Noncommutative probability theory

Independence and Central Limit Theorems

Juan Pablo Vigneaux Ariztía

Leipzig, September 15, 2020

Outline

Algebraic preliminaries

A $*$-ring is a ring R with an anti-automorphism $\iota : r \mapsto r^*$ (i.e. linear, $(rs)^* = s^* r^*$) that is an involution (i.e. $(r^*)^* = r$).

A $*$-algebra A is a $*$-ring, with involution $*$, that is an associative algebra over a commutative $*$-ring R, with involution $'$, and such that $(ra)^* = r' a^*$ for all $r \in R$ and $a \in A$.

A $*$-homomorphism $f : A \to B$ between $*$-algebras is an algebra homomorphism compatible with the involution i.e. $f(a^*) = f(a)^*$ for all $a \in A$.

Examples: any *commutative* ring with $\iota =$ identity is a $*$-ring (trivial involution); the complex numbers with ι=conjugation are a $*$-ring, and also a $*$-algebra over the reals (with trivial involution).

Noncommutative probability spaces

In classical probability theory, complex-valued random variables are measurable functions on a probability space $(\Omega, \mathscr{F}, \mathbb{P})$, they form an algebra over $\mathbb{C}$. The expectation $\mathbb{E}(f) = \int_\Omega f \, \mathrm{d}\mathbb{P}$ defines a linear functional on

Definition

A *noncommutative probability space* is a pair (A, Φ) where $\mathscr{A}$ is a unital $*$-algebra (over $\mathbb{C}$) and Φ is a *state* on A: a $*$-homomorphism $\Phi : A \to \mathbb{C}$

1. Φ is unital: $\Phi(1) = 1$.
2. Φ is positive: for all $a \in A$, $\Phi(aa^*) \geq 0$.

The state is *tracial* if $\Phi(ab) = \Phi(ba)$ for all $a, b \in A$, and *faithful* if $\Phi(aa^*) = 0$ iff $a = 0$.

Examples

1. **Classical probability:** Take $A = \bigcap_{1 \le p < \infty} L^p(\Omega, \mathscr{F}, \mathbb{P}, \mathbb{C})$, $\Phi = \mathbb{E}$, and $*$ as conjugation.
2. **Matrices:** Set $A = M_n(\mathbb{C})$ and $\Phi = \frac{1}{n}\text{trace}$, with $*$ being the conjugate transpose.
3. **Random matrices:** One can mix the previous examples, considering the algebra A of $M_n(\mathbb{C})$-valued random variables, with state $\Phi = \frac{1}{n}\mathbb{E}\text{trace}$ and involution $=$ conjugate transpose.
4. **Quantum mechanics:** In this case, the algebra A is a C^*-algebra i.e. a $\star$-algebra that is also a Banach space, satisfying the identity $||xx^*|| = ||x||^2$. By the GNS construction, A can be represented as a norm-complete $*$-subalgebra of the space of bounded operators on a Hilbert space H. The functional Φ is then given by a unit vector $h \in H$, as $\Phi(a) = \langle ah, h \rangle$ (pure state). The self-adjoint elements are interpreted as physical observables (operators with real spectrum).

Random variables

Definition

Let (A, Φ) be a noncommutative probability space. A *(noncommutative) random variable* is a $*$-homomorphism $j\colon B \to \mathscr{A}$ where B is a (unital) $*$-algebra. The state $\varphi_j := \Phi \circ j$ on B is called *distribution* of j.

The "elementary" random variables are generated by a single element $a \in A$ i.e. the algebra generated over $\mathbb{C}$ by 1 and a. The distribution of a is then a linear map $\mu_a : \mathbb{C}[X] \to \mathbb{C}$ such that $\mu_a(P) = \Phi(P(a))$. The information in μ_a is the collection of moments $(\Phi(a^n))_{n \in \mathbb{N}}$.

Similarly, the distribution of a family of random variables $(a_i)_{i \in I} \subset A$ is a map $\mu_a : \mathbb{C}\langle X_1, ..., X_n \rangle \to \mathbb{C}$ such that $\mu_a(P) = \Phi(P(a_1, .., a_n))$, and it is equivalent to the collection of mixed moments $\Phi(a_{i_1} \cdots a_{i_p})$.

If a is selfadjoint element in a C^*-proba. space, μ_a can be identified with a compactly supported probability measure on $\mathbb{R}$ (comes from the spectral theorem).

Outline

Independence: Basic ideas

Loosely speaking, two random variables $A \to C$ and $B \to C$ are independent if their joint distribution (i.e. of the $*$-subalgebra generated by $A \cup B$) is the *product* of their marginals. We will see that this idea has several realizations in the noncommutative world i.e. there are several ways in which the mixed moments $\Phi(a_1^{k_1} b_1^{l_1} \ldots a_n^{k_n} b_n^{l_n})$ can depend of the moments of the a_i and b_j.

Categorical remarks

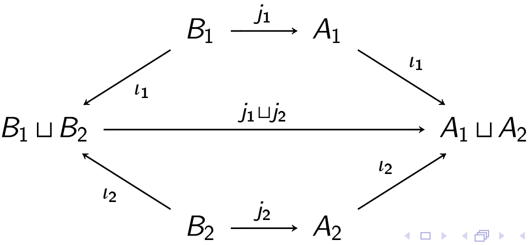
Let A_i be (non necessarily unital) associative algebra over $\mathbb{C}$.

Their coproduct in the category of algebras is the free product

$$A_1 \sqcup A_2 \equiv A_1 * A_2 = \bigoplus_{\varepsilon \in \mathbf{S}} A_{\varepsilon_1} \otimes \cdots A_{\varepsilon_n},$$

where $\mathbf{S}$ is the set of all finite sequences $\varepsilon = (\varepsilon_1, \cdots \varepsilon_n) \in \{1,2\}^n$ such that two consecutive terms differ.

For any pair of algebra homomorphisms $j_1 : B_1 \to A_1$ and $j_2 : B_2 \to A_2$, there exists a unique morphism $j_1 \sqcup j_2$ such that the following diagram commutes



Natural products

A natural product in the category of $*$-algebras equipped with linear functionals is a map $((A_1,\phi_1),(A_2,\phi_2)) \mapsto (A_1 \sqcup A_2, \phi_1 \Box \phi_2)$ that satisfies:

1. Associativity: Under the natural identification $(A_1 \sqcup A_2) \sqcup A_3 \cong A_1 \sqcup (A_2 \sqcup A_3)$, one has $(\phi_1 \Box \phi_2) \Box \phi_3 = \phi_1 \Box (\phi_2 \Box \phi_3)$.
2. Universality: for any $j_i : B_i \to A_i$, one has $(\phi_1 \circ j_1) \Box (\phi_2 \circ j_2) = (\phi_1 \Box \phi_2) \circ (j_1 \sqcup j_2)$.
3. Normalization: $(\phi_1 \Box \phi_2) \circ j_i = \phi_i$, for $i = 1,2$, and

$$(\phi_1 \Box \phi_2)(\iota_1(a_1)\iota_2(a_2)) = (\phi_1 \Box \phi_2)(\iota_2(a_2)\iota_1(a_1)) = \phi_1(a_1)\phi_2(a_2)$$

for any $a_1 \in A_1$ and $a_2 \in A_2$.

Theorem (Ben Ghorbal and Schürmann, Speicher, Muraki cf. [4])

There exist only five natural products: the tensor product $\otimes$, the free product $\star$, the boolean product $\diamond$, the monotone product $\triangleright$ and the anti-monotone product $\triangleleft$.

The natural products

Conventions: Given $(\varepsilon_1, ..., \varepsilon_n)$ in $\mathbf{S}$, $n \geq 2$, introduce $V_1 := V_1(\varepsilon) := \{i \,|\, \varepsilon_i = 1\}$, and similarly V_2. Suppose given $a_1, .., a_n$ in $A_1 \sqcup A_2$ such that for each $k \in [n]$, we have $a_k = i_{\varepsilon_k}(a_k^{(\varepsilon_k)})$ for some $a_k^{(\varepsilon_k)} \in A_{\varepsilon_k}$.

1. **Tensor:**

$$(\phi_1 \otimes \phi_2)(a_1 \cdots a_n) = \phi_1(\overrightarrow{\prod}_{k \in V_1} a_k^{(1)})\phi_2(\overrightarrow{\prod}_{l \in V_2} a_l^{(2)}).$$

2. **Boolean:**

$$(\phi_1 \diamond \phi_2)(a_1 \cdots a_n) = \prod_{k \in V_1} \phi_1(a_k^{(1)}) \prod_{l \in V_2} \phi_2(a_l^{(2)}).$$

3. **Free:**

$$(\varphi_1 \star \varphi_2)(a_1 \cdots a_n) = \sum_{I \subset [n], I \neq [n]} (-1)^{n-|I|+1} \left((\phi_1 \star \phi_2)(\overrightarrow{\prod}_{k \in I} a_k) \right) \prod_{l \notin I} \phi_{\varepsilon_l}(a_l^{(\varepsilon_l)})$$

with $(\phi_1 \star \phi_2)(\overrightarrow{\prod}_{k \in \emptyset} a_k) := 1.$

4. **Monotone:**

$$(\phi_1 \triangleright \phi_2)(a_1 \cdots a_n) = \phi_1(\overrightarrow{\prod}_{k \in V_1} a_k^{(1)}) \prod_{l \in V_2} \phi_2(a_l^{(2)}).$$

5. **Anti-monotone:**

$$(\phi_1 \triangleright \phi_2)(a_1 \cdots a_n) = \prod_{k \in V_1} \phi_1(a_k^{(1)}) \phi_2(\overrightarrow{\prod}_{l \in V_2} a_l^{(2)}).$$

Independence

Definition

Let (A, Φ) be a nc probability space. Variables $i_1 : A_1 \to A$ and $i_2 : A_2 \to A$ are $\Box$-independent if

$$\Phi \circ (i_1 \sqcup i_2) = (\Phi \circ i_1)\Box(\Phi \circ i_2).$$

i.e. if the law of the joint $A_1 \sqcup A_2$ is the $\Box$-product of the marginal laws.

Example: Let us compute $\phi(a_1 a_2 a_1 a_2 a_1)$, for $a_i \in A_i$, when A_1 and A_2 are

1. Tensor independent: $\phi_1(a_1^3)\phi_2(a_2^2)$.
2. Boolean independent: $\phi_1(a_1)^3\phi_2(a_2)^2$.
3. Free independent: $\phi(a_1 a_2 a_1 a_2 a_1) = \ldots\ 2^5 - 1$ terms.
4. Monotone: $\phi(a_1^3)\phi(a_2)^2$.
5. Anti-monotone: $\phi_1(a_1)^3\phi_2(a_2^2)$.

A family of unital $*$-subalgebras $(A_i)_{i\in I}$ of (A,Φ) is *freely independent* if $\Phi(a_1\cdots a_n)=0$, whenever $\Phi(a_j)=0$ for all $j\in[n]$ and two consecutive factors a_i belong to different algebras.

A family of subsets $(S_i)_{i\in I}$ of (A,Φ) is freely independent if the algebras A_i generated by $\{1\}\cup S_i$ are freely independent.

To compute the moments of a product $b_1\cdots b_n$ of freely independent noncentered elements, one considers first

$$\Phi((b_1-\Phi(b_1)1)\cdots(b_n-\phi(b_n)1)),$$

which gives the product formula quoted before.

Free groups

Let G be a discrete group. Let λ be the left regular representation of G on $\ell^2(G)$: $\lambda(g)e_h = e_{gh}$, where e_g are the canonical basis vectors of $\ell^2(G)$. The von Neumann algebra $L(G)$ is defined as the weakly closed linear space generated by $\lambda(G)$; roughly speaking, operators $\sum_{g\in G} c_g\lambda(g)$ that are bounded in ℓ^2. The trace/state is $\tau(\sum_{g\in G} c_g\lambda(g)) = c_e$, with e the neutral element.

A family of subgroups $(G_i)_{i\in I}$ is group-theoretic free if there is no non-trivial algebraic relation among the G_i i.e. $g_1\cdots g_n \neq e$ whenever all $g_i \neq e$ and two consecutive factors belong to different subgroups. Then, the free independence of $(\lambda(G_i))_{i\in I}$ in in $(L(G),\tau)$ is equivalent to the group theoretic freeness of $(G_i)_{i\in I}$ in G.

$$\tau(\lambda(g_1)\cdots\lambda(g_n)) = \tau(\lambda(\underbrace{g_1\cdots g_n}_{\neq e})) = 0.$$

Problem: Is it possible to have $L(\mathbb{F}_n) \cong L(\mathbb{F}_m)$ if $N \neq m$.

Relation with random matrices

As before, $\mathscr{A}_N = \bigcap_{1\le p<\infty} L^p(\Omega,\mathscr{F},\mathbb{P},M_n(\mathbb{C}))$, and $\Phi_N(a) = N^{-1}\mathbb{E}(\mathrm{tr}(a))$.

Let A_N and B_N be symmetric matrices with i.i.d. entries that are $(0,N^{-1})$-Gaussians. Then A_N and B_N are asymptotically free as $N\to\infty$,

$$\lim_{N\to\infty}\Phi_N((A_N^{n_1}-\lim_M \mathrm{tr}(A_M^{n_1})\cdot 1)(B_N^{m_1}-\lim_M \mathrm{tr}(B_M^{m_1})\cdot 1) \\ \cdots(A_N^{n_k}-\lim_M \mathrm{tr}(A_M^{n_k})\cdot 1)(B_N^{m_k}-\lim_M \mathrm{tr}(B_M^{m_k})\cdot 1)) = 0. \quad (1)$$

Holds in much more generality: matrices A_N and B_N with asymptotic eigenvalue distribution (i.e. $\lim_M \mathrm{tr}(A_M^k)$ exists for all $k\in\mathbb{N}$), independent, B_N with unitarily invariant distribution.

Noncommutative probability theory

For each notion of independence, there is a notion of central limit theorem, stochastic processes with independent increments (Brownian motion, Lévy process...) and “quantum” stochastic calculus.

Outline

Given a C^*-probability space (A, Φ), consider a sequence $(a_i)_{i\geq 1}$ of self-adjoint elements ($a_i = a_i^*$ for any i), identically distributed (same moments), centered and normalized i.e. $\Phi(a_i) = 0$ and $\Phi(a_i^2) = 1$ for all $i \geq 1$. Suppose that the variables are independent (for some of the notions introduced above), and the mixed moments are bounded: for all $n \geq 1$,

$$\sup_{i_1\geq 1,\dots,i_n\geq 1} |\Phi(a_{i_1}\cdots a_{i_n})| < \infty.$$

Then for all $k \geq 0$,

$$\lim_{n\to\infty} \Phi\left(\left(\frac{a_1 + \cdots + a_n}{\sqrt{n}}\right)^k\right) = \int_{\mathbb{R}} x^k \, \mathrm{d}\mu =: M_k,$$

where the distribution μ is...

...where the distribution μ is:

1. for tensor independence (commutative), the standard Gaussian distribution $(2\pi)^{-1/2}\exp(-x^2/2)\,\mathrm{d}x$, with moments $M_{2k+1} = 0$ and $M_{2k} = \frac{(2k)!}{2(k!)}$.
2. for free independence, the semicircle distribution on $[-2,2]$, $(2\pi)^{-1}\sqrt{4-x^2}\mathbf{1}_{[-2,2]}(x)\,\mathrm{d}x$, with moments $M_{2k+1} = 0$ and $M_{2k} = \frac{1}{k+1}\binom{2k}{k}$, the Catalan numbers.
3. for boolean independence, the symmetric Bernoulli distribution, $\frac{1}{2}(\delta_{-1} + \delta_1)$, with moments $M_{2k+1} = 0$ and $M_{2k} = 1$.
4. for monotone independence, the arc-sine distribution on $[-\sqrt{2}, \sqrt{2}]$, $\pi^{-1}(2-x^2)^{-1/2}\mathbf{1}_{(-\sqrt{2},\sqrt{2})}(x)\,\mathrm{d}x$, with moments $M_{2k+1} = 0$ and $M_{2k} = 2^{-k}\binom{2k}{k}$ [3].

Outline

Free entropy, free Fisher information

If X is a self-adjoint nc random variable with law μ on $\mathbb{R}$, its free entropy is

$$\chi(X) := \int\int \log|s-t|\,\mathrm{d}\mu(s)\,\mathrm{d}\mu(t) + \frac{3}{4} + \frac{1}{2}\log(2\pi e).$$

Among the centered variables with unit variance, the free entropy is maximized by the semicircular law $(2\pi)^{-1}\sqrt{4-x^2}\mathbf{1}_{[-2,2]}(x)\,\mathrm{d}x$.

Idea: an information theoretic proof of the free CLT: partial sums have increasing entropy, and they convergence to the element S of maximum entropy. See [2].

Infinitesimal theory: Introduce the Fisher information $\Phi(X) = \frac{\mathrm{d}}{\mathrm{d}t}\chi(X+\sqrt{t}S)$. This equals $\frac{2}{3}\int (f(x))^3\,\mathrm{d}x$ when X has law $f\,\mathrm{d}x$.

Then the entropy is

$$\chi^*(X) = \frac{1}{2}\int_0^\infty \left(\frac{1}{1+t} - \Phi(X+\sqrt{t}S)\right)\mathrm{d}t + \frac{1}{2}\log(2\pi e).$$

Free Fisher information

If X is a classical, real-valued random variable with density p, its Fisher information is

$$J(X) = \int (\frac{p'}{p})^2 p \,\mathrm{d}\lambda = \left\| \frac{p'}{p} \right\|^2_{L^2(\mathbb{R}, p\,\mathrm{d}\lambda)}.$$

If p has copact support, the score function $s = p'/p$ is the adjoint of the derivation $-(\mathrm{d}/\mathrm{d}t)^*1$. In other words, $\int sq\,\mathrm{d}\lambda = -\int \frac{\mathrm{d}}{\mathrm{d}t} q\,\mathrm{d}\lambda$ for any polynomial q.

Similarly, one can introduce the derivation $\partial_{X:B} : B[X] \to B[X] \otimes B[X]$ that sends X to $1 \otimes 1$ and B to zero, and introduce the score function $s(X:B)$ through the condition

$$\langle s(X:B), P \rangle = \langle 1 \otimes 1, \partial_{X:B} P \rangle = \tau \otimes \tau(\partial_{X:B} P^*)$$

for any P in $B[X]$. Then $J(X:B) = \|s(X:B)\|^2_{L^2(M)}$, where M is the ambient nc proba. space.

Shannon's conjecture

If $(X_i)_{i\in\mathbb{N}^*}$ is an i.i.d. sequence of square integrable random variables, then

$$H\left(\frac{X_1+\cdots X_n}{\sqrt{n}}\right)$$

is an increasing function of n. This was proved by Artstein *et al* in [1].

Later, Shlyakhtenko proved the analogous result in the free case: if $(X_i)_i$ are freely independent, identically distributed random variables

$$n\mapsto\chi\left(\frac{X_1+\cdots X_n}{\sqrt{n}}\right)$$

is monotonically increasing in n [5].

The key ingredient is a projection formula

$$s(X+Y:\mathbb{C})=\mathbb{E}_{W^*(\mathbb{C}[X+Y])}s(X:\mathbb{C}[Y]).$$

This induces certain inequalities because the conditional expectation is a contraction in L^2.

S. Artstein, K. Ball, F. Barthe, and A. Naor, *Solution of shannon's problem on the monotonicity of entropy*, Journal of the American Mathematical Society, 17 (2004), pp. 975–982.

O. Johnson, *Information Theory and the Central Limit Theorem*, Imperial College Press, 2004.

N. Muraki, *Monotonic independence, monotonic central limit theorem and monotonic law of small numbers*, Infinite Dimensional Analysis, Quantum Probability and Related Topics, 4 (2001), pp. 39–58.

——, *The five independences as natural products*, Infinite Dimensional Analysis, Quantum Probability and Related Topics, 6 (2003), pp. 337–371.

D. Shlyakhtenko, *A free analogue of shannon's problem on monotonicity of entropy*, Advances in Mathematics, 208 (2007), pp. 824–833.

R. Speicher, *Free probability theory*, arXiv preprint arXiv:0911.0087, (2009).

D. Voiculescu, *Free entropy*, Bulletin of the London Mathematical Society, 34 (2002), pp. 257–278.