EXACT LOWER BOUNDS ON THE EXPONENTIAL MOMENTS OF TRUNCATED RANDOM VARIABLES

IOSIF PINELIS,* *Michigan Technological University*

Abstract

Exact lower bounds on the exponential moments of $\min(y, X)$ and $X\,\mathbf{1}\{X < y\}$ are provided given the first two moments of a random variable X. These bounds are useful in work on large deviation probabilities and nonuniform Berry–Esseen bounds, when the Cramér tilt transform may be employed. Asymptotic properties of these lower bounds are presented. Comparative advantages of the so-called Winsorization $\min(y, X)$ over the truncation $X\,\mathbf{1}\{X < y\}$ are demonstrated. An application to option pricing is given.

Keywords: Exponential moments; exact lower bounds; Winsorization; truncation; large deviations; nonuniform Berry–Esseen bounds; Cramér tilt transform; option pricing

2010 Mathematics Subject Classification: Primary 60E15
Secondary 60E10; 60F10; 60F05

1. Introduction

Cramér's tilt transform of a random variable (RV) X is an RV X_c such that

$$\mathrm{E}\, f(X_c) = \frac{\mathrm{E}\, f(X)\mathrm{e}^{cX}}{\mathrm{E}\, \mathrm{e}^{cX}} \tag{1.1}$$

for all nonnegative Borel functions f, where c is a real parameter. This transform is an important tool in the theory of large deviation probabilities $\mathrm{P}(X > x)$, where $x > 0$ is a large number; then the appropriate value of the parameter c is positive. As, e.g. in the proof of [23, Theorem 2.3], one often needs to bound from above the f-moment $\mathrm{E}\, f(X_c)$ of the c-tilted RV X_c for a nonnegative f—and, therefore, we need to bound the denominator $\mathrm{E}\, \mathrm{e}^{cX}$ in (1.1) from below. If $\mathrm{E}\, X = 0$, this can be done quite easily: by Jensen's inequality, $\mathrm{E}\, \mathrm{e}^{cX} \geq 1$.

A usual problem with this approach occurs when the right tail of X is too heavy for $\mathrm{E}\, \mathrm{e}^{cX}$ to be finite and, hence, for the transform to make sense. The standard cure in such situations is to truncate the RV X, say to $T_y(X) := X\,\mathbf{1}\{X \leq y\}$ for some real number $y > 0$, where $\mathbf{1}\{\cdot\}$ is the indicator function. Then, of course, $\mathrm{E}\, \mathrm{e}^{cT_y(X)} < \infty$ for any $c > 0$. However, now instead of the condition $\mathrm{E}\, X = 0$ we have $\mathrm{E}\, T_y(X) \leq 0$, and the inequality $\mathrm{E}\, \mathrm{e}^{cT_y(X)} \geq 1$ (in place of $\mathrm{E}\, \mathrm{e}^{cX} \geq 1$) will not hold in general. In fact, $\mathrm{E}\, \mathrm{e}^{cT_y(X)}$ can be however small for some $c > 0$, even if one imposes a restriction such as $\mathrm{E}\, X^2 \leq \sigma^2$ for a given real $\sigma > 0$; see the discussion in Subsection 2.3.

A much better way to cut off the right tail of the distribution of X is the so-called Winsorization. That is, instead of the truncation $T_y(X)$, we deal with $W_y(X) := y \wedge X = \min(y, X)$. Clearly, $W_y(X) \geq T_y(X)$ and, hence, $\mathrm{E}\, \mathrm{e}^{cW_y(X)} \geq \mathrm{E}\, \mathrm{e}^{cT_y(X)}$ for $c > 0$. Moreover, it turns out that, for any given real $\sigma > 0$ and $y > 0$, the infimum of $\mathrm{E}\, \mathrm{e}^{cW_y(X)}$ over all $c > 0$ and all RVs

X with $\mathrm{E}\,X \geq 0$ and $\mathrm{E}\,X^2 \leq \sigma^2$ is strictly positive; furthermore, it decreases slowly from 1 to 0 as σ increases from 0 to ∞. These properties of Winsorization make it a clear winner over truncation in many relevant situations.

2. Results

Take any real $\sigma > 0$. Let X denote any RV with

$$\mathrm{E}\,X \geq 0 \quad \text{and} \quad \mathrm{E}\,X^2 \leq \sigma^2.$$

For any positive real a and b, let $X_{a,b}$ stand for any zero-mean RV with values in the two-point set $\{-a, b\}$; thus, the distribution of $X_{a,b}$ is uniquely determined by a and b. Note also that $\mathrm{E}\,X_{a,b}^2 = ab$.

2.1. Winsorization

Consider the Winsorization function defined by the formula

$$W(x) := 1 \wedge x. \tag{2.1}$$

The following proposition allows us to define the terms in which to express the exact lower bounds on $\mathrm{E}\,e^{cW(X)}$.

Proposition 2.1. *Take any real $c > 0$.*

(i) *For any real a, let*

$$b_{a,c}^* := \frac{2(e^{c+ac} - 1) - ac}{c}. \tag{2.2}$$

Then there exists a unique $a_{c,\sigma}$ such that

$$\{a_{c,\sigma}\} = \{a > 0\colon ab_{a,c}^* = \sigma^2\}. \tag{2.3}$$

(ii) *The expression*

$$\ell_1(a) := \ell_1(a, \sigma) := \ln\frac{a}{\sigma^2} - \frac{2(a+1)(a - \sigma^2)}{a^2 + \sigma^2} \tag{2.4}$$

changes sign exactly once, from negative to positive, as a increases from 0 to σ^2. Therefore, we can uniquely define a_σ by the formula

$$\{a_\sigma\} = \{a \in (0, \sigma^2)\colon \ell_1(a) = 0\}. \tag{2.5}$$

The proofs are deferred to Section 4.

Now we are ready to define three more symbols:

$$b_{c,\sigma} := \frac{\sigma^2}{a_{c,\sigma}}, \tag{2.6}$$

$$b_\sigma := \frac{\sigma^2}{a_\sigma}, \tag{2.7}$$

$$c_\sigma := \frac{\ln b_\sigma}{1 + a_\sigma}. \tag{2.8}$$

Theorem 2.1. *For any real $c > 0$,*

$$\mathrm{E}\exp\{cW(X)\} \geq L_{W;c,\sigma} := \mathrm{E}\exp\{cW(X_{a_{c,\sigma},b_{c,\sigma}})\} \tag{2.9}$$

$$\geq L_{W;\sigma} := \mathrm{E}\exp\{c_\sigma W(X_{a_\sigma,b_\sigma})\}. \tag{2.10}$$

Moreover, inequality (2.9) is strict unless $X \overset{\mathrm{D}}{=} X_{a_{c,\sigma},b_{c,\sigma}}$, where '$\overset{\mathrm{D}}{=}$' denotes equality in distribution, and inequality (2.10) is strict unless $c = c_\sigma$. Furthermore,

$$a_\sigma = a_{c_\sigma,\sigma} \quad \text{and} \quad b_\sigma = b_{c_\sigma,\sigma} = b^*_{a_\sigma,c_\sigma}, \tag{2.11}$$

so that (2.10) turns into an equality if and only if $c = c_\sigma$.

In addition to being zero mean, each of the RVs $X_{a_{c,\sigma},b_{c,\sigma}}$ and X_{a_σ,b_σ} has variance σ^2, in view of (2.6) and (2.7). Moreover, by (2.5) and (2.7), $b_\sigma > 1$ and, hence, $c_\sigma > 0$ by (2.8). Thus, (2.9) provides an exact lower bound on $\mathrm{E}\exp\{cW(X)\}$ for a fixed $c > 0$, while (2.10) provides an exact lower bound on $\mathrm{E}\exp\{cW(X)\}$ over all $c > 0$.

Let us now describe the asymptotics of the bounds $L_{W;c,\sigma}$ and $L_{W;\sigma}$ for $\sigma \downarrow 0$ and $\sigma \to \infty$. As usual, we write $a \sim b$ if $a/b \to 1$.

Proposition 2.2. (i) *For any real $c > 0$,*

$$L_{W;c,\sigma} - 1 \sim \frac{-c^2}{4(\mathrm{e}^c - 1)}\sigma^2 \quad \text{as } \sigma \downarrow 0, \tag{2.12}$$

$$L_{W;c,\sigma} \sim \frac{4\mathrm{e}^c}{c^2}\frac{\ln^2\sigma}{\sigma^2} \quad \text{as } \sigma \to \infty. \tag{2.13}$$

(ii) *The expression*

$$f(t) := \ln t + 2(1 - t) \tag{2.14}$$

changes sign exactly once, from negative to positive, as t increases from 0 to 1. Therefore, we can uniquely define t_* by the formula

$$\{t_*\} = \{t \in (0, 1) \colon f(t) = 0\}; \tag{2.15}$$

in fact, $t_* = 0.203\ldots$.

(iii) *We have*

$$L_{W;\sigma} - 1 \sim -(1 - t_*)t_*\sigma^2 \quad \text{as } \sigma \downarrow 0, \tag{2.16}$$

$$L_{W;\sigma} \sim \mathrm{e}^2\frac{\ln^2\sigma}{\sigma^2} \quad \text{as } \sigma \to \infty. \tag{2.17}$$

(iv) *To compare (2.12) with (2.16), and (2.13) with (2.17), we observe that*

$$\inf_{c>0} \frac{-c^2}{4(\mathrm{e}^c - 1)} = -(1 - t_*)t_* \quad \text{(attained at } c = -\ln t_* = 1.593\ldots), \tag{2.18}$$

$$\inf_{c>0} \frac{4\mathrm{e}^c}{c^2} = \mathrm{e}^2 \quad \text{(attained at } c = 2). \tag{2.19}$$

Thus, the asymptotic expression in (2.16) is the minimum in $c > 0$ of that in (2.12), and the asymptotic expression in (2.17) is the minimum in $c > 0$ of that in (2.13).

Note that the convergence for $\sigma \to \infty$ in Proposition 2.2 is very slow. For example, the ratio $e^2(\ln^2 \sigma/\sigma^2)/L_{W;\sigma}$ of the terms in (2.17) is $1.201\ldots$ for σ as large as 10^{10}.

The relations $a_\sigma \sim t_* \sigma^2$ as $\sigma \downarrow 0$ and $a_\sigma \sim \frac{1}{2}\ln(\sigma^2)$ as $\sigma \to \infty$, to be established in the proof of part (iii) of Proposition 2.2, suggest, and numerical calculations confirm, that a good initial approximation for solving the equation $\ell_1(a) = 0$ in (2.5) for a_σ is $a = \frac{1}{2}\ln(1 + 2t_*\sigma^2) \approx \frac{1}{2}\ln(1 + 0.406\sigma^2)$.

2.2. Truncation

Consider the truncation function T defined by the formula

$$T(x) := x\,\mathbf{1}\{x < 1\}. \tag{2.20}$$

The following proposition allows us to define the terms in which to express the exact lower bounds on $\mathrm{E}\,e^{cT(X)}$.

Proposition 2.3. *Take any real $c > 0$. For any real a, let*

$$B_{a,c}^* := \frac{2(e^{ac} - 1) - ac}{c};$$

cf. (2.2). Then we can uniquely define A_c and $A_{c,\sigma}$ by the formulae

$$\{A_c\} = \{a > 0 \colon B_{a,c}^* = 1\}, \tag{2.21}$$

$$\{A_{c,\sigma}\} = \{a > 0 \colon a B_{a,c}^* = \sigma^2\}, \tag{2.22}$$

because each of the equations on the right-hand sides of (2.21) and (2.22) has a unique root $a > 0$. Moreover, we have the implication

$$A_c \le \sigma^2 \quad \Longrightarrow \quad A_{c,\sigma} \ge A_c. \tag{2.23}$$

We shall need one more definition:

$$B_{c,\sigma} := \frac{\sigma^2}{A_{c,\sigma}}.$$

Theorem 2.2. *For any real $c > 0$,*

$$\mathrm{E}\exp\{cT(X)\} \ge L_{T;c,\sigma} := \begin{cases} \mathrm{E}\exp\{cT(X_{\sigma^2,1})\} & \text{if } \sigma^2 \le A_c, \\ \mathrm{E}\exp\{cT(X_{A_{c,\sigma},B_{c,\sigma}})\} & \text{if } \sigma^2 \ge A_c. \end{cases} \tag{2.24}$$

Moreover, inequality (2.24) is strict unless X equals $X_{\sigma^2,1}$ or $X_{A_{c,\sigma},B_{c,\sigma}}$ in distribution, depending on whether $\sigma^2 \le A_c$ or $\sigma^2 \ge A_c$.

To complete this subsection, let us describe the asymptotics of the bound $L_{T;c,\sigma}$ for $\sigma \downarrow 0$ and $\sigma \to \infty$; cf. Proposition 2.2.

Proposition 2.4. *For any real $c > 0$,*

$$L_{T;c,\sigma} - 1 \sim -c\sigma^2 \quad \text{as } \sigma \downarrow 0, \tag{2.25}$$

$$L_{T;c,\sigma} \sim \frac{4}{c^2}\frac{\ln^2 \sigma}{\sigma^2} \quad \text{as } \sigma \to \infty. \tag{2.26}$$

2.3. Winsorization and truncation: discussion and comparison

The Winsorization function W and the truncation function T as defined by (2.1) and (2.20) 'cut' a given value x at the level 1. However, by simple rescaling, it is easy to restate the results for any positive 'cut' level y. Indeed, we may consider $W_y(x) := y \wedge x$ and $T_y(x) := x \, \mathbf{1}\{x < y\}$, so that $W_1 = W$ and $T_1 = T$. Then $cW_y(X) = cyW(X/y)$ and $cT_y(X) = cyT(X/y)$. Now we can use the results of Subsections 2.1 and 2.2 with c, X, and σ replaced by cy, X/y, and σ/y, respectively. It should therefore be clear that the 'cut' level was set to be 1 just for the simplicity of presentation.

Observe that, for each $c > 0$, the exact lower bound $L_{W;c,\sigma}$ in (2.9) is no greater than 1, since the zero RV X obviously satisfies the conditions $\mathrm{E}\,X \geq 0$ and $\mathrm{E}\,X^2 \leq \sigma^2$. Hence, the exact lower bounds $L_{W;\sigma}$ and $L_{T;c,\sigma}$, which are no greater than $L_{W;c,\sigma}$, are also no greater than 1. It is also clear that each of these exact lower bounds is nondecreasing in σ—since the exactness is over all RVs X with $\mathrm{E}\,X \geq 0$ and $\mathrm{E}\,X^2 \leq \sigma^2$. However, for any $c > 0$, the exact lower bound $L_{W;c,\sigma}$ for the Winsorized RV $W(X)$ decreases rather slowly from 1 to 0 as σ increases from 0 to ∞. Even the smaller, universal over all $c > 0$ exact lower bound $L_{W;\sigma}$ decreases rather slowly; see Figure 1 and also recall Proposition 2.2. In particular, for the value $\sigma^2 = 1$ (which is of special interest as far as the application in [23] is concerned), the lower bound $L_{W;\sigma}$ is $0.878\ldots$, rather close to 1. Even for $\sigma^2 = 100$, this bound is $0.194\ldots$, not very small.

Moreover, the universal (over all $c > 0$) Winsorization bound $L_{W;\sigma}$ is remarkably close to the fixed-c Winsorization bounds $L_{W;c,\sigma}$, especially if the value of c is in the interval $[1, 3] = \{p/2 \colon 2 \leq p \leq 6\}$—which is of particular interest in [23]. See the top diagram of Figure 2; the $c = 2$ line (cf. (2.19)) of this diagram looks exactly horizontal at level 1, but it is in fact not.

As for the truncation case, it is quite different from the Winsorization case. Indeed, the exact lower bound $L_{T;c,\sigma}$ is significantly smaller than $L_{W;c,\sigma}$, especially for larger values of σ. In the bottom diagram of Figure 2 we plot the ratios of these two bounds for $c = 1, \frac{3}{2}, 2, 3, 5$.

It is also easy to compare the asymptotics for $L_{T;c,\sigma}$ in (2.25) and (2.26) with those for $L_{W;c,\sigma}$ in (2.12) and (2.13). Comparing (2.12) with (2.25), it is easy to see that, for $\sigma \downarrow 0$, $1 - L_{T;c,\sigma}$ is at least four times as large (asymptotically) as $1 - L_{W;c,\sigma}$, and may be infinitely
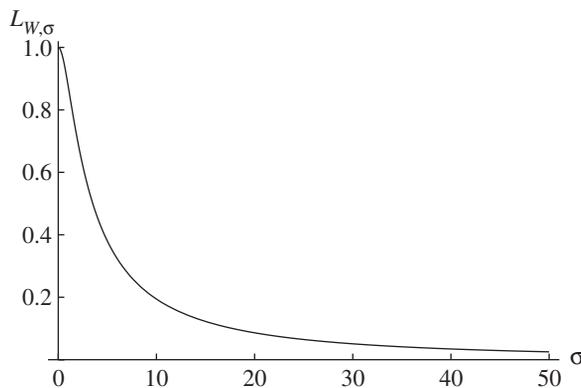
FIGURE 1: The exact lower bound $L_{W;\sigma}$.
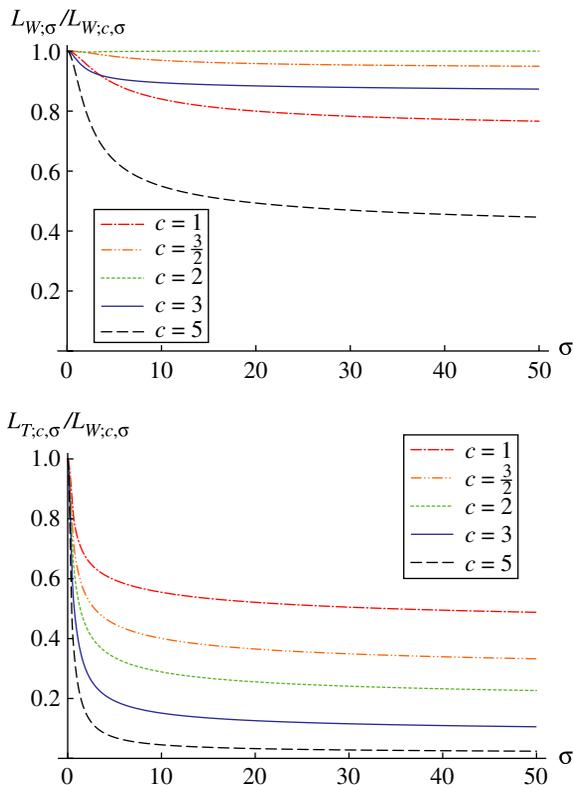
FIGURE 2: The ratios $L_{W;\sigma}/L_{W;c,\sigma}$ (*top*) and $L_{T;c,\sigma}/L_{W;c,\sigma}$ (*bottom*).

many times as large when c goes to 0 or ∞. Similarly, for $\sigma \to \infty$, $L_{W;c,\sigma}$ is e^c times as large (asymptotically) as $L_{T;c,\sigma}$.

Moreover, in contrast with the Winsorization case, there is no nontrivial lower bound in the truncation case that would be universal over all $c > 0$. Namely, for any given $\sigma > 0$, the infimum of $E \exp\{cT(X)\}$ over all $c > 0$ and all RVs X with $E X \geq 0$ and $E X^2 \leq \sigma^2$ is 0; the same holds even if the conditions $E X \geq 0$ and $E X^2 \leq \sigma^2$ are strengthened to $E X = 0$ and $E X^2 = \sigma^2$. Indeed, let $a \downarrow 0$, $b := \sigma^2/a$, and $c := 1/a^2$; then it is easy to see that $E \exp\{cT(X_{a,b})\} \to 0$; cf. (2.25) and (2.26), with large c.

Let us also briefly discuss the case $c < 0$. Then the function $e^{cW(\cdot)}$ is convex and nonincreasing, while $e^{cT(\cdot)} \geq e^{cW(\cdot)}$. These simple observations together with Jensen's inequality lead to the trivial lower bound 1 on both $E e^{cW(X)}$ and $E e^{cT(X)}$ for any RV X with $E X \leq 0$:

$$E e^{cT(X)} \geq E e^{cW(X)} \geq e^{cW(E X)} \geq e^{cW(0)} = 1.$$

This lower bound is exact for any prescribed value $\sigma^2 \in (0, \infty)$ of $E X^2$, which is easily seen by letting $X = X_{a,\sigma^2/a}$ with $a \downarrow 0$; indeed, then eventually

$$E e^{cT(X)} = \frac{a^2 + \sigma^2 e^{-ca}}{a^2 + \sigma^2} \to 1.$$

We may also note that no finite upper bound exists in this case; indeed, again let $X = X_{a,\sigma^2/a}$, but now with $a \to \infty$; then eventually

$$\mathrm{E}\,\mathrm{e}^{cT(X)} \geq \mathrm{E}\,\mathrm{e}^{cW(X)} = \frac{a^2 \mathrm{e}^{c\sigma^2/a} + \sigma^2 \mathrm{e}^{-ca}}{a^2 + \sigma^2} \to \infty.$$

Thus, the case $c < 0$ is of little interest in this setting. However, imposing the additional restriction that $X \geq 0$ with probability 1 and considering the exponential moments $\mathrm{E}\,\mathrm{e}^{cX}$ of the RV X itself (rather than its Winsorization $W(X)$ or truncation $T(X)$), one has the following interesting inequalities, both for $c > 0$ and $c < 0$:

$$\mathrm{E}\,\mathrm{e}^{cX} \geq \mathrm{E}\,\mathrm{e}^{c\hat{X}} \quad \text{for } c > 0 \quad \text{and} \quad \mathrm{E}\,\mathrm{e}^{cX} \leq \mathrm{E}\,\mathrm{e}^{c\hat{X}} \quad \text{for } c < 0.$$

Here X is any nonnegative RV with $\mathrm{E}\,X^2 < \infty$ and $\hat{X}$ is an RV taking values in the set of the form $\{0, b\}$ for some $b \in (0, \infty)$ and such that $\mathrm{E}\,\hat{X}^j = \mathrm{E}\,X^j$ for $j = 1, 2$; these inequalities immediately follow from the more general result [16, Inequalities (2.32)]. The method of [16] is based on s-orderings (rather than on duality); for more on that method, see, e.g. [5]; concerning stochastic orders in general, see [25].

3. Another application: covered call option pricing

A covered (European) call option is an investment portfolio which consists of a unit of stock less a call option for the same amount of the stock. The value of such a portfolio at the expiration time of the option is $S - (S - K)_+ = K \wedge S$, where K is the strike price and S is the stock price at expiry. In the Black–Scholes model,

$$S = S_0 \mathrm{e}^Y,$$

where S_0 is the initial price of the stock (at time moment 0),

$$Y = \sigma B_T + \left(r - \tfrac{1}{2}\sigma^2\right)T, \tag{3.1}$$

σ is the volatility, T is the expiration time, B. is a standard Brownian motion, and r is the interest rate. Here K, S_0, σ, T, and r are assumed to be known nonrandom positive real constants. (It is hoped that the role that the symbol T is given in this section will not be confused with the truncation role T plays in other sections; in the current section we shall deal only with Winsorization and not at all with truncation.) The expected value, say V_{BS}, of $K \wedge S$ in the Black–Scholes model can be easily expressed in terms of the standard normal distribution function; note that the Black–Scholes formula provides such an expression for the expected value, $\mathrm{E}(S - K)_+$, of the call option.

Assume now that the RV $Y = \ln(S/S_0)$ is not necessarily normally distributed, as in the Black–Scholes model (3.1), but may follow any distribution with the same mean $\mu := \mathrm{E}\,Y = (r - \sigma^2/2)T$ and variance $\mathrm{var}\,Y = \sigma^2 T$. Let

$$c := \ln \frac{K}{S_0} - \mu; \tag{3.2}$$

for about a half of all call options encountered in practice, one will have $c > 0$, which we shall assume to be the case. Then $K \wedge S = S_0 \mathrm{e}^\mu \mathrm{e}^{cW(X)}$, where $X := (Y - \mu)/c$. Note also that
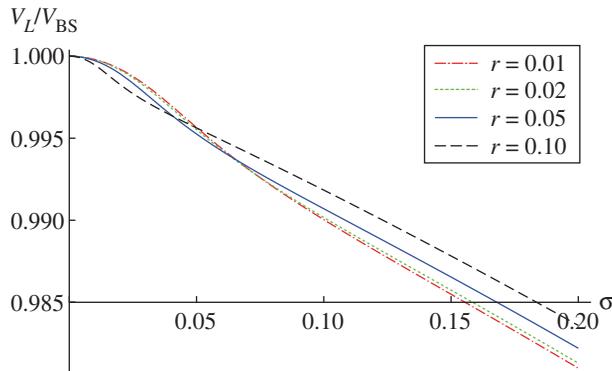
FIGURE 3: The ratios V_L/V_{BS}.

E $X = 0$ and E $X^2 = \sigma^2 T/c^2$. Thus, by (2.9), we have an exact lower bound on the expected value of the covered call:

$$\mathrm{E}(K \wedge S) \geq V_L := S_0 \mathrm{e}^{\mu} L_{W;c,\sigma\sqrt{T}/c}.$$

The ratios V_L/V_{BS} for $K/S_0 = 1.05\mathrm{e}^{rT}$, $T = 0.5$, $r = 0.01, 0.02, 0.05, 0.10$, and $\sigma \in [0, 0.20]$ are shown in Figure 3, which suggests that the lower bound V_L on the expected value of the covered call is not much less than V_{BS}, the expected value for the Black–Scholes model.

Other exact lower and upper bounds for financial assets are known, under various conditions. In particular, Lo [18] obtained an exact upper bound on the expected value of $(S - K)_+$ given the mean and variance of S (rather than of $Y = \ln(S/S_0)$). Grundy [9] provided such a bound given the pth moment of S, for $p \geq 1$. De la Peña et al. [4] gave exact lower and/or upper bounds on $\mathrm{E}(S - K)_+$ given the mean, variance, and other characteristics of the distribution of S: $\mathrm{P}(S > K)$, the p-norm of the density function, or the third moments of S, as well as exact lower and upper bounds on $\mathrm{E}(S - K)_+$ in the binomial model.

4. Proofs

The general problem of finding the maximum or minimum of the generalized moment $\int f \, \mathrm{d}\mu$ over the set of all nonnegative measures μ with given generalized moments $\int f_i \, \mathrm{d}\mu$ ($i \in I$) goes back to Chebyshev and Markov; here a function f and a family of functions $(f_i)_{i \in I}$ are given; see, e.g. [10], [12], [13], [14], [15], and [19]. One group of results in this area is that, for finite I, under general conditions, it may be assumed without loss of generality that the support of μ is also finite, with cardinality no greater than that of I; methods based on such results may be referred to as finite-support methods. Other results, valid for finite or infinite I, concern the following duality: under general conditions, the supremum (say) of $\int f \, \mathrm{d}\mu$ over all μ such that $\int f_i \, \mathrm{d}\mu \leq c_i$ for all $i \in I$ coincides with the infimum of $\int_I c_i \nu(\mathrm{d}i)$ over all nonnegative measures ν on I (say with a finite support) such that $\int_I f_i \nu(\mathrm{d}i) \geq f$.

Such methods were used in, e.g. [1], [2], [3], [11], [21], [24], [26], and [27]. In particular, the supremum of $\int_{\mathbb{R}} \mathrm{e}^x \mu(\mathrm{d}x)$ given $\int_y^{\infty} \mu(\mathrm{d}x) = 0$, $\int_{\mathbb{R}} x \mu(\mathrm{d}x) = 0$, and $\int_{\mathbb{R}} x^2 \mu(\mathrm{d}x) = \sigma^2$ was found (implicitly) in [1] and (explicitly) in [24].

In [11] a similar problem was solved, under the additional restriction that μ is a probability measure. This result was extended in [2] and [3] to the moment functions $(\cdot - t)_+^2$ ($t \in \mathbb{R}$) in place of the exponential function $e^{\cdot}$ in [11]; on the other hand, this was a further development of the line of results obtained in [6], [7], [19], and [20]. The results of [1], [11] and [2], [3] were refined in [24] and [21], respectively, by also taking into account positive-part third moments. The supremum of the moments $\int f \, d\mu$ over all Stein-type moment functions $x \mapsto f(x) := xg(x) - g'(x)$ with Lipschitz-1 functions g' and over all probability measures μ with given mean, variance, and third absolute moment was presented (in an equivalent form) in [8, Theorem 3] and [26, Theorem 3]; see also [17]. Results somewhat related to the mentioned ones were obtained in [22]; see also the bibliography therein. Of course, mentioned above is a very small sample of the work done on the Chebyshev–Markov type of extremal problems.

Concerning our problems of minimizing the exponential moments of $W(X)$ and $T(X)$, we could use the mentioned finite-support methods to reduce the consideration to RVs X taking only three values, since we have here three affine restrictions: on the first two moments and on the total mass of the measure (which has to be a probability measure). Another, more ad hoc kind of approach would be to condition the distribution of the RV X on $\mathbf{1}\{X < 1\}$, which would preserve the mean and would not increase the second moment; also, this conditioning would not increase the exponential moments of $W(X)$ and $T(X)$, since both functions, e^{cW} and e^{cT}, are convex on $(-\infty, 1)$ and on $[1, \infty)$; thus, it would remain to consider RVs X taking only two values. However, the duality-type method that we chose to prove inequalities (2.9) and (2.24) appears more effective, as it immediately reduces the consideration to RVs $X_{a,b}$ that not only take just two values, but also have the first two moments exactly equal to 0 and σ^2, respectively; moreover, this approach appears more convenient in obtaining the strictness conditions for inequalities (2.9) and (2.24).

Let us now turn to proving the specific statements made in Section 2.

Proof of Proposition 2.1. Part (i) follows because $ab^*_{a,c}$ strictly and continuously increases from 0 to ∞ as a increases from 0 to ∞.

(ii) Observe that $a(a^2 + \sigma^2)^2 \ell'_1(a)$ is a quadratic polynomial in σ^2, whence we can see that the system of inequalities $\ell'_1(a) > 0$ and $0 < a < \sigma^2$ can be rewritten as $0 < a < \sqrt{1 + \sigma^2} - 1$. This means that $\ell_1(a)$ switches from increase to decrease over $a \in (0, \sigma^2)$; at that, $\ell_1(0+) = -\infty$ and $\ell_1(\sigma^2) = 0$. Now part (ii) of Proposition 2.1 follows as well.

Proof of Theorem 2.1. The idea of the proof is in accordance with the general duality principle discussed at the beginning of this section. Indeed, take any real $c > 0$. Let

$$F(x) := e^{cW(x)} = e^{c(x \wedge 1)} \quad \text{and} \quad G(x) := \alpha + \beta x + \gamma x^2 \tag{4.1}$$

for all $x \in \mathbb{R}$, where α, β, and γ are real parameters.

We are trying to obtain a tight lower bound on the F-moment $\mathrm{E}\, F(X)$ of an RV X subject to linear (or, more exactly, affine) restrictions on its power moments: $\mathrm{E}\, X^0 = 1$, $\mathrm{E}\, X^1 \geq 0$, and $\mathrm{E}\, X^2 \leq \sigma^2$. Thus, the function G is a linear combination of the power moment functions $x \mapsto x^j$ for $j = 0, 1, 2$. We would like to find values of the parameters α, β, and γ such that the function G is a tightest possible minorant of F on $\mathbb{R}$, as shown in Figure 4. In particular, we may begin by taking arbitrary real numbers $a > 0$ and

$$b > 1, \tag{4.2}$$

and then solving for α, β, and γ the system of three linear (in (α, β, γ)) equations

$$F(b) = G(b), \qquad F'(b) = G'(b), \quad \text{and} \quad F'(-a) = G'(-a). \tag{4.3}$$
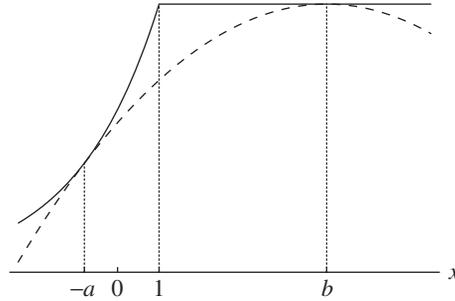
FIGURE 4: A tightest possible minorization of F (*solid line*) by G (*dashed line*).

It is not hard to see that the formulae

$$\alpha := \mathrm{e}^c - \frac{b^2 c \mathrm{e}^{-ac}}{2(a+b)}, \qquad \beta := \frac{bc\mathrm{e}^{-ac}}{a+b}, \qquad \gamma := -\frac{c\mathrm{e}^{-ac}}{2(a+b)} \tag{4.4}$$

indeed define a solution to system (4.3). Moreover, with such α, β, and γ we have $2\mathrm{e}^{ac}(F(-a) - G(-a)) = bc + ac - 2\mathrm{e}^{ac+c} + 2$, whence we find that the condition $F(-a) = G(-a)$ will be satisfied by letting b equal $b_{a,c}^*$, as defined in (2.2); note that then $b > 2$, so that condition (4.2) holds, as well as all the four 'tightness' conditions

$$F(b) = G(b), \qquad F'(b) = G'(b), \qquad F(-a) = G(-a), \qquad F'(-a) = G'(-a).$$

Next, by (4.4), $\gamma < 0$ and, hence, the function G is strictly concave, while the function F is convex on $(-\infty, 1)$ and on $[1, \infty)$; so, the difference $D := F - G$ is strictly convex on $(-\infty, 1)$ and on $[1, \infty)$; at that, $D(-a) = D'(-a) = D(b) = D'(b) = 0$, whence $D > 0$ and $F > G$ on $\mathbb{R} \setminus \{-a, b\}$, while $F = G$ on the two-point set $\{-a, b\}$.

Now specifically choose $a = a_{c,\sigma}$. Then, recalling (2.3) and (2.6), we see that

$$b_{a_{c,\sigma},c}^* = b_{c,\sigma}. \tag{4.5}$$

Therefore, also specifically choosing $b = b_{c,\sigma}$, we have

$$\mathrm{E}\, X_{a,b}^2 = ab = \sigma^2 \geq \mathrm{E}\, X^2 \tag{4.6}$$

and

$$\mathrm{E}\,\mathrm{e}^{cW(X)} = \mathrm{E}\, F(X) \geq \mathrm{E}\, G(X) = \alpha + \beta\, \mathrm{E}\, X + \gamma\, \mathrm{E}\, X^2 \tag{4.7}$$

$$\geq \alpha + \beta\, \mathrm{E}\, X_{a,b} + \gamma\, \mathrm{E}\, X_{a,b}^2 \tag{4.8}$$

$$= \mathrm{E}\, G(X_{a,b}) = \mathrm{E}\, F(X_{a,b}) = \mathrm{E}\exp\{cW(X_{a,b})\}; \tag{4.9}$$

inequality (4.8) takes place because (in view of (4.4)) $\beta > 0 > \gamma$, while $\mathrm{E}\, X \geq 0 = \mathrm{E}\, X_{a,b}$ and, by (4.6), $\mathrm{E}\, X^2 \leq \mathrm{E}\, X_{a,b}^2$. Thus, (2.9) follows. Moreover, because $F > G$ on $\mathbb{R} \setminus \{-a, b\}$, the inequality in (4.7) is strict unless the support of the distribution of X is a subset of $\{-a, b\}$, and inequality (4.8) is strict unless $\mathrm{E}\, X = 0$ and $\mathrm{E}\, X^2 = \sigma^2 = ab$; thus, indeed, inequality (2.9) is strict unless $X \stackrel{\mathrm{D}}{=} X_{a,b}$.

To prove inequality (2.10), it suffices to show that $\mathrm{E}\exp\{c_\sigma W(X_{a_\sigma, b_\sigma})\}$ is a lower bound on $\mathrm{E}\exp\{cW(X_{a,b})\}$ for any a and b such that $a \in (0, \sigma^2)$ and $b = \sigma^2/a$. Take any such a and b.

Observe that $\mathrm{E} \exp\{cW(X_{a,b})\} = (a^2 \mathrm{e}^c + \mathrm{e}^{-ac}\sigma^2)/(a^2 + \sigma^2)$ is strictly convex in $c \in \mathbb{R}$ and attains its minimum,

$$m(a, \sigma) := \frac{a(1+a)(a/\sigma^2)^{-1/(1+a)}}{a^2 + \sigma^2},$$

in c only at $c = \ln(\sigma^2/a)/(1+a) = (\ln b)/(1+a)$; cf. (2.8). Next, the minimum of $m(a, \sigma)$ or, equivalently, of

$$\ell(a) := \ell(a, \sigma) := \ln m(a, \sigma)$$

in $a \in (0, \sigma^2)$ is attained only at the point $a = a_\sigma$ defined by (2.5), because, by part (ii) of Proposition 2.1, $\ell'(a) = \ell_1(a)/(1+a)^2$ changes sign from negative to positive over $a \in (0, \sigma^2)$. Thus, inequality (2.10) is true, and it is strict unless $c = c_\sigma$ and $a_{c,\sigma} = a_\sigma$.

It remains to verify the three equalities in (2.11). In view of (4.5), the last of these equalities is implied by the first one. So, if any of the equalities in (2.11) were false then the equality $X_{a_\sigma, b_\sigma} \stackrel{\mathrm{D}}{=} X_{a_{c_\sigma},\sigma, b_{c_\sigma},\sigma}$ would also be false, and so, by what has been proved, inequality (2.9) with c_σ and X_{a_σ, b_σ} in place of c and X would be strict, which would contradict inequality (2.10).

This completes the proof of Theorem 2.1.

Proof of Proposition 2.2. To prove part (i), consider first the case $\sigma \downarrow 0$. Then, by (2.2) and (2.3), $a_{c,\sigma} \downarrow 0$. Moreover, $b^*_{a,c} \to 2(\mathrm{e}^c - 1)/c$ whenever $a \to 0$. So, by (4.5), $b_{c,\sigma} \to 2(\mathrm{e}^c - 1)/c > 2 > 1$, and so, $a_{c,\sigma} = \sigma^2/b_{c,\sigma} \sim c\sigma^2/(2(\mathrm{e}^c - 1))$. On the other hand,

$$\mathrm{E} \exp\{cW(X_{a,b})\} - 1 = (\mathrm{e}^c - 1)\frac{a}{a+b} + (\mathrm{e}^{-ca} - 1)\frac{b}{a+b} \sim \left(\frac{\mathrm{e}^c - 1}{b} - c\right)a \quad (4.10)$$

whenever $a \downarrow 0$ and $a = o(b)$. This, together with the relations $b_{c,\sigma} \to 2(\mathrm{e}^c - 1)/c$ and $a_{c,\sigma} \sim c\sigma^2/(2(\mathrm{e}^c - 1))$, implies (2.12).

Now consider the case $\sigma \to \infty$. Then, by (2.2) and (2.3), $a_{c,\sigma} \to \infty$. Next,

$$b^*_{a,c} \sim \frac{2\mathrm{e}^c}{c}\mathrm{e}^{ac} \quad \text{as } a \to \infty. \quad (4.11)$$

So, in view of (2.6) and (4.5), for $a = a_{c,\sigma}$, we have

$$\sigma^2 \sim \frac{2\mathrm{e}^c}{c^2}ac\mathrm{e}^{ac} = \mathrm{e}^{ac(1+o(1))}, \quad (4.12)$$

whence $a_{c,\sigma} = a \sim (1/c)\ln(\sigma^2)$ and $b_{c,\sigma} = \sigma^2/a_{c,\sigma} \sim c\sigma^2/\ln(\sigma^2)$. Also, for $a \to \infty$ and $b = b^*_{a,c}$, (4.11) yields $a = o(b)$ and

$$\mathrm{E} \exp\{cW(X_{a,b})\} = \frac{a\mathrm{e}^c + b\mathrm{e}^{-ac}}{a+b} \sim \frac{a\mathrm{e}^c + 2\mathrm{e}^c/c}{a+b} \sim \frac{a\mathrm{e}^c}{b}. \quad (4.13)$$

This, together with the relations $a_{c,\sigma} \sim (1/c)\ln(\sigma^2)$ and $b_{c,\sigma} \sim c\sigma^2/\ln(\sigma^2)$, implies (2.13).

Part (ii) follows because $f(0+) = -\infty$, $f(1) = 0$, and $f'(t) = -2(t - \frac{1}{2})/t$ changes sign from positive to negative as t increases from 0 to 1.

To prove part (iii), consider first the case $\sigma \downarrow 0$. Then, by (2.4) and (2.14), for each fixed $t \in (0, 1)$, we have $\ell_1(t\sigma^2) = \ln t - (2 + o(1))(t - 1) = f(t) + o(1)$. So, by part (ii), for each fixed $t \in (0, t_*)$, we have $\ell_1(t\sigma^2) < 0$, eventually (for all small enough σ); similarly, for each fixed $t \in (t_*, 1)$, eventually $\ell_1(t\sigma^2) > 0$. Therefore, by (2.5), $a_\sigma \sim t_*\sigma^2$ and, hence,

by (2.7) and (2.8), $b_\sigma \to 1/t_*$ and $c_\sigma \to -\ln t_*$. Now (2.16) follows by (4.10), since, for $c = c_\sigma$, we have $e^c \to 1/t_*$ and $c \to -\ln t_* = 2(1 - t_*)$, where the equality follows from (2.15) and (2.14).

The case $\sigma \to \infty$ is considered similarly. Then, by (2.4), $\ell_1(\kappa \ln(\sigma^2)) \sim 2(\kappa - \frac{1}{2}) \ln(\sigma^2)$ for each fixed $\kappa \in (0, \infty) \setminus \{\frac{1}{2}\}$, so that $\ell_1(\kappa \ln(\sigma^2))$ is eventually less than 0 for each $\kappa \in (0, \frac{1}{2})$ and eventually greater than 0 for each $\kappa \in (\frac{1}{2}, \infty)$. Thus, by part (ii) of Proposition 2.1, $a_\sigma \sim \frac{1}{2} \ln(\sigma^2)$, and, hence, $b_\sigma \sim 2\sigma^2/\ln(\sigma^2)$ and $c_\sigma \to 2$. Moreover, by (2.11), we have $b_\sigma = b^*_{a_\sigma, c_\sigma}$. Recall that relations (4.11) and (4.13) were derived assuming that $a \to \infty$, $b = b^*_{a,c}$, and $c > 0$ is fixed. Reasoning quite similarly—with $a_\sigma, b_\sigma, c_\sigma$ in place of such a, b, c—we conclude that $L_{W;\sigma} \sim a_\sigma e^{c_\sigma}/b_\sigma$, and now (2.17) follows since $a_\sigma \sim \frac{1}{2} \ln(\sigma^2)$, $b_\sigma \sim 2\sigma^2/\ln(\sigma^2)$, and $c_\sigma \to 2$.

(iv) The derivative of $-c^2/(4(e^c - 1))$ in $c > 0$ is positive if and only if $f(t) < 0$ for $t := e^{-c}$. Therefore, and by part (ii), $-c^2/(4(e^c - 1))$ attains a minimum in $c > 0$ at $c = -\ln t_* = 2(1 - t_*)$. Replacing c in the denominator of $-c^2/(4(e^c - 1))$ by $-\ln t_*$ and c in the numerator by $2(1 - t_*)$, we obtain (2.18). As for (2.19), it is much more straightforward.

Proof of Proposition 2.3. That each of the equations on the right-hand sides of (2.21) and (2.22) has a unique root $a > 0$ follows because both $B^*_{a,c}$ and $a B^*_{a,c}$ strictly and continuously increase from 0 to ∞ as a does so. Now (2.23) follows because the value of $a B^*_{a,c}$ at $a = A_c$ is A_c.

Proof of Theorem 2.2. Here let $F(x) := e^{cT(x)}$ and let $G(x)$ be defined as in (4.1). Consider first the case $\sigma^2 \le A_c$. Take $G(x)$ with

$$\alpha := \frac{e^{-ac}(a^2 e^{ac} + ca^2 + ac + 2a + 1)}{(a+1)^2},$$

$$\beta := \frac{e^{-ac}(2a(e^{ac} - 1) + c(1 - a^2))}{(a+1)^2}, \qquad \gamma := \frac{e^{-ac}(e^{ac} - ac - c - 1)}{(a+1)^2}.$$

Then, for $D := F - G$ and any $a > 0$, we have $D(-a) = D'(-a) = D(1) = 0$. Now let $a = \sigma^2$, so that (by the current case condition) $0 < a \le A_c$, which implies that $B^*_{a,c} \le 1$ and, hence, $G'(1) \le 0$ (because $G'(1) = (B^*_{a,c} - 1)ce^{-ac}/(1 + a)$); also, $0 < a \le A_c$ implies that $(e^{ac} - 1) - (1 + a)c < 2(e^{ac} - 1) - (1 + a)c = (B^*_{a,c} - 1)c \le 0$, whence $\gamma < 0$, so that G is strictly convex on $\mathbb{R}$; moreover, $\beta > e^{-ac}c(1 + a^2)/(1 + a)^2 > 0$. In turn, the inequality $G'(1) \le 0$ means that $D'(1+) \ge 0$; also, D is strictly convex on $(-\infty, 1)$ and on $[1, \infty)$. Recalling now that $D(-a) = D'(-a) = D(1) = 0$, we have $D > 0$ and $F > G$ on $\mathbb{R} \setminus \{-a, 1\}$, while $F = G$ on the two-point set $\{-a, 1\} = \{-\sigma^2, 1\}$. Now the first line of (2.24) follows; cf. (4.7)–(4.9).

Consider now the case $\sigma^2 \ge A_c$. Take $G(x)$ with

$$\alpha := \frac{e^{-ac}(ca^2 + 2abc + 2a + 2b)}{2(a+b)}, \qquad \beta := \frac{cbe^{-ac}}{a+b}, \qquad \gamma := -\frac{ce^{-ac}}{2(a+b)},$$

where $a > 0$ and $b := B^*_{a,c}$. Assume now that a is so large that $b \ge 1$. Then, again for $D := F - G$, we have $D(-a) = D'(-a) = D(b) = D'(b) = 0$, while $\beta > 0 > \gamma$, so that again D is strictly convex on $(-\infty, 1)$ and on $[1, \infty)$, $D > 0$ and $F > G$ on $\mathbb{R} \setminus \{-a, b\}$, while $F = G$ on the two-point set $\{-a, b\}$; if $b = 1$ then $D'(b)$ is understood as the right derivative of D at point 1. Since the current case is $\sigma^2 \ge A_c$, (2.23) yields $A_{c,\sigma} \ge A_c$ and, hence,

$B_{c,\sigma} = \sigma^2/A_{c,\sigma} = B^*_{A_{c,\sigma},c} \geq B^*_{A_c,c} = 1$. Now, reasoning again similarly to (4.7)–(4.9), we obtain the second line of (2.24).

The proof of the strictness statement on (2.24) is quite similar to that for (2.9), because here too we have $\beta > 0 > \gamma$—in either case, whether $\sigma^2 \leq A_c$ or $\sigma^2 \geq A_c$.

Proof of Proposition 2.4. Here the case $\sigma \downarrow 0$ is quite straightforward. Indeed, by (2.24), we eventually have $L_{T;c,\sigma} - 1 = \mathrm{E}\exp\{cT(X_{\sigma^2,1})\} - 1 = (\mathrm{e}^{-c\sigma^2} - 1)/(1 + \sigma^2) \sim -c\sigma^2$.

As for the case $\sigma \to \infty$, the proof of (2.26) is quite similar to that of relation (2.13) in part (i) of Proposition 2.2: replace all instances of $b^*_{a,c}$, $a_{c,\sigma}$, $b_{c,\sigma}$, W by $B^*_{a,c}$, $A_{c,\sigma}$, $B_{c,\sigma}$, T, respectively, and also drop all instances of the factor e^c in the numerators of the ratios in (4.11), (4.12), and (4.13).

Acknowledgement

References

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

[2] BENTKUS, V. (2002). A remark on the inequalities of Bernstein, Prokhorov, Bennett, Hoeffding, and Talagrand. *Liet. Mat. Rink.* **42,** 332–342 (in Russian). English translation: *Lithuanian Math. J.* **42,** 262–269.

[3] BENTKUS, V. (2004). On Hoeffding's inequalities. *Ann. Prob.* **32,** 1650–1673.

[4] DE LA PEÑA, V. H., IBRAGIMOV, R. AND JORDAN, S. (2004). Option bounds. In *Stochastic Methods and Their Applications* (J. Appl. Prob. Spec. Vol. **41A**), eds J. Gani and E. Seneta, Applied Probability Trust, Sheffield, pp. 145–156.

[5] DENUIT, M., LEFEVRE, C. AND SHAKED, M. (1998). The *s*-convex orders among real random variables, with applications. *Math. Inequal. Appl.* **1,** 585–613.

[6] EATON, M. L. (1970). A note on symmetric Bernoulli random variables. *Ann. Math. Statist.* **41,** 1223–1226.

[7] EATON, M. L. (1974). A probability inequality for linear combinations of bounded random variables. *Ann. Statist.* **2,** 609–613.

[8] GOLDSTEIN, L. (2010). Bounds on the constant in the mean central limit theorem. Preprint. Available at http://arxiv.org/abs/0906.5145v2.

[9] GRUNDY, B. D. (1991). Option prices and the underlying asset's return distribution. *J. Finance* **46,** 1045–1069.

[10] HOEFFDING, W. (1955). The extrema of the expected value of a function of independent random variables. *Ann. Math. Statist.* **26,** 268–275.

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

[12] KARLIN, S. AND STUDDEN, W. J. (1966). *Tchebycheff Systems: With Applications in Analysis and Statistics* (Pure Appl. Math. **XV**). John Wiley, New York.

[13] KARR, A. F. (1983). Extreme points of certain sets of probability measures, with applications. *Math. Operat. Res.* **8,** 74–85.

[14] KEMPERMAN, J. H. B. (1983). On the role of duality in the theory of moments. In *Semi-Infinite Programming and Applications* (Austin, Texas, 1981; Lecture Notes Econom. Math. Systems **215**), Springer, Berlin, pp. 63–92.

[15] KREĬN, M. G. AND NUDEL´MAN, A. A. (1977). *The Markov Moment Problem and Extremal Problems*. American Mathematical Society, Providence, RI.

[16] LEFÈVRE, C. AND UTEV, S. (1996). Comparing sums of exchangeable Bernoulli random variables. *J. Appl. Prob.* **33,** 285–310.

[17] LEFÈVRE, C. AND UTEV, S. (2003). Exact norms of a Stein-type operator and associated stochastic orderings. *Prob. Theory Relat. Fields* **127,** 353–366.

[18] LO, A. W. (1987). Semi-parametric upper bounds for option prices and expected payoffs. *J. Financial Econom.* **19,** 373–387.

[19] PINELIS, I. (1998). Optimal tail comparison based on comparison of moments. In *High Dimensional Probability* (Oberwolfach, 1996; Progress Prob. **43**), Birkhäuser, Basel, pp. 297–314.

[20] PINELIS, I. (1999). Fractional sums and integrals of *r*-concave tails and applications to comparison probability inequalities. In *Advances in Stochastic Inequalities* (Atlanta, GA, 1997; Contemp. Math. **234**), American Mathematical Society, Providence, RI, pp. 149–168.

[21] PINELIS, I. (2009). On the Bennett–Hoeffding inequality. Preprint. Available at http://arxiv.org/abs/0902.4058v1.

[22] PINELIS, I. (2009). Optimal two-value zero-mean disintegration of zero-mean random variables. *Electron. J. Prob.* **14,** 663–727.

[23] PINELIS, I. AND MOLZON, R. (2009). Berry–Esséen bounds for general nonlinear statistics, with applications to Pearson's and non-central Student's and Hotelling's. Preprint. Available at http://arxiv.org/abs/0906.0177v1.

[24] PINELIS, I. S. AND UTEV, S. A. (1989). Sharp exponential estimates for sums of independent random variables. *Theory Prob. Appl.* **34,** 340–346.

[25] SHAKED, M. AND SHANTHIKUMAR, J. G. (2007). *Stochastic Orders*. Springer, New York.

[26] TYURIN, I. (2009). New estimates of the convergence rate in the Lyapunov theorem. Preprint. Available at http://arxiv.org/abs/0912.0726v1.

[27] UTEV, S. A. (1985). Extremal problems in moment inequalities. In *Limit Theorems of Probability Theory* (Trudy Inst. Mat. **5**), 'Nauka' Sibirsk. Otdel., Novosibirsk, pp. 56–75, 175.