Journal of Combinatorial Theory, Series A **91**, 283–304 (2000)
doi:10.1006/jcta.1999.3017, available online at http://www.idealibrary.com on IDEAL®

On the Combinatorics of Cumulants

Gian-Carlo Rota

Department of Mathematics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

and

Jianhong Shen

School of Mathematics, University of Minnesota, Minneapolis, Minnesota 55455
E-mail: jhshen@math.umn.edu

Communicated by the Managing Editors

Received July 26, 1999

DEDICATED TO THE MEMORY OF GIAN-CARLO ROTA

We study cumulants by Umbral Calculus. Various formulae expressing cumulants by umbral functions are established. Links to invariant theory, symmetric functions, and binomial sequences are made.

1. INTRODUCTION

Cumulants were first defined and studied by Danish scientist T. N. Thiele. He called them *semi-invariants*. The importance of cumulants comes from the observation that many properties of random variables can be better represented by cumulants than by moments. We refer to Brillinger [3] and Gnedenko and Kolmogorov [4] for further detailed probabilistic aspects on this topic.

Given a random variable X with the moment generating function $g(t)$, its nth cumulant K_n is defined as

$$K_n = \frac{d^n}{dt^n}\bigg|_{t=0} \log g(t).$$

0097-3165/00 $35.00

That is,

$$\sum_{n\geqslant 0} \frac{m_n}{n!} t^n = g(t) = \exp\left(\sum_{n\geqslant 1} \frac{K_n}{n!} t^n\right), \tag{1}$$

where, m_n is the nth moment of X.

Generally, if σ denotes the standard deviation, then

$$K_1 = m_1, \qquad K_2 = m_2 - m_1^2 = \sigma^2.$$

Cumulants of some important and familiar random distributions are listed as follows. For the *Poisson* distribution with mean λ,

$$K_n \equiv \lambda, \qquad n \geqslant 1.$$

The *exponential* distribution with mean μ has cumulants

$$K_n = (n-1)!\, \mu^n, \qquad n \geqslant 1.$$

The *Gaussian* distribution $N(\mu, \sigma)$ possesses the simplest list of cumulants:

$$K_1 = \mu, \qquad K_2 = \sigma^2, \qquad K_n = 0, \qquad n \geqslant 3.$$

These classical examples clearly demonstrate the simplicity and efficiency of cumulants for describing random variables. It is apparently not fortuitous for cumulants to encode the most important information of the associated random variables. The underlying reason may well reside in the following two invariant properties (which are in fact related to each other).

- (Translation Invariance) Let $K_n(X)$ denote the nth cumulant of a random variable X. Then, for any constant c,

$$K_1(X+c) = c + K_1(X), \qquad K_n(X+c) = K_n(X), \quad n \geqslant 2.$$

- (Additivity) Let X and Y be any two independent random variables. Then,

$$K_n(X+Y) = K_n(X) + K_n(Y), \qquad n \geqslant 1.$$

Our combinatorial interests in cumulants are very much inspired by these algebraic properties, especially the translation invariance. Notice that these are truly *algebraic* relations, which has very little to do with the *positivity* attribute of random variables (though in probability theory or statistics, positivity is extremely important; see, for instance, the famous Moment Problem [1, 2, 8]). Given any sequence of (real) numbers: $a_0 = 1$, a_1, a_2, ..., the associated cumulants $K_1, K_2, \ldots$ are always well defined according to Eq. (1). While enjoying this fresh degree of freedom, we have to face the difficulty in interpreting $X+c$, since for an arbitrary sequence, it is unclear whether there exists a random variable X or not (unless we solve the moment problem).

This difficulty is overcome by the notion of umbrae. An umbra is a *formal*, or *dry*, vividly speaking, random variable, which has no probabilistic flesh, but is indeed endowed with the algebraic spirit of a random variable. It is a powerful tool in dealing with a sequence of numbers (such as the Bell numbers; see Rota [10]), and for studying combinatorial algebraic objects like binomial sequences and algebraic invariants of polynomial systems (Hilbert [5], Kung and Rota [7]). We refer to Roman and Rota [9] and Rota, Kahaner, and Odlyzko [11] for the history and development of Umbral Calculus. An algebraic treatment was given in Joni and Rota [6]. New developments can be found in Rota and Taylor [13] and Rota, Shen, and Taylor [12]. Also see Shen [14] for a recent application in wavelet analysis.

In this paper, we shall employ the full freedom of Umbral Calculus to study cumulants. Umbral Calculus leads to various formulae for the cumulant sequence, each of which reveals one portion of the secret encoded in cumulants. Through these formulae, cumulants are connected to familiar combinatorial objects such as binomial sequences and symmetric functions. In return, the study of cumulants has stimulated new extension of the existing theory of Umbral Calculus. For instance, for the first time in this paper, we discuss umbral derivatives (or the *star algebra*).

Our plan is the following. In Section 2, we survey briefly the literature and recent development of Umbral Calculus. By doing so, we make our readers comfortable with the notations and symbols necessary for the rest of the paper. The following five sections study five different ways of understanding cumulants, among which, four are based on Umbral Calculus. The fith employs partition lattices and similar work can be found in Speed [15].

2. UMBRAL CALCULUS

An umbra α is the generalization of the expectation E of a random variable X. It is a powerful tool for dealing with a sequence of numbers $a_0, a_1, \ldots$, whether positive definite or not. (A sequence of real numbers is said to be *positive definite* if it can be realized as the moment sequence of certain random variables.) In some sense, an umbra can be called *a formal random variable*. The formality makes it much easier to understand operators like $\partial/\partial\alpha$.

2.1. *Fundamentals*

Umbral Calculus is axiomatized by the following definition (See Rota and Taylor [13]).

DEFINITION 2.1. An umbral calculus consists of the following data and rules:

(1) An alphabet A, whose elements are called *umbrae* and denoted by Greek letters.

(2) A commutative integral domain D whose quotient field is of characteristic zero.

(3) A D-linear functional **eval**: $D[A] \to D$ such that $\mathbf{eval}(1) = 1$, and

$$\mathbf{eval}(\alpha^i \beta^j \cdots \gamma^k) = \mathbf{eval}(\alpha^i)\, \mathbf{eval}(\beta^j) \cdots \mathbf{eval}(\gamma^k)$$

whenever $\alpha, \beta, ..., \gamma$ are distinct umbrae.

(4) A distinguished element ε of the alphabet A, such that

$$\mathbf{eval}(\varepsilon^n) = \delta_n,$$

where δ is the Kronecker delta.

If two umbral polynomials p and q are evaluated to a same element in D, then they are said to be *umbrally equivalent* and denoted by $p \simeq q$. A sequence $a_0, a_1, ...$ is said to be *umbrally represented* by an umbra α, if

$$\mathbf{eval}(\alpha^n) = a_n, \qquad \text{for} \quad n = 0, 1, ...$$

According to Axiom (3), a_0 must be 1. If two distinct umbrae α and β represent the same sequence, then they are said to be *exchangeable* and we write

$$\alpha \equiv \beta.$$

Notice the difference between $\alpha \simeq \beta$ and $\alpha \equiv \beta$. The notion of exchangeability plays a significant role in the algebraic development of Umbral Calculus. It models the familiar concept of "i.i.d" (independent and identical distributions) in probability theory.

Umbral Calculus is usually assumed to be *saturated*, which means that any sequence in D can be represented by infinitely many umbrae in A (see Rota and Taylor [13]). Especially, A cannot be countable.

A taste of Umbral Calculus can be sought from the famous example of the *Bell Umbra* β. An umbra β is called a *Bell umbra* if it umbrally represents the sequence of Bell numbers $B_0, B_1, ...$. (B_n is the number of distinct partitions of a set with n elements. See Rota [10].) This umbra is completely characterized by

$$\beta^{n+1} \simeq (\beta+1)^n, \qquad n=0, 1, 2, ...$$

All other properties of the Bell numbers are the derivatives of this simple relation.

By the saturation assumption, any formal power series in $D[[t]]$

$$g(t) = \sum_{n \geqslant 0} \frac{a_n}{n!} t^n, \qquad a_n \in D$$

can be umbrally represented by a formal power series in $D[A][[t]]$. In fact, if α umbrally represents the sequence (a_n), then

$$g(t) \simeq e^{\alpha t},$$

assuming that we naturally extend **eval** to be $D[t]$-linear. We say that $g(t)$ is umbrally represented by α. If $g(t)$ and $f(t)$ are umbrally represented by two distinct umbrae α and β, then $g(t)\, f(t)$ is represented by $\alpha+\beta$ (or any one that is exchangeable with it). This is precisely a restatement of the fact that $g(t)\, f(t)$ defines a convolution in terms of their coefficients.

2.2. *Further Developments*

2.2.1. *The Annihilating Umbra and Auxiliary Umbrae*

For a given umbra $\alpha \in A$, we define its *annihilating umbra* $-.\alpha$ to be the unique umbra (up to umbral exchangeability) whose moment generating function is the reciprocal of α's. It is uniquely characterized by the umbral equivalence

$$\alpha + (-.\alpha) \equiv \varepsilon.$$

More generally, for any integer n, we use $n.\alpha$ to denote the unique (up to exchangeability) umbra whose moment generating function is the nth power of α's. When n is positive, then

$$n.\alpha \equiv \alpha_1 + \alpha_2 + \cdots + \alpha_n$$

for any n distinct umbrae α_i, all exchangeable with α. If n is negative, say $n = -k$, then

$$-k.\alpha \equiv \beta_1 + \beta_2 + \cdots + \beta_k$$

for any k distinct umbrae β_i exchangeable with the annihilating umbra $-.\alpha$. We call $n.\alpha$ an *auxiliary umbra* of α. More discussion can be found in Rota and Taylor [13].

2.2.2. *Partial Derivatives*

In the algebra $D[A]$, for any $\alpha \in A$, let α^* denote the partial derivative operator ∂_α. It is clear that for any $\alpha, \beta \in A$,

$$\alpha^* \circ \beta^* = \beta^* \circ \alpha^*.$$

Hence, if $D[A^*]$ denotes the subalgebra generated by all star elements α^* in the D-linear operator algebra of $D[A]$, then $D[A^*]$ is isomorphic to the polynomial algebra generated by the star alphabet

$$A^* = \{\alpha^* \mid \alpha \in A\}.$$

The mapping $\alpha \to \alpha^*$ from A to A^* introduces a natural isomorphism between $D[A]$ and $D[A^*]$. That is, for any umbral polynomial $p(\alpha, \beta, ..., \gamma) \in D[A]$,

$$p \to p^* \in D[A^*]: \qquad p^* = p(\alpha^*, \beta^*, ..., \gamma^*).$$

We call it *the star isomorphism.*

Furthermore, $D[A]$ can D-linearly act on itself by multiplication. In this sense, the elements in $D[A]$ are called *multipliers* and $D[A]$ is also treated as a subalgebra of its D-linear operator algebra. Let $D[A \mid A^*]$ denote the subalgebra generated by all multipliers and star elements. Because of the uncertainty principle (Strang [16]),

$$\alpha^*\alpha - \alpha\alpha^* = 1,$$

$D[A \mid A^*]$ cannot be commutative. Yet the following result seems to need no proof.

PROPOSITION 2.1. *Algebra $D[A \mid A^*]$ is isomorphic to $(D[A])[A^*]$, the polynomial algebra generated by the star alphabet A^* on the umbral algebra $D[A]$.*

It is now the time for exploring various umbral approaches to understanding cumulants.

3. FIRST REPRESENTATION OF CUMULANTS: CUMULANT UMBRAE

As for the Bell numbers, it is possible to give a clean umbral recursion formula for cumulants of a given sequence $a_0 = 1, a_1, a_2, ...$. The concept of cumulant umbra is coined for this purpose. This first representation is the simplest and most direct, though it may not be optimal for understanding the algebraic properties hidden in cumulants.

DEFINITION 3.1 [The Cumulant Umbra]. For a given umbra α, up to exchangeability, there exists a unique umbra, say κ, such that,

$$\alpha^n \simeq \kappa(\kappa+\alpha)^{n-1} \tag{2}$$

for all $n=1, 2, \ldots$. We call κ the *cumulant umbra* associated with α.

The existence and uniqueness follow readily from the recursive relation

$$\kappa^n \simeq \alpha^n - \sum_{m=0}^{n-2} \binom{n-1}{m} \kappa^{m+1} \alpha^{n-1-m}. \tag{3}$$

On the other hand, suppose α represents a sequence (a_n). Define

$$K_n = \left.\frac{d^n}{dt^n}\right|_0 \ln\left(\sum_{m \geqslant 0} \frac{a_m}{m!} t^m\right),$$

for $n=1, 2, \ldots$. As for random variables, we call K_n the nth cumulant of the sequence (and the umbra α).

PROPOSITION 3.1.

$$\kappa^n \simeq K_n, \qquad n=1, 2, \ldots.$$

Proof. Define the generating functions

$$g(t) = \mathbf{eval}(e^{\alpha t}), \qquad K(t) = \mathbf{eval}(e^{\kappa t}).$$

From Eq. (2),

$$\alpha \sum_{n \geqslant 0} \frac{\alpha^n}{n!} t^n \simeq \kappa \sum_{n \geqslant 0} \frac{(\kappa+\alpha)^n}{n!} t^n,$$

$$\alpha e^{\alpha t} \simeq \kappa e^{(\kappa+\alpha)t} = \kappa e^{\kappa t} e^{\alpha t},$$

$$\frac{de^{\alpha t}}{dt} \simeq \frac{de^{\kappa t}}{dt} e^{\alpha t}.$$

Noticing that d/dt commutes with **eval**, we eventually have the ordinary differential equation

$$g'(t) = K'(t)\, g(t),$$

with the initial condition $g(0)=1$. This gives the unique solution

$$g(t)=\exp(K(t)-1),$$

which proves that $K_n=\mathbf{eval}(\kappa^n)$. ∎

Thus, we can derive the classical result.

COROLLARY 3.1. *Suppose that* α *represents a sequence* a_n, $n=0, 1, \ldots$. *Then its* n *th cumulant* K_n *is given by*

$$K_n=(-1)^{n-1}(n-1)!\det\begin{bmatrix} a_1 & 1 & 0 & 0 & \cdots \\ a_2 & a_1 & 1 & 0 & \cdots \\ a_3/2! & a_2/2! & a_1 & 1 & \cdots \\ a_4/3! & a_3/3! & a_2/2! & a_1 & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \end{bmatrix}_{n\times n}. \tag{4}$$

Proof. According to the preceding proposition, the recursion formula (3), after evaluation, is a linear system of equations on cumulants K_m, with coefficients from the sequence $a_0, a_1, \ldots$. We apply the evaluated formula for $m=1, 2, \ldots, n$. Then K_n can be expressed explicitly using Cramer's formula for linear systems. Formula (4) is finally obtained by moving the last column to the first in Cramer's formula for K_n.

Our next representation starts from a further umbralization of this matrix formula.

4. SECOND REPRESENTATION: UMBRAL SYMMETRIC FUNCTIONS

4.1. *A Notation for Generalized Vandemondes*

Let R be an arbitrary ring (commutative or non-commutative). For any n by n matrix $M=[a_{i,j}]$, $a_{i,j}\in R$, if a_{ij} commutes with a_{kl} whenever $k\neq i$, and $l\neq j$, then we can define the determinant of M by the classical expansion formula. Properties such as skew-symmetry and Laplace expansion still hold.

Especially, suppose $x_1, x_2, \ldots, x_n$ are n mutually commutative elements in R. For any partition λ of length n:

$$\lambda=(\lambda_1, \lambda_2, \ldots, \lambda_n)$$

with non-negative integer parts λ_i, we define

$$(x_1, x_2, ..., x_n \mid \lambda) := \det[x_i^{\lambda_j}]_{n \times n}.$$

For example, the classical Vandemonde $V_n(x_1, x_2, ..., x_n)$ is now written as

$$(x_1, x_2, ..., x_n \mid 0, 1, ..., n-1).$$

We shall call $(x_1, x_2, ..., x_n \mid \lambda)$ the Vandemonde of $(x_1, x_2, ..., x_n)$ with respect to λ. Notice that

$$\frac{(x_1, x_2, ..., x_n \mid \lambda)}{(x_1, x_2, ..., x_n \mid 0, 1, ..., n-1)}$$

is the classical Schur function S_λ.

In what follows, we apply this notation to the umbral algebra $D[A]$ and operator algebra $D[A \mid A^*]$. For star elements, we can even allow the partition λ to have negative parts under the following assumption. For any star element $\alpha^* \in A^*$, and any positive integer k, we define $[\alpha^*]^{-k}$ to be the multiplier α^k. That is, "–" seems to annihilate "*."

Notice that for a star element, negative powers do not commute with positive ones. However, for any two *distinct* umbrae $\alpha, \beta \in A$, and any two integers n, m (negative maybe), $[\alpha^*]^n$ commutes with $[\beta^*]^m$. Hence, for any n distinct umbrae $\alpha_1, \alpha_2, ..., \alpha_n$, and partition λ of length n and with arbitrary integer parts,

$$(\alpha_1^*, \alpha_2^*, ..., \alpha_n^* \mid \lambda)$$

is well defined and is an element in $D[A \mid A^*]$.

4.2. *Symmetric Umbral Representation for Cumulants*

Suppose K_n is the nth cumulant of a given umbra α, which represents a sequence a_n, $n = 0, 1, ...$ Let $\alpha_1, \alpha_2, ..., \alpha_n \in A$ be any n distinct umbrae, each exchangeable with α. The main task of this section is to establish the following theorem:

THEOREM 4.1 (Symmetric and Translation Invariant Representation). *For any* $n = 1, 2, ...,$

$$K_n \simeq C_n(\alpha_1^*, \alpha_2^*, ..., \alpha_n^* \mid -1, 0, ..., n-2) \cdot (\alpha_1, \alpha_2, ..., \alpha_n \mid 0, 1, ..., n-1)$$

with

$$C_n = (-1)^{n-1} [((n-2)!)!\, n!]^{-1}.$$

Here $(k!)!=(1!)(2!)\cdots(k!)$ *for any non-negative integer* k *and* $((-1)!)!$ *is assumed to be* 1.

Proof. In Eq. (4), use α_i to represent the ith column of the matrix involved. Since all columns are now umbrally unrelated, we can exchange the order of the det operator and **eval**. Therefore,

$$K_n \simeq (-1)^{n-1}\,(n-1)!\,\det\begin{bmatrix} \alpha_1 & 1 & 0 & 0 & \cdots \\ \alpha_1^2 & \alpha_2 & 1 & 0 & \cdots \\ \alpha_1^3/2! & \alpha_2^2/2! & \alpha_3 & 1 & \cdots \\ \alpha_1^4/3! & \alpha_2^3/3! & \alpha_3^2/2! & \alpha_4 & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \end{bmatrix}_{n\times n}$$

$$=(-12)^{n-1}\,(n-1)!$$

$$\times\det\begin{bmatrix} \alpha_1\cdot 1 & 1 & \alpha_3^* 1 & [\alpha_4^*]^2\, 1 & \cdots \\ \alpha_1\cdot\alpha_1 & \alpha_2 & \alpha_3^*\alpha_3 & [\alpha_4^*]^2\,\alpha_4 & \cdots \\ \alpha_1\cdot\alpha_1^2/2! & \alpha_2^2/2! & \alpha_3^*\alpha_3^2/2! & [\alpha_4^*]^2\,\alpha_4^2/2! & \cdots \\ \alpha_1\cdot\alpha_1^3/3! & \alpha_2^3/3! & \alpha_3^*\alpha_3^3/3! & [\alpha_4^*]^2\,\alpha_4^3/3! & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \end{bmatrix}_{n\times n}$$

$$=\frac{(-1)^{n-1}}{((n-2)!)!}\,[\alpha_1^*]^{-1}\,[\alpha_2^*]^0\,[\alpha_3^*]^1\cdots[\alpha_n^*]^{n-2}\,V_n(\alpha_1,\alpha_2,...,\alpha_n).$$

The linearity of **eval** and exchangeability of α_k's allow the symmetrization:

$$K_n \simeq \frac{(-1)^{n-1}}{((n-2)!)!\,n!}\sum_{\sigma\in S_n}[\alpha^*_{\sigma[1]}]^{-1}\,[\alpha^*_{\sigma[2]}]^0\,[\alpha^*_{\sigma[3]}]^1\cdots[\alpha^*_{\sigma[n]}]^{n-2}$$

$$\times V_n(\alpha_{\sigma[1]},\alpha_{\sigma[2]},...,\alpha_{\sigma[n]})$$

$$=C_n\sum_{\sigma\in S_n}\operatorname{sign}(\sigma)[\alpha^*_{\sigma[1]}]^{-1}$$

$$\times[\alpha^*_{\sigma[2]}]^0\,[\alpha^*_{\sigma[3]}]^1\cdots[\alpha^*_{\sigma[n]}]^{n-2}\,V_n(\alpha_1,\alpha_2,...,\alpha_n)$$

$$=C_n(\alpha_1^*,\alpha_2^*,...,\alpha_n^*\mid -1,0,...,n-2)\cdot(\alpha_1,\alpha_2,...,\alpha_n\mid 0,1,...,n-1).$$

This completes the proof. ∎

Let $p(\alpha_1,\alpha_2,...,\alpha_n)$ denote the umbral polynomial

$$(\alpha_1^*,\alpha_2^*,...,\alpha_n^*\mid -1,0,...,n-2)\cdot(\alpha_1,\alpha_2,...,\alpha_n\mid 0,1,...,n-1).$$

It is easy to see that p is a symmetric function of $\alpha_1, ..., \alpha_n$. Furthermore, we now show that when $n \geqslant 2$, p is translation invariant. Namely, for any real constant c,

$$p(\alpha_1 + c, \alpha_2 + c, ..., \alpha_n + c) = p(\alpha_1, \alpha_2, ..., \alpha_n).$$

First,

$$(\alpha_1, \alpha_2, ..., \alpha_n \mid 0, 1, ..., n-1) = \prod_{j>i} (\alpha_j - \alpha_i)$$

is apparently translation invariant. Second, the property of translation invariance is preserved under the action of the star algebra $D[A^*]$. Finally, Laplace expansion with respect to the first and second columns represents

$$(\alpha_1^*, \alpha_2^*, ..., \alpha_n^* \mid -1, 0, ..., n-2)$$

by

$$\sum_{j>i} (-1)^{i+j} (\alpha_i, \alpha_j \mid 0, 1)$$
$$\times (\alpha_1^*, ..., \alpha_{i-1}^*, \alpha_{i+1}^*, ..., \alpha_{j-1}^*, \alpha_{j+1}^*, ..., \alpha_n^* \mid 1, 2, ..., n-2).$$

For each index pair j, i,

$$(\alpha_1^*, ..., \alpha_{i-1}^*, \alpha_{i+1}^*, ..., \alpha_{j-1}^*, \alpha_{j+1}^*, ..., \alpha_n^* \mid 1, 2, ..., n-2)$$

is in the star algebra $D[A^*]$ and the multiplier $(\alpha_i, \alpha_j \mid 0, 1) = \alpha_j - \alpha_i$ is clearly translation invariant. All together, these facts confirm the translation invariance of p.

Accordingly, we have established umbrally the translation invariance of cumulants addressed in the Introduction.

COROLLARY 4.1. *Let $K_n(\alpha)$ denote the nth cumulant of an umbra α. Then for any real constant c, and $n \geqslant 2$,*

$$K_n(\alpha + c) = K_n(\alpha).$$

5. THIRD REPRESENTATION: UMBRAL PARAMETRIC FORM

5.1. *The Generalized Umbral Calculus*

The *Generalized Umbral Calculus* (see Rota, Shen, and Taylor [12]) is obtained from the classical one by replacing axiom (3) in Definition 2.1 by

(3′) A D-linear functional **eval**: $D[A] \to D[x]$ (instead of D) such that $\mathbf{eval}(1) = 1$, and

$$\mathbf{eval}(\alpha^i \beta^j \cdots \gamma^k) = \mathbf{eval}(\alpha^i)\, \mathbf{eval}(\beta^j) \cdots \mathbf{eval}(\gamma^k)$$

whenever $\alpha, \beta, ..., \gamma$ are distinct umbrae.

We call α a *scalar* umbra if it represents a sequence of D elements, and χ a *polynomial umbra* if it represents a sequence of $D[x]$ elements. In this paper, we always use Greek letters α, β, γ to represent scalar umbrae, and χ, ϕ, ψ to represent polynomial umbrae.

There is an important operator called *restriction* in the generalized Umbral Calculus that relates polynomial umbrae to scalar ones.

Given an element in D, say b, for any polynomial umbra χ, its *restriction* at b is the unique scalar umbra (up to exchangeability) that represents

$$1, \mathbf{eval}(\chi)(b), \mathbf{eval}(\chi^2)(b), \mathbf{eval}(\chi^3)(b), ...$$

We denote this scalar umbra by $\langle \chi \,|\, b \rangle$.

The following properties hold apparently.

(1) $\langle \phi \,|\, b \rangle + \langle \psi \,|\, b \rangle \equiv \langle \phi + \psi \,|\, b \rangle$ for any two distinct polynomial umbrae ϕ and ψ.

(2) $c \langle \phi \,|\, b \rangle \equiv \langle c\phi \,|\, b \rangle$ for any constant c.

For simplicity, we now assume that D is the real number field.

5.2. *Binomial Sequences and Binomial Umbrae*

There is a close relation between the moment–cumulant pair and the binomial sequence of polynomials. We refer to Rota, Shen, and Taylor [12] for the most recent development in binomial sequences of polynomials.

A sequence of real polynomials $p_n(x)$, $n = 0, 1, ...,$ is called a *binomial sequence* if $p_1 \neq 0$, and for any non-negative integer n and real numbers a and b,

$$p_n(a+b) = \sum_{0 \leqslant k \leqslant n} \binom{n}{k} p_k(a)\, p_{n-k}(b).$$

It is easy to see that $p_0 = 1$ and $p_1 = cx$ for some non-zero constant c.

A polynomial umbra χ representing a binomial sequence $p_n(x)$, $n = 0, 1, ...,$ is called a *binomial umbra*.

PROPOSITION 5.1. *A polynomial umbra χ is a binomial umbra if and only if* $\mathbf{eval}(\chi) \neq 0$, *and for any real numbers a and b,*

$$\langle \chi \,|\, a+b \rangle \equiv \langle \chi \,|\, a \rangle + \langle \chi \,|\, b \rangle.$$

DEFINITION 5.1. A binomial umbra χ is called *unital* if

$$\chi \simeq x.$$

Similarly, we call the associated binomial sequence $p_n(x)$ *unital* when $p_1(x) = x$.

A unital binomial sequence is the most direct generalization of the sequence x^n, $n = 0, 1, \ldots$. Another household binomial sequence is the factorial sequence:

$$p_n(x) = x(x-1)\cdots(x-n+1) = (x)_n, \qquad n = 0, 1, \ldots.$$

The preceding proposition together with the properties of the restriction operator brings us to

COROLLARY 5.1 (Linearity). *Suppose χ_1 and χ_2 are two binomial umbrae, and c_1, c_2 two real numbers such that*

$$\mathbf{eval}(c_1\chi_1 + c_2\chi_2) \neq 0.$$

Then, any polynomial umbra exchangeable with $c_1\chi_1 + c_2\chi_2$ must be a binomial umbra.

Rota, Shen, and Taylor proved the following theorem in [12] to uniformly characterize an arbitrary binomial sequence.

THEOREM 5.1. *A polynomial umbra χ is binomial if and only if there exist a scalar umbra β and a non-zero real number c, such that for any $n = 1, 2, \ldots$*

$$\chi^n \simeq cx(cx + n.\beta)^{n-1}.$$

Furthermore, c and β are unique (up to exchangeability).

5.3. *Moment–Cumulant vs Binomial Sequence*

THEOREM 5.2. *Suppose κ is the cumulant umbra of a given scalar umbra α whose "mean" $\mathbf{eval}(\alpha)$ is not zero. Then there exists a unique binomial umbra (up to exchangeability), say χ_α, such that*

$$\alpha \equiv \langle \chi_\alpha \mid 1 \rangle, \qquad \kappa^n \simeq \left.\frac{d}{dx}\right|_0 \chi_\alpha^n, \qquad n = 1, 2, \ldots,$$

where $d/dx(\chi_\alpha^n)$ stands for $d/dx(\mathbf{eval}(\chi_\alpha^n))$.

Proof. Uniqueness follows from the fact that any binomial umbra can be reconstructed (up to exchangeability) from its restricted scalar umbra at some non-zero value.

Let $g(t)$ and $K(t)$ denote the generating functions of α and κ. Set $f(t) = K(t) - 1$. Then $g(t) = e^{f(t)}$. Assume that

$$e^{f(t)\,x} = \sum_{n \geqslant 0} \frac{p_n(x)}{n!}\, t^n. \tag{5}$$

Then the sequence $p_n(x)$, $n = 0, 1, ...,$ must be a binomial sequence. Let χ denote the binomial umbra representing (p_n). Then $\exp(f(t)\,x) \simeq \exp(t\chi)$ and

$$g(t) = e^{f(t)} \simeq e^{t\,\langle \chi \mid 1 \rangle}.$$

On the other hand, taking derivative with respect to x on both sides of (5) and setting $x = 0$, we have

$$\kappa^n \simeq p_n'(0) \simeq \left.\frac{d}{dx}\right|_0 \chi^n.$$

Hence, χ is the demanded binomial umbra. ▮

As a result, one can easily prove (by noticing that $\chi_{\alpha+\beta} \equiv \chi_\alpha + \chi_\beta$) that

Corollary 5.2 (Additivity). *For any two distinct scalar umbra α and β,*

$$K_n(\alpha + \beta) = K_n(\alpha) + K_n(\beta), \qquad n = 1, 2, ...$$

The preceding two theorems establish the umbral parametric representation for cumulants.

Corollary 5.3 (Umbral Parametric Form). *For any umbra α with a non-zero mean $c = \mathbf{eval}(\alpha)$, there exists a unique umbra, say β, such that*

$$\alpha^n \simeq c(c + n.\beta)^{n-1} \qquad \textit{and} \qquad \kappa^n \simeq c(n.\beta)^{n-1}, \qquad n = 1, 2, ...$$

In this representation, α and κ are "parameterized" by another scalar umbra β.

Example 1. Suppose β is the Bell umbra and $c = 1$. The associated binomial sequence $p_n(x)$ is exactly the factorial sequence $(x)_n$. Hence, $\alpha^n \simeq p_n(1) = 0$ for all $n \geqslant 2$, and $\kappa^n \simeq (-1)^{n-1}\,(n-1)!$. Since $K_2 = D_2 = -1$, α is cannot be positive, or equivalently, α cannot be the moment umbra of a real random variable. However, if we take $c = -1$, then

$$\kappa^n \simeq (-1)^n\,(n-1)!,$$

which is exactly the nth cumulant of the exponential distribution with mean -1 (see Section 1).

EXAMPLE 2. Suppose β represents sequence $1, b, 0, 0, \ldots$ and $c=1$. By the preceding corollary,

$$\kappa^n \simeq (n)_{n-1}\, b^{n-1} = n!\, b^{n-1}.$$

Therefore, the moment generating function of the underlying moment umbra α is

$$g(t) = \exp\left(\frac{t}{1-bt}\right).$$

6. FOURTH REPRESENTATION: PARTITION LATTICE AND MÖBIUS INVERSION

This section parallels Speed's work [15].

6.1. *Sequence Partitions and Set Partitions*

A finite sequence of real numbers $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_k)$ is a *sequence partition* if $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_k$. Another way to represent a sequence partition is the *multiset* form: $(a_1^{m_1}, a_2^{m_2}, \ldots, a_h^{m_h})$ with $a_1 > a_2 > \cdots > a_h$ and $m_i > 0$. λ has exactly m_i a_i's. The following quantities are standard but helpful for the rest of the paper.

(1) $\lambda! = \lambda_1!\, \lambda_2! \cdots \lambda_k!$. $|\lambda| = \lambda_1 + \lambda_2 + \cdots + \lambda_k$. $\prod \lambda = \lambda_1 \lambda_2 \cdots \lambda_k$. $\lambda \vdash n$ means that $|\lambda| = n$.

(2) k is the *length* of λ and is also denoted by l_λ.

(3) Vector $m_\lambda = (m_1, m_2, \ldots, m_h)$ encodes the *multiplicity* of λ. $m_\lambda! = m_1!\, m_2! \cdots m_h!$. Clearly, $l_\lambda = m_1 + m_2 + \cdots + m_h$.

(4) For any two partitions λ and μ, their sum $\lambda + \mu$ is the unique sequence partition whose associated multiset is the union. For example, suppose $\lambda = (322)$, $\mu = (4321)$; then $\lambda + \mu = (4332221)$.

(5) A function g_λ defined for all sequence partitions λ is said to be *decomposable* if

$$g_{\lambda+\mu} = g_\lambda g_\mu, \qquad \text{for any } \lambda, \mu.$$

For example, the product function $\prod \lambda$ is decomposable. Apparently, if g_λ is decomposable, then

$$\operatorname{span}_{\mathbb{Z}}\{g_\lambda \mid \text{all } \lambda\} = \mathbb{Z}[g_1, g_2, g_3, ...].$$

Now we review some basic facts about partitions of a *set*. Let $\prod(S)$ denote the partition lattice of a set S. If $S=[n]=\{1, 2, ..., n\}$, we simply write $\prod_n$ for $\prod(S)$. A partition is denoted by $\sigma=\sigma_1 \mid \sigma_2 \mid \cdots \mid \sigma_k$. Each σ_i is a block of σ.

For a given partition $\sigma=\sigma_1 \mid \sigma_2 \mid \cdots \mid \sigma_k$, we define

$$\lambda(\sigma)=(\lambda_1, \lambda_2, ..., \lambda_k), \qquad l_i = \#\sigma_i, \quad i=1, 2, ..., k.$$

$\lambda(\sigma)$ is called the *type* of σ and k the *size* of σ and sometimes is denoted by b_σ. Hence, $b_\sigma = l_{\lambda(\sigma)}$.

Suppose T and S are two disjoint sets and $\sigma \in \prod(T)$, $\theta \in \prod(S)$. Their sum $\sigma+\theta$ is the unique partition of $T \cup S$ which refines the two-block partition $T \mid S$ and whose restrictions in T and S are σ and t. Then, $\lambda(\sigma+\theta) = \lambda(\sigma)+\lambda(\theta)$.

Let $\prod_\infty = \bigcup_{n \geqslant 1} \prod_n$ be the poset whose order is defined by $\sigma \leqslant \theta$ if and only if for some n, $\sigma, \theta \in \prod_n$, and $\sigma \leqslant \theta$. A function g_σ in $\prod_\infty$ (or $\prod_n$) is *indistinguishable* if

$$g_\sigma = g_\theta, \qquad \text{whenever} \quad \lambda(\sigma)=\lambda(\theta).$$

If g_σ is indistinguishable, then it introduces a type class function g^*_λ such that $g_\sigma = g^*_{\lambda(\sigma)}$. For simplicity, we still denote g^*_λ by g_λ. The size function b_σ is indistinguishable, for instance. A distinguishable function g_σ is said to be *decomposable* if its associated type class function g_λ is so.

6.2. *Möbius Inversion*

The integral f_σ of a given function g_σ on $\prod_\infty$ is

$$f_\sigma = \sum_{\theta \leqslant \sigma} g_\theta.$$

According to the Möbius inversion formula,

$$g_\sigma = \sum_{\theta \leqslant \sigma} \mu(\theta, \sigma) f_\theta,$$

if $\mu(\theta, \sigma)$ denotes the Möbius function of the lattice.

By the order structure of $\prod_\infty$, if $\sigma \in \prod_n$, we can replace $\sum_{\theta \leqslant \sigma}$ by $\sum_{\theta \leqslant \sigma,\, \theta \in \prod_n}$ in the above equations. The following proposition can be checked easily.

PROPOSITION 6.1. *A function g_σ is indistinguishable (or decomposable) if and only if its integral f_σ is.*

For any $\lambda \vdash n$, let $\binom{n}{\lambda}^*$ denote the number of partitions σ in $\prod_n$ whose type is λ. It is not difficult to see that

$$\binom{n}{\lambda}^* = \frac{n!}{\lambda!\, m_\lambda!} = \binom{n}{\lambda} \frac{1}{m_\lambda!}.$$

Denote $[n] \in \prod_n$ by $\mathbf{1}_n$ and $g_{\mathbf{1}_n}$ by g_n for any function g_σ in $\prod_\infty$.

THEOREM 6.1 (Inversion Formula). *Suppose that g_σ is a function on $\prod_\infty$ with integral f_σ.*

(a) *If g is indistinguishable, then*

$$f_n = \sum_{\lambda \vdash n} \binom{n}{\lambda}^* g_\lambda, \qquad \textit{and} \qquad g_n = \sum_{\lambda \vdash n} \binom{n}{\lambda}^* f_\lambda (-1)^{l_\lambda - 1} (l_\lambda - 1)!.$$

(b) *If g is decomposable and $\lambda = (\lambda_1 \lambda_2 \cdots)$, then*

$$f_\lambda = f_{\lambda_1} f_{\lambda_2} \cdots \qquad \textit{and} \qquad g_\lambda = g_{\lambda_1} g_{\lambda_2} \cdots.$$

Especially, $\{f_\lambda \mid \lambda\}$ is a $\mathbb{Z}$ basis for $\mathbb{Z}[g_1, g_2, ...]$.

Proof. (a) Notice that in the lattice $\prod_n$, for any σ,

$$\mu(\sigma, \mathbf{1}_n) = (-1)^{b_\sigma - 1} (b_\sigma - 1)!.$$

(b) Notice that as sequence partitions,

$$\lambda = (\lambda_1 \lambda_2 \cdots \lambda_k) = (\lambda_1) + (\lambda_2) + \cdots + (\lambda_k). \quad \blacksquare$$

Therefore, if g_σ on $\prod_\infty$ is decomposable, then both g_σ and its integral f_σ are uniquely determined by two sequences:

$$g_1, g_2, ... \qquad \text{and} \qquad f_1, f_2, ...$$

The relation between their exponential generating functions is revealed by this simple lemma.

LEMMA 6.1.

$$\exp\left(\sum_{n\geqslant 1}\frac{g_n}{n!}t^n\right)=\sum_{\lambda}\frac{g_\lambda t^{|\lambda|}}{\lambda!\,m_\lambda!}.$$

It can be checked directly by expansion. Comparing it to the preceding theorem, we have thus established

COROLLARY 6.1. *Suppose function g_σ on $\prod_\infty$ is decomposable. Then the exponential generating functions of g_n and its integral f_n satisfy*

$$\sum_{n\geqslant 0}\frac{f_n}{n!}t^n=\exp\left(\sum_{n\geqslant 1}\frac{g_n}{n!}t^n\right).$$

Therefore, f_n's and g_n's in fact embody the exact relation between moments and cumulants, which eventually leads to the fourth formula for cumulants.

THEOREM 6.2. *Let K_n's be the cumulants of a moment sequence (a_n). Then*

$$a_n=\sum_{\lambda\vdash n}\binom{n}{\lambda}^* K_\lambda,\qquad K_n=\sum_{\lambda\vdash n}\binom{n}{\lambda}^*(-1)^{l_\lambda-1}(l_\lambda-1)!\,a_\lambda,\tag{6}$$

where $K_\lambda=K_{\lambda_1}K_{\lambda_2}\dots$, $a_\lambda=a_{\lambda_1}a_{\lambda_2}\dots$, if $\lambda=(\lambda_1\lambda_2\cdots)$.

7. FIFTH REPRESENTATION: UMBRAL "FOURIER" TRANSFORM

In this section, we assume that the integral domain D in umbral calculus is the complex number field.

Suppose that umbra α represents a sequence $a_0, a_1, \dots$ and K_n is its nth cumulant. Let

$$f(t)=\sum_{n\geqslant 1}\frac{K_n}{n!}t^n$$

be the exponential generating function of K_n. Then

$$e^{\alpha t} \simeq e^{f(t)}.$$

THEOREM 7.1 ("Fourier" Transform). *For any integer n, let $\omega = \exp(2\pi i/n)$ be the nth unit root. Then,*

$$K_n \simeq \frac{1}{n} \left(\sum_{i=0}^{n-1} \alpha_i \omega^i \right)^n \tag{7}$$

for any n distinct umbrae $\alpha_0, \alpha_1, ..., \alpha_{n-1}$ exchangeable with α.

Proof. For each $i = 0, 1, ..., n-1$,

$$e^{\alpha_i \omega^i t} \simeq e^{f(\omega^i t)}.$$

Since $e^{\alpha_i \omega^i t}$, $i = 0, 1, ..., n-1$, are unrelated, we have

$$\exp\left(\sum_{i=0}^{n-1} \alpha_i \omega^i t \right) \simeq \exp\left(\sum_{i=0}^{n-1} f(\omega^i t) \right) = \exp\left(n \left(\frac{K_n}{n!} t^n + \frac{K_{2n}}{(2n)!} t^{2n} + \cdots \right) \right). \tag{8}$$

Comparing the nth power coefficients of both sides yields Eq. (7). ∎

Let $\mathbb{N}$ denote the set of all non-negative integers. For any fixed positive integer n and an n-tuple $I = (I_1, ..., I_n) \in \mathbb{N}^n$, define $\int_n I$ to be the weighted sum

$$\int_n I = \sum_{i=0}^{n-1} i I_i.$$

For any sequence partition $\lambda = (\lambda_1, \lambda_2, ..., \lambda_k) \vdash n$, and $I = (I_1, I_2, ..., I_n) \in \mathbb{N}^n$, we write $I \vdash \lambda$ if as multisets,

$$\{I_1, I_2, ..., I_n\} = \{\lambda_1, \lambda_2, ..., \lambda_k, 0^{n-k}\}.$$

COROLLARY 7.1. *For any sequence partition $\lambda \vdash n$,*

$$\frac{1}{n} \sum_{I \in \mathbb{N}^n, I \vdash \lambda} \omega^{\int_n I} = \frac{(-1)^{l_\lambda - 1} (l_\lambda - 1)!}{m_\lambda !}. \tag{9}$$

Proof. By Eq. (7),

$$K_n \simeq \frac{1}{n}(\alpha_0\omega^0+\alpha_1\omega^1+\cdots+\alpha_{n-1}\omega^{n-1})^n$$

$$=\frac{1}{n}\sum_{I\in\mathbb{N}^n,\ |I|=n}\binom{n}{I}\alpha_0^{I_0}\alpha_1^{I_1}\cdots\alpha_{n-1}^{I_{n-1}}\omega^{\int_n I}$$

$$\simeq\sum_{\lambda\vdash n}\binom{n}{\lambda}\left(\frac{1}{n}\sum_{I\in\mathbb{N}^n,\ I\vdash\lambda}\omega^{\int_n I}\right)a_\lambda.$$

Comparing the last expression with Eq. (6), we obtain the identity (9). ∎

For any given positive integer d, let ω denote the dth unit root and θ_d any umbra exchangeable with

$$\alpha_0\omega^0+\alpha_1\omega^1+\cdots+\alpha_{d-1}\omega^{d-1}.$$

Equation (8) provides us with the following information:

- $K_n(\theta_d)=0$ unless when $d\,|\,n$, $K_n=dK_n(\alpha)$.
- $\theta_d^n\simeq 0$ unless $d\,|\,n$.

Therefore, Theorem 6.2 generalizes to

THEOREM 7.2. *For any non-negative integer m,*

$$\theta_d^{dm}\simeq\sum_{\lambda\vdash m}d^{l_\lambda}\binom{dm}{d\lambda}^* K_{d\lambda}, \tag{10}$$

where $d\lambda=(d\lambda_1, d\lambda_2, ..., d\lambda_k)$ if $\lambda=(\lambda_1, ..., \lambda_k)$; and the inversion is given by

$$K_{dm}\simeq\frac{1}{d}\sum_{\lambda\vdash m}\binom{dm}{d\lambda}^*\theta_d^{d\lambda}(-1)^{l_\lambda-1}(l_\lambda-1)!. \tag{11}$$

EXAMPLE 3. Suppose we want to estimate the dth cumulant K_d of a real random variable X. Let ω_d be the dth unit root. Assume that X_0,

$X_1, \ldots, X_{d-1}$ are available d independent occurrences of X. To estimate K_d, we define our statistic T_d to be

$$T_d = \frac{1}{d}(X_0 + X_1\omega + \cdots + X_{d-1}\omega^{d-1})^d.$$

By Theorem 7.1, T_d is an unbiased estimator of K_d. According to the last theorem, the standard error SE is

$$\begin{aligned}
\mathrm{SE} = \sigma(T_d) &= (ET_d^2 - K_d^2)^{1/2} \\
&= \left(\frac{1}{d^2}\left\{d\binom{2d}{2d} K_{2d} + d^2\binom{2d}{d}\frac{1}{2!}K_d^2\right\} - K_d^2\right)^{1/2} \\
&= \left(\left(\frac{1}{2}\binom{2d}{d} - 1\right)K_d^2 + \frac{K_{2d}}{d}\right)^{1/2}.
\end{aligned}$$

Therefore, the standard error contains the information of K_{2d}.

8. A FINAL NOTE FROM THE SECOND AUTHOR

The paper is "distilled" from a longer draft by the authors and was rewritten by the second author after Gian-Carlo's unexpected departure from his beloved world, and his little umbral pets: $\alpha, \beta, \ldots$. Gian-Carlo's initial intention was to understand the positivity property of the moments and cumulants of a real random variable by Umbral Calculus. His major conjecture is

Conjecture. Any positive quantity (i.e., polynomial functions of the moments) of a real random variable can be expressed as a sum of squares of umbral polynomials.

This is different from Hilbert's problem. Umbrae certainly provide more flexibility. The conjecture was very much inspired by the following observation. A real unital sequence (a_n) is non-negative (i.e., can be realized as the moment sequence of a real random variable) if and only if all the Hankel determinants of (a_n) are non-negative. It is not difficult to show that the nth order Hankel determinant is precisely (up to a multiplicative constant) the evaluation of the squared Vandemonde

$$[V(\alpha_1, \alpha_2, \ldots, \alpha_n)]^2,$$

where $\alpha_1, \ldots, \alpha_n$ are distinct umbrae representing the sequence. Expressing cumulants (and even orthogonal polynomials) by umbrae is believed to be a first step. The present paper partially fulfills this goal.

Mixed into the Chinglish writing of the present paper is my evergreen memory of Gian-Carlo, my dear mentor and friend.

REFERENCES

1. N. I. Akhiezer, “The Classical Moment Problem” (N. Kemme, Transl.), Hafner, New York, 1965.
2. C. Berg, J. P. R. Christensen, and C. U. Jensen, A remark on the multidimensional moment problem, *Math. Ann.* **223** (1979), 163–169.
3. D. R. Brillinger, “Time Series, Data Analysis and Theory,” Holt, Rinehart & Winston, New York, 1975.
4. B. V. Gnedenko and A. N. Kolmogorov, “Limit Distributions for Sums of Independent Random Variables,” Addison–Wesley, Reading, MA, 1954.
5. D. Hilbert, “Theory of Algebraic Invariants,” Cambridge Univ. Press, Cambridge, UK, 1993.
6. S. A. Joni and G.-C. Rota, Coalgebras and bialgebras in combinatorics, *Stud. Appl. Math.* **61** (1979), 93–139.
7. J. P. S. Kung and G.-C. Rota, The invariant theory of binary forms, *Bull. Amer. Math. Soc.* **10** (1984), 27–85.
8. H. J. Landau (Ed.), “Moments in Mathematics,” Proceedings of Symposia in Applied Mathematics, Vol. 37, Amer. Math. Soc., Providence, RI, 1987.
9. S. M. Roman and G.-C. Rota, The umbral calculus, *Adv. in Math.* **27**, No. 2 (1978), 95–188.
10. G.-C. Rota, The number of partitions of a set, *Amer. Math. Monthly* **71**, No. 5 (1964), 498–504.
11. G.-C. Rota, D. Kahaner, and A. Odlyzko, On the foundations of combinatorial theory. VIII. Finite operator calculus, *J. Math. Anal. Appl.* **42** (1973), 684–760.
12. G.-C. Rota, J. Shen, and B. Taylor, All polynomials of binomial type are represented by Abel polynomials, *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)* **25** (1997).
13. G.-C. Rota and B. D. Taylor, The classical umbral calculus, *SIAM J. Math. Anal.* **25**, No. 2 (1994), 694–711.
14. J. Shen, Combinatorics for wavelets: The umbral refinement equation, *Stud. Appl. Math.* **103**, No. 2 (1999), 121–147.
15. T. P. Speed, Cumulants and partition lattices, *Austral. J. Statist.* **25**, No. 2 (1983), 378–388.
16. G. Strang, “Introduction to Applied Mathematics,” Wellesley–Cambridge Press, Wellesley, MA, 1986.