WHAT IS... a Free Cumulant?

Jonathan Novak and Piotr Śniady

Dedicated to Roland Speicher on the occasion of his fiftieth birthday.

Cumulants are quantities coupled to combinatorial notions of "connectivity" and probabilistic notions of "independence". There are two principal species of cumulants: classical and free. Our discussion begins with the more widely known classical cumulants.

Suppose one wishes to determine the number c_n of connected graphs on the vertex set $[n] = \{1, \ldots, n\}$. The total number of (possibly disconnected) graphs on $[n]$ is $m_n = 2^{\binom{n}{2}}$, and because any graph is the disjoint union of its connected components, we have

$$m_n = \sum_{\pi \in \mathcal{P}(n)} \prod_{B \in \pi} c_{|B|}, \tag{1}$$

where the sum runs over all partitions $\pi = B_1 \sqcup B_2 \sqcup \ldots$ of $[n]$ into disjoint nonempty subsets. The sequence (m_n) thus recursively determines (c_n) via (1).

More generally, if m_n is the number of "structures" that can be placed on $[n]$, and c_n is the number of "connected structures" on $[n]$ of the same type, then the sequences (m_n) and (c_n) are related as in (1). This fundamental enumerative link, which is often expressed in terms of generating functions, is ubiquitous in mathematics. Prominent examples come from enumerative algebraic geometry, where connected covers of Riemann surfaces are counted in terms of all covers, and quantum field theory, where sums over connected Feynman diagrams are computed in terms of sums over all diagrams.

Piotr Śniady is a professor in the Polish Academy of Sciences and the University of Wroclaw. His email address is `piotr.sniady@math.uni.wroc.pl`.

Jonathan Novak is a postdoctoral fellow at MSRI and the University of Waterloo. His email address is `j2novak@math.uwaterloo.ca`.

Formula (1) is also well known to probabilists. In stochastic applications, $m_n = m_n(X) = \mathbb{E}X^n$ is the moment sequence of a random variable, and the quantities $c_n(X)$ implicitly defined by (1) are called the *cumulants* of X. This term was coined by R. Fisher and J. Wishart in 1932. Cumulants were, however, first investigated by the Danish astronomer T. Thiele as early as 1889. He called them *half-invariants* because

$$X, Y \text{ independent} \implies c_n(X+Y) = c_n(X) + c_n(Y).$$

This linear behavior is what gives cumulants their advantage: in most situations cumulants are the "right" quantities to work with. For example, the universality of the standard Gaussian distribution is reflected in the simplicity of its cumulant sequence $0, 1, 0, 0, 0, \ldots$.

Let us now consider a geometric variation of our initial graph-counting problem. Given a graph G on $[n]$, we may represent the vertices of G as points on a circle and the edges of G as line segments joining these points. The resulting picture may be connected even if G is not (Figure 1). Let κ_n denote the number of "geometrically connected" graphs on $[n]$. As before, we consider a partition π of $[n]$ which tells which vertices of G belong to the same connected component of its geometric realization.

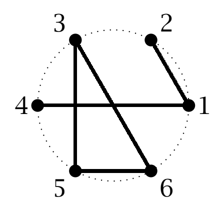

Figure 1. Example of a geometrically connected graph.

It turns out that π is always *noncrossing*: if we represent its blocks as convex polygons, they will be disjoint (Figure 2). It follows that

$$m_n = \sum_{\pi \in \mathcal{NC}(n)} \prod_{B \in \pi} \kappa_{|B|}, \tag{2}$$

where now the summation runs over noncrossing partitions of $[n]$.

Remarkably, just like the c_n's of formula (1), the κ_n's of formula (2) may be realized as half-invariants. Recall that a noncommutative probability space is a complex algebra $\mathcal{A}$ together with a linear functional $\varphi : \mathcal{A} \to \mathbb{C}$ sending $1_{\mathcal{A}}$ to 1. One regards the elements of $\mathcal{A}$ as random variables, with φ playing the role of expectation. The decision to view noncommutative algebras as probability spaces is justified a posteriori by the fact that this framework supports a new notion of independence, called *free independence*, modeled on the free product of algebras. Free independence is the key to a rich noncommutative probability theory—D. Voiculescu's *free probability theory*—which is now widely studied and well known to appear in a wide variety of contexts, for example, in the large dimension limit of random matrix theory.

The realization that formula (2) holds the key to free half-invariants is due to R. Speicher. Given a random variable X in a noncommutative probability space, we associate with its moment sequence $m_n(X) = \varphi(X^n)$ the *free cumulant* sequence $\kappa_n(X)$ via (2). This is completely analogous to the construction of classical cumulants, the only difference being that the lattice of all partitions has been replaced by the lattice of noncrossing partitions. Fundamentally, this means that set-theoretic connectivity has been replaced with geometric connectivity. Speicher showed that

$$X, Y \text{freely indep.} \implies \kappa_n(X+Y) = \kappa_n(X) + \kappa_n(Y).$$

Indeed, the relationship between free independence and free cumulants mirrors the relationship between classical independence and classical cumulants in every conceivable way. For example, the analogue of the Gaussian distribution in free probability theory is the Wigner semicircle distribution, and its universality is reflected in the simplicity of its free cumulant sequence $0, 1, 0, 0, 0, \ldots$.

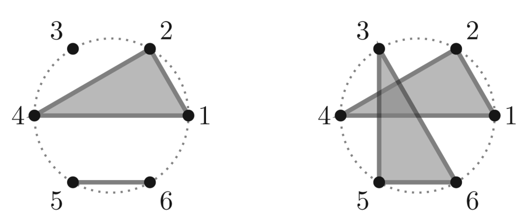

Figure 2. Noncrossing (left) versus a crossing (right) partition.

Free probability theory was initiated by Voiculescu in the 1980s, who used it to solve several previously intractable problems in operator algebras. Thus it is a very new field of mathematics, albeit an exceptionally successful one. It therefore came as a surprise when P. Biane and S. Kerov discovered that free cumulants play an important role in a very classical theory, namely the representation theory of the symmetric groups.

Given a finite group G and a conjugacy class C of G (viewed as an element of the group algebra $\mathbb{C}[G]$), one knows that C acts as a scalar operator in any irreducible representation of $\mathbb{C}[G]$. The value of this scalar defines a function on conjugacy classes, called the *central character* of the irreducible representation. Understanding an irreducible representation amounts to computing its central character which, typically, is very difficult. In the case of the symmetric group $G = S_N$, free cumulants shed substantial light on this question. One begins by considering the *Jucys-Murphy element* $X_N = (1\, N+1) + \cdots + (N\, N+1) \in \mathbb{C}[S_{N+1}]$, which is the sum of all transpositions interchanging $N+1$ with a smaller number. The spectrum of X_N can be completely understood in any irreducible representation; thus X_N and simple functions of X_N may be considered "known". One then defines the nth moment $m_n = m_n(X_N)$ of X_N to be that part of the expansion of X_N^n which belongs to $\mathbb{C}[S_N]$. These moments are not scalars but central elements in $\mathbb{C}[S_N]$, and they give rise to free cumulants $\kappa_n = \kappa_n(X_N)$ via (2). Remarkably, the conjugacy class C_k of k-cycles can be expressed as a polynomial in the free cumulants of X_n, and furthermore one has the first-order approximation $C_k = \kappa_{k+1}$ + lower order terms.

This is not the only surprising appearance of the free cumulant concept. For example, R. Stanley has shown that the free cumulants of M. Haiman's parking function symmetric functions are precisely the complete symmetric functions. M. Lassalle has formulated conjectures linking free cumulants to Jack polynomials. Where will free cumulants appear next?

References

[1] Philippe Biane, Free probability and combinatorics. In *Proceedings of the International Congress of Mathematicians, Vol. II (Beijing, 2002)*, pages 765–774, Beijing, 2002, Higher Ed. Press.

[2] ———, Free probability for probabilists. In *Quantum probability communications, Vol. XI (Grenoble, 1998)*, QP-PQ, XI, pages 55–71, World Sci. Publ., River Edge, NJ, 2003.

[3] Alexandru Nica and Roland Speicher, *Lectures on the combinatorics of free probability*, volume 335 of London Mathematical Society Lecture Note Series, Cambridge University Press, Cambridge, 2006.