SOME BOUNDS ON THE DEVIATION PROBABILITY FOR SUMS OF NONNEGATIVE RANDOM VARIABLES USING UPPER POLYNOMIALS, MOMENT AND PROBABILITY GENERATING FUNCTIONS

STEVEN G. FROM

Abstract. We present several new bounds for certain sums of deviation probabilities involving sums of nonnegative random variables. These are based upon upper bounds for the moment generating functions of the sums. We compare these new bounds to those of Maurer [?], Bernstein [?], Pinelis [?], and Bentkus [?]. We also briefly discuss the infinitely divisible distributions case.

1. Introduction

In this paper, we shall use upper polynomial bounds to moment-generating functions to improve an upper bound for a certain deviation probability derived in Pinelis and Utev [?] and rediscovered in Maurer [?]. In some cases, it is also an improvement on a bound given in Bentkus [?].

Theorem 1.1. (Pinelis and Utev [?], Maurer [?]). *Let $X_1, X_2, \ldots, X_m$ be independent random variables with $P(X_i \geq 0) = 1$, $E(X_i^2) < \infty$, $i = 1, 2, \ldots, m$. Let $S = \sum_{i=1}^m X_i$ and suppose $t > 0$. Then*

$$Pr\{E(S) - S \geq t\} \leq \exp\left(\frac{-t^2}{2\sum_{i=1}^{m} E(X_i^2)}\right). \tag{1.1}$$

Corollary 1.2. *Let $X_1, X_2, \ldots, X_m$ be independent random variables with $E(X_i^2) < \infty$, $i = 1, 2, \ldots, m$. Let $S = \sum_{i=1}^m X_i$ and let $t \geq 0$. Let $b_1, b_2, \ldots, b_m$ be real constants with $P(X_i \leq b_i) = 1$, $i = 1, 2, \ldots, m$. Let $\sigma_i^2 = E(X_i^2) - (E(X_i))^2$, $i = 1, 2, \ldots, m$. Then for $t > 0$,*

$$Pr\{S - E(S) \geq t\} \leq \exp\left(\frac{-t^2}{2\sum_{i=1}^{m}\sigma_i^2 + 2\sum_{i=1}^{m}(b_i - E(X_i))^2}\right). \tag{1.2}$$

In [?], the following result is given.

Theorem 1.3. (Bernstein's inequality). *Let $X_i, X_2, \ldots, X_m$ be independent nonnegative random variables. Let $d = \max_i\{E(X_i)\}$. Let $S = \sum_{i=1}^{m} X_i$. Let $\sigma_i^2 = E(X_i^2) - (E(X_i))^2$, $i = 1, 2, \ldots, m$. Then, for $t > 0$*

$$Pr\{E(S) - S \geq t\} \leq \exp\left(\frac{-t^2}{2\sum_{i=1}^{m}\sigma_i^2 + \frac{2td}{3}}\right). \tag{1.3}$$

In [?], Maurer compared the inequalities (**??**) and (**??**), with $d = \max_i(b_i - E(X_i))$. Both results are trivial unless $t < B_1 = \sum_{i=1}^{m}(b_i - E(X_i))$. Let

$$B_2 = \sum_{i=1}^{m}(b_i - E(X_i))^2, \quad B_\infty = \max_i(b_i - E(X_i))\,.$$

Then Bernstein's inequality is stronger if $0 < \epsilon < \frac{3B_2}{B_1 B_\infty}$, where $t = \epsilon B_1$, $0 < \epsilon < 1$. Maurer's inequality is stronger if $\frac{3B_2}{B_1 B_\infty} < \epsilon < 1$. To paraphrase Maurer in [?], Maurer's inequality is stronger when there is a significant range in the b_i deviations, such as $b_i = \frac{1}{i}$ and $m \to \infty$, as shown by Maurer in [?].

Theorem **??** below presents a better bound than the one in Theorem **??**.

Theorem 1.4. ([?], Theorem 2.2). *Under the conditions of Theorem **??**,*

$$Pr\{E(S) - S \geq t\} \leq \exp\left(\frac{-t^2}{2\sum_{i=1}^{m} W_i}\right), \tag{1.4}$$

where $W_i = \max((E(X_i))^2,\ V(X_i))$, and $V(X_i) = E(X_i^2) - (E(X_i))^2$, $i = 1, 2, \ldots, m$.

It will be seen that the new bound of Theorem **??** in the next section always improves on the bound (**??**), as does bound (**??**). However, it will be shown that the new bound of Theorem **??** given by (**??**) is better than (**??**) in some cases.

For related work on bounds for deviation probabilities, see also [?, ?, ?].

In the next section, we shall embed upper bound (??) into a more general family of bounds. In addition, we shall make use of probability generating functions to greatly improve upon the bounds of [?] and [?] in some cases. We shall also show that the bounds (??) and (??) are identical in the limit for infinitely divisible distributions.

2. Main Results

First, we need the following lemma.

Lemma 2.1. *Let $P_{a,b}(t)$ denote the polynomial in t given by*

$$P_{a,b}(t) = 1 - at + bt^2\,. \tag{2.1}$$

Let $\mathcal{P}$ denote the set of all polynomials of the form (??) where $0 \le a \le 1$, $b \ge 0$ and $a^2 + 4b^2 \le 4b$. Then for $t \ge 0$,

$$e^{-t} \le P_{a,b}(t) \tag{2.2}$$

for all $P_{a,b}(t)$ in $\mathcal{P}$. Thus, $\mathcal{P}$ is a collection of upper bound polynomials for e^{-t} on $[0,\infty)$.

Proof. Consider the auxiliary difference function $g(t) = e^{-t} - P_{a,b}(t)$. Then $g(0) = 0$ and $g(\infty) = \lim_{t\to\infty} g(t) = -\infty$, since $a^2+4b^2 \le 4b$ implies $b \ge 0$. Let's show that $g(t) \le 0$ for all $t \ge 0$. The derivative of $g(t)$ is

$$g'(t) = -e^{-t} + a - 2bt\,. \tag{2.3}$$

At any value of t, call it t_0, with $g'(t) = 0$, if any, we have

$$g'(t_0) = -e^{-t_0} + a - 2bt_0 = 0\,. \tag{2.4}$$

Thus,

$$\begin{aligned} g(t_0) &= e^{-t_0} - P_{a,b}(t_0) = a - 2bt_0 - P_{a,b}(t_0) \\ &= a - 2bt_0 - (1 - at_0 + bt_0^2) \\ &= (a-1) + (a-2b)t_0 - bt_0^2\,. \end{aligned}$$

Let $h(u)$ denote the auxiliary function

$$h(u) = (a-1) + (a-2b)u + (-b)u^2, \quad u \ge 0\,.$$

Then $h(0) = a - 1 < 0$ and $h(\infty) = \lim_{u\to\infty} h(u) = -\infty$, since $b \ge 0$. Also, $h'(u) = (a-2b) - 2bu = 0$ when $u = -1 + \frac{a}{2b}$. Since

$$h\left(-1 + \frac{a}{2b}\right) = \frac{4b^2 + a^2 - 4b}{4b} \le 0\,,$$

it follows that $h(u) \le 0$ for $u \ge 0$. So $g(t_0) \le 0$ for any t_0 with $g'(t_0) = 0$. We may therefore conclude that $g(t) \le 0$ for $t \ge 0$, as claimed. This completes the proof. □

Remark 2.2. When $a^2+4b^2=4b$ holds in Lemma **??**, then $b=\frac{1-\sqrt{1-a^2}}{2}$. In particular, $a=1$ gives $b=\frac{1}{2}$ to obtain $e^{-t}\le 1-t+\frac{1}{2}t^2$. This is the first step of Maurer's proof of (**??**) in [**?**]. We shall replace this first step of Maurer's proof by (**??**) above.

Theorem 2.3. *Suppose* $0\le a\le 1$. *Under the conditions of Theorem* **??**, *we have, for* $t\ge 0$,

$$Pr\{E(S)-S\ge t\}\le \exp\left(-\frac{(t-(1-a)E(S))^2}{4b\sum_{i=1}^{m}E(X_i^2)}\right), \tag{2.5}$$

where

$$\frac{E(S)-t}{E(S)}\le a\le 1, \quad \textit{and} \quad b=\frac{1-\sqrt{1-a^2}}{2}.$$

Proof. We shall proceed as in the proof of (**??**) given in Theorem 2.1 of Maurer in [**?**], except we shall apply Lemma **??** above. Let $\beta\ge 0$. We obtain

$$\begin{aligned} E(e^{-\beta X_i}) &\le 1-a\beta E(X_i)+b\beta^2E(X_i^2)\\ &\le \exp(-a\beta E(X_i)+b\beta^2E(X_i^2)) \end{aligned}$$

which gives

$$E(e^{-\beta S})\le \exp\left(-a\beta E(S)+b\beta^2\sum_{i=1}^{m}E(X_i^2)\right)$$

and

$$Ln(E(e^{-\beta S}))\le -a\beta E(S)+b\beta^2\sum_{i=1}^{m}E(X_i^2)\,.$$

Let χ be the characteristic function of $[0,\infty)$, as done in [**?**]. Then

$$\begin{aligned} Ln(Pr\{E(S)-S\ge t\}) &= Ln(E[\chi(-t+E(S)-S)])\\ &\le LnE[\exp(\beta(-t+E(S)-S))]\\ &\le -\beta t+\beta E(S)+Ln(E(e^{-\beta S}))\\ &\le -\beta t+\beta E(S)-a\beta E(S)+b\beta^2\sum_{i=1}^{m}E(X_i^2)\\ &= -\beta t+(1-a)\beta E(S)+b\beta^2\sum_{i=1}^{m}E(X_i^2)\,. \end{aligned} \tag{2.6}$$

Expression (**??**) is minimized by

$$\beta = \frac{t-(1-a)E(S)}{2b\sum_{i=1}^{m} E(X_i^2)} . \tag{2.7}$$

Since $a \geq \frac{E(S)-t}{E(S)}$, the choice for β in (**??**) ensures $\beta \geq 0$. Substitution of (**??**) into (**??**) completes the proof, upon exponentiation. □

Remark 2.4. The choice of a in Theorem **??** will now be discussed. The choice $a = 1$ and $b = 1/2$ gives inequality (**??**), for

$$\beta = \frac{t-(1-a)E(S)}{2b\sum_{i=1}^{m} E(X_i^2)} = \frac{t}{\sum_{i=1}^{m} E(X_i^2)} .$$

Let $q = E(S)$. Straightforward, but messy differentiation of (**??**) above with respect to a using (**??**) and solving for a eventually gives the equation

$$(t-q+aq)\cdot(2q\sqrt{1-a^2}-2q+qa^2-at+aq) = 0 . \tag{2.8}$$

By applying Tartaglia's method for solving the quartic equation equivalent of the second factor in (**??**) for the unknown a, we arrive at the following simple procedure to calculate the three roots of (**??**). Let $q = E(S)$. Then (**??**) has the three real roots $a = a_i^*$, $i = 1, 2, 3$: $a = a_1^* = \frac{q-t}{q}$, $a = a_2^* = z_3 + z_6$, and $a = a_3^* = (a = a_3^* = 0)$, where $z_3 = \frac{2(t-q)}{3q}$, $z_6 = \frac{z_5}{3q} + \frac{z_4}{3qz_5}$, $z_5 = (z_1 + 6\sqrt{3}z_2)^{1/3}$, $z_1 = -t^3 + 3t^2q - 57tq^2 + 55q^3$, $z_2 = q\sqrt{t^4 - 4t^3q + 33t^2q^2 - 58tq^3 + 28q^4}$, and $z_4 = t^2 - 2tq + q^2$. Note that all quantities exist as real numbers since for $t < E(S) = q$, the condition corresponding to a nontrivial bound, it is guaranteed that the radicand of z_2 is nonnegative, since the sum of coefficients of this radicand is $1 - 4 + 33 - 58 + 28 = 0$. It is easily shown that the root $a = a_1^*$ corresponds to $\beta = 0$ and a maximum upper bound of 1.0, so this can be eliminated from consideration when attempting to minimize the upper bound of Theorem **??**. Since $a_3^* = 0 < a_1^*$, it can be eliminated from consideration as a minimizer of (**??**). Two of the roots of (**??**) can be shown to be the strictly complex-values:

$$\left(\frac{-z_5}{6q} - \frac{(t-q)^2}{6qz_5} + z_3\right) \pm \frac{z_5}{6q}\sqrt{3}I ,$$

where $I = \sqrt{-1}$.

Theorem 2.5. *The choice $a = a_2^*$ given in Remark* **??** *satisfies $a_1^* \leq a_2^* < 1$. Also, it minimizes the bound (**??**). In particular, bound (**??**) is less than bound (**??**).*

Proof. For $a_1^* \le a \le 1$, let $g(a)$ denote the natural logarithm of bound (**??**) considered as a function of a, using $b = \frac{1-\sqrt{1-a^2}}{2}$. Let's show that $g(a)$ is decreasing on some right neighborhood of a_1^* and that $g(a)$ is increasing on some left neighborhood of $a = 1$. This will give the existence of a value of a, call it $\hat{a}_2$ in $[a_1^*, 1]$, such that $g(a)$ has a relative minimum at $a = \hat{a}_2$ on the interval $a_1^* \le a \le 1$, with $g'(\hat{a}_2) = 0$. By Remark **??** discussion, and since $\hat{a}_2$ cannot equal either of a_1^* or 1, we must have $\hat{a}_2 = a_2^*$. In this case, the absolute minimum of $g(a)$ on $[a_1^*, 1]$ must occur at $a = a_2^*$, which would prove $a_1^* \le a_2^* \le 1$, as desired.

To this end, let's first show that $g(a)$ is decreasing on some right neighborhood of a_1^*. Let $w = \sum_{i=1}^m E(X_i^2)$, $q = E(S)$. Then $g(a)$ has derivatives:

$$g'(a) = \frac{(t-q+qa)q}{w(-1+\sqrt{1-a^2})} + \frac{(t-q+qa)^2 a}{2(-1+\sqrt{1-a^2})^2 w\sqrt{1-a^2}}$$

and

$$g''(a) = \frac{q^2}{w(-1+\sqrt{1-a^2})} + \frac{2(t-q+qa)qa}{(-1+\sqrt{1-a^2})^2 w\sqrt{1-a^2}} + \frac{a^2(t-q+qa)^2}{(-1+\sqrt{1-a^2})^2 w(1-a^2)}\,.$$

Thus, we have that $g'(a_1^*) = 0$ and

$$g''(a_1^*) = \frac{q^2}{w(-1+\sqrt{1-(a_1^*)^2})} < 0\,.$$

So $g(a)$ is decreasing on some right neighborhood of a_1^*. This is also true for bound (**??**), upon exponentiation.

Now let's demonstrate that $g(a)$ is increasing on some left neighborhood of $a = 1$. Then $\lim_{a\to 1^-} g'(a) = +\infty$, so $g(a)$ and bound (**??**) are both increasing to the left of $a = 1$, as claimed. This completes the proof. □

We now compare bounds (**??**) and (**??**). It will be seen that neither bound is uniformly better than the other bound.

Theorem 2.6. *Let $q = E(S)$. Then there exist positive real numbers δ_1 and δ_2 such that $\delta_1 + \delta_2 < q$ and:*

a) *for all t in $(q-\delta_1, q)$, bound (**??**) is smaller than bound (**??**)*
b) *for all t in $(0, \delta_2)$, bound (**??**) is smaller than bound (**??**).*

Proof. Let $\sigma_i^2 = E(X_i^2) - (E(X_i))^2$, $i = 1, 2, \ldots, m$. Subtraction of the natural logarithms of bounds (**??**) and (**??**) gives that bound (**??**) is smaller than bound (**??**) if and only if

$$h(t) = \frac{(t-(1-a)q)^2}{2bt^2} \geq \frac{\displaystyle\sum_{i=1}^{m} E(X_i^2)}{\displaystyle\sum_{i=1}^{m} \max((E(X_i))^2, \sigma_i^2)} \equiv \theta\,.$$

It is easily shown that $1 < \theta < 2$ holds. Now $a = a_2^*$ and $b = \frac{1-\sqrt{1-a^2}}{2}$ are functions of t and q, so using formulas given in Remark **??** we obtain: $\lim_{t\to q^-} h(t) \equiv 2$, for all q. Thus, $\lim_{t\to q^-}(h(t)-\theta) > 0$ and part (a) of the result follows from continuity of $h(t)$ near $t = q$. Similarly, $\lim_{t\to 0^+} h(t) = 1$, using L'Hospital's Rule, since $a = a_2^*$ approaches one as t approaches 0^+. Thus, $\lim_{t\to 0^+}(h(t)-\theta) < 0$ and part (b) is proven. □

Remark 2.7. The new bound (**??**) can be shown to be better than Bernstein's bound (**??**) near $t = q$ if $r(q) < 2$, where

$$r(t) = \frac{\displaystyle\sum_{i=1}^{m} E(X_i^2)}{\left[\displaystyle\sum_{i=1}^{m} \sigma_i^2 + \frac{t\max\{E(X_i)\}}{3}\right]}\,.$$

Since $r(0) > 1$, for t near zero, (**??**) is better than (**??**). The proof of this is very similar to the proof of Theorem **??** above and is omitted.

Besides the new bounds considered thus far, other bounds have also been considered by using upper bounds for moment-generating functions and probability generating functions (instead of Lemma 2.1) given in [**?**], [**?**], and [**?**]. Although these bounds are usually better than the new bounds discussed thus far, they sometimes give only the trivial upper bound 1. Moreover, the choice of the corresponding β parameter is only of closed form in the special case of IID (independent and identically distributed) random variables, and is not easy to determine in general. Research is ongoing in this case.

3. Numerical Study

In this section, we compute the actual deviation probability and upper bounds for the deviation probability of interest in this paper. We assume that $X_1, X_2, \ldots, X_m$ have distributions that are members of the Weibull distribution with cumulative distribution function (cdf) ($X = X_i$ for some *i)*

$$P(X \le x) = 1 - \exp\left(-\left(\frac{x}{\theta_2}\right)^{\theta_1}\right), \ x \ge 0, \ \theta_1 > 0, \ \theta_2 > 0\,. \qquad (3.1)$$

Table 3.1 below gives numerical values of the upper bounds discussed in this paper for family (**??**). First, we describe notations for these bounds.

Let U_P denote the upper bound of Pinelis and Utev given by (**??**) in Theorem **??**. Let U_B denote the upper bound based upon the Bernstein inequality, (**??**) of Theorem **??**. Let U_V = upper bound (**??**) discussed by V. Bentkus. These are previously proposed bounds.

The new bound is: U_1 = upper bound (**??**) of Theorem **??** (upon exponentiation) using $a = a_2^*$ given in Remark **??**.

Although the new bound (**??**) with $a = a_2^*$ is harder to compute than any of the old bounds, the new bound is the best for t not too far from $q = E(S)$. Table 3.1 below gives the bounds for various choices of θ_1 and θ_2 for each X_i random variable. The value of p is the approximated value of the left-hand side of deviation probability (**??**) and was found using one million runs of Monte Carlo simulation.

Weibull Family (3.1)

(a) $m = 5$, θ_1 values: 0.5, 0.75, 1.0, 2.0, 3.0 ($\theta_2 = 1$ for all θ_1).

t	p	U_P	U_B	U_V	U_1
5.0	0.0007	0.676	0.632	0.609	0.566
3.0	0.1906	0.868	0.841	0.683	0.793
0.1	0.6489	0.999524	0.999792	0.999048	0.999520

For $t = 5.0$, $a_2^* = 0.76$.

The bound U_V is best and is better than U_1 for $0 \le t \le 3.86$.

The new bound U_1 is best and is better than U_V for $3.86 \le t \le 5.97$, where $q = E(S) = 5.97$. This is in line with Theorem **??**.

(b) $m = 3$, $(\theta_1, \theta_2) = (2,1), (3, \frac{1}{2}), (5,1)$, $q = E(S) = 2.25$.

t	p	U_P	U_B	U_V	U_1
2.00	0	0.388	0.108	0.335	0.234
1.50	0.0001	0.587	0.221	0.540	0.505
0.50	0.2219	0.943	0.752	0.934	0.938

For $t = 2.00$, $a_2^* = 0.69$. The Bernstein bound U_B is best. Comparing U_1 to U_V, we see that

U_1 is better than U_V for $1.04 \le t \le 2.25 = E(S)$.

U_V is better than U_1 if $0 \leq t \leq 1.04$, in agreement with Theorem **??**.

Table 3.1

Many other parametric families were considered besides (**??**) above. Also considered were the beta binomial and logarithmic series distributions discussed in Johnson, Kotz and Kemp [**?**], the gamma, beta, binomial, and a generalization of the Poisson distribution given in [**?**]. Although we do not give numerical comparisons for these families, the conclusions reached are the same as those reached in the Weibull family (**??**) case. They are:

(1) For small t ($t << E$), $U_P \approx U_1$ with $U_1 < U_P$. For $t < E$ but not much smaller than t, U_1 significantly improves on U_P.
(2) The closer t is to $q = E(S)$ with $t < q$, the better the new bound U_1 does compared to U_V and U_P. In this case, $U_1 < U_V < U_P$. If t is substantially smaller than q, however, we have $U_V < U_1 < U_P$. As t approaches 0, U_P and U_1 are virtually indistinguishable.
(3) The old bounds U_P, U_B, and U_V do have the advantage of being simplest to compute, however.

It should be mentioned that also considered in the numerical comparisons were the upper bounds for tail probabilities using the standard normal cumulative distribution function with $c_0 = \frac{2e^3}{9}$ discussed in [**?**], [**?**], [**?**], and [**?**]. In most cases, this provides bounds which are larger than all bounds discussed in this paper. For example, in Table 3.1, part (a) above, this upper bound was 0.712 for $t = 5.0$ and 1.879 for $t = 1.0$.

4. Infinitely Divisible Cases

In this section, we point out some connections between the bounds of (**??**) and (**??**) given in Section 1. We also discuss how some of the new bounds given in Sections 1 and 2 can be improved if it is known that $X_1, X_2, \ldots, X_m$ have moment or probability generating functions (mgf and pgf, respectively) that are of the infinite divisibility type.

Definition 4.1. *Let X be a random variable with characteristic function $\varphi(u) = E(e^{IuX})$, where $I = \sqrt{-1}$. Then X is infinitely divisible if for every positive integer n, $\varphi(u)$ has the form*

$$\varphi(u) = [\varphi_n(u)]^n \,, \tag{4.1}$$

*where $\varphi_n(u)$ is itself a characteristic function. If X has a moment-generating function (mgf) or probability generating function (pgf), then (**??**) holds with $\varphi(u)$ denoting either the mgf or pgf. It is well-known that in these latter two cases, $\varphi(u)$ must have the form*

$$\varphi(u) = \exp(-\lambda[1 - h(u)]), \ \text{where } 0 < \lambda < \infty$$

and $h(u)$ is itself a pgf. For a discussion of infinitely divisible random variables, see Feller [**?**] *or Johnson, Kotz, and Kemp* [**?**], *which contains many more references. Infinitely divisible distributions are important in models used in mathematical science areas, including biology, chemistry, physics, economics, and queueing theory.*

Next, we point out a connection between Maurer's inequality given in Theorem **??** and Bernstein's Inequality given Theorem **??**. First, we need Theorems **??** and **??**.

Theorem 4.2. (Inequality (**??**) for infinitely divisible case.) *Suppose $X_1, X_2, \ldots, X_m$ are nonnegative random variables. Suppose that the mgfs of $X_1, \ldots, X_L$ are infinitely divisible and that the mgfs of $X_{L+1}, \ldots, X_m$ are not, where $0 \leq L \leq m$. If $\sigma_i^2 < \infty$, where $\sigma_i^2 = E(X_i^2) - \mu_i^2$, $\mu_i = E(X_i)$, $i = 1, 2, \ldots, m$, then*

$$Pr\{E(S) - S \geq t\} \leq \exp\left[\frac{-t^2}{2\left(\displaystyle\sum_{i=1}^{L}\sigma_i^2 + \sum_{j=L+1}^{m} E(X_j^2)\right)}\right], \; t \geq 0\,. \quad (4.2)$$

Proof. Since X_i is infinitely divisible, $0 \leq i \leq L$, there exist random variables $\{W_{ij}\}$, $i = 1, \ldots, L, j = 1, \ldots, N$ such that

$$X_i \stackrel{\text{st}}{=} \sum_{j=1}^{N} W_{ij}\,,$$

for any given positive integer N, no matter how large, where $\stackrel{\text{st}}{=}$ denotes stochastically equal. Moreover, $\{W_{i1}, W_{i2}, \ldots, W_{iN}\}$ are IID random variables, $i = 1, 2, \ldots, L$. Thus

$$E(W_{ij}^2) = \frac{NE(X_i^2) + \mu_i^2(1-N)}{N^2}, \; 1 \leq i \leq L, \; 1 \leq j \leq N\,.$$

Applying Theorem **??** to the double array $\{W_{ij}\}$ along with $X_{L+1}, \ldots, X_m$, and using

$$\sum_{i=1}^{m} X_i = \sum_{i=1}^{L}\sum_{j=1}^{N} W_{ij} + \sum_{i=L+1}^{m} X_i\,,$$

we obtain

$$Pr\{E(S)-S \geq t\} \leq \exp\left(\frac{-t^2}{2\left[\sum_{i=1}^{L}\left(E(X_i^2)+\left(\frac{1-N}{N}\right)\mu_i^2\right)+\sum_{j=L+1}^{m}E(X_j^2)\right]}\right).$$

Now let $N \to \infty$. Then

$$Pr\{E(S) - S \geq t\} \leq \exp\left(\frac{-t^2}{2\left[\sum_{i=1}^{L}\sigma_i^2+\sum_{j=L+1}^{m}E(X_j^2)\right]}\right)$$

as desired. This completes the proof. $\square$

Theorem 4.3. (Bernstein's inequality for the infinitely divisible case). *Under the conditions of Theorem* **??**, *for* $t \geq 0$,

$$Pr\{E(S) - S \geq t\} \leq \exp\left[\frac{-t^2}{\left(2\sum_{i=1}^{L}\sigma_i^2\right)+\left(2\sum_{j=L+1}^{m}\sigma_j^2+\frac{2tD}{3}\right)}\right], \quad (4.3)$$

where

$$D = \max_{L+1\leq j\leq m} E(X_j)\,.$$

Proof. We construct the $\{W_{ij}\}_{j=1,\ldots,N}^{i=1,\ldots,L}$ double array of random variables as done in the proof of Theorem **??**. Using

$$E(W_{ij}) = \frac{E(X_i)}{N},\ j = 1, \ldots, N\,,$$

we obtain

$$d_N = \max_{\substack{1\leq i\leq L\\ 1\leq j\leq N}}[E(W_{ij})] = \frac{1}{N}\max_{1\leq i\leq L} E(X_i)\,.$$

Let $N \to \infty$. Then $d_N \to 0$ as $N \to \infty$, giving

$$d = \max_{1\leq i\leq m} E(X_i) \to \max_{L+1\leq j\leq m} E(X_j) \text{ as } N \to \infty\,.$$

This completes the proof, upon application of Theorem **??**. $\square$

Corollary 4.4. *Suppose $X_1, X_2, \dots, X_m$ are nonnegative infinitely divisible random variables with $E(X_i^2) < \infty$, $i = 1, 2, \dots, m$. Then Theorems* **??** *and* **??** *give the same upper bound for the deviation probability given by*

$$Pr\{E(S) - S \geq t\} \leq \exp\left(\frac{-t^2}{2\sum_{i=1}^{m} \sigma_i^2}\right), \ t \geq 0\,. \tag{4.4}$$

Proof. Set $L = m$ in Theorem **??**, with impossible sums interpreted as zero. Also, $D = 0$ in the proof of Theorem **??**, if $L = m$. Then the rhs values of (**??**) and (**??**) are both equal to

$$\exp\left(\frac{-t^2}{2\sum_{i=1}^{m} \sigma_i^2}\right).$$

□

Remark 4.5. Corollary **??** states that inequalities (**??**) and (**??**) are the same for infinitely divisible distributions 'in the limit'.

Next, we indicate how the bound given in Theorem **??** can be improved in the case where $X_1, \dots, X_L$ are infinitely divisible.

Theorem 4.6. *Under the conditions of Theorem* **??**, *we have, for $t \geq 0$,*

$$Pr\{E(S) - S \geq t\} \leq \exp\left(-\frac{(t-(1-a)E(S))^2}{4bW}\right), \tag{4.5}$$

where

$$b = \frac{1-\sqrt{1-a^2}}{2} \ \text{and} \ W = \sum_{i=1}^{L} \sigma_i^2 + \sum_{j=L+1}^{m} E(X_j^2)\,.$$

Proof. Let $w = \sum_{i=1}^{L} \sigma_i^2 + \sum_{j=L+1}^{m} E(X_j^2)$, $R = t - (1-a)E(S)$. The proof of Theorem **??** can be followed to give $\beta = \frac{R}{2bw}$ giving

$$\begin{aligned} -\beta t + (1-a)\beta E(S) + b\beta^2 w &= \frac{-R^2}{2bw} + (1-a)\left(\frac{R}{2bw}\right)E(S) + b\left(\frac{R}{2bw}\right)^2 w \\ &= \frac{-2R^2 b + 2Rb(1-a)E(S) + bR^2}{4b^2 w}, \end{aligned}$$

which is nondecreasing in w if we can show that

$$-2R^2 b + 2Rb(1-a)E(S) + bR^2 \leq 0\,. \tag{4.6}$$

Now $R \geq 0$ since $\beta \geq 0$ is required. So $t \geq (1-a)E(S)$ gives $-R \leq -2(1-a)E(S)$. Multiplication by $bR \geq 0$ gives $-bR^2+2bR(1-a)E(S) \leq 0$. Thus, $-2R^2b+2Rb(1-a)E(S)+bR^2 = -bR^2+2br(1-a)E(S) \leq 0$. Thus, expression (**??**) is nonpositive. The rest of the proof proceeds as in the proof of Theorem **??** and is omitted. □

The numerical comparisons and conclusions given earlier in Section 3 still hold in the infinitely divisible case, (with all bounds smaller, of course!)

The author is currently investigating the application of the methods given in this paper and in [**?**] and [**?**] to probability inequalities of the type given in [**?**], [**?**], and [**?**].

References

[1] G. Bennett, *Probability inequalities for the sum of independent random variables*, J. Amer. Statist. Assoc., **57** (1962), 33–45.

[2] V. Bentkus, N. Kalosha, and M. Van Zuijlen, *On domination of tail probabilities of (super) martingales: explicit bounds*, Lithuanian Mathematical Journal, **46** (1) (2003), 1–43.

[3] V. Bentkus, *An inequality for tail probabilities of martingales with differences bounded from one side*, Journal of Theoretical Probability, **16** (1) (2003), 161–173.

[4] S. Bernstein, *Theory of Probability*, Moscow, 1927.

[5] H. Braun, *Polynomial bounds for probability generating functions*, J. Appl. Prob., **12** (1975), 507–514.

[6] D. Brook, *Bounds for moment-generating functions and for extinction probabilities*, J. Appl. Prob., **3** (1966), 171–178.

[7] A. Dembo and O. Zeitoumi, *Large deviation techniques and applications*, Springer-Verlag, 1998.

[8] W. Feller, *An introduction to probability theory and its applications*, Vol. 1, 3rd ed., John Wiley and Sons, New York, 1967.

[9] S. From, *Approximating the distribution of a renewal process using generalized Poisson distributions*, J. Statist. Comp. and Sim., **74** (9) (2004), 667–681.

[10] W. Hoeffding, *Probability inequalities for sums of bounded random variables*, J. Amer. Statist. Assoc., **58** (1963), 13–30.

[11] N. L. Johnson, S. Kotz, and A. W. Kemp, *Univariate discrete distributions*, 2nd ed., Wiley, New York, 1992.

[12] A. Maurer, *A bound on the deviation probability for sums of nonnegative random variables*, Journal of Inequalities in Pure and Applied Mathematics (JIPAM), **4** (1) (2003), Article 15.

[13] P. Narayan, *On bounds for probability generating functions*, Austral. J. Statist., **23** (1) (1981), 80–90.

[14] I. Pinelis, *Toward the best constant factor for the Rademacher-Gaussian tail comparison*, 2006 preprint, Vilnius Institute of Mathematics and Informatics.

[15] I. Pinelis, *On Normal Domination of (Super) Martingales*, Electronic J. Prob., **11**, Ms. 39 (2006), 1049–1070.

[16] I. S. Pinelis and S. A. Utev, *Exact Exponential Bounds for Sums of Independent Random Variables*, Theory Probab. Appl. **34** (1989), 340–346.

Department of Mathematics, University of Nebraska at Omaha, Omaha, Nebraska 68182-0243

E-mail address: sfrom@mail.unomaha.edu or sfrom@unomaha.edu