

GEOMETRIC AND PROBABILISTIC ASPECTS OF SOME COMBINATORIAL IDENTITIES

T. P. SPEED

(Received June 9, 1975; revised October 20, 1975)

Communicated by W. D. Wallis

Abstract

For positive integers a , b and n define the combinational expression

$$A_n(a, b) = \frac{a}{a + bn} \binom{a + bn}{n}.$$

We give geometric and probabilistic interpretations of these expressions (and their multidimensional extensions) and find new, simple proofs of the convolution identities known to hold for such expressions.

1. Introduction

For non-negative integers a , b with $a + bn > 0$ let us define the combinatorial expression

$$(1) \quad A_n(a, b) = \frac{a}{a + bn} \binom{a + bn}{n}.$$

In two papers written some twenty years ago Gould (1956, 1957) discussed the above (and related) expressions. He obtained, amongst other results, the following convolution identity: for positive integral c

$$(2) \quad \sum_{m=0}^n A_m(a, b) A_{n-m}(c, b) = A_n(a + c, b).$$

Gould's two papers contain different approaches to this identity whilst in his recent article Gould (1974) gives yet another. We also note that Riordan (1958) presents an inductive proof of (2), as do Gould and Kaucky (1966) where further comments and extensions can be found. The proof of Blackwell and Dubins (1966) is perhaps closest in spirit to the one given below. Mohanty (1966a)

extended the argument of Gould's first paper, stating and proving multinomial analogues of (2) and the related identities. He also gave a probabilistic interpretation of these facts. To formulate these results we use bold letters to denote k -tuples of non-negative integers, $\mathbf{b} = (b_1, b_2, \dots, b_k)$ and $\mathbf{n} = (n_1, n_2, \dots, n_k)$. Further we use the usual dot-product notation $\mathbf{b} \cdot \mathbf{n} = b_1 n_1 + b_2 n_2 + \dots + b_k n_k$ and write $\mathbf{1} = (1, 1, \dots, 1)$. With these preliminaries we can extend the notation above when $a + \mathbf{b} \cdot \mathbf{n} > 0$ writing

$$(1') \quad A_{\mathbf{n}}(a, \mathbf{b}) = \frac{a}{a + \mathbf{b} \cdot \mathbf{n}} \binom{a + \mathbf{b} \cdot \mathbf{n}}{\mathbf{n}}$$

where $\binom{N}{\mathbf{n}} = N(N-1) \dots (N - \mathbf{1} \cdot \mathbf{n} + 1) / \prod_1^k n_i!$ denotes the usual multinomial coefficient. In this notation one of Mohanty's results (1966a, equation (9) p. 502), the generalisation of (2) above, can be written

$$(2') \quad \sum_{\mathbf{m}=0}^{\mathbf{n}} A_{\mathbf{m}}(a, \mathbf{b}) A_{\mathbf{n}-\mathbf{m}}(c, \mathbf{b}) = A_{\mathbf{n}}(a + c, \mathbf{b}).$$

Here is the summation from $m_1 = 0$ to $m_1 = n_1, \dots, m_k = 0$ to $m_k = n_k$ as the notation suggests, and $\mathbf{n} - \mathbf{m} = (n_1 - m_1, n_2 - m_2, \dots, n_k - m_k)$.

It is the purpose of this note to provide new proofs of these identities, the first, it is believed, that involve the geometrical interpretation of the expression (1'). After doing this we reconsider the probabilistic aspects of (2'), being somewhat more concrete than Mohanty in obtaining a random walk whose first passage probabilities to a certain hyperplane provide yet another interpretation and proof of (2').

2. Geometric interpretation of $A_{\mathbf{n}}(a, \mathbf{b})$

It is hoped that the notation will enable us to deal with the general case (arbitrary k) almost as easily as one would the case $k = 1$, but this will involve some slightly unusual temporary usages. We will be working in the positive orthant of the integer lattice in $k + 1$ dimensions, the coordinate variables being denoted by X_0, X_1, \dots, X_k and an arbitrary element will be denoted by (x_0, \mathbf{x}) where $\mathbf{x} = (x_1, x_2, \dots, x_k)$. The first coordinate will be treated differently, and all all bold letters will be k -tuples of non-negative integers.

Given any k -tuple \mathbf{b} and non-negative integer a we can define a hyperplane by the equation

$$(P) \quad X_0 = a + (\mathbf{b} - \mathbf{1}) \cdot \mathbf{X}.$$

Clearly the point $(a + (\mathbf{b} - \mathbf{1}) \cdot \mathbf{n}, \mathbf{n})$ lies on (P) for any k -tuple \mathbf{n} , and we may now state the desired interpretation as follows:

PROPOSITION 1 (Mohanty). *The number of minimal lattice paths from $(0, 0)$ to $(a + (b - 1) \cdot n, n)$ which do not touch the plane (P) until the last point, is $A_n(a, b)$.*

For the case $k = 1$ this result is implicit in Mohanty and Narayana (1961) (following by duality from their Corollary on p. 256), and appears in the present generality in Mohanty (1972). The following proof is essentially Mohanty's but we include it for completeness.

PROOF. The minimal lattice paths from $(0, 0)$ to $(a + (b - 1) \cdot n, n)$ can be put into one-one correspondence with N -tuples $L = (\lambda_1, \lambda_2, \dots, \lambda_N)$, where $N = a + b \cdot n$; for each $i, 1 \leq i \leq N$, λ_i is one of the symbols S_0, S_1, \dots, S_k ; for each $j, 1 \leq j \leq k$ there are precisely n_j symbols S_j , and there are $a + (b - 1) \cdot n$ symbols S_0 .

Given such an N -tuple L we can build up a minimal lattice path, starting at either end, by interpreting a symbol S_j to mean 'move one unit along the X_j -axis towards the other end'. Conversely any minimal lattice path defines a unique such N -tuple in the obvious way.

It is also clear that there are precisely $\binom{N}{n}$ such N -tuples and so this is the total number of minimal lattice paths connecting $(0, 0)$ with $(a + (b - 1) \cdot n, n)$. But we want the number of these which do not touch the plane (P) other than at the last point. To express this requirement as a property of the N -tuple L we need a little more notation. For each $h, 1 \leq h \leq N$ and $j, 0 \leq j \leq k$ define

$$\sigma_{hj} = \begin{cases} 1 & \text{if } \lambda_h = S_j \\ 0 & \text{otherwise.} \end{cases}$$

Clearly $\sum_{j=0}^k \sigma_{hj} = 1$. Also put $\xi_{ij} = \sum_{h=1}^i \sigma_{hj}$, this being the number of times the symbol S_j appears in the first i positions of L , and finally write $\xi_i = (\xi_{i1}, \xi_{i2}, \dots, \xi_{ik})$.

We will build up the lattice path by working backwards from the endpoint using L . After i steps have been incorporated, the X_j coordinate has reduced by $\xi_{ij} (0 \leq j \leq k)$ and so we are at the point $(a + (b - 1) \cdot n - \xi_{i0}, n - \xi_i)$. This point lies in the half-space defined by (P) which contains the origin for all $i, 1 \leq i \leq N$, if, and only if,

$$a + (b - 1) \cdot n - \xi_{i0} < a + (b - 1) \cdot (n - \xi_i) \quad (1 \leq i \leq N),$$

equivalently, upon expanding and using the fact that $\xi_{i0} + 1 \cdot \xi_i = i$, if and only if

$$(C) \quad b \cdot \xi_i < i \quad (1 \leq i \leq N).$$

Now $b \cdot \xi_i = \sum_{j=1}^k b_j \xi_{ij}$ is simply a partial sum along L of numerical terms if we replace S_j by the integer $b_j, 1 \leq j \leq k$, and S_0 by 0. With this interpretation we can immediately recognise the condition (C) and use a well-known result to

deduce that of the N cyclic permutations of the N -tuple L (with the numerical components just indicated), precisely $a = N - b \cdot n$ have the property (C); that is, satisfy the condition that for all i ($1 \leq i \leq N$) the partial sums of the first i terms are less than i .

We refer to Takacs (1967) p. 4 for this result; for a geometric proof more in the spirit of the present paper, see Mohanty (1966b).

This completes the proof that the number of minimal lattice paths from $(0, 0)$ to $(a + (b - 1) \cdot n, n)$ not touching (P) before their endpoint is $\frac{a}{N} \binom{N}{n}$ where $N = a + b \cdot n$.

3. Derivation of identities

Let us consider the hyperplane (P) defined above, and the parallel hyperplane (c being another non-negative integer)

$$(P') \quad X_0 = a + c + (b - 1) \cdot X.$$

Clearly any minimal lattice path from $(0, 0)$ to $(a + c + (b - 1) \cdot n, n)$ on (P') must hit (P) for the first time at $X = m$ for some m , $0 \leq m \leq n$. Indeed there are precisely $A_m(a, b)$ such paths. Each can be completed in $A_{n-m}(c, b)$ ways, as can be seen by viewing (P') relative to the coordinate system (X'_0, X') where $X'_0 = X_0 - a$ and $X' = X - m$. This, plus an obvious counting argument, completes the proof of (2').

Another identity which can be derived in a similar way is:

$$(3) \quad \sum_{m=0}^n A_m(a, b) A_{n-m}(d \cdot m, b + d) = A_n(a, b + d).$$

To get this one we consider the hyperplane (P) and the 'steeper' plane (P'') having the same X_0 -intercept viz:

$$(P'') \quad X_0 = a + (b + d - 1) \cdot X.$$

Any minimal lattice path from $(0, 0)$ to $(a + (b + d - 1) \cdot n, n)$ on (P'') must hit the hyperplane (P) for the first time at $X = m$ for some m , $0 \leq m \leq n$. Again there are $A_m(a, b)$ such, and each can be completed in $A_{n-m}(d \cdot m, b + d)$ ways, as we can see by viewing (P'') relative to the coordinate system (X''_0, X'') where $X''_0 = X_0 - a - (b - 1) \cdot m$, $X'' = X - m$. Thus (3) follows in the same way as (2').

The general identity in Mohanty (1966a) is seen to be a combination of (2') and (3). Another identity derived in Gould's papers involves the expressions

$$(4) \quad A_n(a, b) = \frac{a}{a + bn} \frac{(a + bn)^n}{n!}.$$

Gould (1957 equation 6) shows that (2) holds with this definition of $A_n(a, b)$ and one might wonder whether a geometric interpretation exists for the entities

(4) similar to that derived for (2). I have been unable to find such an interpretation although a probabilistic one exists, and Raney (1964) gives the combinatorial interpretations of closely related expressions which lead to the proof of (2) in this case. The definition (4) also suggests a generalisation not discussed by Mohanty, namely

$$(4') \quad A_n(a, b) = \frac{a}{a + b \cdot n} \frac{(a + b \cdot n)^{1 \cdot n}}{n!}$$

where $n! = n_1! n_2! \cdots n_k!$ and $1 \cdot n = n_1 + n_2 + \cdots + n_k$. The coefficients $A_n(a, b)$ defined by (1') approximate those defined in (4') when a and b are large so it is reasonable to suppose that the convolution identity (2') also holds in this case. This is indeed true, the result being deducible (with a little effort) from Raney (1964).

4. Associated probability distributions

Let (p_0, p) be a $(k + 1)$ -tuple with $p_0 > 0, p_1 > 0, \dots, p_k > 0$ and $p_0 + p_1 + \dots + p_k = 1$. Then if $b \cdot p \leq 1$, Mohanty (1966a) proved that

$$(5) \quad \sum_{n=0}^{\infty} A_n(a, b) p_0^{a+(b-1) \cdot n} p^n = 1$$

where $p^n = p_1^{n_1} p_2^{n_2} \cdots p_k^{n_k}$. We will offer an alternative derivation of (5) based upon a random walk interpretation. To do this we consider the random walk on the lattice points in the positive orthant in $(k + 1)$ -space which begins at $(0, \mathbf{0})$ and at each step moves along the X_j axis one unit in the positive direction with probability p_j ($0 \leq j \leq k$), steps being mutually independent and identically constructed.

PROPOSITION 2. *The probability that the above random walk ever hits the hyperplane (P) is π^a , where π is the smallest positive root of*

$$(6) \quad \sum_1^k p_j x^{b_j} - x + p_0 = 0.$$

PROOF. Let us define the function, in fact a probability generating function:

$$(7) \quad f(x) = p_0 + \sum_1^k p_j x^{b_j}$$

We will see that the probability that the walk ever hits (P) is π^a where π is a probability that the walk ever hits (P) when $a = 1$, and that π is the smallest positive root of the equation $f(x) = x$. The first assertion is an immediate consequence of the assumed independence of the steps in the walk, as the passage from $(0, \mathbf{0})$ to (P) can be viewed as a succession of independent and

probabilistically identical passages from $(0, \mathbf{0})$ to $X_0 = 1 + (\mathbf{b} - \mathbf{1}) \cdot \mathbf{X}$, from this plane to $X_0 = 2 + (\mathbf{b} - \mathbf{1}) \cdot \mathbf{X}$, and so on up to (P) .

Let $\pi_a(m)$ denote the probability that the walk hits the hyperplane (P) in less than m steps. Clearly $\pi_1(m) \uparrow \pi$ as $m \rightarrow \infty$. If $m > 1$ we may condition upon the outcome of the first (random) step and find that

$$(8) \quad \pi_1(m) = p_0 + \sum_1^k p_i \pi_{b_i}(m-1).$$

Now $\pi_{b_i}(m-1) \leq \pi_{b_i}(m) \leq [\pi_1(m)]^{b_i}$ and so we find that $\pi_1(m)$ satisfies the inequality

$$(9) \quad 0 \leq \pi_1(m) \leq f(\pi_1(m)).$$

Letting $m \rightarrow \infty$ we see from (8) and the remarks opening this proof that $\pi = f(\pi)$ and it follows from (9) that π is the smallest such positive root.

COROLLARY. $\pi = 1$ if and only if $\mathbf{b} \cdot \mathbf{p} \leq 1$.

PROOF. This is easily derived using methods well known in the theory of branching processes. See for example Harris (1963).

Let us define T to be random time, possibly infinite, which the walk takes to hit the hyperplane (P) . Then we have the distribution of T involving our coefficients.

PROPOSITION 2. (i) $P(T = a + \mathbf{b} \cdot \mathbf{n}) = A_n(a, \mathbf{b}) p_0^{a + (\mathbf{b} - \mathbf{1}) \cdot \mathbf{n}} \mathbf{p}^{\mathbf{n}}$.

(ii) $P(T < \infty) = \pi^a$ where π is defined above.

PROOF. Result (i) follows from Proposition 1 and the definition of the walk, whereas (ii) follows from the previous proposition.

COROLLARY 2. Identity (5) holds if $\mathbf{b} \cdot \mathbf{p} \leq 1$.

If we denote by T_a the above random variable, then it is probabilistically obvious that the first passage time T_{a+c} should be distributed as the sum of a r.v. T_a and another, independent, r.v. T_c . This convolution property is equivalent to (2') as is easily checked. Thus an alternative, probabilistic, proof of (2') could be constructed. The details are left to the reader.

Finally we note that $E\{T_a\} = a/(1 - \mathbf{b} \cdot \mathbf{p})$ can be proved in a manner analogous to that used to obtain the equation for π . That is, by first deriving the equation $E\{T_a\} = aE\{T_1\}$, and then conditioning upon the outcome of the first step obtaining

$$E\{T_1\} = p_0 + \sum_1^k p_i (1 + E\{T_{b_i}\}).$$

The variance formula for T_a can be derived in a similar way.

5. Acknowledgements

I am indebted to H. W. Gould and G. N. Raney for providing references to some relevant work.

References

- David Blackwell and Lester Dubins (1966), 'An elementary proof of an identity of Gould's', *Bol. Soc. Mat. Mexicana Ser. 2* **11**, 108–110.
- H. W. Gould (1956), 'Some generalisations of Vandermonde's convolution', *Amer. Math. Monthly* **63**, 84–91.
- H. W. Gould (1957), 'Final analysis of Vandermonde's convolution', *Amer. Math. Monthly* **64**, 409–415.
- H. W. Gould (1974), 'Coefficient identities for powers of Taylor and Dirichlet series', *Amer. Math. Monthly* **81**, 3–13.
- H. W. Gould and J. Kaucky (1966), 'Evaluation of a class of binomial coefficient summations', *J. Combinatorial Theory* **1**, 233–247.
- Theodore E. Harris (1963), *The theory of branching processes*. Springer-Verlag Berlin, Göttingen, Heidelberg.
- S. G. Mohanty (1966a), 'Some convolutions with multinomial coefficients and related probability distributions', *SIAM Review* **8**, 501–509.
- S. G. Mohanty (1966b), 'An urn problem related to the ballot problem', *Amer. Math. Monthly* **73**, 526–528.
- S. G. Mohanty (1972), 'On queues involving batches', *J. Applied Probability* **9**, 430–435.
- S. G. Mohanty and T. V. Narayana (1961), 'Some properties of compositions and their application to probability and statistics I', *Biom. Zeit.* **3**, 252–258.
- George N. Raney (1964), 'A formal solution of $\sum_{i=1}^{\infty} A_i e^{B_i X} = X$ ', *Canad. J. Math.* **16**, 755–762.
- John Riordan (1958), *An introduction to combinatorial analysis*. John Wiley and Sons, New York, London, Sydney.
- Lajos Takács (1967) *Combinatorial methods in the theory of stochastic processes*. John Wiley and Sons, Inc., New York, London, Sydney.

Department of Mathematics,
University of Western Australia,
Nedlands, W.A. 6009.