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 such as

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

appears throughout the undergraduate curriculum. At the end of the previous century, the evaluation of these sums was trivialized by the work of H. Wilf, D. Zeilberger and M. Petkovšek [9]. The method of *creative telescoping*, described in the charming book [9], provides an automatic tool for the verification of this type of identities.

On the other hand, 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 function $f_X(x)$, such that

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

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 measurable function g of the random variable X is computed as

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

Date: February 24, 2014.

1991 *Mathematics Subject Classification.* Primary 05A10, Secondary 33B15, 60C99.

Key words and phrases. binomial sums, gamma distributed random variables, Vandermonde identity, orthogonal polynomials.

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

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}, \quad \text{for } x \in \mathbb{R}, \tag{1.5}$$

and to exponential random variables, with probability density function

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

with $\lambda > 0$.

The examples employed in the arguments presented here include random variables with a gamma distribution of shape parameter $a > 0$ and scale parameter $\theta > 0$, written as $X \sim \Gamma(a,\theta)$. These are defined by the density function

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

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 $\operatorname{Re} s > 0$. The reader will find in [2] extensive information about this special function. The exponential distribution is the special case of the gamma distribution with shape parameter $a = 1$. Recall that for a random variable X, the n-th moment is defined by $\mathbb{E}(X^n)$. Observe that if $X \sim \Gamma(a,\theta)$, then $X = \theta Y$ where $Y \sim \Gamma(a,1)$. Moreover

$$\mathbb{E}X^n = \theta^n (a)_n, \tag{1.9}$$

where

$$(a)_n = \frac{\Gamma(a+n)}{\Gamma(a)} = a(a+1)\cdots(a+n-1) \tag{1.10}$$

is the Pochhammer symbol. The main property of this family of Gamma random variables is that it is closed under addition: assume $X_i \sim \Gamma(a_i,\theta)$ are independent, then

$$X_1 + \cdots + X_m \sim \Gamma(a_1 + \cdots + a_m, \theta). \tag{1.11}$$

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

Another distribution will be useful in the following, namely the beta distribution denoted as $Be(a,b)$ with density

$$f_X(x;a,b) = \begin{cases} x^{a-1}(1-x)^{b-1}/B(a,b) & \text{for } 0 \leq x \leq 1; \\ 0 & \text{otherwise.} \end{cases} \tag{1.12}$$

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

$$B(a,b) = \int_0^1 x^{a-1}(1-x)^{b-1}\, dx \tag{1.13}$$

and also the Pearson type II distribution [16] denoted as $Pe(c)$ with density

$$\begin{cases} \frac{1}{B(1/2,c)}(1-x^2)^{c-1} & \text{for } -1 \leq x \leq 1; \\ 0 & \text{otherwise.} \end{cases} \tag{1.14}$$

The uniform distribution on $[0,1]$ appears as the special case $a=b=1$ of the beta distribution. A random variable $Z_{a,b}$ with distribution $Be(a,b)$ can be generated as

$$Z_{a,b} := \frac{X_a}{X_a + X_b}, \tag{1.15}$$

where X_a and X_b are independent gamma distributed with shape parameters a and b, respectively; and a random variable Z_c with $Pe(c)$ distribution can be generated as $1-2Z_{c,c}$, that is,

$$Z_c := 1 - \frac{2X_c}{X_c + X_c'} = \frac{X_c - X_c'}{X_c + X_c'}, \tag{1.16}$$

where X_c and X_c' are independent gamma distributed with shape parameter c. A well-known result is that $Z_{a,b}$ and $X_a + X_b$ are independent in (1.15); similarly, $X_c + X_c'$ and Z_c are independent in (1.16). The reader will find information about these random variables and detailed proofs of the statements employed here in Chapter 2 of [7].

The central idea of the paper is simple. Suppose a sequence of interest $\{a_k\}$ is identified as the moments of a random variable X, so that $\mathbb{E}(X^k) = a_k$. Suppose also that if $X_1, X_2, \cdots, X_m$ are independent random variables, identically distributed like X, then the moments of the sum $Y = X_1 + X_2 + \cdots + X_m$ can also be computed, say $\mathbb{E}(Y^k) = b_k$. Then the multinomial theorem and the linearity of the expected value operator give

$$\begin{aligned} b_n &= \mathbb{E}(X_1 + \cdots + X_m)^n \\ &= \sum_{k_1 + \cdots k_m = n} \binom{n}{k_1,\, k_2, \cdots, k_m} \mathbb{E}(X_1^{k_1})\, \mathbb{E}(X_2^{k_2}) \cdots \mathbb{E}(X_m^{k_m}) \\ &= \sum_{k_1 + \cdots k_m = n} \binom{n}{k_1,\, k_2, \cdots, k_m} a_{k_1} a_{k_2} \cdots a_{k_m}. \end{aligned}$$

In terms of probability density functions, this could be rephrased as saying that if one can compute the integral $\int_{\mathbb{R}} x^k f(x)\, dx$ as well as the integral $\int_{\mathbb{R}} x^k f^{*m}(x)\, dx$ of the m-th convolution f^{*m} of f with itself, then the multinomial theorem gives interesting identities. This formulation hides the probability setting of the method.

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 $h(x)^2$, where

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

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 $h(x)^m$. 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 $1-4x$ 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 (2.5) appeared in 2006 in [6]. 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

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

If $X \sim \Gamma(a,\theta)$, then the moments are given by (1.9). Thus, for each random variable X_i, the moments are given by

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

iterating the functional equation $\Gamma(a+1) = a\Gamma(a)$ to obtain the second form. The expression

$$\binom{n}{k_1, \cdots, k_m} = \frac{n!}{k_1!\, k_2! \,\cdots\, k_m!} \tag{2.10}$$

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

$$\frac{n!}{2^{2n}} \sum_{k_1+\cdots+k_m=n} \binom{2k_1}{k_1} \cdots \binom{2k_m}{k_m}. \tag{2.11}$$

The evaluation of the left-hand side of (2.8) employs basic probabilistic results about the pdf of the sum of independent, gamma distributed random variables.

From (1.11), the sum of m independent random variables $\Gamma\left(\frac{1}{2},1\right)$ has a distribution $\Gamma(\frac{m}{2},1)$. Therefore, the left-hand side of (2.8) is

$$\frac{\Gamma(\frac{m}{2}+n)}{\Gamma(\frac{m}{2})}. \tag{2.12}$$

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 volume 1 of [10]. The proof presented here employs the famous multisection technique, first introduced by Simpson [11] 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 Pearson type II distribution $Pe(\frac{1}{2})$ [1] with density $f_{Z_{1/2}}(z) = 1/(\pi\sqrt{1-z^2})$. In particular, the even moments of $Z_{1/2}$ are proportional to the central binomial coefficients:

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

Therefore,

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

[1] the Pearson type II distribution with parameter $c=\frac{1}{2}$ is also called the arcsine distribution

It follows that [2]

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

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, let $\omega = e^{2\pi i/p}$ be a complex p-th root of unity. Then a natural extension of (4.1) is given by

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

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 the Rodrigues formula

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

The Legendre polynomials are examples of orthogonal polynomials and their properties may be found in a variety of texts. The authors's favorite ones include [2], [8], [13] and [14] as well as Chapter 4 in the recent book [3].

In a classical subject, like the one treated in this paper, it is hard to state that a result is new. The authors have not been able to find the next theorem in the literature.

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

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

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,

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

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 to above.

[2]we note that this moment could be equally easily computed using the generating function (2.7)

Now proceed as in the previous section to obtain the moments

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

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). The next result appears in [12].

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 moment generating functions. Compute the sum

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

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. See page 146 of [3]. □

Corollary 4.3. Let x be a variable and X_1, X_2 as before. Then

$$\mathbb{E}(X_1+x^2X_2)^n = n!x^nP_n\left(\tfrac{1}{2}(x+x^{-1})\right). \tag{4.11}$$

Proof. This result follows from Proposition 4.2 by the change of variables $u = \frac{1}{2}(x+x^{-1})$ and the identity $u^2-1 = (\frac{1}{2}(x-x^{-1}))^2$. □

Equation (4.11) is a polynomial identity in the variable x. Hence we can replace x by the random variable $W^{1/2}$ and average over the values of W. This yields

$$\mathbb{E}(X_1+x^2X_2)^{np} = (np)!\frac{1}{p}\sum_{l=0}^{p-1}e^{\imath\frac{2\pi}{p}\frac{npl}{2}}P_{np}\left(\frac{1}{2}\left(e^{\imath\frac{\pi}{p}}+e^{-\imath\frac{\pi}{p}}\right)\right) = \frac{(np)!}{p}\sum_{l=0}^{p-1}(-1)^{nl}P_{np}\left(\cos(\frac{\pi l}{p})\right). \tag{4.12}$$

The proof of Theorem 4.1 is complete. □

5. Chu-Vandermonde and other classical identities

This section contains a selection of identities from the area of Special Functions that can be derived by the method described in this paper. For example, 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

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

This is a well-known result and the reader will find in [2] a more traditional proof. The paper [15] describes how to find and prove this identity in automatic form.

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

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

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

Proof. Consider m independent Gamma random variables $X_i\sim\Gamma\left(a_i,1\right)$. Then (1.9) gives

$$\mathbb{E}X_i^k=\frac{\Gamma\left(a_i+k_i\right)}{\Gamma\left(a_i\right)}=\left(a_i\right)_{k_i}, \tag{5.3}$$

and with $X=X_1+\cdots+X_n$,

$$\mathbb{E}\left[X^n\right]=\sum_{k_1+\cdots+k_m=n}\frac{n!}{k_1!\ldots k_m!}\left(a_1\right)_{k_1}\ldots\left(a_m\right)_{k_m}. \tag{5.4}$$

To obtain the result recall (1.11): the sum $X=X_1+\cdots+X_n$ is a Gamma random variable with $X\sim\Gamma\left(a_1+\cdots+a_n,1\right)$. Therefore

$$\mathbb{E}\left[X^n\right]=\left(a_1+\cdots+a_m\right)_n. \tag{5.5}$$

□

The final stated result presents a generalization of Theorem 4.1. This statement involves the Gegenbauer polynomial $C_n^{(a)}(x)$ of degree n and parameter $a>0$, defined by the Rodrigues' formula [2, 6.4.14]

$$C_n^{(a)}(x)=\frac{(2a)_n}{2^n n!\left(a+\frac{1}{2}\right)_n}(-1)^n(1-x^2)^{\frac{1}{2}-a}\frac{d^n}{dx^n}(1-x^2)^{n+a-\frac{1}{2}}. \tag{5.6}$$

The reader will find in [2] and [3] information about these polynomials and in Section 18.3 of [8] a collection of formulas for them.

The authors have been unable to find the next result and the note following it in the literature.

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

$$\sum_{k=0}^{n} \frac{(a)_{kp}}{(kp)!} \frac{(a)_{(n-k)p}}{((n-k)p)!} z^{2kp} = \frac{1}{p} \sum_{\ell=0}^{p-1} (-1)^{\ell n} z^{np} C_{np}^{(a)} \left(\tfrac{1}{2} (z\omega^{\ell} + z^{-1}\omega^{-\ell}) \right). \tag{5.7}$$

Proof. Start with the moment representation for the Gegenbauer polynomials

$$C_n^{(a)}(x) = \frac{1}{n!} \mathbb{E} \left(U(x + \sqrt{x^2-1}) + V(x - \sqrt{x^2-1}) \right)^n \tag{5.8}$$

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 $-\frac{1}{2}$ by the 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 [12]. □

Note 5.4. The value $z = 1$ in (5.7) gives

$$\sum_{k=0}^{n} \frac{(a)_{kp}}{(kp)!} \frac{(a)_{(n-k)p}}{((n-k)p)!} = \frac{1}{p} \sum_{\ell=0}^{p-1} (-1)^{\ell n} C_{np}^{(a)} \left(\cos\left(\frac{\pi \ell}{p} \right) \right). \tag{5.9}$$

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

$$H_n(x) = (-1)^n e^{x^2} \left(\frac{d}{dx} \right)^n e^{-x^2}. \tag{5.10}$$

The textbook [3] provides extensive information about this classical family of orthogonal polynomials.

The next theorem appears as entry 4.5.2.9 in volume 2 of [10].

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

$$\frac{1}{n!} H_n \left(\frac{x_1 + \cdots + x_m}{\sqrt{m}} \right) = m^{-n/2} \sum_{k_1 + \cdots + k_m = n} \frac{H_{k_1}(x_1)}{k_1!} \cdots \frac{H_{k_m}(x_m)}{k_m!}. \tag{5.11}$$

Proof. Let N be a normal random variable with mean 0 and variance $\frac{1}{2}$. The proof starts with the moment representation for the Hermite polynomials

$$H_n(x) = 2^n \mathbb{E}(x + iN)^n, \tag{5.12}$$

that appears as Exercise 6.8 on page 167 of [13]. The details are left to the reader. □

The moment representation for the Gegenbauer polynomials (5.8) and the same probabilistic technique as before yield the final result presented here. The reader will find the following statement as entry 5.18.2.7 in [4].

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

$$C_n^{(a_1 + \cdots + a_m)}(x) = \sum_{k_1 + \cdots + k_m = n} C_{k_1}^{(a_1)}(x) \cdots C_{k_m}^{(a_m)}(x). \tag{5.13}$$

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

$$\lim_{a\to\infty} \frac{1}{a^{n/2}} C_n^{(a)}\left(\frac{x}{\sqrt{a}}\right) = \frac{1}{n!} H_n(x). \tag{5.14}$$

This relation allows us to recover easily identity (5.11) from identity (5.13).

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 these methods in his/her favorite identity. For example, he/she may want to prove the Pfaff-Kummer transformation formula

$$ {}_2F_1\left(a,b;c;z\right) = \left(1-z\right)^{-a} {}_2F_1\left(a, c-b; c; \frac{z}{z-1}\right) \tag{5.15}$$

where

$$ {}_2F_1\left(a,b;c;z\right) = \sum_{k=0}^{+\infty} \frac{\left(a\right)_k \left(b\right)_k}{\left(c\right)_k} z^k \tag{5.16}$$

is the hypergeometric function, by remarking that

$$ {}_2F_1\left(a,b;c;z\right) := \mathbb{E}\left(\exp\left(zX_a Z_{b,c-b}\right)\right) = \mathbb{E}\left(1 - zZ_{b,c-b}\right)^{-a} \tag{5.17}$$

where $X_a \sim \Gamma\left(a,1\right)$ and $Z_{b,c-b} \sim B\left(b, c-b\right)$ and by using the symmetry $1 - Z_{b,c-b} \sim Z_{c-b,b}$. The reader will find in [2] a proof of (5.15).

Acknowledgements. The work of the second author was partially supported by NSF-DMS 0070567. The authors wish to thank the referee for his detailed review of the manuscript.

References

[1] T. Amdeberhan, V. De Angelis, M. Lin, V. Moll, and B. Sury. A pretty binomial identity. *Elem. Math.*, 67:1–8, 2012.
[2] G. E. Andrews, R. Askey, and R. Roy. *Special Functions*, volume 71 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, New York, 1999.
[3] R. Beals and R. Wong. *Special Functions. A Graduate Text*, volume 126 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, New York, 2010.
[4] Y. A. Brychkov. *Handbook of Special Functions. Derivatives, Integrals, Series and Other Formulas*. Taylor and Francis, Boca Raton, Florida, 2008.
[5] G. Chang and C. Xu. Generalization and probabilistic proof of a combinatorial identity. *Amer. Math. Monthly*, 118:175–177, 2011.
[6] V. De Angelis. Pairings and signed permutations. *Amer. Math. Monthly*, 113:642–644, 2006.
[7] R. Durrett. *Probability: Theory and Examples*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 2010.
[8] F. W. J. Olver, D. W. Lozier, R. F. Boisvert, and C. W. Clark, editors. *NIST Handbook of Mathematical Functions*. Cambridge University Press, 2010.
[9] M. Petkovšek, H. Wilf, and D. Zeilberger. *A=B*. A. K. Peters, Ltd., 1st edition, 1996.
[10] A. P. Prudnikov, Yu. A. Brychkov, and O. I. Marichev. *Integrals and Series*. Gordon and Breach Science Publishers, 1992.
[11] 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.
[12] P. Sun. Moment representation of Bernoulli polynomial, Euler polynomial and Gegenbauer polynomials. *Stat. and Prob. Letters*, 77:748, 2007.
[13] N. M. Temme. *Special Functions. An introduction to the Classical Functions of Mathematical Physics*. John Wiley and sons, New York, 1996.
[14] E. T. Whittaker and G. N. Watson. *Modern Analysis*. Cambridge University Press, 1962.

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

[16] L. Devroye. *Non-Uniform Random Variate Generation*. Springer-Verlag, New York, 1986.

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`