SCHUR FUNCTIONS IN STATISTICS. III: 
A STOCHASTIC VERSION OF WEAK MAJORIZATION, 
WITH APPLICATIONS

BY

S. E. NEVIUS\textsuperscript{1}, F. PROSCHAN\textsuperscript{1}, and J. SETHURAMAN\textsuperscript{2}

FSU Statistics Report M307
AFOSR Technical Report No. 33
USARO-D Technical Report No. 8

September, 1974
The Florida State University
Department of Statistics
Tallahassee, Florida

\textsuperscript{1}Research sponsored by the Air Force Office of Scientific Research, AFSC, USAF, under Grant No. AFOSR-74-2581. The United States Government is authorized to reproduce and distribute reprints for governmental purposes notwithstanding any copyright notation hereon.

\textsuperscript{2}Research supported by the United States Army Research Office, Durham, under Grant No. DAHC04-74-G-0188.

AMS Classification Numbers: 26A51
26A86
60B05
62H99
Schur Functions in Statistics. III: A Stochastic Version of Weak Majorization, with Applications

Abstract

This is Part III of a series of papers on Schur functions in statistics. The main purpose of this part is to (a) develop the new concept of stochastic weak majorization, and (b) to develop techniques to utilize this concept in the field of stochastic inequalities. We introduce a stochastic version of weak majorization, develop its properties, and obtain multivariate applications. New results are obtained for the multinomial, multivariate negative binomial, and multivariate hypergeometric distributions.
Schur Functions in Statistics. III: A Stochastic Version of Weak Majorization, with Applications

by

S. E. Nevius, F. Proschan, and J. Sethuraman

1. Introduction and Summary.

In Part I, a basic theorem concerning the preservation of a Schur function under a certain integral transformation was derived. A stochastic version of majorization was introduced in Part II and multivariate applications of this new concept were developed. (Definitions from Parts I and II will generally not be repeated.) In Part III we introduce a stochastic version of weak majorization (see Section 2 for new definitions), develop its properties, and obtain multivariate applications.


The concept of weak majorization has been mentioned by several authors (Mitrinović, 1970; Marshall, Walkup, and Wets, 1967; Beckenbach and Bellman, 1965), but has not as yet been applied to statistical problems.

The deterministic definition of weak majorization will first be stated. Given a vector \( \mathbf{x} = (x_1, \ldots, x_n) \), let \( x_{[1]} \geq \cdots \geq x_{[n]} \) denote a decreasing rearrangement of \( x_1, \ldots, x_n \). \( \mathbf{x} \) is said to weakly majorize \( \mathbf{x}' \) if

\[
\sum_{i=1}^{j} x_{[i]} \geq \sum_{i=1}^{j} x'_{[i]}, \quad j = 1, \ldots, n;
\]

in symbols, \( \mathbf{x} \geq^m \mathbf{x}' \).

The following lemma will enable us to derive an important characterization of weak stochastic majorization.

**Lemma 2.1.** If \( \mathbf{x} \geq^m \mathbf{x}' \), then there exists a vector \( \mathbf{x}^* \) such that \( \mathbf{x} \geq^m \mathbf{x}^* \) and \( \mathbf{x}^* \geq \mathbf{x}' \).
Proof. Without loss of generality, assume \( x = (x_1, \ldots, x_n) \) and \( x' = (x'_1, \ldots, x'_n) \), where \( x_1 \geq x_2 \geq \cdots \geq x_n \), and \( x'_1 \geq \cdots \geq x'_n \). Let \( i_0 \) be the largest \( j, 1 \leq j \leq n \), such that \( \frac{1}{j} \sum_{i=1}^{j} x_i = \frac{1}{j} x'_i \); however, if

\[
\sum_{i=1}^{j} x_i > \sum_{i=1}^{j} x'_i
\]

for all \( j, 1 \leq j \leq n \), set \( i_0 = 0 \). If \( i_0 = n \), let \( x^* = x' \).

If \( i_0 < n \), let \( d_1 = \min \{ \frac{1}{j} x_i - \frac{1}{j} x'_i : j = i_0 + 1, \ldots, n \} \) and

\[
i_1 = \max \{ j, i_0 + 1 \leq j \leq n \}, \text{ such that } \frac{1}{j} x_i - \frac{1}{j} x'_i = d_1 \} \). Set

\[
x^{(1)}_{i_0 + 1} = x'_{i_0 + 1} + d_1, \quad x^{(1)}_i = x'_i \text{ for } i \neq i_0 + 1. \text{ If } i_1 = n, \text{ let } x^* = (x^{(1)}_1, \ldots, x^{(1)}_n).
\]

If \( i_1 < n \), let \( d_2 = \min \{ \frac{1}{j} x_i - \frac{1}{j} x^{(1)}_i : j = i_1 + 1, \ldots, n \} \) and

\[
i_2 = \max \{ j, i_1 + 1 \leq j \leq n \}, \text{ such that } \frac{1}{j} x_i - \frac{1}{j} x^{(1)}_i = d_2 \} \). Set

\[
x^{(2)}_{i_1 + 1} = x^{(1)}_{i_1 + 1} + d_2 \text{ and } x^{(2)}_i = x^{(1)}_i \text{ for } i \neq i_1 + 1. \text{ If } i_2 = n, \text{ let } x^* = (x^{(2)}_1, \ldots, x^{(2)}_n).
\]

Continuing in this manner, we get in a finite number, say \( m \), of steps such that \( i_m = n \) and \( x^* = (x^{(m)}_1, \ldots, x^{(m)}_n) \).

By construction, we have \( x^* \geq x' \) and \( x \succeq^m x^* \).

We may now prove the following characterization of weak majorization.

Theorem 2.2. \( x \succeq^m x' \) if and only if \( f(x) \geq f(x') \) for all nondecreasing Schur-convex functions.

Proof. Suppose \( f(x) \geq f(x') \) for all nondecreasing Schur-convex functions \( f \).

Consider the nondecreasing Schur-convex functions \( f_j(x) = \sum_{i=1}^{j} x[i] \), \( j = 1, \ldots, n \). By assumption, \( f_j(x) \geq f_j(x') \), \( j = 1, \ldots, n \). Thus
\( x \gg^m x' \) by definition.

Conversely, suppose \( x \gg^m x' \). By Lemma 2.1, there exists a vector \( x^* \) such that \( x \geq^m x^* \) and \( x^* \geq x' \). Let \( f \) be a nondecreasing Schur-convex function. Since \( f \) is Schur-convex and \( x \geq^m x^* \), we have by the definition of Schur-convexity that \( f(x) \geq f(x^*) \). Since \( f \) is nondecreasing and \( x^* \geq x' \), then \( f(x^*) \geq f(x') \). Thus \( f(x) \geq f(x') \).

This equivalent version of weak stochastic majorization leads us to the following stochastic version. Throughout, let \( X \) and \( X' \) have probability measures \( P \) and \( P' \), respectively, and distribution functions \( G \) and \( G' \), respectively.

Definition. (a) \( X \) is said to stochastically weakly majorize \( X' \) if

\[
f(X) \geq^* f(X')
\]

for every nondecreasing Schur-convex function \( f \); in symbols, \( X \gg^* X' \).

(b) We say \( P \gg^* P' \) if \( X \gg^* X' \).

(c) We say \( G \gg^* G' \) if \( X \gg^* X' \).

We need the following definition to state Theorem 2.3 which characterizes weak stochastic majorization.

Definition. A subset \( S \) of \( \mathbb{R}_n \) is said to be increasing if \( x \in S \) and \( x' \geq x \) imply \( x' \in S \).

Theorem 2.3. The following statements are equivalent.

(i) \( X \gg^* X' \).

(ii) \( Ef(X) \geq Ef(X') \) for every nondecreasing Schur-convex function \( f \) for which both these expectations exist.
(iii) \( \mathbf{E} f(X) \geq \mathbf{E} f(X') \) for every bounded nondecreasing Schur-convex function \( f \).

(iv) \( P(A) \geq P'(A) \) for all measurable increasing Schur-convex sets.

**Proof.** The proof is similar to the proof of Theorem 2.2 of Part II; the proof of the implication (ii) \( \Rightarrow \) (i) requires the additional observation that a nondecreasing function of a nondecreasing Schur-convex function is itself nondecreasing and Schur-convex, while the proof of (iv) \( \Rightarrow \) (ii) uses the fact that \( \{ x : f(x) \geq t \} \) is an **increasing** Schur-convex set for every nondecreasing Schur-convex function \( f \).

Next, we relate stochastic weak majorization to deterministic weak majorization in terms of a stochastic comparison of vectors. See Veinott (1965) for a characterization of the following stochastic ordering of vectors.

**Definition.** A random vector \( X \) is **stochastically larger** than a random vector \( X' \) if \( h(X) \geq^{st} h(X') \) for all nondecreasing functions \( h \); in symbols, \( X \succeq^{st} X' \).

An additional characterization of stochastic weak majorization may be stated in terms of the stochastic comparison of the vectors of partial sums of the reverse order statistics \( X[1] \geq \cdots \geq X[n] \). Define the map \( T \) from \( R^n \) into \( R^n \) by \( T(x) = (y_1, \ldots, y_n) \), where \( y_i = \sum_{j=1}^{i} x[j] \), \( i = 1, \ldots, n \). Thus, \( T \) gives the partial sums of the reverse order statistics of \( x_1, \ldots, x_n \). Let \( TR_n = C \). The following characterization of nondecreasing Schur-convex functions will prove useful in the characterization of stochastic weak majorization developed in Theorem 2.5 below.
Lemma 2.4. For any permutation-invariant function \( f \) on \( \mathbb{R}^n \), define the function \( g \) on \( C \) by putting \( g(y) = f(x) \) whenever \( y = Tx \). This defines a 1-1 correspondence, \( f \leftrightarrow g \), between permutation-invariant functions on \( \mathbb{R}^n \) and functions on \( C \). Moreover, \( f \) is a nondecreasing Schur-convex function if and only if \( g \) is nondecreasing.

Proof. The first statement is easily verified. The second statement is an immediate consequence of the fact that \( x \gtrsim^m x' \) if and only if \( y_i \geq y_i' \), \( i = 1, \ldots, n \), where \( y = Tx \) and \( y' = Tx' \).

Theorem 2.5. Let \( X \) and \( X' \) be random vectors in \( \mathbb{R}^n \). Set \( Y = TX \) and \( Y' = TX' \). Then \( X \gtrsim^{st,m} X' \) if and only if \( Y \gtrsim^{st} Y' \).

Proof. This theorem is an immediate consequence of Lemma 2.4.

As a direct consequence of Theorem 2.5, we obtain:

Corollary 2.6. If \( X \gtrsim^{st,m} X' \), then \( \frac{1}{j} \sum_{i=1}^{j} x_{[i]} \gtrsim^{st} \frac{1}{j} \sum_{i=1}^{j} x'_{[i]} \) for \( j = 1, \ldots, n \).

Note that Corollary 2.6 also follows directly from the definition of stochastic weak majorization since \( \frac{1}{j} \sum_{i=1}^{j} x_{[i]} \) is a nondecreasing Schur-convex function of \( x \) for \( j = 1, \ldots, n \). To see that the converse does not hold, consider the following example. Let \( X \) take on the values \((4,2)\) and \((3,1)\), each with probability \( \frac{1}{2} \), and let \( X' \) take on the values \((4,0)\) and \((3,2)\), each with probability \( \frac{1}{2} \). Then \( X_{[1]} = X'_{[1]} \) and \( X_{[1]} + X'_{[2]} = X'_{[1]} + X'_{[2]} \). Hence \( \frac{1}{j} \sum_{i=1}^{j} x_{[i]} \gtrsim^{st} \frac{1}{j} \sum_{i=1}^{j} x'_{[i]} \), \( j = 1, 2 \). Consider the non-decreasing Schur-convex function \( f(x) = x_1^3 + x_2^3 \). Note that
P(f(X) ≤ 28) > P(f(X') ≤ 28). Thus it is not true that $X \gg_{st.m.} X'$.
(A version of this example appears in Marshall and Olkin (in preparation).)

In the deterministic case, we know that $x \geq x'$ implies that $X \gg_{m} X'$. Theorem 2.7 gives a similar result for the stochastic case.

**Theorem 2.7.** If $X \geq_{st} X'$, then $X \gg_{st.m.} X'$.

**Proof.** Every nondecreasing Schur-convex function of $X$ is itself a non-decreasing function of $X$. Hence the result holds. ||

To see that the converse does not hold even in the deterministic case, let $X = (5,3,3)$ with probability one and $X' = (4,4,2)$ with probability one. Then $X \gg_{st.m.} X'$, but $X$ is not stochastically larger than $X'$.

The following statement is equivalent to the definition of majorization in the deterministic case.

**Equivalent Definition.** A vector $X$ weakly majorizes a vector $X'$ if $X \gg_{m} z$ for every $z$ such that $X' \gg_{m} z$.

We have the following stochastic analogue of this definition.

**Theorem 2.8.** If $X \gg_{st.m.} X'$, then $P(X \gg_{m} z) > P(X' \gg_{m} z)$ for each vector $z$.

**Proof.** The result follows by choosing the bounded nondecreasing Schur-convex function $g(x) = I_{[y \geq_{m} x]}$ and applying Theorem 2.3. ||

Intuitively, Theorem 2.8 states that if $X \gg_{st.m.} X'$, then $X$ is more likely to weakly majorize any vector $z$ than is $X'$. The converse,
however, is not true. To see this, consider the example following Corollary 2.8 of Part II, where $X$ takes on the values $(1, 0, 0), (\frac{1}{2}, \frac{1}{2}, 0), (\frac{2}{3}, \frac{1}{3}, \frac{1}{6})$, and $(\frac{1}{2}, \frac{1}{3}, \frac{1}{6})$ with probabilities .1, .1, .1, and .7, respectively, and $X'$ takes on these values with probabilities 0, .2, .2, and .6, respectively. The fact that $P[X \gg^m z] \geq P[X' \gg^m z]$ for all $z$ follows from the easily verified fact that a vector $z$ is weakly majorized by both $(\frac{1}{2}, \frac{1}{2}, 0)$ and $(\frac{2}{3}, \frac{1}{3}, \frac{1}{6})$ only if it is weakly majorized by $(\frac{1}{2}, \frac{1}{3}, \frac{1}{6})$. Since $f(x)$ is not stochastically larger than $f(x')$ for the nondecreasing Schur-convex function $f(x) = x_1^3 + x_2^2 + x_3^3$, it follows that $X$ does not stochastically weakly majorize $X'$.


The applicability of our new notion of stochastic weak majorization is greatly expanded by the degree to which it is preserved under various standard mathematical, probabilistic, and statistical operations. In this section we display operations which preserve stochastic weak majorization.

First, we show that stochastic weak majorization is preserved under mixtures of distributions.

**Theorem 3.1.** Let $X$ and $X'$ be two random vectors and let $U$ be a random variable such that the conditional distribution of $X$ given $U = u$ stochastically weakly majorizes the conditional distribution of $X'$ given $U = u$ for each $u$. Then $X \gg^* X'$.

The proof of Theorem 3.1 is essentially the same as the proof of the analogous theorem for stochastic majorization in Part II.
Another preservation result involving sums of independent random vectors may be shown utilizing the following two lemmas. Corresponding to a random vector $X$, denote by $X^*$ the vector of reverse order statistics $(X_{[1]}, X_{[2]}, \ldots, X_{[n]})$.

**Lemma 3.2.** $X \gg_{st.m.} X'$ if and only if $X^* \gg_{st.m.} X'^*$.

The proof utilizes the fact that a Schur-convex function is necessarily permutation-invariant. The details are omitted.

**Lemma 3.3.** If $Z \geq_{st.} Z'$, $W \geq_{st.} X'$, $Z$ and $W$ are independent, and $Z'$ and $W'$ are independent, then $Z + W \geq_{st.} Z' + W'$.

We may now state the following preservation result for sums.

**Theorem 3.4.** If $X \gg_{st.m.} X'$, $Y \geq_{st.m.} Y'$, $X$ and $Y$ are independent, and $X'$ and $Y'$ are independent, then $X^* + Y^* \geq_{st.m.} X'^* + Y'^*$.

**Proof.** By Theorem 2.6, it suffices to show $(X_{[1]} + Y_{[1]}', X_{[1]} + X_{[2]} + Y_{[1]} + Y_{[2]}'), \ldots, \sum_{i=1}^{n} X_{[i]} + \sum_{i=1}^{n} Y_{[i]} \geq_{st.} (X'_{[1]} + Y'_{[1]}', X'_{[1]} + X'_{[2]} + Y'_{[1]} + Y'_{[2]}'), \ldots,

\[ \sum_{i=1}^{n} X'_{[i]} + \sum_{i=1}^{n} Y'_{[i]} \].

But this follows directly from Lemma 3.3.

To see that the result does not hold even in the deterministic case without rearranging vectors before adding, let $X = (6, 2)$, $Y = (3, 6)$, $X' = Y' = (1, 5)$ with probability one. Then $X \gg_{st.m.} X'$, $Y \geq_{st.m.} Y'$, but $X + Y$ does not stochastically weakly majorize $X' + Y'$. 
Next, we show that stochastic weak majorization is also preserved under multiplication by an increasing function of \( \sum_{i=1}^{n} X_i \).

**Theorem 3.5.** Let \( X \gg_{st,m} X' \) and \( f \) be a nonnegative nondecreasing Borel-measurable function. Then \( f(\sum_{i=1}^{n} X_i)X \gg_{st,m} f(\sum_{i=1}^{n} X'_i)X' \).

**Proof.** Let \( g \) be a nondecreasing Schur-convex function. Then \( g(f(\sum_{i=1}^{n} x_i)x) \) is a nondecreasing Schur-convex function of \( x \). The desired result follows. |

In proving the next theorem, which gives the most important and useful operation preserving stochastic majorization, we require the following lemmas. The first lemma, originally proven by Esary and Proschan (unpublished), is presented in Fledger and Proschan (1973).

**Lemma 3.6.** Let \( X_1, \ldots, X_k \) be mutually independent random vectors, and similarly \( X'_1, \ldots, X'_k \) be mutually independent. Suppose \( X_i \geq_{st} X'_i \) for \( 1 \leq i \leq k \). Then \( (X_1, \ldots, X_k) \geq_{st} (X'_1, \ldots, X'_k) \).

We also require the following deterministic lemma.

**Lemma 3.7 (Deterministic).** \( f \) is a nondecreasing Schur-convex function if and only if

\[
(3.1) \quad x \geq_{m} x' \Rightarrow f(x) \geq f(x').
\]

**Proof.** Let \( f \) be a nondecreasing Schur-convex function and \( x \geq_{m} x' \). By Lemma 2.1, \( f(x) \geq f(x') \).

Conversely, suppose (3.1) holds. Since \( x \geq_{m} x' \) implies \( x \geq_{m} x' \), it follows that \( x \geq_{m} x' \Rightarrow f(x) \geq f(x') \). Thus \( f \) is Schur-convex by
definition. Since $x \succeq x'$ implies $x \succeq^m x'$, we have $f(x) \succeq f(x')$.

Thus $f$ is nondecreasing. ||

We may now prove the following preservation theorem.

**Theorem 3.8.** Let $\phi(\lambda, x)$ be TP$_2$ in $\lambda > 0$ and $x \geq 0$, and satisfy the semigroup property. (See Part I for definitions.) Let $X_\lambda = (X_{\lambda_1}, \ldots, X_{\lambda_n})$ be a random vector of independent components, where $X_{\lambda_i}$ is a nonnegative random variable with density $\phi(\lambda_i, x)$, $i = 1, \ldots, n$. If $\lambda \succeq^m \lambda'$, then $X_\lambda \succeq^m X_{\lambda'}$.

**Proof.** Let $g$ be a nondecreasing bounded Schur-convex function. Define $h(\lambda) = E_g(X_{\lambda_1}, \ldots, X_{\lambda_n})$. We want to show $\lambda \succeq^m \lambda'$ implies $h(\lambda) \geq h(\lambda')$.

Thus, by Lemma 3.7 it suffices to show that $h$ is a nondecreasing Schur-convex function of $\lambda$. By Theorem 3.3 of Part II, $h$ is Schur-convex.

To show that $h$ is nondecreasing, let $\lambda \geq \lambda'$. Then for each set of corresponding components $X_{\lambda_i}$ and $X_{\lambda'_i}$ of $X_\lambda$ and $X_{\lambda'}$, respectively, we have $X_{\lambda_i} \succeq^s X_{\lambda'_i}$. Since $g$ is nondecreasing, we have by Lemma 3.6 that $g(X_{\lambda_i}) \succeq^s g(X_{\lambda'_i})$. Thus $h(\lambda) \geq h(\lambda')$. ||

Thus a deterministic property (weak majorization) of the parameter vector $\lambda$ is transformed into a corresponding stochastic property (stochastic weak majorization) of the random vector $X_\lambda$. This result will be exploited in Section 4 to obtain useful multivariate applications of stochastic weak majorization.

An immediate corollary of Theorem 3.8 is an application to stochastic processes.
Corollary 3.9. Let \((X(t), 0 \leq t < \infty)\) be a stochastic process with stationary, independent, and nonnegative increments. Let the density \(\phi(\lambda, x)\) of \(X(t + \lambda) - X(t)\) be \(TP_2\) in \(\lambda > 0\) and \(x \geq 0\). Let \(0 = t_0 < t_1 < \cdots < t_n\), \(0 = t'_0 < t'_1 < \cdots < t'_n\), \(\lambda_i = t_i - t_{i-1}\), and \(\lambda'_i = t'_i - t'_{i-1}\). Then \(\lambda \gg \lambda'\) implies that

\[
(X(t_1) - X(t_0), \ldots, X(t_n) - X(t_{n-1})) \gg_{st.m.} \quad (X(t'_1) - X(t'_0), \ldots, X(t'_n) - X(t'_{n-1})).
\]

The proof is analogous to the proof of Corollary 3.4 in Part II.

Theorem 3.8 gives sufficient conditions on the distribution of a random vector \(X_\lambda\) to state that \(\lambda \gg \lambda'\) implies that \(X_\lambda \gg_{st.m.} X_{\lambda'}\). We now consider prior distributions on the parameter vector \(\lambda\) with probability distributions \(G\) and \(G'\), where \(G \gg_{st.m.} G'\). The next theorem shows that stochastic weak majorization is preserved under a mixing operation.

Theorem 3.10. Let \(\{X_\lambda\}\) be a family of random vectors indexed by \(\lambda \in \Theta \subset \mathbb{R}_n\) such that \(\lambda \gg \lambda'\) implies \(X_\lambda \gg_{st.m.} X_{\lambda'}\). Let \(G\) and \(G'\) be two probability measures on \(\Theta\) such that \(G \gg_{st.m.} G'\). Let \(Q(\lambda) = \int_{\Theta} P(X_\Lambda \in A) dG(\lambda)\) and \(Q'(\lambda) = \int_{\Theta} P(X_\Lambda \in A) dG'(\lambda)\) for all Borel sets \(A\) in \(\mathbb{R}_n\). Then \(Q \gg_{st.m.} Q'\).

Proof. Let \(f\) be a bounded nondecreasing Schur-convex function. Let \(X\) and \(X'\) have probability measures \(Q\) and \(Q'\), respectively. Then

\[
Ef(X) = \int_{\Theta} Ef(X_\lambda) dG(\lambda).
\]

By assumption and Theorem 2.3, \(\lambda \gg \lambda'\) implies \(Ef(X_\lambda) \geq Ef(X_{\lambda'}\). By Lemma 3.7, \(Ef(X_\Lambda)\) is a nondecreasing Schur-convex function of \(\lambda\). Since \(G_1 \gg_{st.m.} G_2\), it follows by Theorem 2.3 that \(\int_{\Theta} Ef(X_\lambda) dG_1(\lambda) \geq \int_{\Theta} Ef(X_\lambda) dG_2(\lambda)\), that is, \(Ef(X) \geq Ef(X')\). Thus \(X \gg_{st.m.} X'\).
A type of preservation theorem under conditioning may be proven for certain types of discrete random vectors. We will use the following definitions from Barlow and Proschan (1974).

**Definition.** A random variable $T$ is **stochastically increasing** in a random variable $S$ if $P(T > t \mid S = s)$ is increasing in $s$ for all $t$. We write $T \triangleright \triangleright \triangleright_{st} S$.

**Definition.** A random vector $T$ is **stochastically increasing** in $S$ if $f(T) \triangleright \triangleright \triangleright_{st} S$ for all nondecreasing functions $f(t)$. We write $T \triangleright \triangleright \triangleright_{st} S$.

Let $\frac{X}{N}$ have the conditional probability measure of the random vector $X$ given that $\sum_{i=1}^{n} X_i = N$. The following preservation theorem may now be stated.

**Theorem 3.11.** Let $X$ and $X'$ be discrete random vectors and $X \triangleright \triangleright \triangleright_{st} X'$ in $\sum_{i=1}^{n} X_i$. Let $P(\sum_{i=1}^{n} X_i = N) > 0$ and $P(\sum_{i=1}^{n} X_i = N') > 0$. If $X \geq \text{st.m.} X'$ and $N \geq N'$, then $\frac{X}{N} \geq \text{st.m.} \frac{X'}{N'}$.

**Proof.** Since $X \triangleright \triangleright \triangleright \sum_{i=1}^{n} X_i$, we have $\frac{X}{N} \geq \text{st.m.} \frac{X}{N}$, and thus by Theorem 2.7, $\frac{X}{N} \geq \text{st.m.} \frac{X'}{N'}$. Since $X \geq \text{st.m.} X'$, we have by Corollary 2.10 of Part II that $\frac{X'}{N} \geq \text{st.m.} \frac{X'}{N'}$, and hence $\frac{X}{N} \geq \text{st.m.} \frac{X'}{N'}$. Thus $\frac{X}{N} \geq \text{st.m.} \frac{X'}{N'}$.

Next, we give a sufficient condition for $X \triangleright \triangleright \triangleright \sum_{i=1}^{n} X_i$ which will provide applications of Theorem 3.11 in the next section.
Theorem 3.12. Let $X$ be a random vector such that $X_{N+1}$ has the same distribution as $X_N + \Delta(X_N)$, where $\Delta(X_N)$ is a nonnegative random vector whose distribution may depend on $X_N$. Then $X + st \in \sum_{i=1}^{n} X_i$.

Proof. Note that $X_N + \Delta(X_N) \geq X_N$ a.s. Thus $X_{N+1} \geq st X_N$. I.e., $X + st \in \sum_{i=1}^{n} X_i$.

Let $X_N$ represent a random vector whose conditional distribution given $N = N$ is that of $X_N$ as defined above. In other words, the distribution of $N$ is a mixing distribution for the parameter $N$. We may now state the following theorem.

Theorem 3.13. Let $X$ and $X'$ be discrete random vectors with $X + st \in \sum_{i=1}^{n} X_i$. If $X \geq st.m. X'$ and $N \geq st N'$, then $X_N \geq st.m. X'_N$.

Proof. Let $f$ be a nondecreasing Schur-convex function. Then

$$Ef(X_N) = EE[f(X_N)|N]$$

$$\geq EE[f(X'_N)|N] \quad \text{[by Theorem 2.10 of Part II]}$$

$$\geq EE[f(X'_N)|N'] \quad \text{[by Theorem 3.11 and the hypothesis that } N \geq st N'].$$

We may combine Theorems 3.10 and 3.13 to obtain the following result.

Theorem 3.14. Let $\{X_{\lambda N}\}$ be a family of discrete random vectors indexed by $\lambda \in \theta \subset R_n$ such that $\lambda \geq m \lambda'$ implies $X_{\lambda} \geq st.m. X_{\lambda'}$, and $X_{\lambda} + st \in \sum_{i=1}^{n} X_{i,\lambda}$. Let $X_{\lambda N}$ denote a random vector with the conditional distribution
of $\mathbf{X}_\mathbf{\Lambda}$ given that $\sum_{i=1}^{n} X_{i,\mathbf{\Lambda}} = N$, and $\mathbf{X}_{\mathbf{\Lambda},\mathbf{N}}$ denote the resultant mixture as in Theorem 3.13. Finally, let $\mathbf{X}_{\mathbf{\Lambda},\mathbf{N}}$ have probability measure $\mu(A) = \int P(X_{\mathbf{\Lambda},\mathbf{N}} \in A) d\mathbf{\Lambda}$ for all Borel sets $A$ in $\mathbb{R}^{n}$, where $\mathbf{\Lambda}$ is a random vector taking values in $\Theta$ a.s. and with distribution $\mathbf{\Lambda}$. 

If $N \succeq_{st} N'$ and $\mathbf{\Lambda} \succeq_{st.m.} \mathbf{\Lambda}'$, then $\mathbf{X}_{\mathbf{\Lambda},\mathbf{N}} \succeq_{st.m.} \mathbf{X}_{\mathbf{\Lambda}',\mathbf{N}'}$.

**Proof.** Let $g$ be a bounded nondecreasing Schur-convex function. Then

$$
\text{E}g(\mathbf{X}_{\mathbf{\Lambda},\mathbf{N}}) = \text{E}E[g(\mathbf{X}_{\mathbf{\Lambda},\mathbf{N}})|\mathbf{N}]
$$

$$
\geq \text{E}E[g(\mathbf{X}_{\mathbf{\Lambda}',\mathbf{N}'}|\mathbf{\Lambda}')|\mathbf{N}] \quad \text{[by Theorem 3.10]}
$$

$$
= \text{E}g(\mathbf{X}_{\mathbf{\Lambda}',\mathbf{N}})
$$

$$
= \text{E}E[g(\mathbf{X}_{\mathbf{\Lambda}',\mathbf{N}'})|\mathbf{\Lambda}']
$$

$$
\geq \text{E}E[g(\mathbf{X}_{\mathbf{\Lambda}',\mathbf{N}'})|\mathbf{\Lambda}'] \quad \text{[by Theorem 3.13]}
$$

$$
= \text{E}g(\mathbf{X}_{\mathbf{\Lambda}',\mathbf{N}'})
$$

By Theorem 2.3, $\mathbf{X}_{\mathbf{\Lambda},\mathbf{N}} \succeq_{st.m.} X_{\mathbf{\Lambda}',\mathbf{N}'}$.

4. **Applications of Stochastic Weak Majorization.**

In this section we present applications of stochastic weak majorization to obtain general inequalities for several well known multivariate distributions.
Application 4.1. Let \( X_{\lambda} \) be a random vector of independent components, where \( X_{\lambda_i} \) has a density of the form given in (a), (b), or (c) below. Let \( \lambda \gg m \lambda' \). Then \( X_{\lambda} \gg_{st.m.} X_{\lambda'} \).

(a) Poisson. \( f_{\lambda}(x) = \frac{(\lambda \theta)^x}{x!} e^{-\lambda \theta} \), \( x = 0, 1, \ldots, \lambda > 0 \), and fixed \( \theta > 0 \).

(b) Binomial. \( f_{\lambda}(x) = \binom{\lambda}{x} p^x (1 - p)^{\lambda - x} \), \( x = 0, 1, \ldots, n; n = 1, 2, \ldots \); and \( 0 < p < 1 \).

(c) Gamma. \( f_{\lambda}(x) = \frac{\theta^\lambda x^{\lambda - 1}}{\Gamma(\lambda)} e^{-\theta x} \), \( x \geq 0, \lambda > 0 \), and fixed \( \theta > 0 \).

The result follows from Theorem 3.8 by noting that \( \phi(\lambda, x) = f_{\lambda}(x) \) is TP\(_2\) and satisfies the semi-group property in each of the cases (a), (b), and (c).

Application 4.1 is next used to obtain stochastic comparisons for the multivariate negative binomial distribution by applying Theorem 3.1.

Application 4.2. Let \( Y_{\lambda} = (Y_{1,\lambda}, \ldots, Y_{n,\lambda}) \) have a multivariate negative binomial distribution with density

\[
 f_{\lambda}(y) = \frac{\Gamma(N + \sum_{i=1}^{n} y_i)}{\Gamma(N)} \frac{\lambda_1^{y_1}}{y_1!} \left(1 + \sum_{i=1}^{n} \lambda_i\right)^{-N} \frac{\lambda_i}{y_i!} \lambda_1^{y_1} (1 + \sum_{i=1}^{n} \lambda_i)^{-N - \sum_{i=1}^{n} y_i},
\]

where \( y_i = 0, 1, \ldots, \lambda_i > 0, i = 1, \ldots, n, N > 0 \). If \( \lambda \gg m \lambda' \), then \( Y_{\lambda} \gg_{st.m.} Y_{\lambda'} \).

Proof. The result follows from Application 4.1(a), Theorem 3.1, and the fact utilized in Part II that the multivariate negative binomial distribution is a mixture of independent Poisson random variables with densities 4.1(a) under a gamma distribution for \( \theta \) with density \( g(\theta) = \frac{\theta^{N-1}}{\Gamma(N)} e^{-\theta} \).*
Using the fact that \( g(x) = -I_{\max(x_1, \ldots, x_n) \leq r} (x) \) is a nondecreasing Schur-convex function, we obtain the following application.

**Application 4.3.** Let \( X_{\lambda} \) have any one of the multivariate distributions given in Applications 4.1 or 4.2. If \( \lambda \gg_m \lambda' \), then

\[
P(Y_{1,\lambda} \leq r, \ldots, Y_{n,\lambda} \leq r) \leq P(Y_{1,\lambda'} \leq r, \ldots, Y_{n,\lambda'} \leq r).
\]

Application 4.3 generalizes a similar majorization theorem for the multinomial distribution proven by Olkin (1972).

Applications of stochastic weak majorization to the multinomial and multivariate hypergeometric distributions with random sample size are possible using the preservation results of Section 3. First, we present the following result.

**Theorem 4.4.** Let \( X_{N,\mathbf{p}} \) be a multinomial random vector with parameters \( N \) and \( \mathbf{p} \). Let \( \mathbf{p} \geq_m \mathbf{p}' \) and \( N \geq N' \). Then \( X_{N,\mathbf{p}} \gg_{ST} X_{N',\mathbf{p}'} \).

**Proof.** Recall that the multinomial may be derived by conditioning a vector of independent Poisson random variables \( X \) on \( \sum_{i=1}^{n} X_i \). By Theorem 3.12 and Application 4.1 of Part II, the hypotheses of Theorem 3.11 are satisfied. Thus we may apply Theorem 3.11 to conclude \( X_{N,\mathbf{p}} \gg_{ST} X_{N',\mathbf{p}'} \).

An analogous theorem holds for the multivariate hypergeometric distribution.

**Theorem 4.5.** Let \( X_{\lambda} \) be a multivariate hypergeometric random vector with density
\[ f(y) = \frac{\sum_{i=1}^{n} \binom{\lambda_i}{y_i}}{\sum_{i=1}^{n} \lambda_i} , \]

where \( y_1, \ldots, y_n \) are integers satisfying \( 0 \leq y_i \leq \lambda_i, \ i = 1, \ldots, n, \)

\[ \sum_{i=1}^{n} y_i = N; \lambda_1, \ldots, \lambda_n \text{ are positive integers; and integer } N \text{ satisfies } 0 < N \leq \sum_{i=1}^{n} \lambda_i. \]

If \( \lambda \geq_m \lambda' \) and \( N \geq N' \), then \( X_{N, \lambda} \geq_{st.m.} X_{N', \lambda'} \).

\[ \text{Proof. Recall that the multivariate hypergeometric distribution may be derived by conditioning a vector of independent binomial random vectors } X \text{ on } \sum_{i=1}^{n} X_i. \]

By Theorem 3.12 and Application 4.1 of Part II, the hypotheses of Theorem 3.11 are satisfied. Thus we may apply Theorem 3.11 as in the proof of Theorem 4.4 to obtain the desired result. ||

Similar results for the multinomial and multivariate hypergeometric with random sample size are possible by utilizing Theorem 3.13. The following theorem is for the multinomial; an analogous result holds for the multivariate hypergeometric.

**Theorem 4.6.** Let \( N \) be a random variable taking values in \( \{1, 2, \ldots\} \) and let the conditional distribution of \( X_{\frac{N}{2}, \lambda} \) given \( N = r \) be multinomial with parameters \( r \) and \( \pi \). If \( N \geq_{st} N' \) and \( \pi \geq_{m} \pi' \), then \( X_{\frac{N}{2}, \lambda} \geq_{st.m.} X_{\frac{N'}{2}, \lambda'} \).

\[ \text{Proof. The result follows directly from Theorem 3.13 by similar arguments as in the proof of Theorem 4.4. ||} \]
We may also apply Theorem 3.14 to get a weak stochastic majorization result for the multinomial with both $N$ and $p$ random. An analogous result holds for the multivariate hypergeometric.

**Theorem 4.7.** Let $X_{\mathbf{p},N}$ be as defined in Theorem 4.6. Let $\mathbf{\Lambda}$ and $\mathbf{\Lambda}'$ be random vectors with distributions $G$ and $G'$, respectively, and which take values in $\{\mathbf{p}: \ p_i > 0, \ i = 1, \ldots, n, \ \sum_{i=1}^{n} p_i = 1\}$ a.s. Let $X_{\mathbf{\Lambda},N}$ have probability measure

$$\mu(A) = \int P(X_{\mathbf{p},N} \in A) dG(\mathbf{p})$$

for all Borel sets $A$ in $\mathbb{R}_n^n$; similarly for $X_{\mathbf{\Lambda}',N'}$. If $N \geq_{\text{st}} N'$ and $\mathbf{\Lambda} \geq_{\text{st.m.}} \mathbf{\Lambda}'$, then $X_{\mathbf{\Lambda},N} \geq_{\text{st.m.}} X_{\mathbf{\Lambda}',N'}$.

**Proof.** The result follows directly from Theorem 3.14. ||
REFERENCES


**Schur Functions in Statistics. III. A Stochastic Version of Weak Majorization with Applications**

**AUTHOR(s)**
- S. Edward Nevius
- Frank Proschan
- Jayaram Sethuraman

**PERFORMING ORGANIZATION NAME & ADDRESS**
The Florida State University
Department of Statistics
Tallahassee, Florida 32306

**CONTROLLING OFFICE NAME & ADDRESS**
- United States Air Force
  - Air Force Office of Scientific Research
  - 1400 Wilson Boulevard
  - Arlington, Virginia 22209
- U.S. Army Research Office
  - Box CM, Duke Station
  - Durham, North Carolina 27706

**REPORT DATE**
September, 1974

**NUMBER OF PAGES**
20

**DISTRIBUTION STATEMENT**
Releasable without limitation on dissemination.

**SUPPLEMENTARY NOTES**

**KEY WORDS**
- Majorization
- Weak majorization
- Schur-convex
- Schur function
- Transformation
- Stochastic majorization
- Stochastic weak majorization
- Multivariate distributions
- Stochastic inequalities
20. ABSTRACT

This is Part III of a series of papers on Schur functions in statistics. The main purpose of this part is to (a) develop the new concept of stochastic weak majorization, and (b) to develop techniques to utilize this concept in the field of stochastic inequalities. We introduce a stochastic version of weak majorization, develop its properties, and obtain multivariate applications. New results are obtained for the multinomial, multivariate negative binomial, and multivariate hypergeometric distributions.