Summary of Commutative and Non Commutative Probability

Vikus JvRensburg

Abstract

A unifying theme in both probability theory and operator theory is the study of *functions of elements*—random variables or operators—and what happens when we apply a *linear functional* such as expectation or trace to them. In analogy with probability, the more functions of a random variable you can integrate, the "nicer" the variable; likewise, in operator theory, the broader the functional calculus available, the more regular the operator. This exposition is intended as a big-picture overview highlighting the parallel structures between probability and operator theory. For clarity of exposition, many technical details and subtleties are omitted.

Functions of Random Variables and Operators

1. Linear functionals and algebras

Both settings begin with an algebra (vector space) of "observables" equipped with a distinguished linear functional.

- In **classical probability**, the algebra is the commutative C^*-algebra

$$L^{-\infty}(\Omega, \mathcal{F}, \mathbb{P}),$$

whose elements are random variables with all finite moments. The expectation

$$\mathbb{E}[\cdot] : L^\infty \to \mathbb{C}$$

is a positive, normalized linear functional. One does not have to assume all finte moments, but if not one might reduce from algebra-structure to vector space structure.

- In **operator theory**, we study a (noncommutative) C^*-algebra $\mathcal{A} \subset B(H)$, and a *state*

$$\varphi : \mathcal{A} \to \mathbb{C}, \qquad \varphi \text{ linear, positive, } \varphi(1) = 1.$$

  The pair $(\mathcal{A}, \varphi)$ plays the role of a "noncommutative probability space".

2. Functions of elements

In both settings, we can apply scalar functions to our basic objects.

Setting	**Object**	**Functional calculus**
Probability	X (random variable)	$f(X)$ defined by composition
Operator theory	A (self-adjoint operator)	$f(A)$ defined by holomorphic, continuous, or Borel functional calculus

In probability theory, finite expectation of $f(X)$ gives valuable information about the distribution $\mathbb{P}_X$ of X. Indeed for example:

1. $\mathbb{E}[|X|] < \infty$ if, and only if $x \in L^1(\mathbb{R}, \mathbb{P}_X)$.
2. $\mathbb{E}[|X|^n] < \infty$ if, and only if $x^n \in L^1(\mathbb{R}, \mathbb{P}_X)$.
3. $\mathbb{E}[e^X] < \infty$ if, and only if $e^x \in L^1(\mathbb{R}, \mathbb{P}_X)$.

3. Applying linear functionals

Once functions of elements make sense, we examine the quantities

$$\mathbb{E}[f(X)] \quad \text{or} \quad \varphi(f(A)).$$

In the commutative case,

$$\mathbb{E}[f(X)] = \int f(x)\, d\mu_X(x),$$

where μ_X is the law of X. In the operator case,

$$\varphi(f(A)) = \int f(x)\, d\mu_A(x),$$

where μ_A is the *spectral measure* of A with respect to φ. Thus, expectations of functions of random variables correspond to traces or states applied to functions of operators.

4. Moments and distributions

Moments store information about the "distribution" in both worlds:

$$\mathbb{E}[X^n] \quad \leftrightarrow \quad \varphi(A^n).$$

Multivariate case: In the commutative setting, the product space is a natural habitat for the joint-distribution of independent random variables $X_1, \cdots X_n$ which is the product measure of the distributions of X_i. This is the law of $X = (X_1, \cdots, X_n)$. In the noncommutative setting, the definition of $A_1, \cdots, A_n$ being free is such that the mixed moments

$$\varphi(A_{i_1} A_{i_2} \cdots A_{i_k})$$

play the role of the joint distribution of the tuple $(A_1, \ldots, A_g)$, in that they only depend on the individual moments of $A_1, \cdots, A_n$. Although the dependence is very complicated.

The takeaway is that free-independence allows one to compute the "joint-distribution" in terms of the individual distribution.

5. Commutative vs. non-commutative probability

The commutative and non-commutative theory becomes visible when we work with a n-tuple of observables which is either classically independent random variables, or freely independent bounded operators.

- **Classical probability** corresponds to the *commutative* case, where all random variables commute:

$$ab = ba \quad \text{for all } a, b \in L^\infty.$$

  Then, the Gelfand representation identifies $L^\infty(\Omega)$ with continuous functions on a measure space, and expectation is integration with respect to a classical probability measure.

- **Free (or noncommutative) probability** generalizes this to *noncommutative* algebras $(\mathcal{A}, \varphi)$, where the elements ("noncommutative random variables") need not commute. Here the distribution of a variable is determined by all its noncommutative moments, and notions of independence are replaced by noncommutative ones, such as *freeness.*

6. Convergence of expectations

A major theme in both subjects is understanding convergence through expectations.

- In probability theory:

  $X_n \xrightarrow{d} X \quad \iff \quad \mathbb{E}[f(X_n)] \to \mathbb{E}[f(X)] \quad \forall f$ bounded, continuous.

- In operator algebras:

  $\mathrm{tr}(p(A_1^{(n)}, \ldots, A_g^{(n)})) \to \mathrm{tr}(p(A_1, \ldots, A_g)) \quad \forall p$ noncommutative polynomial.

7. Transforms of Our Observables

Laplace Transform (Moment Generating Function, MGF)

For a real random variable X, the *Laplace transform* is defined by

$$M_X(t) = \mathbb{E}[e^{tX}],$$

and assume M_X exists on open neiborhood of 0. Expanding e^{tX} gives

$$M_X(t) = \sum_{n=0}^{\infty} \frac{t^n}{n!} \mathbb{E}[X^n],$$

so derivatives at $t = 0$ yield the moments of X. However, the MGF may fail to exist for many distributions (for example, the Cauchy distribution). When it exists in a neighborhood of 0, it uniquely determines the law of X.

Fourier Transform (Characteristic Function)

The *characteristic function* of X is

$$\phi_X(t) = \mathbb{E}[e^{itX}],$$

which always exists since $|e^{itX}| = 1$. It uniquely determines the distribution of X, and its derivatives at $t = 0$ (when they exist) recover the moments:

$$\phi_X^{(n)}(0) = i^n \mathbb{E}[X^n].$$

Characteristic functions play a central role in limit theorems such as the Central Limit Theorem via Lévy's continuity theorem.

Cauchy (Stieltjes) Transform

For a real-valued random variable X with law μ, the *Cauchy transform* (or Stieltjes transform) is defined by

$$G_\mu(z) = \int_{\mathbb{R}} \frac{1}{z-x}\, d\mu(x), \qquad z \in \mathbb{C}^+.$$

This transform is analytic on the upper half-plane and determines the measure via the boundary limit

$$\mu(dx) = -\frac{1}{\pi} \lim_{\varepsilon \downarrow 0} \Im G_\mu(x + i\varepsilon)\, dx.$$

Expanding at infinity gives

$$G_\mu(z) = \frac{1}{z} + \frac{\mathbb{E}[X]}{z^2} + \frac{\mathbb{E}[X^2]}{z^3} + \cdots,$$

so the Cauchy transform also encodes all moments. In the noncommutative setting, this definition only makes sense for *self-adjoint* operators X, ensuring the existence of a spectral measure and that $(z-X)^{-1}$ is well-defined for $z \in \mathbb{C}^+$.

Voiculescu's Use in Free Probability

Voiculescu's key insight was to use the Cauchy transform to build an analytic framework for distributions of *noncommutative (free)* random variables.

For a self-adjoint X in a tracial W^*-probability space, define

$$G_X(z) = \varphi[(z-X)^{-1}], \qquad z \in \mathbb{C}^+.$$

From G_X one defines the reciprocal $F_X(z) = 1/G_X(z)$ and the *Voiculescu transform*

$$\varphi_X(z) = F_X^{-1}(z) - \frac{1}{z}.$$

The crucial property is that for freely independent X and Y,

$$\varphi_{X+Y}(z) = \varphi_X(z) + \varphi_Y(z).$$

This mirrors how the logarithm of the characteristic function linearizes classical convolution, but in the free (noncommutative) setting it is the Cauchy transform that plays this role.

8. Summary

Concept	**Classical probability**	**Operator / free probability**
Algebra of observables	$L^{-\infty(\Omega)}$ (commutative)	C^*-algebra $\mathcal{A}$ (possibly non-commutative)
Element	Random variable X	Operator A
Linear functional	Expectation $\mathbb{E}$	State or trace φ
Function of element	$f(X)$ pointwise	$f(A)$ via functional calculus
Distribution	Law of X	Spectral measure / φ-moments
Independence	Classical independence	Freeness
Convergence	Weak convergence of laws	Convergence in $*$-moments / trace

In this way, classical probability theory and operator theory can be viewed as different aspects of the same overarching idea: studying how linear functionals—expectations or traces—interact with functions of the elements in their respective algebras.