Adv. Appl. Prob. **40,** 1223–1226 (2008)
Printed in Northern Ireland

LETTER TO THE EDITOR

Dear Editor,

On an inequality of Karlin and Rinott concerning weighted sums of i.i.d. random variables

This note delivers an entropy comparison result concerning weighted sums of independent and identically distributed (i.i.d.) random variables. The main result, Theorem 1, confirms a conjecture of Karlin and Rinott (1981).

For a continuous random variable X with density $f(x)$, $x \in \mathbb{R}$, the (differential) entropy is defined as

$$H(X) = -\int f(x) \log f(x) \, \mathrm{d}x$$

and the more general α-entropy, $\alpha > 0$, is defined as

$$H_\alpha(X) = \frac{1}{1-\alpha} \log G_\alpha(X),$$

where

$$G_\alpha(X) = \int f^\alpha(x) \, \mathrm{d}x. \tag{1}$$

It is convenient to define $H(X) = H_\alpha(X) = -\infty$ when X is discrete, e.g. degenerate. (Our notation differs from that of Karlin and Rinott (1981) here.)

We study the entropy of a weighted sum, $S = \sum_{i=1}^n a_i X_i$, of i.i.d. random variables X_i, assuming that the density f of X_i is *log-concave*, i.e. $\operatorname{supp}(f) = \{x : f(x) > 0\}$ is an interval and $\log f$ is a concave function on $\operatorname{supp}(f)$. The main result is that $H(S)$ (or $H_\alpha(S)$ with $0 < \alpha < 1$) is smaller when the weights $a_1, \dots, a_n$ are more 'uniform' in the sense of *majorization*. A real vector $\boldsymbol{b} = (b_1, \dots, b_n)^\top$ is said to majorize $\boldsymbol{a} = (a_1, \dots, a_n)^\top$, denoted $\boldsymbol{a} \prec \boldsymbol{b}$, if there exists a doubly stochastic matrix $\boldsymbol{T}$, i.e. an $n \times n$ matrix (t_{ij}) where $t_{ij} \ge 0$, $\sum_i t_{ij} = 1$, $j = 1, \dots, n$, and $\sum_j t_{ij} = 1$, $i = 1, \dots, n$, such that

$$\boldsymbol{T}\boldsymbol{b} = \boldsymbol{a}.$$

A function $\phi(\boldsymbol{a})$ symmetric in the coordinates of $\boldsymbol{a} = (a_1, \dots, a_n)^\top$ is said to be *Schur convex* if

$$\boldsymbol{a} \prec \boldsymbol{b} \quad \Longrightarrow \quad \phi(\boldsymbol{a}) \le \phi(\boldsymbol{b}).$$

Basic properties and various applications of these two notions can be found in Hardy *et al.* (1964) and Marshall and Olkin (1979).

Theorem 1. *Let $X_1, \dots, X_n$ be i.i.d. continuous random variables having a log-concave density on $\mathbb{R}$. Then $H(\sum_{i=1}^n a_i X_i)$ is a Schur convex function of $(a_1, \dots, a_n) \in \mathbb{R}^n$. The same holds for $H_\alpha(\sum_{i=1}^n a_i X_i)$ if $0 < \alpha < 1$.*

As an immediate consequence of Theorem 1, we have the following corollary.

Received 9 April 2008.

Corollary 1. *In the setting of Theorem 1, subject to a fixed $\sum_{i=1}^n a_i$, the entropy $H(\sum_{i=1}^n a_i X_i)$ is minimized when all the a_is are equal. The same holds if H is replaced by H_α with $\alpha \in (0, 1)$.*

Note that Corollary 1 and, hence, Theorem 1 need not hold without the assumption that the density of X_i is log-concave. For example, if $X_i \sim \text{gamma}(1/n, 1)$, i.e. a gamma distribution with shape parameter $1/n$, then the equally weighted $\sum_{i=1}^n X_i$, which has an exponential distribution, maximizes rather than minimizes the entropy H among $\sum_{i=1}^n a_i X_i$ with $\sum_{i=1}^n a_i = n$. For more entropy comparison results where log-concavity plays a role, see Yu (2008a), (2008b).

Karlin and Rinott (1981) conjectured Theorem 1 (their Remark 3.1) and proved a special case (their Theorem 3.1) assuming that (i) $a_i > 0$ and (ii) $f(x)$, the density of the X_is, is supported on $[0, \infty)$ and admits a Laplace transform of the form

$$\int_0^\infty \mathrm{e}^{-sx} f(x)\,\mathrm{d}x = \left(\prod_{i=1}^\infty (1 + \beta_i s)^{\alpha_i}\right)^{-1},$$

where $\alpha_i \geq 1$, $\beta_i \geq 0$, and $0 < \sum_{i=1}^\infty \alpha_i \beta_i < \infty$. Their proof of this special case, however, is somewhat complicated and does not extend easily when the additional assumptions are relaxed. A short proof of the general case is presented below.

We shall make use of the *convex order* between random variables. For random variables X and Y on $\mathbb{R}$ with finite means, we say that X is smaller than Y in the convex order, denoted $X \leq_{\text{cx}} Y$, if

$$\mathrm{E}\,\phi(X) \leq \mathrm{E}\,\phi(Y)$$

for every convex function ϕ. Properties of the convex order and many other stochastic orders can be found in Shaked and Shanthikumar (1994).

Lemma 1, below, relates the convex order and the log-concavity to entropy comparisons. The basic idea is due to Karlin and Rinott (1981). See Yu (2008b) for a discrete version that is used to compare the entropy between compound distributions on nonnegative integers.

Lemma 1. *Let X and Y be continuous random variables on $\mathbb{R}$. Assume that $X \leq_{\text{cx}} Y$ and that the density of Y is log-concave. Then $H(X) \leq H(Y)$ and $H_\alpha(X) \leq H_\alpha(Y)$, $0 < \alpha < 1$.*

Proof. Denote the density functions of X and Y by f and g, respectively. Note that because g is log-concave, $\mathrm{E}Y^2 < \infty$, which implies that $H(Y) < \infty$, as $H(Y)$ is bounded from above by the entropy of a normal variate with the same variance as Y. Also, $X \leq_{\text{cx}} Y$ implies that $\mathrm{E}X^2 \leq \mathrm{E}\,Y^2 < \infty$, which gives $H(X) < \infty$.

Using $X \leq_{\text{cx}} Y$ and Jensen's inequality, we obtain

$$\begin{aligned} H(Y) &= -\int g(x) \log g(x)\,\mathrm{d}x \\ &\geq -\int f(x) \log g(x)\,\mathrm{d}x \\ &\geq -\int f(x) \log f(x)\,\mathrm{d}x \\ &= H(X). \end{aligned}$$

All integrals are effectively over $\operatorname{supp}(g)$ as $X \leq_{\text{cx}} Y$ implies that f assigns zero mass outside of $\operatorname{supp}(g)$ when $\operatorname{supp}(g)$ is an interval.

To show that $H_\alpha(Y) \geq H_\alpha(X)$, we can equivalently show that $G_\alpha(Y) \geq G_\alpha(X)$, where G_α is given in (1). From the log-concavity of g and $\alpha < 1$, it follows that $(\alpha - 1)\log g$ and, hence, $g^{\alpha-1} = \exp[(\alpha-1)\log g]$ are convex. We may use this, $X \leq_{\text{cx}} Y$, and Hölder's inequality to obtain

$$\begin{aligned} G_\alpha(Y) &= \bigg(\int g(x) g^{\alpha-1}(x)\,\mathrm{d}x\bigg)^{\alpha} \bigg(\int g^{\alpha}(x)\,\mathrm{d}x\bigg)^{1-\alpha} \\ &\geq \bigg(\int f(x) g^{\alpha-1}(x)\,\mathrm{d}x\bigg)^{\alpha} \bigg(\int g^{\alpha}(x)\,\mathrm{d}x\bigg)^{1-\alpha} \\ &\geq \int f^{\alpha}(x)\,\mathrm{d}x \\ &= G_\alpha(X). \end{aligned}$$

Lemma 2, below, compares weighted sums of exchangeable random variables in the convex order.

Lemma 2. *Let X_i, $i = 1, \ldots, n$, be exchangeable random variables with a finite mean. Assume that $(a_1, \ldots, a_n) \prec (b_1, \ldots, b_n)$, $a_i, b_i \in \mathbb{R}$. Then*

$$\sum_{i=1}^{n} a_i X_i \leq_{\text{cx}} \sum_{i=1}^{n} b_i X_i.$$

Theorem 1 then follows from Lemmas 1 and 2 and the well-known fact that convolutions of log-concave densities are also log-concave.

Remark. Lemma 2 can be traced back to Marshall and Proschan (1965) (see also Eaton and Olshen (1972) and Bock *et al.* (1987)). When the X_is are i.i.d., Lemma 2 is given in Arnold and Villaseñor (1986) for $a_1 = \cdots = a_n = 1/n$, $b_1 = 0$, and $b_2 = \cdots = b_n = 1/(n-1)$, and in O'Cinneide (1991) for $a_1 = \cdots = a_n = 1/n$ and general b. Further discussions and generalizations of Lemma 2 can be found in Ma (2000). Some recent applications of Lemma 2 in the context of wireless communications can be found in Jorswieck and Boche (2007).

References

Arnold, B. C. and Villaseñor, J. A. (1986). Lorenz ordering of mean and medians. *Statist. Prob. Lett.* **4,** 47–49.

Bock, M. E., Diaconis, P., Huffer, H. W. and Perlman, M. D. (1987). Inequalities for linear combinations of gamma random variables. *Canad. J. Statist.* **15,** 387–395.

Eaton, M. L. and Olshen, R. A. (1972). Random quotients and the Behrens–Fisher problem. *Ann. Math. Statist.* **43,** 1852–1860.

Hardy, G. H., Littlewood, J. E. and Pólya, G. (1964). *Inequalities.* Cambridge University Press.

Jorswieck, E. and Boche, H. (2007). Majorization and matrix-monotone functions in wireless communications. *Foundations Trends Commun. Inf. Theory* **3,** 553–701.

Karlin, S. and Rinott, Y. (1981). Entropy inequalities for classes of probability distributions. I. The univariate case. *Adv. Appl. Prob.* **13,** 93–112.

Ma, C. (2000). Convex orders for linear combinations of random variables. *J. Statist. Planning Infer.* **84,** 11–25.

Marshall, A. W. and Olkin, I. (1979). *Inequalities: Theory of Majorization and Its Applications*. Academic Press, New York.

Marshall, A. W. and Proschan, F. (1965). An inequality for convex functions involving majorization. *J. Math. Anal. Appl.* **12,** 87–90.

O'Cinneide, C. A. (1991). Phase-type distributions and majorization. *Ann. Appl. Prob.* **1,** 219–227.

Shaked, M. and Shanthikumar, J. G. (1994). *Stochastic Orders and Their Applications*. Academic Press, New York.

YU, Y. (2008a). Log-concave distributions and a pair of triangle inequalities. Tech. Rep., Department of Statistics, University of California, Irvine.

YU, Y. (2008b). On the entropy of compound distributions on nonnegative integers. Tech. Rep., Department of Statistics, University of California, Irvine.

Yours sincerely,

YAMING YU*

* Department of Statistics,
University of California,
Irvine, CA, 92697-1250,
USA.
Email address: yamingy@ics.uci.edu

https://doi.org/10.1239/aap/1231340171 Published online by Cambridge University Press