SCHUR-OSTROWSKI THEOREMS FOR FUNCTIONALS ON $L_1(0,1)$

by

Wai Chan*, Frank Proschan†, and Jayaram Sethuraman*

FSU Statistics Report No. M684
USARO Technical Report No. D-69
AFOSR Technical Report No. 84-171

August, 1984

The Florida State University
Department of Statistics
Tallahassee, Florida 32306

†Research supported by the Air Force Office of Scientific Research, System Command, USAF, under Grant number AFOSR 82-K-0007.

*Research supported by the U.S. Army Research Office under Grant number DAAG 29-82-K-0168.
The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation thereon.

Key Words and phrases: Inequalities, majorization, Muirhead's Theorem, peakedness in symmetric distributions, rearrangement, Schur functions, Schur-Ostrowski's Theorem.

AMS (1980) Subject Classifications: 26D10; 60E15
Schur-Ostrowski Theorems for Functionals on $L_1(0,1)$

by

Wai Chan, Frank Proschan, and Jayaram Sethuraman

ABSTRACT

Hardy, Littlewood and Polya (1934) introduced the partial ordering of majorization among n-dimensional real vectors. Many well known inequalities can be recast as the statement that certain functions are increasing with respect to this ordering. Such functions are said to be Schur-convex. An important result in the theory of majorization is the Schur-Ostrowski Theorem, which characterizes Schur-convex functions. The concept of majorization has been extended to elements of $L_1(0,1)$ by Ryff (1963). A function on $L_1(0,1)$ that is increasing with respect to the ordering of majorization is said to be Schur-convex. In this paper, we prove an analogue of the Schur-Ostrowski condition which characterizes Schur-convex functionals in terms of their Gâteaux differentials. We also introduce another partial ordering in $L_1(0,1)$ called unrestricted majorization. This partial ordering is similar to majorization but does not involve the use of decreasing rearrangements. We establish a characterization of non-decreasing functions on $L_1(0,1)$ with respect to the partial ordering of unrestricted majorization through another analogue of the Schur-Ostrowski condition.
1. Introduction.

Hardy, Littlewood and Polya (1934) introduced the following partial order in n-dimensional Euclidean spaces: a n-vector \( x = (x_1, \ldots, x_n) \) majorizes \( y = (y_1, \ldots, y_n) \), \( (x \preceq y \text{ in symbols}) \), whenever

\[
\frac{1}{k} \sum_{i=1}^{k} x_i^* \geq \frac{1}{k} \sum_{i=1}^{k} y_i^*, \quad k = 1, \ldots, n-1
\]

and

\[
\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i,
\]

where \( x^*, y^* \) are the vectors obtained from \( x \) and \( y \) by rearranging their components in decreasing order.

This partial order has been extended to elements of \( L_1(0,1) \) Ryff (1963) and is given in Definition 1.2 below. Before giving the definition, we develop some notation to be used in defining a decreasing rearrangement of a function. Let \( x \) be a measurable, real valued function on \( (0,1) \) and \( m \) be the Lebesgue measure. For each \( x \), one can associate a function \( d_x \) on \( (-\infty, \infty) \) defined by

\[
d_x(s) = m(\{t: x(t) > s\}), \quad -\infty < s < \infty.
\]

This function \( d_x \), called the distribution function of \( x \), is non-increasing and right continuous. Two functions \( x \) and \( y \) are said to be equiprobable if \( d_x = d_y \). The right continuous inverse of \( d_x \), denoted \( x^* \), is defined by

\[
x^*(t) = \inf \{s: d_x(s) \leq t\}.
\]

The function \( x^* \), which is non-increasing, right continuous and has the same distribution function as \( x \), is called the decreasing rearrangement of \( x \). The functions \( x \) and \( x^* \) are simultaneously integrable (or not integrable), and their integrals are related by
\[
\int_0^s x^*(t) \, dt \geq \int_0^s x(t) \, dt \quad 0 \leq s < 1,
\]

and
\[
\int_0^1 x^*(t) \, dt = \int_0^1 x(t) \, dt.
\]

The following theorem due to Ryff (1970) shows that by composing the decreasing rearrangement of a function with a measure preserving transformation, one can recover the original function.

1.1. Theorem. To each \( x \in L_1(0,1) \), there corresponds a measure preserving transformation \( \sigma : (0,1) \to (0,1) \) such that \( x(t) = x^*[\sigma(t)] \), where \( \sigma \) is defined by
\[
\sigma(s) = \mathcal{M}(t: x(t) > x(s)) + \mathcal{M}(t \leq s: x(t) = x(s)).
\]

The definition of the partial ordering of majorization of elements in \( L_1(0,1) \), due to Ryff (1963), is given below.

1.2. Definition. Let \( x, y \in L_1(0,1) \). We say that \( x \) majorizes \( y \), \( (x \geq y \) in symbols) if

\[
\int_0^s x^*(t) \, dt \geq \int_0^s y^*(t) \, dt, \quad 0 \leq s < 1,
\]

and
\[
\int_0^1 x(t) \, dt = \int_0^1 y(t) \, dt,
\]

where \( x^* \) and \( y^* \) are the decreasing rearrangements of \( x \) and \( y \), respectively.

Several authors [see, eg. Day (1975), Chong (1976)] have obtained interesting results using this partial ordering. It is also related to the variability ordering of Ross (1982).

By removing the rearrangement requirement in Definition 1.2, we obtain a different ordering called unrestricted majorization, as described below.
1.3. Definition. Let \( x, y \in L_1(0,1) \). We say that \( x \) dominates \( y \) in the ordering of unrestricted majorization, \((x \triangleright y \) in symbols), if

\[
\int_0^s x(t)dt \geq \int_0^s y(t)dt, \quad 0 \leq s < 1,
\]

and

\[
\int_0^1 x(t)dt = \int_0^1 y(t)dt.
\]

The ordering of unrestricted majorization as applied to the class of density functions leads to the usual stochastic ordering as seen below:

Let \( X \) and \( Y \) be random variables on \((0,1)\) with densities \( f \) and \( g \) respectively. If \( f \triangleright g \), then \( \int_0^s f \geq \int_0^s g \) for all \( 0 < s < 1 \), or \( P(X \leq s) \geq P(Y \leq s) \). Thus the condition \( X \leq^{st} Y \) is equivalent to \( f \geq g \).

Many inequalities that arise from majorization in the finite dimensional case can be extended for elements of \( L_1(0,1) \). Ryff (1967) proved the following analogue of Muirhead's inequality.

1.4. Theorem. Let \( x \) and \( y \) be bounded measurable functions on \((0,1)\).

If \( x \triangleright y \) and \( u \) is a positive function such that \( u \in L_1(0,1) \) for all \( p, -\infty < p < \infty \), then

\[
\int_0^1 \log^p \left[ \int_0^1 u(t)x(s)dt \right] ds \geq \int_0^1 \log^p \left[ \int_0^1 u(t)y(s)dt \right] ds
\]

Conversely, if the inequality holds for all such \( u \), then \( x \triangleright y \).

In the discrete case, Muirhead's inequality can be reformulated by identifying an appropriate function which preserves the ordering of majorization. Such functions are said to be Schur-convex. Schur (1923) and Ostrowski (1952) gave necessary and sufficient conditions.
for a function to be Schur-convex in terms of their partial derivatives. We quote from Marshall and Olkin (1979) about the importance of this result, "it is difficult to overemphasize the usefulness of the (Schur-Ostrowski) condition, ...., many or even most of the theorems giving Schur-convexity were first discovered by checking (the Schur-Ostrowski condition)." In the next section, we will present an analogue of this result for Schur-convex functionals on $L_\infty(0,1)$. This result, given in Theorem 2.9, is then used to characterize Schur-convex functionals on $L_1(0,1)$. We also characterize non-decreasing functionals on $L_1(0,1)$ with respect to the partial ordering of unrestricted majorization through another analogue of the Schur-Ostrowski condition. These results will be used to prove the generalized Muirhead's Theorem (Proshan and Sethuraman, 1976) in Section 3. An application to peakedness comparisons of distributions is discussed in Section 3.

2. Main Theorems.

We first proceed with some definitions.

2.1. Definition. A functional $\phi$ defined on a set $A \subseteq L_1(0,1)$ is said to be Schur-convex on $A$ if $y_1, y_2 \in A$ and $y_1 \geq y_2$ imply that $\phi(y_1) \geq \phi(y_2)$.

A Schur-convex functional is necessarily constant over functionals that are equivalent in distribution. Thus for a Schur-convex functional $\phi$, the value $\phi(x)$ depends only on the distribution function of $x$. A set $A$ is said to be invariant if $x \in A$ and $x$ and $y$ are equivalent in distribution imply that $y \in A$. Henceforth, we shall only consider Schur-convex functionals on an invariant set.

For a characterization of Schur-convex functionals, we need the following notion of directional derivative.
2.2. Definition. Let $\phi$ be a functional defined on a convex set $A \subseteq L_1(0,1)$. Let $y \in A$ and $h$ be such that $y + \theta h \in A$ for all sufficiently small $\theta$. The Gâteaux differential of $\phi$ at $y$ in the direction of $h$ is defined to be

$$\frac{\partial \phi}{\partial h}(y) = \lim_{\theta \to 0} \frac{\phi(y+\theta h) - \phi(y)}{\theta}$$

if the limit exists.

Note that $\frac{\partial \phi}{\partial h}(y)$ is simply the derivative, at $\theta = 0$, of the real-valued function on $[0,1]$ defined by $\psi(\theta) = \phi(y + \theta h)$.

Let $D_1$ be the class of decreasing functions in $L_1(0,1)$, let $T$ be the class of decreasing functions in $L_\infty(0,1)$. Let

$$T = \{ h : h = \lambda_1 I_{(a,b)} + \lambda_2 I_{(c,d)} \text{, where } 0 < a < b < c < d \leq 1, \lambda_1 \geq 0 \geq \lambda_2, \lambda_1(b-a) + \lambda_2(d-c) = 0 \}.$$  

The class $T$ consists of step functions $h$ which take at most two non-zero values, are decreasing on its support and satisfy $\int_0^1 h(t) dt = 0$. Note that $h \in T$ implies $h \geq 0$.

Let $y \in D_1$ and $h \in T$. Then $y + h$ need not be decreasing. However, we have $y + h \geq y$, as given in the next lemma.

2.3 Lemma. Let $y \in D_1$ and $h \in T$, then $y + h \geq y$.

Proof. 

Note that

$$\int_0^s (y+h)^* \geq \int_0^s (y+h)$$

$$= \int_0^s y^* + \int_0^s h$$

$$\geq \int_0^1 y^*, \quad 0 \leq s < 1.$$
and

\[ \int_0^{(y+h)} = \int_0^y + \int_0^h = \int_0^y. \]

Hence, \( y + h \geq y. \)

In the following Theorem, we give a necessary condition for functionals increasing in the ordering of unrestricted majorization.

2.4. Theorem. Let \( A \) be an open subset of \( L_1(0,1) \). Let \( \phi \) be a functional defined on \( A \) such that \( \phi \) is non-decreasing with respect to the ordering of unrestricted majorization. Let \( y \in A \) and \( h \in T \). Suppose that the Gâteaux differential \( \frac{\partial \phi}{\partial h}(y) \exists \). Then \( \frac{\partial \phi}{\partial h}(y) \geq 0. \)

Proof.

Since \( A \) is open, \( y + \theta h \in A \) for all sufficiently small \( \theta \). For all sufficiently small positive \( \theta \), \( y + \theta h \) and \( y \) are elements of \( A \) and \( y + \theta h \succneq y \). This implies that

\[ \phi(y + \theta h) \geq \phi(y) \]

and

\[ \frac{\partial \phi}{\partial h}(y) = \lim_{\theta \to 0} \frac{1}{\theta} [\phi(y + \theta h) - \phi(y)] \]

\[ \geq 0. \]

Next, we consider Schur-convex functionals defined on an invariant set \( A \).

2.5. Theorem. Let \( A \) be an open invariant subset of \( L_1(0,1) \). Let \( \phi \) be a Schur-convex functional defined on \( A \). Let \( y \in D_\infty \cap A \) and \( h \in T \). Suppose that the Gâteaux differential \( \frac{\partial \phi}{\partial h}(y) \exists \). Then \( \frac{\partial \phi}{\partial h}(y) \geq 0. \)
Proof.

Since $A$ is open, $y + \theta h \in A$ for all sufficiently small $\theta$. Furthermore, for sufficiently small positive $\theta$, $y + \theta h \supseteq y$ from Lemma 2.3. Hence $\phi(y + \theta h) \geq \phi(y)$ and

$$\frac{d\phi}{d\theta}(y) = \lim_{\theta \to 0} \frac{1}{\theta} [\phi(y + \theta h) - \phi(y)]$$

$$\geq 0. \quad ||$$

To show that this condition is also sufficient, which is the content of the main theorem, Theorem 2.10 of this section, we need the following lemmas.

2.6. Lemma. Let $A$ be a convex subset of $L_1(0,1)$. Let $\phi$ be a functional defined on an open set containing $A$. Let $\frac{d\phi}{dh}(y) \geq 0$ for $y \in A$ and $h \in T$. Then $y_1, y_2 \in A$ and $y_2 - y_1 \in T$ imply that $\phi(y_2) \geq \phi(y_1)$.

Proof.

Let $h = y_2 - y_1 \in T$. For $\theta \in [0,1]$, define

$$y_\theta = y_1 + \theta(y_2 - y_1) = \theta y_2 + (1-\theta)y_1,$$

and

$$\psi(\theta) = \phi(y_\theta).$$

Note that $y_\theta$ is in $A$. Now,

$$\frac{d}{d\theta} \psi(\theta) = \lim_{\epsilon \to 0} \frac{1}{\epsilon} [\psi(\theta + \epsilon) - \psi(\theta)]$$

$$= \lim_{\epsilon \to 0} \frac{1}{\epsilon} [\phi(y_\theta + \epsilon h) - \phi(y_\theta)]$$

$$= \frac{d\phi}{dh}(y_\theta) \geq 0 \quad \text{for } 0 \leq \theta \leq 1.$$
Hence,

\[
\phi(y_2) - \phi(y_1) = \int_0^1 \frac{d}{d\theta} \psi(\theta) d\theta
\]

\[
= \int_0^1 \frac{d\phi}{dh} (y_0) d\theta \geq 0. \| \|
\]

In the next lemma, we show that if \( y_1, y_2 \) are step functions such that \( y_2 \uplus y_1 \), then \( y_2 - y_1 \) can be written as the sum of functions in \( T \).

2.7. Lemma. Let \( y_1, y_2 \) be step functions on (0,1) such that \( y_2 \uplus y_1 \). Then there exist \( h_1, \ldots, h_N \) in \( T \) such that

\[
y_2 = y_1 + \sum_{i=1}^N h_i.
\]

(2.1)

Proof. There is nothing to prove if \( y_1 = y_2 \).

Let \( y_1 \neq y_2 \). Since \( y_1 \) and \( y_2 \) are step functions, there is an integer \( n \geq 2 \), such that

\[
y_2(t) - y_1(t) = \sum_{i=1}^n a_i I(c_i, d_i)(t) \quad \text{where} \quad a_i \neq 0,
\]

\((c_i, d_i)\) are disjoint intervals and \( 0 \leq c_1 < d_1 < \ldots < c_n < d_n \leq 1 \). Note that \( y_2 \uplus y_1 \) implies that \( a_1 > 0 \) and \( a_n < 0 \). We will prove that (2.1) holds with \( N \leq n-1 \), by an induction on \( n \).

Note that the lemma is immediate when \( n = 2 \). Assume that the lemma is true for \( n = 2, \ldots, k-1 \). We will now prove that the lemma is true for \( n = k \).
holds for \( n = k \). Let \( a_j \) be the first negative term such that either \( a_{j+1} > 0 \) or \( j = k \). Define a function in \( T \) by

\[
h = e_1 I(c_1', c_1') + a_j I(d_j', d_j'),
\]

where \( c_1' \leq d_1 \) and \( d_j' \geq c_j \) are chosen so that \( a_1 (c_1' - c_1) + a_j (d_j' - d_j) \) and one of the following holds:

1) \( c_1' = d_1 \) and \( d_j' = c_j \) if \( a_1 (d_1 - c_1) + a_j (d_j - c_j) = 0 \),
2) \( c_1' < d_1 \) and \( d_j' = c_j \) if \( a_1 (d_1 - c_1) + a_j (d_j - c_j) > 0 \),
3) \( c_1' = d_1 \) and \( d_j' > c_j \) if \( a_1 (d_1 - c_1) + a_j (d_j - c_j) < 0 \).

We will now establish that \( y_2 \leq y_1 + h \) by showing that \( \int_0^s (y_2 - y_1 - h) \geq 0 \) for all \( 0 < s \leq 1 \). Note that \( h = 0 \) on the intervals \( (d_j, 1) \).

Let \( s > d_j \), then \( \int_0^s h = \int_0^1 h = 0 \). Thus,

\[
\int_0^s (y_2 - y_1 - h) = \int_0^1 (y_2 - y_1) \geq 0.
\]

Let \( 0 < s \leq d_j \). Then either \( y_2(t) - y_1(t) \geq 0 \) for all \( 0 < t < s \) or \( y_2(t) - y_1(t) \leq 0 \) for all \( s < t \leq d_j \), since there is only one sign change among \( a_1, \ldots, a_j \) and the sign changes from positive to negative. Note that \( h \) agrees with \( y_2 - y_1 \) on the intervals \( (c_1', c_1') \) and \( (d_j', d_j) \), and that \( h \) is identically zero outside these intervals.
If \( y_2(t) - y_1(t) \geq 0 \) for all \( 0 < t < s \), then \( y_2 - y_1 \geq h \geq \) on the interval \((0,s)\). This implies that

\[
\int_0^s (y_2 - y_1 - h) \geq 0.
\]

If \( y_2(t) - y_1(t) \leq 0 \) for all \( s < t < d_j \), then \( y_2 - y_1 \leq h \leq \) on the interval \((s,d_j)\). This implies that

\[
\int_0^{d_j} (y_2 - y_1 - h) \geq \int_0^{d_j} (y_2 - y_1 - h)
\]

\[
= \int_0^{d_j} (y_2 - y_1)
\]

\[\geq 0.\]

Hence we have \( y_2 \overset{\mu}{\geq} y_1 + h \). Since \( y_2 - y_1 - h \) is a step function which takes at most \( k-1 \) nonzero values, it follows from the induction hypothesis that \( y_2 - y_1 - h = \sum_{i=1}^{N} h_i \) where \( h_i \in \mathcal{T} \) for \( i = 1, \ldots, N \), and \( N \leq k-2 \). This completes the proof. \( \Box \)

In Lemma 2.7, if we assume that \( y_1, y_2 \) are decreasing step functions, then the condition \( y_2 \overset{\mu}{\geq} y_1 \) is equivalent to \( y_2 \overset{m}{\geq} y_1 \). In addition, we can choose \( y_1 + h_1, y_1 + h_1 + h_2, \ldots, y_1 + \sum_{i=1}^{N-1} h_i \) to be decreasing functions as shown in the following lemma.

2.8. Lemma. Let \( y_1, y_2 \) be decreasing step functions on \((0,1)\) such that \( y_2 \overset{m}{\geq} y_1 \). Then there exist \( h_1, \ldots, h_N \) in \( \mathcal{T} \) such that
i) \( y_2 = y_1 + \sum_{i=1}^{N} h_i \),

and

ii) \( y_1 + h_1, \ldots, y_1 + \sum_{i=1}^{N-1} h_i \) are decreasing functions.

Proof.

Define \( h = a_1 I(c_1, c_1^{'}) + a_j I(d_j^{'}, d_j) \) as in the proof of Lemma 2.7. We need to show that \( y_1 + h \) is decreasing. Note that

\[
y_1(t) + h(t) = \begin{cases} 
y_2(t) & \text{if } 0 < t < c_1^{'}, 
y_1(t) & \text{if } c_1^{'} \leq t \leq d_j^{'}, 
y_2(t) & \text{if } d_j^{'} < t \leq d_j, 
y_1(t) & \text{if } d_j \leq t < 1.
\end{cases}
\]

Since \( a_1 > 0 \), \( y_1 + h \) is decreasing on a neighborhood of \( c_1^{'}. \)

Similarly, \( a_j < 0 \) implies that \( y_1 + h \) is decreasing on a neighborhood of \( d_j^{'}. \) Suppose that \( d_j < 1 \), then the choice of \( a_j \) indicates that for \( \epsilon > 0 \) sufficiently small, \( y_2 - y_1 \geq 0 \) on \( (d_j, d_j + \epsilon) \). Since \( y_1 + h = y_2 \) on \( (d_j^{'}, d_j) \), it follows that \( y_1 + h \) is decreasing on the interval \( (d_j^{'}, d_j + \epsilon) \). Thus \( y_1 + h \) is decreasing on the interval \( (0, 1) \).

Note that \( h \in T \) implies \( y_1 + h \preceq y_1 \). Since \( y_1 + h \), \( y_1 \) are decreasing functions, this is equivalent to \( y_1 + h \preceq y_1 \). Following the same induction argument as in Lemma 2.7, we conclude that there exist \( h_1, \ldots, h_N \) in \( T \) such that \( y_2 = y_1 + h + \sum_{i=1}^{N} h_i \) and that \( y_1 + h + h_1, \ldots, y_1 + h + \sum_{i=1}^{N-1} h_i \) are decreasing functions. This proves the lemma. \( \Box \)
In the next theorem, we give a sufficient condition for a functional of \( L_\infty(0,1) \) to be Schur-convex.

2.9. **Theorem.** Let \( A \) be an invariant open convex subset of \( L_\infty(0,1) \).

Let \( \phi \) be a continuous functional defined on \( A \) such that \( \phi \) is constant over functions that are equivalent in distribution. If the \( \text{Gâteaux} \) differential \( \frac{\partial \phi}{\partial h}(y) \geq 0 \) for each \( y \in D_\infty \cap A \) and \( h \in T \), then \( \phi \) is Schur-convex on \( A \).

**Proof.**

Since \( \phi \) is constant over functions that are equivalent in distribution, it suffices to prove that \( \phi \) is Schur-convex on \( D_\infty \cap A \).

Let \( y_1, y_2 \in D_\infty \cap A \) be right continuous and \( y_2 \preceq y_1 \). Let \( \epsilon \) be arbitrary. Then for \( i = 1, 2 \), the sets \( \{ t \colon y_i(t^-) - y_i(t) > \epsilon \} \) are finite, where \( y_i(t^-) = \inf_{t < t_0} y_i(t) \). Hence there exists a partition

\[
0 < a_1 < \ldots < a_n < 1
\]

such that

\[
y_i(a_k) - y_i(a_{k+1}) < \epsilon, \quad i = 1, 2; \quad k = 1, \ldots, n-1.
\]

Define

\[
y_{i\epsilon}(t) = \frac{1}{a_1} \left[ y_1(s)ds \right]_{0}^{a_1} I_{(0,a_1)}(t) + \sum_{k=1}^{n-1} \frac{1}{a_{k+1} - a_k} \left[ y_1(s)ds \right]_{a_k}^{a_{k+1}} I_{[a_k,a_{k+1})}(t) + \frac{1}{1-a_n} \left[ y_1(s)ds \right]_{a_n}^{1} I_{[a_n,1)}(t), \quad i = 1, 2.
\]
Then \( y_{1\varepsilon}, y_{2\varepsilon} \) are decreasing step functions satisfying
\[
a_k = a_k, \quad \int_0 y_{1\varepsilon}(s) ds = \int_0 y_{1}(s) ds \quad \text{for } k = 1, \ldots, n.
\]
This implies \( y_{2\varepsilon} \geq y_{1\varepsilon} \).

Since \( A \) is open and \( \| y_{1\varepsilon} - y_{i\varepsilon} \|_{\infty} < \varepsilon \), for sufficiently small positive \( \varepsilon \), \( y_{1\varepsilon}, y_{2\varepsilon} \) are in \( A \).

By Lemma 2.8,
\[
y_{2\varepsilon} - y_{1\varepsilon} = \sum_{i=1}^{N} h_i \quad \text{for some } \{h_1, \ldots, h_N\} \subset T,
\]
where \( y_{1\varepsilon} + h_1, \ldots, y_{1\varepsilon} + \sum_{i=1}^{N-1} h_i \) are decreasing functions. The functions \( y_{1\varepsilon} + h_1, \ldots, y_{1\varepsilon} + \sum_{i=1}^{N-1} h_i \) need not be elements of \( A \). Since \( A \) is open for sufficiently small positive \( \theta \), \( y_{1\varepsilon} + \theta h_1, \ldots, y_{1\varepsilon} + \theta \sum_{i=1}^{N-1} h_i \) are decreasing functions in \( A \) satisfying
\[
y_{1\varepsilon} + \theta \sum_{i=1}^{N-1} h_i \geq y_{1\varepsilon} + \theta \sum_{i=1}^{N-1} h_i \geq \ldots \geq y_{1\varepsilon}.
\]
It now follows from Lemma 2.8 that \( \phi(y_{1\varepsilon} + \theta \sum_{i=1}^{N-1} h_i) \geq \phi(y_{1\varepsilon} + \theta \sum_{i=1}^{N-1} h_i) \geq \ldots \geq \phi(y_{1\varepsilon}) \).

Next, we shall show that this implies \( \phi(y_{2\varepsilon}) \geq \phi(y_{1\varepsilon}) \). Note that we have just demonstrated that the set \( \Theta = \{0 \leq \theta \leq 1 : \phi(y_{1\varepsilon} + \theta \sum_{i=1}^{N} h_i) \geq \phi(y_{1\varepsilon}) \} \) is non-empty. Let \( \theta_0 = \sup(\Theta : \theta \in \Theta) \).

Since \( \phi \) is continuous, we have
\[
\phi(y_{1\varepsilon} + \theta_0 \sum_{i=1}^{N} h_i) \geq \phi(y_{1\varepsilon}), \quad \text{which shows that } \theta_0 \in \Theta.
\]
Now suppose \( \theta_0 < 1 \). The preceding arguments show that for sufficiently small positive \( r, y_{1\varepsilon} + \theta_0 \sum_{i=1}^{N} h_i + rh_i, \ldots, y_{1\varepsilon} + \theta_0 \sum_{i=1}^{N} h_i + r \sum_{i=1}^{N} h_i \) are decreasing functions in \( A \) and satisfy
\[
\phi(y_{1\varepsilon} + \theta_0 \sum_{i=1}^{N} h_i) \leq \phi(y_{1\varepsilon} + \theta_0 \sum_{i=1}^{N} h_i + rh_i) \leq \ldots \leq \phi(y_{1\varepsilon} + \theta_0 \sum_{i=1}^{N} h_i + r \sum_{i=1}^{N} h_i).
\]
Thus \( \theta_0 + r \in \Theta \), which provides a
contradiction to the assumption that $\theta_0 < 1$. We therefore conclude that $\theta_0 = 1$ and thus, $\phi(y_{2\epsilon}) \geq \phi(y_{1\epsilon})$.

Since the functional is continuous with respect to $L_\infty$-norm, we conclude that $\phi(y_2) \geq \phi(y_1)$ by letting $\epsilon \to 0$. \|

We now use this theorem to establish a sufficient condition for Schur-convex functional of $L_1(0,1)$.

2.10. Theorem. Let $A$ be an invariant open convex subset of $L_1(0,1)$. Let $\phi$ be a continuous functional defined on $A$ such that $\phi$ is constant over functions that are equivalent in distribution. If the Gâteaux differential $\frac{\partial \phi}{\partial h}(y) \geq 0$ for each $y \in D_\infty \cap A$ and $h \in T$, then $\phi$ is Schur-convex on $A$.

Proof.

Since $\phi$ is constant over functions that are equivalent in distribution, it suffices to prove that $\phi$ is Schur-convex on $D_1 \cap A$.

Let $y_1, y_2 \in D_1 \cap A$ be right continuous and $y_2 \geq y_1$. Let $\epsilon > 0$ be arbitrary, then $\exists \delta > 0$ such that

$$\int_0^{y_1(t)} |y_1(t)| dt < \frac{\epsilon}{4} \quad \text{and} \quad \int_{1-\delta}^{y_1(t)} |y_1(t)| dt < \frac{\epsilon}{4}, \quad i = 1, 2.$$

Since the $y_i$'s are in $D_1$, they are bounded on the interval $[\delta, 1-\delta]$. Define
\[ y_{i\epsilon}(t) = \frac{1}{\epsilon} \left[ \int_{0}^{\delta} y_{i}(s) ds \right] I_{(0,\delta)}(t) + \frac{1}{\epsilon} \int_{1-\delta}^{1} y_{i}(s) ds I_{(1-\delta,1)}(t), \quad i = 1, 2. \]

Then \( y_{i\epsilon} \in D_{\infty}, \ i = 1, 2 \) and \( y_{2\epsilon} \geq y_{1\epsilon} \). We also have

\[ \| y_{i\epsilon} - y_{i} \|_{1} = \int_{0}^{\delta} | y_{i\epsilon} - y_{i} | + \int_{1-\delta}^{1} | y_{i\epsilon} - y_{i} | \leq \epsilon, \quad i = 1, 2. \]

Hence, for sufficiently small \( \epsilon \), \( y_{i\epsilon} \in D_{\infty} \cap A \). It now follows from Theorem 2.9 that \( \phi(y_{2\epsilon}) \geq \phi(y_{1\epsilon}) \). Since \( \phi \) is a continuous functional, we obtain that \( \phi(y_{2}) \geq \phi(y_{1}) \) by letting \( \epsilon \to 0 \). This completes the proof. \|

The following lemma is used to prove Theorem 2.12, which is an analogue of Theorem 2.10 for functionals on \( L_{1}(0,1) \) which are non-decreasing with respect to the ordering of unrestricted majorization.

**Lemma 2.11.** Let \( y_{1}, y_{2} \in L_{1}(0,1) \) such that \( y_{2} \geq y_{1} \). For each \( \epsilon > 0 \), there exists a partition \( 0 < a_{1} < \ldots < a_{n} < 1 \) such that the step functions defined by

\[ y_{i\epsilon}(t) = \frac{1}{a_{1}} \int_{0}^{a_{1}} y_{i}(s) ds I_{(0,a_{1})}(t) \]

\[ + \sum_{k=1}^{n-1} \frac{1}{a_{k+1} - a_{k}} \int_{a_{k}}^{a_{k+1}} y_{i}(s) ds I_{[a_{k},a_{k+1})}(t) \]

\[ + \frac{1}{1-a_{n}} \int_{a_{n}}^{1} y_{i}(s) ds I_{[a_{n},1)}(t), \quad i = 1, 2, \]

satisfy the following:
(i) $\|y_{1\epsilon} - y_1\|_1 < \epsilon$, $i = 1, 2$

and

(ii) $y_{2\epsilon} \not\subseteq y_{1\epsilon}$.

Proof.

Note that if $y_1$, $y_2$ are continuous functions on the interval $[0,1]$, then (i) follows from the uniform continuity of $y_1$ and $y_2$.

If $y_i$'s are not continuous on $[0,1]$, we first approximate $y_i$'s by continuous functions $x_i$ on $[0,1]$ such that $\|y_i - x_i\|_1 < \frac{\epsilon}{3}$, $i = 1, 2$.

Next, we find a partition $0 < a_1 < \ldots < a_n < 1$ such that the step functions defined by

$$x_{i\epsilon}(t) = \frac{1}{a_1} \left[ \int_0^{a_1} y_i(s)ds \right] I(0,a_1)(t)$$

$$+ \sum_{k=1}^{n-1} \frac{1}{a_{k+1} - a_k} \left[ \int_{a_k}^{a_{k+1}} y_i(s)ds \right] I[a_k,a_{k+1}](t)$$

$$+ \frac{1}{1-a} \left[ \int_{a_n}^{1} y_i(s)ds \right] I[a_n,1](t), \ i = 1, 2$$

satisfy $\|x_{i\epsilon} - x_i\|_1 < \frac{\epsilon}{3}$.

Now, define the step functions $y_{1\epsilon}$, $y_{2\epsilon}$ by

$$y_{i\epsilon}(t) = \frac{1}{a_1} \left[ \int_0^{a_1} y_i(s)ds \right] I(0,a_1)(t)$$

$$+ \sum_{k=1}^{n-1} \frac{1}{a_{k+1} - a_k} \left[ \int_{a_k}^{a_{k+1}} y_i(s)ds \right] I[a_k,a_{k+1}](t)$$

$$+ \frac{1}{1-a} \left[ \int_{a_n}^{1} y_i(s)ds \right] I[a_n,1](t), \ i = 1, 2.$$
Then

\[ \|y_{i\epsilon} - x_{i\epsilon}\|_1 = \left| \int_0^{a_1} [y_i(s) - x_i(s)]ds \right| \]

\[ + \sum_{k=1}^{n-1} \int_{a_k}^{a_{k+1}} [y_i(s) - x_i(s)]ds \]

\[ + \int_{a_n}^{1} [y_i(s) - x_i(s)]ds \]

\[ \leq \|x_i - y_i\|_1, \quad i = 1, 2. \]

Thus

\[ \|y_{1\epsilon} - y_{2\epsilon}\|_1 \leq \|y_1 - x_1\|_1 \leq \|x_1 - x_{1\epsilon}\|_1 + \|x_{1\epsilon} - y_{1\epsilon}\|_1 \]

\[ \leq \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon, \quad i = 1, 2. \]

This proves the first part of the lemma.

Let \( y_2 \not\leq y_1 \). Then for any partition \( 0 < a_1 < \ldots < a_n < 1 \), the step functions \( y_{1\epsilon} \), \( y_{2\epsilon} \) satisfy that \( y_{2\epsilon} \not\leq y_{1\epsilon} \). This proves (ii).

2.12 Theorem. Let \( A \) be an open convex subset of \( L_1(0,1) \) and let \( \phi \) be a continuous functional on \( A \) such that the Gâteaux differentiable 

\[ \frac{\partial \phi}{\partial h}(y) \geq 0 \] for \( y \in A \) and \( h \in \mathcal{T} \). Then \( y_1, y_2 \in A \) and \( y_2 \not\leq y_1 \) imply that \( \phi(y_2) \geq \phi(y_1) \).

Proof.

We shall first prove the theorem for step functions. Let \( y_1, y_2 \) be step functions in \( A \) and \( y_2 \not\leq y_1 \). Then by Lemma 2.7, \( y_2 - y_1 = \sum_{i=1}^{N} h_i \) for some \( \{h_1, \ldots, h_N\} \subseteq \mathcal{T} \). Since \( A \) is open, for
sufficiently small positive $\theta$, $y_1 + \theta h_1$, $\ldots$, $y_1 + \theta \sum_{i=1}^{N} h_i$ are functions in $A$ and satisfy that

$$y_1 + \theta \sum_{i=1}^{N} h_i \leq \ldots \leq y_1 + \theta \sum_{i=1}^{N-1} h_i \leq \ldots \leq y_1 + \theta h_1 \leq y_1.$$ 

It now follows from Lemma 2.6 that

$$\phi(y_1 + \theta \sum_{i=1}^{N} h_i) \geq \ldots \geq \phi(y_1 + \theta h_1) \geq \phi(y_1). \quad \text{Define}$$

$$\Theta = \{0 \leq \theta \leq 1: \phi(y_1 + \theta \sum_{i=1}^{N} h_i) \geq \phi(y_1)\} \quad \text{and} \quad \theta_0 = \sup\{\theta: \theta \in \Theta\}.$$ 

Following the same argument as in the proof of Theorem 2.9, we can show that $\theta_0 = 1$. Hence,

$$\phi(y_2) = \phi(y_1 + \sum_{i=1}^{N} h_i) \geq \phi(y_1).$$

In general, let $y_1, y_2 \in L_1(0,1)$ and $y_2 \geq y_1$. Let $\epsilon > 0$. By Lemma 2.11 there exist step functions $y_{1\epsilon}, y_{2\epsilon}$ such that $\|y_{1\epsilon} - y_i\|_1 < \epsilon$ for $i = 1, 2$ and $y_{2\epsilon} \geq y_{1\epsilon}$. Since $A$ is open, for sufficiently small $\epsilon$, $y_{1\epsilon}, y_{2\epsilon}$ are functions in $A$. Thus $\phi(y_{2\epsilon}) \geq \phi(y_{1\epsilon})$. Since $\phi$ is continuous, we conclude that $\phi(y_2) \geq \phi(y_1)$ by letting $\epsilon \to 0$. ||

3. Applications.

The inequality given in Theorem 1.4 can be reformulated as the statement that the functional defined by

$$\phi(x) = \int_{0}^{1} \log(\int_{0}^{1} u(t)x(s) \, dt) \, ds$$

is Schur-convex. By Theorem 2.10, this is equivalent to the condition $\frac{\partial \phi}{\partial h}(x) \geq 0 \quad \forall \, x \in D_{\alpha}, \, \forall \, h \in T$. This condition can be verified as fol
Using Holder's inequality, we note that the function
\[ M(\alpha) = \log \|u\|^\alpha_\alpha \] is convex in \( \alpha \), and thus
\[ M'(x) = \int_0^1 u(t)^\alpha \log u(t) dt / \int_0^1 u(t)^\alpha dt \]
is increasing in \( \alpha \). Let \( x \in D_\infty \) and \( h \in T \), then both \( x \) and \( h \) are functions decreasing on their supports, and \( \int_0^1 h(t) dt = 0 \). This implies that
\[
\frac{\partial \phi}{\partial h}(x) = \int_0^1 \left[ \frac{1}{\int_0^1 u(t) x(s) dt} \right] h(s) ds \geq 0.
\]

More generally, we can replace the function \( u(t) x(s) \) by functions of the form \( \psi(t,z) \) which are log convex in \( z \) for fixed \( t \). This is the result of Proschan and Sethuraman (1976), which we will state below.

3.1. Theorem. Let the function \( \psi(t,z) \) on \((0,1) \times (-\infty,\infty)\) be a log convex function in \( z \) for fixed \( t \), and the partial derivative \( \psi_z(t,z) = \frac{\partial}{\partial z} \psi(t,z) \) exists. Also let \( \sup |\psi(t,z)| \) belong to \( L_1(0,1) \) for each \( k < \infty \). For any bounded measurable function \( x \) on \((0,1)\), define
\[ M_\psi(x) = \int_0^1 \log \left[ \frac{1}{\int_0^1 \psi(t,x(s)) dt} \right] ds. \]

Then \( M_\psi \) is Schur-convex.
Proof.

If follows from Artin's Theorem (1931) that the positive linear combination \[
\int_0^1 \psi(t,z) dt \text{ is log convex in } z. \text{ Let } x \in D_{\infty} \text{ and } h \in T,
\]

\[
\frac{\partial M_\psi(x)}{\partial h} = \int_0^1 \left[ \frac{1}{\psi(t,x(s)) dt} \right] h(s) ds \geq 0,
\]

which implies that \( M_\psi \) is Schur-convex. ||

Next, we shall study an application of unrestricted majorization to peakedness ordering of symmetric distributions.

Let \( X \) and \( Y \) be random variables possessing densities symmetric about zero. According to the definition of peakedness given by Birnbaum (1948), \( X \) is more peaked than \( Y \), \((X \overset{P}{\geq} Y)\) in symbols), if \( P(X \leq t) \geq P(Y \leq t) \) for all \( t \geq 0 \). Let \( f \) and \( g \) be the densities of \( X \) and \( Y \) respectively. Then the condition \( X \overset{P}{\geq} Y \) is equivalent to \( f \overset{P}{\underset{\cup}{\geq}} g \) on the interval \((0, \infty)\).

Birnbaum (1948) showed that under appropriate conditions, \( X_1 \overset{P}{\geq} Y_1 \) and \( X_2 \overset{P}{\geq} Y_2 \) imply that \( X_1 + X_2 \overset{P}{\geq} Y_1 + Y_2 \). This result can be obtained by considering certain order preserving functionals. First introduce some simplifying notations.

For \( s > 0 \), define

\[
\chi_s(x) = I(|x| < s)
\]

and, for a symmetric function \( h \), define

\[
h(s,x) = (h * \chi_s)(x) = \int h(x-y) \chi_s(y) dy,
\]
and

\[ h(s,0) = \int h(-y) \chi_s(y) dy \]

\[ = \int h(y) \chi_s(y) dy. \]

Note that

\[ h(s,x) = \begin{cases} \frac{1}{2}[h(x+s,0) - h(x-s,0)] & \text{if } s \leq x \leq s, \\ \frac{1}{2}[h(x+s,0) + h(-x+s,0)] & \text{if } -s \leq x \leq s, \\ \frac{1}{2}[h(-x+s,0) - h(-x-s,0)] & \text{if } x \leq -s. \end{cases} \]

We need the following lemma.

**3.2. Lemma.** Let \( C = \{ h : h \text{ symmetric and } h(s,0) \geq 0 \text{ for all } s > 0 \}. \) Let \( g \) be symmetric and decreasing on \((0, \infty)\). Then \( h * g \in C \) for all \( h \in C \), i.e., \((h * g * \chi_s)(0) \geq 0\) for \( h \in C \) and \( s > 0 \).

**Proof.**

Let \( h \in C \) and \( s > 0 \). Then

\[ (h * g \chi_s)(0) = \int (h \chi_s)(x) g(-x) dx \]

\[ = \int h(s,x) g(-x) dx \]

\[ = \frac{1}{2} \left\{ \int_{x \geq s} [h(x+s,0) - h(x-s,0)] g(x) dx ight. \\

\[ + \int_{-s \leq x \leq s} [h(x+s,0) - h(-x+s,0)] g(x) dx \\

\[ + \int_{x < -s} [h(-x+s,0) - h(-x-s,0)] g(x) dx \right\} \]

\[ = \frac{1}{2} \int_{x \geq s} h(x+s,0) g(x) dx + \int_{x < -s} h(-x-s,0) g(x) dx \\

\[ + \int_{x < -s} h(-x-s,0) g(x) dx - \int h(x-s,0) g(x) dx. \]
Let $y = x + s$ in the first integral, $y = -x + s$ in the second integral, $y = -x - s$ in the third integral and $y = x - s$ in the fourth integral.

We get

$$
(h \ast g \ast \chi_s)(0) = \frac{1}{2} \int_{y>0} h(y,0)g(y-s)dy
$$

$$
+ \int_{y>0} h(y,0)g(-y+s)dy - \int_{y>0} h(y,0)g(-y-s)dy
$$

$$
- \int_{y>0} h(y,0)g(y+s)dy.
$$

By the symmetry of $g$,

$$
(h \ast g \ast \chi_s)(0) = \int_{y>0} h(y,0)[g(y-s) - g(y+s)]dy.
$$

Since $h(y,0) \geq 0$ for all $y > 0$, and $g(y-s) - g(y+s) \geq 0$ for $y > 0$ and $s > 0$, we conclude that $(h \ast g \ast \chi_s)(0) \geq 0$.

We may now obtain the following result.

3.3. Theorem. Let $X_1, X_2, Y_1, Y_2$ be independent symmetric random variables on $(-1,1)$ with densities $f_1, f_2, g_1, g_2$, respectively. Let $f_1, g_2$ be non-increasing on $(0,1)$. Let $X_i \overset{p}{\geq} Y_i$, for $i = 1, 2$.

Then $X_1 + X_2 \overset{p}{\geq} Y_1 + Y_2$.

Proof.

We will first establish that $X_1 + X_2 \overset{p}{\geq} X_1 + Y_2$.

Fix $f_1$. For each $f \in L_1(0,1)$, let $f_s$ be the symmetric function on $(-1,1)$ defined by

$$
f_s(t) = f(|t|).
$$
For each $s > 0$, define a functional on $L_1(0,1)$ by

$$
\phi_s(f) = \int \int I(|x_1 + x_2| \leq s) f_1(x_1) f_S(x_2) \, dx_1 \, dx_2.
$$

Let $T(0,1)$ be the class of nonnegative functions $u$ on $(0,1)$ with

$$
\int_0^1 u(t) \, dt = \frac{1}{2}.
$$

Note that for $f \in T(0,1)$,

$$
\phi_s(f) = P(|X_1 + Z| \leq s),
$$

where $X_1$, $Z$ are independent random variables with densities $f_1$, $f_S$ respectively.

We shall show that for each $s > 0$, $\phi_s$ is non-decreasing with respect to the ordering of unrestricted majorization on $T(0,1)$. $s > 0$. Let $f \in T(0,1)$ and $h \in T$. Then,

$$
\frac{\partial \phi_s}{\partial h}(f) = \lim_{\theta \to 0} \frac{1}{\theta} \int \int I(|x_1 + x_2| \leq s) \left\{ f_1(x_1) [f_S(x_2) + \theta h_S(x_2)] - f_1(x_1) f_S(x_2) \right\} \, dx_1 \, dx_2
$$

$$
= \int \int I(|x_1 + x_2| \leq s) f_1(x_1) h_S(x_2) \, dx_1 \, dx_2
$$

$$
= (f_1 h_S * h_S)(0).
$$

Since $h \in T$, $h_S(t,0) \geq 0$ for all $t > 0$. By Lemma 3.2, $(f_1 h_S * h_S)(0) \geq 0$. It now follows from Theorem 2.12 that $\phi_s$ is nondecreasing with respect to the ordering of unrestricted majorization on $T(0,1)$.

Note that $X_2 \overset{P}{\preceq} Y_2$ implies that $f_2 \overset{P}{\preceq} g_2$ when these are considered as elements of $T(0,1)$. We now have

$$
P(|X_1 + X_2| \leq s) = \phi_s(f_2)
$$

$$
\geq \phi_s(g_2)
$$

$$
= P(|X_1 + Y_2| \leq s), \text{ for all } s > 0.
$$
Thus $x_1 + x_2 \geq x_1 + y_2$. Similarly, we can establish that
$x_1 + y_2 \geq y_1 + y_2$. Hence $x_1 + x_2 \geq y_1 + y_2$. ||
REFERENCES


Schur-Ostrowski Theorems for Functionals on $L_1(0,1)$.

Wai Chan, Frank Proschan, Jayaram Sethuraman

The Florida State University
Department of Statistics
Tallahassee, Florida 32306

U.S. Army Research Office-Durham
P.O. Box 12211
Research Triangle Park, NC 27709
and
U.S. Air Force
Air Force Office of Scientific Research
Bolling Air Force Base, DC 20332

Approved for public release; distribution unlimited

Inequalities, majorization, Muirhead's Theorem, peakedness in symmetric distributions, rearrangement, Schur functions, Schur-Ostrowski's Theorem.

Hardy, Littlewood and Pólya (1934) introduced the partial ordering of majorization among $n$-dimensional real vectors. Many well known inequalities can be recast as the statement that certain functions are increasing with respect to this ordering. Such functions are said to be Schur-convex. An important result in the theory of majorization is the Schur-Ostrowski Theorem.
ABSTRACT:

which characterizes Schur-convex functions. The concept of majorization has been extended to elements of \( L_1(0,1) \) by Ryff (1963). A functional of \( L_1(0,1) \) that is increasing with respect to the ordering of majorization is said to be Schur-convex. In this paper, we prove an analogue of the Schur-Ostrowski condition which characterized Schur-convex functionals in terms of their Gâteaux differentials. We also introduce another partial ordering in \( L_1(0,1) \) called unrestricted majorization. This partial ordering is similar to majorization but does not involve the use of decreasing rearrangements. We establish a characterization of non-decreasing functionals on \( L_1(0,1) \) with respect to the partial ordering of unrestricted majorization through another analogue of the Schur-Ostrowski condition.