The Annals of Probability
2003, Vol. 31, No. 2, 914–925

MEAN ABSOLUTE DEVIATIONS OF SAMPLE MEANS AND MINIMALLY CONCENTRATED BINOMIALS

BY LUTZ MATTNER

Universität zu Lübeck

Dedicated to Georg Neuhaus on his sixtieth birthday

This is a contribution to the theory of sums of independent random variables at the level of optimal explicit inequalities: we compute the optimal constants in Hornich's lower bounds for the mean absolute deviations of sample means. This is done by reducing the original problem to the elementary one of determining the minimally concentrated binomial distributions $B_{n,p}$ with fixed sample size parameter n.

1. Introduction and results.

A classical and important goal of probability theory is to describe the distributions of sums of independent random variables via a few simple functionals of their terms. At the levels of limit theorems, rate of convergence results and asymptotic expansions, this goal has been achieved remarkably well, with several results apparently having reached their final forms, often decades ago. See, for example, Petrov (1995). However, at the more difficult and more useful level of optimal explicit inequalities, almost no final results appear to be known.

The principal aim of this paper is to provide one simple such result, namely the sharp version of Hornich's inequality stated as Theorem 1.1 below. A possibly new auxiliary result on binomial distributions is given in Lemma 1.4(c).

Results stated in this section are discussed in the form of various remarks in Section 2 and proved in Section 3.

For $n \in \mathbb{N} \cup \{0\}$, $p \in [0, 1]$ and $k \in \{0, \dots, n\}$, we write

$$b(n, p; k) := \binom{n}{k} p^k (1-p)^{n-k}$$

for the density of the binomial distribution $B_{n,p}$. Let $\lfloor x \rfloor$ denote the integer part of the real number x and let us put, for $n \in \mathbb{N}$,

$$c_n := b\left(n, \frac{\lfloor n/2 \rfloor}{n}; \lfloor n/2 \rfloor\right) \tag{1}$$

$$= b\left(n-1, \frac{\lfloor n/2 \rfloor}{n}; \lfloor n/2 \rfloor\right). \tag{2}$$

Received July 2001; revised January 2002.

AMS 2000 subject classifications. 60E15, 62G05, 60G50.

Key words and phrases. Binomial distribution, concentration function, Hornich, moment inequality, sums of independent random variables.

THEOREM 1.1. *Let $n \in \mathbb{N}$ and let $X, X_1, \dots, X_n$ be independent and identically distributed random variables with existing expectation $\mathbb{E}X$. Then*

$$\mathbb{E}\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \mathbb{E}X\right| \geq c_n \mathbb{E}|X - \mathbb{E}X|. \tag{3}$$

Equality holds in (3) *iff $n = 1$ or X can be written as $aY + b$ with real numbers a and b and with a random variable Y satisfying $\mathbb{P}(Y = 1) = 1 - \mathbb{P}(Y = 0) = \lfloor n/2 \rfloor / n$.*

REMARK 1.2. The constants c_n from (1) are of the order $1/\sqrt{n}$. More precisely,

$$1 \geq c_n\sqrt{n} \geq \frac{1}{\sqrt{2}} = 0.7071\dots, \tag{4}$$

$$\lim_{n\to\infty} c_n\sqrt{n} = \sqrt{\frac{2}{\pi}} = 0.7978\dots. \tag{5}$$

The first few of the c_n are given by

$$c_1 = 1, \quad c_2 = \tfrac{1}{2}, \quad c_3 = \tfrac{4}{9}, \quad c_4 = \tfrac{3}{8}, \quad c_5 = \tfrac{216}{625},$$
$$c_6 = \tfrac{5}{16}, \quad c_7 = \tfrac{34560}{117649}, \quad c_8 = \tfrac{35}{128},$$

and the following table gives best lower decimal bounds with the stated accuracy.

n	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**99**	**100**	**∞**
c_n	1	0.500	0.444	0.375	0.345	0.312	0.293	0.273	0.260	0.246	0.0799	0.0795	0
$c_n\sqrt{n}$	1	0.707	0.769	0.750	0.772	0.765	0.777	0.773	0.780	0.778	0.7959	0.7958	0.7978

COROLLARY 1.3. *Under the assumptions of Theorem* 1.1, *we have*

$$\mathbb{E}\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \mathbb{E}X\right| \geq \frac{1}{\sqrt{2n}}\mathbb{E}|X - \mathbb{E}X|. \tag{6}$$

Equality holds in (6) *iff X is constant almost surely, or $n = 2$ and $X = aY + b$ with $\mathbb{P}(Y = 0) = \mathbb{P}(Y = 1) = 1/2$.*

The proof of inequality (3) given below combines a conditioning argument method, as described by Bshouty, Hengartner, Rohatgi and Székely [(1993), Appendix] and attributed to Burgess Davis, with the following properties of binomial distributions, of which the third one appears to be new.

LEMMA 1.4 (On binomial distributions). *Let $n \in \mathbb{N}$.*

(a) (The maximal terms). *For $p \in [0,1]$, we have*

$$k_0 \in \underset{k\in\{0,\dots,n\}}{\arg\max}\, b(n,p;k) \iff k_0 \in \{0,\dots,n\} \quad \text{and} \quad (n+1)p-1 \le k_0 \le (n+1)p. \tag{7}$$

(b) (De Moivre's mean absolute deviation identity). *For $p \in [0,1]$, we have*

$$\sum_{k=0}^{n} |k-np|\, b(n,p;k) = 2np(1-p) \max_{k\in\{0,\dots,n-1\}} b(n-1,p;k). \tag{8}$$

(c) (The minimally concentrated binomials with given sample size). *We have*

$$\min_{p\in[0,1]} \max_{k\in\{0,\dots,n\}} b(n,p;k) = b\left(n, \frac{\lfloor (n+1)/2 \rfloor}{n+1}; \lfloor (n+1)/2 \rfloor\right), \tag{9}$$

$$\underset{p\in[0,1]}{\arg\min} \max_{k\in\{0,\dots,n\}} b(n,p;k) = \left\{ \frac{\lfloor (n+1)/2 \rfloor}{n+1}, \frac{\lceil (n+1)/2 \rceil}{n+1} \right\}. \tag{10}$$

For the proof of the discussion of equality in Theorem 1.1, we also need the following simple fact.

LEMMA 1.5 (On equality in the conditional triangle inequality). *Let Z be an integrable random variable on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ and let $\mathcal{C} \subset \mathcal{A}$ be a sub-σ-algebra. Then we have the implications*

$$|\mathbb{E}[Z|\mathcal{C}]| = \mathbb{E}[|Z||\mathcal{C}] \text{ a.s.} \tag{11}$$

$$\iff \quad Z \ge 0 \text{ a.s. on } \{\mathbb{E}[Z|\mathcal{C}] \ge 0\} \quad \text{and} \quad Z \le 0 \text{ a.s. on } \{\mathbb{E}[Z|\mathcal{C}] \le 0\} \tag{12}$$

$$\implies \quad Z = 0 \text{ a.s. on } \{\mathbb{E}[Z|\mathcal{C}] = 0\}. \tag{13}$$

2. Remarks.

2.1. *Earlier versions of Theorem* 1.1. Theorem 1.1 with

$$\tilde{c}_n = \tfrac{1}{2} b\big(n-1, \tfrac{1}{2}; \lfloor n/2 \rfloor\big) \sim \tfrac{1}{2} c_n$$

in place of c_n is due to Hornich (1941). Birnbaum and Zuckerman (1944) give a more readable presentation. These three authors first consider the problem under the additional assumption that X is symmetrically distributed. Denoting the corresponding optimal constants replacing the c_n in (3) by $c_{n,\mathrm{sym}}$, they show that

$$c_{n,\mathrm{sym}} = b\big(n-1, \tfrac{1}{2}; \lfloor n/2 \rfloor\big). \tag{14}$$

Then they apply a simple symmetrization argument to obtain (3) with $\tilde{c}_n$ in place of c_n.

While Hornich's result (14) can obviously be obtained from a specialization of the proof of Theorem 1.1 given in Section 3, the discussion of equality, not given by Hornich or Birnbaum and Zuckerman, becomes a bit more complicated. For even n, we have equality iff $X = aY$ with $a \in \mathbb{R}$ and $\mathbb{P}(Y = -1) = \mathbb{P}(Y = 1) = 1/2$. This follows from Theorem 1.1, since $c_{n,\mathrm{sym}} = c_n$ for even n. For odd $n = 2k + 1 \geq 3$, however, we have equality in the analogue of (3) under symmetry iff $X = aY$ with Y symmetric and $\mathbb{P}(Y \in [1, 1 + 1/k]) = 1/2$. This can be shown as in the proof of Theorem 1.1, but using the equivalence (11) $\Leftrightarrow$ (12) rather than the implication (11) $\Rightarrow$ (13).

Tukey (1946) generalizes the symmetric case differently by showing: For $X_1, \dots, X_n$ independent but not necessarily identically distributed, we have for any choice of the medians $m(X_i)$

$$\left|\frac{1}{n}\sum_{i=1}^{n}\big(X_i - m(X_i)\big)\right| \geq c_{n,\mathrm{sym}}\frac{1}{n}\sum_{i=1}^{n}\mathbb{E}|X_i - m(X_i)|.$$

Bshouty, Hengartner, Rohatgi and Székely (1993) consider the analogue of (3) without centring. Part of their Theorem 1 can be stated as follows: For $X, X_1, \dots, X_n$ i.i.d., we have

$$\mathbb{E}\left|\frac{1}{n}\sum_{i=1}^{n}X_i\right| \geq c_n^*\mathbb{E}|X|,$$

with the optimal constant

$$c_n^* = \frac{1}{n}\min\left\{\frac{\sum_{k=0}^{n}|k - na|b(n, p; k)}{\sum_{k=0}^{1}|k - a|b(1, p; k)} : p \in]0, 1[\,,\ a \in \mathbb{R}\right\}.$$

They further show (page 8) that $c_3^* = \frac{1}{3}\cdot 1.3316$ and they state (in the abstract) that $c_2^* = \frac{1}{2}$, which can be verified. Obviously we must have $c_n^* \leq c_n$. Indeed $c_1^* = c_1$ and $c_2^* = c_2$, but $c_3^* < \frac{4}{9} = c_3$.

2.2. *The case* $n = 2$. For $n = 2$, inequality (3) has been proved in various ways in the literature, for example, by Cox and Kemperman [(1983), Theorem 2.6, the case $p = 1$]. They observe that

$$|x + y| \geq \tfrac{1}{2}(|x| + |y| + x\operatorname{sgn} y + y\operatorname{sgn} x) \qquad (x, y \in \mathbb{R}), \tag{15}$$

with equality iff either $y = -x$ or $xy > 0$. (Our "$xy > 0$" here corrects an error in the cited paper.) Replacing x, y by i.i.d. mean zero random variables and taking expectations on both sides of (15) easily yields the present Theorem 1.1 for $n = 2$, with discussion of equality.

An obvious generalization of (15) to an arbitrary number of reals $x_1, \dots, x_n$ is

$$\left|\sum_{i=1}^{n}x_i\right| \geq \frac{1}{n}\Big(\sum x_i\Big)\Big(\sum \operatorname{sgn} x_i\Big) = \frac{1}{n}\left(\sum_i |x_i| + \sum_{i\neq j} x_i \operatorname{sgn} x_j\right), \tag{16}$$

with equality throughout iff either $\sum x_i = 0$ or each x_i is positive or each x_i is negative. However, inequality (16) is not good enough to yield Theorem 1.1 for any $n \geq 3$.

2.3. *The trivial upper bound.* Under the assumptions of Theorem 1.1 we have from the triangle inequality the corresponding upper bound

$$\mathbb{E}\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \mathbb{E}X\right| \leq \mathbb{E}|X - \mathbb{E}X|, \tag{17}$$

without using either the independence assumption or the fact that the $X_i - \mathbb{E}X_i$ have zero means. Using independence, however, it is easy to see that equality in (17) occurs only in the trivial cases, namely if $n = 1$ or if X is constant almost surely. Nevertheless (17) is unimprovable, in the sense that the supremum of the left-hand side given the right-hand side is the right-hand side, even if we assume in addition that X is symmetrically distributed. To show this, we may consider $X_i = Y_i - Z_i$ with the Y_i, Z_i being i.i.d. Bernoulli with success probability p. Then we have $\mathbb{E}|X| = 2p(1-p)$ and it is not difficult to show that, for fixed $n \in \mathbb{N}$, $\mathbb{E}|\sum_{i=1}^{n} X_i| = 2np + O(p^2)$. [Ramasubban (1958) gives explicit formulae for $\mathbb{E}|\sum_{i=1}^{n} X_i|$.] Scaling the X_i to fix $\mathbb{E}|X|$ and letting p tend to zero yields the claim.

2.4. *A "right order upper bound" under minimal conditions.* If we suppose that $X, X_1, X_2, \ldots$ are i.i.d. with existing expectation but with infinite second moment, then we have

$$\lim_{n\to\infty} \sqrt{n}\,\mathbb{E}\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \mathbb{E}X\right| = \infty,$$

by Theorem 2 and Remark 1.3 of Esseen and Janson (1985). This shows that we cannot get an upper bound of order $1/\sqrt{n}$ for $\mathbb{E}|\frac{1}{n}\sum_{i=1}^{n} X_i - \mathbb{E}X|$ unless we assume the finiteness of $\mathbb{E}X^2$. Thus a natural analogue of (3) from Theorem 1.1, yielding an upper bound of the same order in n under the minimal condition of finiteness of the second moment, appears to be

$$\mathbb{E}\left|\frac{1}{n}\sum_{j=1}^{n} X_j - \mathbb{E}X\right| \leq C_n\big(\mathbb{E}(X - \mathbb{E}X)^2\big)^{1/2}, \tag{18}$$

with the optimal constants C_n still to be determined. By Lyapunov's inequality, we have $C_n \leq 1/\sqrt{n}$ in general. By considering X to be symmetric Bernoulli, we see that $C_1 = 1$. But already C_2 appears to be unknown. By considering X to be uniformly distributed on an interval, we can see that $C_2 \geq 1/\sqrt{3}$, whereas X Bernoulli yields at best, namely again in the symmetric case, the weaker inequality $C_2 \geq 1/2$.

The preceding choice of the uniform distribution is motivated by a consideration of the problem under the additional assumption that X is symmetrically distributed. Denoting the corresponding optimal constants replacing the C_n from (18) by $C_{n,\mathrm{sym}}$, we have $C_{n,\mathrm{sym}} \le C_n$, $C_{1,\mathrm{sym}} = C_1 = 1$ and $C_{2,\mathrm{sym}} = 1/\sqrt{3}$. Here the last equality follows from a theorem of Plackett (1947) which states in part that, for X_1, X_2 i.i.d., we have $\mathbb{E}|X_1 - X_2| \le (2/\sqrt{3})\,\mathrm{Var}(X_1)$, with nondegenerate equality occuring for uniform distributions on intervals only. Under the assumption of symmetry about zero, we have $\mathbb{E}|X_1 - X_2| = \mathbb{E}|X_1 + X_2|$, and hence it follows that $C_{2,\mathrm{sym}} = 1/\sqrt{3}$. We refer to Mattner (1992) and Buja, Logan, Reeds and Shepp (1994) for modern proofs and multivariate generalizations of the result of Plackett used above. The constants $C_{n,\mathrm{sym}}$ with $n \ge 3$ appear to be unknown.

2.5. *Other bounds for the mean absolute deviations of sample means.* Lower and upper bounds differing essentially from (3) and (18) have been given by Klass [(1980), Theorem 7, inequality (1.21)]. At the price of depending on the distribution of X in a somewhat more complicated way than via absolute moments, these bounds have the following two virtues: first, they are applicable without assuming anything beyond the finiteness of $\mathbb{E}|X|$; second, not only are the lower and upper bounds of the same order for any given distribution of X, but also the constants involved do not depend on X. See Hall (1981) for extensions of this result.

Asymptotically more accurate bounds apply when finiteness of higher moments is assumed. See von Bahr [(1965), Theorem 2, the case $\nu = 1$] for Berry–Esseen or Edgeworth type results.

2.6. *Equality in the conditional Jensen inequality.* For discussions of equality in the conditional Jensen inequality in situations more general than that of Lemma 1.5, we refer to Kozek and Suchanecki (1980) and to Mussmann (1988). For example, the present implication (11) $\Rightarrow$ (13) is a special case of Corollary 6.2(iii) in Kozek and Suchanecki (1980).

One might be tempted to think that (11) implies some measurability restrictions on Z, such as $\{Z \ge 0\}$ belonging to $\mathcal{C}$ up to almost sure equality. To see that this is not the case, it suffices to consider the example $\Omega = \{1, 2, 3, 4\}$ with the discrete σ-algebra, $\mathbb{P} =$ uniform distribution, $Z = (2, 0, -2, 0)$ and $\mathcal{C} = \{\varnothing, \{1, 2\}, \{3, 4\}, \Omega\}$. Then $\mathbb{E}[Z|\mathcal{C}] = (1, 1, -1, -1)$ and $|\mathbb{E}[Z|\mathcal{C}]| = (1, 1, 1, 1) = \mathbb{E}[|Z||\mathcal{C}]$, but no Z-measurable sets apart from $\varnothing$ and Ω are in $\mathcal{C}$ up to almost sure equality.

3. Proofs. The results stated in Section 1 are proved here in the order 1.4, 1.5, 1.1, 1.2 and 1.3.

PROOF OF LEMMA 1.4. Part (a) is a standard exercise. Part (b) is also well known; see Diaconis and Zabell (1991).

Proof of (c). We have to minimize the function f defined by

$$f(p) := \max_{k\in\{0,\dots,n\}} b(n,p;k) \qquad (p \in [0,1]).$$

By (7), we have for every $k \in \{0,\dots,n\}$

$$f(p) = b(n,p;k) \qquad \left(\frac{k}{n+1} \le p \le \frac{k+1}{n+1}\right). \tag{19}$$

As a function of $p \in]0,1[$, $\log b(n,p;k)$ is strictly concave, as follows from $\partial_p \log b(n,p;k) = k/p - (n-k)/(1-p)$ and hence $\partial_p^2 \log b(n,p;k) = -(k/p^2 + (n-k)/(1-p)^2) < 0$. Hence the minimum of f over any of the intervals $[k/(n+1),(k+1)/(n+1)]$ can be attained at boundary points only, so that (19) together with $b(n,0;0) = b(n,1;n) = 1 > f(1/2)$ yields

$$\operatorname*{arg\,min}_{p\in[0,1]} f(p) \subset \left\{\frac{1}{n+1},\dots,\frac{n}{n+1}\right\}.$$

Since $a_k := ((k+1)/k)^k$ is strictly increasing in k, we get for $k \in \{1,\dots,n-1\}$

$$\frac{f\left(\frac{k+1}{n+1}\right)}{f\left(\frac{k}{n+1}\right)} = \frac{b\left(n,\frac{k+1}{n+1};k\right)}{b\left(n,\frac{k}{n+1};k\right)} = \frac{\left(\frac{k+1}{k}\right)^k}{\left(\frac{n-k+1}{n-k}\right)^{n-k}} = \frac{a_k}{a_{n-k}} \begin{Bmatrix} < \\ = \\ > \end{Bmatrix} 1 \iff k \begin{Bmatrix} < \\ = \\ > \end{Bmatrix} n-k.$$

Now (10) and (9) easily follow. $\square$

PROOF OF LEMMA 1.5. Assume (11). Then $\mathbb{E}[Z|\mathcal{C}] = \mathbb{E}[|Z||\mathcal{C}]$ a.s. on $C_+ := \{\mathbb{E}[Z|\mathcal{C}] \ge 0\} \in \mathcal{C}$. Hence

$$\int_{C_+} Z\,d\mathbb{P} = \int_{C_+} \mathbb{E}[Z|\mathcal{C}]\,d\mathbb{P} = \int_{C_+} \mathbb{E}[|Z||\mathcal{C}]\,d\mathbb{P} = \int_{C_+} |Z|\,d\mathbb{P}. \tag{20}$$

Using $Z \le |Z|$, this implies $Z = |Z| \ge 0$ a.s. on C_+. Similarly we get $Z = -|Z| \le 0$ a.s. on $C_- := \{\mathbb{E}[Z|\mathcal{C}] \le 0\}$, yielding (12).

Assume (12). We then get (20) in the permuted order 2143, yielding $\mathbb{E}[Z|\mathcal{C}] = \mathbb{E}[|Z||\mathcal{C}]$ a.s. on C_+. Similarly we get $\mathbb{E}[Z|\mathcal{C}] = -\mathbb{E}[|Z||\mathcal{C}]$ a.s. on C_-, yielding (11).

The implication (12) $\Rightarrow$ (13) is trivial. $\square$

PROOF OF THEOREM 1.1. The case $n = 1$ being trivial, we may and do assume that $n \ge 2$. For convenience and without loss of generality, we also assume that $\mathbb{E}X = 0$ and $\mathbb{E}|X| > 0$, so that in particular $\mathbb{P}(X > 0) > 0$ and $\mathbb{P}(X < 0) > 0$.

For each $i \in \{0,\dots,n\}$, let $\mathcal{C}_i$ denote the σ-algebra generated by the event $\{X_i > 0\}$, and let Y_i be a conditional expectation given $\mathcal{C}_i$:

$$\begin{aligned} \mathcal{C}_i &:= \sigma(\{X_i > 0\}), \\ Y_i &\in \mathbb{E}[X_i|\mathcal{C}_i]. \end{aligned} \tag{21}$$

Of course, Y_i is unique and given by

$$Y_i = \frac{(X_i > 0)}{\mathbb{P}(X_i > 0)} \int_{\{X_i > 0\}} X_i \, d\mathbb{P} + \frac{(X_i \le 0)}{\mathbb{P}(X_i \le 0)} \int_{\{X_i \le 0\}} X_i \, d\mathbb{P}, \tag{22}$$

with (statement) $:= 1$ or 0 according to whether "statement" is true or false. We deduce that $|Y_i| \in \mathbb{E}[|X_i| \mid \mathcal{C}_i]$, by observing that the modulus of the right-hand side of (22) can be obtained by replacing the integrand X_i by $|X_i|$ in both cases. [Alternatively, we could have applied the implication (12) $\Rightarrow$ (11) to $Z = X_i$.] Hence

$$\mathbb{E}|Y_i| = \mathbb{E}|X_i|. \tag{23}$$

Let now $\mathcal{C}$ denote the σ-algebra generated by all the $\mathcal{C}_i$:

$$\mathcal{C} := \sigma(\mathcal{C}_1, \dots, \mathcal{C}_n).$$

Writing $S_n := \sum_{i=1}^n X_i$, we have

$$\begin{aligned}
\mathbb{E}|S_n| &= \mathbb{E}\big[\mathbb{E}[|S_n| \mid \mathcal{C}]\big] && \\
&\ge \mathbb{E}|\mathbb{E}[S_n \mid \mathcal{C}]| && \text{[by the conditional triangle inequality]} \qquad (24)\\
&= \mathbb{E}\left|\sum_{i=1}^n \mathbb{E}[X_i \mid \mathcal{C}]\right| && \\
&= \mathbb{E}\left|\sum_{i=1}^n \mathbb{E}[X_i \mid \mathcal{C}_i]\right| && \text{[by independence of the } X_i\text{]} \\
&= \mathbb{E}\left|\sum_{i=1}^n Y_i\right| && \text{[by (21)]} \\
&\ge n c_n \mathbb{E}|Y_1| && \text{[see below]} \qquad (25)\\
&= n c_n \mathbb{E}|X| && \text{[by (23)],}
\end{aligned}$$

yielding (3) up to the justification of (25).

To prove (25), we observe that the Y_i are i.i.d. two-valued mean zero random variables, say with values $-\alpha < 0$ and $\beta > 0$. Thus the random variables $Z_i := (Y_i + \alpha)/(\alpha + \beta)$ are i.i.d. Bernoulli with success probability $p := \alpha/(\alpha + \beta)$. By homogeneity, (25) is equivalent to

$$\mathbb{E}\left|\sum_{i=1}^n Z_i - np\right| \ge n c_n \mathbb{E}|Z_1 - p|. \tag{26}$$

Now

$$\mathbb{E}\left|\sum_{i=1}^{n} Z_i - np\right| = 2np(1-p) \max_{k\in\{0,\dots,n-1\}} b(n-1,p;k) \qquad \text{[by (8)]}$$

$$\geq 2np(1-p) \min_{p\in[0,1]} \max_{k\in\{0,\dots,n-1\}} b(n-1,p;k) \tag{27}$$

$$= 2np(1-p) b\left(n-1, \frac{\lfloor n/2\rfloor}{n}; \lfloor n/2\rfloor\right) \qquad \text{[by (9)]}$$

$$= n\mathbb{E}|Z_1 - p| c_n \qquad \text{[by (2) and by (8) with } n=1\text{]}.$$

This completes the proof of (3). It is easily checked that equality holds under the stated condition.

Conversely, let us now assume that equality holds in (3). Then we have equality in (24) and (25). Equality in (25) yields equality in (27). Hence, by (10) with $n-1 \in \mathbb{N}$ in place of n, we have

$$p \in \left\{\lfloor n/2\rfloor/n, \lceil n/2\rceil/n\right\}. \tag{28}$$

In particular, using $n \geq 2$, there is a $k \in \{1,\dots,n-1\}$ with $p = k/n$ and we have

$$\left\{\sum_{i=1}^{n} Y_i = 0\right\} = \left\{\sum_{i=1}^{n} Z_i = np\right\} = \left\{\sum_{i=1}^{n} (X_i > 0) = k\right\}. \tag{29}$$

Now equality in (24) yields

$$\left|\mathbb{E}\left[\sum_{i=1}^{n} X_i \middle| \mathcal{C}\right]\right| = \mathbb{E}\left[\left|\sum_{i=1}^{n} X_i\right| \middle| \mathcal{C}\right] \qquad \text{a.s.}$$

which is (11) with $Z := \sum_{i=1}^{n} X_i$ and $\mathbb{E}[Z|\mathcal{C}] = \sum_{i=1}^{n} Y_i$. Using (13) and (29), we deduce that

$$\sum_{i=1}^{n} X_i = 0 \text{ holds almost surely on } \left\{\sum_{i=1}^{n} (X_i > 0) = k\right\}. \tag{30}$$

Hence each X_i is at most two-valued almost surely. To prove this claim, let us assume the contrary. Then there are three pairwise disjoint Borel sets A, B, C with $\mathbb{P}(X \in A)\mathbb{P}(X \in B)\mathbb{P}(X \in C) > 0$. Since $\mathbb{E}X = 0$, we can choose A, B, C such that additionally either $A, B \subset]-\infty, 0]$ and $C \subset]0, \infty[$, or $A, B \subset]0, \infty[$ and $C \subset]-\infty, 0]$. In either case, using that $k \in \{1,\dots,n-1\}$, we may choose $D_1,\dots,D_{n-1} \in \{A, B, C\}$ with

$$(A \subset]0,\infty[) + \sum_{i=1}^{n-1} (D_i \subset]0,\infty[) = k, \tag{31}$$

and then (31) is also true with B in place of A. Let us denote the underlying probability measure by $\mathbb{P}$ and let us write $Y \,\square\, \mathbb{P}$ for the corresponding distribution of any random variable Y. For $(X_1,\ldots,X_n)\square\mathbb{P}$ almost every $(x_1,\ldots,x_n)\in D_1\times\cdots\times D_{n-1}\times A$ we have, by (31) and (30), $\sum_{i=1}^n x_i=0$. Hence, using independence, Fubini and the fact that $X_n\square\mathbb{P}(A)>0$, we deduce the following: For $(X_1,\ldots,X_{n-1})\square\mathbb{P}$ almost every $(x_1,\ldots,x_{n-1})\in D_1\times\cdots\times D_{n-1}$ there is an $x_n\in A$ with $\sum_{i=1}^n x_i=0$. Since this statement is also true with B in place of A, we get, by combining two null sets: For $(X_1,\ldots,X_{n-1})\square\mathbb{P}$ almost every $(x_1,\ldots,x_{n-1})\in D_1\times\cdots\times D_{n-1}$ there is a $y\in A$ and a $z\in B$ with $y+\sum_{i=1}^{n-1}x_i=0=z+\sum_{i=1}^{n-1}x_i$. Since $(X_1,\ldots,X_{n-1})\square\mathbb{P}(D_1\times\cdots\times D_{n-1})>0$ and since A and B are disjoint, we have arrived at the desired contradiction.

Hence we must in fact have $X_i=Y_i$ almost surely. Together with (28) the final claim of Theorem 1.1 follows. (Take $a<0$ if $p=\lceil n/2\rceil/n$.) $\square$

PROOF OF REMARK 1.2. The stated values of the c_n are easily verified using (2). Relation (5) follows easily from the local central limit theorem for binomial distributions $B_{n,p}$, if in the formulation of the latter the success probability p is allowed to vary in a compact subset of $]0,1[$.

To prove (4), we show that each of the two sequences $(c_{2k}\sqrt{2k}:k\in\mathbb{N})$ and $(c_{2k+1}\sqrt{2k+1}:k\in\mathbb{N})$ is strictly increasing. Relation (4) then follows using (5) and the values for $c_n\sqrt{n}$ with $n\in\{1,2,3\}$.

For $k\in\mathbb{N}$ we have

$$c_{2k}=b\left(2k,\frac{1}{2};k\right)=\binom{2k}{k}2^{-2k}$$

and hence

$$\frac{c_{2k+2}}{c_{2k}}=\frac{2k+1}{2k+2},\qquad \frac{c_{2k+2}\sqrt{2k+2}}{c_{2k}\sqrt{2k}}=\frac{2k+1}{\sqrt{2k(2k+2)}}>1.$$

Hence the sequence $(c_{2k}\sqrt{2k}:k\in\mathbb{N})$ is strictly increasing.

Again for $k\in\mathbb{N}$ we have, using (2),

$$c_{2k+1}=b\left(2k,\frac{k}{2k+1};k\right)=\binom{2k}{k}\frac{k^k(k+1)^k}{(2k+1)^{2k}}$$

and hence

$$\frac{c_{2k+3}}{c_{2k+1}}=\frac{\left(k+\frac{1}{2}\right)^{2k+1}(k+2)^{k+1}}{\left(k+\frac{3}{2}\right)^{2k+2}k^k}$$

so that

$$\frac{c_{2k+3}\sqrt{2k+3}}{c_{2k+1}\sqrt{2k+1}}=\frac{\left(1+\frac{1}{2k}\right)^{2k+1/2}\left(1+\frac{2}{k}\right)^{k+1}}{\left(1+\frac{3}{2k}\right)^{2k+3/2}}.\tag{32}$$

The logarithm of the right-hand side of (32) is

$$l_k:=\left(2k+\frac{1}{2}\right)\log\left(1+\frac{1}{2k}\right)+(k+1)\log\left(1+\frac{2}{k}\right)-\left(2k+\frac{3}{2}\right)\log\left(1+\frac{3}{2k}\right).$$

Now

$$l_k = kf\left(\frac{1}{k}\right) \qquad (k \in \mathbb{N})$$

with

$$f(x) := \left(2 + \frac{x}{2}\right)\log\left(1 + \frac{x}{2}\right) + (1+x)\log(1+2x)$$
$$- \left(2 + \frac{3}{2}x\right)\log\left(1 + \frac{3}{2}x\right) \qquad (x \in [0,1]).$$

One can easily check the following three facts: (i) $f(0) = f'(0) = 0$, (ii) the second derivative

$$f''(x) = 3x\frac{4 - 19x^2 - 12x^3}{(2+x)^2(1+2x)^2(2+3x)^2}$$

changes sign at exactly one $x_0 \in]0,1]$, from plus to minus, (iii) $f(1) > 0$. By (i) and by $f'' > 0$ on $]0, x_0[$, we have $f > 0$ on $]0, x_0]$. By $f(x_0) > 0$, by (iii), and by $f'' < 0$ on $]x_0, 1]$, we have $f > 0$ on $[x_0, 1]$ as well. Hence we have $l_k > 0$ for $k \in \mathbb{N}$, so that the left-hand side of (32) is > 1. Hence the sequence $(c_{2k+1}\sqrt{2k+1} : k \in \mathbb{N})$ is strictly increasing. $\square$

PROOF OF COROLLARY 1.3. Inequality (6) is obvious from (3) and (4). The discussion of equality follows from the corresponding one in Theorem 1.1, since the above proof of Remark 1.2 shows that equality in $c_n\sqrt{n} \geq 1/\sqrt{2}$ occurs iff $n = 2$. $\square$

Acknowledgments. I thank Leonid Bogachev, John Kent, Sergey Utev and a referee for helpful discussions or remarks.

REFERENCES

BIRNBAUM, Z. W. and ZUCKERMAN, H.S. (1944) An inequality due to H. Hornich. *Ann. Math. Statist.* **15** 328–329.

BSHOUTY, D., HENGARTNER, W., ROHATGI, V. K. and SZÉKELY, G. J. (1993). On the extrema of the expected values of functions of independent identically distributed random variables. *Ulam Quart.* **2** no. 1 (electronic only).

BUJA, A., LOGAN, B. F., REEDS, J. R. and SHEPP, L. A. (1994). Inequalities and positive-definite functions arising from a problem in multidimensional scaling. *Ann. Statist.* **22** 406–438.

COX, D. C. and KEMPERMAN, J. H. B. (1983). Sharp bounds on the absolute moments of a sum of two i.i.d. random variables. *Ann. Probab.* **11** 765–771.

DIACONIS, P. and ZABELL, S. (1991). Closed form summation for classical distributions: Variations on a theme of De Moivre. *Statist. Sci.* **6** 284–302.

ESSEEN, C.-G. and JANSON, S. (1985). On moment conditions for normed sums of independent variables and martingale differences. *Stochastic Process. Appl.* **19** 173–182.

HALL, P. (1981). Order of magnitude of moments of sums of random variables. *J. London Math. Soc.* **24** 562–568.

HORNICH, H. (1941). Zur Theorie des Risikos. *Monatsh. Math.* **50** 142–150.

KLASS, M. J. (1980). Precision bounds for the relative error in the approximation of $E|S_n|$ and extensions. *Ann. Probab.* **8** 350–367.

KOZEK, A. and SUCHANECKI, Z. (1980). Multifunctions of faces for conditional expectations of selectors and Jensen's inequality. *J. Multivariate Anal.* **10** 579–598.

MATTNER, L. (1992). Extremal problems for probability distributions: A general method and some examples. In *Stochastic Inequalities* (M. Shaked and Y. L. Tong, eds.) 274–283. IMS, Hayward, CA.

MUSSMANN, D. (1988). Sufficiency and Jensen's inequality for conditional expectations. *Ann. Inst. Statist. Math.* **40** 715–726.

PETROV, V. V. (1995). *Limit Theorems of Probability Theory. Sequences of Independent Random Variables.* Clarendon, Oxford.

PLACKETT, R. L. (1947). Limits of the ratio of mean range to standard deviation. *Biometrika* **34** 120–122.

RAMASUBBAN, T. A. (1958). The mean difference and mean deviation of some discontinuous distributions. *Biometrika* **45** 549–556.

TUKEY, J. W. (1946). An inequality for deviations from medians. *Ann. Math. Statist.* **17** 75–78.

VON BAHR, B. (1965). On the convergence of moments in the central limit theorem. *Ann. Math. Statist.* **36** 808–818.

INSTITUT FÜR MATHEMATIK
UNIVERSITÄT ZU LÜBECK
WALLSTRASSE 40
23560 LÜBECK
GERMANY
E-MAIL: MattnerL@member.ams.org
WEB PAGE: www.math.uni-luebeck.de/mattner