Turkish Journal of Mathematics

http://journals.tubitak.gov.tr/math/

Research Article

Turk J Math
(2014) 38: 558 – 575

doi:10.3906/mat-1301-6

Moment equalities for sums of random variables via integer partitions and Faà di Bruno's formula

Dietmar FERGER[*]
Department of Mathematics, Dresden University of Technology, Dresden, Germany

Received: 03.01.2013 • **Accepted:** 01.10.2013 • **Published Online:** 14.03.2014 • **Printed:** 11.04.2014

Abstract: We give moment equalities for sums of independent and identically distributed random variables including, in particular, centered and specifically symmetric summands. Two different types of proofs, combinatorial and analytical, lead to 2 different types of formulas. Furthermore, the combinatorial method allows us to find the optimal lower and upper constants in the Marcinkiewicz–Zygmund inequalities in the case of even moment-orders. Our results are applied to give elementary proofs of the classical central limit theorem (CLT) and of the CLT for the empirical bootstrap. Moreover, we derive moment and exponential inequalities for self-normalized sums.

Key words: Moments, integer partitions, Faà di Bruno's chain rule, Marcinkiewicz–Zygmund inequalities, bootstrap, self-normalized sums

1. Introduction and main results

Let $X_1, \ldots, X_n$ be $n \in \mathbb{N}$ independent and identically distributed (i.i.d.) copies of a real random variable X with $S_n := \sum_{1 \leq i \leq n} X_i$ the pertaining sum. If for some $m \in \mathbb{N}$ the mth moment $\mu_m := \mathbb{E}[X^m]$ is a real number, then we wish to compute the mth moment

$$M_n(m) := \mathbb{E}[(\sum_{i=1}^{n} X_i)^m]$$

of S_n explicitly in terms of $\mu_1, \ldots, \mu_m$. This paper focuses on centered ($\mu_1 = 0$) and also more specifically on symmetric summands. In these cases there exists a host of publications dealing with inequalities for $M_n(m)$, but to the best of our knowledge, we are not aware of exact formulas for $M_n(m)$.

If $\mu_1 = 0$, first simple calculations show that $M_n(1) = 0$, $M_n(2) = n\mu_2$, $M_n(3) = n\mu_3$, and $M_n(4) = n\mu_4 + 3n(n-1)\mu_2^2$. Indeed, we will formalize these calculations of $M_n(m)$ for general $m \in \mathbb{N}$ and give an explicit formula involving integer partitions as specified below. It allows for a simple implementation with formal computer algebra programs. Another method is based on characteristic functions and Faà di Bruno's extension of the chain rule for differentiation. It results in a second expression that mathematically has a more appealing form. The idea for this approach can be traced back to Lukacs [11], who used Faà di Bruno's formula for deriving the well-known relations between moments and cumulants of a random variable; refer to Theorem 3

[*]Correspondence: dietmar.ferger@tu-dresden.de
Dietmar Ferger, Department of Mathematics, Dresden University of Technology, Dresden, Germany
2010 *AMS Mathematics Subject Classification:* Primary 60E10, 60E15; Secondary 60-08, 60F05, 62G09.

there. For a historical survey on Faà di Bruno's formula we recommend the work of Johnson [10]. Constantine and Savits [3] generalized Faà di Bruno's formula to functions with several arguments and with it they extended Lukacs' [11] relations to multivariate moments and cumulants; refer to Theorem 4.5 there.

The representation of our formulas requires some minimal notation. For $r \in \mathbb{N}$ and $\mathbf{k} = (k_1, \ldots, k_r) \in \mathbb{N}^r$, let

$$J_{m,r} := \{\mathbf{k} \in \mathbb{N}^r : k_1 + \ldots + k_r = m\}$$

be the set of all ordered integer partitions of m into exactly r summands (parts) and let

$$I_{m,r} := \{\mathbf{k} \in \mathbb{N}^r : k_1 + \ldots + k_r = m, k_1 \geq \cdots \geq k_r\}$$

denote the corresponding subset of all unordered partitions, where the parts are decreasingly arranged and thus the order of the components is not taken into account. For example: $J_{5,3} = \{$(3,1,1), (1,3,1), (1,1,3), (2,2,1), (2,1,2), (1,2,2)$\}$ with pertaining set of unordered partitions $I_{5,3} = \{(3,1,1),(2,2,1)\}$. To each integer partition $\mathbf{k} \in J_{m,r}$ there belongs the unique tuple $(\kappa_1, \ldots, \kappa_a)$ of a distinct parts $\kappa_1 > \cdots > \kappa_a$ and their respective (positive) multiplicities $(n_1, \ldots, n_a)$. Thus, for every $\mathbf{k} = (k_1, \ldots, k_r) \in J_{m,r}$, one has $k_1 + \ldots + k_r = n_1\kappa_1 + \ldots + n_a\kappa_a = m$ with $1 \leq a \leq r$ and $n_1 + \ldots + n_a = r$. For example, $\mathbf{k} = (7,7,5,5,5,2,1,1,1) \in I_{34,9} \subseteq J_{34,9}$ has $a = 4$ distinct parts $(7,5,2,1)$ with pertaining multiplicities $(2,3,1,3)$. Finally, for any set $K \subseteq \mathbb{N}^r$ we use the traditional convention

$$K + \mathbf{1} := \{(k_1 + 1, \ldots, k_r + 1) : \mathbf{k} \in K\}.$$

Observe that there is a one-to-one correspondence between the set $I_{m,r}$ and the set

$$L_{m,r} := \{\mathbf{l} = (l_1, \ldots, l_m) \in \mathbb{N}_0^m : \sum_{i=1}^m i l_i = m, \ \sum_{i=1}^m l_i = r\}.$$

Indeed, given a $\mathbf{k} \in I_{m,r}$ with distinct parts $\kappa_1 < \ldots < \kappa_a$ (increasingly ordered) and pertaining multiplicities $(n_1, \ldots, n_a)$, put $l_{\kappa_i} := n_i, 1 \leq i \leq a$ and $l_j := 0$ for $j \in \{1, \ldots, m\} \setminus \{\kappa_1, \ldots, \kappa_a\}$. Then $\sum_{i=1}^m i l_i = \sum_{i=1}^a \kappa_i l_{\kappa_i} = \sum_{i=1}^a \kappa_i n_i = m$ and $\sum_{i=1}^m l_i = \sum_{i=1}^a l_{\kappa_i} = \sum_{i=1}^a n_i = r$, whence $\mathbf{l} = (l_1, \ldots, l_m) \in L_{m,r}$. The inverse map is as follows: given an $\mathbf{l} = (l_1, \ldots, l_m) \in L_{m,r}$, consider the indices i with $l_i \neq 0$, say a many, and let them be decreasingly ordered, which gives the distinct parts $\kappa_1 > \ldots > \kappa_a$. Put $n_i := l_{\kappa_i}$ for $i = 1, \ldots, a$. Then $\mathbf{k} := (\kappa_1, \ldots, \kappa_1, \kappa_2, \ldots, \kappa_2, \ldots, \kappa_a, \ldots, \kappa_a)$ with exactly n_i entries of κ_i an element of $I_{m,r}$, because $\sum_{i=1}^a n_i = \sum_{i=1}^a l_{\kappa_i} = \sum_{j=1}^m l_j = r$ and $\sum_{i=1}^r k_i = \sum_{i=1}^a n_i\kappa_i = \sum_{i=1}^a l_{\kappa_i}\kappa_i = \sum_{j=1}^m j l_j = m$.

The starting point for the derivation of our main results on centered or symmetric summands is the following Lemma 1.1. In the first part thereof, we consider more generally summands that are not necessarily centered and possibly are not identically distributed.

Denote by $N_n := \{1, \ldots, n\}$ the set of the first n integers and let us agree for a nonempty subset $\{a_1, \ldots, a_r\} \subseteq N_n$, $1 \leq r \leq n$, that $a_i \neq a_j \ \forall\, i \neq j$. Moreover, recall the usual definition $\binom{n}{r} := 0$, if $r > n$.

Lemma 1.1 *Let m be any positive integer and assume that $X_1, \ldots, X_n$ are random variables with $\mathbb{E}[|X_i|^m] < \infty$ for each $1 \leq i \leq n$.*

(1) If $X_1, \ldots, X_n$ are independent, then

$$\mathbb{E}[(\sum_{i=1}^{n} X_i)^m] = \sum_{r=1}^{m} \sum_{\{a_1,\ldots,a_r\}\subseteq N_n} \sum_{\boldsymbol{k}\in J_{m,r}} \binom{m}{k_1 \ldots k_r} \prod_{s=1}^{r} \mathbb{E}[X_{a_s}^{k_s}] \,. \tag{1.1}$$

(2) If $X_1, \ldots, X_n$ are i.i.d., then

$$\mathbb{E}[(\sum_{i=1}^{n} X_i)^m] = \sum_{1\leq r\leq m} A_{r,m} \binom{n}{r} \tag{1.2}$$

with constants

$$A_{r,m} = A_{r,m}(\mu_1, \ldots, \mu_m) = \sum_{\boldsymbol{k}\in I_{m,r}} \binom{m}{k_1 \ldots k_r} \binom{r}{n_1 \ldots n_a} \mu_{\kappa_1}^{n_1} \cdots \mu_{\kappa_a}^{n_a} \tag{1.3}$$

$$= r!m! \sum_{\boldsymbol{l}\in L_{m,r}} \prod_{j=1}^{m-r+1} \frac{\{\frac{1}{j!}\mu_j\}^{l_j}}{l_j!} \,. \tag{1.4}$$

Equation (1.1) is a rather simple consequence of the multinomial theorem in combination with linearity of the expectation and the multiplication rule for products with independent factors. As for formulas (1.2) and (1.3), Packwood [12] used a different combinatorial method, coming up with another formula. However, using the concept of integer partitions, his expression can be broken down such that it coincides with (1.2) and (1.3).

In the case of centered random variables X_i, the key problem lies in figuring out exactly those summands in (1.3) or (1.4), respectively, that vanish by virtue of $\mu_1 = 0$. Indeed, Packwood [12] pointed out that this is still an open problem ("While there appears to be no general formula ...", see p.8). The solution of this main issue is stated in the following:

Theorem 1.2 *Let m be any positive integer and assume that $X_1, \ldots, X_n$ are i.i.d. copies of some mean zero random variable X with $\mathbb{E}[|X|^m] < \infty$. Then for every $n \in \mathbb{N}$ the following equality holds:*

$$\mathbb{E}[(\sum_{i=1}^{n} X_i)^m] = \sum_{1\leq r\leq m/2} B_{r,m} \binom{n}{r} \tag{1.5}$$

with constants

$$B_{r,m} = B_{r,m}(\mu_2, \mu_3, \ldots, \mu_m) = \sum_{\boldsymbol{k}\in I_{m-r,r}+\boldsymbol{1}} \binom{m}{k_1 \ldots k_r} \binom{r}{n_1 \ldots n_a} \mu_{\kappa_1}^{n_1} \cdots \mu_{\kappa_a}^{n_a} \tag{1.6}$$

$$= r!m! \sum_{\boldsymbol{l}\in L_{m-r,r}} \prod_{j=1}^{m-r} \frac{\{\frac{1}{(j+1)!}\mu_{j+1}\}^{l_j}}{l_j!}. \tag{1.7}$$

The implementation of (1.3) and (1.6) in MATHEMATICA is very simple, especially by using the built-in function IntegerPartitions$[m, \{r\}]$, which generates the sets $I_{m,r}$. Moreover, a representation of $M_n(m)$ as polynomial in the variable n is easily possible; see the example in the Appendix. Indeed, by (1.5) the mth moment $M_n(m)$ as a function in n is a polynomial of degree $[m/2]$ with $[\cdot]$ denoting the floor-function. Since

$$(2p)!/(2^p p!) = (2p-1)!! := \prod_{i=1}^{p}(2i-1) \quad \text{for every } p \in \mathbb{N} \tag{1.8}$$

as follows easily by induction, we obtain from (1.6) or (1.7) that the leading coefficient for $m \geq 2$ is given by

$$\frac{1}{p!} B_{p,m} = (m-1)!!\, \mu_2^p, \text{ if } m = 2p \text{ is even}$$

and

$$\frac{1}{(p-1)!} B_{p-1,m} = \frac{1}{6}\, m!!\, (m-1)\mu_2^{p-2}\mu_3, \text{ if } m = 2p-1 \text{ is odd.}$$

In Section 3.1 below we will apply Theorem 1.2 to give elementary proofs of the classical central limit theorem (CLT) and of the CLT for the empirical bootstrap, which plays an important role in probability theory as well as in statistics.

If the summands are actually symmetric, then not only the first moment but also all odd moments of X are equal to zero. This leads to a further considerable reduction of the nonvanishing summands. In fact, $\mathbb{E}[S_n^m] = 0$ for every odd m, because symmetry of the summands entails symmetry of the whole sum S_n. In contrast to this extreme case in which actually all summands are equal to zero, we obtain for m even:

Theorem 1.3 *Let p be any positive integer and assume that $X_1, \ldots, X_n$ are i.i.d. copies of some symmetric random variable X with $\mathbb{E}[|X|^{2p}] < \infty$. Then for every $n \in \mathbb{N}$ the following equality holds:*

$$\mathbb{E}[(\sum_{i=1}^{n} X_i)^{2p}] = \sum_{1 \leq r \leq p} C_{r,p} \binom{n}{r} \tag{1.9}$$

with constants

$$C_{r,p} = C_{r,p}(\mu_2, \mu_4, \ldots, \mu_{2p}) = \sum_{\boldsymbol{k} \in I_{p,r}} \binom{2p}{2k_1 \ldots 2k_r} \binom{r}{n_1 \ldots n_a} \mu_{2\kappa_1}^{n_1} \cdots \mu_{2\kappa_a}^{n_a} \tag{1.10}$$

$$= r!(2p)! \sum_{\boldsymbol{l} \in L_{p,r}} \prod_{j=1}^{p-r+1} \frac{\{\frac{\mu_{2j}}{(2j)!}\}^{l_j}}{l_j!}\,. \tag{1.11}$$

For symmetric variables we can also easily derive the (lower and upper) Marcinkiewicz–Zygmund inequalities restricted to even moment-orders but with best constants. In Section 3.2 we will use these to derive moment and exponential inequalities for self-normalized sums.

Theorem 1.4 *Let p be any positive integer. If*

$$X_1, \ldots, X_n \text{ are independent and symmetric with } \mathbb{E}|X_i|^{2p} < \infty,\ 1 \leq i \leq n, \tag{1.12}$$

then

$$\mathbb{E}\big[\{\sum_{i=1}^{n} X_i^2\}^p\big] \leq \mathbb{E}[(\sum_{i=1}^{n} X_i)^{2p}] \leq (2p-1)!!\, \mathbb{E}\big[\{\sum_{i=1}^{n} X_i^2\}^p\big]. \tag{1.13}$$

Here, the constant 1 *in the lower inequality and the constant* $(2p-1)!!$ *in the upper inequality are optimal in the following sense: if there are constants* D_p *and* C_p *such that for each* $n \in \mathbb{N}$ *and every sequence* $(X_1, \ldots, X_n)$ *with (1.12) the inequalities*

$$D_p\, \mathbb{E}\big[\{\sum_{i=1}^{n} X_i^2\}^p\big] \leq \mathbb{E}[(\sum_{i=1}^{n} X_i)^{2p}] \leq C_p\, \mathbb{E}\big[\{\sum_{i=1}^{n} X_i^2\}^p\big] \tag{1.14}$$

hold, then $D_p \leq 1$ *and* $C_p \geq (2p-1)!!$.

Remark 1.5 *(1) Under the assumption of the above theorem, Egorov [7,8] derived the upper inequality in (1.14) with* $C_p = Ce^{-p}(2p)^p$, *where* C *is some absolute constant with* $C \geq \sqrt{2}$. *Note that* $C = \sqrt{2}$ *cannot be taken for granted so far. However, with (1.8) and with the sharp inequalities of Batir [1] for the factorial, one verifies easily that* $\sqrt{2}e^{-p}(2p)^p > (2p-1)!!$ *for all* $p \in \mathbb{N}$, *whence Theorem 1.4 confirms that Egorov's bound* C_p *is valid for* $C = \sqrt{2}$.

Egorov [7,8] actually states that

$$\mathbb{E}[|\sum_{i=1}^{n} X_i|^s] \leq C_s\, \mathbb{E}\big[\{\sum_{i=1}^{n} X_i^2\}^{s/2}\big] \quad \textit{for every real } s \geq 2 \tag{1.15}$$

and justifies this simply by the monotonicity of the L_s*-norm. Even though we expect (1.15) to be true, his argument requires some more clarification.*

(2) Let $f_i : \mathbb{R}^n \to \mathbb{R}$ *be the* i*th projection,* $\nu := \nu_1 \otimes \ldots \otimes \nu_n$ *with* ν_i *denoting the distribution of* X_i, $s := 2p$, *and* $c := (s-1)!!^{1/s}$. *The upper inequality in (1.13) can then be restated as follows:*

$$(\int (\sum_{i=1}^{n} f_i)^s d\nu)^{1/s} \leq c\, (\int (\sum_{i=1}^{n} f_i^2)^{s/2} d\nu)^{1/s}. \tag{1.16}$$

Defant and Junge [4] considered linear and continuous operators $T : L_q(\mu) \to L_s(\nu)$ *(with* μ *and* ν *arbitrary measures) and* $f_1, \ldots, f_n \in L_q(\mu)$. *Given a triple* $(s, q, r) \in [1, \infty]^3$, *the problem is to find the minimal constant* $c = c(s, q, r) \geq 0$ *such that for each* T *the following inequality holds:*

$$(\int (\sum_{i=1}^{n} |Tf_i|^r)^{s/r} d\nu)^{1/s} \leq c||T||(\int (\sum_{i=1}^{n} |f_i|^r)^{q/r} d\mu)^{1/q}. \tag{1.17}$$

For $r = 2$ *and* $q = s$ *we see that the integrals on the right-hand side of (1.16) and (1.17), respectively, coincide. However, then the corresponding integrals on the left-hand sides do not fit together even if we choose* T *to be the identity map. Indeed, these integrals in each case are incompatible because the summands* f_i *in (1.16) have the counterparts* $|Tf_i|$ *in (1.17) and it is the absolute value that makes the main difference here. Unfortunately, Defant and Junge [4] also called (1.17) the Marcinkiewicz–Zygmund inequality, and this might cause some confusion. However, as we just have pointed out, there is no intrinsic link between these 2 results up to the same denomination.*

2. The proofs

We start with the following version of the multinomial theorem, which is better suited for our intentions.

Theorem 2.1 (Modified multinomial theorem) *Let* $x_1, \ldots, x_n$ *be* $n \in \mathbb{N}$ *real numbers (or any elements of some field). Then for every* $m \in \mathbb{N}$*:*

$$(\sum_{i=1}^{n} x_i)^m = \sum_{r=1}^{m} \sum_{\{a_1,\ldots,a_r\}\subseteq N_n} \sum_{\boldsymbol{k}\in J_{m,r}} \binom{m}{k_1 \ldots k_r} \prod_{s=1}^{r} x_{a_s}^{k_s}.$$

Proof Expanding the left term gives $(\sum_{i=1}^{n} x_i)^m = \sum_{\mathbf{j}\in N_n^m} \prod_{s=1}^{m} x_{j_s}$. Now, for every summation index $\mathbf{j} = (j_1, \ldots, j_m)$, there is the unique set $\{a_1, \ldots, a_r\} \subseteq N_n$ of $r \in \{1, \ldots, m\}$ pairwise distinct components with pertaining multiplicities $k_1, \ldots, k_r$. Notice that $k_1 + \ldots + k_r = m$, whence $\mathbf{k} = (k_1, \ldots, k_r) \in J_{m,r}$. Thus, we can decompose the last sum into

$$\sum_{\mathbf{j}\in N_n^m} \prod_{s=1}^{m} x_{j_s} = \sum_{r=1}^{m} \sum_{\{a_1,\ldots,a_r\}\subseteq N_n} \sum_{\mathbf{k}\in J_{m,r}} \sum_{\mathbf{j}}{}' \prod_{s=1}^{r} x_{a_s}^{k_s},$$

where $\sum'$ extends over all $\mathbf{j} \in \{a_1, \ldots, a_r\}^m$ with each a_i occurring exactly k_i times, $1 \le i \le r$. Since $\sum'$ has $\binom{m}{k_1\ldots k_r}$ summands that do not depend on the summation index $\mathbf{j}$, the assertion follows. □

Proof of Lemma 1.1 From Theorem 2.1 we can infer that

$$\begin{aligned}
&\mathbb{E}[(\sum_{i=1}^{n} X_i)^m] \\
&= \sum_{r=1}^{m} \sum_{\{a_1,\ldots,a_r\}\subseteq N_n} \sum_{\mathbf{k}\in J_{m,r}} \binom{m}{k_1 \ldots k_r} \mathbb{E}[\prod_{s=1}^{r} X_{a_s}^{k_s}] \quad \text{by linearity} \\
&= \sum_{r=1}^{m} \sum_{\{a_1,\ldots,a_r\}\subseteq N_n} \sum_{\mathbf{k}\in J_{m,r}} \binom{m}{k_1 \ldots k_r} \prod_{s=1}^{r} \mathbb{E}[X_{a_s}^{k_s}] \quad \text{by independence,} \qquad (2.1)
\end{aligned}$$

which shows the first part (1).

From our assumption in part (2) it follows that $\mathbb{E}[X_{a_s}^{k_s}] = \mu_{k_s}$ for all $1 \le s \le r$ and so the right-hand side in (2.1) simplifies to

$$\sum_{r=1}^{m} \sum_{\{a_1,\ldots,a_r\}\subseteq N_n} \sum_{\mathbf{k}\in J_{m,r}} \binom{m}{k_1 \ldots k_r} \prod_{s=1}^{r} \mu_{k_s} = \sum_{r=1}^{m} \binom{n}{r} \sum_{\mathbf{k}\in J_{m,r}} \binom{m}{k_1 \ldots k_r} \prod_{s=1}^{r} \mu_{k_s}. \qquad (2.2)$$

Next note that in the inner sum on the right-hand side of (2.2) none of the summands

$$s(\mathbf{k}) = s(k_1, \ldots, k_r) := \binom{m}{k_1 \ldots k_r} \prod_{s=1}^{r} \mu_{k_s}$$

changes if the arguments $k_1, \ldots, k_r$ are permuted arbitrarily. Consider an arbitrary summation index $\mathbf{k} \in J_{m,r}$. It is an ordered integer partition of m with, let us say, a distinct parts $(\kappa_1, \ldots, \kappa_a)$ and pertaining multiplicities

$(n_1,\dots,n_a)$. Since summation is invariant with respect to interchanging the summands, there are exactly $q := \binom{r}{n_1 \dots n_a}$ different ordered integer partitions $\mathbf{k}_1,\dots,\mathbf{k}_q$ (including the original $\mathbf{k}$) with the same distinct parts $(\kappa_1,\dots,\kappa_a)$ and multiplicities $(n_1,\dots,n_a)$. In particular, every $\mathbf{k}_i$ emerges from $\mathbf{k}$ by a permutation so that all summands $s(\mathbf{k}_i)$ coincide, $1 \le i \le q$. Finally, among the set $\{\mathbf{k}_1,\dots,\mathbf{k}_q\}$ there is exactly one element in $I_{m,r}$ that without loss of generality is $\mathbf{k}_1$. Thus, $\sum_{1\le i\le q} s(\mathbf{k}_i) = q\ s(\mathbf{k}_1) = \binom{r}{n_1 \dots n_a}\binom{m}{k_1 \dots k_r}\ \mu_{\kappa_1}^{n_1}\cdots\mu_{\kappa_a}^{n_a}$, whence

$$\sum_{\mathbf{k}\in J_{m,r}} \binom{m}{k_1 \dots k_r} \prod_{s=1}^{r} \mu_{k_s} = \sum_{\mathbf{k}\in I_{m,r}} \binom{m}{k_1 \dots k_r} \binom{r}{n_1 \dots n_a} \prod_{s=1}^{a} \mu_{\kappa_s}^{n_s}$$

and the assertion (1.2) with constants $A_{r,m}$ given in (1.3) follows by (2.1) and (2.2).

For the proof of (1.4), let ϕ_n and ϕ denote the characteristic functions of S_n and X, respectively. Then it is well known from probability theory (see, e.g., Resnick [13], chapter 9) that

$$\phi_n(t) = \{\phi(t)\}^n$$

and that

$$\phi_n^{(m)}(0) = i^m\ \mathbb{E}(S_n^m) \quad \text{where} \quad \phi_n^{(m)}(t) := \frac{d^m}{dt^m}\phi_n(t).$$

Now, the formula of Faà di Bruno gives the mth derivative of a composite function $f\circ g$; see, e.g., Johnson [10]:

$$(f\circ g)^{(m)}(t) = \sum_{\mathbf{l}\in L_m} \frac{m!}{l_1!\cdots l_m!} f^{(l_1+\ldots+l_m)}(g(t)) \prod_{j=1}^{m} \{\frac{1}{j!} g^{(j)}(t)\}^{l_j},$$

where

$$L_m := \{\mathbf{l}\in\mathbb{N}_0^m : \sum_{i=1}^{m} i\ l_i = m\} = \biguplus_{r=1}^{m} L_{m,r}.$$

An application of this formula with $f(t) = t^n$ and $g(t) = \phi(t)$ and taking into account that $\phi(0) = 1$ and $\phi^{(j)}(0) = i^j\mu_j$ yields:

$$\begin{aligned} \mathbb{E}(S_n^m) &= \sum_{\mathbf{l}\in L_m} \frac{m!}{l_1!\cdots l_m!} f^{(l_1+\ldots+l_m)}(1) \prod_{j=1}^{m} \{\frac{1}{j!}\mu_j)\}^{l_j} \\ &= \sum_{r=1}^{m} r!\binom{n}{r} \sum_{\mathbf{l}\in L_{m,r}} \frac{m!}{l_1!\cdots l_m!} \prod_{j=1}^{m} \{\frac{1}{j!}\mu_j)\}^{l_j}. \end{aligned} \tag{2.3}$$

From this, (1.4) follows immediately, once we have shown that

$$\mathbf{l}\in L_{m,r} \quad\Rightarrow\quad l_j = 0 \quad \forall\, j > m-r+1\ . \tag{2.4}$$

In fact, assume that there is some $j > m-r+1$ with $l_j \ge 1$. Then we can conclude that

$$m = \sum_{i=1}^{m} i l_i \ge \sum_{1\le i\le m, i\ne j} l_i\ + j l_j = \sum_{i=1}^{m} l_i + (j-1)l_j = r + (j-1)l_j \ge r + (j-1) > r + (m-r) = m,$$

which results in the contradiction $m > m$. With (2.3) and (2.4), the proof of (1.4) is complete.

Proof of Theorem 1.2 According to (2.1) and (2.2) we have that

$$\mathbb{E}(S_n^m) = \sum_{r=1}^{m} \binom{n}{r} \sum_{\mathbf{k}\in J_{m,r}} \binom{m}{k_1 \dots k_r} \prod_{s=1}^{r} \mu_{k_s}. \tag{2.5}$$

Recall that $\mu_1 = 0$, whence the product $\prod_{s=1}^{r} \mu_{k_s}$ vanishes if there exists at least one index such that $k_s = 1$. At any rate, this applies to each $r > m/2$, because otherwise $k_s \geq 2 \; \forall \; 1 \leq s \leq r$ and therefore $k_1 + \ldots + k_r \geq 2r > m$ in contradiction to $\mathbf{k} \in J_{m,r}$. Thus, the sum in (2.5) reduces to

$$\sum_{1\leq r\leq m/2} \binom{n}{r} {\sum_{\mathbf{k}\in J_{m,r}}}' \binom{m}{k_1 \dots k_r} \prod_{s=1}^{r} \mu_{k_s} = \sum_{1\leq r\leq m/2} \binom{n}{r} \sum_{\mathbf{k}\in J_{m-r,r}+\mathbf{1}} \binom{m}{k_1 \dots k_r} \prod_{s=1}^{r} \mu_{k_s}, \tag{2.6}$$

where $\sum'$ extends over all $\mathbf{k} \in J_{m,r}$ with $k_s \geq 2 \; \forall \; 1 \leq s \leq r$, which is equivalent to $\mathbf{k} \in J_{m-r,r} + \mathbf{1}$.

Now, note that in the inner sum on the right-hand side of (2.6), none of the summands

$$s(\mathbf{k}) = s(k_1, \dots, k_r) := \binom{m}{k_1 \dots k_r} \prod_{s=1}^{r} \mu_{k_s}$$

change if the components $k_1, \dots, k_r$ are arranged in an arbitrary order. Therefore, the same arguments as in the derivation of (1.3) (with $J_{m,r}$ there replaced by $J_{m-r,r} + \mathbf{1}$ here) lead to

$$\sum_{\mathbf{k}\in J_{m-r,r}+\mathbf{1}} \binom{m}{k_1 \dots k_r} \prod_{s=1}^{r} \mu_{k_s} = \sum_{\mathbf{k}\in I_{m-r,r}+\mathbf{1}} \binom{m}{k_1 \dots k_r} \binom{r}{n_1 \dots n_a} \prod_{s=1}^{a} \mu_{\kappa_s}^{n_s}$$

and the assertion (1.5) with constants $B_{r,m}$ given in (1.6) follows by (2.5) and (2.6).

As to the proof of (1.7), recall that by (2.3)

$$\mathbb{E}(S_n^m) = \sum_{r=1}^{m} r! \binom{n}{r} \sum_{\mathbf{l}\in L_{m,r}} \frac{m!}{l_1! \cdots l_m!} \prod_{j=1}^{m} \{\frac{1}{j!}\mu_j)\}^{l_j} .$$

In view of $\mu_1 = 0$, the product vanishes for all $\mathbf{l} \in L_{m,r}$ with $l_1 \geq 1$. The latter is true for all $r > m/2$, because otherwise $l_1 = 0$ and one had $m < 2r = 2\sum_{1\leq i\leq m} l_i = \sum_{2\leq i\leq m} 2l_i \leq \sum_{2\leq i\leq m} il_i = \sum_{1\leq i\leq m} il_i = m$, a contradiction.

Thus, in the last displayed formula the summation can be reduced to $1 \leq r \leq m/2$ and we obtain:

$$\mathbb{E}(S_n^m) = \sum_{1\leq r\leq m/2} r! \binom{n}{r} \sum_{\mathbf{l}\in L_{m,r}} \frac{m!}{l_1! \cdots l_m!} \prod_{j=1}^{m} \{\frac{1}{j!}\mu_j)\}^{l_j} . \tag{2.7}$$

Recall that in the sum over $\mathbf{l} \in L_{m,r}$, all those summands vanish with $\mathbf{l}$ such $l_1 \geq 1$. In view of the remaining summands observe that

$$\mathbf{l} = (0, l_2, \dots, l_m) \in L_{m,r} \Leftrightarrow \mathbf{l} = (0, l_2, \dots, l_{m-r+1}, 0, \dots, 0) \text{ with } (l_2, \dots, l_{m-r+1}) \in L_{m-r,m}. \tag{2.8}$$

Indeed, for $\mathbf{l} = (0, l_2, \dots, l_m) \in L_{m,r}$ we see from (2.4) that $\mathbf{l} = (0, l_2, \dots, l_{m-r+1}, 0, \dots, 0)$ and consequently

$$m - r = \sum_{i=2}^{m}(il_i - l_i) = \sum_{i=2}^{m}(i-1)l_i = \sum_{2 \le i \le m-r+1}(i-1)l_i = \sum_{1 \le j \le m-r} jl_{j+1}.$$

Moreover, $\sum_{i=1}^{m-r} l_{i+1} = \sum_{i=1}^{m} l_i = r$, whence $(l_2, \dots, l_{m-r+1}) \in L_{m-r,r}$. The reverse direction in (2.8) follows, because the last 2 equalities also show that if $(l_2, \dots, l_{m-r+1}) \in L_{m-r,r}$ then $\mathbf{l} = (0, l_2, \dots, l_{m-r+1}, 0, \dots, 0) \in L_{m,r}$.

From (2.7) and (2.8) we immediately obtain that

$$\begin{aligned}
\mathbb{E}(S_n^m) &= \sum_{1 \le r \le m/2} r!\binom{n}{r} \sum_{(l_2,\dots,l_{m-r+1}) \in L_{m-r,r}} \frac{m!}{l_2! \cdots l_{m-r+1}!} \prod_{j=2}^{m-r+1} \{\frac{1}{j!}\mu_j\}^{l_j} \\
&= \sum_{1 \le r \le m/2} r!\binom{n}{r} \sum_{(l_1,\dots,l_{m-r}) \in L_{m-r,r}} \frac{m!}{l_1! \cdots l_{m-r}!} \prod_{j=1}^{m-r} \{\frac{1}{(j+1)!}\mu_{j+1}\}^{l_j} \\
&= \sum_{1 \le r \le m/2} \binom{n}{r} r! m! \sum_{\mathbf{l} \in L_{m-r,r}} \prod_{j=1}^{m-r} \frac{\{\frac{1}{(j+1)!}\mu_{j+1}\}^{l_j}}{l_j!},
\end{aligned}$$

which yields the desired result, (1.7).

Proof of Theorem 1.3 By (2.5) and (2.6) we have that

$$\mathbb{E}[(\sum_{i=1}^{n} X_i)^{2p}] = \sum_{1 \le r \le p} \binom{n}{r} {\sum_{\mathbf{k} \in J_{2p,r}}}' \binom{2p}{k_1 \dots k_r} \prod_{s=1}^{r} \mu_{k_s}, \tag{2.9}$$

where $\sum'$ extends over all $\mathbf{k} \in J_{2p,r}$ with $k_s \ge 2 \ \forall \ 1 \le s \le r$. Since all odd moments of X vanish by symmetry we can further infer that the product $\mu_{k_1} \cdots \mu_{k_r}$ vanishes if at least one of the indices $k_1, \dots, k_r$ is odd. Therefore, the summation reduces to all tuples $\mathbf{k} = (2l_1, \dots, 2l_r) = 2\mathbf{l}$ with $2l_1 + \dots + 2l_r = 2p$ and $2l_s \ge 2 \, \forall \, 1 \le s \le r$, i.e. $\mathbf{l} \in J_{p,r}$. Conclude that

$${\sum_{\mathbf{k} \in J_{2p,r}}}' \binom{2p}{k_1 \dots k_r} \prod_{s=1}^{r} \mu_{k_s} = \sum_{\mathbf{k} \in J_{p,r}} \binom{2p}{2k_1 \dots 2k_r} \prod_{s=1}^{r} \mu_{2k_s}.$$

Again, all involved summands in the last sum are permutation-invariant so that arguing analogously as in the proofs above leads to

$$\sum_{\mathbf{k} \in J_{p,r}} \binom{2p}{2k_1 \dots 2k_r} \prod_{s=1}^{r} \mu_{2k_s} = \sum_{\mathbf{k} \in I_{p,r}} \binom{2p}{2k_1 \dots 2k_r} \binom{r}{n_1 \dots n_a} \mu_{2\kappa_1}^{n_1} \dots \mu_{2\kappa_a}^{n_a},$$

which gives (1.9) and (1.10).

For the derivation of (1.11) we may use (2.7) with $m = 2p$ as the starting point:

$$\mathbb{E}(S_n^{2p}) = \sum_{1 \le r \le p} r!(2p)!\binom{n}{r} \sum_{\mathbf{l} \in L_{2p,r}} \prod_{j=1}^{2p} \frac{\{\frac{1}{j!}\mu_j\}^{l_j}}{l_j!}. \tag{2.10}$$

Since $\mu_j = 0$ for odd indices j, all those summands vanish that belong to an $\mathbf{l} \in L_{2p,r}$ such that there is at least one nonzero component l_j with odd j. Thus, the remaining summands possess indices $\mathbf{l} = (0, l_2, 0, l_4, 0, \ldots, 0, l_{2p})$. For such an index $\mathbf{l}$ we observe that:

$$r = \sum_{i=1}^{2p} l_i = \sum_{i=1}^{p} l_{2i} \quad \text{and} \quad 2p = \sum_{i=1}^{2p} i l_i = 2\sum_{i=1}^{p} i l_{2i},$$

whence $(l_2, l_4, \ldots, l_{2p}) \in L_{p,r}$. On the other hand, if $(l_2, l_4, \ldots, l_{2p}) \in L_{p,r}$ then by the above equalities we know that $(0, l_2, 0, l_4, 0, \ldots, 0, l_{2p}) \in L_{2p,r}$. Therefore, equality (2.10) can be restated as

$$\begin{aligned} \mathbb{E}(S_n^{2p}) &= \sum_{1 \le r \le p} r!(2p)! \binom{n}{r} \sum_{(l_2, l_4, \ldots, l_{2p}) \in L_{p,r}} \prod_{j=1}^{p} \frac{\{\frac{1}{(2j)!}\mu_{2j}\}^{l_{2j}}}{l_{2j}!} \\ &= \sum_{1 \le r \le p} r!(2p)! \binom{n}{r} \sum_{\mathbf{l} \in L_{p,r}} \prod_{j=1}^{p} \frac{\{\frac{1}{(2j)!}\mu_{2j}\}^{l_j}}{l_j!} \\ &= \sum_{1 \le r \le p} r!(2p)! \binom{n}{r} \sum_{\mathbf{l} \in L_{p,r}} \prod_{j=1}^{p-r+1} \frac{\{\frac{1}{(2j)!}\mu_{2j}\}^{l_j}}{l_j!}, \end{aligned}$$

where the last equality holds by (2.4). This finishes our proof.

Proof of Theorem 1.4 It follows from (1.1) in Lemma 1.1 that:

$$\mathbb{E}[(\sum_{i=1}^{n} X_i)^{2p}] = \sum_{r=1}^{2p} \sum_{\{a_1,\ldots,a_r\} \subseteq N_n} \sum_{\mathbf{k} \in J_{2p,r}} \binom{2p}{k_1 \ldots k_r} \prod_{s=1}^{r} \mathbb{E}[X_{a_s}^{k_s}] \,.$$

By assumption, $\mathbb{E}[X_{a_s}] = 0$ for all $1 \le s \le r$ and consequently the product $\prod_{s=1}^{r} \mathbb{E}[X_{a_s}^{k_s}]$ vanishes if there is at least one index k_s equal to one. As already pointed out in the treatment of (2.5), this is true for every $\mathbf{k} \in J_{2p,r}$ as long as $r > p$. Therefore, we obtain

$$\mathbb{E}[(\sum_{i=1}^{n} X_i)^{2p}] = \sum_{r=1}^{p} \sum_{\{a_1,\ldots,a_r\} \subseteq N_n} \sum_{\mathbf{k} \in J_{2p,r}} \binom{2p}{k_1 \ldots k_r} \prod_{s=1}^{r} \mathbb{E}[X_{a_s}^{k_s}] \,.$$

Next, observe that by symmetry $\mathbb{E}[X_{a_s}^{k_s}] = 0$ whenever k_s is odd, so that the product $\prod_{s=1}^{r} \mathbb{E}[X_{a_s}^{k_s}]$ vanishes if there is at least one odd index k_s. Thus, in the inner sum only those summands pertaining to $\mathbf{k} = (2l_1, \ldots, 2l_r) \in J_{2p,r}$ with solely even components do remain. Since $(2l_1, \ldots, 2l_r) \in J_{2p,r}$ is equivalent to $(l_1, \ldots, l_r) \in J_{p,r}$, we see that

$$\mathbb{E}[(\sum_{i=1}^{n} X_i)^{2p}] = \sum_{r=1}^{p} \sum_{\{a_1,\ldots,a_r\} \subseteq N_n} \sum_{\mathbf{k} \in J_{p,r}} \binom{2p}{2k_1 \ldots 2k_r} \prod_{s=1}^{r} \mathbb{E}[X_{a_s}^{2k_s}] \,.$$

Notice that the product is nonnegative and that

$$\binom{2p}{2k_1 \ldots 2k_r} = \frac{\binom{2p}{2k_1 \ldots 2k_r}}{\binom{p}{k_1 \ldots k_r}} \binom{p}{k_1 \ldots k_r} \le M_p \binom{p}{k_1 \ldots k_r},$$

where

$$M_p := \max\{\frac{\binom{2p}{2k_1 \ldots 2k_r}}{\binom{p}{k_1 \ldots k_r}} : \mathbf{k} \in J_{p,r},\ 1 \le r \le p\},$$

and so

$$\begin{aligned}
\mathbb{E}[(\sum_{i=1}^{n} X_i)^{2p}] &\le M_p \sum_{r=1}^{p} \sum_{\{a_1,\ldots,a_r\} \subseteq N_n} \sum_{\mathbf{k} \in J_{p,r}} \binom{p}{k_1 \ldots k_r} \prod_{s=1}^{r} \mathbb{E}[X_{a_s}^{2k_s}] \\
&= M_p \sum_{r=1}^{p} \sum_{\{a_1,\ldots,a_r\} \subseteq N_n} \sum_{\mathbf{k} \in J_{p,r}} \binom{p}{k_1 \ldots k_r} \mathbb{E}[\prod_{s=1}^{r} X_{a_s}^{2k_s}] \quad \text{by independence} \\
&= M_p \, \mathbb{E}\big[\sum_{r=1}^{p} \sum_{\{a_1,\ldots,a_r\} \subseteq N_n} \sum_{\mathbf{k} \in J_{p,r}} \binom{p}{k_1 \ldots k_r} \prod_{s=1}^{r} X_{a_s}^{2k_s}\big] \quad \text{by linearity} \\
&= M_p \, \mathbb{E}\big[\{\sum_{i=1}^{n} X_i^2\}^p\big] \quad \text{by Theorem 2.1.}
\end{aligned}$$

If in the above arguments we replace M_p by

$$m_p := \min\{\frac{\binom{2p}{2k_1 \ldots 2k_r}}{\binom{p}{k_1 \ldots k_r}} : \mathbf{k} \in J_{p,r},\ 1 \le r \le p\},$$

we obtain

$$\mathbb{E}[(\sum_{i=1}^{n} X_i)^{2p}] \ge m_p \, \mathbb{E}\big[\{\sum_{i=1}^{n} X_i^2\}^p\big].$$

For the computation of the constant M_p observe that for every $\mathbf{k} \in J_{p,r}$ and each $1 \le r \le p$ one has by (1.8) that

$$\begin{aligned}
\frac{\binom{2p}{2k_1 \ldots 2k_r}}{\binom{p}{k_1 \ldots k_r}} &= \frac{(2p)!}{(2k_1)! \ldots (2k_r)!} \frac{k_1! \ldots k_r!}{p!} = \frac{(2p)!}{2^{k_1 + \ldots k_r} p!} \frac{k_1! 2^{k_1}}{(2k_1)!} \cdots \frac{k_r! 2^{k_r}}{(2k_r)!} \\
&= \frac{(2p)!}{2^p p!} \, \{\prod_{j=1}^{r} \frac{(2k_j)!}{k_j! 2^{k_j}}\}^{-1} \\
&= \frac{(2p-1)!!}{\prod_{j=1}^{r} (2k_j - 1)!!}.
\end{aligned}$$

Since each factor $(2k_j - 1)!!$ is greater or equal to one, so is the product. In fact, this lower bound is attained for $\mathbf{k} = (1, \ldots, 1) \in J_{p,p}$, whence

$$M_p = (2p-1)!!.$$

As to the constant m_p, note that $\{\frac{\binom{2p}{2k_1 \ldots 2k_r}}{\binom{p}{k_1 \ldots k_r}} : \mathbf{k} \in J_{p,r},\ 1 \le r \le p\}$ is a set of positive integers, which in particular contains 1 (take $r = 1$ and observe that $J_{p,1} = \{p\}$.) This shows (1.13).

It remains to show optimality. To this end consider X_i with $\mathbb{P}(X_i = 1) = \mathbb{P}(X_i = -1) = \frac{1}{2}, 1 \le i \le n$. Then the upper inequality in (1.14) can be restated as

$$\mathbb{E}[(\sum_{i=1}^{n} X_i)^{2p}] \le C_p\, n^p$$

and it follows from the CLT, the continuous mapping theorem, and Theorem 3.4 of Billingsley [2] that:

$$C_p \ge \liminf_{n\to\infty} \mathbb{E}[(n^{-\frac{1}{2}} \sum_{i=1}^{n} X_i)^{2p}] \ge \mathbb{E}[N(0,1)^{2p}] = (2p-1)!!\,.$$

Finally, it follows immediately from the lower inequality in (1.14) with $n = 1$ that $D_p \le 1$, and so the proof is complete.

3. Applications

3.1. The CLT for the empirical bootstrap

The empirical bootstrap is a well-established method in statistics. Here, one generates a new random sample $X_1^*, \dots, X_n^*$ from the empirical distribution function, say F_n, pertaining to the original sample $X_1, \dots, X_n$ with distribution function, say, F. To understand the basic idea, consider a certain statistic (e.g., a standardized estimator) $T_n = T_n(X_1, \dots, X_n, F)$, which is known to have some distributional limit T, e.g., the normal distribution. Then, if T_n is replaced by its bootstrap version $T_n^* = T_n(X_1^*, \dots, X_n^*, F_n)$, it very often turns out that the distribution of T_n^* gives a better approximation to the distribution of T_n than the limit T does. Now, even though the distribution of T_n^* usually is analytically intractable, it can be approximated with arbitrary accuracy by using the Monte-Carlo method. For an introduction to the bootstrap method we recommend the textbook of Efron and Tibshirani [6], whereas the monograph of Hall [9] gives advanced insight into the theory.

The precise formal description can be as follows: let $(X_i, i \in \mathbb{N})$ be a sequence of i.i.d. random variables defined on some probability space $(\Omega, \mathcal{A}, \mathbb{P})$. The empirical distribution function F_n pertaining to $X_1, \dots, X_n$ is defined by

$$F_n(\omega, x) := \frac{1}{n} \sum_{i=1}^{n} 1_{\{X_i(\omega) \le x\}}, \quad \omega \in \Omega,\ x \in \mathbb{R}.$$

Furthermore, let $(U_i, i \in \mathbb{N})$ be a sequence of i.i.d. random variables with the uniform distribution on $(0,1)$ defined on some other probability space $(\Omega^*, \mathcal{A}^*, \mathbb{P}^*)$. If F_n^{-1} denotes the quantile function of F_n, then we put for every $1 \le i \le n$ and each $n \in \mathbb{N}$:

$$X_{in}^*(\omega, \omega^*) := F_n^{-1}(\omega, U_i(\omega^*)), \quad \omega \in \Omega,\ \omega^* \in \Omega^*.$$

Then our (canonical) construction ensures that for every $n \in \mathbb{N}$ and for all $\omega \in \Omega$

$$X_{1n}^*(\omega, \cdot), \dots, X_{nn}^*(\omega, \cdot) \quad \text{are i.i.d. with respect to } \mathbb{P}^* \tag{3.1}$$

and have distribution function $F_n(\omega, \cdot)$. In the sequel we will shortly write X_i^* for $X_{in}^*(\omega, \cdot)$ and simply $F_n(x)$ instead of $F_n(\omega, x)$, but still keep in mind that these quantities depend on ω.

The prototype of an example for a statistic T_n as mentioned above is

$$T_n = T_n(X_1, \dots, X_n, F) = \frac{1}{\sqrt{n\sigma^2}} \sum_{i=1}^{n} (X_i - \mu)$$

with $\mu = \mathbb{E}[X_1] = \int x \; F(dx)$ and $\sigma^2 = \mathrm{Var}(X_1) = \int (x-\mu)^2 \; F(dx)$. Since

$$\mathbb{E}^*[X_1^*] = \int x \; F_n(dx) = \frac{1}{n} \sum_{i=1}^{n} X_i =: \bar{X}_n \tag{3.2}$$

and

$$\mathrm{Var}^*[X_1^*] = \int (x - \bar{X}_n) \; F_n(dx) = \frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X}_n)^2 =: s_n^2,$$

the corresponding bootstrap version is given by

$$T_n^* = T_n(X_1^*, \dots, X_n^*, F_n) = \frac{1}{\sqrt{n s_n^2}} \sum_{i=1}^{n} (X_i^* - \bar{X}_n). \tag{3.3}$$

Now, we will give a simple derivation of:

Theorem 3.1 (CLT for bootstrap) *If all moments of X_1 exist and are finite, then*

$$T_n^* \stackrel{\mathcal{L}}{\to} N(0,1), \; n \to \infty, \quad \textit{for } \mathbb{P} - \textit{almost all } \omega \in \Omega. \tag{3.4}$$

Proof By the method of moments (see, e.g., Shorack [15], Theorem 8.2 in chapter 11), it suffices to show that

$$\mathbb{E}^*[\{T_n^*\}^m] \;\to\; \mathbb{E}[N(0,1)^m], \; n \to \infty \quad \text{ for all } \; m \in \mathbb{N} \quad \mathbb{P}\text{-almost surely}, \tag{3.5}$$

where it is well known that $\mathbb{E}[N(0,1)^m] = (m-1)!!$ for even moment-order m and $\mathbb{E}[N(0,1)^m] = 0$ for odd m.

Now,

$$\mathbb{E}^*[\{T_n^*\}^m] = \{n s_n^2\}^{-\frac{m}{2}} \; \mathbb{E}^*[\{\sum_{i=1}^{n} (X_i^* - \bar{X}_n)\}^m] \tag{3.6}$$

and by (3.1) and (3.2) we may apply Theorem 1.2 to the variables $X_i^* - \bar{X}_n$ for computing the expectation on the right side in (3.6). Here, recall that the constants in (1.6) or (1.7), respectively, depend on the moments μ_k there: $B_{r,m} = B_{r,m}(\mu_2, \dots, \mu_m)$. In our case, these moments are given by

$$\mathbb{E}^*[(X_1^* - \bar{X}_n)^k] = \frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X}_n)^k \; =: \mu_{k,n}, \quad 2 \le k \le m.$$

Thus with

$$B_{r,m}^{(n)} := B_{r,m}(\mu_{2,n}, \dots, \mu_{m,n}),$$

Theorem 1.2 and (3.6) yield

$$\mathbb{E}^*[\{T_n^*\}^m] = \{s_n^2\}^{-\frac{m}{2}} \sum_{r=1}^{[m/2]} B_{r,m}^{(n)} \; n^{-\frac{m}{2}} \binom{n}{r}. \tag{3.7}$$

Using the strong law of large numbers, the sample moments $\mu_{k,n}$ are easily seen to be strongly consistent; see, e.g., Serfling [14]. Consequently,

$$\mu_{k,n} \to \mu_k = \mathbb{E}[(X_1-\mu)^k],\ n \to \infty, \quad \text{for all } \ 2 \le k \le m \quad \mathbb{P}\text{-almost surely.} \tag{3.8}$$

So, by continuity,

$$B_{r,m}^{(n)} \to B_{r,m}(\mu_2,\ldots,\mu_m),\ n \to \infty, \quad \text{for all } \ 1 \le r \le [m/2] \quad \mathbb{P}\text{-almost surely} \tag{3.9}$$

and (because $s_n^2 = \mu_{2,n}$),

$$\{s_n^2\}^{-\frac{m}{2}} \to \{\mu_2\}^{-\frac{m}{2}} \quad \mathbb{P}\text{-almost surely.} \tag{3.10}$$

Finally, in view of (3.7) note that

$$n^{-\frac{m}{2}}\binom{n}{r} = \frac{1}{r!}n^{-\frac{m}{2}+r}(1+o(1)) \ \text{ as } n \to \infty. \tag{3.11}$$

Thus, if m is an odd integer, then $[m/2] < m/2$ so that by (3.9) and (3.11) every summand in (3.7) vanishes $\mathbb{P}$-almost surely as $n \to \infty$ and therefore by (3.10):

$$\mathbb{E}^*[\{T_n^*\}^m] \to 0,\ n \to \infty \quad \text{for all odd integers } m \quad \mathbb{P}\text{-almost surely.}$$

However, if $m = 2p, p \in \mathbb{N}$, is an even integer, then the (last) summand in (3.7) pertaining to $r = p$ remains and converges to $\frac{1}{p!}B_{p,m} = (m-1)!!\mu_2^p$. Consequently, by (3.10) we obtain that

$$\mathbb{E}^*[\{T_n^*\}^m] \to (m-1)!!,\ n \to \infty \quad \text{for all even integers } m \quad \mathbb{P}\text{-almost surely.}$$

Hence, we have shown (3.5) and the proof is finished. □

Remark 3.2 *(1) If we apply Theorem 1.2 to the centered random variables $X_i - \mu$, we obtain*

$$\mathbb{E}[\{T_n\}^m] = \{\sigma^2\}^{-\frac{m}{2}} \sum_{r=1}^{[m/2]} B_{r,m}\, n^{-\frac{m}{2}}\binom{n}{r}. \tag{3.12}$$

It follows immediately from (3.11) and (3.12) that

$$\mathbb{E}[\{T_n\}^m] \to 0,\ n \to \infty \quad \textit{for all odd integers } m \tag{3.13}$$

and

$$\mathbb{E}[\{T_n\}^m] \to (m-1)!!,\ n \to \infty \quad \textit{for all even integers } m. \tag{3.14}$$

Thus, Theorem 1.2 in combination with the method of moments also gives a very short proof of the classical CLT:

$$T_n \xrightarrow{\mathcal{L}} N(0,1),\ n \to \infty. \tag{3.15}$$

(2) We would like to mention that indeed our proofs of the CLTs both are rather brief and in particular are elementary. On the other hand, however, our moment assumption is more restrictive than in the classical formulations, where only the finiteness of the second moment $\mathbb{E}[X_1^2]$ is required; see, e.g., van der Vaart [16].

(3) Packwood [12] followed the same strategy to derive the classical CLT, but first he considered only symmetric summands, so that (3.13) was trivially fulfilled. Second, he stated (3.14), but gave a formal proof only for $m \in \{2,4,6\}$ and left the complete solution open.

3.2. Inequalities for self-normalized sums

In this section, we assume that the $X_1, \ldots, X_n$ are independent and symmetric with positive variances. The latter condition is only to exclude the degenerated case in which all X_i vanish with probability of one. Recall that $S_n := \sum_{i=1}^n X_i$ and put $V_n := \sqrt{\sum_{i=1}^n X_i^2}$, which by assumption is positive almost surely. Then

$$T_n := \frac{S_n}{V_n}$$

is called a self-normalized sum. The nomenclature comes from the fact that the variance of T_n is equal to one, provided that the variances of the X_i are finite.

Theorem 3.3 *Let p be any positive integer and assume that $X_1, \ldots, X_n$ are independent, symmetric, and $2p$-integrable. Then*

$$\mathbb{E}[T_n^{2p}] \leq (2p-1)!! \,. \tag{3.16}$$

Proof We use a so-called symmetrization argument. Let $(\epsilon_1, \ldots, \epsilon_n)$ be i.i.d. with

$\mathbb{P}(\epsilon_i = 1) = \mathbb{P}(\epsilon_i = -1) = 1/2$ and such that $(\epsilon_1, \ldots, \epsilon_n)$ and $(X_1, \ldots, X_n)$ are independent. Then by independence and symmetry it follows that

$$(X_1, \ldots, X_n) \stackrel{\mathcal{L}}{=} (\epsilon_1 X_1, \ldots, \epsilon_n X_n),$$

and so (upon noticing that $\epsilon_i^2 = 1$ a.s.)

$$T_n \stackrel{\mathcal{L}}{=} \frac{\sum_{i=1}^n \epsilon_i X_i}{\sqrt{\sum_{i=1}^n (\epsilon_i X_i)^2}} = \frac{\sum_{i=1}^n \epsilon_i X_i}{\sqrt{\sum_{i=1}^n X_i^2}} = \frac{\sum_{i=1}^n \epsilon_i X_i}{V_n}. \tag{3.17}$$

Thus, conditioning on $(X_1, \ldots, X_n)$ yields

$$\begin{aligned} &\mathbb{E}[T_n^{2p}] \\ = \;& \mathbb{E}\Big[\frac{\{\sum_{i=1}^n X_i \,\epsilon_i\}^{2p}}{V_n^{2p}}\Big] = \mathbb{E}\Big[\mathbb{E}\Big(\frac{\{\sum_{i=1}^n X_i \,\epsilon_i\}^{2p}}{V_n^{2p}} | X_1, \ldots, X_n\Big)\Big] \\ = \;& \mathbb{E}\Big[V_n^{-2p}\, \mathbb{E}\Big(\{\sum_{i=1}^n X_i \,\epsilon_i\}^{2p} | X_1, \ldots, X_n\Big)\Big]. \end{aligned} \tag{3.18}$$

As to the conditional expectation in (3.18), notice that by independence

$$\mathbb{E}\Big[\{\sum_{i=1}^n X_i \,\epsilon_i\}^{2p} | X_1 = x_1, \ldots, X_n = x_n\Big] = \mathbb{E}\Big[\{\sum_{i=1}^n x_i \,\epsilon_i\}^{2p}\Big] \tag{3.19}$$

for ν-almost every $(x_1, \ldots, x_n) \in \mathbb{R}^n$, where $\nu := \mathbb{P} \circ (X_1, \ldots, X_n)^{-1}$ denotes the distribution of the vector $(X_1, \ldots, X_n)$. An application of Theorem 1.4 to the variables $x_i \epsilon_i$ ensures that

$$\mathbb{E}\Big[\{\sum_{i=1}^n x_i \,\epsilon_i\}^{2p}\Big] \leq (2p-1)!! \; \mathbb{E}[\{\sum_{i=1}^n (x_i \epsilon_i)^2\}^p] = (2p-1)!! \; \{\sum_{i=1}^n x_i^2\}^p, \tag{3.20}$$

recalling that $\epsilon_i^2 = 1$ $\mathbb{P}$-almost surely. Consequently, by (3.19)

$$\mathbb{E}\Big(\{\sum_{i=1}^n X_i\ \epsilon_i\}^{2p}|X_1,\ldots,X_n\Big)\Big] \le (2p-1)!!\ \{\sum_{i=1}^n X_i^2\}^p = (2p-1)!!\ V_n^{2p} \quad \mathbb{P}-\text{a.s.},$$

which in view of (3.18) yields the desired result. □

Theorem 3.3 improves the result of Egorov [7,8], who found the weaker upper bound $Ce^{-p}(2p)^p \ge \sqrt{2}e^{-p}(2p)^p > (2p-1)!!$; refer to our Remark 1.5 (1).

Another application of the symmetrization in combination with Theorem 1.4 gives the exponential inequality of Efron [5]. Here, we do not require any moment condition.

Theorem 3.4 (Efron) *Let $X_1,\ldots,X_n$ be independent and symmetric with positive variances, which need not be finite. Then*

$$\mathbb{P}(T_n \ge x) \le \exp\{-\frac{1}{2}x^2) \quad \textit{for all } x > 0. \tag{3.21}$$

Proof According to (3.17) we have that

$$\mathbb{P}(T_n \ge x) = \mathbb{P}\Big(\frac{\sum_{i=1}^n X_i\ \epsilon_i}{V_n} \ge x\Big) = \mathbb{E}\Big[\mathbb{P}\Big(\sum_{i=1}^n X_i\ \epsilon_i \ge xV_n\ |X_1,\ldots,X_n\Big)\Big]. \tag{3.22}$$

For every fixed point $(x_1,\ldots,x_n) \in \mathbb{R}^n$, we put $Z_n := \sum_{i=1}^n x_i\epsilon_i$ and $v_n := \sqrt{\sum_{i=1}^n x_i^2}$. It follows by independence that for all $t > 0$:

$$\begin{aligned}
&\mathbb{P}\Big(\sum_{i=1}^n X_i\ \epsilon_i \ge xV_n\ |X_1 = x_1,\ldots,X_n = x_n\Big)\\
&= \mathbb{P}\Big(\sum_{i=1}^n x_i\ \epsilon_i \ge xv_n\Big) = \mathbb{P}\Big(Z_n \ge xv_n\Big)\\
&\le e^{-txv_n}\mathbb{E}[e^{tZ_n}] \qquad \text{by Markov's inequality}\\
&= e^{-txv_n}\mathbb{E}[\sum_{k=0}^\infty \frac{(tZ_n)^k}{k!}]\\
&= e^{-txv_n}\sum_{k=0}^\infty \frac{t^k}{k!}\mathbb{E}[Z_n^k]
\end{aligned}$$

by the dominated convergence theorem. In fact, observe that for every $m \in \mathbb{N}_0$ the partial sums $Y_m := \sum_{k=0}^m \frac{(tZ_n)^k}{k!}$ are bounded:

$$|Y_m| \le \sum_{k=0}^m \frac{|tZ_n|^k}{k!} \le \sum_{k=0}^\infty \frac{|tZ_n|^k}{k!} = e^{|t||Z_n|} \le e^{|t|\sum_{i=1}^n |x_i|} \quad \forall\, m \in \mathbb{N}_0,$$

where the last inequality holds, because $|Z_n| = |\sum_{i=1}^n x_i \epsilon_i| \le \sum_{i=1}^n |x_i||\epsilon_i| \le \sum_{i=1}^n |x_i|$. Thus, the constant $e^{|t| \sum_{i=1}^n |x_i|}$ is a $\mathbb{P}$-integrable majorant for (Y_m) and therefore the dominated convergence theorem is applicable.

It follows that

$$\begin{aligned}
&\mathbb{P}\Big(\sum_{i=1}^n X_i \, \epsilon_i \ge x V_n \, |X_1 = x_1, \ldots, X_n = x_n\Big) \\
&= \quad e^{-txv_n} \sum_{p=0}^\infty \frac{t^{2p}}{(2p)!} \mathbb{E}[Z_n^{2p}] \qquad \text{by symmetry of } Z_n \\
&\le \quad e^{-txv_n} \sum_{p=0}^\infty \frac{t^{2p}}{(2p)!} (2p-1)!! v_n^{2p} \qquad \text{by (3.20)} \\
&= \quad e^{-txv_n} \sum_{p=0}^\infty \frac{(\frac{1}{2} t^2 v_n^2)^p}{p!} \qquad \text{by (1.8)} \\
&= \quad e^{-txv_n} e^{\frac{1}{2} t^2 v_n^2} = e^{-txv_n + \frac{1}{2} t^2 v_n^2} \, . \qquad (3.23)
\end{aligned}$$

The last expression in (3.23) is minimal for $t = x/v_n$, whence we obtain:

$$\mathbb{P}\Big(\sum_{i=1}^n X_i \, \epsilon_i \ge x V_n \, |X_1, \ldots, X_n\Big) \le \exp\{-\frac{1}{2} x^2\} \quad \mathbb{P} - a.s.$$

and the assertion follows from (3.22). □

Appendix

Example For $m = 11$ we obtain the following polynomial expansion:

$$\begin{aligned}
M_n(11) \quad = \quad & n^5 \, 17325 \mu_2^4 \mu_3 \\
+ \quad & n^4 (15400 \mu_2 \mu_3^3 - 173250 \mu_2^4 \mu_3 + 34650 \mu_2^2 \mu_3 \mu_4 + 6930 \mu_2^3 \mu_5) \\
+ \quad & n^3 (606375 \mu_2^4 \mu_3 - 92400 \mu_2 \mu_3^3 - 207900 \mu_2^2 \mu_3 \mu_4 + 5775 \mu_3 \mu_4^2 - 41580 \mu_2^3 \mu_5 + 4620 \mu_3^2 \mu_5 \\
+ \quad & \quad 6930 \mu_2 \mu_4 \mu_5 + 4620 \mu_2 \mu_3 \mu_6 + 990 \mu_2^2 \mu_7) \\
+ \quad & n^2 (169400 \mu_2 \mu_3^3 - 866250 \mu_2^4 \mu_3 + 381150 \mu_2^2 \mu_3 \mu_4 - 17325 \mu_3 \mu_4^2 + 76230 \mu_2^3 \mu_5 \\
- \quad & \quad 13860 \mu_3^2 \mu_5 - 20790 \mu_2 \mu_4 \mu_5 - 13860 \mu_2 \mu_3 \mu_6 + 462 \mu_5 \mu_6 - 2970 \mu_2^2 \mu_7 + 330 \mu_4 \mu_7 \\
+ \quad & \quad 165 \mu_3 \mu_8 + 55 \mu_2 \mu_9) \\
+ \quad & n \, (415800 \mu_2^4 \mu_3 - 92400 \mu_2 \mu_3^3 - 207900 \mu_2^2 \mu_3 \mu_4 + 11550 \mu_3 \mu_4^2 - 41580 \mu_2^3 \mu_5 + 9240 \mu_3^2 \mu_5 \\
+ \quad & \quad 13860 \mu_2 \mu_4 \mu_5 + 9240 \mu_2 \mu_3 \mu_6 - 462 \mu_5 \mu_6 + 1980 \mu_2^2 \mu_7 - 330 \mu_4 \mu_7 - 165 \mu_3 \mu_8 \\
- \quad & \quad 55 \mu_2 \mu_9 + \mu_{11}).
\end{aligned}$$

For $m = 25$, the corresponding r-th, coefficient of $\binom{n}{r}$ with, e.g., $r = 11$, is given by

$$\begin{aligned} B_{11,25} &= 4628453517704016000000\mu_2^8\mu_3^3 + 23142267588520080000000\mu_2^9\mu_3\mu_4 \\ &+ 1388536055311204800000\mu_2^{10}\mu_5. \end{aligned}$$

References

[1] Batir N. Sharp inequalities for factorial n. Proyecciones 2008; 27: 97–102.

[2] Billingsley P. Convergence of Probability Measures. 2nd ed. New York, NY, USA: John Wiley & Sons, 1999.

[3] Constantine GM, Savits TH. A multivariate Faà di Bruno formula with applications. Trans Amer Math Soc 1996; 348: 503–520.

[4] Defant A, Junge M. Best constants and asymptotics of Marcinkiewicz–Zygmund inequalities. Stud Math 1997; 125: 271–287.

[5] Efron B. Student's t-test under symmetry condition. J Amer Statist Assoc 1969; 64: 1278–1302.

[6] Efron B, Tibshirani RJ. An Introduction to the Bootstrap. New York, NY, USA: Chapman & Hall, 1993.

[7] Egorov VA. Two-sided estimates for constants in the Marcinkiewicz inequalities. J Math Sci 2009; 159: 305–310 (translated from Zap Nauchn Semin POMI 2008; 361: 45–56).

[8] Egorov VA. On the growth rate of moments of random sums. PUB IRMA LILLE 1997; 44: 1–8.

[9] Hall P. The Bootstrap and Edgeworth Expansion. New York, NY, USA: Springer-Verlag, 1992.

[10] Johnson WP. The curious history of Faà di Bruno's formula. Amer Math Monthly 2002; 109: 217–234.

[11] Lukacs E. Applications of Faà di Bruno's formula in mathematical statistics. Amer Math Monthly 1995; 62: 340–348.

[12] Packwood DM. Moments of sums of independent and identically distributed random variables. http://arxiv.org/abs/1105.6283.

[13] Resnick SI. A Probability Path. Boston, MA, USA: Birkhäuser, 1999.

[14] Serfling RJ. Approximation Theorems of Mathematical Statistics. New York, NY, USA: John Wiley & Sons, 1980.

[15] Shorack GR. Probability for Statisticians. New York, NY, USA: Springer-Verlag, 2000.

[16] van der Vaart AW. Asymptotic Statistics. Cambridge, UK: Cambridge University Press, 1998.