On the tails of linear combinations of Rademacher random variables through exponential tilting

Fredrik Jonsson
On the tails of linear combinations of Rademacher random variables through exponential tilting

Fredrik Jonsson, Uppsala University

Abstract
Consider linear combinations \( S = \sum_{i=1}^{n} a_i \varepsilon_i \) of independent Rademacher random variables \( \varepsilon_1, \ldots, \varepsilon_n \) for real numbers \( a_1, \ldots, a_n \) such that \( \sum_{i=1}^{n} a_i^2 = 1 \). Let \( N \) refer to the normal tail function, \( N(x) = 1 - \Phi(x) \). We prove that, for some constant \( C_1 \), all \( x > 0 \), and all \( S \),
\[
P(S \geq x) \leq N(x)(1 + C_1 x^{-2}).
\]
Moreover, we verify the existence of divergent sequences \( \{x_n\}_{n \geq 1} \) such that, regarding the subset of standardized binomial random variables \( B_n = \sum_{i=1}^{n} \varepsilon_i / \sqrt{n} \),
\[
\lim_{n \to \infty} x_n^2 \left( P(B_n \geq x_n) / N(x_n) - 1 \right) = 3.
\]
Finally, we verify that, for some constant \( C_2 \),
\[
P(S_n \geq x) \leq N(x), \text{ for } x \geq C_2,
\]
regarding all self-normalized sums \( S_n = \sum_{i=1}^{n} X_i / \left( \sum_{i=1}^{n} X_i^2 \right)^{1/2} \) of i.i.d. \( N(0, 1) \) random variables \( X_1, \ldots, X_n \).

1 Introduction
Let \( \{\varepsilon_i\}_{i \geq 1} \) refer to an i.i.d. sequence of Rademacher random variables, that is, with \( P(\varepsilon_1 = 1) = 1/2 \) and \( P(\varepsilon_1 = -1) = 1/2 \). Moreover, for integers \( n \), consider the class of vectors \( (a_1, \ldots, a_n) \in \mathbb{R}^n_+ \) satisfying \( \sum_{i=1}^{n} a_i^2 = 1 \). Finally, for given \( n \) and \( a \), consider the linear combination \( S = a_1 \varepsilon_1 + \cdots + a_n \varepsilon_n \). Efron [4] refers to \( S \) as a generalized binomial random variable. Indeed, with \( a_1 = a_2 = \ldots = a_n = 1/\sqrt{n} \), \( S \) refers to a binomial random variable with parameters \( n \) and \( 1/2 \), standardized to zero mean and unit variance. For convenience, we also define \( B_n = (\varepsilon_1 + \cdots + \varepsilon_n) / \sqrt{n} \).

Starting with [4], Efron noted that
\[
P(S \geq x) \leq e^{-x^2/2}.
\]
The proof given in [4] relies upon bounds on the moments of \( S \). Confer also [8, Section 2.2] for an alternative proof of this result. With reference to the normal tails \( N(x) = 1 - \Phi(x) \), one may note (cf. for instance [5, Chapter VII, Lemma 2]) that,
\[
\lim_{x \to \infty} \sqrt{2\pi} xe^{x^2/2} N(x) = 1.
\]

AMS 2000 subject classifications. Primary 60E15; Secondary 60F10, 60G50, 60C05.
Keywords and phrases. Rademacher random variables, sums of independent random variables, self-normalized sums, tail probabilities, exponential tilting, large and moderate deviations.
Date. August 30, 2011
Therefor, bounds of the following kind, given some constant $C_1$,

$$P(S \geq x) \leq C_1 N(x),$$

are more attractive than (1), as $x$ becomes large. Recently, Pinelis [10] proved that (3) holds with $C_1 \approx 3.22$, and that (3) does not hold for $C_1 < 3.18$. The optimal

global constant for the given kind of inequality is thus almost known. The proof in [10] is a refinement of the method used in [2], where a short proof for $C_1 \approx 12$ is given. Both proofs proceed inductively regarding the number of summands $n$.

The main result of the present study (Theorem 3.1) is that, for some constant $C_2$,

$$P(S \geq x) \leq N(x)(1 + C_2 x^{-2}).$$

Clearly, (4) has a more attractive form than (3), as $x$ becomes large. We also verify (Corollary 4.2) that the above form is optimal in the given setting, in the sense that there exists a sequence $\{x_n\}_{n \geq 1}$ such that

$$\lim_{n \to \infty} x_n^2 \left( P(B_n \geq x_n) / N(x_n) - 1 \right) = 3.$$

As a remark, the order of deviation of the right hand side in (4) from $N(x)$ is similar to the order of deviation of the approximation $\sqrt{2\pi x e^{x^2/2}}$ to $N(x)$. Indeed, (cf. [5, Chapter VII]),

$$\lim_{x \to \infty} x^2 \left( \sqrt{2\pi x e^{x^2/2}} - N(x) \right) = -1.$$  

Theorem 3.1 is proved through the technique known as exponential change of measure, or exponential tilting, which is the basic tool in the theory of large deviations (confer Bucklew’s comment in [3, page 13]). Thus, the starting point of the analysis follows the tradition in classical methodology (confer for instance Feller [6, Section XVI.7] or Petrov [9, Chapter VIII]) for proving results in the theory of large deviations for sums of independent random variables.

Theorem 4.1, from which the previously mentioned Corollary 4.2 follows, is concerned with more detailed results regarding binomial distributions. The analysis rests upon the lattice structure associated with binomial distributions. Thus, the main ingredient in the proof of Theorem 4.1 is the classical Stirling approximation applied to binomial coefficients (confer e.g. Feller’s account in [5, Chapters II and VII]).

It is a well–known fact that there exists a connection between random variables $S$ of the above form and $t$–statistic random variables of symmetric, independent summands $X_1, \ldots, X_n$. Indeed, there is, on the one hand, a close connection between self–normalized sums $S_n = \sum_{i=1}^n X_i / (\sum_{i=1}^n X_i^2)^{1/2}$ and $t$–statistic random variables (confer e.g. [8, Chapter 1]). On the other hand (cf. [8, Section 2.2]), any such self–normalized sum can be expressed as a mixture distribution, with respect to sums of the form $S$.

Thus, global bounds on the tails of random variables $S$ are also applicable as bounds on the tails of self–normalized sums of symmetric summands. However, (4) may not be equally relevant as a global bound on the tails of self–normalized sums of absolutely continuous random variables. Indeed, we end the present study (Section 5) by verifying, regarding self–normalized sums of i.i.d. $N(0,1)$ random variables, that there exists a constant $C_3$ such that $P(S_n \geq x) \leq N(x)$, for $x \geq C_3$.  

2
2 Large and moderate deviations

We begin by a brief review of large and moderate deviation results regarding the symmetric binomial random variables \( \{B_n\}_{n\geq 1} \).

The previously mentioned perspective of exponential tilting corresponds to the following representation of tail probabilities (cf. Lemma 3.1 below)

\[
P(B_n \geq x) = \exp \left( \frac{x^2}{2} + g_n(h) - hx \right) \left( \sqrt{2\pi x} \int_0^\infty e^{-hy} \tilde{G}_h(dy) \right) \left( \sqrt{2\pi x} e^{x^2/2} \right)^{-1},
\]

with \( g_n \) referring to the cumulant generating function of \( B_n \),

\[
g_n(h) = n \log(\cosh(\frac{h}{\sqrt{n}})),
\]

and \( h \) chosen such that \( x = g_n'(h) \). Moreover, \( \tilde{G}_h \) is referred to as the exponentially tilted measure. Thus, introducing \( t = h/\sqrt{n} \), it follows that

\[
g_n(h) - hx = n \left( \log(\cosh(t)) - t \sinh(t)/\cosh(t) \right).
\]

In the classical large deviation perspective (cf. e.g. [3, Chapter 2]), one considers \( \lambda = x/\sqrt{n} \) fixed, \( 0 < \lambda < 1 \), with \( x \) tending to infinity. For the considered case, with exponential tilting, this corresponds to \( \lambda = \sinh(t)/\cosh(t) \), or in other words, \( t = \log((1 + \lambda)/(1 - \lambda))/2 \). Thus, we may rewrite (8) as

\[
g_n(h) - hx = -(x^2/2)\lambda^{-2} \left( (1 + \lambda) \log(1 + \lambda) + (1 - \lambda) \log(1 - \lambda) \right),
\]

which, by series representation of the logarithms, transfers into

\[
g_n(h) - hx = \frac{-x^2}{2} \left( 1 + \frac{\lambda^2}{6} + \sum_{k=2}^{\infty} 2 \left( \frac{1}{2k+1} - \frac{1}{2k+2} \right) \lambda^{2k} \right).
\]

Here, one may in particular note the subgaussian behaviour, with positive coefficients in the series representation of \( (1 + \lambda) \log(1 + \lambda) + (1 - \lambda) \log(1 - \lambda) \). Thus, for \( \lambda = x/\sqrt{n} \) fixed, it follows that \( P(B_n \geq x)/N(x) \to 0 \) as \( x \) tends to infinity, due to the behaviour of the exponent \( x^2/2 + g_n(h) - hx \).

For the perspective of moderate deviations (cf. e.g. [8, Section 2.1.3]) or [9, Chapter VIII]), i.e. the perspective of \( x \) tending to infinity with \( x = o(\sqrt{n}) \), it follows that (cf. Proposition 3.2 below or [6, Section XVI.7]),

\[
\sqrt{2\pi x} \int_0^\infty e^{-hy} \tilde{G}_h(dy) - 1 = O\left( \frac{x}{\sqrt{n}} \right) + O(x^{-2}).
\]

Going back to (7), it may thus be seen that

\[
\lim_{n \to \infty} P(B_n \geq x_n)/N(x_n) = 1,
\]

occurs if and only if \( x^2/2 + g_n(h) - hx \to 0 \). Moreover, by (9), this corresponds to \( x\lambda \to 0 \), which in turn corresponds to \( x^4 = o(n) \). For an alternative proof of this result, confer [1, Corollary 3].
Now, regarding the remainder term in (11) and \( x^4 = o(n) \), it follows from (9), (10) and the normal tail approximation (6), that it can be represented as

\[
- \frac{x^4}{12n} + o\left(\frac{x^4}{n}\right) + O\left(\frac{x}{\sqrt{n}}\right) + O(x^{-2}).
\]

(12)

Notably, \( x^4/n \) is proportional to \( x/\sqrt{n} \) whenever \( x^6 \) is proportional to \( n \), in which case \( x/\sqrt{n} \) is also proportional to \( x^{-2} \). Moreover, \( x^4/n \) dominates \( x/\sqrt{n} \), unless \( x \) is sufficiently small. Thus, it can be expected that an inequality of the form

\[
P(B_n \geq x) \leq N(x)(1 + C_1 x^{-2}),
\]

(13)

should hold, for all \( n \) and \( x \) and some constant \( C_1 \). Moreover, asymptotic sharpness of the given kind of result should take place in the perspective of having \( cn \leq x^6 \leq Cn \), given some constants \( c \) and \( C \).

In the following, Section 3 is concerned with proving (13) for the generalized setting, whereas Section 4 is concerned with the behaviour of \( C_1 \) in (13) for \( cn \leq x^6 \leq Cn \). Numerical results regarding lower bounds on the constant \( C_1 \) in (13) is demonstrated in Section 5.

3 Inequalities through exponential tilting

The main result of the present section is Theorem 3.1. We recall (cf. Section 1) that \( N \) refers to the normal tail function, and \( S \) to a linear combination of Rademacher random variables.

**Theorem 3.1.** There exists a constant \( C_1 \) such that, for any given \( x \) and \( S \),

\[
P(S \geq x) \leq N(x)(1 + C_1 x^{-2}).
\]

The proof of Theorem 3.1 is based on Lemmas 3.1–3.4 and Propositions 3.1–3.2. We begin by introducing the considered setting of cumulant generating functions and exponential tilting. Thus, for the remaining part of the section, let a linear combination \( S \) of independent Rademacher random variables be given. Moreover, let \( h \mapsto g(h), 0 < h < \infty \), refer to its cumulant generating function, \( g(h) = \log \mathbb{E}(e^{hS}) \).

Thus, by differentiation,

\[
g(h) = \sum_{i=1}^{n} \log (\cosh(a_i h)), \quad g'(h) = \sum_{i=1}^{n} \frac{a_i \sinh(a_i h)}{\cosh(a_i h)}, \quad (14)
\]

\[
g''(h) = \sum_{i=1}^{n} \left( \frac{a_i}{\cosh(a_i h)} \right)^2, \quad 1 - g''(h) = \sum_{i=1}^{n} \left( \frac{a_i \sinh(a_i h)}{\cosh(a_i h)} \right)^2. \quad (15)
\]

Then, given \( h > 0 \), consider exponentially tilted and mean centered random variables \( \tilde{X}_1(h), \ldots, \tilde{X}_n(h) \),

\[
\tilde{X}_i(h) = \begin{cases} 
  a_i \frac{e^{-a_i h}}{e^{2a_i h} + e^{-a_i h}}, & \text{with probab. } \frac{e^{a_i h}}{e^{2a_i h} + e^{-a_i h}}, \\
  -a_i \frac{e^{a_i h}}{e^{2a_i h} + e^{-a_i h}}, & \text{with probab. } \frac{e^{-a_i h}}{e^{2a_i h} + e^{-a_i h}}, 
\end{cases}
\]

with sum \( \tilde{S}(h) = \sum_{i=1}^{n} \tilde{X}_i(h) \), whose distribution we refer to as \( \tilde{G}_h \).
The connection between cumulant generating function, exponential tilting and tail probabilities of $S$ is presented in Lemma 3.1. Note that $g'(h)$ increases continuously from 0 to $\sum_{i=1}^n a_i$, as $h$ increases from 0 to $\infty$. The implicit claim in Lemma 3.1 regarding a unique solution to $g'(h) = x$ is thus valid for $0 < x < \sum_{i=1}^n a_i$.

**Lemma 3.1.** For any given $S$ and $x > 0$ such that $\sum_{i=1}^n a_i > x$, define $h = h(x)$ as the unique solution to $g'(h) = x$. Then,

$$P(S \geq x) = \exp \left( x^2/2 + g(h) - hx \right) \left( \sqrt{2\pi x} \int_0^\infty e^{-hy} \tilde{G}_h(dy) \right)^{-1} \left( \sqrt{2\pi x}e^{x^2/2} \right)^{-1}.$$ 

As noted in the introductory section, the third factor $\left( \sqrt{2\pi x}e^{x^2/2} \right)^{-1}$ in the above representation of $P(S \geq x)$ can be associated with the corresponding normal tail $N(x)$. Indeed, Lemma 3.2 presents explicit bounds for this association. For the proof of Lemma 3.2, cf. [5, Chapter VII, Lemma 2].

**Lemma 3.2.** For any $x > 0$,

$$(1 - x^{-2}) \leq N(x) \sqrt{2\pi x} e^{x^2/2} \leq 1.$$ 

For the proof of Theorem 3.1, it thus suffices to analyze the two remaining factors in the above representation of $P(S \geq x)$. Proposition 3.1 presents inequalities for $\exp \left( x^2/2 + g(h) - hx \right)$, whereas Proposition 3.2 is concerned with inequalities for $\sqrt{2\pi x} \int_0^\infty e^{-hy} \tilde{G}_h(dy)$. All of these bounds are expressed in terms of $g''(h), h$ and $x$.

There is a close connection between Propositions 3.1–3.2 and the perspective of moderate deviations in Section 2. Indeed, the fundamental quantity $\lambda = x/\sqrt{n}$ in Section 2 can be reexpressed as $\lambda = \sqrt{1 - g''(h)}$. Thus, the main content of (9) for the perspective of moderate deviations can be expressed, by truncation of the series representation, as

$$x^2/2 + g(h) - hx \leq -(1 - g''(h))x^2/12.$$ 

We have not been able to generalize this inequality to general sums (a success would have simplified the remaining part of the proof of Theorem 3.1). However, a weaker form of the inequality is given in Proposition 3.1(i), which is based on Lemma 3.3.

The second corner stone of the moderate deviation analysis in Section 2, statement (10) regarding the exponentially tilted measure, also generalizes through the connection $\lambda = \sqrt{1 - g''(h)}$. Indeed, for $\lambda \to 0$, Proposition 3.4 implies that

$$\sqrt{2\pi x} \int_0^\infty e^{-hy} \tilde{G}_h(dy) - 1 = O(\lambda) + O(x^{-2}),$$

which is the analogue of (10). The basic tool for the proof of Proposition 3.4 is the Berry–Esseen inequality applied to the tilted sum $\tilde{S}(h)$. Its application leads to bounds in terms of sums of absolute moments of third order. Relevant bounds on these sums are in turn given by Lemma 3.4.

**Lemma 3.3.** For any $0 < h < \infty$,

$$g(h) - h^2/2 \leq -(1 - g''(h))h^2/12.$$
Proposition 3.1. For any $0 < h < \infty$, with $x = g'(h)$,

(i) $x^2/2 + g(h) - hx \leq -(1 - g''(h))g''(h)h^2/12$;

(ii) $x^2/2 + g(h) - hx \leq -x^2(\log 2 - 1/2) + 2hx \sqrt{g''(h)}$.

Lemma 3.4. For any $0 < h < \infty$,

$$h \sum_{i=1}^{n} \mathbb{E}[	ilde{X}_i(h)]^3 \leq 2 \sqrt{(1 - g''(h))g''(h)}.$$  

Proposition 3.2. There exists a constant $C_2$ such that, for any $0 < h < \infty$ and $x = g'(h)$,

(i) $\sqrt{2\pi} \int_0^\infty e^{-hy} \check{G}_h(dy) - 1 \leq C_2(g''(h))^{-1}(1 - g''(h))^{1/2}$;

(ii) $1 - \sqrt{2\pi} x \int_0^\infty e^{-hy} \check{G}_h(dy) \leq C_2(g''(h))^{-1}(1 - g''(h))^{1/2} + (g''(h))^{-3/2}h^{-2}$.

Proof of Theorem 3.1. Let $m$, $x$ and $a$ be given. It suffices to verify that

$$\int_x^\infty G(dy) \leq \frac{1}{\sqrt{2\pi} x} e^{-x^2/2}(1 + C_1x^{-2}), \quad (17)$$

due to Lemma 3.2. We may assume that $1 \leq x \leq \sum_{i=1}^{n} a_i$, since $\int_x^\infty G(dy) = 0$ for $x > \sum_{i=1}^{n} a_i$. Regarding $x = \sum_{i=1}^{n} a_i$, it follows that $\int_x^\infty G(dy) = 2^{-m}$, whereas

$$x = \sum_{i=1}^{n} a_i \leq \sqrt{m} \left( \sum_{i=1}^{n} a_i^2 \right)^{1/2} = \sqrt{m},$$

so that

$$x^{-1}e^{-x^2/2} \geq m^{-1/2}e^{-m/2} = 2^{-m}m^{-1/2} \exp(m(\log 2 - 1/2)).$$

Thus, (17) holds for $x = \sum_{i=1}^{n} a_i$, since

$$1 \leq \frac{1}{\sqrt{2\pi m}} \exp(m(\log 2 - 1/2)),$$

for $m$ sufficiently large.

In view of Lemma 3.1 and (17) we thus consider verifying that, for $1 \leq x < \sum_{i=1}^{n} a_i$, with $h$ such that $g'(h) = x$,

$$\exp \left( x^2/2 + g(h) - hx \right) \left( \sqrt{2\pi} x \int_0^\infty e^{-hy} \check{G}_h(dy) \right) \leq 1 + C_1x^{-2}.$$

By taking logarithms, we may equally well verify that

$$x^2/2 + g(h) - hx + \log \left( \sqrt{2\pi} x \int_0^\infty e^{-hy} \check{G}_h(dy) \right) \leq C_1x^{-2}. \quad (18)$$

For the proof of (18), introduce a number $\lambda$ satisfying

$$0 < \lambda < (\log 2 - 1/2)^2/4.$$
As a first step, consider the case \( h^2 g''(h) \leq \lambda x^2 \). Then, by Proposition 3.1(ii),

\[
x^2/2 + g(h) - hx \leq -x^2 (\log 2 - 1/2) + 2x^2 \sqrt{\lambda} = -\varepsilon x^2,
\]

with

\[
\varepsilon = \log 2 - 1/2 - 2\sqrt{\lambda} > 0,
\]

by the upper bound on \( \lambda \). Thus,

\[
x^2/2 + g(h) - hx \leq -x^2 (\log 2 - 1/2) + 2x^2 \sqrt{\lambda} = -\varepsilon x^2,
\]

which is sufficient for (18) regarding this case.

Secondly, consider the alternative case \( h^2 g''(h) > \lambda x^2 \), and \( x^2 (1 - g''(h)) > 12\lambda^{-1} \log (\sqrt{2\pi} x) \). (19)

Then, by Proposition 3.1(i),

\[
x^2/2 + g(h) - hx \leq -(1 - g''(h))\lambda x^2/12 \leq -\log (\sqrt{2\pi} x),
\]

and thus,

\[
x^2/2 + g(h) - hx + \log (\sqrt{2\pi} x) \int_0^\infty e^{-hy} \tilde{G}_h(dy) \leq 0.
\]

As a third and final step, consider the case

\[
x^2 (1 - g''(h)) \leq 12\lambda^{-1} \log (\sqrt{2\pi} x). \tag{20}
\]

From Proposition 3.2(i),

\[
\sqrt{2\pi} x \int_0^\infty e^{-hy} \tilde{G}_h(dy) \leq 1 + C_2 (g''(h))^{-1} \sqrt{1 - g''(h)}. \tag{21}
\]

It suffices, in view of Proposition 3.1 (i) and (21), with \( \log(1 + x) \leq x \), for \( x \geq 0 \), to verify that

\[
-(1 - g''(h))g''(h)h^2/12 + C_2 (g''(h))^{-1} \sqrt{1 - g''(h)} \leq C_1 x^{-2},
\]

which may be rewritten as

\[
(g''(h))^{-1} \sqrt{1 - g''(h)} \left( -\sqrt{1 - g''(h)} (g''(h))^2 h^2/12 + C_2 \right) \leq C_1 x^{-2}. \tag{22}
\]

Here, it indeed suffices to consider the case

\[
\sqrt{1 - g''(h)} (g''(h))^2 h^2/12 \leq C_2. \tag{23}
\]

Now, (20) implies that \( g''(h) \) approaches 1, as \( x \) tends to infinity. Thus, from (20) and (23), noting that \( h \geq x \), we deduce that, for some constant \( C \),

\[
\sqrt{1 - g''(h)} \leq C x^{-2}. \tag{24}
\]
Next, (20) and (24) imply that

$$(g''(h))^{-1} \sqrt{1 - g''(h)} \leq C x^{-2} \left(1 - 12 \lambda^{-1} x^{-2} \log \left(\sqrt{2 \pi x} \right) \right)^{-1} \leq \tilde{C} x^{-2},$$

for some constant $\tilde{C}$. Thus,

$$C_2 (g''(h))^{-1} \sqrt{1 - g''(h)} \leq C_2 \tilde{C} x^{-2},$$

which proves that (22) holds for the case (20).

\[\square\]

**Proof of Lemma 3.1.** It suffices to verify that

$$e^{g(h) - h x} \int_0^{\infty} e^{-hy} \tilde{G}_h(dy) = \int_x^{\infty} G(dy),$$

with $G$ referring to the distribution of $S$. Exponential tilting corresponds to the operation

$$G^*(d(y)) = e^{h_y - g(h)} G(dy),$$

which gives a distribution with mean $g'(h) = x$, cf. [6, Section XVI.7]. Thus, since $\tilde{G}$ is the mean centered version of $G^*$,

$$\tilde{G}_h(dy - x) = e^{h_y - g(h)} G(dy).$$

Therefore, by a change of variables,

$$\int_x^{\infty} G(dy) = \int_x^{\infty} e^{g(h) - hy} \tilde{G}_h(dy - x)$$

$$= \int_0^{\infty} e^{g(h) - hy - hx} \tilde{G}_h(dy),$$

which verifies (25).

\[\square\]

**Proof of Lemma 3.3.** Note that the considered inequality can be rewritten as

$$\sum_{i=1}^{n} \log (\cosh(a_i h)) - (a_i h)^2 / 2 + \frac{(a_i h)^2}{12} \left( \frac{\sinh(a_i h)}{\cosh(a_i h)} \right)^2 \leq 0.$$

Thus, it suffices to consider the case $n = 1$, i.e. to prove that, for any $t \geq 0$,

$$\log (\cosh(t)) - \frac{t^2}{2} \leq -\left( \frac{\sinh(t)}{\cosh(t)} \right)^2 \frac{t^2}{12}.$$

To begin with, consider the restricted statement

$$\log (\cosh(t)) - \frac{t^2}{2} \leq -\left( \frac{\sinh(t)}{\cosh(t)} \right)^2 \frac{t^2}{12} \text{ for } t \geq 6/5. \quad (26)$$

Clearly,

$$0 \leq \sinh(t) \leq \cosh(t). \quad (27)$$
Thus, (26) follows from
\[ \log(\cosh(t)) - \frac{t^2}{2} \leq -\frac{t^2}{12}, \quad \text{for} \quad t \geq 6/5, \]
which we rewrite as
\[ 2 \exp \left( \frac{5t^2}{12} \right) \geq e^t + e^{-t}, \quad \text{for} \quad t \geq 6/5, \]
or in other words,
\[ e^{2t} \left( 2 \exp \left( \frac{5t^2}{12} - t \right) - 1 \right) \geq 1, \quad \text{for} \quad t \geq 6/5. \quad (28) \]
Differentiation of \( 5t^2/12 - t \) with respect to \( t \) gives \( \frac{5t}{6} - 1 \geq 0 \), and thus that \( 5t^2/12 - t \) is non-decreasing within the considered interval. Thus, the left hand side of the inequality in (28) is a non-decreasing function of \( t \). Moreover, for \( t = 6/5 \),
\[ e^{12/5} \left( 2 \exp \left( -3/5 \right) - 1 \right) \approx 1.08 > 1. \]
Thus, we conclude that (28), and therefore also (26) holds.

It now remains to verify that
\[ \log(\cosh(t)) - \frac{t^2}{2} \leq -\left( \frac{\sinh(t)}{\cosh(t)} \right)^2 \frac{t^2}{12}, \quad \text{for} \quad t \leq 6/5, \]
which we rewrite as
\[ \cosh(t) \exp \left( \left( \frac{\sinh(t)}{\cosh(t)} \right)^2 \frac{t^2}{12} \right) \leq e^{t^2/2}, \quad \text{for} \quad t \leq 6/5. \quad (29) \]
First, in view of (27),
\[ \left( \frac{\sinh(t)}{\cosh(t)} \right)^2 \frac{t^2}{12} \leq (6/5)^2/12 = 0.12 < 1. \]
With the inequality \( e^x \leq 1 + x + x^2, \quad x \leq 1, \) (cf. [7, Appendix A, Lemma 1.1]), and (27) we thus deduce that
\[
\exp \left( \left( \frac{\sinh(t)}{\cosh(t)} \right)^2 \frac{t^2}{12} \right) \leq 1 + \left( \frac{\sinh(t)}{\cosh(t)} \right)^2 \frac{t^2}{12} + \left( \frac{\sinh(t)}{\cosh(t)} \right)^4 \frac{t^4}{144} \\
\leq 1 + \left( \frac{\sinh(t)}{\cosh(t)} \right)^2 \left( \frac{t^2}{12} + \frac{t^4}{144} \right).
\]
Noting that \( e^{t^2/2} \geq 1 + \frac{t^2}{2} + \frac{t^4}{8} + \frac{t^6}{48} \), by truncation of the corresponding Taylor series, it hence suffices to verify that
\[ \cosh(t) + \left( \frac{\sinh(t)}{\cosh(t)} \right)^2 \left( \frac{t^2}{12} + \frac{t^4}{144} \right) \leq 1 + \frac{t^2}{2} + \frac{t^4}{8} + \frac{t^6}{48}, \quad \text{for} \quad t \leq 6/5. \quad (30) \]
We first verify that
\[ \left( \frac{\sinh(t)}{\cosh(t)} \right)^2 \leq t^2 \cosh(t), \quad \text{for} \quad t \leq 6/5. \quad (31) \]
Indeed, by Taylor expansion, with $\lambda = 6/5$,

$$
\sinh(t) = t + \frac{t^3}{6} + \frac{t^5}{\lambda^5} \sum_{k=2}^{\infty} \frac{\lambda^5 t^{2k-4}}{(2k+1)!}
$$

$$
\leq t + \frac{t^3}{6} + \frac{t^5}{\lambda^5} \sum_{k=2}^{\infty} \frac{\lambda^5}{(2k+1)!}
$$

$$
= t + \frac{t^3}{6} + \frac{t^5}{\lambda^5} \left( \sinh(\lambda) - (\lambda + \frac{\lambda^3}{6}) \right)
$$

$$
\leq t + \frac{t^3}{6} + \frac{t^5}{\lambda^5} \left( \sinh(\lambda) - (\lambda + \frac{\lambda^3}{6}) \right) \leq t + 0.18t^3,
$$

whereas $\cosh(t) \geq 1 + \frac{t^2}{2} + \frac{t^4}{24}$, by truncation of the corresponding Taylor series. We arrive at (31), since

$$
(t + 0.18t^3)^2 = t^2 (1 + 0.36t^2 + 0.0324t^4) \leq t^2 (1 + t^2/2 + t^4/24).
$$

Similarly, with $t \leq \lambda = 6/5$,

$$
\cosh(t) = 1 + \frac{t^2}{2} + \frac{t^4}{24} + \frac{t^6}{\lambda^6} \sum_{k=3}^{\infty} \frac{\lambda^6 t^{2k-6}}{2k!}
$$

$$
\leq 1 + \frac{t^2}{2} + \frac{t^4}{24} + \frac{t^6}{\lambda^6} \left( \cosh(\lambda) - (1 + \frac{\lambda^2}{2} + \frac{\lambda^4}{24}) \right)
$$

$$
\leq 1 + \frac{t^2}{2} + \frac{t^4}{24} + 0.0015t^6.
$$

By combining (31) and (32) we conclude that (30) holds, since

$$
(1 + \frac{t^2}{2} + \frac{t^4}{24} + 0.0015t^6) + \left( \frac{t^4}{12} + \frac{t^6}{144} \right) = 1 + \frac{t^2}{2} + \frac{t^4}{8} + t^6 \left( 0.0015 + \frac{1}{144} \right)
$$

$$
\leq 1 + \frac{t^2}{2} + \frac{t^4}{8} + \frac{t^6}{48}.
$$

This completes the proof of the lemma. \qed

Proof of Proposition 3.1. Differentiation of $x^2/2 + g(h) - hx$ with respect to $h$ gives

$$
g'(h)g''(h) + g'(h) - g'(h) - hg''(h) = -g''(h)(h - g'(h)).
$$

Thus, since $(g'(0))^2/2 + g(0) - hg'(0) = 0$, and $g'(0) = 0$,

$$
(g'(h))^2/2 + g(h) - hg'(h) = - \int_0^h g''(y)(y - g'(y)) dy.
$$

Now, (cf. (15)),

$$
g''(0) = 1, \quad g'''(h) = -2 \sum_{i=1}^{n} a_i^3 \frac{\sinh(a_i h)}{(\cosh(a_i h))^3} \leq 0.
$$
Thus, \( h \mapsto g''(h) \) is a non-increasing function, and \( g'(h) \leq h \). We infer that
\[
\int_0^h g''(y)(y - g'(y)) \, dy \geq g''(h) \int_0^h y - g'(y) \, dy = g''(h)(h^2/2 - g(h)).
\] (34)

Combining (33) with (34) and Lemma 3.3 gives
\[
(g'(h))^2/2 + hg'(h) \leq -(1 - g''(h))g''(h)h^2/12,
\]
which proves part (i).

Regarding part (ii), with a slight reformulation, it suffices to verify that
\[
\sum_{i=1}^n \log (\cosh(a_i h)) - a_i h \frac{\sinh(a_i h)}{\cosh(a_i h)} \leq -(\log 2)x^2 + 2hx \sqrt{g''(h)}.
\] (35)

To begin with,
\[
\log (\cosh(t)) = t - \frac{t^3}{6} + \cdots \leq t - \frac{t^3}{6},
\]
whereas
\[
\sinh(t) \cosh(t) = 1 - \frac{t^2}{2} + \cdots \geq 1 - \frac{t^2}{2}.
\]
Thus,
\[
\sum_{a_i h \geq \lambda} \log (\cosh(a_i h)) - a_i h \frac{\sinh(a_i h)}{\cosh(a_i h)} \leq -(\log 2)x^2 + 2hx \sqrt{g''(h)}.
\] (36)

On the other hand,
\[
h^2g''(h) = \sum_{i=1}^n \left( \frac{a_i h}{\cosh(a_i h)} \right)^2,
\] (37)
whereas

\[
\begin{align*}
x &= \sum_{i=1}^{n} a_i \frac{\sinh(a_i h)}{\cosh(a_i h)} \\
&= \sum_{a_i h < \lambda} a_i \frac{\sinh(a_i h)}{\cosh(a_i h)} + \sum_{a_i h \geq \lambda} a_i \frac{\sinh(a_i h)}{\cosh(a_i h)} \\
&\leq \sum_{a_i h < \lambda} a_i^2 h + \sum_{a_i h \geq \lambda} a_i \\
&\leq \left( \sum_{a_i h < \lambda} a_i^2 \right)^{1/2} \left( \sum_{a_i h < \lambda} a_i^2 h^2 \right)^{1/2} + \left( \sum_{a_i h \geq \lambda} a_i^2 \right)^{1/2} \left( \sum_{a_i h \geq \lambda} \right)^{1/2} \\
&\leq \cosh(\lambda) \left( \sum_{a_i h < \lambda} \left( \frac{a_i h}{\cosh(a_i h)} \right)^2 \right)^{1/2} + \left( \sum_{a_i h \geq \lambda} \right)^{1/2} \\
&= \cosh(\lambda) h \sqrt{g''(h)} + \left( \sum_{a_i h \geq \lambda} \right)^{1/2}.
\end{align*}
\]

Thus, rewriting the above inequality, combined with the fact that \( \log 2 \cosh(\lambda) \leq 1 \), we obtain

\[
- \log 2 \sum_{a_i h \geq \lambda} \leq - \log 2 \left( x^2 + 2 \cosh(\lambda) h x \sqrt{g''(h)} - (\cosh(\lambda))^2 h^2 g''(h) \right) \\
&\leq - \log 2 x^2 + 2 h x \sqrt{g''(h)} - \log 2(\cosh(\lambda))^2 h^2 g''(h).
\]

(38)

Turning to the second part of the right hand side in (36), first note that \((\log 2)e^{2\lambda} > 4\), while \((1 + 2t) \leq 4t^2\), for \( t \geq 0.9 \). Thus,

\[
(1 + 2t) \leq (\log 2)e^{2\lambda} t^2, \quad \text{for } t \geq \lambda,
\]

Hence, for \( t \geq \lambda \),

\[
(1 + 2t)e^{-2t} \leq (\log 2)t^2 e^{2\lambda - 2t} \\
= (\log 2)t^2 \left( \frac{\cosh(\lambda)}{\cosh(t)} \right)^2 \left( \frac{1 + e^{-2t}}{1 + e^{-2\lambda}} \right)^2 \leq (\log 2)t^2 \left( \frac{\cosh(\lambda)}{\cosh(t)} \right)^2.
\]

Thus,

\[
\sum_{a_i h \geq \lambda} (1 + 2a_i h)e^{-2a_i h} \leq \log 2(\cosh(\lambda))^2 \sum_{a_i h \geq \lambda} \left( \frac{a_i h}{\cosh(a_i h)} \right)^2 \\
= \log 2(\cosh(\lambda))^2 h^2 g''(h).
\]

(39)

Finally, combining (36), (38) and (39) gives the desired inequality (35).

\[\square\]

Proof of Lemma 3.4. By definition,

\[
E|\tilde{X}_i|^3 = a_i^3 \frac{e^{2a_i h} + e^{-2a_i h}}{2(\cosh(a_i h))^4} = a_i^3 \frac{(2 \cosh(a_i h))^2 - 2}{2(\cosh(a_i h))^4} \leq \frac{2a_i^3}{(\cosh(a_i h))^2}.
\]

12
Now, since \( \sinh(t) \geq t \), for \( t \geq 0 \),
\[
\sum_{i=1}^{n} E[|\tilde{X}_i|^3] \leq 2 \sum_{i=1}^{n} a_i h \frac{a_i^2}{(\cosh(a_i h))^2}
\]
\[
\leq 2 \sum_{i=1}^{n} \sinh(a_i h) \frac{a_i^2}{(\cosh(a_i h))^2}
\]
\[
\leq 2 \left( \sum_{i=1}^{n} \left( \frac{a_i \sinh(a_i h)}{\cosh(a_i h)} \right)^2 \right)^{1/2} \left( \sum_{i=1}^{n} \left( \frac{a_i}{\cosh(a_i h)} \right)^2 \right)^{1/2}
\]
\[
= 2\sqrt{(1 - g''(h))g''(h)},
\]
which proves the desired inequality.

Proof of Proposition 3.2. First, we rewrite (cf. [7, Theorem 12.1, Chapter 2]),
\[
\int_{0}^{\infty} e^{-hy} \tilde{G}(dy) = \int_{0}^{1} P(y \leq e^{-h\tilde{S}(h)} \leq 1) dy = \int_{0}^{1} P(0 \leq h\tilde{S}(h) \leq -\log y) dy. \tag{40}
\]
We recall, cf. [6, Section XVI.7], that \( \tilde{S}(h) \) has mean zero and variance \( g''(h) \). For brevity, set \( \sigma = \sqrt{g''(h)} \). By the Berry–Esseen theorem (cf. [7, Theorem 6.2, Chapter 7]), for some constant \( C \),
\[
\left| P(0 \leq h\tilde{S}(h) \leq -\log y) - \Phi(-\sigma h^{-1}\log y) + 1/2 \right| \leq C\sigma^{-2} \sum_{i=1}^{n} E[|\tilde{X}_i|^3]. \tag{41}
\]
Combining (40), (41) and Lemma 3.4 gives
\[
\left| \int_{0}^{\infty} e^{-hy} \tilde{G}(dy) - \int_{0}^{1} \Phi(-\sigma h^{-1}\log y) dy + 1/2 \right|
\]
\[
\leq \int_{0}^{1} \left| P(0 \leq h\tilde{S}(h) \leq -\log y) - \Phi(-\sigma h^{-1}\log y) + 1/2 \right| dy
\]
\[
\leq C(hg''(h))^{-1} \sqrt{1 - g''(h)}. \tag{42}
\]
Now, with a reversed reasoning compared to (40),
\[
\int_{0}^{1} \Phi(-\sigma h^{-1}\log y) dy - 1/2 = \int_{0}^{\infty} e^{-hy} \frac{e^{-y^2/2\sigma}}{\sqrt{2\pi \sigma}} dy
\]
\[
= e^{(h\sigma)^2/2} \int_{0}^{\infty} e^{-(y+h\sigma)^2/2} \sqrt{2\pi} dy
\]
\[
= e^{(h\sigma)^2/2} N(h\sigma). \tag{43}
\]
Combining (42)–(43) gives,
\[
\sqrt{2\pi x} \left| \int_{0}^{\infty} e^{-by} \tilde{G}(dy) - e^{(h\sigma)^2/2} N(h\sigma) \right| \leq C\sqrt{2\pi x}(hg''(h))^{-1} \sqrt{1 - g''(h)}.
\]
Here, \( x \leq h \), due to (14), which proves that
\[
\sqrt{2\pi x} \left| \int_{0}^{\infty} e^{-by} \tilde{G}(dy) - e^{(h\sigma)^2/2} N(h\sigma) \right| \leq C_2(g''(h))^{-1} \sqrt{1 - g''(h)},
\]
13
for some constant $C_2$. It now suffices to verify that
\[
\sqrt{2\pi} xe^{(h\sigma)^2/2} N(h\sigma) - 1 \leq (g''(h))^{-1} \sqrt{1 - g''(h)},
\] (44)
and that
\[
1 - \sqrt{2\pi} xe^{(h\sigma)^2/2} N(h\sigma) \leq \sqrt{1 - g''(h)} + (g''(h))^{-3/2} h^{-2}.
\] (45)
Regarding (44), it follows from Lemma 3.2, recalling that $\sigma = \sqrt{g''(h)}$, that
\[
\sqrt{2\pi} xe^{(h\sigma)^2/2} N(h\sigma) - 1 \leq (x/h)(g''(h))^{-1/2} - 1.
\]
Now, since $0 < g''(h) \leq 1$,
\[
(g''(h))^{-1/2} - 1 \leq (g''(h))^{-1} - 1 = (g''(h))^{-1}(1 - g''(h)) \leq (g''(h))^{-1} \sqrt{1 - g''(h)},
\]
which proves (44). Similarly, due to Lemma 3.2,
\[
1 - \sqrt{2\pi} xe^{(h\sigma)^2/2} N(h\sigma) \leq 1 - (x/h)(g''(h))^{-1/2} + x(h^2 g''(h))^{-3/2}.
\]
Applying $x \leq h$, $x \geq g''(h)h$ and $0 < g''(h) \leq 1$ gives
\[
1 - (x/h)(g''(h))^{-1/2} + x(h^2 g''(h))^{-3/2} \leq 1 - (g''(h))^{1/2} + h^{-2}(g''(h))^{-3/2} \leq \sqrt{1 - (g''(h)) + h^{-2}(g''(h))^{-3/2}},
\]
which proves (45).

\[\square\]

4 The tails of binomial $B_n$ for $cn^{1/6} \leq x \leq Cn^{1/6}$

For the present section, define $x_{k,n} = (2k - n)/\sqrt{n}$ and $p_{k,n} = 2^{-n} \binom{n}{k}$, for $k = 0, \ldots, n$. Thus, $B_n$ assumes values $\{x_{k,n}\}_{k=0}^n$ with probabilities $\{p_{k,n}\}_{k=0}^n$.

**Theorem 4.1.** Let $R_{k,n}$ be given by
\[
P(B_n \geq x_{k,n}) = N(x_{k,n}) \left(1 + \frac{3}{x_{k,n}^2} - \frac{1}{12x_{k,n}^4} \left(6 - \frac{x_{k,n}^2}{\sqrt{n}}\right)^2 + R_{k,n}\right).
\]
Then, for any positive constants $c$ and $C$,
\[
\lim_{n \to \infty} \sup \left\{x_{k,n}^2 R_{k,n} : \text{with } k \text{ such that } cn^{1/6} \leq x_{k,n} \leq Cn^{1/6}\right\} = 0.
\]
Noting that $x \mapsto P(B_n \geq x)$ attains local maxima at $\{x_{k,n}\}_k$, and by omitting, or cancelling the term $\frac{1}{12x_{k,n}^4} \left(6 - \frac{x_{k,n}^2}{\sqrt{n}}\right)^2$, respectively, we arrive at the following two corollaries.

**Corollary 4.1.** For any positive constants $c$ and $C$,
\[
\lim_{n \to \infty} \sup \left\{x^2 (P(B_n \geq x)/N(x) - 1) : \text{with } x \text{ s. t. } cn^{1/6} \leq x \leq Cn^{1/6}\right\} \leq 3.
\]
Corollary 4.2. For any sequence \( \{x_{k,n}\} \) such that \( \lim_{n \to \infty} x_{k,n}^6/n = 36 \),
\[
\lim_{n \to \infty} x_{k,n}^2 P(B_n \geq x_{k,n})/N(x_{k,n}) - 1 = 3.
\]

We refer to Section 5 for numerical evaluations relating to Corollaries 4.1 and 4.2. The proof of Theorem 4.1 is based on approximation of \( p_{k,n} \) for corresponding \( x_{k,n} \) of relevant order, and on approximation of corresponding sums of approximate probabilities (cf. Lemmas 4.1 and 4.2).

Lemma 4.1. Let \( R_{k,n} \) be given by
\[
p_{k,n} = \frac{2}{\sqrt{2\pi n}} \left( 1 + R_{k,n} \right).
\]
Then, for any positive constant \( c \),
\[
\lim_{n \to \infty} \sup \left\{ x_{k,n}^2 R_{k,n} : \text{with } k \text{ such that } cn^{1/6} \leq x_{k,n} \leq n^{1/5} \right\} = 0.
\]

Lemma 4.2. Consider pairs \( (k,l) \) with \( 0 \leq k \leq n - 1 \) and \( k + 1 \leq l \leq n \), and let \( R_{k,l,n} \) be given by
\[
\lim_{n \to \infty} \sup \left\{ x_{k,n}^2 R_{k,l,n} : \text{with } cn^{1/6} \leq x_{k,n} \leq Cn^{1/6}, \ cn^{1/5} \leq x_{l,n} \leq n^{1/5} \right\} = 0.
\]
A consequence of Lemma 4.1 is that \( x_{k,n}^3 \sim \sqrt{n}C \), for some constant \( C \), yields
\[
p_{k,n} \sim \frac{2}{\sqrt{2\pi n}} e^{-\frac{1}{2} x_{k,n}^2} \sim N(x_{k,n}) \frac{2x_{k,n}^2}{\sqrt{n}} \sim N(x_{k,n}) \frac{2C}{x_{k,n}^2}.
\]
For instance, when \( C = 6 \), as in Corollary 4.2, then \( p_{k,n} \sim 12N(x_{k,n})x_{k,n}^{-2} \), whereas \( P(B_n \geq x_{k,n}) - N(x_{k,n}) \sim 3N(x_{k,n})x_{k,n}^{-2} \). Thus, it is essential to Corollary 4.2 that tail probabilities for points of discontinuity of the distribution function of \( B_n \) are considered.

Proof of Theorem 4.1. Let \( x = x_{k,n} \) be given such that \( cn^{1/6} \leq x \leq Cn^{1/6} \), given some constants \( c \) and \( C \). Then introduce \( x_{l,n} \) such that \( cn^{1/5} \leq x_{l,n} \leq n^{1/5} \). To begin with,
\[
N(x) = \int_{x}^{x_{l,n}} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} y^2} dy + N(x_{l,n}).
\]
Here, \( \int_{x}^{x_{l,n}} e^{-\frac{1}{2} y^2} dy \) is proportional to \( x_{l,n}^{-1} e^{-\frac{1}{2} x_{l,n}^2} \), according to Lemma 3.2, and thus negligible compared to \( x_{l,n}^{-2} N(x) \), which is proportional to \( x_{l,n}^{-3} e^{-\frac{1}{2} x_{l,n}^2} \). Similarly,
\[
P(S_n \geq x) = \sum_{r=k}^{l-1} \int_{x_{r,n}}^{x_{r+1,n}} \frac{\sqrt{n}}{2} P_{r,n} dy + P(S_n \geq x_{l,n}).
\]
By Theorem 3.1, \( P(S_n \geq x_{l,n}) \leq N(x_{l,n})(1 + C_1 x_{l,n}^{-2}) \), so that also \( P(S_n \geq x_{l,n}) \) is negligible compared to \( x^{-2}N(x) \). Thus, it suffices to verify that

\[
\sum_{r=k}^{l-1} \int_{x_{r,n}}^{x_{r+1,n}} \left( \frac{\sqrt{n}}{2} p_{r,n} - e^{-\frac{x^2}{2}} \right) dy = N(x) \left( 3x^{-2} - \frac{1}{12} \left( \frac{x^2}{\sqrt{n}} - 6x^{-1} \right)^2 + R \right),
\]

with a remainder term \( R \) negligible compared to \( x^{-2} \). Applying Lemma 4.1, note that, by Theorem 3.1,

\[
o(x^{-2}) \sum_{r=k}^{l-1} p_{r,n} \leq o(x^{-2})N(x)(1 + C_1 x^{-2}) = o(x^{-2})N(x).
\]

Thus, (46) follows from

\[
\sum_{r=k}^{l-1} \int_{x_{r,n}}^{x_{r+1,n}} \left( e^{-\frac{x^2}{2}} - e^{-\frac{x^2}{2}} \right) dy = \sqrt{2\pi}N(x) \left( \frac{3}{x^2} - \frac{1}{12} \left( \frac{x^2}{\sqrt{n}} - \frac{6}{x} \right)^2 + R \right),
\]

i.e., from Lemma 4.2. \( \square \)

**Proof of Lemma 4.1.** By definition,

\[
p_{k,n} = 2^{-n} \frac{n!}{(n-k)!k!},
\]

with

\[
n - k = n/2 - \sqrt{n}x/2 \geq n/2(1 + n^{-3/10}),
\]

and

\[
k = n/2 + \sqrt{n}x/2 \geq n/2.
\]

Thus, the relative errors from Stirling approximation in (47) will be of size \( n^{-1} \) (cf. [5, Chapter II.9]), and thus negligible compared to \( x^{-2} \geq n^{-2/5} \). With Stirling’s formula in (47),

\[
\frac{\sqrt{2\pi n}}{2} p_{k,n}(1 + R) = \left( \frac{2k}{n} \right)^{-k-\frac{1}{2}} \left( \frac{2(n-k)}{n} \right)^{-(n-k)-\frac{1}{2}} \left( 1 + \frac{x}{\sqrt{n}} \right)^{-k-\frac{1}{2}} \left( 1 - \frac{x}{\sqrt{n}} \right)^{-(n-k)-\frac{1}{2}}.
\]

Here,

\[
\left( 1 + \frac{x}{\sqrt{n}} \right)^{-\frac{1}{2}} \left( 1 - \frac{x}{\sqrt{n}} \right)^{-\frac{1}{2}} = \left( 1 - \frac{x^2}{n} \right)^{-\frac{1}{2}},
\]

with \( x^2/n \leq x^{-3} \), which is negligible compared to \( x^{-2} \). Thus, taking logarithms, it suffices to verify that

\[
k \log \left( 1 + \frac{x}{\sqrt{n}} \right) + (n-k) \log \left( 1 - \frac{x}{\sqrt{n}} \right) = \frac{x^2}{2} + \frac{x^4}{12n} + R.
\]

(48)
By the relation $2k - n = x/\sqrt{n}$ and Taylor expansion of the two logarithms in (48), for some $0 \leq \xi_1, \xi_2 \leq x/\sqrt{n}$,

\[
k \log \left(1 + \frac{x}{\sqrt{n}}\right) + (n-k) \log \left(1 - \frac{x}{\sqrt{n}}\right) = (2k - n) \left(\frac{x}{\sqrt{n}} + \frac{x^3}{3 \cdot n^{3/2}}\right)
- n \left(\frac{1}{2} \frac{x^2}{n} + \frac{1}{4} \frac{x^4}{n^2}\right) + \frac{1}{5} \left(k \xi_1^5 - (n-k) \xi_2^5\right).
\]

\[
= \frac{x^2}{2} + \frac{x^4}{12n} + \frac{1}{5} \left(k \xi_1^5 - (n-k) \xi_2^5\right).
\]

(49)

Here, since $n \geq x^5$,

\[
n \frac{x^5}{n^{3/2}} = x^5 n^{-3/2} \leq x^{-(15/2 - 5)} = x^{-(15/2 - 5)} = x^{-5/2},
\]

which is negligible compared to $x^{-2}$. Thus, the two remainder terms in (49) are also negligible compared to $x^{-2}$, which proves (48).

\[\square\]

Proof of Lemma 4.2. First, consider a given $r$, $k \leq r \leq l-1$, and set $z = x_{r,n}$. Note that

\[
e^{-\frac{x^2}{n}} - e^{-\frac{x^2}{r,n}} = e^{-\frac{x^2}{r,n}} \left(1 + 1 - e^{-\frac{x^2}{r,n}}\right).
\]

Here,

\[
e^{-\frac{x^4}{r,n}} - 1 = -\frac{z^4}{12n} + R,
\]

with $R$ being of order $z^5 n^{-2} \leq n^{5/5 - 2} = n^{-2/5}$, and thus negligible compared to $x^{-2} \geq n^{-2/5} c^{-2}$. Similarly,

\[
1 - e^{-\frac{x^2}{r,n}} = \frac{y^2}{2} - \frac{z^2}{2} + R,
\]

with $R$ being of order

\[
(y^2 - z^2)^2/16 \leq x_{r+1,n}/n \leq n^{2/5 - 1} = n^{-3/5},
\]

and thus negligible compared to $x^{-2}$. On the other hand,

\[
\sum_{r=k}^{l-1} \int_{x_{r,n}}^{x_{r+1,n}} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2n}} dy = \sum_{r=k-1}^{l-2} \int_{x_{r,n}}^{x_{r+1,n}} \frac{1}{\sqrt{2\pi}} e^{-\frac{x_{r+1,n}^2}{2n}} dy \leq N(x_{k-1,n}).
\]

(50)

Thus, noting that

\[
3x^{-2} = \frac{1}{12} \left(\frac{x^2}{\sqrt{n}} - 6x^{-1}\right)^2 = \frac{x}{\sqrt{n}} - \frac{x^4}{12n},
\]

it suffices to verify that

\[
\sum_{r=k}^{l-1} \int_{x_{r,n}}^{x_{r+1,n}} \frac{1}{\sqrt{2\pi}} \left(y^2 - x_{r,n}^2 - \frac{x_{r,n}^4}{6n}\right) dy = \sqrt{2\pi} N(x) \left(\frac{x}{\sqrt{n}} - \frac{x^4}{12n} + R\right).
\]

17
Now, introducing \( \Delta = x_{r+1,n} - x_{r,n} = 2/\sqrt{n} \),
\[
\int_{x_{r,n}}^{x_{r+1,n}} \left( \frac{y^2}{2} - \frac{x_{r,n}^2}{2} - \frac{x_{r,n}^4}{12n} \right) dy = \frac{x_{r+1,n}^3}{6} - \frac{x_{r,n}^3}{6} - \frac{\Delta x_{r,n}^2}{2} - \frac{\Delta x_{r,n}^4}{12n},
\]
where
\[
x_{r+1,n}^3 - x_{r,n}^3 - 3\Delta x_{r,n}^2 = \Delta^3 - 3x_{r,n}^2(\Delta + x_{r,n}) + 3x_{r,n}(\Delta + x_{r,n})^2 - 3\Delta x_{r,n}^2
to
= \Delta^3 + x_{r,n}^3(3 - 3) + \Delta x_{r,n}^2(6 - 3 - 3) + 3\Delta^2 x_{r,n}
= \Delta^3 + 3\Delta^2 x_{r,n}.
\]
By (50) and \( n^{-1} \leq x^{-6}C^6 \),
\[
\sum_{r=k}^{l-1} e^{-\frac{x_{r,n}^2}{2}} \frac{\Delta^3}{\sqrt{n}} \leq \sqrt{2\pi} \Delta^2 N(x_{k-1,n}) = o(x^{-2})N(x),
\]
so that the contribution from \( \Delta^3 \) is negligible. Thus, with reference to the two remaining terms, it suffices to verify that
\[
\sum_{r=k}^{l-1} \Delta e^{-\frac{x_{r,n}^2}{2}} \left( \frac{x_{r,n}}{\sqrt{n}} - \frac{x_{r,n}^4}{12n} \right) = \sqrt{2\pi} N(x) \left( \frac{x}{\sqrt{n}} - \frac{x^4}{12n} + R \right),
\]
with \( R = o(x^{-2}) \).

By partial integration,
\[
\int_{x}^{\infty} y^4 e^{-\frac{y^2}{2}} dy = x^3 e^{-\frac{x^2}{2}} + 3x^2 e^{-\frac{x^2}{2}} + 3N(x) = \sqrt{2\pi} N(x)x^4(1 + R),
\]
with \( R = o(1) \). Thus,
\[
\sum_{r=k}^{l-1} \Delta e^{-\frac{x_{r,n}^2}{2}} \frac{x_{r,n}^4}{12n} = \sqrt{2\pi} N(x) \frac{x^4}{12n} \left( 1 + R \right),
\]
with \( R = o(1) \). Note that \( e^{x^2} \leq x^4/n \leq C^4x^{-2} \), so that \( |x^4 R/n| \leq o(x^{-2}) \). Similarly,
\[
\int_{x}^{\infty} y e^{-\frac{y^2}{2}} dy = e^{-\frac{x^2}{2}} = x\sqrt{2\pi} N(x)(1 + R),
\]
with \( R = o(1) \). Thus, analogously,
\[
\sum_{r=k}^{l-1} \Delta e^{-\frac{x_{r,n}^2}{2}} \frac{x_{r,n}}{\sqrt{n}} = \sqrt{2\pi} N(x) \frac{x}{\sqrt{n}} \left( 1 + R \right),
\]
and \( cx^{-2} \leq x/\sqrt{n} \leq Cx^{-2} \). Combining (52) and (53) yields (51). \( \square \)
Starting with the results of Section 4, Figure 1 complements Corollaries 4.1 and 4.2 regarding the behaviour of

\[ C_{k,n} = x_{k,n}^2 \left( \frac{P(B_n \geq x_{k,n})}{N(x_{k,n})} - 1 \right). \]

Figure 1 suggests that \( C_{k,n} \leq 8 \), with instances of \( C_{k,n} > 3 \) over the entire range of \( x \)-values. Moreover, the figure is in agreement with Corollary 4.1, suggesting that \( C_{k,n} \) is approximately bounded by 3 as \( x_{k,n} \) becomes large.

We recall that Corollary 4.2 suggests that \( C_{k,n} \) is maximized approximately for \( n = \frac{x_{k,n}^6}{36} \). As an example, \( x_{k,n} = 4 \) gives \( n \approx 113 \) through this relation. The behaviour observed in Figure 1 is in agreement with this. Indeed, a global bowing starting approximately at \( x = 4 \) indicates the insufficiency of \( n \leq 100 \) for obtaining the maximal behaviour of \( C_{k,n} \).

Next, consider the more general family of functions \( x \mapsto C(x) \),

\[ C(x) = x^2 \left( \frac{P(S \geq x)}{N(x)} - 1 \right), \]

for the considered family of linear combinations \( S \). Theorem 3.1 stated that \( C \) is bounded above by some universal constant \( C_1 \). It is thus natural to ask, is the sharpness of this result characteristic for the binomial random variables \( B_n \), or is it a phenomenon of a more global character?

As mentioned previously, there are indications that the results in Section 4 and Figure 1 are intimately connected with the discrete nature of binomial distributions. Now, the random variables \( B_n \) hold a rather extreme position in this sense within the general family of random variables \( S \). Indeed, the support of a non-binomial \( S = \sum_{i=1}^{n} a_i \xi_i \) tends to be more dispersed compared to the support of \( B_n \). Thus, for variables \( S \) with a continuously varying behaviour, it may be questioned whether bounds \( C \leq C_1 \) are equally relevant (for \( x \) large).
For the remaining part of the section, consider, as an example, self–normalized sums

\[ S_n = \sum_{i=1}^{n} X_i / \left( \sum_{i=1}^{n} X_i^2 \right)^{1/2}, \]

with independent \( N(0, 1) \) random variables \( X_1, \ldots, X_n \). Recall that Theorem 3.1 applies to \( S_n \) (cf. Section 1). Also note that \( S_n \) has an absolutely continuous distribution with respect to the Lebesgue measure. Analogously, set

\[ C_n(x) = x^2 \left( \frac{P(S_n \geq x)}{N(x)} - 1 \right). \]

The counterpart of Figure 1, with \( B_n \) replaced by \( S_n \), is given by Figure 2. Note that the scaling for the axes differs in Figure 1 and 2.

Figure 2: Scatterplots of \( C_n(x) = x^2 \left( \frac{P(S_n \geq x)}{N(x)} - 1 \right) \) against \( x \), for \( n = 2, \ldots, 100 \), \( 1 \leq x \leq 3 \), and \( C \geq -1 \).

Figure 2 suggests that \( C_n(x) \leq 0 \), for all \( x \) sufficiently large. Theorem 5.1 confirms this conjecture. Indeed, we prove that \( C_n(x) \leq 0 \) for \( x \geq \sqrt{6} \). In fact, it is likely that \( C_n(x) \leq 0 \) holds already for \( x \geq \sqrt{3} \), as Figure 2 suggests.

**Theorem 5.1.** Consider self–normalized sums \( S_n = \sum_{i=1}^{n} X_i / \left( \sum_{i=1}^{n} X_i^2 \right)^{1/2} \), with \( X_1, \ldots, X_n \) i.i.d. \( N(0, 1) \). Then, for \( x \geq \sqrt{6} \), \( P(S_n \geq x) \leq N(x) \).

**Proof.** To begin with, write

\[ P(S_n \geq x) = P(T_n \geq x \sqrt{(n-1)/(n-x^2)}), \]

for the associated \( t \)–statistic random variables \( T_n \) (cf. [8, Chapter 1]). \( T_n \) has density function \( f_n \) given by (cf. [6, Section II.3]),

\[ f_n(y) = \frac{\Gamma\left(\frac{\nu}{2}\right)}{\sqrt{\pi(n-1)}\Gamma\left(\frac{n-1}{2}\right)} \left( 1 + \frac{y^2}{n-1} \right)^{-\frac{n}{2}}, \quad -\infty < y < \infty. \]
Thus, since
\[
\frac{d}{dx} x \sqrt{(n-1)/(n-x^2)} = \left( \frac{n-1}{n-x^2} \right)^{1/2} \left( 1 + \frac{x^2}{n-x^2} \right) = \left( \frac{n-1}{n} \right)^{1/2} \left( 1 - \frac{x^2}{n} \right)^{-3/2},
\]

\(S_n\) has density function \(g_n\) given by
\[
g_n(x) = \frac{\Gamma \left( \frac{n}{2} \right)}{\sqrt{\pi n} \Gamma \left( \frac{n-1}{2} \right)} \left( 1 - \frac{x^2}{n} \right)^{\frac{n-3}{2}}, \quad -\sqrt{n} < x \leq \sqrt{n}.
\]

(54)

It suffices to verify that, for \(x^2 \geq 6\) and \(n \geq 7\),
\[
g_n(x) \leq (2\pi)^{-1/2} e^{-x^2/2}.
\]

We first note that, for \(x^2 \geq 6\),
\[
\left( 1 - \frac{x^2}{n} \right)^{\frac{n-3}{2}} \leq e^{-x^2/2}.
\]

(55)

Indeed,
\[
\frac{n-3}{2} \log \left( 1 - \frac{x^2}{n} \right) = -\frac{x^2}{2} \frac{n-3}{n} \sum_{k=0}^{\infty} \left( \frac{x^2}{n} \right)^k \frac{1}{k+1} \leq -\frac{x^2}{2} \frac{n-3}{n} \sum_{k=0}^{\infty} \left( \frac{6}{n} \right)^k \frac{1}{k+1}.
\]

Moreover,
\[
\sum_{k=0}^{\infty} \left( \frac{6}{n} \right)^k \frac{1}{k+1} \geq \sum_{k=0}^{\infty} \left( \frac{3}{n} \right)^k = \frac{n}{n-3},
\]

since \(k+1 \leq 2^k\). In view of (54) and (55), it now suffices to verify that
\[
\Gamma \left( \frac{n}{2} \right) \leq \sqrt{\frac{n}{2}} \Gamma \left( \frac{n-1}{2} \right).
\]

(56)

First, assume that \(n\) is odd, and set \(n-1 = 2m\) (for \(m \geq 3\)). Then, by definition of the gamma function,
\[
\frac{\Gamma \left( \frac{n}{2} \right)}{\sqrt{\pi} \Gamma \left( \frac{n-1}{2} \right)} = \frac{(2m)! \sqrt{\pi m}}{(m!)^{2m}} \sqrt{\frac{n-1}{n}}.
\]

By Stirling’s formula (cf. [5, Section II.9]),
\[
\frac{(2m)! \sqrt{\pi m}}{(m!)^{2m}} \leq \exp \left( \frac{1}{24m} - \frac{2}{12m+1} \right) \leq 1,
\]

which verifies that (56) holds.
Next, assume that \( n \) is even, and set \( n = 2(m + 1) \), (for \( m \geq 3 \)). Then,
\[
\frac{\Gamma\left(\frac{n}{2}\right)}{\sqrt{\pi} \Gamma\left(\frac{n-1}{2}\right)} = \frac{(m!)^{2^m \sqrt{\pi m}}}{(2m)! \sqrt{m + 2}}.
\]
By Stirling’s formula,
\[
\frac{(m!)^{2^m}}{(2m)! \sqrt{\pi m}} \leq \exp \left( \frac{2}{12m} - \frac{1}{24m + 1} \right) \leq \exp \left( \frac{1}{6m} \right).
\]
Now,
\[
\exp \left( \frac{1}{6m} \right) - \frac{m + 2}{m} \leq \frac{1}{3m} + \frac{1}{9m} - \frac{2}{m} \leq 0,
\]
which verifies that \( \exp \left( \frac{1}{6m} \sqrt{\frac{m}{m+2}} \right) \leq 1 \), and thus, that (56) holds.

\[\square\]

References


