SOME MAJORIZATION INEQUALITIES FOR FUNCTIONS OF EXCHANGEABLE RANDOM VARIABLES

By Philip J. Boland,[1] Frank Proschan,[2] and Y.L. Tong[3]

University College, Dublin, Florida State University, and Georgia Institute of Technology

This paper contains inequalities for the expectations of permutation-invariant concave functions and Schur-concave functions of the partial sums of nonnegative exchangeable random variables. Two majorization inequalities are derived, and an application in reliability theory is presented.

1. Introduction and Summary. For fixed $n > 1$ let $\mathbf{X} = (X_1, \ldots, X_n)$ denote an n-dimensional random vector with density function $f(\mathbf{x})$ that is absolutely continuous w.r.t. the Lebesgue measure or the product measure of counting measures. $X_1, \ldots, X_n$ are said to be exchangeable† if f is invariant under permutations of its arguments. This paper develops inequalities for the expectations of functions of partial sums of $X_1, \ldots, X_n$.

The notion of majorization defines a partial ordering of the diversity of the components of vectors. Let $\mathbf{a} = (a_1, \ldots, a_n)$, $\mathbf{b} = (b_1, \ldots, b_n)$ be two n-dimensional vectors and let $a_{[1]} \geq \cdots \geq a_{[n]}$, $b_{[1]} \geq \cdots \geq b_{[n]}$ denote their ordered components. $\mathbf{a}$ is said to *majorize* $\mathbf{b}$ (in symbols $\mathbf{a} \succ \mathbf{b}$) if

$$\Sigma_1^h a_{[i]} \geq \Sigma_1^h b_{[i]} \quad \text{for } h = 1, \ldots, n-1$$

and $\Sigma_1^n a_i = \Sigma_1^n b_i$. It is known that $\mathbf{a} \succ \mathbf{b}$ iff there exists a doubly stochastic matrix Q such that $\mathbf{b} = \mathbf{a}Q$, i.e., $\mathbf{b}$ is an "average" of $\mathbf{a}$. A function $\psi : R^n \to R$ is said to be a Schur-concave function if $\mathbf{a} \succ \mathbf{b}$ implies $\psi(\mathbf{a}) \leq \psi(\mathbf{b})$. For a comprehensive treatment of majorization and Schur functions, see Marshall and Olkin (1979).

---

[1]Research partially supported by the Air Force Office of Scientific Research, AFSC, USAF, under Grant AFOSR 88-0040.

[2]Research supported by the Air Force Office of Scientific Research, AFSC, USAF, under Grant AFOSR 88-0040.

[3]Research partially supported by NSF Grants DMS-8502346 and DMS-8801327.

AMS 1980 subject classifications. 60E15, 62H99.

Key words and phrases. Majorization inequalities, exchangeable random variables, concave and Schur-concave functions, moment inequalities.

† More precisely, $X_1, \ldots, X_n$ are *finitely* exchangeable instead of exchangeable. For the minor distinction between finite exchangeability and exchangeability see e.g., Tong ((1980), p. 96).

In an earlier paper Marshall and Proschan (1965) proved the following inequality: Let $X_1, \ldots, X_n$ be exchangeable and let $\phi : R^n \to R$ be Borel measurable, permutation invariant, and concave. If $a_1, \ldots, a_n$ are more diverse than $b_1, \ldots, b_n$ in the sense of majorization, i.e., if $\mathbf{a} \succ \mathbf{b}$, then $E\phi(b_1X_1, \ldots, b_nX_n) \geq E\phi(a_1X_1, \ldots, a_nX_n)$ provided the expectation exists. This inequality yields a number of useful results and implies many previously-known results as special cases (see, e.g., Corollaries 1–3 in their paper). In this paper, we prove some related results and discuss an application in reliability theory. The results (Theorems 1 and 2) involve the expectations of functions of partial sums of exchangeable random variables, and depend on the notion of majorization in a different fashion. For fixed $k < n$, let $\mathbf{r} = (r_1, \ldots, r_k)$ be a vector of positive integers such that $\Sigma_1^k r_i = n$. Let $X_1, \ldots, X_n$ be exchangeable random variables and let $\mathbf{Y_r} = (Y_1^{(\mathbf{r})}, \ldots, Y_k^{(\mathbf{r})})$ denote a k-dimensional random vector such that

$$Y_1^{(\mathbf{r})} = \Sigma_1^{r_1} X_i, \quad Y_2^{(\mathbf{r})} = \Sigma_{r_1+1}^{r_1+r_2} X_i, \ldots, \quad Y_k^{(\mathbf{r})} = \Sigma_{r_1+\cdots+r_{k-1}+1}^{n} X_i;$$

that is, $Y_j^{(\mathbf{r})}$ is the sum of r_j such X_i's and $Y_1^{(\mathbf{r})}, \ldots, Y_k^{(\mathbf{r})}$ do not contain any common elements. Let $\mathbf{s} = (s_1, \ldots, s_k)$ denote another such vector and $\mathbf{Y}^{(s)}$ be defined similarly. Let $\phi(\mathbf{y}) = \phi(y_1, \ldots, y_k)$ denote a real-valued function that is permutation invariant and concave. We show that (Theorem 1) if $\mathbf{s} \succ \mathbf{r}$ and if the X_i's are nonnegative exchangeable random variables, then $E\phi(\mathbf{Y}^{(\mathbf{r})}) \geq E\phi(\mathbf{Y}^{(\mathbf{s})})$. The reasons for considering such a random vector $\mathbf{Y}$ and for studying inequalities of this type arise from certain applications. One such application concerns the optimal arrangement policy for parallel and series systems in reliability theory, and is given in Section 4. In Theorem 2 we show that, by imposing an additional condition on the joint density f, the same inequality holds for all Schur-concave functions ϕ.

Since the theorems apply to nonnegative random variables only, a natural question is whether the same statements hold for random variables which may take negative values. We show in Section 3 that the answer is negative even for i.i.d. normal variables.

2. The Main Results. For the theorems stated in this section, the density function f of $\mathbf{X} = (X_1, \ldots, X_n)$ is assumed to be absolutely continuous w.r.t. the Lebesgue measure or the product measure of the counting measures. The proofs will be given for the former. For the product of counting measures, simply change the integral signs to summation signs.

THEOREM 1. *If (i) f is permutation invariant and $f = 0$ for any $x_i < 0$ $(i = 1, \ldots, n)$, (ii) $\phi(y_1, \ldots, y_k)$ is a permutation invariant concave function, and (iii) $\mathbf{s} \succ \mathbf{r}$, then*

$$E\phi(Y_1^{(\mathbf{r})}, \ldots, Y_k^{(\mathbf{r})}) \geq E\phi(Y_1^{(\mathbf{s})}, \ldots, Y_k^{(\mathbf{s})}) \tag{1}$$

holds provided that the expectations exist.

PROOF. It is well-known (Marshall and Olkin (1979), Chapter 2) that it suffices to assume that

$$s_1 > r_1 \geq r_2 > s_2 \equiv t, \quad r_1 + r_2 = s_1 + s_2 \equiv d$$

and $r_j = s_j$ for $j = 3, \ldots, k$. Let us define

$$Z_1 = \Sigma_1^t X_i, \quad Z_2 = \Sigma_{s_1+1}^d X_i$$

and $Y_j = Y_j^{(\mathbf{r})} = Y_j^{(\mathbf{s})}$ for $j = 3, \ldots, k$. Let

$$g(z_1, z_2) = g(z_1, z_2 \mid \mathbf{x}_0, y_3, \ldots, y_k) \tag{2}$$

denote the conditional density of (Z_1, Z_2) given $\mathbf{X}_0 \equiv (X_{t+1}, \ldots, X_{s_1}) = \mathbf{x}_0$ and $Y_j = y_j$ $(j = 3, \ldots, k)$. Then it is easy to check that $g(z_1, z_2) = g(z_2, z_1)$ and that

$$\begin{aligned}
& E[\phi(Y_1^{(\mathbf{r})}, \ldots, Y_k^{(\mathbf{r})}) \mid (\mathbf{X}_0, Y_3, \ldots, Y_k) = (\mathbf{x}_0, y_3, \ldots, y_k)] \\
= & \int\int \phi(z_1 + u_1, z_2 + u_2, y_3, \ldots, y_k) g(z_1, z_2) dz_1 dz_2 \\
= & \iint_{z_1 \geq z_2} \phi(z_1 + u_1, z_2 + u_2, y_3, \ldots, y_k) g(z_1, z_2) dz_1 dz_2 \\
& + \iint_{z_1 < z_2} \phi(z_1 + u_1, z_2 + u_2, y_3, \ldots, y_k) g(z_1, z_2) dz_1 dz_2 \\
= & \iint_{z_1 \geq z_2} \{\phi(z_1 + u_1, z_2 + u_2, y_3, \ldots, y_k) \\
& \qquad + \phi(z_1 + u_2, z_2 + u_1, y_3, \ldots, y_k)\} g(z_1, z_2) dz_1 dz_2
\end{aligned}$$

where $(u_1, u_2) = \left(\Sigma_{i=t+1}^{r_1} x_i, \Sigma_{i=r_1+1}^{s_1} x_i\right)$. Now let $(v_1, v_2) = \left(\Sigma_{i=t+1}^{s_1} x_i, 0\right)$. Since $x_i \geq 0$, there exists an $\alpha = \frac{u_1}{v_1} \in [0, 1]$ which satisfies

$$\begin{aligned}
(z_1 + u_1, z_2 + u_2) &= \alpha(z_1 + v_1, z_2 + v_2) + (1 - \alpha)(z_1 + v_2, z_2 + v_1), \\
(z_1 + u_2, z_2 + u_1) &= (1 - \alpha)(z_1 + v_1, z_2 + v_2) + \alpha(z_1 + v_2, z_2 + v_1)
\end{aligned}$$

for every point in $\{(z_1, z_2) : z_1 \geq z_2\}$. Thus for every fixed $(\mathbf{x}_0, y_3, \ldots, y_k)$ and every such (z_1, z_2),

$$\begin{aligned}
& \phi(z_1 + u_1, z_2 + u_2, y_3, \ldots, y_k) + \phi(z_1 + u_2, z_2 + u_1, y_3, \ldots, y_k) \\
\geq & \ \alpha\phi(z_1 + v_1, z_2 + v_2, y_3, \ldots, y_k) + (1 - \alpha)\phi(z_1 + v_2, z_2 + v_1, y_3, \ldots, y_k) \\
& + (1 - \alpha)\phi(z_1 + v_1, z_2 + v_2, y_3, \ldots, y_k) + \alpha\phi(z_1 + v_2, z_2 + v_1, y_3, \ldots, y_k).
\end{aligned}$$

Consequently, we have

$$\begin{aligned} & E\left[\phi(Y_1^{(\mathbf{r})},\ldots,Y_k^{(\mathbf{r})}) \mid (\mathbf{X}_0,Y_3,\ldots,Y_k)=(\mathbf{x}_0,y_3,\ldots,y_k)\right] \\ \geq & \iint_{z_1\geq z_2} \{\phi(z_1+v_1,z_2+v_2,y_3,\ldots,y_k)+ \\ & \qquad \phi(z_1+v_2,z_2+v_1,y_3,\ldots,y_k)\}g(z_1,z_2)dz_1dz_2 \\ = & E\left[\phi(Y_1^{(\mathbf{s})},\ldots,Y_k^{(\mathbf{s})}) \mid (\mathbf{X}_0,Y_3,\ldots,Y_k)=(\mathbf{x}_0,y_3,\ldots,y_k)\right], \end{aligned}$$

and the conclusion follows by unconditioning. ||

In the next theorem we change the condition on ϕ to be any measurable Schur-concave function, and impose a stronger condition on the conditional density g.

THEOREM 2. *If (i) f is permutation invariant, $f=0$ for any $x_i<0$ ($i=1,\ldots,n$), and such that the conditional density $g(z_1,z_2)$ defined in (2) is a Schur-concave function of (z_1,z_2) for every fixed $(\mathbf{x}_0,y_3,\ldots,y_k)$ and every $t>0$, (ii) $\phi(y_1,\ldots,y_k)$ is a Borel-measurable Schur-concave function, and (iii) $\mathbf{s}\succ\mathbf{r}$, then (1) holds provided the expectations exist.*

PROOF. We shall follow the notation developed in the proof of Theorem 1 and compare $E\phi(\mathbf{Y}^{(\mathbf{r})})$ with $E\phi(\mathbf{Y}^{(\mathbf{s})})$ for $\mathbf{s}\succ\mathbf{r}$. Again it suffices to assume that $s_1>r_1\geq r_2>s_2$ and $r_j=s_j$ for $j>2$. Then the conditional expectation of

$$\phi(Y_1^{(\mathbf{r})},Y_2^{(\mathbf{r})},Y_3,\ldots,Y_k)-\phi(Y_1^{(\mathbf{s})},Y_2^{(\mathbf{s})},Y_3,\ldots,Y_k)$$

given $(\mathbf{X}_0,Y_3,\ldots,Y_k)=(\mathbf{x}_0,y_3,\ldots,y_k)$ is

$$\Delta\equiv\int\int\{\phi^*(z_1+u_1,z_2+u_2)-\phi^*(z_1+u_1+u_2,z_2)\}g(z_1,z_2)dz_1dz_2$$

where

$$\phi^*(y_1,y_2)=\phi(y_1,y_2,y_3,\ldots,y_k)$$

and g is the conditional density of (Z_1,Z_2). It is straightforward to verify that, after following the same steps as in the proof of Theorem J.1 in Marshall and Olkin (1979, p. 100), we have

$$\Delta=\iint_{z_1\geq z_2}\{\phi^*(z_1,z_2+u_2)-\phi^*(z_1+u_2,z_2)\}\{g(z_1-u_1,z_2)-g(z_1,z_2-u_1)\}dz_1dz_2.$$

Since ϕ^* and g are Schur-concave functions and $u_i>0$ ($i=1,2$), we have

$$\begin{aligned} (z_1+u_2,z_2) &\succ (z_1,z_2+u_2), \\ (z_1,z_2-u_1) &\succ (z_1-u_1,z_2) \end{aligned}$$

and $\Delta\geq 0$. Thus the conclusion follows by unconditioning. ||

REMARK. Proschan and Sethuraman (1977) previously proved that if $X_1, \ldots, X_n$ are i.i.d. nonnegative random variables with a common density that is log-concave, then the conclusion in Theorem 2 holds. Their proof depends on an application of the main theorem in their paper and on a TP_2 property of the convolution of log-concave densities given in Karlin and Proschan (1960). It is noted here that their result now follows immediately from Theorem 2. This is so because if $X_1, \ldots, X_n$ are i.i.d. random variables with a common density that is log-concave, then $\Sigma_1^{s_2} X_i$ and $\Sigma_{s_1+1}^{s_1+s_2} X_i$ are independent random variables with a common density that is also log-concave (see e.g., Das Gupta (1973, Theorem 4.2)). Consequently, the joint density of $\left(\Sigma_1^{s_2} X_i, \Sigma_{s_1+1}^{s_1+s_2} X_i\right)$ is a Schur-concave function and Theorem 2 applies.

In most applications, the assumption on the Schur-concavity of the conditional density $g(z_1, z_2)$ is not easy to verify. It is clear that if the following conjecture concerning the convolution of Schur-concave random variables is true, then the assumption holds when f (the joint density of $\mathbf{X}$) is a Schur-concave function. We state the conjecture in a more general form without assuming that the random variables are nonnegative.

CONJECTURE. For $n = mk$ and $\mathbf{X} = (X_1, \ldots, X_n)$ let

$$Z_j = \Sigma_{(j-1)m+1}^{jm} X_i, \quad j = 1, 2, \ldots, k.$$

If the joint density of $\mathbf{X}$ is a Schur-concave function of $\mathbf{x}$ for $\mathbf{x} \in \Re^n$, then the joint density of $\mathbf{Z} = (Z_1, \ldots, Z_k)$ is a Schur-concave function of $\mathbf{z}$ for $\mathbf{z} \in \Re^k$ for all positive integers k and m.

It is not yet known to us whether this conjecture is true for continuous random variables. However, the following counterexample shows that at least for the discrete case, it is not true.

EXAMPLE. Consider $k = m = 2$, and assume that (X_1, X_2, X_3, X_4) takes only integer values $0, 1, 2, 3$. Let $Z_1 = X_1 + X_2$, $Z_2 = X_3 + X_4$. Then $P[Z_1 = 4,\ Z_2 = 2]$ is the probability of the set of the following points:

(3,1,1,1), (1,3,1,1), (2,2,1,1), (2,2,2,0) (2,2,0,2)
(3,1,2,0), (3,1,0,2), (1,3,2,0), (1,3,0,2)

Similarly $P[Z_1 = Z_2 = 3]$ is the probability of the set consisting of

(2,1,2,1), (2,1,1,2), (1,2,2,1), (1,2,1,2),
(2,1,3,0), (2,1,0,3), (1,2,3,0), (1,2,0,3),
(3,0,2,1), (3,0,1,2), (0,3,2,1), (0,3,1,2),
(3,0,3,0), (3,0,0,3), (0,3,3,0), (0,3,0,3).

If the joint density of (X_1, X_2, X_3, X_4) takes values of $1/14$ for each of the points (3,1,1,1), (2,2,2,0), (2,2,1,1) and each of their permutations, and zero otherwise, then it is a Schur-concave function on the product of integer space, and we have

$$P[Z_1 = 4, Z_2 = 2] = \frac{5}{14} > \frac{4}{14} = P[Z_1 = 3, Z_2 = 3]. \quad \|$$

3. An Example For Random Variables Which Are Not Nonnegative. It might be tempting to think that results similar to our Theorems 1 and 2 also hold when the condition that $X_i \geq 0$ a.s. is removed. This is not true even for i.i.d. normal variables. In the following, we give an example to show that the conclusion of Theorem 2 does not hold without this condition. An example for Theorem 1 can be obtained similarly.

Consider, for $n = 2m$, independent normal variables $X_1, \ldots, X_n$ with mean zero and variance one. For $t \leq m$ consider

$$Y_1 = \Sigma_{i=1}^{t} X_i, \quad Y_2 = \Sigma_{t+1}^{n} X_i,$$

and denote $W = Y_1 - Y_2$, $V = Y_1 + Y_2 = \Sigma_1^n X_i$. Then (W, V) has a bivariate normal distribution with means zero, variances n, and correlation coefficient $(2t/n) - 1$. Thus the conditional distribution of W given $V = v$ is normal with mean $\frac{1}{n}(2t - n)v$ and variance $\sigma^2_{W|V=v} = 4t(n-t)/n$. Now choose $n = 4$, $\mathbf{s} = (3, 1)$, $\mathbf{r} = (2, 2)$. Clearly the conditional density function $g(z_1, z_2)$ of (X_1, X_4) given $\mathbf{X}_0 = (X_2, X_3)$ is Schur-concave. For an arbitrary but fixed $\epsilon > 0$ let us define

$$\phi(y_1, y_2) = \begin{cases} -(y_1 - y_2)^2 & \text{for } 0 \leq |\, y_1 + y_2 \,| \leq \epsilon \\ 0 & \text{otherwise,} \end{cases}$$

then ϕ is also Schur-concave. From the joint distribution of (W, V) clearly we have

$$\begin{aligned} & E\left[\{\phi(X_1, \Sigma_2^4 X_i) - \phi(\Sigma_1^2 X_i, \Sigma_3^4 X_i)\} \mid \Sigma_1^4 X_i = 0\right] \\ = \;& -\mathrm{Var}((X_1 - \Sigma_2^4 X_i) \mid V = \Sigma_1^4 X_i = 0) + \mathrm{Var}((\Sigma_1^2 X_i - \Sigma_3^4 X_i) \mid V = 0) \\ = \;& (-3 + 4) = 1 > 0. \end{aligned}$$

Thus by continuity there exists a small $\epsilon > 0$ such that

$$\begin{aligned} & E\left[\phi(X_1, \Sigma_2^4 X_i) - \phi(\Sigma_1^2 X_i, \Sigma_3^4 X_i)\right] \\ = & \int_{-\epsilon}^{\epsilon} E\left[\{- \mid X_1 - \Sigma_2^4 X_i \mid^2 + \mid \Sigma_1^2 X_i - \Sigma_3^4 X_i \mid^2\} \mid \Sigma_1^4 X_i = v\right] dP\left[\Sigma_1^4 X_i \leq v\right] > 0. \end{aligned}$$

4. An Application in Reliability Theory. In this section we state an application of Theorem 1 in reliability theory. Consider n exchangeable components with life lengths $X_1, \ldots, X_n$ which are obviously nonnegative. If the components are manufactured independently, then the joint density f of the X_i's is the product of the common marginal densities; otherwise if they are manufactured under the influence of some common factors or under a common environment, then it is well-known that f is a mixture and the random variables are conditionally i.i.d. In either case f is permutation invariant.

Suppose that a system consists of k subsystems, and that the j-th subsystem, consisting of $r_j \geq 1$ such components, is required to operate properly with one component in operation and the others in a standby capacity ($j = 1, \ldots, k$). Then the life length Y_j of the j-th subsystem is $\Sigma_{r_1+\cdots+r_{j-1}+1}^{r_1+\cdots+r_j} X_i$. Let $Y_{(1)} \leq Y_{(2)} \leq \cdots \leq Y_{(k)}$ denote the order statistics of $Y_1, \ldots, Y_k$ and $\mathbf{r} = (r_1, \ldots, r_k)$ be an allocation vector such that $r_j \geq 1$ and $\Sigma_1^k r_j = n$. When the subsystems are connected in series, then the life length of the system is $Y_{(1)}$. On the other hand if they are connected in parallel, then it is $Y_{(k)}$. Now for fixed $c_i \geq 0$, $\Sigma_1^k c_j y_{(j)}$ is a permutation invariant and concave (convex) function of $(y_1, \ldots, y_k)$ if $c_1 \geq \cdots \geq c_k$ (if $c_1 \leq \cdots \leq c_k$). Consequently, Theorem 1 provides a partial ordering for the expected life length of the system for series and parallel systems. In particular, for series systems the optimal allocation policy is such that $| r_j - r_{j'} | \leq 1$ for all $j \neq j'$, and for parallel systems an optimal policy is that $r_1 = n - k + 1$ and $r_2 = \cdots = r_k = 1$.

REFERENCES

DAS GUPTA, S. (1973). S-unimodal functions: Related inequalities and statistical applications. *Sankhyā Ser. B* **38** 301–314.

KARLIN, S. and PROSCHAN, F. (1960). Pólya type distributions of convolutions. *Ann. Math. Statist.* **31** 721–736.

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.

PROSCHAN, F. and SETHURAMAN, J. (1977). Schur functions in statistics. I. The preservation theorem. *Ann. Math. Statist.* **5** 256–262.

TONG, Y.L. (1980). *Probability Inequalities in Multivariate Distributions.* Academic Press, New York.

DEPARTMENT OF STATISTICS
UNIVERSITY COLLEGE, DUBLIN
BELFIELD, DUBLIN 4, IRELAND

DEPARTMENT OF STATISTICS
FLORIDA STATE UNIVERSITY
TALLAHASSEE, FL 32306-3033

SCHOOL OF MATHEMATICS
GEORGIA INSTITUTE OF TECHNOLOGY
ATLANTA, GA 30332