Journal of Mathematical Research & Exposition
Vol.24, No.3, 391–399, August, 2004

A Probabilistic Interpretation to Umbral Calculus *

SUN Ping[1], *WANG Tian-ming*[2]

(1. Dept. of Math., Northeastern University, Shenyang 110004, China;
2. Dept. of Math., Dalian University of Technology, Liaoning 116024, China)

Abstract: A natural simple interpretation to the umbral calculus is given. As an application, a simple umbral proof of Abel identity is obtained.

Key words: Umbral calculus; moment of random variable; Abel identity.

Classification: AMS(2000) 05A40/CLC number: O156.41

Document code: A **Article ID:** 1000-341X(2004)03-0391-09

1. Introduction

The umbral calculus, as a simple device, yields insight into a number of formulas and often suggests the right proofs in combinatorics. But this kind of proofs have an air of witchcraft over them. Mathematicians view the umbral calculus as a baffling device. Sylvester held the umbral calculus in high esteem but made no attempt to present it; E.T.Bell attempted to display the full power of the methods and to give it a presentation that would meet the standards of algebraic rigor of the twentieth century, but he had not succeeded.

There exists unavoidable difficulty to establish rigorous mathematical theory of the umbral calculus. The main reason is that people don't know what mathematical meaning hidden behind the umbral calculus.

G.-C.Rota first considered that the umbral calculus can be displayed by linear functional on the space of polynomials[1], but as he stated "could not alone explain the stunts of calculation the stunts of calculations performed by the umbral mathematicians"[2]. Rota's another effort is to introduce the language of Hopf algebras, and the newly found rigorous language made the method altogether unwieldy and unmanageable[3]. About the presentation of the umbral calculus, S.Roman follows the point of view in [3] and published a monograph in 1980'[4].

Until 1994 Rota and Taylor presented an interpretation of the umbral calculus, which satisfies rigor of mathematics and is more simple and appliable[2]. Our presentation mostly

*Received date: 2001-03-09
Foundation item: Supported by the Mathematical Tianyuan Foundation (A0324645)
Biography: SUN Ping (1968-), Ph.D., Associate Professor.

coincides with the rigorous presentation given by Rota and can be viewed as one kind of concrete realization of the Rota and Taylor's presentation.

The first baffling difficulty in the calculus of umbral is the strange 'rule', that is if we have two umbrae α, β representing the same sequence $\{\alpha^n \equiv a_n, \beta^n \equiv a_n, n \geq 0\}$, then the umbra $\alpha + \beta$ represents the sequences $\{\sum_{k=0}^{n} \binom{n}{k} a_k a_{n-k},\ n \geq 0\}$ instead of the sequence $\{2^n a_n, n \geq 0\}$. People have not found a reasonable ground for this rule.

It is necessary to give umbral calculus a simple natural explanation because the calculus is an effective device and a simple skill in proving and computing but just lack of a legal mathematical coat.

Instead of develop of a whole of rigorous theory of umbral calculus, that may appeal to a expansion of the probability space, we present a point of view that the umbral calculus (especially in proof and computation of combinatorial sums) may be probabilistic calculus without mathematical expectation symbol.

The reasons are as follows:

1. Many combinatorial numbers and polynomials have probabilistic representations, which are moments of random variables. Among them there are well known Stirling numbers, Bell numbers, Bernoulli numbers, and Hermite polynomials, Euler polynomials and so on.

We can use the representations to compute combinatorial summations and to prove combinatorial identities, especially to discover identities involving this special numbers and orthogonal polynomials effectively. The calculus process omitting the symbol of mathematical expectation is just standard umbral calculus[5]. For example, the factorial sequence $\{n!, n \geq 0\}$ is just sequence $\{\mathbf{E}X^n, n \geq 0\}$, where X is a random variable obeying distribution $\Gamma(1, 1)$ (see Section 4), $\mathbf{E}$ is the symbol of mathematical expectation.

2. In this presentation of umbrae, it is easy and natural to explain the strange rule without introducing heavy mathinery[2]. For given sequence $\{a_n, n \geq 0\}$ if there is a random variable X such that $a_n = \mathbf{E}X^n, n \geq 0$, then two different umbrae representing the same sequence can be understood as two independent random variables X_1 and X_2 with the identical distribution of X. Naturally $X_1 + X_2$ has a probabilistic meaning different from $2X_1$(much less $\mathbf{E}X_1^n = \mathbf{E}X_2^n$ for all $n \geq 0$ can not guarantee that X_1 and X_2 are identically distributed in probability theory). So that the sequence represented by X_1+X_2 must be $\{\mathbf{E}(X_1 + X_2)^n = \sum_{k=0}^{n} \binom{n}{k} a_k a_{n-k}, n \geq 0\}$ instead $\{\mathbf{E}(2X_1)^n = 2^n a_n, n \geq 0\}$.

3. The presentation maintains simplicity and skill in calculus process. In the last section we shall give a probabilistic (umbral) proof of Abel identity. The reader will find what we have done is just like Riordon did in [6], but is considerably simplified proof when you omit the mathematical expectation symbol. In addition, in [6], umbral calculus was used extensively, no justifications were given. We found probabilistic calculus that can take the place of the most of all umbral calculus used in [6] and [7]. The air of witchcraft hovered over the umbral calculus is just a probabilistic characters of random variables and some skills in probability theory.

4. Our presentation basically agree with the idea of umbral calculus in [1] and [2]. The set of all random variables can be taken as the alphabet set in [2], mathematical expectation as linear functional operator. In fact, mathematical expectation of random variables is a definite integral or a summation and can be viewed as a functional.

We have found the corresponding random variables of the Bernoulli umbra and it is inverse umbra, and the most of umbral calculus in [2] can be taken place by probabilistic calculus.

For example, the Bernoulli umbra β is the random variable $X = iL_e - \frac{1}{2}$, where $i^2 = -1, L_e = \sum_{k\geq 1} \frac{L_k}{2k\pi}, L_1, L_2, \cdots$ are independent identically distributed (abbreviated as i.i.d) random variables and their common distribution is Laplace distribution $L(0,1)$. The inverse of Bernoulli umbra γ is a random variable u which obeys the uniform distribution $U(0,1)$ ($u \sim U(0,1)$ in notion). And for any function $f(x)$ we have[5]

$$f^n(x) = \mathbf{E}[f(x) + X + u]^n, \quad n \geq 0. \tag{1.1}$$

Then

(i) Theorem 4.2 in [2] can be translated into probabilistic language as follows.

$$\mathbf{E}(X+1)^n = \mathbf{E}(-X)^n, \quad n \geq 0,$$

that is

$$\mathbf{E}(iL_e + \frac{1}{2})^n = \mathbf{E}(-iL_e + \frac{1}{2})^n, \quad n \geq 0,$$

this is obvious because the random variable L_e is symmetric.

(ii) The simplification of the proof of Proposition 8.3 in [2].

Notice the random variable $u + a \sim U(a, a+1)$ when $u \sim U(0,1)$, and

$$\mathbf{E}\xi^k = \frac{b^{k+1} - a^{k+1}}{(k+1)(b-a)}, \quad k \geq 0, \tag{1.2}$$

where the random variable $\xi \sim U(a,b)$.

Let the random variables X and Y i.i.d and $u \sim U(0,1)$,

$$\begin{aligned}
&\sum_{j=0}^{m-1} B_k(x + \frac{j}{m}) = \sum_{j=0}^{m-1} \mathbf{E}(X + x + \frac{j}{m})^k = m^{-k} \sum_{j=0}^{m-1} \mathbf{E}(mX + mx + j)^k \\
&\overset{(*)}{=} m^{-k} \sum_{j=0}^{m-1} \mathbf{E}(mX + Y + mx + j + u)^k \\
&\overset{(**)}{=} m^{-k}\mathbf{E}\Big[\sum_{j=0}^{m-1} \frac{(mX + Y + mx + j + 1)^{k+1} - (mX + Y + mx + j)^{k+1}}{k+1}\Big] \\
&= m^{-k}\mathbf{E}\Big[\frac{(mX + Y + mx + m)^{k+1} - (mX + Y + mx)^{k+1}}{k+1}\Big] \\
&= m\mathbf{E}\Big[\frac{(X + \frac{Y}{m} + x + 1)^{k+1} - (X + \frac{Y}{m} + x)^{k+1}}{k+1}\Big] \overset{(**)}{=} m\mathbf{E}(X + \frac{Y}{m} + x + u)^k \\
&\overset{(*)}{=} m\mathbf{E}(\frac{Y}{m} + x)^k = m^{-k+1}\mathbf{E}(Y + mx)^k = m^{-k+1}B_k(mx),
\end{aligned}$$

where the first and the last equations hold because of the probabilistic representation of Bernoulli polynomials (see [5]), the steps $(*)$ and $(**)$ hold by (1.1) and (1.2) respectively.

2. The probabilistic umbra

In this paper, the object such as combinatorial numbers and polynomials are called combinatorial variables.

Definition 2.1 *$\{f_n(x), n \geq 0\}$ is said to be a probabilistic umbra, if the combinatorial variable sequences $\{f_n(x), n \geq 0\}$ satisfy the condition that there is a closed form $G(n)$ such that*

$$f_n(x) = G(n)\mathbf{E}[X^n(x)], \quad n \geq 0, \tag{2.1}$$

where $X(x)$ is a random variable with parameter x, $G(n)$ is closed form which means $G(n+1)/G(n)$ is a rational function of n.

$\{f_n(x), n \geq 0\}$ is said to be a probabilistic umbral polynomial when the right side of equation (2.1) is replaced with a polynomial whose undetermined variables are $G_i(n)\mathbf{E}[X_i^n(x)]$, $i \geq 1$.

The following combinatorial variables (see [8]) are probabilistic umbrae, presented in [5]:

1. Two kinds of Stirling numbers,
2. Bell numbers,
3. Harmonic numbers,
4. Fibonacci numbers,
5. The numbers of derangements,
6. Two kinds of Cauchy numbers,
7. Bernoulli numbers,
8. Euler polynomials,
9. Hermite polynomials,
10. Gegenbauer polynomials.

Remark

1. The reason $G(n)$ appears in (2.1) is as follows.

For some combinatorial variables, such as Harmonic numbers $H_n = \sum_{k=1}^{n} \frac{1}{k} = n\mathbf{E}(1-u_1u_2)^{n-1}, n \geq 1$, where u_1, u_2 i.i.d $\sim U(0,1)$, if we directly denote an umbra $\alpha : \alpha^{n-1} = H_n$ for $n \geq 1$ it seems no useful for the summation involving H_n. This may be the reason that umbral calculus was seldom used to treat these combinatorial variables in papers. However if the probabilistic umbra α is denoted as $\alpha^{n-1} = \frac{H_n}{n}, n \geq 1$, for the summation

$$S_n = \sum_k a_{n,k} H_k$$

let us consider the following two kinds of operations:

(i) $\sum_k a_{n,k} H_k = \sum_k k a_{n,k} \alpha^{k-1} = A_n(\alpha)$,

$A_n(\alpha)$ can be simplified by using the properties of umbra α and S_n is solved, which is familiar classical umbral calculus.

(ii) $\sum_k a_{n,k} H_k = \sum_k k a_{n,k} \mathbf{E}(1-u_1u_2)^{k-1} = \mathbf{E}[\sum_k k a_{n,k}(1-u_1u_2)^{k-1}] = \mathbf{E}[A_n(X)]$, where X is the random variable $1 - u_1u_2$.

It can be seen that the two operations (i) and (ii) are almost the same except the last step, by using the properties of umbra α to simplify $A_n(\alpha)$ is equivalently replaced by using the properties of random variable X to compute the mathematical expectation of $A_n(X)$ in (ii).

2. Relating to umbral calculus, we introduce a concept of probabilistic umbral calculus, a linear functional on an integral domain of polynomials whose undetermined elements are random variables and coefficients are the closed form functions. In fact the linear functional is mathematical expectation.

Comparing with the foundation of umbral calculus in [2] we have:

(i) The argumentation ϵ for which $\text{eval}(\varepsilon^n) = \delta_{n,0}$ where $\delta_{n,0}$ is the Kronecker delta does not exit in our representation. In other words, if an operator Y is determined as an inverse of random variable $X : \mathbf{E}(X+Y)^n = \delta_{n,0}$, Y may not be a random variable. This is a stumbling block to a rigorous presentation of the umbral calculus because the concept of inverse umbra should not be abandoned in view of practice.

(ii) The concept of exchangeability just is i.i.d of random variables, so that the term of umbrally equivalent is not necessary to be introduced.

3. The above ten examples are some combinatorial variables which are probabilistic umbrae and only need to deal with the most common random variables. We believe that all combinatorial variables are probabilistic umbrae but some of them are not convenient in use because the form of some $G_i(n)$ in (2.1) is more complex or the probabilistic characteristics of the related random variables $X_i(x)$ are known more less. In this case we would rather use other methods. The situation is the same with umbral calculus, that is the reason that umbral calculus is not all-powerful and is only effective in some special cases. To the best of our knowledge the most of umbral calculus which treats the summation such as the restrictive condition of summation is partition of integer can be equivalently replaced by probabilistic umbra. More importantly, when the probabilistic presentation of a combinatorial variable has been discovered, we shall have more understanding and further inspection about it. The properties of the combinatorial variable are revealed to us by the random variable, so that we can obtain some surprising results easily. For a instance, a new method computing the Riemann zeta function $\zeta(2n)$ can be discovered only by the probabilistic presentation of Bernoulli numbers[5].

3. The Hamburger moment problem

From the point of practicing, it is important and difficult to find the best probabilistic presentation of a combinatorial variable, because the presentation is not unique. Fortunately, we have found a lot of famous combinatorial numbers and orthogonal polynomials are probabilistic umbral polynomials and $G(\cdot)$ in (2.1) is a linear function.

When $G(\cdot) = 1$ it is the Hamburger moment problem in probability theory[9].

Proposition 3.1 *For a number sequence* $\{\mu_n, n \geq 0\}$, *there exits a bounded non-decreasing function* $\varphi(x)$ *such that* $\mu_n = \int_{-\infty}^{\infty} x^n \mathrm{d}\varphi(x), n \geq 0$, *if and only if the quadratic forms* $Q_n(x) = \sum_{i,j=0}^{n} \mu_{i+j} x_i x_j, n \geq 0$ *are non-negative.*

Remark $\int_{-\infty}^{\infty} \mathrm{d}\varphi(x) = 1$ if we demand that $\varphi(x)$ is a probability distribution function, that means $\mu_0 = 1$, the restriction of a_0 in classical umbral calculus[2].

4. A simple proof of Abel identity

Let us remind some basic probability knowledge. A random variable X obeys Gamma distribution $\Gamma(a,1), a>0$ (we write it $X \sim \Gamma(a,1)$) means that

$$\mathbf{P}(X \le y) = \int_0^y p(x)\mathrm{d}x = \int_0^y \frac{x^{a-1}e^{-x}}{\Gamma(a)}\mathrm{d}x, \quad 0 < y < \infty,$$

where $\Gamma(a) = \int_0^\infty x^{a-1}e^{-x}\mathrm{d}x$ and $p(x)$ is called the density function of X.

It is well-known that the moments of order k of random variable X are

$$\mathbf{E}X^k = \int_0^\infty x^k p(x)\mathrm{d}x = a(a+1)\cdots(a+k-1), k \ge 0$$

when $X \sim \Gamma(a,1)$. Therefore

$$\mathbf{E}X^k = k!, k \ge 0, \quad \text{when} \quad X \sim \Gamma(1,1),$$

$$\mathbf{E}X^k = \frac{(n+k-1)!}{(n-1)!}, k \ge 0, \quad \text{when} \quad X \sim \Gamma(n,1), n \in \mathbf{N}.$$

Theorem 4.1 *Suppose the random variable $X \sim \Gamma(1,1)$. We have*

$$\sum_{k=0}^{n} \binom{n}{k} (x-kt)^k (y+kt)^{n-k} = \mathbf{E}(x+y-tX)^n, \quad n \ge 0, \tag{4.1}$$

where $0^0 = 1$ and x, y, t are arbitrary .

Proof

$$\begin{aligned}\text{The LHS of (4.1)} &= \sum_{k=0}^{n}\binom{n}{k}(x-kt)^k[x+y-(x-kt)]^{n-k}\\ &= \sum_{k=0}^{n}\binom{n}{k}(x-kt)^k\sum_{j=0}^{n-k}(-1)^j\binom{n-k}{j}(x+y)^{n-k-j}(x-kt)^j\\ &= \sum_{k=0}^{n}\sum_{r=k}^{n}\binom{n}{r}\binom{r}{k}(-1)^{r-k}(x+y)^{n-r}(x-kt)^r\\ &= \sum_{r=0}^{n}\binom{n}{r}(-1)^r(x+y)^{n-r}\Big[\sum_{k=0}^{r}\binom{r}{k}(-1)^k(x-kt)^r\Big].\end{aligned}$$

Notice

$$\sum_{k=0}^{r}\binom{r}{k}(-1)^{r-k}k^j = \Big[\frac{t^j}{j!}\Big]\sum_{k=0}^{r}\binom{r}{k}(-1)^{r-k}e^{kt} = \Big[\frac{t^j}{j!}\Big](e^t-1)^r = \begin{cases} 0 & j<r, \\ r! & j=r, \end{cases}$$

where $[t^n]f(t)$ means the coefficient of t^n in the formal series $f(t)$, and in fact we know this summation is equal to $r!S(j,r)$, $S(j,r)$ is Stirling number of the second kind. Then

$$\sum_{k=0}^{r}\binom{r}{k}(-1)^k(x-kt)^r=\sum_{k=0}^{r}\binom{r}{k}(-1)^k\sum_{j=0}^{r}\binom{r}{j}(-1)^j k^j t^j\cdot x^{r-j}$$
$$=\sum_{j=0}^{r}\binom{r}{j}(-1)^{r-j}t^j x^{r-j}\Big[\sum_{k=0}^{r}\binom{r}{k}(-1)^{r-k}k^j\Big]=r!t^r=\mathbf{E}(tX)^r,$$

where the random variable $X\sim\Gamma(1,1)$. So

$$\text{The LHS of (4.1)}=\sum_{r=0}^{n}\binom{n}{r}(-1)^r(x+y)^{n-r}\mathbf{E}(tX)^r$$
$$=\mathbf{E}\Big[\sum_{r=0}^{n}\binom{n}{r}(x+y)^{n-r}(-tX)^r\Big]=\mathbf{E}(x+y-tX)^n.$$

Remark

1. Variables x,y,t in (4.1) can be random ones. This will broaden our horizons to identities[5].

2. There are many applications of Theorem 4.1. For example, the number of derangements $d(n)$ has the probabilistic presentation:

$$d(n)=\mathbf{E}(X-1)^n,\quad n\geq 0$$

with $X\sim\Gamma(1,1)$. Taking $x+y=-1,t=-1$ in (4.1), we have

$$d(n)=\sum_{k=0}^{n}\binom{n}{k}(-1)^{n-k}(x+k)^k(x+k+1)^{n-k},\tag{4.2}$$

where $0^0=1$ and x is arbitrary, and the recursive relations

$$d(n+1)=(n+1)d(n)+(-1)^{n+1},\quad d(0)=1$$

are implied when $x=0$ and $x=1$ in (4.2).

Using Theorem 4.1 we can give a simple proof of Abel identity.

Proposition 4.2 (Abel) *For any complex numbers x,y,t*

$$(x+y)^n=\sum_{k=0}^{n}\binom{n}{k}x(x-kt)^{k-1}(y+kt)^{n-k}.\tag{4.3}$$

Proof Let $n\geq 1$. Suppose a random variable $X\sim\Gamma(1,1)$

$$\mathbf{E}(x+y-tX)^n+nt\mathbf{E}(x+y-tX)^{n-1}$$
$$=(x+y)^n+n!\sum_{k=1}^{n}\Big[\frac{(x+y)^{n-k}(-t)^k}{(n-k)!}\Big]+n!\sum_{r=0}^{n-1}\Big[\frac{(x+y)^{n-r-1}(-1)^r t^{r+1}}{(n-r-1)!}\Big]$$
$$=(x+y)^n,\tag{4.4}$$

and by Theorem 4.1

$$\begin{aligned}\mathbf{E}(x+y-tX)^{n-1} &= \mathbf{E}[(x-t)+(y+t)-tX)]^{n-1}\\ &= \sum_{r=0}^{n-1}\binom{n-1}{r}[x-(r+1)t]^r[y+(r+1)t]^{n-r-1}.\end{aligned}$$

Thus

$$\begin{aligned}(x+y)^n &= \sum_{k=0}^{n}\binom{n}{k}(x-kt)^k(y+kt)^{n-k} + nt\sum_{r=0}^{n-1}\binom{n-1}{r}[x-(r+1)t]^r[y+(r+1)t]^{n-r-1}\\ &= \sum_{k=0}^{n}\binom{n}{k}(x-kt)^k(y+kt)^{n-k} + \sum_{k=0}^{n}\binom{n}{k}kt(x-kt)^{k-1}(y+kt)^{n-k}\\ &= \sum_{k=0}^{n}\binom{n}{k}x(x-kt)^{k-1}(y+kt)^{n-k}.\end{aligned}$$

By the theory of probability, $X+Y\sim\Gamma(a+b,1)$ when $X\sim\Gamma(a,1), Y\sim\Gamma(b,1)$ and X,Y are independent, therefore Theorem 4.1 can be extended to

Theorem 4.3 *Suppose the random variable $X\sim\Gamma(k-1,1), k\geq 2$,*

$$\begin{aligned}&\sum_{r_1+\cdots+r_k=n}\binom{n}{r_1\cdots r_k}(x_1-r_1t)^{r_1}\cdots(x_{k-1}-r_{k-1}t)^{r_{k-1}}[x_k+(n-r_k)t]^{r_k}\\ &= \mathbf{E}(x_1+\cdots+x_k-tX)^n \qquad (4.5)\end{aligned}$$

is true for arbitrary $x_1,\cdots,x_k,t$.

The following multiple Abel identity can be derived from Theorem 4.3 similarly.

Proposition 4.4 *For any complex numbers $x_1,\cdots,x_k,t$, we have*

$$\begin{aligned}(x_1+\cdots+x_k)^n = &\sum_{r_1+\cdots+r_k=n}\binom{n}{r_1\cdots r_k}x_1(x_1-r_1t)^{r_1-1}\cdots\\ &x_{k-1}(x_{k-1}-r_{k-1}t)^{r_{k-1}-1}[x_k+(n-r_k)t]^{r_k}. \qquad (4.6)\end{aligned}$$

Remark

1. Taking $t=-1, x_k=x_k+n$ in (4.6) and (4.5), we can obtain the first and the second Hurwitz identities[6].

2. One might like to delete the symbol $\mathbf{E}$ of mathematical expectation in this section. Then Theorem 4.1 can be rewritten as

$$\sum_{k=0}^{n}\binom{n}{k}(x-kt)^k(y+kt)^{n-k} = (x+y-t\alpha)^n, n\geq 0,$$

where $0^0=1, x, y, t$ are arbitrary, and α is an umbra $\alpha^k=k!, k\geq 0$, and so on, we will see a standard rigorous umbral calculus.

References:

[1] ROTA G C. *The number of Partition of a set* [J]. Amer. Math. Monthly, 1964, **71**: 498–504.
[2] ROTA G C, TAYLOR B D. *The Classical umbral calculus* [J]. SIAM J. Math. ANAL, 1994, **25**: 694–711.
[3] ROMAN S, ROTA G C. *The umbral calculus* [J]. Adv. Math., 1978, **27**: 95–188.
[4] ROMAN S. *The Umbral Calculus* [M]. Academic Press, New York, 1981.
[5] SUN Ping. *Probabilistic methods in combinatorial analysis* [D]. Ph.D. Thesis, DUT., 1998.
[6] RIORDAN J. *Combinatorial Identities* [M]. John Wiley, New York, 1968.
[7] RIORDAN J. *An Introduction to Combinatorial Analysis* [M]. John Wiley, New York, 1958.
[8] COMTET L. *Advanced Combinatorics* [M]. D.Reidel Pub Com, Boston, 1974.
[9] SHOHAT J A, TAMARKIN J D. *The Problem of Moments (Mathematical Surveys 1)* [M]. A.M.S, New York, 1970.

哑运算的概率解释

孙 平[1], 王 天 明[2]

(1. 东北大学数学系, 辽宁 沈阳 110004; 2. 大连理工大学应用数学系, 辽宁 大连 116024)

摘 要: 利用随机变量的矩以及期望运算, 给出了哑运算一种简单、自然的概率解释, 并且得到了 Abel 恒等式的一个广泛哑运算证明.

关键词: 哑运算; 随机变量; 矩; Abel 恒等式.