University of Pennsylvania
ScholarlyCommons

Statistics Papers

Wharton Faculty Research

11-2006

Several Colorful Inequalities

Ingram Olkin
Stanford University

Larry A. Shepp
University of Pennsylvania

Follow this and additional works at: https://repository.upenn.edu/statistics_papers

Part of the Statistics and Probability Commons

Recommended Citation

Olkin, I., & Shepp, L. A. (2006). Several Colorful Inequalities. *The American Mathematical Monthly, 113* (9), 817-822. http://dx.doi.org/10.2307/27642063

This paper is posted at ScholarlyCommons. https://repository.upenn.edu/statistics_papers/188
For more information, please contact repository@pobox.upenn.edu.

Several Colorful Inequalities

Abstract

For a positive random variable X, let $\mu_\alpha(X)$ be the αth moment of X. Mark Brown proves that for positive and independent random variables X, Y, F(X + Y) ≥ F(X) + F(Y), where *F* (X) is the ratio μ_{-1} over μ_{-2}. We prove this inequality and several generalizations by a method which can be used to prove the Schwarz inequality, but which is not widely appreciated.

Disciplines

Statistics and Probability

This journal article is available at ScholarlyCommons: https://repository.upenn.edu/statistics_papers/188

SEVERAL COLORFUL INEQUALITIES

by

Ingram Olkin
Larry Shepp

Technical Report No. 2004-24
October 2004

Department of Statistics
STANFORD UNIVERSITY
Stanford, California 94305-4065

SEVERAL COLORFUL INEQUALITIES

by

Ingram Olkin
Department of Statistics
Stanford University

Larry Shepp
Department of Statistics
Rutgers University

Technical Report No. 2004-24
October 2004

Department of Statistics
STANFORD UNIVERSITY
Stanford, California 94305-4065

http://www-stat.stanford.edu

Several Colorful Inequalities

Ingram Olkin

Stanford University

Larry Shepp

Rutgers University

e-mail: iolkin@stat.stanford.edu, shepp@stat.rutgers.edu

Abstract. For a positive random variable X, let $\mu_\alpha(X)$ be the $\alpha\underline{\text{th}}$ moment of X. Mark Brown proves that for positive and independent random variables X, Y, $F(X+Y) \geq F(X) + F(Y)$, where $F(X)$ is the ratio μ_{-1} over μ_{-2}. We prove this inequality and several generalizations by a method which can be used to prove the Schwarz inequality, but which is not widely appreciated.

1. Proof of the Schwarz inequality by the method of cloning.

The Schwarz inequality is better known and is more basic than the Brown inequality, but they can be proved by a similar method. The Schwarz inequality can be stated for a general measure space but easily reduces to the statement that

$$(EX^2)(EY^2) \geq (EXY)^2, \tag{1}$$

where X and Y are any random variables on a common probability space. Equality holds if and only if X and Y are proportional.

Brown (2001) proves two interesting inequalities that refer to *positive and independent* random variables on a common probability space:

$$\frac{E\frac{1}{X+Y}}{E(\frac{1}{X+Y})^2} \geq \frac{E\frac{1}{X}}{E\frac{1}{X^2}} + \frac{E\frac{1}{Y}}{E\frac{1}{Y^2}}, \tag{2}$$

$$\frac{1}{E\frac{1}{X+Y}} \geq \frac{1}{E\frac{1}{X}} + \frac{1}{E\frac{1}{Y}}. \tag{3}$$

Equality holds if and only if both X and Y are constants.

The Schwarz inequality can be proved by constructing the product probability space with the product measure so that X_1, Y_1 and X_2, Y_2 are two *independent* pairs of random variables on the product space each with the joint distribution of X, Y. Thus, (X_1, Y_1) and (X_2, Y_2) are independent clones. Note that for any four numbers, x_1, y_1, x_2, y_2, the homogeneous polynomial

$$x_1^2 y_2^2 + x_2^2 y_1^2 - 2x_1 y_1 x_2 y_2 = (x_1 y_2 - y_1 x_2)^2 \geq 0.$$

Now substitute X_i, Y_i for $x_i, y_i, i = 1, 2$ and take expectations, using the independence of the random variables with different subscripts to obtain that

$$2EX^2 EY^2 \geq 2(EXY)^2;$$

the proof is complete after dividing by 2. The only case of equality is when $X_1Y_2 - X_2Y_1 \equiv 0$, that is when the ratios X_i/Y_i are constant because they are independent for $i = 1, 2$.

Inequality (2) is next proved by the same method, but the proof is a bit more delicate. We "clear of fractions" by multiplying by $EX^{-2}EY^{-2}E(X+Y)^{-2}$, and so inequality (2) is equivalent to

$$E\frac{1}{X+Y}EX^{-2}EY^{-2} - E\frac{1}{(X+Y)^2}(EX^{-1}EY^{-2} + EX^{-2}EY^{-1}) \geq 0. \tag{4}$$

To use the method of cloning, we again construct the product probability space on which two independent pairs of independent random variables. X_i, Y_i are defined. If we could show that for any four numbers x_1, y_1, x_2, y_2 the function

$f(x_1, y_1, x_2, y_2) \equiv \frac{1}{x_1+y_1}\frac{1}{x_2^2}\frac{1}{y_2^2} - \frac{1}{(x_1+y_1)^2}(x_2^{-1}y_2^{-2} + x_2^{-2}y_2^{-1})$

is everywhere nonnegative, then it would follow that

$$Ef(X_1, Y_1, X_2, Y_2) \geq 0$$

which would then prove inequality (2). But (alas) f takes negative values. Alternatively, if we could show that $f(x_1, y_1, x_2, y_2) + f(x_2, y_2, x_1, y_1)$ is everywhere nonnegative, then the same proof would give the Brown inequality (2) because upon substituting random variables X_i, Y_i for x_i, y_i we would obtain the desired inequality after dividing by two. Again (alas), there are numbers $x_i, y_i, i = 1, 2$ for which this form is also negative. Fortunately, we have one last resort. If we can show that the doubly mixed (symmetric in x_1, x_2 and in y_1, y_2) form

$$(5) \qquad f(x_1, y_1, x_2, y_2) + f(x_1, y_2, x_2, y_1) + f(x_2, y_1, x_1, y_2) + f(x_2, y_2, x_1, y_1) \geq 0$$

for all positive values of $x_i, y_i, i = 1, 2$, then substituting X_i, Y_i for x_i, y_i, taking expecations and using the independence of X_1, X_2, Y_1, Y_2 we obtain the Brown inequality (2) after dividing by 4. We prove the above inequality, which is not as straight forward as it appears unless one does it in the right way, in the next section. In particular, the method of the proof not only provides a new proof of the Brown inequality (2), but also allows for some generalizations given in the following section. As to possible alternative proofs of the nonnegativity of the above form, Hardy, Littlewood and Pólya [HLP] discuss in Chapter 1 conditions under which a homogeneous form in positive variables is everywhere positive and [HLP] gives a necessary and sufficient condition on a homogeneous form in positive variables to be positive among other results, but it appears difficult to use any of their conditions to prove positivity for our particular form. One can expand our homogeneous form by clearing of fractions into a homogeneous polynomial (of eleventh degree) but there are so many terms that even with an algebraic package we were unable to see how to show the form is a positive one. [HLP] show that a sufficiently high power of $x_1 + y_1 + x_2 + y_2$ multiplied by epsilon plus our polynomial will have all nonnegative coefficients if our form is positive (which it is as is proved in the next section) but we have no idea how large this power will be (it seems to be quite large) so we could not use this condition which is necessary and sufficient. On the other hand choosing $x_i, y_i, i = 1, 2$ repeatedly at random with a random number generator left no doubt that the above form is a positive one. The usual rules of mathematical publication demand us to provide what it is usually considered a proof, though one could argue that there is a less chance for error in the Monte Carlo pseudo-proof based on a short program than in a mathematical proof. After some effort we found the proof in the next section and now consider it to be readily checkable. It is better not to expand the form into a polynomial.

2. Proof of the nonnegativity of the homogeneous form.

The proof is complicated but not difficult. Define

$$(6) \quad F(x_1, y_1, x_2, y_2) = f(x_1, y_1, x_2, y_2) + f(x_1, y_2, x_2, y_1) + f(x_2, y_1, x_1, y_2) + f(x_2, y_2, x_1, y_1)$$

and prove that $F \geq 0$ by exhibiting it as the sum of two squares. This type of proof of the positivity of a homogeneous form is discussed in [HLP, Chapter 1], but the present example is apparently not discussed there and it would seem that this paper contributes to the discussion in [HLP].

Indeed, we can write

(7)
$$F(x_1, y_1, x_2, y_2) = \frac{x_1 + y_1 - x_2 - y_2}{(x_1 + y_1)^2 x_2^2 y_2^2} + \frac{x_2 + y_1 - x_1 - y_2}{(x_1 + y_1)^2 x_1^2 y_2^2} + \frac{x_1 + y_2 - x_2 - y_1}{(x_1 + y_2)^2 x_2^2 y_1^2} + \frac{x_2 + y_2 - x_1 - y_1}{(x_2 + y_2)^2 x_1^2 y_1^2}.$$

Gathering terms with the factor $x_1 - x_2$ and $y_1 - y_2$, we can write (7) as

$$F(x_1, y_1, x_2, y_2) = (x_1 - x_2)\left[\frac{1}{(x_1 + y_1)^2 x_2^2 y_2^2} - \frac{1}{(x_2 + y_1)^2 x_1^2 y_2^2} + \frac{1}{(x_1 + y_2)^2 x_2^2 y_1^2} - \frac{1}{(x_2 + y_2)^2 x_1^2 y_1^2}\right]$$

(8)
$$+(y_1 - y_2)\left[\frac{1}{(x_1 + y_1)^2 x_2^2 y_2^2} - \frac{1}{(x_1 + y_2)^2 x_2^2 y_1^2} + \frac{1}{(x_2 + y_1)^2 x_1^2 y_2^2} - \frac{1}{(x_2 + y_2)^2 x_1^2 y_1^2}\right].$$

Now we can factor out another $x_1 - x_2$ and $y_1 - y_2$, respectively, to yield

(9)
$$F(x_1, y_1, x_2, y_2) = (x_1 - x_2)^2\left[\frac{1}{y_2^2}\frac{y_1((x_2+y_1)x_1+(x_1+y_1)x_2)}{(x_1+y_1)^2(x_2+y_1)^2 x_1^2 x_2^2} + \frac{1}{y_1^2}\frac{y_2((x_2+y_2)x_1+(x_1+y_2)x_2)}{(x_1+y_2)^2(x_2+y_2)^2 x_1^2 x_2^2}\right] +$$
$$(y_1 - y_2)^2\left[\frac{1}{x_2^2}\frac{x_1((x_1+y_2)y_1+(x_1+y_1)y_2)}{(x_1+y_1)^2(x_1+y_2)^2 y_1^2 y_2^2} + \frac{1}{x_1^2}\frac{x_2((x_2+y_2)y_1+(x_2+y_1)y_2)}{(x_2+y_2)^2(x_2+y_1)^2 y_1^2 y_2^2}\right].$$

This bit of algebra shows the needed result and proves the Brown inequality (2). Because $F(x_1, y_1, x_2, y_2) = 0$ only when $x_1 = x_2$ and $y_1 = y_2$ it follows that equality holds in (2) if and only if the random variables X and Y are both constant. Examples of non-independent X, Y can be constructed where the Brown inequality (2) fails so that the assumption of independence, used in the proof, is also necessary.

3. A First Generalization.

We now investigate whether functions other than moments can be considered.

To answer this, we prove that if $g = g(x) > 0$ is log convex, that is, if for $x > 0$, $d^2 log\ g(x)/dx^2 \geq 0$, then for any pair of independent random variables X, Y which are each *strictly positive-valued,*

(10)
$$\frac{E(X+Y)g(X+Y)}{Eg(X+Y)} \geq \frac{EXg(X)}{Eg(X)} + \frac{EYg(Y)}{Eg(Y)}.$$

Proof. For any four positive numbers x_1, y_1, x_2, y_2, define

$f(x_1, y_1, x_2, y_2) = (x_1+y_1)g(x_1+y_1)g(x_2)g(y_2) - g(x_1+y_1)[x_2 g(x_2)g(y_2) + g(x_2)y_2 g(y_2)]$.

It is enough to prove that

$f(x_1, y_1, x_2, y_2) + f(x_1, y_2, x_2, y_1) + f(x_2, y_1, x_1, y_2) + f(x_2, y_2, x_1, y_1) \geq 0$.

After dividing by the positive quantity $g(x_1)g(y_1)g(x_2)g(y_2)$, algebra similar to (7) - (9) reduces the form to:

(11)

$$(x_1 - x_2)^2\{[\frac{g(x_1+y_1)}{g(x_1)} - \frac{g(x_2+y_1)}{g(x_2)}]\frac{1}{(x_1-x_2)g(y_1)} + [\frac{g(x_1+y_2)}{g(x_1)} - \frac{g(x_2+y_2)}{g(x_2)}]\frac{1}{(x_1-x_2)g(y_2)}\}+$$
$$+(y_1 - y_2)^2\{[\frac{g(x_1+y_1)}{g(y_1)} - \frac{g(x_1+y_2)}{g(y_2)}]\frac{1}{(y_1-y_2)g(x_1)} + [\frac{g(x_2+y_1)}{g(y_1)} - \frac{g(x_2+y_2)}{g(y_2)}]\frac{1}{(y_1-y_2)g(x_2)}\}.$$

Each of the four terms inside the curly brackets will be positive if for all $a_1 \geq a_2 > 0$,

$$\frac{g(a_1+b)}{g(a_1)} \geq \frac{g(a_2+b)}{g(a_2)}.$$

But this means that $g(x+y)/g(x) \uparrow$ in x for each y, or

$$\frac{g'(x+y)}{g(x)} \geq \frac{g'(x)g(x+y)}{g^2(x)} \geq 0,$$

for each $y > 0$. Rearranging terms shows that an equivalent condition is that for $y > 0$,

$$\frac{g'(x+y)}{g(x+y)} \geq \frac{g'(x)}{g(x)}$$

or that $g'(x)/g(x)$ increases in x, which is equivalent to log convexity of $g(x)$.

Again, equality holds if and only if X and Y are both constant random variables.

The case $g(x) = 1/x^2$ gives Brown's inequality (2), and the case $g(x) = 1/x$ gives Brown's inequality (3).

4. A further generalization.

Our most general inequality is: if $g(x) > 0$ is log convex, $h(x) > 0$ is superadditive, that is, $h(x+y) \geq h(x) + h(y)$, then for strictly positive random variables X and Y

$$\frac{E\ h(X+Y)g(X+Y)}{E\ g(X+Y)} \geq \frac{E\ h(X)g(X)}{E\ g(X)} + \frac{E\ h(Y)g(Y)}{E\ g(Y)}. \tag{12}$$

The proof is similar to the previous proof and requires using the superadditive inequality. Inequality (12) is sharp when $h(x) = x$ but not in general. The case $h(x) = x^s, 0 < x \leq 1$, yields the inequality.

$$\frac{E\ (X+Y)^{-r}}{E\ (X+Y)^{-r-s}} \geq \frac{EX^{-r}}{EX^{-r-s}} + \frac{EY^{-r}}{EY^{-r-s}},$$

which for $r = s = 1$ is inequality (2).

REMARK. It is immediate that inequalities (2), (3), (10), (11) can, by interation, be extended to n variables., as for example, from (10)

$$\frac{E(\sum X_i)g(\sum X_i)}{E\ g(\sum X_i)} \geq \sum \frac{E\ X_i\ g(X_i)}{E\ g(X_i)}.$$

5. A Schur convex inequality

For any nonnegative random variable X the function

$$\varphi(r_1, \ldots, r_n) = \prod_1^n EX^{r_i} \tag{13}$$

is Schur-convex. See Marshall and Olkin (1979) for a discussion of this result in the context of other moment inequalities. We now give a new proof of (13).

__**Proof.**__ It sufficies to prove (13) for $n = 2$ because we can move by steps of two. The function φ is Schur-convex if for $r \geq s$

$$\varphi(r, s) \leq \varphi(r - \Delta, s + \Delta), \tag{14}$$

for any Δ that maintains the order, that is, $r - \Delta \geq s + \Delta$. Let x_1 and x_2 be real numbers and form the function

$$f(x_1, x_2) = x_1^r\ x_2^s + x_1^s\ x_2^r - x_1^{r-\Delta} x_2^{s+\Delta} - x_1^{s+\Delta} x_2^{r-\Delta} \tag{15}$$

$$= x_1^r x_2^r (1 - z^{\Delta})(1 - z^{r-s-\Delta}),$$

where $z = x_2/x_1$. Because $r - s - \Delta \geq \Delta \geq 0$, it follows that $f(x_1, x_2) \geq 0$. In (15) replace x_1 and x_2 by the independent identically distributed random variables X_1 and X_2, and take expectations to yield

(16) $$EX^r\, EX^s \geq EX^{r-\triangle} EX^{s+\triangle},$$

which is (14).

References

Brown, M. (2001). Exploiting the waiting time paradox: Applications of the renewal length transformation. Preprint, Accepted for publication in the Annals of Applied Probability.

Hardy, G. H., Littlewood, J. E. and Pólya, G. *Inequalities*, 1st ed., 2nd ed. Cambridge University Press, 1951.