A probabilistic approach to some binomial identities

Victor H. Moll, Christophe Vignat

A PROBABILISTIC APPROACH TO SOME BINOMIAL IDENTITIES

CHRISTOPHE VIGNAT AND VICTOR H. MOLL

Abstract. Classical binomial identities are established by giving probabilistic interpretations to the summands. The examples include Vandermonde identity and some generalizations.

1. Introduction

The evaluation of finite sums involving binomial coefficients appears throughout the undergraduate curriculum. Students are often exposed to the identity

$$\sum_{k=0}^{n} \binom{n}{k} = 2^n. \tag{1.1}$$

Elementary proofs abound: simply choose $x = y = 1$ in the binomial expansion of $(x+y)^n$. The reader is surely aware of many other proofs, including some combinatorial in nature.

At the end of the previous century, the evaluation of these sums was trivialized by the work of H. Wilf and D. Zeilberger [7]. In the preface to the charming book [7], the authors begin with the phrase

`You've been up all night working on your new theory, you found the answer, and it is in the form that involves factorials, binomial coefficients, and so on, ...`

and then proceed to introduce the method of *creative telescoping*. This technique provides an automatic tool for the verification of this type of identities.

Even in the presence of a powerful technique, such as the WZ-method, it is often a good pedagogical idea to present a simple identity from many different points of view. The reader will find in [1] this approach with the example

$$\sum_{k=0}^{m} 2^{-2k} \binom{2k}{k} \binom{2m-k}{m} = \sum_{k=0}^{m} 2^{-2k} \binom{2k}{k} \binom{2m+1}{2k}. \tag{1.2}$$

The current paper presents probabilistic arguments for the evaluation of certain binomial sums. The background required is minimal. The continuous random variables X considered here have a probability density function. This is a nonnegative

Date: July 1, 2018.
1991 *Mathematics Subject Classification.* Primary 05A10, Secondary 33B15, 60C99.
Key words and phrases. binomial sums, gamma distributed random variables, Vandermonde identity, orthogonal polynomials.

function $f_X(x)$, such that

$$\Pr(X < x) = \int_{-\infty}^{x} f_X(y)\,dy. \tag{1.3}$$

In particular, f_X must have total mass 1. Thus, all computations are reduced to the evaluation of integrals. For instance, the expectation of a function of the random variable X is computed as

$$\mathbb{E}g(X) = \int_{-\infty}^{\infty} g(y) f_X(y)\,dy. \tag{1.4}$$

In elementary courses, the reader has been exposed to normal random variables, written as $X \sim N(0,1)$, with density

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}, \tag{1.5}$$

and exponential random variables, with probability density function

$$f(x;\lambda) = \begin{cases} \lambda e^{-\lambda x} & \text{for } x \geq 0; \\ 0 & \text{otherwise.} \end{cases} \tag{1.6}$$

The examples employed in the arguments presented here have a gamma distribution with shape parameter k and scale parameter θ, written as $X \sim \Gamma(k,\theta)$. These are defined by the density function

$$f(x;k,\theta) = \begin{cases} x^{k-1} e^{-x/\theta}/\theta^k \Gamma(k), & \text{for } x \geq 0; \\ 0 & \text{otherwise.} \end{cases} \tag{1.7}$$

Here $\Gamma(s)$ is the classical gamma function, defined by

$$\Gamma(s) = \int_0^{\infty} x^{s-1} e^{-x}\,dx \tag{1.8}$$

for $\operatorname{Re} s > 0$. Observe that if $X \sim \Gamma(a,\theta)$, then $X = \theta Y$ where $Y \sim \Gamma(a,1)$. Moreover $\mathbb{E}X^n = \theta^n (a)_n$, where

$$(a)_n = \frac{\Gamma(a+n)}{\Gamma(a)} = a(a+1)\cdots(a+n-1) \tag{1.9}$$

is the Pochhammer symbol. The main property of these random variables employed in this paper is the following: assume $X_i \sim \Gamma(k_i,\theta)$ are independent, then

$$X_1 + \cdots + X_n \sim \Gamma(k_1 + \cdots + k_n, \theta). \tag{1.10}$$

This follows from the fact that that the density probability function for the sum of two independent random variables is the convolution of the individual ones.

Related random variables include those with a beta distribution

$$f_{a,b}(x) = \begin{cases} x^{a-1}(1-x)^{b-1}/B(a,b) & \text{for } 0 \leq x \leq 1; \\ 0 & \text{otherwise.} \end{cases} \tag{1.11}$$

Here $B(a,b)$ is the beta function defined by

$$B(a,b) = \int_0^1 x^{a-1}(1-x)^{b-1}\,dx \tag{1.12}$$

and also the symmetric beta distributed random variable Z_c, with density proportional to $(1-x^2)^{c-1}$ for $-1 \le x \le 1$. The first class of random variables can be generated as

$$B_{a,b} = \frac{\Gamma_a}{\Gamma_a + \Gamma_b}, \tag{1.13}$$

where Γ_a and Γ_b are independent gamma distributed with shape parameters a and b, respectively and the second type is distributed as $1 - 2B_{c,c}$, that is,

$$Z_c = 1 - \frac{2\Gamma_c}{\Gamma_c + \Gamma'_c} = \frac{\Gamma_c - \Gamma'_c}{\Gamma_c + \Gamma'_c}, \tag{1.14}$$

where Γ_c and Γ'_c are independent gamma distributed with shape parameter c. A well-known result is that $B_{a,b}$ and $\Gamma_a + \Gamma_b$ are independent in (1.13); similarly, $\Gamma_c + \Gamma'_c$ and Z_c are independent in (1.14).

2. A sum involving central binomial coefficients

Many finite sums may be evaluated via the generating function of terms appearing in them. For instance, a sum of the form

$$S_2(n) = \sum_{i+j=n} a_i a_j \tag{2.1}$$

is recognized as the coefficient of x^n in the expansion of $f(x)^2$, where

$$f(x) = \sum_{j=0}^{\infty} a_j x^j \tag{2.2}$$

is the generating function of the sequence $\{a_i\}$. Similarly,

$$S_m(n) = \sum_{k_1+\cdots+k_m=n} a_{k_1} \cdots a_{k_m} \tag{2.3}$$

is given by the coefficient of x^n in $f(x)^m$. The classical example

$$\frac{1}{\sqrt{1-4x}} = \sum_{j=0}^{\infty} \binom{2j}{j} x^j \tag{2.4}$$

gives the sums

$$\sum_{i=0}^{n} \binom{2i}{i} \binom{2n-2i}{n-i} = 4^n \tag{2.5}$$

and

$$\sum_{k_1+\cdots+k_m=n} \binom{2k_1}{k_1} \cdots \binom{2k_m}{k_m} = \frac{2^{2n}}{n!} \frac{\Gamma(\frac{m}{2}+n)}{\Gamma(\frac{m}{2})}. \tag{2.6}$$

The powers of $(1-4x)^{-1/2}$ are obtained from the binomial expansion

$$(1-4x)^{-a} = \sum_{j=0}^{\infty} \frac{(a)_j}{j!} (4x)^j, \tag{2.7}$$

where $(a)_j$ is the Pochhammer symbol.

The identity (2.5) is elementary and there are many proofs in the literature. A nice combinatorial proof of (2.6) appeared in 2006 in this journal [3]. In a more recent contribution, G. Chang and C. Xu [5] present a probabilistic proof of these identities. Their approach is elementary: take m independent Gamma random variables $X_i \sim \Gamma(\frac{1}{2}, 1)$ and write

$$\mathbb{E}\left(\sum_{i=1}^{m} X_i\right)^n = \sum_{k_1+\cdots+k_m=n} \binom{n}{k_1, \cdots, k_m} \mathbb{E}X_1^{k_1} \cdots \mathbb{E}X_m^{k_m} \tag{2.8}$$

where $\mathbb{E}$ denotes the expectation operator. For each random variable X_i, the moments are given by

$$\mathbb{E}X_i^{k_i} = \frac{\Gamma(k_i + \frac{1}{2})}{\Gamma(\frac{1}{2})} = 2^{-2k_i} \frac{(2k_i)!}{k_i!} = \frac{k_i!}{2^{2k_i}} \binom{2k_i}{k_i}, \tag{2.9}$$

using Euler's duplication formula for the gamma function

$$\Gamma(2z) = \frac{1}{\sqrt{\pi}} 2^{2z-1} \Gamma(z) \Gamma(z + \tfrac{1}{2}) \tag{2.10}$$

(see [6], 5.5.5) to obtain the second form. The expression

$$\binom{n}{k_1, \cdots, k_m} = \frac{n!}{k_1!\, k_2! \cdots k_m!} \tag{2.11}$$

for the multinomial coefficients shows that the right-hand side of (2.8) is

$$\frac{n!}{2^{2n}} \sum_{k_1+\cdots+k_m=n} \binom{2k_1}{k_1} \cdots \binom{2k_m}{k_m}. \tag{2.12}$$

To evaluate the left-hand side of (2.8), recall that the sum of m independent $\Gamma\left(\frac{1}{2}, 1\right)$ has a distribution of $\Gamma(\frac{m}{2}, 1)$. Therefore, the left-hand side of (2.8) is

$$\frac{\Gamma(\frac{m}{2} + n)}{\Gamma(\frac{m}{2})}. \tag{2.13}$$

This gives (2.6). The special case $m = 2$ produces (2.5).

3. More sums involving central binomial coefficients

The next example deals with the identity

$$\sum_{k=0}^{n} \binom{4k}{2k} \binom{4n-4k}{2n-2k} = 2^{4n-1} + 2^{2n-1} \binom{2n}{n} \tag{3.1}$$

that appears as entry 4.2.5.74 in [4]. The proof presented here employs the famous dissection technique, first introduced by Simpson [8] in the simplification of

$$\frac{1}{2} \left(\mathbb{E}(X_1 + X_2)^{2n} + \mathbb{E}(X_1 - X_2)^{2n} \right), \tag{3.2}$$

where X_1, X_2 are independent random variables distributed as $\Gamma\left(\frac{1}{2}, 1\right)$.

The left-hand side is evaluated by expanding the binomials to obtain

$$\frac{1}{2}(\mathbb{E}(X_1 + X_2)^{2n} + \mathbb{E}(X_1 - X_2)^{2n}) =$$
$$\frac{1}{2} \sum_{k=0}^{2n} \binom{2n}{k} \mathbb{E}X_1^k \, \mathbb{E}X_2^{2n-k} + \frac{1}{2} \sum_{k=0}^{2n} (-1)^k \binom{2n}{k} \mathbb{E}X_1^k \, \mathbb{E}X_2^{2n-k}$$

This gives

$$\frac{1}{2}(\mathbb{E}(X_1+X_2)^{2n}+\mathbb{E}(X_1-X_2)^{2n}) = \sum_{k=0}^{n} \binom{2n}{2k} \mathbb{E}X_1^{2k}\,\mathbb{E}X_2^{2n-2k}.$$

Using (2.9), this reduces to

$$\frac{1}{2}\left(\mathbb{E}(X_1+X_2)^{2n}+\mathbb{E}(X_1-X_2)^{2n}\right) = \frac{(2n)!}{2^{4n}} \sum_{k=0}^{n} \binom{4k}{2k}\binom{4n-4k}{2n-2k}. \tag{3.3}$$

The random variable X_1+X_2 is $\Gamma(1,1)$ distributed, so

$$\mathbb{E}(X_1+X_2)^{2n} = (2n)!, \tag{3.4}$$

and the random variable X_1-X_2 is distributed as $(X_1+X_2)Z_{1/2}$, where $Z_{1/2}$ is independent of X_1+X_2 and has a symmetric beta distribution with density $f_{Z_{1/2}}(z) = 1/\pi\sqrt{1-z^2}$. In particular, the even moments are given by

$$\frac{1}{\pi}\int_{-1}^{1} \frac{z^{2n}\,dz}{\sqrt{1-z^2}} = \frac{1}{2^{2n}}\binom{2n}{n}. \tag{3.5}$$

Therefore,

$$\mathbb{E}(X_1-X_2)^{2n} = \mathbb{E}(X_1+X_2)^{2n}\,\mathbb{E}Z_{1/2}^{2n} = \frac{(2n)!}{2^{2n}}\binom{2n}{n}. \tag{3.6}$$

It follows that

$$\mathbb{E}(X_1+X_2)^{2n}+\mathbb{E}(X_1-X_2)^{2n} = (2n)! + \frac{(2n)!}{2^{2n}}\binom{2n}{n}. \tag{3.7}$$

The evaluations (3.3) and (3.7) imply (3.1).

4. An extension related to Legendre polynomials

A key point in the evaluation given in the previous section is the elementary identity

$$1+(-1)^k = \begin{cases} 2 & \text{if } k \text{ is even;} \\ 0 & \text{otherwise.} \end{cases} \tag{4.1}$$

This reduces the number of terms in the sum (3.3) from $2n$ to n. A similar cancellation occurs for any $p \in \mathbb{N}$. Indeed, the natural extension of (4.1) is given by

$$\sum_{j=0}^{p-1} \omega^{jr} = \begin{cases} p & \text{if } r \equiv 0 \pmod{p}; \\ 0 & \text{otherwise;} \end{cases} \tag{4.2}$$

Here $\omega = e^{2\pi i/p}$ is a complex p-th root of unity. Observe that (4.2) reduces to (4.1) when $p=2$.

The goal of this section is to discuss the extension of (3.1). The main result is given in the next theorem. The Legendre polynomials appearing in the next theorem are defined by

$$P_n(x) = \frac{1}{2^n\,n!}\left(\frac{d}{dx}\right)^n (x^2-1)^n. \tag{4.3}$$

Theorem 4.1. Let n, p be positive integers. Then

$$\sum_{k=0}^{n} \binom{2kp}{kp} \binom{2(n-k)p}{(n-k)p} = \frac{2^{2np}}{p} \sum_{\ell=0}^{p-1} e^{i\pi \ell n} P_{np}\left(\cos\left(\frac{\pi \ell}{p} \right) \right). \tag{4.4}$$

Proof. Replace the random variable $X_1 - X_2$ considered in the previous section, by $X_1 + WX_2$, where W is a complex random variable with uniform distribution among the p-th roots of unity. That is,

$$\Pr\left\{ W = \omega^{\ell} \right\} = \frac{1}{p}, \quad \text{for } 0 \leq \ell \leq p-1. \tag{4.5}$$

The identity (4.2) gives

$$\mathbb{E}W^r = \begin{cases} 1 & \text{if } r \equiv 0 \pmod{p}; \\ 0 & \text{otherwise.} \end{cases} \tag{4.6}$$

This is the cancellation alluded above.

Now proceed as in the previous section to obtain the moments

$$\begin{aligned} \mathbb{E}(X_1 + WX_2)^{np} &= \sum_{k=0}^{n} \binom{np}{kp} \mathbb{E}X_1^{(n-k)p}\, \mathbb{E}X_2^{kp} \\ &= \frac{(np)!}{2^{2np}} \sum_{k=0}^{n} \binom{2kp}{kp} \binom{2(n-k)p}{(n-k)p}. \end{aligned} \tag{4.7}$$

A second expression for $\mathbb{E}(X_1 + WX_2)^{np}$ employs an alternative form of the Legendre polynomial $P_n(x)$ defined in (4.3).

Proposition 4.2. The Legendre polynomial is given by

$$P_n(x) = \frac{1}{n!} \mathbb{E}\left[(x + \sqrt{x^2-1})X_1 + (x - \sqrt{x^2-1})X_2 \right]^n, \tag{4.8}$$

where X_1 and X_2 are independent $\Gamma\left(\frac{1}{2}, 1\right)$ random variables.

Proof. The proof is based on characteristic functions. Compute the sum

$$\mathbb{E}e^{t(x+\sqrt{x^2-1})\, X_1}\, \mathbb{E}e^{t(x-\sqrt{x^2-1})\, X_2} = \sum_{k=0}^{\infty} \frac{t^n}{n!} \mathbb{E}\left[(x + \sqrt{x^2-1})\, X_1 + (x - \sqrt{x^2-1})\, X_2 \right]. \tag{4.9}$$

The moment generating function for a $\Gamma\left(\frac{1}{2}, 1\right)$ random variable is

$$\mathbb{E}e^{tX} = (1-t)^{-1/2}. \tag{4.10}$$

This reduces (4.9) to

$$\left(1 - t(x + \sqrt{x^2-1})\right)^{-1/2} \left(1 - t(x - \sqrt{x^2-1})\right)^{-1/2} = (1 - 2tx + t^2)^{-1/2}$$

which is the generating function of the Legendre polynomials. □

This concludes the proof of Theorem 4.1. □

Corollary 4.3. Let x be a variable and Γ_1, Γ_2 as before. Then

$$\mathbb{E}(\Gamma_1 + x^2\Gamma_2)^n = n!x^n P_n\left(\tfrac{1}{2}(x + x^{-1})\right). \tag{4.11}$$

Proof. This result follows from Proposition 4.2 and the change of variables $x \mapsto \frac{1}{2}(x+x^{-1})$, known as the Joukowsky transform. □

Replacing x by $W^{1/2}$ in (4.11) and averaging over the values of W gives the second expression for $\mathbb{E}(X_1+WX_2)^{np}$. The proof of Theorem 4.1 is complete.

5. Chu-vandermonde

The arguments presented here to prove (2.5) can be generalized by replacing the random variables $\Gamma\left(\frac{1}{2},1\right)$ by two random variables $\Gamma(a_i,1)$ with shape parameters a_1 and a_2, respectively. The resulting identity is the Chu-Vandermonde theorem.

Theorem 5.1. Let a_1 and a_2 be positive real numbers. Then

$$\sum_{k=0}^{n} \frac{(a_1)_k}{k!}\frac{(a_2)_{n-k}}{(n-k)!} = \frac{(a_1+a_2)_n}{n!}. \tag{5.1}$$

The reader will find in [2] a more traditional proof. The paper [10] describes how to find and prove this identity in automatic form.

Exactly the same argument for (2.6) provides a multivariable generalization of the Chu-Vandermonde identity.

Theorem 5.2. Let $\{a_i\}_{1\le i\le m}$ be a collection of m positive real numbers. Then

$$\sum_{k_1+\cdots+k_m=n} \frac{(a_1)_{k_1}}{k_1!}\cdots\frac{(a_m)_{k_m}}{k_m!} = \frac{1}{n!}(a_1+\cdots+a_m)_n. \tag{5.2}$$

The final stated result presents a generalization of Theorem 4.1.

Theorem 5.3. Let $n,\, p\in\mathbb{N}, a\in\mathbb{R}^+$ and $\omega = e^{i\pi/p}$. Then

$$\sum_{k=0}^{n} \frac{(a)_{kp}}{(kp)!}\frac{(a)_{(n-k)p}}{((n-k)p)!}z^{2kp} = \frac{1}{p}\sum_{\ell=0}^{p-1} e^{i\pi\ell n} z^{np} C_{np}^{(a)}\left(\tfrac{1}{2}(z\omega^{\ell}+z^{-1}\omega^{-\ell})\right). \tag{5.3}$$

Here $C_n^{(a)}(x)$ is the Gegenbauer polynomial of degree n and parameter a.

Proof. Start with the moment representation for the Gegenbauer polynomials

$$C_n^{(a)}(x) = \frac{1}{n!}\mathbb{E}_{U,V}\left(U(x+\sqrt{x^2-1})+V(x-\sqrt{x^2-1})\right)^n \tag{5.4}$$

with U and V independent $\Gamma(a,1)$ random variables. This representation is proved in the same way as the proof for the Legendre polynomial, replacing the exponent $-1/2$ by and exponent $-a$. Note that the Legendre polynomials are Gegenbauer polynomials with parameter $a=\frac{1}{2}$. This result can also be found in Theorem 3 of [9]. □

Note 5.4. The value $z=1$ in (5.3) gives

$$\sum_{k=0}^{n} \frac{(a)_{kp}}{(kp)!}\frac{(a)_{(n-k)p}}{((n-k)p)!} = \frac{1}{p}\sum_{\ell=0}^{p-1} e^{i\pi\ell n} C_{np}^{(a)}\left(\cos\left(\frac{\pi\ell}{p}\right)\right). \tag{5.5}$$

This is a generalization of Chu-Vandermonde.

The techniques presented here may be extended to a variety of situations. Two examples illustrate the type of identities that may be proven. They involve the Hermite polynomials defined by

$$H_n(x) = (-1)^n e^{x^2} \left(\frac{d}{dx} \right)^n e^{-x^2}. \tag{5.6}$$

Theorem 5.5. Let $m \in \mathbb{N}$. The Hermite polynomials satisfy

$$\frac{1}{n!} H_n \left(\frac{x_1 + \cdots + x_m}{\sqrt{m}} \right) = m^{-n/2} \sum_{k_1 + \cdots + k_m} \frac{H_{k_1}(x_1)}{k_1!} \cdots \frac{H_{k_m}(x_m)}{k_m!}. \tag{5.7}$$

Proof. Start with the moment representation for the Hermite polynomials

$$H_n(x) = 2^n \mathbb{E}(x + iN)^n, \tag{5.8}$$

where N is normal with mean 0 and variance $\frac{1}{2}$. The details are left to the reader. □

The moment representation for the Gegenbauer polynomials (5.4) yields the final result presented here.

Theorem 5.6. Let $m \in \mathbb{N}$. The Gegenbauer polynomials $C_n^{(a)}(x)$ satisfy

$$C_n^{(a_1 + \cdots + a_m)}(x) = \sum_{k_1 + \cdots + k_m = n} C_{k_1}^{(a_1)}(x) \cdots C_{k_m}^{(a_m)}(x). \tag{5.9}$$

Remark 5.7. A relation between Gegenbauer and Hermite polynomials is given by

$$\lim_{a \to \infty} \frac{1}{a^{n/2}} C_n^{(a)} \left(\frac{x}{\sqrt{a}} \right) = \frac{1}{n!} H_n(x). \tag{5.10}$$

This relation allows to recover easily identity (5.7) from identity (5.9).

The examples presented here, show that many of the classical identities for special functions may be established by probabilistic methods. The reader is encouraged to try this method in his/her favorite identity.

Acknowledgements. The work of the second author was partially supported by nsf-dms 0070567.

References

[1] T. Amdeberhan, V. De Angelis, M. Lin, V. Moll, and B. Sury. A pretty binomial identity. *Elem. Math.*, to appear, 2012.

[2] G. Andrews, R. Askey, and R. Roy. *Special Functions*, volume 71 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, New York, 1999.

[3] V. De Angelis. Pairings and signed permutations. *Amer. Math. Monthly*, 113:642–644, 2006.

[4] Y. A. Brychkov. *Handbook of Special Functions. Derivatives, Integrals, Series and Other Formulas.* Taylor and Francis, Boca Raton, Florida, 2008.

[5] Guisong Chang and Chen Xu. Generalization and Probabilistic Proof of a Combinatorial Identity. *Amer. Math. Monthly*, 118:175–177, 2011.

[6] F. W. J. Olver, D. W. Lozier, R. F. Boisvert, and C. W. Clark, editors. *NIST Handbook of Mathematical Functions*. Cambridge University Press, 2010.

[7] M. Petkovsek, H. Wilf, and D. Zeilberger. *A=B*. A. K. Peters, Ltd., 1st edition, 1996.

[8] T. Simpson. The invention of a general method for determining the sum of every second, third, fourth, or fifth, etc term of a series, taken in order; the sum of the whole series being known. *Phil. Trans. Royal Soc. London*, 50:757–769, 1759.

[9] P. Sun. Moment representation of Bernoulli polynomial, Euler polynomial and Gegenbauer polynomials. *Stat. and Prob. Letters*, 77:748, 2007.

[10] D. Zeilberger. Three recitations on holonomic functions and hypergeometric series. *Jour. Symb. Comp.*, 20:699–724, 1995.

Information Theory Laboratory, E.P.F.L., 1015 Lausanne, Switzerland
E-mail address: `christophe.vignat@epfl.ch`

Department of Mathematics, Tulane University, New Orleans, LA 70118
E-mail address: `vhm@math.tulane.edu`