*Inequalities in Statistics and Probability*
IMS Lecture Notes–Monograph Series Vol. 5 (1984), 4–12

# STOCHASTIC VERSIONS OF REARRANGEMENT INEQUALITIES[1]

BY CATHERINE D'ABADIE and FRANK PROSCHAN
*AT&T Technologies, Inc. and Florida State University*

This paper develops a unified way of obtaining stochastic versions of deterministic rearrangement inequalities. Rearrangement inequalities compare the value of a function of vector arguments with the value of the same function after the components of the vectors have been rearranged. The classical example of a rearrangement inequality is the well-known inequality of Hardy, Littlewood, and Pólya for sums of products. They show that if $a_1 \geqslant \dots \geqslant a_n$ and $b_1 \geqslant \dots \geqslant b_n$ are positive numbers, then for every permutation $(\pi(1), \dots, \pi(n))$ of $(1, \dots, n)$ the inequalities $\sum_{i=1}^n a_i b_i \geqslant \sum_{i=1}^n a_i b_{\pi(i)} \geqslant \sum_{i=1}^n a_i b_{n-i+1}$ hold.

The function $\Sigma xy$ is an example from a class of functions called arrangement increasing functions for which such rearrangement inequalities hold. Given two nonnegative random vectors $\mathbf{X}$ and $\mathbf{Y}$ with joint density $f(\mathbf{x},\mathbf{y})$ we determine conditions on $f$ for the stochastic rearrangment inequalities $g(X_1, \dots, X_n; Y_1, \dots, Y_n) \overset{st}{\geqslant} g(X_1, \dots, X_n; Y_{\pi(1)}, \dots, Y_{\pi(n)}) \overset{st}{\geqslant} g(X_1, \dots, X_n; Y_n, \dots, Y_1)$ to hold for every permutation $\pi$ and arrangement increasing function $g$. We present a number of examples of densities which satisfy the condition.

**1. Introduction.** The development of stochastic versions of deterministic concepts arising in mathematics has, in the past, led to important new results in probabiliy and statistics. The subject of this paper is in this spirit.

Specifically, we obtain *stochastic versions of rearrangement inequalities*. Rearrangement inequalities compare the value of a function of vector arguments with the value of the same function after the components of the vectors have been rearranged.

The classical example of a rearrangement inequality involving a function of two vector arguments is the well-known inequality of Hardy, Littlewood, and Pólya (1952) for sums of products. For vectors $\mathbf{a} = (a_1, \dots, a_n)$ and $\mathbf{b} = (b_1, \dots, b_n)$ of positive numbers, Hardy, Littlewood, and Pólya show that the function $f(\mathbf{a},\mathbf{b}) = \sum_{i=1}^n a_i b_i$ takes its largest value when the components of each $\mathbf{a}$ and $\mathbf{b}$ are arranged in increasing (or, equivalently, decreasing) order, and that $f$ takes its smallest value when the components of one of the vectors are arranged in increasing order and those of the other vector are arranged in decreasing order. In symbols they show that if $a_1 \geqslant a_2 \geqslant \dots \geqslant a_n$ and $b_1 \geqslant b_2 \geqslant \dots \geqslant b_n$ (after relabelling, say) for every permutation $(\pi(1), \dots, \pi(n))$ of $(1, \dots, n)$, then the inequalities:

$$\sum_{i=1}^n a_i b_i \geqslant \sum_{i=1}^n a_i b_{\pi(i)} \geqslant \sum_{i=1}^n a_i b_{n-i+1} \tag{1.1}$$

hold.

*The original idea which motivated this work was to obtain a stochastic version of the inequalities in (1.1).* More explicitly, given two random vectors $\mathbf{X} \equiv (X_1, \dots, X_n)$ and $\mathbf{Y} \equiv (Y_1, \dots, Y_n)$, we wished to determine conditions to impose on $\mathbf{X}$ and $\mathbf{Y}$ to yield for every permutation $\pi$ the stochastic inequalities:

$$\sum_{i=1}^n X_i Y_i \overset{st}{\geqslant} \sum_{i=1}^n X_i Y_{\pi(i)} \overset{st}{\geqslant} \sum_{i=1}^n X_i Y_{n-i+1}, \tag{1.2}$$

where $X \overset{st}{\geqslant} Y$ means $P(X \geqslant t) \geqslant P(Y \geqslant t)$ for all $t$.

---

[1] The research of both authors was supported under Contract AFOSR 82–K–0007 with the Air Force Office of Scientific Research, AFSC, U.S.A.F.

*AMS 1980 subject classifications.* 60C05.

Key words and phrases: Rearrangement inequalities, stochastic rearrangement inequalities, arrangement increasing, Hardy-Littlewood-Pólya inequality, permutations, preservation of properties, positive set function, monotone power function.

As it happens, it is not hard to obtain such a stochastic version of the Hardy–Littlewood–Pólya inequality. Sufficient conditions on $\mathbf{X}$ and $\mathbf{Y}$ for (1.2) to hold can be easily stated. Suppose that $\mathbf{X}$ and $\mathbf{Y}$ are nonnegative random vectors having a joint density $f(\mathbf{x},\mathbf{y})$. For a given vector $\mathbf{x}$ we write $\mathbf{x} \geqslant^{t_{ij}} \mathbf{x}'$ if $i<j$, $x_i \leqslant x_j$, and $\mathbf{x}'$ is obtained from $\mathbf{x}$ by interchanging $x_i$ and $x_j$ and leaving the other components fixed. Then inequality (1.2) holds for $\mathbf{X}$ and $\mathbf{Y}$ if for all pairs $i, j$, $1 \leqslant i < j \leqslant n$, $f$ satisfies:

$$f(\mathbf{x},\mathbf{y}) + f(\mathbf{x}',\mathbf{y}') - f(\mathbf{x}',\mathbf{y}) - f(\mathbf{x},\mathbf{y}') \geqslant 0, \tag{1.3}$$

where $\mathbf{x} \geqslant^{t_{ij}} \mathbf{x}'$ and $\mathbf{y} \geqslant^{t_{ij}} \mathbf{y}'$.

Since the work of Hardy, Littlewood, and Pólya (1952), papers on inequalities involving rearrangements of vectors in $\mathscr{R}^n$ have appeared widely in the literature. Marshall and Olkin (1979) present a unified approach to the study of deterministic rearrangement inequalities.

We develop a theory which offers a *unified approach* to the task of obtaining stochastic versions of rearrangement inequalities. Our work generalizes that of previous authors in that we obtain their deterministic inequalities as special cases. In this paper *we present an overview of the theory we develop and some applications to statistics.*

**2. Deterministic Rearrangement Inequalities.** A deterministic rearrangement inequality compares the value of a function of vector arguments with the value of the same function after the components of the vectors have been rearranged. In the case of two vectors, rearrangement inequalities have the form

$$f(\overset{\uparrow}{\mathbf{x}},\overset{\uparrow}{\mathbf{y}}) = f(\overset{\downarrow}{\mathbf{x}},\overset{\downarrow}{\mathbf{y}}) \geqslant f(\mathbf{x},\mathbf{y}) \geqslant f(\overset{\uparrow}{\mathbf{x}},\overset{\downarrow}{\mathbf{y}}) = f(\overset{\downarrow}{\mathbf{x}},\overset{\uparrow}{\mathbf{y}}), \tag{2.1}$$

where $\overset{\uparrow}{\mathbf{z}}$ ($\overset{\downarrow}{\mathbf{z}}$) denotes the vector whose components are those of $\mathbf{z}$ arranged in increasing (decreasing) order.

The classical rearrangement inequality is the inequality of Hardy, Littlewood, and Pólya (1952) where $f(\mathbf{x},\mathbf{y}) = \Sigma x_i y_i$. As we have noted above the inequality states that if $x_1 \geqslant x_2 \geqslant \ldots \geqslant x_n$ and $y_1 \geqslant y_2 \geqslant \ldots \geqslant y_n$ are nonnegative numbers, then for every permutation $\pi$ of the subscripts of $\mathbf{y}$, the inequalities

$$\Sigma_{i=1}^n x_i y_i \geqslant \Sigma_{i=1}^n x_i y_{\pi(i)} \geqslant \Sigma_{i=1}^n x_i y_{n-1+1}$$

hold.

Rearrangement inequalities involving functions of vectors in $\mathscr{R}^n$ have been widely studied in the literature. Jurkat and Ryser (1966) obtained rearrangement inequalities for functions of $\min(x,y)$. They show that for nonnegative $n$-tuples $\overset{\downarrow}{\mathbf{x}}$ and $\overset{\downarrow}{\mathbf{y}}$

$$\Pi_{i=1}^n \min(x_i,y_i) \geqslant \Pi_{i=1}^n \min(x_i,y_{\pi(i)}) \geqslant \Pi_{i=1}^n \min(x_i,y_{n-i+1})$$

and

$$\Sigma_{i=1}^n \min(x_i,y_i) \geqslant \Sigma_{i=1}^n \min(x_i,y_{\pi(i)}) \geqslant \Sigma_{i=1}^n \min(x_i,y_{n-1+1})$$

for all permutations $\pi$.

Minc (1971) obtained similar rearrangement inequalities for products and sums of $\max(x,y)$.

Rearrangement inequalities also hold for a number of well-known test statistics. An example is Pearson's product moment correlation coefficient given by

$$r(\mathbf{x},\mathbf{y}) = \Sigma_{i,j}(x_i - x_j)(y_i - y_j)\,[\Sigma_{i,j}(x_i - x_j)^2\, \Sigma_{i,j}(y_i - y_j)^2]^{-1/2}$$

Spearman's $\rho$ and Kendall's correlation coefficient $\tau$ also yield rearrangement inequalities.

Rearrangement inequalities can be obtained for Blomquist's quadrant test. Blomquist (1950) proposed the following test for positive association:

$$\beta(\mathbf{x},\mathbf{y}) = \Sigma_{i=1}^{n}[a_1(x_i)b_1(y_i) + a_2(x_i)b_2(y_i)],$$

where

$$a(x_i) = 0, \text{ if } x_i \leqslant x_{\text{med}}$$
$$= 1, \text{ if } x_i > x_{\text{med}};$$

$b(y_i)$ is similary defined, $a_2 = 1 - a_1$, and $b_2 = 1 - b_1$.

**3. The Arrangement Ordering and Arrangement Increasing Functions.** All rearrangement inequalities, such as the ones just described, are examples of functions which are increasing in a partial ordering on $\mathcal{R}^n \times \mathcal{R}^n$. This partial ordering, implicit in the work of Hollander, Proschan, and Sethuraman (1977), is defined in Marshall and Olkin (1979). They refer to the partial ordering as the arrangement ordering. Using this ordering they obtain refinements of rearrangement inequalities involving many more comparisons than given in the examples in the previous section.

To define the arrangement ordering we need some terminology and notation.

Let $S_n$ denote the group of all permutations of $\{1,2, \dots, n\}$. An element of $S_n$ will be denoted by $\pi \equiv (\pi(1), \dots, \pi(n))$. Let $\pi$ and $\pi'$ be elements of $S_n$. We say that $\pi'$ is a *simple transposition* of $\pi$ if there exist positive integers $1 \leqslant i < j \leqslant n$ such that $\pi(i) = \pi'(j) < \pi'(i) = \pi(j)$ and $\pi(k) = \pi'(k)$ for $k \neq i, j$. We write this as $\pi <^{t_{ij}} \pi'$. For $\pi,\pi'$ in $S_n$ we say that $\pi'$ is a *transposition* of $\pi$, written $\pi \overset{t}{\leqslant} \pi$, if $\pi = \pi'$ or if $\pi'$ can be obtained from $\pi$ by a sequence of simple transpositions.

For a vector $\mathbf{x}$ in $\mathcal{R}^n$, we define $\mathbf{x}\pi$ to be the vector $(x_{\pi(1)}, \dots, x_{\pi(n)})$. Recall that we denote by $\vec{\mathbf{x}}$ the vector obtained from $\mathbf{x}$ by arranging the components of $\mathbf{x}$ in increasing order. We say that $\mathbf{x}'$ is a *transposition* of $\mathbf{x}$ if $\mathbf{x} = \vec{\mathbf{x}}\pi$, $\mathbf{x}' = \vec{\mathbf{x}}\pi'$ where $\pi \overset{t}{\leqslant} \pi'$. We write $\mathbf{x} \overset{t}{\leqslant} \mathbf{x}'$. We note that this defines a partial ordering of $\mathcal{R}^n$. This partial ordering has been studied by Savage (1957), Lehmann (1966), and Hollander, Proschan, and Sethuraman (1977), among others.

Let $(\mathbf{x},\mathbf{y}) \in \mathcal{R}^n \times \mathcal{R}^n$. The *orbit* of $(\mathbf{x},\mathbf{y})$ is the set $0_{\mathbf{x},\mathbf{y}} = \{(\mathbf{x}\pi,\mathbf{y}\sigma)\colon \pi,\sigma \in S_n\}$. For a vector $\mathbf{x} \in \mathcal{R}^n$ the orbit of $\mathbf{x}$ is defined similarly.

*Definition 3.1.* Let $(\mathbf{x},\mathbf{y})$ and $(\mathbf{x}',\mathbf{y}')$ be two elements of $\mathcal{R}^n \times \mathcal{R}^n$ belonging to the same orbit. We say that $(\mathbf{x},\mathbf{y})$ is *more similarly arranged* than $(\mathbf{x}',\mathbf{y}')$ if there exist $\pi,\sigma \in S_n$ such that $\mathbf{x}\pi = \mathbf{x}'\sigma = \vec{\mathbf{x}}$ and $\mathbf{y}\pi \overset{t}{\leqslant} \mathbf{y}'\sigma$. We write $(\mathbf{x},\mathbf{y}) \overset{a}{\leqslant} (\mathbf{x}',\mathbf{y}')$.

This partial ordering of $\mathcal{R}^n \times \mathcal{R}^n$ is referred to as the *arrangement ordering*. We write $(\mathbf{x},\mathbf{y}) \overset{a}{=} (\mathbf{x}',\mathbf{y}')$ if $(\mathbf{x},\mathbf{y}) \overset{a}{\leqslant} (\mathbf{x}',\mathbf{y}')$ and $(\mathbf{x}',\mathbf{y}') \overset{a}{\leqslant} (\mathbf{x},\mathbf{y})$.

Figure 3.1 illustrates the arrangement ordering when $\vec{\mathbf{x}} = (.5,1,3)$ and $\vec{\mathbf{y}} = (2,3.5,4)$. An arrow in the diagram from an element $(\vec{\mathbf{x}},\mathbf{y})$ to an element $(\vec{\mathbf{x}},\mathbf{y}')$ means that $(\vec{\mathbf{x}},\mathbf{y}) \overset{a}{\leqslant} (\vec{\mathbf{x}},\mathbf{y}')$.

$$\begin{tikzcd}
 & ((.5,1,3),(2,3.5,4)) \arrow[dl] \arrow[dr] & \\
((.5,1,3),(3.5,2,4)) \arrow[d] \arrow[drr] & & ((.5,1,3),(2,4,3.5)) \arrow[d] \arrow[dll] \\
((.5,1,3),(3.5,4,2)) \arrow[dr] & & ((.5,1,3),(4,2,3.5)) \arrow[dl] \\
 & ((.5,1,3),(4,3.5,2)) &
\end{tikzcd}$$

FIGURE 3.1. An Illustrative Arrangement Ordering

*Remark.* Let $(\mathbf{x},\mathbf{y})$ denote the largest element of its orbit in the arrangement ordering, that is, $(\mathbf{x},\mathbf{y}) \overset{a}{\geqslant} (\mathbf{x}\pi,\mathbf{y}\sigma)$ for all $\pi,\sigma \in S_n$. Then it is easy to see that $(x_i-x_j)(y_i-y_j) \geqslant 0$ for all pairs $i,j$. In this case we say that $\mathbf{x}$ and $\mathbf{y}$ are *similarly arranged*. (Hardy, Littlewood, and Pólya (1952) use the expression "similarly ordered".) We write $\mathbf{x} \overset{s}{=} \mathbf{y}$.

Functions which are order-preserving with respect to the arrangement ordering were introduced by Hollander, Proschan, and Sethuraman (1977).

*Definition 3.2.* A function $f$ from $\mathcal{R}^n \times \mathcal{R}^n$ into $R$ is said to be *arrangement increasing* if $(\mathbf{x},\mathbf{y}) \overset{a}{\geqslant} (\mathbf{x}',\mathbf{y}')$ implies $f(\mathbf{x},\mathbf{y}) \geqslant f(\mathbf{x}',\mathbf{y}')$ for all $(\mathbf{x},\mathbf{y}) \in \mathcal{R}^n \times \mathcal{R}^n$.

Functions which are arrangement increasing play an important role in the theory we develop. Their properties and many useful applications were first studied by Hollander, Proschan, and Sethuraman (1977). In their 1977 paper they gave an alternative definition of an arrangement increasing function which they call a function "decreasing in transposition". The present name is due to Marshall and Olkin.

PROPOSITION 3.3. (Marshall and Olkin (1979).) *A function $f$ from $\mathcal{R}^n \times \mathcal{R}^n$ into $\mathcal{R}$ is arrangement increasing if and only if* (i) $f(\mathbf{x},\mathbf{y}) = f(\mathbf{x}\pi,\mathbf{y}\pi)$ *for* $(\mathbf{x},\mathbf{y}) \in \mathcal{R}^n \times \mathcal{R}^n$, $\pi \in S_n$, *and* (ii) $f(\vec{\mathbf{x}},\mathbf{y}) \geqslant f(\vec{\mathbf{x}},\mathbf{y}')$, *where* $\mathbf{y} \overset{t}{\geqslant} \mathbf{y}'$.

A function satisfying (i) of Proposition 3.3 is called *permutation invariant*.

Hollander, Proschan, and Sethuraman (1977) give many examples of arrangement increasing functions including a number of well-known densities in statistics. Some of these examples are presented in Section 5.

**4. Stochastic Rearrangement Inequalities.** In this section we obtain stochastic versions of the rearrangement inequalities of Section 2. Specifically, we show the following. Let $\mathbf{X}$ and $\mathbf{Y}$ be nonnegative random $n$-vectors with joint density $f$ satisfying the conditions of Def. 4.1 below. Then for any arrangement increasing function $g$ and permutation $\pi$ we have

$$g(X_1, \ldots, X_n; Y_1, \ldots, Y_n) \overset{st}{\geqslant} g(X_1, \ldots, X_n; Y_{\pi(1)}, \ldots, Y_{\pi(n)}) \overset{st}{\geqslant} g(X_1, \ldots, X_n; Y_n, \ldots, Y_1).$$

These stochastic rearrangement inequalities follow as a corollary to our main result presented in Theorem 4.3. The condition we need on the joint density of $\mathbf{X}$ and $\mathbf{Y}$ to obtain stochastic rearrangement inequalities is defined as follows.

*Definition 4.1.* A function $f$ from $\mathcal{R}^n \times \mathcal{R}^n$ into $\mathcal{R}$ is called a *positive set function in arrangement* (PSA) if $\mathbf{x} \geqslant^{t_{ij}} \mathbf{x}'$ and $\mathbf{y} \geqslant^{t_{ij}} \mathbf{y}'$ for any pair $i<j$ imply

$$f(\mathbf{x},\mathbf{y}) - f(\mathbf{x}',\mathbf{y}) - f(\mathbf{x},\mathbf{y}') + f(\mathbf{x}',\mathbf{y}') \geqslant 0.$$

We note that a function $f$ is arrangement increasing if and only if $f$ is PSA and permutation invariant.

Before we state the main theorem, we introduce the notion of an arrangement preserving kernel.

*Definition 4.2.* A function $K$ from $(\mathcal{R}^n \times \mathcal{R}^n) \times (\mathcal{R}^n \times \mathcal{R}^n)$ into $\mathcal{R}$ is called an *arrangement preserving (AP) kernel* if: (i) $K(\mathbf{u},\mathbf{x}; \mathbf{v},\mathbf{y})$ is permutation invariant in $(\mathbf{u},\mathbf{x})$ and in $(\mathbf{v},\mathbf{y})$, and (ii) For all $\mathbf{u},\mathbf{v} \in \mathcal{R}^n$, $K(\mathbf{u},\mathbf{x}; \mathbf{v},\mathbf{y})$ is PSA in $(\mathbf{x},\mathbf{y})$.

In our main result Theorem 4.3 below we state that the arrangement increasing property is preserved under an integral transform defined by an arrangement preserving kernel.

THEOREM 4.3. *Let $f(\mathbf{x},\mathbf{y})$ be arrangement increasing and let $K(\mathbf{u},\mathbf{x}; \mathbf{v},\mathbf{y})$ be an arrangement preserving kernel. Then under mild conditions on the measure $m$ and the assumption that the integral exists finitely, the function*

$$g(\mathbf{u},\mathbf{v}) = \int\int f(\mathbf{x},\mathbf{y})K(\mathbf{u},\mathbf{x};\mathbf{v},\mathbf{y})m(d\mathbf{x},d\mathbf{y})$$

*is arrangement increasing.*

A corollary to Theorem 4.3 yields stochastic versions of deterministic rearrangement inequalities.

COROLLARY 4.4. *Let* $(\mathbf{X},\mathbf{Y})$ *have a PSA density. Then for all arrangement increasing functions f and all permutations* $\pi$ *we have*

$$f(X_1, \ldots, X_n; Y_1, \ldots, Y_n) \overset{st}{\geqslant} f(X_1, \ldots, X_n; Y_{\pi(1)}, \ldots, Y_{\pi(n)}) \overset{st}{\geqslant} f(X_1, \ldots, X_n; Y_n, \ldots, Y_1).$$

Since the function $g(\mathbf{x},\mathbf{y}) = \Sigma x_i y_i$ is arrangement increasing we have as a consequence of Corollary 4.4 a stochastic version of the Hardy, Littlewood, and Pólya inequality, namely that

$$\Sigma X_i Y_i \overset{st}{\geqslant} \Sigma X_i Y_{\pi(i)} \overset{st}{\geqslant} \Sigma X_i Y_{n-i+1}.$$

A similar result holds for all the other rearrangement inequalities in Section 2.

**5. Examples of PSA Functions and AP Kernels.** The results in the previous section allow us to obtain stochastic versions of rearrangement inequalities for a large class of random vectors which contains those pairs $(\mathbf{X},\mathbf{Y})$ having PSA and AP densities. The purpose of this section is to show that many multivariate densities of interest in statistical practice fall into these two classes of functions.

A function $\phi$ is called a *positive set function* if

$$\phi(x_1,y_1) - \phi(x_1,y_2) - \phi(x_2,y_1) + \phi(x_2,y_2) \geqslant 0 \text{ for all } x_1 < x_2 \text{ and } y_1 < y_2.$$

Positive set functions can be used to construct AP kernels as we state in Theorem 5.1.

THEOREM 5.1. *Let* $\phi$ *be a positive set function and let* $g_1$ *and* $g_2$ *be arrangement increasing. Then* $\phi(g_1(\mathbf{u},\mathbf{x}), g_2(\mathbf{v},\mathbf{y}))$ *is an AP kernel.*

Some examples of positive set functions are (i) $\phi(x,y) = xy$, (ii) $\phi(x,y) = F(x,y)$ where $F$ is a c.d.f., and (iii) $\phi(x,y) = h(x-y)$ where $h$ is concave.

As a consequence of Theorem 5.1 and the fact that the product $xy$ is a positive set function we have the following important example showing how to construct AP kernels

*Example 5.2.* Let $\mathbf{X}$ and $\mathbf{Y}$ be independent random vectors each having arrangement increasing density. Then the joint density of $(\mathbf{X},\mathbf{Y})$ is an AP kernel.

The following examples of AI densities can be used to construct AP densities. (See Hollander, Proschan, and Sethuraman (1977) for proofs.)

*5.3.a. Multinomial:* $g_1(\mathbf{u},\mathbf{x}) = N!\,\Pi_{i=1}^n(u_i^{x_i}/x_i!)$, where $0 < u_i < 1$, $x_i = 0,1,2, \ldots$, $i = 1, \ldots, n$, $\Sigma_{i=1}^n u_i = 1$, and $\Sigma_{i=1}^n x_i = N$.

*5.3.b. Negative multinomial:*

$$g_2(\mathbf{u},\mathbf{x}) = (\Gamma(N))^{-1}\Gamma(N + \Sigma_{i=1}^n x_i)(1 + \Sigma_{i=1}^n u_i)^{-N-\Sigma_{i=1}^n x_i}\Pi_{i=1}^n(u_i^{x_i}/x_i!),$$

where $u_i > 0$, $x_i = 0,1, \ldots, i = 1, \ldots, n$, and $N > 0$.

*5.3.c. Multivariate hypergeometric:* $g_3(\mathbf{u},\mathbf{x}) = \Pi_{i=1}^n\binom{u_i}{x_i} \Big/ \binom{\Sigma_{i=1}^n u_i}{N}$, where $u_i > 0$, $x_i = 0,1, \ldots, \Sigma_{i=1}^n x_i = N < \Sigma_{i=1}^n u_i$.

*5.3.d. Dirichlet:* $g_4(\mathbf{u},\mathbf{x}) = (\Gamma(\theta)\Pi_{i=1}^n G(u_i))^{-1}\Gamma(\theta + \Sigma_{i=1}^n u_i)(1 - \Sigma_{i=1}^n x_i)^{\theta-1}\Pi_{i=1}^n x_i^{u_i-1}$, where $u_i > 0$, $x_i \geqslant 0$, $i = 1, \ldots, n$. $\Sigma_{i=1}^n x_i \leqslant 1$, and $\theta > 0$.

*5.3.e. Inverted Dirichlet:*

$g_5(\mathbf{u},\mathbf{x}) = (\Gamma(\theta)\Sigma_{i=1}^n\Gamma(u_i))^{-1}\Gamma(\theta + \Sigma_{i=1}^n u_i)\Pi_{i=1}^n x_i^{u_i-1}/((1 + \Sigma_{i=1}^n x_i)^\theta + \Sigma_{i=1}^n x_i)$,

where $u_i > 0, x_i \geq 0, i = 1, \ldots, n$, *and* $\theta > 0$.

5.3.f. Negative multivariate hypergeometric:

$$g_6(\mathbf{u},\mathbf{x}) = (\Pi_{i=1}^n x_i!\Gamma(N + \Sigma_{i=1}^n u_i))^{-1}N!\,(\Sigma_{i=1}^n u_i)\Pi_{i=1}^n(\Gamma(x_i+u_i)/\Gamma(u_i)),$$

where $u_i > 0, x_i = 0,1, \ldots, N, \Sigma_{i=1}^n x_i = N$, and $N = 1,2, \ldots$ .

*5.3.g. Dirichlet compound negative multinomial:* $g_7(\mathbf{u},\mathbf{x}) = (\Pi_{i=1}^n x_i!\Gamma(N)\Gamma(\theta)\Gamma(N + \theta + \Sigma_{i=1}^n u_i + \Sigma_{i=1}^n x_i))^{-1}\,\Gamma(N - \Sigma_{i=1}^n x_i)\,\Gamma(\theta + \Sigma_{i=1}^n u_i)\,\Gamma(N + \theta)\,\Pi_{i=1}^n(\Gamma(x_i + u_i)/\Gamma(u_i))$, where $u_i > 0, x_i = 0,1, \ldots, i = 1, \ldots, n, \theta > 0$, and $N = 1,2, \ldots$ .

*5.3.h. Multivariate logarithmic series distribution:*

$$g_8(\mathbf{u},\mathbf{x}) = (\log(1 + \Sigma_{i=1}^n u_i))^{-1}(\Sigma_{i=1}^n x_i - 1)!\,(1 + \Sigma_{i=1}^n u_i)^{-\Sigma_{i=1}^n x_i}\Pi_{i=1}^n(u_i^{x_i}/x_i!),$$

where $u_i > 0, x_i = 0,1, \ldots, i = 1, \ldots, n$.

*5.3.i. Multivariate F distribution:*

$$g_9(\mathbf{u},\mathbf{x}) = (2\Pi_{i=0}^n\Gamma(\lambda_i)(\lambda_0 + \Sigma_{i=1}^n\lambda_i x_i)^\lambda)^{-1}\,\Gamma(u)\Pi_{i=0}^n(2u_i)^{u_i}\Pi_{i=1}^n x_i^{u_i-1},$$

where $u_i > 0, i = 0,1, \ldots, n, u = \Sigma_{i=0}^n\lambda_i, x_i \geq 0, i = 0,1, \ldots, n$.

*5.3.j. Multivariate Pareto distribution:* $g_{10}(\mathbf{u},\mathbf{x}) = (\Pi_{i=1}^n u_i)^{-1}(\Sigma_{i=1}^n\lambda_i^{-1}x_i - n + 1)^{-(a+n)}$, where $x_i > u_i > 0, i = 1, \ldots, n$, and $a > 0$.

*5.3.k. Multivariate normal distribution with common variance and common covariance:* $g_{11}(\mathbf{u},\mathbf{x}) = (2\pi)^{n/2}|\Sigma|^{-1/2}\exp(-\tfrac{1}{2}(\mathbf{x}-\mathbf{u})\Sigma^{-1}(\mathbf{x}-\mathbf{u})^{-1})$, where $\Sigma$ is the positive definite covariance matrix with elements $\sigma^2$ along the main diagonal and elements $\rho\sigma^2$ elsewhere, $\rho > -1/(n-1)$.

AP densities can also be constructed from independent $TP_2$ densities using the next result. Recall that $f$ is $TP_2$ if

$$f(x_1,y_1)f(x_2,y_2) \geq f(x_1,y_2)f(x_2,y_1) \text{ for all } x_1 < x_2 \text{ and } y_1 < y_2.$$

THEOREM 5.4. *Let $f$ be $TP_2$. Then $\Pi f(\theta_i,y_i)$ is arrangement increasing.*

Some examples of $TP_2$ densities are (i) normal with variance 1, (ii) exponential, and (iii) Poisson.

In the next example we use a result of Karlin for $TP_2$ densities to obtain more PSA densities.

Suppose that components have lifelengths $X$ with $TP_2$ density $g(\theta,x)$ and $Y$ with $TP_2$ density $f(\theta,y)$. Further suppose that $X$ and $Y$ are independent and that $\theta$ depends on the environment with distribution $\pi(\theta)$. Then Karlin (1968) has shown that the joint distribution of $X,Y$ given by $K(x,y) = \int f(\theta,y)g(\theta,y)d\pi(\theta)$ is $TP_2$. From this result we get the following:

THEOREM 5.5. *Let $(X_1,Y_1), \ldots, (X_n,Y_n)$ be independent with $TP_2$ density. Then the joint density of $(\mathbf{X},\mathbf{Y}) = (X_1, \ldots, X_n; Y_1, \ldots, Y_n)$ is arrangement increasing and hence PSA.*

In the next two results we state that certain operations on pairs of random vectors which are PSA (AP) preserve the PSA (AP) property.

A vector $\mathbf{x} \in \mathcal{R}^n$ *majorizes* $\mathbf{y} \in \mathcal{R}^n$ if $\Sigma_{i=1}^k x_{[i]} \geq \Sigma_{i=1}^k y_{[i]}$ for $k = 1, \ldots, n$ and $\Sigma_{i=1}^n x_i = \Sigma_{i=1}^n y_i$ where $x_{[1]} \geq \ldots \geq x_{[n]}$. A function $f$ is *Schur-concave if $f(\mathbf{x}) \leq f(\mathbf{y})$* whenever $\mathbf{x}$ majorizes $\mathbf{y}$. In Theorem 5.6 we show that Schur-concave densities can be used to construct PSA densities.

THEOREM 5.6. *Let* $(\mathbf{X},\mathbf{Y})$ *have a PSA (AP) density. Let* $\mathbf{W}$ *and* $\mathbf{Z}$ *be mutually independent and independent of* $(\mathbf{X},\mathbf{Y})$ *each having a Schur-concave density. Then* $(\mathbf{X} + \mathbf{W},\mathbf{Y} + \mathbf{Z})$ *has a PSA (AP) density.*

Marshall and Olkin (1974) give a number of interesting examples of Schur-concave densities. We list some of them in Example 5.7 below.

*Example 5.7.* The following multivariate densities are Schur-concave:

*5.7.a. Multivariate normal:* Let $X_1, \ldots, X_n$ be exchangeable and jointly normally distributed. Then the joint density of $\mathbf{X}$ is Schur-concave.

*5.7.b. Multivariate "t" distribution:* Let $U_1, \ldots, U_n$ be exchangeable and jointly normally distributed. Let $Z^2$ be chi-square distributed for $Z \geqslant 0$. Then $\mathbf{X} = (U_1/Z, \ldots, U_m/Z)$ has a Schur-concave density.

*5.7.c. Multivariate beta distribution:* Let $U_1, \ldots, U_n$ be independent, identically distributed chi-square random variables and let $Z$ be a chi-square random variable, indpendent of $U_1, \ldots, U_n$. Then $\mathbf{X} = (U_1/(\Sigma U_i+Z), \ldots, U_m/(\Sigma U_i+Z)$ has a Schur-concave density.

*5.7.d. Multivariate "F" distribution:* Let $U_1, \ldots, U_n$ each have a chi-square distribution with $r \geqslant 2$ degress of freedom, and let $Z$ have a chi-square distribution with $s$ degrees of freedom. Then $\mathbf{X} = (U_1/Z, \ldots, U_n/Z)$ has a Schur-concave density.

For random variables $X_1, \ldots, X_n$, denote by $R_i$ the rank of $X_i$ among $X_1, \ldots, X_n$. The random vector $\mathbf{R} \equiv (R_1, \ldots, R_n)$ is called the *rank order* of $(X_1, \ldots, X_n)$. We have the following useful result for the rank orders of PSA (AP) random vectors $(\mathbf{X},\mathbf{Y})$.

THEOREM 5.8. *Let* $(\mathbf{X},\mathbf{Y})$ *be PSA (AP). Let* $\mathbf{R}$ *be the rank order of* $\mathbf{X}$ *and let* $\mathbf{S}$ *be the rank order of* $\mathbf{Y}$. *Then the random vectors* $(\mathbf{R},\mathbf{S})$ *are PSA (AP).*

**6. Applications to Statistics.** The theory of stochastic rearrangement inequalities has applications in a number of areas in statistics. In this section we present two of these applications.

In the first example we show how the stochastic version of the Hardy, Littlewood, and Pólya inequality may be applied to reliability theory. This generalizes a result of Derman, Lieberman, and Ross (1972) in the case of two vectors.

*Application 6.1.* Suppose that we have two stockpiles of $n$ components each, stockpile one of type 1 components, stockpile two of type 2 components. From these stockpiles we are to construct $n$ systems, each composed of a component of type 1 and a component of type 2 arranged in series. A component $i$ of type $j$ has a random reliability $p_i^j$, $j = 1,2$; $i = 1, \ldots, n$. We assume that $\mathbf{P}^1 \equiv (P_1^1, \ldots, P_n^1)$ and $\mathbf{P}^2 \equiv (P_1^2, \ldots, P_n^2)$ are independent, each having an AI density with parameters $\alpha_1 \leqslant \ldots \leqslant \alpha_n$ and $\beta_1 \leqslant \ldots \leqslant \beta_n$, respectively. Then, as we have seen in Section 4, $(\mathbf{P}^1,\mathbf{P}^2)$ is AP.

For the assembly which pairs the $i$-th component of type 1 with the $\pi(i)$-th component of type 2, the average reliability of the $n$ system is $1/n\, \Sigma_{i=1}^n p_i^1 p_{\pi(i)}^2$.

Thus by the stochastic Hardy, Littlewood, and Pólya inequality, the optimal assembly, in terms of average reliability of the $n$ systems, is achieved when the $i$-th component of type 1 is paired with the $i$-th component of type 2. □

Let $(\mathbf{X},\mathbf{Y})$ be AP with parameters $(\alpha,\beta)$ Let $\alpha_0$ be a fixed vector of $\mathcal{R}^n$ in the orbit of $\alpha$. The theory we have developed can be used to study the problem of testing the hypothesis

$$H_0\colon \beta \overset{s}{=} \alpha_0 \quad \text{against}\, H_a\colon \beta \overset{s}{\neq} \alpha_0. \tag{6.1}$$

Let $f$ be an AI function and define the test $T_f$ by

$$T_f(\mathbf{x},\mathbf{y}) = \begin{cases} 1, & \text{if}\, f(\mathbf{x},\mathbf{y}) < \nu_\alpha \\ \gamma, & \text{if}\, f(\mathbf{x},\mathbf{y}) = \nu_\alpha \\ 0, & \text{otherwise.} \end{cases} \tag{6.2}$$

The null hypothesis is rejected with probability $T_f(\mathbf{x},\mathbf{y})$ if $(\mathbf{x},\mathbf{y})$ is observed. Note that the numbers $\nu_\alpha$ and $(0 < \gamma < 1)$ are determined to give size $\alpha$ to the test.

Let $B_{T_f}(\alpha,\beta)$ be the power function of the above test against alternatives $(\alpha,\beta)$, that is,

$$B_{T_f}(\alpha,\beta) = ET_f(X(\alpha),Y(\beta)).$$

We shall need the following definition (see Barlow, Bartholomew, Bremner, and Brunk (1972), Chapter 6).

*Definition 6.2.* Let $(\alpha_0,\beta_0) \in \mathcal{R}^n \times \mathcal{R}^n$ be given. A test $T$ has *isotonic power* against alternative $(\alpha,\beta) \overset{a}{\leq} (\alpha_0,\beta_0)$ (with respect to the ordering "$\overset{a}{\leq}$") if for any $(\alpha_1,\beta_1)$ and $(\alpha_2,\beta_2)$ in $\mathcal{R}^n \times \mathcal{R}^n$ such that

$$(\alpha_2,\beta_2) \overset{a}{<} (\alpha_1,\beta_1) \overset{a}{<} (\alpha_0,\beta_0),$$

we have

$$B_{T_f}(\alpha_2,\beta_2) \geq B_{T_f}(\alpha_1,\beta_1).$$

*Remark 6.3.* It is a consequence of Definition 6.2 that any test $T$ which is isotonic with respect to the "$\overset{a}{\leq}$" ordering is unbiased for testing

$$\begin{aligned} H_0\colon & (\alpha_0,\beta) \overset{a}{=} (\alpha_0,\beta_0) \quad \text{against} \\ H_a\colon & (\alpha_0,\beta) \overset{a}{<} (\alpha_0,\beta_0),\, (\alpha_0,\beta) \overset{a}{\neq} (\alpha_0,\beta_0). \end{aligned} \tag{6.3}$$

Note that by the remark in Section 3, the hypotheses in (6.3) are equivalent to those in (6.1).

It follows from Theorem 4.3, that tests of the form given in (6.2) are isotonic with respect to the arrangement ordering and, consequently, by Remark 6.3 that such tests will be unbiased for testing $H_0$ against $H_a$. We state this formally in Theorem 6.4.

THEOREM 6.4. *Let* $(\mathbf{X},\mathbf{Y})$ *be SSA like* $(\alpha_o,\beta)$. *Consider testing the hypothesis*

$$H_0\colon (\alpha_0,\beta) \overset{a}{=} (\alpha_0,\beta_0)$$

*against*

$$H_a\colon (\alpha_0,\beta) \overset{a}{<} (\alpha_0,\beta),\, (\alpha_0,\beta) \overset{a}{\neq} (\alpha_0,\beta_0).$$

*Let f be an AI function and let* $T_f$ *be the test given in (6.2). Then the test* $T_f$ *has isotonic power against alternatives* $(\alpha_0,\beta) \overset{a}{\leq} (\alpha_0,\beta_0)$. *Consequently, a test based on* $T_f$ *is unbiased for testing* $H_0\colon \alpha_0 \overset{s}{=} \beta$ *against* $H_a\colon \alpha_0 \overset{s}{\neq} \beta$.

A number of well-known statistics, including those given in Section 2 such as Spearman's $\rho$ and Kendall's $\tau$ are AI functions and hence can be used to test the hypotheses in (6.1).

*Remark 6.5.* In Theorem 5.8 we showed that if $(\mathbf{X},\mathbf{Y})$ is PSA (AP) then its rank order $(\mathbf{R},\mathbf{S})$ is PSA (AP). Thus Theorem 6.4 also holds for test statistics $T_f$ based on the rank order of $(\mathbf{X},\mathbf{Y})$. A useful application of the above remark arises in testing for the existence of positive dependence between two time series. An example is described below.

*Application 6.6.* Studies of air pollution have shown that automobile exhaust is the major source of lead elemental air pollution in many urban areas. It is believed that automobile exhaust is also the major source of bromine pollution in the atmosphere. For a particular city, we wish to determine whether automobile exhaust is the predominant source of both of these two pollutants or, alternatively, whether other sources are responsible for bromine pollution. Suppose that $\lambda_i$, the concentration of lead at time $i$, $i = 1, \ldots, n$, is known. Let $\lambda_0 = (\lambda_1, \ldots, \lambda_n)$.

To help in distinguishing between the two alternative hypotheses, we test $H_0$: $\lambda_0 \overset{s}{=} \beta$ against $H_a$: $\lambda_0 \overset{s}{\neq} \beta$, where $\beta_i$ is the true concentration of bromine at time $i$, $i = 1, \ldots, n$, and $\beta = (\beta_1, \ldots, \beta_n)$. Rejection of $H_0$ would indicate that sources other than automobile exhaust contribute to the bromine pollution.

Observations $\mathbf{L}$ on lead and $\mathbf{B}$ on bromine are assumed to be governed by a joint AP density with parameters $(\lambda_0, \beta)$. By Theorem 6.4 we conclude that a test using *an AI test statistic based on the ranks of* $\mathbf{L}$ *and the ranks of* $\mathbf{B}$ *is isotonic and is consequently unbiased against* $H_a$.

*Remark 6.7.* Suppose that the measurements $\mathbf{L}$ and $\mathbf{B}$ are subject to errors $\mathbf{X}$ and $\mathbf{Y}$ with $\mathbf{X} \sim \mathrm{MVN}(\mathbf{0}, \Sigma(\rho_1))$ and $\mathbf{Y} \sim \mathrm{MVN}(\mathbf{0}, \Sigma(\rho_2))$, where

$$\Sigma(\rho) = \begin{pmatrix} \sigma^2 & & \rho\sigma^2 \\ & \ddots & \\ \rho\sigma^2 & & \sigma^2 \end{pmatrix}$$

for $0 \leqslant \rho \leqslant 1$. Since the density of each of $\mathbf{X}$ and $\mathbf{Y}$ is Schur-concave by Theorem 5.6, $(\mathbf{L}+\mathbf{X}, \mathbf{B}+\mathbf{Y})$ is PSA (AP) and, as before, a test using an AI test statistic based on the ranks of $\mathbf{L}+\mathbf{X}$ and of $\mathbf{B}+\mathbf{Y}$ is isotonic.

## REFERENCES


BARLOW, R.E., BARTHOLOMEW, D.J., BREMNER, J.M., and BRUNK, H.D. (1972). *Statistical Inference Under Order Restrictions*. Wiley, New York.

BLOMQUIST, N. (1950). On a measure of dependence between two random variables. *Ann. Math. Statist.* *21* 593–600.

DERMAN, C., LIEBERMAN, G.J., and ROSS, S.M. (1972). On optimal assembly of systems. *Nav. Res. Logist. Quart.* *19* 569–573.

HARDY, G.H., LITTLEWOOD, J.E., and PÓLYA, G. (1952). *Inequalities*. 2nd. ed. Cambridge University Press, London and New York.

HOLLANDER, M., PROSCHAN, F., and SETHURAMAN, J. (1977). Functions decreasing in transposition and their applications in ranking problems. *Ann. Statist.* *5* 722–733.

JURKAT, F.B., and RYSER, H.J. (1966). Matrix factorization of determinants and permanents. *J. Algebra.* *15* 60–77.

KARLIN, S. (1968). *Total Positivity*. Stanford University Press, Stanford, CA.

LEHMANN, E.L. (1966). Some concepts of dependence. *Ann. Math. Statist.* *37* 1137–1153.

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

MINC, H. (1971). Rearrangements. *Trans. Amer. Math. Soc.* *159* 497–504.

SAVAGE, I.R. (1957). Contributions to the theory of rank order statistics—the "trend" case. *Ann. Math. Statist.* *28* 968–977.