one of these must be true:

- \( e = k = v \); graph has a cycle or loop.
- \( e = k - 1 = v - 1 \); graph is a tree; one edge has multiplicity 4;
- \( e = k - 1 = v - 1 \); graph is a tree; two edges have multiplicities 3.
Second attempt: $v = k, k$ odd

As $v = k$, one of these must be true:

- $e = k = v$; graph has a cycle or loop.
- $e = k - 1 = v - 1$; graph is a tree; one edge has multiplicity 4;
- $e = k - 1 = v - 1$; graph is a tree; two edges have multiplicities 3.
Count for $\nu = k$, $k$ odd, graph has a cycle

Exactly as before; total contribution is

$$\sum_{r=3}^{k} \frac{2k^2}{r} \left( \sum_{\kappa \vdash (k-r)/2} \prod_{i=1}^{r} C_i \right)^2.$$
Count for $v = k$, $k$ odd, graph has a loop

The loop must be the overlap. There are $k$ choices where to attach it to a closed walk on a tree with $(k - 1)/2$ vertices (at each step). Two trees, so total number of choices is $k^2 \binom{(k-1)/2}{2}$.

The existence of the loop means that $\mathbb{E}(w_I) = \mathbb{E}(w_J) = 0$.

There are $n^k(1 + o(1))$ choices for vertices, and each edge appears twice for a weight of $(1/n)^k$. 
Asymptotical variance for \( k \) odd

Putting it all together for \( k \) odd:

\[
\sigma^2_k := \text{Var} \left( W_n^k \right) = k^2 C^2_{(k-1)/2} + \sum_{r=3}^{k} \frac{2k^2}{r} \left( \sum_{\kappa \vdash (k-r)/2} \prod_{i=1}^{r} C_i \right)^2.
\]
We have shown that, for any \( k \geq 0 \),

\[
X_{n,k} = \text{tr}(W_n^k) - \mathbb{E}(\text{tr}(W_n^k))
\]

has, asymptotically, finite variance. It follows then that

\[
\frac{1}{n} \text{tr}(W_n^k) - \frac{1}{n} \mathbb{E}(\text{tr}(W_n^k)) \to 0,
\]

in other words

\[
\frac{1}{n} \text{tr}(W_n^k) \to \delta_{k \text{ even}} C_{k/2}.
\]
This means that the average distribution of an eigenvalue converges not just in distribution, but in *probability*.

In fact, we can get more: the variance of $\frac{1}{n} \text{tr}(W_n^k)$ is proportional to $\frac{1}{n^2}$.

Assume we have a sequence of matrices $W_n$, for $n \geq 1$, on the same probability space (e.g., a doubly infinite, symmetric array of iid variables.)
Almost Sure Convergence

The Chebyshev Inequality together with the Borel-Cantelli Lemma say that, for any $\epsilon > 0$ and $k \geq 1$, the number of matrices for which

$$\left| \frac{1}{n} \text{tr}(W_n^k) - \delta_{k \text{ even}} C_k/2 \right| > \epsilon$$

is finite.

This means that the convergence to the semicircle happens \textit{almost surely}, and explains the experiment.