Lecture 10: April 22, 2011

So, now we know that if μ, ν are any two compactly-supported probability measures on $\mathbb{R}$, $\exists$ NCPS $(\mathcal{A}, \varphi)$ with self-adjoint elements $x, y \in \mathcal{A}$ with $\mu_x = \mu$, $\mu_y = \nu$, and x, y are free. Thus, we can generally define

$$\mu \boxplus \nu \equiv \mu_{x+y} \leftarrow \text{depends only on the distributions.}$$

But how is this "free convolution" determined? To begin telling this story, it is actually worthwhile to look at the bigger question of an infinite sum.

Let's go back to classical probability one more time for motivation.

Limit Laws in Probability

Let $X_1, X_2, X_3, \ldots$ be a sequence of (classical) random variables. The (simplest forms of the) two fundamental limit theorems are:

The Law of Large Numbers: If $\{X_n\}$ are iid (independent, and identically distributed – i.e. $\mu_{X_n} = \mu_{X_1}$ $\forall n$) and $\mathbb{E}(X_1) = m$, then

$$\frac{X_1 + \cdots + X_n}{n} \longrightarrow m \quad \text{a.s. as } n \to \infty.$$

The Central Limit Theorem: If $\{X_n\}$ are iid with $\mathbb{E}(X_1) = m$ and

$$\operatorname{Var} X_1 \equiv \mathbb{E}(X_1^2) - \mathbb{E}(X_1)^2 = \sigma^2$$

then

$$\frac{X_1 + \cdots + X_n}{\sqrt{n}} \Longrightarrow \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-m)^2/2\sigma^2} \, dx \quad \text{as } n \to \infty.$$

$\uparrow$ convergence in distribution.

Remarks: (1) The CLT is a refinement of the LLN; it says that the exact rate of convergence in the LLN is $\asymp n^{-1/2}$, and that at this exact scale a universal distribution – the Gaussian – emerges for all sequences.

(2) By translation and scaling, it is sufficient to state and prove both theorems in the special case $\mathbb{E}(X_1) = 0$, and the assumption $\operatorname{Var} X_1 = 1$ in the CLT.

Now, convergence in distribution implies convergence of moments. A simple calculus exercise shows that

$$\int_{\mathbb{R}} x^k \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\,dx = \begin{cases} (k-1)!! & ; \ k \text{ even} \\ 0 & ; \ k \text{ odd} \end{cases}$$

(Here $(k-1)!! = (k-1)(k-3)\cdots(3)(1)$.) Thus, a consequence of the CLT is the following a priori weaker statement.

<u>Theorem</u>: If $\{X_n\}$ are iid., $\mathbb{E}(X_1)=0$, $\operatorname{Var} X_1 = \mathbb{E}(X_1^2)=1$, and $|\mathbb{E}(X_1^k)|<\infty$ $\forall k\geq 0$, then

$$\mathbb{E}\left[\left(\frac{X_1+\cdots+X_n}{\sqrt{n}}\right)^k\right] \longrightarrow \begin{cases} (k-1)!! , & k \text{ even} \\ 0 , & k \text{ odd} \end{cases}$$

<u>Remark</u>: Convergence of moments is generally weaker than convergence in distribution. However, as we will see a little later, if all the measures are supported in a single compact interval, the two notions are equivalent; so for our purposes later, this kind of convergence is relevant.

<u>Pf</u>. We begin by expanding the k^{th} moment. We could write this in terms of multinomial coefficients, but it is actually more convenient to ignore the fact that these classical random variables commute.

$$\mathbb{E}\left[\left(\frac{X_1+\cdots+X_n}{\sqrt{n}}\right)^k\right] = n^{-k/2} \sum_{\substack{i_1,\ldots,i_k \\ \in\{1,\ldots,n\}}} \mathbb{E}(X_{i_1}X_{i_2}\cdots X_{i_k})$$

Now, looking at the terms in this sum, the iid nature of the variables means that many of these terms are the same. For example:

$\mathbb{E}(X_1X_2X_1X_3X_2X_3X_3)$ is a joint moment in X_1, X_2, X_3.
– Since these are independent, it is $\therefore$ determined by $\{\mathbb{E}(X_1^k)\}_{k=1}^\infty, \{\mathbb{E}(X_2^k)\}_{k=1}^\infty, \{\mathbb{E}(X_3^k)\}_{k=1}^\infty$

$$\{\mathbb{E}(X_1^k)\} = \mathbb{E}(X_3^k), \quad \{\mathbb{E}(X_2^k)\} = \mathbb{E}(X_7^k), \quad \{\mathbb{E}(X_3^k)\} = \mathbb{E}(X_5^k) \quad \text{by i.d.}$$

$\therefore$ $\mathbb{E}(X_1X_2X_1X_3X_2X_3X_3) = \mathbb{E}(X_3X_7X_3X_5X_7X_5X_5)$

[The salient point is: the algorithm by which individual moments determine joint moments is uniform — it does not depend on the particular random variables at hand.]

This suggests a relabeling of the sum according to how the indices are partitioned.

Def: Let $[k] = \{1, 2, \ldots, k\}$. A (set) partition π of $[k]$ is an equivalence relation on $[k]$: $\pi = \{B_1, B_2, \ldots, B_r\}$ where $B_j \subseteq [k]$, $B_i \cap B_j = \emptyset$ if $i \neq j$, $B_j \neq \emptyset$, and $B_1 \cup \cdots \cup B_r = [k]$. The B_j are called blocks.

The set of all partitions of $[k]$ is denoted $\mathcal{P}(k)$.

The observation about joint moments above can be expressed as follows. To each k-tuple $(i_1, \ldots, i_k)$, we assign a partition $\pi \in \mathcal{P}(k)$:

$a \sim_\pi b$ iff $i_a = i_b$.

1 2 1 3 2 3 3

← partition diagram

Express this relationship as $(i_1, \ldots, i_k) \mapsto \pi$. Then if there is a single partition π s.t. $(i_1, \ldots, i_k) \mapsto \pi$ & $(j_1, \ldots, j_k) \mapsto \pi$

then $\mathbb{E}(X_{i_1} \cdots X_{i_k}) = \mathbb{E}(X_{j_1} \cdots X_{j_k})$

So, for each $\pi \in \mathcal{P}(k)$, among all k-tuples of indices with partition π, there is a single value of the joint moment, determined therefore only by π

$$\{\mathbb{E}(X_{i_1} \cdots X_{i_k}) : (i_1, \ldots, i_k) \mapsto \pi\} = \{c(\pi)\}$$

a single number; i.e. $c : \mathcal{P}(k) \to \mathbb{C}$ is a function determined by the common distribution of the X_i.

Thus, we can reindex the sum as follows:

$$\mathbb{E}\left[\left(\frac{X_1 + \cdots + X_n}{\sqrt{n}}\right)^k\right] = n^{-k/2} \sum_{i_1, \ldots, i_k \in [n]} \mathbb{E}(X_{i_1} \cdots X_{i_k})$$

$$= n^{-k/2} \sum_{\pi \in \mathcal{P}(k)} \sum_{\substack{(i_1, \ldots, i_k) \in [n] \\ \mapsto \pi}} c(\pi) = n^{-k/2} \sum_{\pi \in \mathcal{P}(k)} c(\pi) \, \#\{(i_1, \ldots, i_k) \mapsto \pi\}$$

does not depend on the indices

It is easy to count $\#\{(i_1,\dots,i_k)\in[n]^k \mapsto \pi\}$ for a fixed π. Suppose $\pi=\{B_1,\dots,B_r\}$. We must select an element of $[n]$ to be the label of all indices in B_1 – there are n such choices. Then we must select a <u>different</u> element of $[n]$ to be the label of all the indices in B_2 – there are $n-1$ such choices. Continuing, we see that

$$\#\{(i_1,\dots,i_k)\mapsto\pi\} = n(n-1)\cdots(n-r+1).$$

<u>Notation</u>: for $\pi\in\mathcal{P}(k)$, $|\pi| = \#\text{blocks in } \pi$.

Thus, our expansion for the k^{th} moment of the normalized sum is

$$\boxed{\mathbb{E}\left[\left(\frac{X_1+\cdots+X_n}{\sqrt{n}}\right)^k\right] = n^{-k/2}\sum_{\pi\in\mathcal{P}(k)} n(n-1)\cdots(n-|\pi|+1)\, c(\pi).}$$

<u>Remark</u>: to get here, we did not assume the X_i's commute, nor did we use the specific form of independence. All we required was that all the X_i have the same moments, and that joint moments are (uniformly) determined by individual ones. (The function c is determined by the common distribution of the X_i, and also by the form of the independence rule in question.) In particular, had we assumed the X_i were <u>free</u>, we would arrive at the same formula (with a different function c).

Now, to evaluate the limit as $n\to\infty$, the following lemma is crucial.

<u>Lemma</u>: Let $\pi\in\mathcal{P}(k)$, and suppose $\exists$ block in π containing only one element. Then $c(\pi)=0$.

Pf. Let $(i_1,\dots,i_k)\mapsto\pi$. Since π has a block with only one element, say s, this means $i_s\notin\{i_1,\dots,i_{s-1},i_{s+1},\dots,i_k\}$. By independence, this means that

$$c(\pi)=\mathbb{E}(X_{i_1}\cdots X_{i_s}\cdots X_{i_k}) = \underbrace{\mathbb{E}(X_{i_s})}_{0}\,\mathbb{E}(X_{i_1}\cdots X_{i_{s-1}}X_{i_{s+1}}\cdots X_{i_k}).$$

///

<u>Notation</u>: $\mathcal{P}_{\geq 2}(k)=\{\pi\in\mathcal{P}(k);\ \text{each block of } \pi \text{ has} \geq 2 \text{ elements}\}$.

$\mathcal{P}_2(k)=\{\pi\in\mathcal{P}(k);\ \text{each block of } \pi \text{ has exactly } 2 \text{ elements}\}$.

Thus, from the lemma, the only terms that contribute are $\pi \in \mathcal{P}_{\geq 2}(k)$. Note: if $\pi \in \mathcal{P}_{\geq 2}(k)$, $|\pi| \leq k/2$; and among these, $|\pi| = k/2$ iff $\pi \in \mathcal{P}_2(k)$.

$$\therefore \mathbb{E}\left[\left(\frac{X_1 + \cdots + X_n}{\sqrt{n}}\right)^k\right] = n^{-k/2} \sum_{\pi \in \mathcal{P}_{\geq 2}(k)} n(n-1)\cdots(n-|\pi|+1)\, c(\pi).$$

Now, we are interested in the limiting moments. The sum is independent of n, so we can compute

$$\lim_{n\to\infty} \mathbb{E}\left[\left(\frac{X_1 + \cdots + X_n}{\sqrt{n}}\right)^k\right] = \sum_{\pi \in \mathcal{P}_{\geq 2}(k)} \lim_{n\to\infty}\left[n^{-k/2} \cdot n(n-1)\cdots(n-|\pi|+1)\right] c(\pi)$$

$n(n-1)\cdots(n-|\pi|+1) \sim n^{|\pi|}$ as $n \to \infty$

So we are taking

$$\lim_{n\to\infty} n^{|\pi| - k/2} = \begin{cases} 1, & |\pi| = k/2 \\ 0, & |\pi| > k/2 \end{cases}$$

where $|\pi| \leq k/2$

$$\therefore = \sum_{\pi \in \mathcal{P}_2(k)} c(\pi).$$

Now, we must calculate $c(\pi)$ for $\pi \in \mathcal{P}_2(k)$. Note, if k is odd, then $\mathcal{P}_2(k) = \phi$; so we have proved that the limiting odd moments are 0. Now, suppose k is even. If $\pi \in \mathcal{P}_2(k)$ and $(i_1, \ldots, i_k) \mapsto \pi$, this means the indices come in $k/2$ pairs. So, $\exists!\, s > 1$ s.t. $i_1 = i_s$, and so

$$c(\pi) = \mathbb{E}(X_{i_1} X_{i_2} \cdots X_{i_s} \cdots X_{i_k}) = \underset{\substack{\| \\ 1}}{\mathbb{E}(X_{i_1}^2)}\, \underbrace{\mathbb{E}(X_{i_2} \cdots X_{i_{s-1}} X_{i_{s+1}} \cdots X_{i_k})}_{\text{Continue by induction.}}$$

$$= 1^{k/2} = 1.$$

Thus, for $\pi \in \mathcal{P}_2(k)$, $c(\pi) = 1$. So

$$\lim_{n\to\infty} \mathbb{E}\left[\left(\frac{X_1 + \cdots + X_n}{\sqrt{n}}\right)^k\right] = \sum_{\pi \in \mathcal{P}_2(k)} c(\pi) = \#\mathcal{P}_2(k) = (k-1)!!$$

(For the block containing 1, choose 1 of the remaining $k-1$ el'ts of $[k]$; now there are $k-2$ remaining el'ts so induction gives $(k-1)(k-3)\cdots(3)(1)$.) ///

Not only did we prove that the limiting moments of the central limit theorem are those of a standard normal Gaussian, we showed (directly) they count a certain combinatorial object:

$$k^{th} \text{ moment } = \# \mathcal{P}_2(k).$$

Now, let's see what happens if we change "independence" to "freeness".

Let $(\mathcal{A},\varphi)$ be a NCPS, and let $x_1, x_2, x_3, \ldots$ be a sequence of f.i.d. self-adjoint elements (free and identically distributed) with $\varphi(x_i)=0$ and $\varphi(x_i^2)=1$. (Such sequences exist, for example in infinite free product algebras.) What can we say about

$$\lim_{n\to\infty} \varphi\left[\left(\frac{x_1+\cdots+x_n}{\sqrt{n}}\right)^k\right] \quad ?$$

By the preceding discussion, there is a function (different from before, since freeness $\neq$ independence) $K:\mathcal{P}(k)\to\mathbb{C}$ such that

$$\varphi\left[\left(\frac{x_1+\cdots+x_n}{\sqrt{n}}\right)^k\right] = n^{-k/2}\sum_{\pi\in\mathcal{P}(k)} n(n-1)\cdots(n-|\pi|+1)\,K(\pi).$$

The function K is determined by the common distribution of the x_i's. We are going to spend the next 2-3 weeks thoroughly understanding the function K, via the combinatorics of the lattice $\mathcal{P}(k)$.

<u>Lemma</u>: If π has a singleton block, then $K(\pi)=0$.

<u>Pf</u>. If $(i_1,\ldots i_k)\mapsto\pi$, this means there $\exists s$ s.t. $i_s\notin\{i_1,\ldots,i_{s-1},i_{s+1},\ldots i_k\}$. Now, let $\mathcal{A}_s=\langle x_{i_s}\rangle$ and $\mathcal{A}_{\neq s}=\langle x_{i_1},\ldots,x_{i_{s-1}},x_{i_{s+1}},\ldots,x_{i_k}\rangle$. By assumption (and associativity of freeness – see exercises from Lecture 3), $\mathcal{A}_s,\mathcal{A}_{\neq s}$ are free. Let $a_1=x_{i_1}\cdots x_{i_{s-1}}$ and $a_2=x_{i_{s+1}}\cdots x_{i_k}$. Then

$$K(\pi)=\varphi(x_{i_1}\cdots x_{i_k}) = \varphi(a_1 x_{i_s} a_2) \underset{\substack{\text{Exercise from}\\ \text{Lecture 3}}}{=} \underbrace{\varphi(x_{i_s})}_{0}\varphi(a_1a_2).$$

///

As above, it therefore follows that $K(\pi)=0$ unless $\pi\in\mathcal{P}_{\geq 2}(k)$. Then, the very same asymptotic analysis as above shows that

$$\lim_{n\to\infty}\varphi\left[\left(\frac{x_1+\cdots+x_n}{\sqrt{n}}\right)^k\right]=\sum_{\pi\in\mathcal{P}_2(k)}K(\pi).$$

Since $\mathcal{P}_2(k)=\phi$ when k is odd, once again we have that the odd limiting moments are all 0. For the even moments, we must calculate K on $\mathcal{P}_2(k)$. Because freeness is involved, we must consider two cases.

- Suppose $(i_1,\ldots,i_k)\mapsto\pi$ where $i_1\neq i_2\neq\cdots\neq i_k$. Since the x_i are free and centered, it follows (by definition) that

$$K(\pi)=\varphi(x_{i_1}\cdots x_{i_k})=0.$$

- The other possibility is that there is $s\in[k-1]$ s.t. $i_s=i_{s+1}$. Since $\pi\in\mathcal{P}_2(k)$, there are no other equal indices, and so

$$i_s=i_{s+1}\notin\{i_1,\ldots,i_{s-1},i_{s+2},\ldots,i_k\}$$

So let $a_1=x_{i_1}\cdots x_{i_{s-1}}$ and $a_2=x_{i_{s+2}}\cdots x_{i_k}$; then x_{i_s} is free from $\{a_1,a_2\}$. Therefore

$$K(\pi)=\varphi(x_{i_1}\cdots x_{i_k})=\varphi(a_1x_{i_s}x_{i_{s+1}}a_2)=\varphi(a_1x_{i_s}^2a_2)=\varphi(x_{i_s}^2)\varphi(a_1a_2).$$

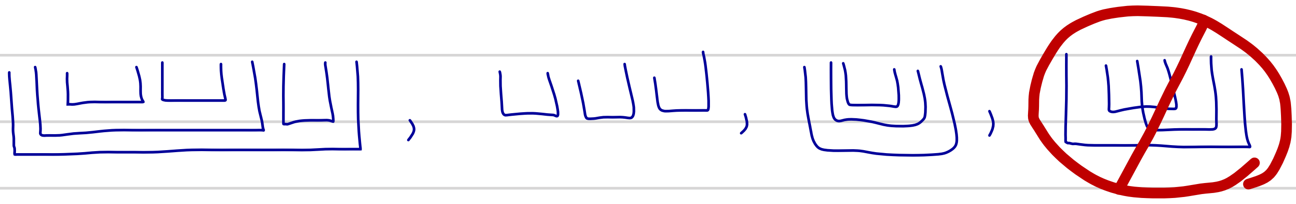
The normalization assumption was $\varphi(x_i^2)=1$, so in this case,

$$K(\pi)=\varphi(a_1a_2)=\varphi(\underbrace{x_{i_1}\cdots x_{i_{s-1}}x_{i_{s+2}}\cdots x_{i_k}}).$$

Now we are set up for an induction. This word is determined by a partition π' of its indices $\{1,\ldots,s-1,s+2,\ldots,k\}\cong[k-2]$. By the same argument we just made, we get 0 unless π' has an adjacent pairing, in which case $K(\pi)=\varphi(\text{a word of length }k-4)$. Continuing, we see

$$K(\pi)=\begin{cases}1, & \pi\in N_2(k)\\ 0, & \pi\notin N_2(k)\end{cases}$$

$\longleftarrow$ $=\{\pi\in\mathcal{P}_2(k)$ that have the following recursive property: removing any adjacent pair produces a new pairing with an adjacent pair.$\}$

Eg's 

Def: A partition $\pi \in \mathcal{P}(k)$ has a crossing if there are two blocks $B, B' \in \pi$ and elements $i, j \in B$, $i', j' \in B'$ with $i < i' < j < j'$.

Exercise: Show that $\pi \in N_2(k)$ iff $\pi \in \mathcal{P}_2(k)$ and π has no crossings.

Def: The set of partitions in $\mathcal{P}(k)$ with no crossings is denoted $NC(k)$ – non-crossing partitions. The non-crossing pairings are

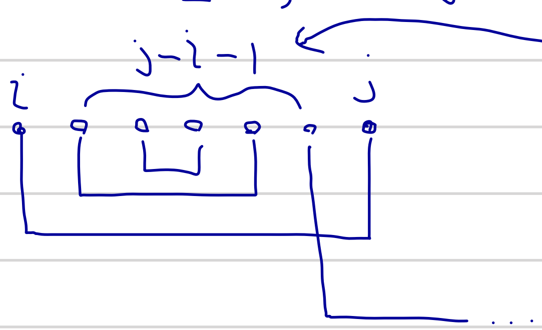
$$NC_2(k) \equiv NC(k) \cap \mathcal{P}_2(k).$$

Thus, $N_2(k) = NC_2(k)$, and we have

$$\lim_{n\to\infty} \varphi\left[\left(\frac{x_1 + \cdots + x_n}{\sqrt{n}}\right)^k\right] = \sum_{\pi \in \mathcal{P}_2(k)} K(\pi) = \# NC_2(k).$$

So, we must enumerate the non-crossing pairings. This should look familiar.

Observation: a pairing in $NC_2(k)$ is parity-reversing: if $\{i,j\} \in \pi \in NC_2(k)$, then $j-i$ is odd. For if not:



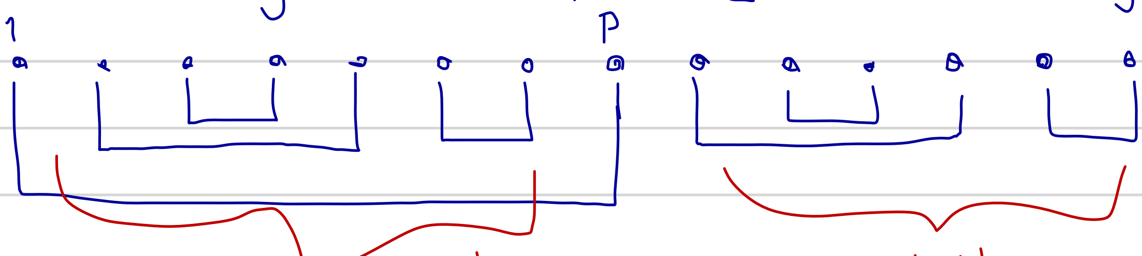
odd if i,j have the same parity; thus at least one must pair outside, producing a crossing.

In particular, the block containing 1 in π must have an even element: $\{1, 2p\} \in \pi$ for $2p \le k$. Say $k = 2m$. Define

$$NC_2^{(p)}(2m) = \{\pi \in NC_2(2m) \,;\, \{1,2p\} \in \pi\}.$$

Thus $NC_2(2m) = \bigsqcup_{p=1}^{m} NC_2^{(p)}(2m).$

Now, any element of $NC_2^{(p)}(2m)$ naturally decomposes:



inside is an arbitrary element of $NC_2(2(p-1))$ an arbitrary element of $NC_2(2(m-p))$.

Thus $\#NC_2^{(p)}(2m) = \#NC_2(2(p-1)) \cdot \#NC_2(2(m-p))$.

Let $N_m = \#NC_2(2m)$. Thus, we have proved

$$N_m = \sum_{p=1}^{m} N_{p-1} N_{m-p} \leftarrow \text{the Catalan recurrence!}$$

Since $NC_2(2) = \{\sqcup\}$, $N_1 = 1$, so indeed $N_m = C_m$.

$$\text{I.e. } \lim_{n\to\infty} \varphi\left[\left(\frac{x_1+\cdots+x_n}{\sqrt{n}}\right)^k\right] = \#NC_2(k) = \begin{cases} C_{k/2}, & k \text{ even} \\ 0, & k \text{ odd} \end{cases}.$$

Ie.

<u>Theorem</u>: If $\{x_n\}_{n=1}^{\infty}$ are self-adjoint in a NCPS $(\mathcal{A},\varphi)$, and are f.i.d. with $\varphi(x_i)=0$ and $\varphi(x_i^2)=1$, then

$$\frac{x_1+\cdots+x_n}{\sqrt{n}} \underset{\substack{\uparrow \\ \text{convergence of all moments}}}{\Longrightarrow} \text{Wigner's Semicircle Law}$$

This is the Free CLT: the "normal" distribution is the semicircle law.

<u>Remark</u>: The laws of $\frac{x_1+\cdots+x_n}{\sqrt{n}}$ and the semicircle law are all compactly-supported. If we could prove they are all supported in a single compact interval, the Weierstrass approximation theorem would imply convergence in distribution. This actually <u>is</u> true, but we need a bunch more machinery to prove it.

<u>Exercise</u>: Find an explicit bijection from Dyck paths of length k to $NC_2(k)$. [Hint: up means "open", down means "close".]