Bull. Aust. Math. Soc. **92** (2015), 149–158
doi:10.1017/S0004972715000350

EXACT UPPER AND LOWER BOUNDS ON THE DIFFERENCE BETWEEN THE ARITHMETIC AND GEOMETRIC MEANS

IOSIF PINELIS

(Received 5 December 2014; accepted 28 February 2015; first published online 4 May 2015)

Abstract

Exact upper and lower bounds on the difference between the arithmetic and geometric means are obtained. The inequalities providing these bounds may be viewed, respectively, as a reverse Jensen inequality and an improvement of the direct Jensen inequality, in the case when the convex function is the exponential.

2010 *Mathematics subject classification*: primary 60E15; secondary 26D15, 90C46.

Keywords and phrases: arithmetic mean, geometric mean, Jensen inequality, reverse Jensen inequality, exact bounds, Tchebycheff–Markoff systems.

1. Summary and discussion

Let $\mathscr{X}_+$ denote the set of all nonnegative random variables (r.v.s) X with $\mathsf{E}\,X < \infty$. Take any $X \in \mathscr{X}_+$ and let

$$\begin{aligned} V_X &:= \mathsf{Var}\,\sqrt{X}, \quad m_X := \inf \operatorname{supp} X, \quad M_X := \sup \operatorname{supp} X, \\ E_X &:= \mathsf{E}(\sqrt{X} - \sqrt{m_X}\,)^2, \quad F_X := \mathsf{E}(\sqrt{M_X} - \sqrt{X}\,)^2, \end{aligned} \tag{1.1}$$

where, as usual, $\operatorname{supp} X$ denotes the support of (the distribution of) the r.v. X.

It will be shown in this note that

$$(2V_X) \wedge \frac{F_X V_X}{F_X - V_X} \leqslant \mathsf{E}\,X - \exp \mathsf{E} \ln X \leqslant (2V_X) \vee E_X \tag{1.2}$$

and that each of these two bounds on $\mathsf{E}\,X - \exp \mathsf{E} \ln X$ is exact, in terms of V_X and E_X for the upper bound and in terms of V_X and F_X for the lower bound. As usual, for any real numbers $z_1, \ldots, z_n$, we write $z_1 \vee \cdots \vee z_n$ and $z_1 \wedge \cdots \wedge z_n$ for their maximum and minimum, respectively.

Since the r.v. X is nonnegative, clearly $m_X \in [0, \infty)$. However, concerning the value of M_X, one can then only say that $M_X \in [m_X, \infty]$, with the case $M_X = \infty$ certainly possible. Next, given the condition $\mathsf{E}\,X < \infty$, the values of E_X and V_X are necessarily

© 2015 Australian Mathematical Publishing Association Inc. 0004-9727/2015 $16.00

https://doi.org/10.1017/S0004972715000350 Published online by Cambridge University Press

finite and hence so is the upper bound in (1.2). On the other hand, $F_X = \infty$ if $M_X = \infty$. Even then, the lower bound in (1.2) will of course be finite. Concerning the ratio $F_X V_X/(F_X - V_X)$ in the lower bound in (1.2), for any $V \in \mathbb{R}$, $E \in \mathbb{R}$ and $F \in (-\infty, \infty]$, we assume the conventions that $FV/(F - V)$ equals V if $F = \infty$ and equals 0 if $F = V$. It will be seen that these conventions are the appropriate ones in the present context.

That the upper and lower bounds in (1.2) hold and are exact will be established in Theorem 1.4 below. The statement of Theorem 1.4 is preceded by three propositions, which complement and help in understanding the main result. All the necessary proofs are given in Section 2.

Take any $V \in \mathbb{R}$, $E \in \mathbb{R}$ and $F \in (-\infty, \infty]$. Introduce the sets

$$\mathscr{X}_{\mathsf{sup};V,E} := \{X \in \mathscr{X}_+ : V_X = V, E_X = E\}, \tag{1.3}$$

$$\mathscr{X}_{\mathsf{inf};V,F} := \{X \in \mathscr{X}_+ : V_X = V, F_X = F\}. \tag{1.4}$$

Proposition 1.1. *One has* $\mathscr{X}_{\mathsf{sup};V,E} \neq \emptyset$ *if and only if*

$$\text{either } E = V = 0 \quad \text{or} \quad E > V > 0. \tag{1.5}$$

Similarly, $\mathscr{X}_{\mathsf{inf};V,F} \neq \emptyset$ *if and only if*

$$\text{either } F = V = 0 \quad \text{or} \quad F > V > 0. \tag{1.6}$$

Values of V and E as in (1.5), as well as values of V and F as in (1.6), may be referred to as admissible.

Proposition 1.2. *If* $\mathscr{X}_{\mathsf{inf};V,F} \neq \emptyset$, *then*

$$E_{V,F} := \frac{FV}{F - V} = \inf\{E_X : X \in \mathscr{X}_{\mathsf{inf};V,F}\}. \tag{1.7}$$

If, moreover, $F < \infty$, *then the latter infimum is attained, and it is attained at a r.v.* $X \in \mathscr{X}_{\mathsf{inf};V,F}$ *if and only if* $\operatorname{supp} X = \{m_X, M_X\}$, *that is, if and only if* $\operatorname{supp} X$ *contains at most two points. If* $F = \infty$, *then the infimum in (1.7) is not attained.*

Proposition 1.3. *Take any* $X \in \mathscr{X}_+$. *Then both inequalities in (1.2) turn simultaneously into equalities if and only if the distribution of the r.v.* $\sqrt{X}$ *is the symmetric distribution on a set of at most two points in* $[0, \infty)$.

Theorem 1.4. *Let*

$$D_X := \mathsf{E}\, X - \exp \mathsf{E} \ln X. \tag{1.8}$$

Then

$$S_{V,E} := \sup\{D_X : X \in \mathscr{X}_{\mathsf{sup};V,E}\} = (2V) \vee E \quad \text{if } \mathscr{X}_{\mathsf{sup};V,E} \neq \emptyset, \tag{1.9}$$

$$I_{V,F} := \inf\{D_X : X \in \mathscr{X}_{\mathsf{inf};V,F}\} = (2V) \wedge E_{V,F} \quad \text{if } \mathscr{X}_{\mathsf{inf};V,F} \neq \emptyset. \tag{1.10}$$

These equalities hold if the sets $\mathscr{X}_{\mathsf{sup};V,E}$ *and* $\mathscr{X}_{\mathsf{sup};V,E}$ *are replaced by their respective subsets consisting of the r.v.s in* $\mathscr{X}_{\mathsf{sup};V,E}$ *and* $\mathscr{X}_{\mathsf{sup};V,E}$ *taking at most two values.*

https://doi.org/10.1017/S0004972715000350 Published online by Cambridge University Press

Clearly, inequalities (1.2) and the exactness of the upper and lower bounds in (1.2) immediately follow from Theorem 1.4.

REMARK 1.5. Note that $(2V) \vee E$ is nondecreasing in V and E, whereas $(2V) \wedge E_{V,F}$ is nondecreasing in V and nonincreasing in F (from $E_{V,V+} = 2V$ down to $E_{V,\infty} = V$). So, (1.9) will hold if the equalities $V_X = V$ and $E_X = E$ in the definition (1.3) of $\mathscr{X}_{\mathsf{sup};V,E}$ are replaced by the inequalities $V_X \leqslant V$ and $E_X \leqslant E$. Similarly, (1.10) will hold if the equalities $V_X = V$ and $F_X = F$ in the definition of (1.4) of $\mathscr{X}_{\mathsf{inf};V,F}$ are replaced by $V_X \geqslant V$ and $F_X \leqslant F$.

Moreover, it is now clear that inequalities (1.2) will hold if m_X and M_X in the definitions of E_X and F_X in (1.1) are replaced, respectively, by any nonnegative a and b such that $\operatorname{supp} X \subseteq [a, b]$.

It also follows from the mentioned monotonicity of the exact lower bound $(2V) \wedge E_{V,F}$ in F that the values of this bound are always between V and $2V$. □

The lower bound in (1.2) is an improvement of the zero bound, which follows immediately by the Jensen inequality for the (convex) exponential function. In particular, the condition $\mathsf{E}\, X < \infty$ implies $\mathsf{E} \ln X < \infty$; however, it is possible that $\mathsf{E} \ln X = -\infty$; we use the standard conventions $\ln 0 := -\infty$ and $\exp(-\infty) := 0$.

As for the second inequality in (1.2), one may consider it as a reverse Jensen inequality (cf. [3]). Indeed, one can write $\mathsf{E}\, X - \exp \mathsf{E} \ln X$ as $\mathsf{E}\, e^Y - e^{\mathsf{E}\, Y}$ for $Y := \ln X$. In contrast with the upper bound in (1.2), the bounds in [3] will be finite only when $M_X - m_X < \infty$. On the other hand, the bounds in (1.2) are only for the case when the convex function is the exponential function.

In the case when the r.v. X is a continuous function on the interval $[0, 1]$ endowed with the Lebesgue measure, obtaining the upper bound $(\sqrt{M_X} - \sqrt{m_X})^2$ on $\mathsf{E}\, X - \exp \mathsf{E} \ln X$ was presented as [5, Problem 11800]. Note that $2V_X = 2 \operatorname{\mathsf{Var}} \sqrt{X}$ can be rewritten as $\mathsf{E}(\sqrt{X} - \sqrt{\tilde{X}}\,)^2$, where $\tilde{X}$ is an independent copy of the r.v. X. Therefore, the upper bound in (1.2) is strictly less than that in [5] unless $\operatorname{supp} X = \{m_X, M_X\}$. In the case when X is a continuous function on the interval $[0, 1]$, the latter condition on $\operatorname{supp} X$ simply means that X is a constant, and then the difference $\mathsf{E}\, X - \exp \mathsf{E} \ln X$ and the upper bound on it in (1.2) (as well as the lower one) are each 0.

Given any nonnegative real numbers $x_1, \ldots, x_n$, let X be any r.v. with the distribution defined by the formula

$$\mathsf{E} f(X) = \frac{1}{n} \sum_{i=1}^{n} f(x_i) \quad \text{for any function } f : \mathbb{R} \to \mathbb{R}. \tag{1.11}$$

(So, in the case when the numbers $x_1, \ldots, x_n$ are pairwise distinct, any such r.v. X takes each of the values $x_1, \ldots, x_n$ with probability $1/n$.) In this case,

$$\mathsf{E}\, X = \frac{x_1 + \cdots + x_n}{n} \quad \text{and} \quad \exp \mathsf{E} \ln X = \sqrt[n]{x_1 \cdots x_n}.$$

Thus, for any r.v. X with $\mathsf{E}\, X < \infty$, the terms $\mathsf{E}\, X$ and $\exp \mathsf{E} \ln X$ in (1.2) can be referred to, respectively, as the arithmetic and geometric means of the r.v. X. Since any

https://doi.org/10.1017/S0004972715000350 Published online by Cambridge University Press

bounded nonnegative r.v. can be approximated in distribution by uniformly bounded r.v.s each taking finitely many nonnegative real values with equal probabilities, the upper and lower bounds in (1.2) will each remain exact in an appropriate sense if one considers only the r.v.s with such discrete uniform distributions. In particular, one has the following immediate corollary from Theorem 1.4 and Remark 1.5.

Corollary 1.6. *For any $n \in \mathbb{N}$, any $z = (z_1, \dots, z_n) \in \mathbb{R}^n$ and any function $f : \mathbb{R} \to \mathbb{R}$, let*

$$\overline{z} := \frac{1}{n}(z_1 + \cdots + z_n), \qquad z^{\mathsf{g}} := \sqrt[n]{|z_1 \cdots z_n|},$$
$$z_{\max} := z_1 \vee \cdots \vee z_n, \quad z_{\min} := z_1 \wedge \cdots \wedge z_n, \quad f(z) := (f(z_1), \dots, f(z_n)).$$

Then, for any real V, E, F such that $0 < V < E \wedge F$,

$$\sup\{\overline{x} - x^{\mathsf{g}} : x = y^2, y \in \mathbb{R}_+^n, n \in \mathbb{N}, \overline{(y - \overline{y})^2} \leqslant V, \overline{(y - y_{\min})^2} \leqslant E\} = (2V) \vee E,$$
$$\inf\{\overline{x} - x^{\mathsf{g}} : x = y^2, y \in \mathbb{R}_+^n, n \in \mathbb{N}, \overline{(y - \overline{y})^2} \geqslant V, \overline{(y_{\max} - y)^2} \leqslant F\} = (2V) \wedge E_{V,F}.$$

The proof of Theorem 1.4, given in Section 2, relies on the theory of Tchebycheff–Markoff systems. Major expositions of this theory and its applications are given in the monographs by Karlin and Studden [4] and Kreĭn and Nudel'man [6]. A brief review of the theory, which contains all the definitions and facts necessary for the proof in the present paper, is given in [7]. A condensed version of [7] can be found in [8, Appendix A].

2. Proofs

Proof of Proposition 1.1. Take any $X \in \mathscr{X}_+$. Clearly, $E_X \geqslant V_X \geqslant 0$. If $V_X = 0$, then $\mathsf{P}(X = c) = 1$ for some $c \in [0, \infty)$, whence $E_X = 0$, so that $E_X = V_X = 0$. If $V_X > 0$, then $\mathsf{E}\sqrt{X} > \sqrt{m_X}$ and hence $E_X > V_X > 0$. So, condition (1.5) is necessary for $\mathscr{X}_{\mathsf{sup};V,E} \neq \emptyset$. *Vice versa*, suppose now that (1.5) holds. For any real u and v such that $0 \leqslant u < v$ and any $p \in [0, 1]$, let $Y_{u,v,p}$ denote any r.v. such that

$$\mathsf{P}(Y_{u,v,p} = u) = p = 1 - \mathsf{P}(Y_{u,v,p} = v). \tag{2.1}$$

If $E = V = 0$, then $0 \in \mathscr{X}_{\mathsf{sup};V,E}$ and so $\mathscr{X}_{\mathsf{sup};V,E} \neq \emptyset$. If now $E > V > 0$, let $X = Y_{u,v,p}^2$ with

$$p = \frac{V}{E} \quad \text{and any } u \text{ and } v \text{ such that } 0 \leqslant u < v \text{ and } v - u = \frac{E}{\sqrt{E - V}}. \tag{2.2}$$

Then $X \in \mathscr{X}_{\mathsf{sup};V,E}$ and so $\mathscr{X}_{\mathsf{sup};V,E} \neq \emptyset$ in this case as well. Thus, the equivalence of the condition $\mathscr{X}_{\mathsf{sup};V,E} \neq \emptyset$ and (1.5) is checked. The equivalence of the condition $\mathscr{X}_{\mathsf{inf};V,F} \neq \emptyset$ and (1.6) is checked quite similarly; here, in the case when $F > V > 0$, (2.2) is replaced by

$$q := 1 - p = \frac{V}{F} \quad \text{and any } u \text{ and } v \text{ such that } 0 \leqslant u < v \text{ and } v - u = \frac{F}{\sqrt{F - V}}. \tag{2.3}$$

Thus, Proposition 1.1 is proved. □

https://doi.org/10.1017/S0004972715000350 Published online by Cambridge University Press

Before proceeding to the proofs of Propositions 1.2 and 1.3, let us state the following observation.

Lemma 2.1. *Take any r.v. Z such that $\mathsf{E}\,Z = 0$ and $\operatorname{supp} Z \subseteq [c, d]$ for some real c and d. Then $c \leqslant 0 \leqslant d$, $\mathsf{Var}\,Z \leqslant |c|d$ and $\mathsf{Var}\,Z = |c|d$ if and only if $\operatorname{supp} Z = \{|c|, d\}$.*

This follows immediately on noting that $c \leqslant \mathsf{E}\,Z = 0 \leqslant d$ and $\mathsf{Var}\,Z = \mathsf{E}\,Z^2 = \mathsf{E}(Z - c)(Z - d) - cd \leqslant -cd = |c|d$.

Being very simple, Lemma 2.1 seems to be a piece of common mathematical lore. For example, the inequality $\mathsf{Var}\,Z \leqslant |c|d$ in Lemma 2.1 follows immediately from [2, Lemma 2.2], by shifting and rescaling. In the case when Z has a discrete distribution of the form given by (1.11), Lemma 2.1 was presented as Theorem 1 and the second part of Proposition 1 in [1].

Proof of Proposition 1.2. Suppose that $\mathscr{X}_{\mathsf{inf};V,F} \neq \emptyset$ and take any $X \in \mathscr{X}_{\mathsf{inf};V,F}$. Let $Y := \sqrt{X}$, $a := m_Y$ and $b := M_Y$. By Lemma 2.1 with $Z := Y - \mathsf{E}\,Y$, $c := a - \mathsf{E}\,Y$ and $d := b - \mathsf{E}\,Y$,

$$E_X = \mathsf{E}(Y - a)^2 = \mathsf{Var}\,Y + (a - \mathsf{E}\,Y)^2 \geqslant \mathsf{Var}\,Y + \frac{(\mathsf{Var}\,Y)^2}{(b - \mathsf{E}\,Y)^2} = E_{V_X,F_X} = E_{V,F} \tag{2.4}$$

provided that $\infty > F > V$, with the inequality in (2.4) turning into the equality if and only if $\operatorname{supp} Y = \{m_Y, M_Y\}$, that is, if and only if $\operatorname{supp} X = \{m_X, M_X\}$. This verifies Proposition 1.2 in the case when $\infty > F > V$.

If now $F = V$, then, by Proposition 1.1, $F = V = 0$. In this case, by the convention, $E_{V,F} = 0$. On the other hand, for any $X \in \mathscr{X}_{\mathsf{inf};V,F}$, one has $\operatorname{supp} X = c$ for some $c \in \mathbb{R}$, which implies that $E_X = 0$. So, Proposition 1.2 holds as well in the case when $F = V$.

Consider the remaining case, with $F = \infty$. Then, by the convention, $E_{V,F} = V$. For each $\varepsilon \in (0, 1)$, let U_ε be any r.v. whose distribution is (a mixture of a Bernoulli distribution and an exponential distribution) defined by the condition that

$$\mathsf{E}\,f(U_\varepsilon) = (1 - \varepsilon)f(0) + (\varepsilon - \varepsilon^2)f(1) + \varepsilon^2 \int_0^\infty f(x)e^{-x}\,dx$$

for all nonnegative Borel functions f on $\mathbb{R}$. Then $\mathsf{E}\,U_\varepsilon = \varepsilon = \mathsf{Var}\,U_\varepsilon$ and $F_{U_\varepsilon} = \infty$. Let now $X_\varepsilon := (V/\varepsilon)\,U_\varepsilon^2$. Then $X_\varepsilon \in \mathscr{X}_{\mathsf{inf};V,\infty} = \mathscr{X}_{\mathsf{inf};V,F}$ and $E_{X_\varepsilon} = (1 + \varepsilon)V$. So,

$$\inf\{E_X : X \in \mathscr{X}_{\mathsf{inf};V,F}\} \leqslant \inf\{(1 + \varepsilon)V : \varepsilon \in (0, 1)\} = V = E_{V,F}.$$

On the other hand,

$$E_X = V_X + (m_X - \mathsf{E}\,X)^2 \geqslant V_X = V = E_{V,F} \tag{2.5}$$

for all $X \in \mathscr{X}_{\mathsf{inf};V,F}$. Now (1.7) follows as well in the case $F = \infty$. However, in this case the infimum in (1.7) is not attained. Indeed, otherwise the inequality in (2.5) would for some $X \in \mathscr{X}_{\mathsf{inf};V,F}$ turn into an equality, which would imply that $\mathsf{E}\,X = m_X$ and hence $F_X = 0$, which would contradict the assumption $F = \infty$. Thus, Proposition 1.2 is completely verified. □

https://doi.org/10.1017/S0004972715000350 Published online by Cambridge University Press

Proof of Proposition 1.3. The ‘if’ side of Proposition 1.3 is quite straightforward to check. Let us verify the ‘only if’ side. Suppose that the inequalities in (1.2) turn simultaneously into equalities, so that the upper and lower bounds there are equal to each other, which is in turn equivalent to the statement that

$$E_X \leqslant 2V_X \leqslant \frac{F_X V_X}{F_X - V_X}. \tag{2.6}$$

If $F_X = V_X$, then, by Proposition 1.1, $V_X = 0$ and hence $\operatorname{supp} X = \{c\}$ for some $c \in [0,\infty)$, that is, the distribution of $\sqrt{X}$ is the (necessarily) symmetric distribution on the singleton set $\{\sqrt{c}\} \subset [0,\infty)$.

It remains to consider the case $F_X > V_X$. Then the double inequality (2.6) can be rewritten as $2V_X \geqslant E_X \vee V_X$, which can be further rewritten as

$$\mathsf{Var}\, Y \geqslant \max[(\mathsf{E}\, Y - a)^2, (b - \mathsf{E}\, Y)^2],$$

where $Y := \sqrt{X}$, $a := m_Y$ and $b := M_Y$, so that $a \leqslant \mathsf{E}\, Y \leqslant b$. Therefore,

$$2\,\mathsf{Var}\, Y \geqslant (\mathsf{E}\, Y - a)^2 + (b - \mathsf{E}\, Y)^2 \geqslant 2(\mathsf{E}\, Y - a)(b - \mathsf{E}\, Y) \geqslant 2\,\mathsf{Var}\, Y, \tag{2.7}$$

where the last inequality follows by Lemma 2.1 (with $Z = Y - \mathsf{E}\, Y$). Hence, all the inequalities in (2.7) are actually equalities. In particular, the equality $(\mathsf{E}\, Y - a)^2 + (b - \mathsf{E}\, Y)^2 = 2(\mathsf{E}\, Y - a)(b - \mathsf{E}\, Y)$ implies $\mathsf{E}\, Y = (a + b)/2$. Also, again by Lemma 2.1, the equality $2(\mathsf{E}\, Y - a)(b - \mathsf{E}\, Y) = 2\,\mathsf{Var}\, Y$ implies $\operatorname{supp} Y = \{a, b\}$. This, together with the condition $\mathsf{E}\, Y = (a + b)/2$, shows that the distribution of the r.v. $Y = \sqrt{X}$ is the symmetric distribution on the set $\{a, b\} \subset [0,\infty)$. This completes the proof of Proposition 1.3. □

The proof of Theorem 1.4 will be preceded by more notation and two lemmas. Take any a and b such that $0 < a < b < \infty$ and introduce

$$Q_{\mathsf{sup};V,E} := \{(\beta_1, \beta_2) \in (0,\infty)^2 : \beta_2 - \beta_1^2 = V,\ \beta_2 - 2a\beta_1 + a^2 = E\},$$
$$Q_{\mathsf{inf};V,F} := \{(\beta_1, \beta_2) \in (0,\infty)^2 : \beta_2 - \beta_1^2 = V,\ \beta_2 - 2b\beta_1 + b^2 = F\}$$

and then

$$\mathscr{Y}_{\beta_1,\beta_2} := \{Y \in \mathscr{X}_+ : \operatorname{supp} Y \subseteq [a, b],\ \mathsf{E}\, Y = \beta_1,\ \mathsf{E}\, Y^2 = \beta_2\},$$
$$S_{\beta_1,\beta_2} := \sup\{D_{Y^2} : Y \in \mathscr{Y}_{\beta_1,\beta_2}\},$$
$$I_{\beta_1,\beta_2} := \inf\{D_{Y^2} : Y \in \mathscr{Y}_{\beta_1,\beta_2}\}$$

for $(\beta_1, \beta_2) \in (0,\infty)^2$, with the definition of D_X in (1.8) in mind; for brevity, the dependence on a and b is not made explicit in this notation.

Lemma 2.2. *Take any* $(\beta_1, \beta_2) \in Q_{\mathsf{sup};V,E}$ *such that* $\mathscr{Y}_{\beta_1,\beta_2} \neq \emptyset$. *Then*

$$S_{\beta_1,\beta_2} \leqslant (2V) \vee E.$$

Lemma 2.3. *Take any* $(\beta_1, \beta_2) \in Q_{\mathsf{inf};V,F}$ *such that* $\mathscr{Y}_{\beta_1,\beta_2} \neq \emptyset$. *Then*

$$I_{\beta_1,\beta_2} \geqslant (2V) \wedge E_{V,F}.$$

https://doi.org/10.1017/S0004972715000350 Published online by Cambridge University Press

Proof of Lemma 2.2. Note that

$$S_{\beta_1,\beta_2} = \beta_2 - \exp(2I_{\ln;\beta_1,\beta_2}), \quad \text{where } I_{\ln;\beta_1,\beta_2} := \inf\{\mathsf{E}\ln Y : Y \in \mathscr{Y}_{\beta_1,\beta_2}\}. \tag{2.8}$$

Using [7, Proposition 1], it is easy to see that the sequence of functions $(1, \#, \#^2, \ln\#)$ is an M_+-system on $[a, b]$. Hence, by [7, part (II)(a) of Proposition 2] (with $n = 2$), the infimum $I_{\ln;\beta_1,\beta_2}$ is attained at a r.v. of the form $Y = Y_{u,v,p} \in \mathscr{Y}_{\beta_1,\beta_2}$ with $0 < u = a < v < \infty$ and $p \in [0, 1]$, whose distribution is defined by (2.1). These conditions on $Y_{u,v,p}$, u and v, together with the condition $(\beta_1, \beta_2) \in Q_{\mathsf{sup};V,E}$, allow one to express u, v, p, $D_{Y^2_{u,v,p}}$, $V_{Y^2_{u,v,p}}$ and $E_{Y^2_{u,v,p}}$ uniquely in terms of a, V and E, in accordance with (2.2):

$$u = a, \quad v = u + \frac{E}{\sqrt{E - V}}, \quad p = \frac{V}{E}, \tag{2.9}$$

$$D_{Y^2_{u,v,p}} = pu^2 + qv^2 - u^{2p}v^{2q}, \tag{2.10}$$

$$V_{Y^2_{u,v,p}} = pq(v - u)^2 = V, \quad E_{Y^2_{u,v,p}} = q(v - a)^2 = q(v - u)^2 = E, \tag{2.11}$$

where

$$q := 1 - p.$$

It follows that

$$S_{\beta_1,\beta_2} = \psi(0) \leqslant \sup_{c\in[-u,\infty)} \psi(c), \quad \text{where} \tag{2.12}$$

$$\psi(c) := D_{Y^2_{u,v,p}} = p(u + c)^2 + q(v + c)^2 - (u + c)^{2p}(v + c)^{2q} \tag{2.13}$$

and u, v and p are as in (2.9) (cf. (2.10)). The supremum in (2.12) is easy to find, and it depends only on V and E. Indeed,

$$\psi'''(c) = 4pq(p - q)(v - u)^3(u + c)^{2p-3}(v + c)^{2q-3}$$

equals $p - q$ in sign for all $c \in (-u, \infty)$. To find, for each $j \in \{0, 1, 2\}$, the limit $\psi^{(j)}(\infty-)$ of the derivative $\psi^{(j)}(c)$ as $c \to \infty$, for any $\gamma \in \mathbb{R}$ write $(v + c)^\gamma = (u + c)^\gamma(1 + \varepsilon)^\gamma$, where $\varepsilon := (v - u)/(u + c) \sim (v - u)/c \to 0$, and then write

$$(1 + \varepsilon)^\gamma = \sum_{i=0}^{2-j} \gamma(\gamma - 1)\cdots(\gamma - i + 1)\frac{\varepsilon^i}{i!} + o(c^{j-2}).$$

Thus, one finds $\psi(\infty-) = 2pq(u - v)^2 = 2V$ and $\psi'(\infty-) = \psi''(\infty-) = 0$. Therefore and because ψ''' equals $p - q$ in sign, one sees that ψ' equals $p - q$ in sign, on the interval $(-u, \infty)$, which implies that the function ψ is monotonic on the interval $[-u, \infty)$, with $\psi(-u) = q(u - v)^2 = E$ and $\psi(\infty-) = 2V$. Thus, the supremum in (2.12) equals $(2V) \vee E$, which completes the proof of Lemma 2.2. □

Proof of Lemma 2.3. This proof is similar to that of Lemma 2.2. Here, instead of the infimum $I_{\ln;\beta_1,\beta_2}$ defined in (2.8), one deals with $S_{\ln;\beta_1,\beta_2} := \sup\{\mathsf{E}\ln Y : Y \in \mathscr{Y}_{\beta_1,\beta_2}\}$.

https://doi.org/10.1017/S0004972715000350 Published online by Cambridge University Press

This supremum is attained at a r.v. of the form $Y = Y_{u,v,p} \in \mathscr{Y}_{\beta_1,\beta_2}$ with

$$v = b, \quad u = v - \frac{F}{\sqrt{F-V}}, \quad q = 1 - p = \frac{V}{F},$$

$$E_{Y^2_{u,v,p}} = q(b-u)^2 = q(v-u)^2 = \frac{V}{F}\Big(\frac{F}{\sqrt{F-V}}\Big)^2 = \frac{VF}{F-V} = E_{V,F}$$

and $D_{Y^2_{u,v,p}}$ as in (2.10), $V_{Y^2_{u,v,p}} = V$ as in (2.11) and $F_{Y^2_{u,v,p}} = p(v-u)^2 = F$. The proof of Lemma 2.3 is concluded with the observation that $\inf_{c\in[-u,\infty)} \psi(c) = (2V) \wedge E_{V,F}$ (cf. the last sentence in the proof of Lemma 2.2). □

Proof of Theorem 1.4. Suppose that $\mathscr{X}_{\mathrm{sup};V,E} \neq \emptyset$, so that condition (1.5) holds. Both sides of (1.9) are obviously 0 if $E = V = 0$. To verify (1.9) in the remaining case $E > V > 0$, fix any $X_* \in \mathscr{X}_{\mathrm{sup};V,E}$. Consider first the case

$$a := \sqrt{m_{X_*}} > 0 \quad \text{and} \quad b := \sqrt{M_{X_*}} < \infty. \tag{2.14}$$

Letting now $Y_* := \sqrt{X_*}$ and $(\beta_1^*, \beta_2^*) := (\mathsf{E}\, Y_*, \mathsf{E}\, Y_*^2)$, one has $(\beta_1^*, \beta_2^*) \in Q_{\mathrm{sup};V,E}$ and $Y_* \in \mathscr{Y}_{\beta_1^*,\beta_2^*}$. Also, $D_{X_*} = D_{Y_*^2} \leqslant S_{\beta_1^*,\beta_2^*} \leqslant (2V) \vee E$, by Lemma 2.2. So,

$$D_{X_*} \leqslant (2V) \vee E \tag{2.15}$$

for any r.v. $X_* \in \mathscr{X}_{\mathrm{sup};V,E}$ satisfying conditions (2.14).

If a r.v. $X_* \in \mathscr{X}_{\mathrm{sup};V,E}$ is such that $m_{X_*} = 0$, then $D_{X_*} \leqslant \mathsf{E}\, X_* = E_{X_*} = E \leqslant (2V) \vee E$, so that inequality (2.15) still holds.

Take now any r.v. $X_* \in \mathscr{X}_{\mathrm{sup};V,E}$ such that $m_{X_*} > 0$ and $M_{X_*} = \infty$. Take any $t \in (m_{X_*}, \infty)$ and let $X_t := X_* \wedge t$, so that $M_{X_t} \leqslant t < \infty$, whence, by (2.15) with X_t in place of X_*, one has $D_{X_t} \leqslant (2V_{X_t}) \vee E_{X_t}$. On the other hand, by dominated convergence with $t \to \infty$, one has $V_{X_t} \to V_{X_*} = V$, $E_{X_t} \to E_{X_*} = E$, $\mathsf{E}\, X_t \to \mathsf{E}\, X_*$ and $\mathsf{E} \ln X_t \to \mathsf{E} \ln X_*$ and so $D_{X_t} \to D_{X_*}$.

Thus, inequality (2.15) holds for all $X_* \in \mathscr{X}_{\mathrm{sup};V,E}$. That is,

$$S_{V,E} \leqslant (2V) \vee E,$$

in the case $E > V > 0$, where $S_{V,E}$ is as in (1.9). On the other hand, again in the case $E > V > 0$, for any u, v, p as in (2.2) and any $c \in [-u, \infty)$, the r.v. $Y^2_{u+c,v+c,p}$ is in $\mathscr{X}_{\mathrm{sup};V,E}$ and so

$$S_{V,E} \geqslant \sup_{c\in[-u,\infty)} \psi(c) = (2V) \vee E, \tag{2.16}$$

with $\psi(c)$ as in (2.13). This concludes the proof of (1.9).

The proof of (1.10) is similar. Suppose that $\mathscr{X}_{\mathrm{inf};V,F} \neq \emptyset$, so that condition (1.6) holds. Both sides of (1.10) are obviously 0 if $F = V = 0$. Consider the remaining case $F > V > 0$.

Fix any $X_* \in \mathscr{X}_{\mathrm{inf};V,F}$. Consider first the case when conditions (2.14) hold.

Letting now $Y_* := \sqrt{X_*}$ and $(\beta_1^*, \beta_2^*) := (\mathsf{E}\, Y_*, \mathsf{E}\, Y_*^2)$, one has $(\beta_1^*, \beta_2^*) \in Q_{\mathrm{inf};V,F}$ and $Y_* \in \mathscr{Y}_{\beta_1^*,\beta_2^*}$. Also, $D_{X_*} = D_{Y_*^2} \geqslant I_{\beta_1^*,\beta_2^*} \geqslant (2V) \wedge E_{V,F}$, by Lemma 2.3. So,

$$D_{X_*} \geqslant (2V) \wedge E_{V,F} \tag{2.17}$$

for any r.v. $X_* \in \mathscr{X}_{\mathrm{inf};V,F}$ satisfying conditions (2.14).

Take now any s and t such that $0 < s < t < \infty$ and let $X_{s,t} := s \vee (t \wedge X_*)$, so that conditions (2.14) are satisfied with $X_{s,t}$ in place of X_*. Hence, one will have $D_{X_{s,t}} \geqslant (2V_{X_{s,t}}) \wedge E_{V_{X_{s,t}},F_{X_{s,t}}}$. Let now $s \downarrow 0$ and $t \uparrow \infty$. Then $X_{s,t} \to X_*$ pointwise, $m_{X_{s,t}} \to m_{X_*}$ and $M_{X_{s,t}} \to M_{X_*}$. By dominated convergence, $\mathsf{E}\, X_{s,t} \to \mathsf{E}\, X_*$ and $V_{X_{s,t}} \to V_{X_*} = V$. If $F_{X_*} < \infty$, then $F_{X_{s,t}} \to F_{X_*}$, again by dominated convergence. If $F_{X_*} = \infty$, then clearly $F_{X_{s,t}} \leqslant F_{X_*}$. Thus, in any case, $\limsup F_{X_{s,t}} \leqslant F_{X_*} = F$. Moreover, by the Fatou lemma, $\mathsf{E} \ln X_* \leqslant \liminf \mathsf{E} \ln X_{s,t}$, whence $D_{X_*} \geqslant \limsup D_{X_{s,t}} \geqslant \limsup[(2V_{X_{s,t}}) \wedge E_{V_{X_{s,t}},F_{X_{s,t}}}] \geqslant (2V) \wedge E_{V,F}$, since $E_{V,F}$ is nonincreasing in F and continuous in (V, F) such that $F > V > 0$.

Thus, inequality (2.17) holds for all $X_* \in \mathscr{X}_{\mathsf{inf};V,F}$. That is,

$$I_{V,F} \geqslant (2V) \wedge E_{V,F},$$

in the case $F > V > 0$, where $I_{V,F}$ is as in (1.10). On the other hand, again in the case $F > V > 0$, for any u, v, p as in (2.3) and any $c \in [-u, \infty)$, the r.v. $Y^2_{u+c,v+c,p}$ is in $\mathscr{X}_{\mathsf{inf};V,F}$ and so

$$I_{V,F} \leqslant \inf_{c\in[-u,\infty)} \psi(c) = (2V) \wedge E_{V,F},$$

with $\psi(c)$ still as in (2.13). This concludes the proof of (1.10).

Concerning the last sentence of Theorem 1.4, let $\mathscr{X}_{\mathsf{sup},2;V,E}$ denote the set of all r.v.s in $\mathscr{X}_{\mathsf{sup};V,E}$ taking at most two values, and then let $S_{2;V,E} := \sup\{D_X : X \in \mathscr{X}_{\mathsf{sup},2;V,E}\}$. Suppose that $\mathscr{X}_{\mathsf{sup};V,E} \neq \emptyset$, as is done in (1.9), so that (1.5) holds.

If $E = V = 0$, then $S_{V,E} = 0$ and, on the other hand, $0 \in \mathscr{X}_{\mathsf{sup},2;V,E}$ and hence $0 = D_0 \leqslant S_{2;V,E} \leqslant S_{V,E} = 0$, so that $S_{2;V,E} = S_{V,E} = (2V) \vee E$.

Suppose now that $E > V > 0$. Then, for any u, v, p as in (2.2) and any $c \in [-u, \infty)$, one has $Y^2_{u,v,p} \in \mathscr{X}_{\mathsf{sup},2;V,E}$ and, hence, by (2.16) and (2.13), $(2V) \vee E = \sup_{c\in[-u,\infty)} \psi(c) = \sup_{c\in[-u,\infty)} D_{Y^2_{u,v,p}} \leqslant S_{2;V,E} \leqslant S_{V,E} = (2V) \vee E$ and so the conclusion $S_{2;V,E} = S_{V,E} = (2V) \vee E$ holds.

That is, the equality in (1.9) holds if the set $\mathscr{X}_{\mathsf{sup};V,E}$ is replaced there by $\mathscr{X}_{\mathsf{sup},2;V,E}$. The corresponding statement concerning the equality in (1.10) and the set $\mathscr{X}_{\mathsf{inf};V,F}$ is verified quite similarly.

Thus, Theorem 1.4 is completely proved. □

References

[1] R. Bhatia and C. Davis, 'A better bound on the variance', *Amer. Math. Monthly* **107**(4) (2000), 353–357.

[2] S. W. Dharmadhikari and K. Joag-Dev, 'Upper bounds for the variances of certain random variables', *Comm. Statist. Theory Methods* **18**(9) (1989), 3235–3247.

[3] S. S. Dragomir, 'Some reverses of the Jensen inequality with applications', *Bull. Aust. Math. Soc.* **87**(2) (2013), 177–194.

[4] S. Karlin and W. J. Studden, *Tchebycheff Systems: With Applications in Analysis and Statistics*, Pure and Applied Mathematics, XV (Interscience–John Wiley, New York–London–Sydney, 1966).

[5] O. Klurman, 'Problem 11800', *Amer. Math. Monthly* **121**(8) (2014), 739.

[6] M. G. Kreĭn and A. A. Nudel'man, *The Markov Moment Problem and Extremal Problems* (American Mathematical Society, Providence, RI, 1977), [Ideas and problems of P. L. Čebyšev and A. A. Markov and their further development. Translated from the Russian by D. Louvish, Translations of Mathematical Monographs, Vol. 50].

https://doi.org/10.1017/S0004972715000350 Published online by Cambridge University Press

[7] I. Pinelis, 'Tchebycheff systems and extremal problems for generalized moments: a brief survey', arXiv:1107.3493, 2011.
[8] I. Pinelis, 'An asymptotically Gaussian bound on the Rademacher tails', *Electron. J. Probab.* **17** (2012), 1–22.

IOSIF PINELIS, Department of Mathematical Sciences,
Michigan Technological University, Houghton, Michigan 49931, USA
e-mail: ipinelis@mtu.edu

https://doi.org/10.1017/S0004972715000350 Published online by Cambridge University Press