A MONOTONICITY PROPERTY OF VARIANCES

J. M. ALDAZ

ABSTRACT. We prove that variances of non-negative random variables have the following monotonicity property: For all $0 < r < s \leq 1$, and all $0 \leq X \in L^2$, we have $\text{Var}(X^r)^{1/r} \leq \text{Var}(X^s)^{1/s}$. We also discuss the real valued case.

1. Introduction

Here, statements such as $X \geq 0$ or $X = Y$, are always meant in the almost sure sense. It is immediate from either Hölder’s or Jensen’s inequality that for every random variable $X \geq 0$ and all $0 < r < s < \infty$, we have $(EX^r)^{1/r} \leq (EX^s)^{1/s}$. In this note we obtain an analogous result for non-negative random variables $X \in L^2$ and variances. As in the case of norms, this inequality helps to clarify the strength of hypotheses that might be made on $\text{Var}(X^r)$. An application to a recent refinement of the AM-GM inequality $\prod_{i=1}^n x_i^{\alpha_i} \leq \sum_{i=1}^n \alpha_i x_i$ is presented. Lastly, this monotonicity property can be used when dealing with real valued random variables, by decomposing them into their positive and negative parts, since the variance of $X$ is always comparable to the sum of the variances of $X_+$ and $X_-$. 

2. Monotonicity of $\text{Var}(X^s)^{1/s}$, and the AM-GM inequality.

Let $0 \leq X \in L^2$, so $\text{Var}(X)$ is well defined. Since for all $0 < s \leq 1$ we have $\|X\|_{2s} \leq \|X\|_2$, all variances $\text{Var}(X^s)$ are also well defined, and thus it is natural to ask how these quantities behave as $s$ changes. In order to be able to compare them, we need to have the same homogeneity on both sides of the inequality, so we consider $\text{Var}(X^s)^{1/s}$, which always is homogeneous of order 2: For all $t \geq 0$, $\text{Var}((tX)^s)^{1/s} = t^2 \text{Var}(X^s)^{1/s}$.

Theorem 2.1. Let $0 \leq X \in L^2$ and let $0 < r < s \leq 1$. Then

(1) $\text{Var}(X^r)^{1/r} \leq \text{Var}(X^s)^{1/s}$.

Proof. Observe first that it is enough to prove the case $\text{Var}(X^s)^{1/s} \leq \text{Var}(X)$ whenever $0 < s < 1$. The fact that $\text{Var}(X^s)^{1/s}$ is increasing in $s$ then follows immediately by making the change of variables $Y = X^s$: $\text{Var}(X^r)^{s/r} = \text{Var}(Y^{r/s})^{s/r} \leq \text{Var}(Y) = \text{Var}(X^s)$.

Next, we assume that $\|X\|_2 = 1$. This can be done by homogeneity, since writing $Y = X/\|X\|_2$, we see that $\text{Var}(X^s)^{1/s} \leq \text{Var}(X)$ is equivalent to $\text{Var}(Y^{r/s})^{1/s} \leq \text{Var}(Y)$. Under the

Key words and phrases. Variance, Arithmetic-Geometric inequality.

The author was partially supported by Grant MTM2012-37436-C02-02 from the MINECO of Spain.

2010 Mathematical Subject Classification. 60A10, 26D15.
condition $\|X\|_2 = 1$, we always have, for every $0 < s \leq 1$ and every $t > 0$, $\|X\|_2^t \leq 1$, and hence, $\text{Var}(X^s)^t \leq 1$.

We shall use the following well known (and direct) interpolation consequence of Hölder’s inequality (cf., for instance, [Fo, Proposition 6.10, p. 177]) which is valid for both finite and infinite measure spaces: If $0 < r < s < p$, and $f \in L^r \cap L^p$, then $f$ belongs to all intermediate spaces $L^t$, and furthermore, $\|f\|_s \leq \|f\|_r^{1-t} \|f\|_p^t$, where $t \in (0, 1)$ is defined by the equation $1/s = (1 - t)/r + t/p$.

Using the indices $0 < s < 2s < 2$, together with $\|X\|_2 = 1$, yields

$$t = 1/(2 - s)$$

(2)

$$E(X^{2s}) \leq (EX^s)^{(2-2s)/(2-s)},$$

while the indices $0 < s < 1 < 2$ give $t = (2 - 2s)/(2 - s)$ and

(3)

$$E(X) \leq (EX^s)^{1/(2-s)}.$$ 

Now, by the preceding assumptions on the size of norms and variances (in particular, by $\|X^s\|_2^2 = \|X\|_2^2 \leq 1$) together with $1/s > 1$, we have

$$\text{Var}(X^s)^{1/s} \leq \text{Var}(X^s) = \|X^s\|_2^2 \text{Var} \left( \frac{X^s}{\|X^s\|_2} \right) \leq \text{Var} \left( \frac{X^s}{\|X^s\|_2} \right) = 1 - \frac{(EX^s)^2}{E(X^{2s})}.$$

Thus, it suffices to show that

$$1 - \frac{(EX^s)^2}{E(X^{2s})} \leq \text{Var}(X) = 1 - (EX)^2,$$

or equivalently, that

$$(EX)^2 E(X^{2s}) \leq (EX^s)^2.$$

But this follows from (3) and (2), since

$$(EX)^2 E(X^{2s}) \leq (EX^s)^{2/(2-s)} (EX^s)^{(2-2s)/(2-s)} = (EX^s)^2.$$

□

Remark 2.2. The interpolation result noted above is useful in a probability context since, instead of the usual bound $\|X\|_s \leq \|X\|_p$, whenever $0 < s < p$, it yields the stronger inequality $\|X\|_s \leq \|X\|_r^t \|X\|_p^{1-t}$ for each $0 < r < s$, with $t$ defined by $1/s = (1 - t)/r + t/p$.

Of course, under different integrability conditions ($X \in L^p$ instead of $X \in L^2$) the analogous inequalities hold, by using the change of variables $Y = X^{p/2} \in L^2$.

Corollary 2.3. Let $p > 0$, let $0 \leq X \in L^p$, and let $0 < r < s \leq p/2$. Then

(4)

$$\text{Var}(X^r)^{1/r} \leq \text{Var}(X^s)^{1/s}.$$

Next we apply the preceding result to a recent refinement of the inequality between arithmetic and geometric means (the AM-GM inequality) proven in [A1] (the reader interested in some probabilistic aspects of the AM-GM inequality, may want to consult [A3] and the references contained therein; for non-variance bounds, see [A4] and its references).
Let us recall the notation used in [A1]: $X$ denotes the vector with non-negative entries $(x_1, \ldots, x_n)$, and $X^{1/2} = (x_1^{1/2}, \ldots, x_n^{1/2})$. Given a sequence of weights $\alpha = (\alpha_1, \ldots, \alpha_n)$ with $\alpha_i > 0$ and $\sum_{i=1}^n \alpha_i = 1$, and a vector $Y = (y_1, \ldots, y_n)$, its arithmetic mean is denoted by $E_\alpha(Y) := \sum_{i=1}^n \alpha_i y_i$, its geometric mean, by $\Pi_\alpha(Y) := \prod_{i=1}^n y_i^{\alpha_i}$, and its variance, by

$$\text{Var}_\alpha(Y) = \sum_{i=1}^n \alpha_i \left( y_i - \frac{\sum_{k=1}^n \alpha_k y_k}{\sum_{k=1}^n \alpha_k} \right)^2 = \sum_{i=1}^n \alpha_i y_i^2 - \left( \sum_{k=1}^n \alpha_k y_k \right)^2.$$

Finally, $Y_{\max}$ and $Y_{\min}$ respectively stand for the maximum and the minimum values of $Y$.

Conceptually, variance bounds for $E_\alpha X - \Pi_\alpha X$ represent the natural extension of the equality case in the AM-GM inequality (zero variance is equivalent to equality). From a more applied viewpoint, the variance is used in the Economics literature to estimate the difference between these means (cf., for instance, [Si, Chapter 1, Appendix 2]; both the arithmetic and geometric means are used when reporting on the performance of a portfolio).

The bounds for the difference in the AM-GM appearing in [A1] involve $\text{Var}(X^{1/2})$, rather than $\sigma(X) = \text{Var}_\alpha(X)^{1/2}$. Using Theorem 2.1 or Corollary 2.3, the following upper bound follows: $E_\alpha X - \Pi_\alpha X \leq \frac{1}{\alpha_{\min}} \sigma(X)$. More generally, by putting together [A1, Theorem 4.2] with Corollary 2.3, we obtain the next result.

**Theorem 2.4.** For $n \geq 2$ and $i = 1, \ldots, n$, let $X = (x_1, \ldots, x_n)$ be such that $x_i \geq 0$, and let $\alpha = (\alpha_1, \ldots, \alpha_n)$ satisfy $\alpha_i > 0$ and $\sum_{i=1}^n \alpha_i = 1$. Then for all $r \in (0, 1]$ and all $s \in [1, \infty)$ we have

$$\frac{1}{1 - \alpha_{\min}} \text{Var}_\alpha(X^{r/2})^{1/r} \leq E_\alpha X - \Pi_\alpha X \leq \frac{1}{\alpha_{\min}} \text{Var}_\alpha(X^{s/2})^{1/s}. \tag{5}$$

These bounds are optimal (cf. [A1, Examples 2.1 and 2.3]). Theorem 4.2 from [A1], and its proof, were suggested by [CaFi, Theorem], which states that if $0 < X_{\min}$, then

$$\frac{1}{2X_{\max}} \text{Var}_\alpha(X) \leq E_\alpha X - \Pi_\alpha X \leq \frac{1}{2X_{\min}} \text{Var}_\alpha(X). \tag{6}$$

A drawback of (6) is that the bounds depend explicitly on $X_{\max}$ and $X_{\min}$, something that makes it unsuitable for some standard applications, such as, for instance, refining Hölder’s inequality (see [A1] for more details). Of course, since the variance is homogeneous of degree 2, dividing by $X_{\max}$ and $X_{\min}$ in (6), gives the left and right hand sides the same homogeneity as the middle term. We also point out that the inequality $\text{Var}_\alpha(X^{1/2}) \leq E_\alpha X - \Pi_\alpha X$, appeared in [A2, Theorem 1]; this inequality is trivial, useful, and as $n \to \infty$, asymptotically optimal, since $(1 - \alpha_{\min})^{-1} \to 1$.

3. REAL VALUED RANDOM VARIABLES.

The monotonicity result applies to $X \geq 0$ only: If $X < 0$ with positive probability, then $X^*$ may fail to be defined as a real valued function, for certain values of $s > 0$. While trivially $\text{Var}(X) \geq \text{Var}(|X|)$, in general these two quantities are not comparable, so it is not possible to simply replace $X$ with $|X|$. However, monotonicity can be used on $\text{Var}(X_+)$ and $\text{Var}(X_-)$,
where \( X_+ := \max \{X, 0\} \) and \( X_- := -\min \{X, 0\} \) denote the positive and negative parts of \( X \), respectively. Thus, indirectly it also applies to \( \text{Var}(X) \), since the latter is indeed comparable to \( \text{Var}(X_+) + \text{Var}(X_-) \). We have not found this result in the literature, so we include it here for completeness. Essentially, the next theorem says that

\[
\text{Var}(X_+) + \text{Var}(X_-) \leq \text{Var}(X) \leq 2(\text{Var}(X_+) + \text{Var}(X_-)),
\]

and the extremal cases occur, for the left hand side inequality, when either \( X \geq 0 \) or \( X \leq 0 \), and for the right hand side inequality, when \( X = c(1_D - 1_{D^c}) \), where \( c \in \mathbb{R} \) and \( D \) is a measurable set.

**Theorem 3.1.** Let \( X \in L^2 \) be real valued, and denote by \( B \) the sub-\( \sigma \)-algebra

\[
B := \{0, \Omega, \{X > 0\}, \{X = 0\}, \{X < 0\}\}.
\]

Then

\[
(7) \quad \text{Var}(X_+) + \text{Var}(X_-) \leq \text{Var}(X)
\]

\[
(8) \quad \leq \text{Var}(X_+) + \text{Var}(E(X_+|B)) + \text{Var}(E(X_-|B)) \leq 2(\text{Var}(X_+) + \text{Var}(X_-)).
\]

Furthermore, equality holds in the first inequality if and only if either \( X \geq 0 \) or \( X \leq 0 \); in the second, if and only if either \( X > 0 \), or \( X < 0 \), or \( 0 < P(\{X > 0\}) < P(\{X < 0\}) \), \( 0 = P(\{X > 0\}) \), and \( E(X_+|\{X > 0\}) = E(X_-|\{X < 0\}) \); and in the third, if and only if \( X = E(X|B) \).

**Proof.** The first inequality follows directly from the definitions, the second, from the convexity of \( \phi(x) = x^2 \), and the third, from the law of total variance. More precisely,

\[
\text{Var}(X_+) + \text{Var}(X_-) \leq \text{Var}(X_+) + \text{Var}(X_-) + 2EX_+EX_- = E(X_+^2) - (EX_+)^2 + E(X_-^2) - (EX_-)^2 + 2EX_+EX_- = E(X^2) - (EX_+ - EX_-)^2 = \text{Var}(X),
\]

and we have equality if and only if \( EX_+EX_- = 0 \), which happens if and only if either \( X \geq 0 \) or \( X \leq 0 \).

Since, as we just saw, \( \text{Var}(X) = \text{Var}(X_+) + \text{Var}(X_-) + 2EX_+EX_- \), to prove the middle inequality in (7)-(8), it is enough to show that

\[
(9) \quad 2EX_+EX_- \leq \text{Var}(E(X_+|B)) + \text{Var}(E(X_-|B)).
\]

Observe that if either \( X \geq 0 \) or \( X \leq 0 \), then

\[
2EX_+EX_- = 0,
\]

and if additionally either \( X > 0 \) or \( X < 0 \), then

\[
0 = \text{Var}(E(X_+|B)) + \text{Var}(E(X_-|B)).
\]

Next, assume that both \( A := P\{X > 0\} > 0 \) and \( B := P\{X < 0\} > 0 \), and write \( C := P\{X = 0\} \), so \( 0 < A + B = 1 - C \leq 1 \). Then \( E(X|B) \) takes exactly two values different from 0, say \( E(X|B) = a > 0 \) on \( \{X > 0\} \), and \( E(X|B) = -b < 0 \) on \( \{X < 0\} \). With this notation, in order to obtain the middle inequality it suffices to show that

\[
2EX_+EX_- = 2Ab \leq \text{Var}(E(X_+|B)) + \text{Var}(E(X_-|B)) = Aa^2 - (Aa)^2 + Bb^2 - (Bb)^2,
\]
or equivalently, that

\[(Aa + Bb)^2 \leq Aa^2 + Bb^2.\]

But this is follows from the convexity of \(\phi(x) = x^2\), since

\[(Aa + Bb)^2 = (A + B)^2 \left( \frac{A}{A + B} a + \frac{B}{A + B} b \right)^2 \leq (A + B)^2 \left( \frac{A}{A + B} a^2 + \frac{B}{A + B} b^2 \right)
\]

\[= (A + B) (Aa^2 + Bb^2) \leq Aa^2 + Bb^2.\]

Furthermore, \((Aa + Bb)^2 = Aa^2 + Bb^2\) if and only if both \(a = b\) (by the strict convexity of \(\phi\)) and \(A + B = 1\).

Finally, the law of total variance \(\text{Var}(X) = \text{Var}(E(X|\mathcal{B}))+E(\text{Var}(X|\mathcal{B}))\), applied to both \(X_+\) and \(X_-\), tells us that \(\text{Var}(X_+) \geq \text{Var}(E(X_+|\mathcal{B}))\) and \(\text{Var}(X_-) \geq \text{Var}(E(X_-|\mathcal{B}))\), with equality if and only if \(E(\text{Var}(X_+|\mathcal{B})) = 0 = E(\text{Var}(X_-|\mathcal{B}))\), which happens if and only if both \(X_+\) and \(X_-\) are constant on \(\{X > 0\}\) and on \(\{X < 0\}\) respectively. This yields the last inequality, together with the equality condition \(X = E(X|\mathcal{B})\). \(\square\)

**Remark 3.2.** Instead of \(\mathcal{B} = \{\emptyset, \Omega, \{X > 0\}, \{X = 0\}, \{X < 0\}\}\), either of the simpler algebras \(\mathcal{B}_1 = \{\emptyset, \Omega, \{X \geq 0\}, \{X < 0\}\}\) or \(\mathcal{B}_2 = \{\emptyset, \Omega, \{X > 0\}, \{X \leq 0\}\}\) could have been used in the preceding theorem, and the inequalities stated there would still hold. But the equality conditions would be less symmetric. For instance, if \(X \geq 0\), then \(\mathcal{B}_1\) is trivial up to sets of measure zero (that is, as a measure algebra), so \(E(X_+|\mathcal{B}_1) = EX_+ = EX\), and \(\text{Var}(E(X_+|\mathcal{B}_1)) = 0\). Thus, the middle inequality in (7)-(8), is actually an equality in this case. However, if \(X = -1_D \leq 0\), where \(0 < P(D) < 1\), then \(X = X_- = E(X_-|\mathcal{B}_1)\), and \(\text{Var}(X) < \text{Var}(X_-) + \text{Var}(E(X_-|\mathcal{B}_1)) = 2 \text{Var}(X)\).

**Corollary 3.3.** Let \(p \geq 2\), let \(X \in L^p\) be real valued, and let \(0 < r \leq 2 \leq s \leq p\). Then

\[
\text{Var}(X_+^{r/2})^{2/r} + \text{Var}(X_-^{r/2})^{2/r} \leq \text{Var}(X) \leq 2 \left( \text{Var}(X_+^{s/2})^{2/s} + \text{Var}(X_-^{s/2})^{2/s} \right). 
\]

**References**


Departamento de Matemáticas, Universidad Autónoma de Madrid, Cantoblanco 28049, Madrid, Spain.

E-mail address: jesus.munarriz@uam.es