Statistics and Probability Letters 122 (2017) 86–89

Contents lists available at ScienceDirect

Statistics and Probability Letters

journal homepage: www.elsevier.com/locate/stapro

A self-improvement to the Cauchy–Schwarz inequality

Stephen G. Walker

Department of Mathematics, University of Texas at Austin, TX, USA

ARTICLE INFO

Article history:
Received 7 October 2016
Received in revised form 27 October 2016
Accepted 1 November 2016
Available online 17 November 2016

Keywords:
Cauchy–Schwarz
Cramer–Rao inequality
Wasserstein distance

ABSTRACT

We present a self improvement to the Cauchy–Schwarz inequality, which in the probability case yields

$$[E(XY)]^2 \leq E(X^2)\,E(Y^2) - \left(|E(X)|\sqrt{\mathrm{Var}(Y)} - |E(Y)|\sqrt{\mathrm{Var}(X)}\right)^2.$$

It is to be noted that the additional term to the inequality only involves the marginal first two moments for X and Y, and not any joint property. We also provide the discrete improvement to the inequality.

1. Introduction

The Cauchy–Schwarz inequality is arguably one of the most widely used inequalities in mathematics; see Steele (2004). Let $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$ be two real sequences; then

$$\left(\sum_{i=1}^{n} x_i y_i\right)^2 \leq \sum_{i=1}^{n} x_i^2 \sum_{i=1}^{n} y_i^2 \tag{1}$$

with equality if and only if there exists a constant c such that $x_i = c\,y_i$ for all $i = 1, \ldots, n$. There are a number of proofs available in the literature and textbooks.

There are refinements of (1), such as Alzer (1992), but are quite specific in the details on the x and y. A survey of the inequality is given in Dragomir (2003) and a recent generalization is presented in Tuo (2015).

The aim in this paper is to provide a new and elementary (self) improvement of the Cauchy–Schwarz inequality which only uses the terms

$$\sum_{i=1}^{n} x_i, \qquad \sum_{i=1}^{n} x_i^2, \qquad \sum_{i=1}^{n} y_i, \quad \text{and} \quad \sum_{i=1}^{n} y_i^2.$$

These are the first two marginal moments for x and y. The self improvement is similar in spirit to one appearing in Walker (2014), which worked on the Jensen inequality, and also in Bobkov et al. (2014), which worked on the logarithmic Sobolev inequality.

Other improvements of the Cauchy–Schwarz inequality use more than these first two moments; see for example Aldaz (2009), which use joint properties of x and y.

E-mail address: s.g.walker@math.utexas.edu.

http://dx.doi.org/10.1016/j.spl.2016.11.001
0167-7152/

In the following we will use $\bar{x}$ and $\bar{y}$ to denote $n^{-1}\sum_{i=1}^n x_i$ and $n^{-1}\sum_{i=1}^n y_i$, respectively, and V_x and V_y to denote

$$\sum_{i=1}^n x_i^2 - n\bar{x}^2 \quad \text{and} \quad \sum_{i=1}^n y_i^2 - n\bar{y}^2,$$

respectively. It is clear that V_x and V_y are nonnegative. The new result is that

$$\left(\sum_{i=1}^n x_i y_i\right)^2 \le \sum_{i=1}^n x_i^2 \sum_{i=1}^n y_i^2 - n\left(|\bar{x}|\sqrt{V_y} - |\bar{y}|\sqrt{V_x}\right)^2. \tag{2}$$

We prove this in Section 2 where we also present the improved inequality for the probability setting. Section 3 contains some applications.

2. Proof of (2)

Now (2) is a self-improvement of the Cauchy–Schwarz inequality. To see this, for some as yet unspecified h_1 and h_2, we have

$$\left(\sum_{i=1}^n (x_i - h_1)(y_i - h_2)\right)^2 \le \sum_{i=1}^n (x_i - h_1)^2 \sum_{i=1}^n (y_i - h_2)^2,$$

so

$$\left(\sum_{i=1}^n x_i y_i - h_1 n\bar{y} - h_2 n\bar{x} + n h_1 h_2\right)^2 \le \left(\sum_{i=1}^n x_i^2 - 2h_1 n\bar{x} + n h_1^2\right)\left(\sum_{i=1}^n y_i^2 - 2h_2 n\bar{y} + n h_2^2\right).$$

Setting $h_1 h_2 = h_1\bar{y} + h_2\bar{x}$, to isolate the $\sum_{i=1}^n x_i y_i$ term from the left side, and after some algebra on the right side, we arrive at

$$\left(\sum_{i=1}^n x_i y_i\right)^2 \le \sum_{i=1}^n x_i^2 \sum_{i=1}^n y_i^2 - n h_1 V_y[2\bar{x} - h_1] - n h_2 V_x[2\bar{y} - h_2].$$

Now setting $h_1 = h$ and $h_2 = h\bar{y}/(h - \bar{x})$ for some h, we have

$$h_1 V_y[2\bar{x} - h_1] + h_2 V_x[2\bar{y} - h_2] = h(2\bar{x} - h)\left[V_y - V_x\frac{\bar{y}^2}{(h - \bar{x})^2}\right]. \tag{3}$$

Putting $h = \bar{x} + t$ the aim now would be to maximize,

$$F(t) = (\bar{x}^2 - t^2)(V_y - V_x\bar{y}^2/t^2).$$

Noting that $F(\infty) = -\infty$ and $F(0+) = -\infty$, we see the maximizing t^2 is given by $\widehat{t^2} = |\bar{x}\bar{y}|\sqrt{V_x/V_y}$, and so (3) becomes

$$\left(|\bar{x}|\sqrt{V_y} - |\bar{y}|\sqrt{V_x}\right)^2.$$

This completes the proof.

In the special case when $x_i = 1$ for all i, we obtain an equality in (2).

Corollary. *It is that*

$$\frac{1}{2}\sum_{i=1}^n\sum_{j=1}^n (x_i y_j - x_j y_i)^2 \ge n\left(|\bar{x}|\sqrt{V_y} - |\bar{y}|\sqrt{V_x}\right)^2.$$

The Lagrange identity (see Steele, 2004) has

$$\sum_{i=1}^n x_i^2 \sum_{i=1}^n y_i^2 - \left(\sum_{i=1}^n x_i y_i\right)^2 = \frac{1}{2}\sum_{i=1}^n\sum_{j=1}^n (x_i y_j - x_j y_i)^2$$

and hence the statement of the Corollary follows.

The probability version of the Cauchy–Schwarz inequality is given by

$$[E(XY)]^2 \le E(X^2)\,E(Y^2) \tag{4}$$

and is an equality if and only if $Y = cX$ for some constant c.

The self improvement for this inequality follows the same outline as in th proof of (2) and we omit the details. The new inequality is given by

$$[E(XY)]^2 \le E(X^2)\,E(Y^2) - \left(|E(X)|\sqrt{\mathrm{Var}(Y)} - |E(Y)|\sqrt{\mathrm{Var}(X)}\right)^2. \tag{5}$$