JOURNAL OF MULTIVARIATE ANALYSIS 10, 467–498 (1980)

Classes of Orderings of Measures and Related Correlation Inequalities. I. Multivariate Totally Positive Distributions*

SAMUEL KARLIN

Department of Mathematics, Stanford University, Stanford, California 94305

AND

YOSEF RINOTT

Department of Mathematics, Stanford University, Stanford, California 94305; and Department of Statistics, The Hebrew University, Jerusalem, Israel

Communicated by P. R. Krishnaiah

A function $f(\mathbf{x})$ defined on $\mathscr{X} = \mathscr{X}_1 \times \mathscr{X}_2 \times \cdots \times \mathscr{X}_n$ where each $\mathscr{X}_i$ is totally ordered satisfying $f(\mathbf{x} \vee \mathbf{y}) f(\mathbf{x} \wedge \mathbf{y}) \geqslant f(\mathbf{x}) f(\mathbf{y})$, where the lattice operations $\vee$ and $\wedge$ refer to the usual ordering on $\mathscr{X}$, is said to be multivariate totally positive of order 2 (MTP_2). A random vector $\mathbf{Z} = (Z_1, Z_2, \ldots, Z_n)$ of n-real components is MTP_2 if its density is MTP_2. Classes of examples include independent random variables, absolute value multinormal whose covariance matrix Σ satisfies $-D\Sigma^{-1}D$ with nonnegative off-diagonal elements for some diagonal matrix D, characteristic roots of random Wishart matrices, multivariate logistic, gamma and F distributions, and others. Composition and marginal operations preserve the MTP_2 properties. The MTP_2 property facilitate the characterization of bounds for confidence sets, the calculation of coverage probabilities, securing estimates of multivariate ranking, in establishing a hierarchy of correlation inequalities, and in studying monotone Markov processes. Extensions on the theory of MTP_2 kernels are presented and amplified by a wide variety of applications.

1. INTRODUCTION AND OVERVIEW

Many classical and constantly arising inequalities in mathematical analysis, probability, and statistics merely express the property that a given function or signed measure belongs to a certain convex cone or its dual.

Received October 29, 1979; revised May 12, 1980

* Supported in part by NIH Grant 2R01 GM10452-16 and NSF Grant MCS76-80624-A02.

0047-259X/80/040467–32$02.00/0

More specifically, two finite measures σ_1 and σ_2 defined on a set Δ in Euclidean n-space (R^n) fulfilling the relations

$$\int_\Delta f\, d\sigma_1 \leqslant \int_\Delta f\, d\sigma_2 \tag{1.1}$$

for all bounded functions f belonging to some prescribed convex cone $\mathscr{F}$ asserts that the signed measure $\sigma_2 - \sigma_1$ belong to the "dual cone" of $\mathscr{F}$. We write (1.1) symbolically as

$$\sigma_2 \succ^{\mathscr{F}} \sigma_1. \tag{1.2}$$

When $\mathscr{F}$ comprises the collection of monotone functions on R^n then the common terminology is that σ_2 is *stochastically larger than* σ_1. When $\mathscr{F}$ consists of all convex functions then the ordering (1.1) is often referred to as σ_2 *dilates* σ_1. Many extensions and abstract versions of these concepts serve in studies of Banach spaces and for the exposition of Choquet theory (e.g., see Phelps [41], Alfsen [2]). Important contributions directed to the objective of classifying and characterizing the ordering relation of (1.2) include works of Hardy, Littlewood, and Pólya [16], Strassen [52], Meyer [39, part 3], and Preston [43], among others.

Dilation of measures in its one-dimensional version is synonymous to the notion of "majorization." For a concrete setting of "majorization," including a wealth of applications and references, see Marshall and Olkin [38]. Dilation and comparisons of distributions for symmetric sampling schemes from finite populations appear in Kemperman [31] and Karlin [25].

Measures of association and related probabilistic inequalities provide further notions and methodologies for comparing measures. These ideas and techniques play an important role in many diverse areas including reliability systems (e.g., Esary, Proschan, and Walkup [12], Barlow and Proschan [4]), in problems of multivariate hypotheses testing and monotonicity criteria of power functions (e.g., Perlman and Olkin [40]), in securing bounds for confidence sets and coverage probabilities (e.g., Das Gupta *et al.* [6], Sidak [50], Jogdeo [19], Dykstra [8]), slippage problems (Karlin and Truax [30]) and ranking and selection procedures (e.g., Rinott and Santner [47]). These concepts also relate to multivariate Schur functions (e.g., Rinott [46]), log concave functions (e.g., Prékopa [42], Borell [5]), and multivariate totally positive kernels.

Recent years have witnessed many new developments and discoveries of classes of stochastic monotone processes $\{X_t, t \geqslant 0, t \text{ discrete or continuous}\}$ where the distribution measure P_t of X_t for appropriate initial conditions satisfies

$$P_t \prec^{\mathscr{F}} P_s \quad \text{for} \quad t < s \tag{1.3}$$

especially where $\mathscr{F}$ is the collection of monotone functions on R^n, e.g., see Harris [17], Kirstein, Franken, and Stoyan [34], Kamae, Krengel, and O'Brein [22], and Kamae and Krengel [21]. These authors clarified and initiated a series of results concerning representations and convergence properties on the realizations of these "monotone" processes. However, the abstract characterizations of the orderings (1.2) and (1.3) are difficult to deal with in concrete statistical contexts. For these purposes the condition of *multivariate total positivity* is germane. A *nonnegative* kernel $f(x, y)$ of two real variables defined on $\mathscr{X} \times \mathscr{Y}$ ($\mathscr{X}$ and $\mathscr{Y}$ each totally ordered) is said to be *totally positive of order* 2, abbreviated TP_2, if the determinant of the second-order matrix $\|f(x_i, y_j)\|_{i,j=1}^2$ is nonnegative for all choices $x_1 < x_2$, $y_1 < y_2$ (Karlin [24]).

A natural approach in defining multivariate total positivity in terms of orderings on lattices is the following: Consider a kernel $f(\mathbf{x})$ defined on $\mathscr{X} = \mathscr{X}_1 \times \mathscr{X}_2 \times \cdots \times \mathscr{X}_n$ where each $\mathscr{X}_i$ is totally ordered satisfying

$$f(\mathbf{x} \vee \mathbf{y})\, f(\mathbf{x} \wedge \mathbf{y}) \geqslant f(\mathbf{x})\, f(\mathbf{y}), \tag{1.4}$$

where $\vee$ and $\wedge$ are the corresponding lattice operations, i.e., for $\mathbf{x} = (x_1, x_2, \ldots, x_n)$, $\mathbf{y} = (y_1, y_2, \ldots, y_n)$

$$\mathbf{x} \vee \mathbf{y} = (\max(x_1, y_1), \max(x_2, y_2), \ldots, \max(x_n, y_n))$$

and (1.5)

$$\mathbf{x} \wedge \mathbf{y} = (\min(x_1, y_1), \min(x_2, y_2), \ldots, \min(x_n, y_n)).$$

A kernel with the property (1.4) is called *multivariate totally positive of order* 2 (MTP_2).

It is convenient to introduce the designation that

$$\begin{gathered}\text{A random vector } \mathbf{X} = (X_1, \ldots, X_n) \text{ of } n\text{-components} \\ \text{is } \mathrm{MTP}_2 \text{ if its density is } \mathrm{MTP}_2.\end{gathered} \tag{1.6}$$

In order to check (1.4) it suffices to show that $f(x_1, x_2, \ldots, x_n) > 0$ is TP_2 in every pair of variables where the remaining variables are kept fixed, (Lorentz [37], Rinott [46], Kemperman [32]).

A density $f(\mathbf{x})$ with the inequality direction reversed in (1.4) is called *multivariate reverse rule of order* 2 (MRR_2). The companion paper II [27] elaborates various correlation type inequalities for MRR_2 densities featuring a number of prominent examples and applications.

When $f(\mathbf{x})$ is an MTP_2 density on R^n with respect to a product measure

$d\sigma(\mathbf{x})$ (which we write for convenience $d\mathbf{x}$), then the multivariate Tchebycheff rearrangement inequality holds, namely,

$$\int f(\mathbf{x})\,\varphi(\mathbf{x})\,\psi(\mathbf{x})\,d\mathbf{x} \geqslant \left(\int f(\mathbf{x})\,\varphi(\mathbf{x})\,d\mathbf{x}\right)\left(\int f(\mathbf{x})\,\psi(\mathbf{x})\,d\mathbf{x}\right) \tag{1.7}$$

provided ψ and φ are simultaneously monotone increasing or decreasing (Sarkar [48]).

In random variable format (1.7) can be compactly written in the form

$$E[\varphi(\mathbf{X})\,\psi(\mathbf{X})] \geqslant (E[\varphi(\mathbf{X})])(E[\psi(\mathbf{X})]) \qquad (E \text{ is the expectation operator}), \tag{1.8}$$

that is,

$$\operatorname{Cov}(\varphi(\mathbf{X}), \psi(\mathbf{X})) \geqslant 0. \tag{1.9}$$

Thus, an MTP_2 random vector $\mathbf{X}$ induces an *associated set of random variables* in the sense of Esary, Proschan, and Walkup [12], signifying that (1.8) holds. We can express (1.9) equivalently as follows. Let A be a monotone set in R^n i.e., if $\mathbf{x} \in A$ and $\mathbf{x}' \geqslant \mathbf{x}$, then $\mathbf{x}' \in A$. Now for A and B *monotone*

$$P\{\mathbf{X} \in A \cap B\} \geqslant \Pr\{\mathbf{X} \in A\}\, P\{\mathbf{X} \in B\}. \tag{1.10}$$

The span of MTP_2 densities incorporate the following classical cases:

(i) The negative multinomial discrete density is characterized by the probability generating function

$$\sum_{\mathbf{k}} q(\mathbf{k})\, s_1^{k_1} s_2^{k_2} \cdots s_n^{k_n} = \left(\frac{\theta_0}{1-\theta_1 s_1 - \theta_2 s_2 - \cdots - \theta_n s_n}\right)^{\alpha},$$
$$\alpha > 0, \quad \theta_i > 0, \quad \sum \theta_i = 1 \tag{1.11}$$

(see Sections 3 and 5 for details and a discussion of the multiparameter negative multinomial).

(ii) Consider the density of $\mathbf{X} = (X_1, X_2, \ldots, X_n) \sim N(0, \Sigma)$ (normally distributed with covariance matrix Σ). This is MTP_2 if and only $-\Sigma^{-1}$ exhibits nonnegative off-diagonal elements (Sarkar [48], Barlow and Proschan [4]).

(iii) If $\mathbf{X} = (X_1, \ldots, X_n) \sim N(0, \Sigma)$, then the density of the absolute value components $|\mathbf{X}| = (|X_1|, |X_2|, \ldots, |X_n|)$ is MTP_2 if and only if there exists a diagonal matrix $D = \operatorname{diag}(d_1, d_2, \ldots, d_n)$, $d_i = +1$ or -1, with the property that $-D\Sigma^{-1}D$ exhibits nonnegative off-diagonal elements. (Karlin and Rinott [28], the case $n=3$ is due to Abdel-Hameed and Sampson [3].)

(iv) The density of the eigenvalues of certain Wishart matrices is MTP_2 (Dykstra and Hewett [9]).

Other examples are described in Section 3.

Among the operations that preserve the MTP_2 property, we have:

$$\text{If } f(\mathbf{x}) = f(x_1, x_2,\ldots, x_n) \text{ is } \mathrm{MTP}_2, \tag{1.12}$$

then

$$g(\mathbf{x}) = \left(\prod_{i=1}^{n} a_i(x_k)\right) f(b_1(x_1), b_2(x_2),\ldots, b_n(x_n)) \text{ is } \mathrm{MTP}_2, \tag{1.13}$$

where each $a_k(x_k)$ is a positive function and b_i are all monotone increasing (or decreasing).

If

$$f(\mathbf{x}) \quad \text{and} \quad g(\mathbf{x}) \text{ are } \mathrm{MTP}_2, \qquad \text{then} \quad f(\mathbf{x})\, g(\mathbf{x}) \text{ is } \mathrm{MTP}_2. \tag{1.14}$$

Composition formula. Suppose $f(\mathbf{x}, \mathbf{y})$ is MTP_2 over $\mathscr{X} \times \mathscr{Y}$ and $g(\mathbf{y}, \mathbf{z})$ is MTP_2 over $\mathscr{Y} \times \mathscr{Z}$, where $\mathscr{X}$, $\mathscr{Y}$, and $\mathscr{Z}$ are subsets of possibly different Euclidean spaces, then

$$h(\mathbf{x}, \mathbf{z}) = \int f(\mathbf{x}, \mathbf{y})\; g(\mathbf{y}, \mathbf{z})\, d\mathbf{y} \text{ is } \mathrm{MTP}_2 \text{ over } \mathscr{X} \times \mathscr{Z}. \tag{1.15}$$

A useful consequence of (1.15) concerns marginal densities. Accordingly, if

$$\mathbf{X} = (X_1, X_2,\ldots, X_n) \textit{ is } \mathrm{MTP}_2,$$

then any subset, say for definiteness,

$$\mathbf{X}_k = (X_1, X_2,\ldots, X_k) \textit{ is also } \mathrm{MTP}_2 \qquad (2 \leqslant k < n). \tag{1.16}$$

Combinations of the above properties coupled with (1.8) lead to the following coverage probability estimates.

Let $f(\mathbf{x})$ *and* $g(\mathbf{x})$ *be* MTP_2 *and* $\varphi(\mathbf{x})$, $\psi(\mathbf{x})$ *concordantly monotone, then*

$$\left(\int f(\mathbf{x})\; g(\mathbf{x})\, d\mathbf{x}\right)\left(\int f(\mathbf{x})\; g(\mathbf{x})\, \varphi(\mathbf{x})\, \psi(\mathbf{x})\, d\mathbf{x}\right)$$

$$\geqslant \left(\int f(\mathbf{x})\; g(\mathbf{x})\, \varphi(\mathbf{x})\, d\mathbf{x}\right)\left(\int f(\mathbf{x})\; g(\mathbf{x})\, \psi(\mathbf{x})\, d\mathbf{x}\right). \tag{1.17a}$$

In particular, if

$$g(\mathbf{x}) = \begin{cases} 1 & \text{for} \quad \mathbf{x} \in C \\ 0 & \text{for} \quad \mathbf{x} \notin C, \end{cases} \qquad C = \text{any rectangle (bounded or unbounded)}$$

$\mathbf{X}$ following the density $f(\mathbf{x})$ which is MTP_2, then for A and B monotone sets

$$P\{\mathbf{X} \in C\} \, P\{\mathbf{X} \in C \cap A \cap B\} \geqslant P\{\mathbf{X} \in C \cap A\} \, P\{\mathbf{X} \in C \cap B\}.$$

Let A and B be sets in R^n and denote $A \vee B = \{\mathbf{u} = \mathbf{a} \vee \mathbf{b}$ for $\mathbf{a} \in A$, $\mathbf{b} \in B\}$ and $A \wedge B = \{\mathbf{v} = \mathbf{a} \wedge \mathbf{b}$ for $\mathbf{a} \in A$, $\mathbf{b} \in B\}$. Then if $\mathbf{X}$ is MTP_2 we have

$$P\{\mathbf{X} \in A \vee B\} \, P\{\mathbf{X} \in A \wedge B\} \geqslant P\{\mathbf{X} \in A\} \, P\{\mathbf{X} \in B\}. \tag{1.17b}$$

Inequalities (1.17) are sharper than property (1.8) of association.

It may be generally difficult to check that a set of random variables $\mathbf{X} = (X_1, X_2, \dots, X_n)$ is associated while the verification of the stronger property that $\mathbf{X}$ is MTP_2 is often easier. We will display a broad spectrum of MTP_2 densities and by further conditioning and compounding devices we will construct new classes of MTP_2 densities and concomitant associated sets of random variables. To illustrate, if $f(\mathbf{x})$ and $g(\mathbf{x})$ are MTP_2 densities, then $f(\mathbf{x}) \, g(\mathbf{x})$ is MTP_2 by (1.14), but if $f(\mathbf{x})$ and $g(\mathbf{x})$ correspond to associated random vectors, then we cannot in general conclude that $f(\mathbf{x}) \, g(\mathbf{x})$ underlies an associated set of random variables.

In order to ascertain broad classes of stochastic orderings among probability measures we need to deal with pairs of densities fulfilling an appropriate multivariate version of the monotone likelihood ratio property. The hypothesis of the next result is of this kind.

When f_1 and f_2 are densities on R^n satisfying

$$f_2(\mathbf{x} \vee \mathbf{y}) \, f_1(\mathbf{x} \wedge \mathbf{y}) \geqslant f_1(\mathbf{x}) \, f_2(\mathbf{y}) \tag{1.18}$$

for all $\mathbf{x}, \mathbf{y} \in R^n$, then

$$\int \varphi(\mathbf{x}) \, f_1(\mathbf{x}) \, dx \leqslant \int \varphi(\mathbf{x}) \, f_2(\mathbf{x}) \, d\mathbf{x} \qquad \textit{holds for all increasing } \varphi \tag{1.19}$$

(Holley [18], Preston [43], Kemperman [32]).

We use the following notation for (1.18):

$$f_2 \succ_{\mathrm{TP}_2} f_1. \tag{1.18a}$$

In the univariate case the relation (1.18a) is that f_1/f_2 where defined is monotone increasing, i.e., f_1 has a monotone likelihood ratio with respect to f_2.

Consider $f(\mathbf{x}; \boldsymbol{\lambda})$, $\mathbf{x} \in X$, $\boldsymbol{\lambda} \in \Lambda$, depending on a parameter set Λ such that $f(\mathbf{x}, \boldsymbol{\lambda})$ is jointly MTP_2 in the variables $\mathbf{x}$ and $\boldsymbol{\lambda}$. In this situation, appeal to (1.18) under $\boldsymbol{\lambda}^{(2)} \geqslant \boldsymbol{\lambda}^{(1)}$, abbreviating $f_i(\mathbf{x}) = f(\mathbf{x}, \boldsymbol{\lambda}^{(i)})$, $i = 1, 2$, we secure

$$\int \varphi(\mathbf{x})\, f_2(\mathbf{x})\, d\mathbf{x} \geqslant \int \varphi(\mathbf{x})\, f_1(\mathbf{x})\, d\mathbf{x} \qquad \text{for all increasing } \varphi(\mathbf{x}). \tag{1.20}$$

An extension of (1.18) involving four functions was featured in Ahlswede and Daykin [1].

We close the introduction by reviewing the layout of the paper. In Section 2 we present a simple approach to a result of Ahlswede and Daykin from which inequalities such as (1.19), (1.17a, 1.17b), and (1.7)–(1.9) are deduced easily. A number of applications of stochastic monotonicity in comparing probability densities in R^n are indicated. The proof of the composition formula (1.15) and ramifications are presented in Section 3.

A spectrum of important classes of MTP_2 densities are described in Section 3 including the generation of totally positive transition stochastic kernels. We already remarked that a MTP_2 density conveys the fact that the underlying random variables are associated. This entails a host of relevant statistical and probabilistic inequalities. By conditioning and compounding devices we extend the association inequalities to cover new cases (Section 4). Applying the ideas to sequences of random variables we further derive a number of moment-type inequalities extending some work of Sidak [51] and Tong [54].

In our discussion of the multiparameter negative multinomial density we extend a univariate representation formula of Hardy, Littlewood, and Pólya. More explicitly, a multiparameter negative multinomial density $q(\mathbf{k}) = q(\mathbf{k}, \mathbf{p}) = P\{Y_1 = k_1,\ Y_2 = k_2, \ldots, Y_n = k_n\}$ (Y_i is the number of times event category i appears before r occurrences of event type 0) having generating function

$$\sum_{\mathbf{k}} q(\mathbf{k})\, s_1^{k_1} s_2^{k_2} \cdots s_n^{k_n} = \prod_{\nu=1}^{r} \frac{p_{0,\nu}^{\alpha}}{(1 - p_{(\nu)1} s_1 - p_{(\nu)2} s_2 - \cdots - p_{(\nu)n} s_n)},$$

$$p_{(\nu)i} > 0, \qquad \sum_{i=0}^{n} P_{(\nu)i} = 1,$$

satisfies

$$\begin{aligned}P\{Y_1 = k_1, Y_2 = k_2, \dots, Y_n = k_n\} \\ &= q(\mathbf{k}) \\ &= \frac{\prod_{\nu=0}^{r} p_{(\nu)0}}{\prod_{i=1}^{n} k_i!} \int_0^\infty \cdots \int_0^\infty \left(\sum_{\nu=1}^{r} p_{(\nu)1} t_\nu\right)^{k_1} \\ &\quad \times \left(\sum_{\nu=1}^{r} p_{(\nu)2} t_\nu\right)^{k_2} \cdots \left(\sum_{\nu=1}^{r} p_{(\nu)n} t_\nu\right)^{k_n} \exp\left[-\sum_{\nu=1}^{r} t_\nu\right] dt_1 \cdots dt_r.\end{aligned} \tag{1.21}$$

In Section 5 this formula and another variants of (1.21) are proved and interpreted in relation to MTP_2 and log concavity properties of $q(\mathbf{k})$.

Some comments on the implication of MTP_2 densities in terms of positive correlation preservation properties for monotone Markov processes are discussed in Section 6. Several inequalities discussed in this paper are quite useful in the area of simultaneous test procedures; see Krishnaiah [35].

2. Concept and Properties of Multivariate Monotone Likelihood Ratio Relationships

Let $\mathscr{X} = \prod_{i=1}^{n} \mathscr{X}_i$ be a product of totally ordered space $\mathscr{X}_i$, $i = 1,\dots, n$, endowed with the natural partial ordering: for $\mathbf{x}, \mathbf{y} \in \mathscr{X}$ we write $\mathbf{x} \leqslant \mathbf{y}$ if $\mathbf{x} = (x_1,\dots, x_n)$, $\mathbf{y} = (y_1,\dots, y_n)$ satisfy $x_i \leqslant y_i$ in $\mathscr{X}_i$ for $i = 1,\dots, n$. Let $\sigma = \sigma_1 \times \cdots \times \sigma_n$ denote a product measure on $\mathscr{X}$ where σ_i are σ-finite measures on $\mathscr{X}_i$, $i = 1,\dots, n$. We shall generally use the abbreviations $d\sigma_i(x_i) = dx_i$ and $d\sigma(\mathbf{x}) = d\mathbf{x}$ without ambiguity and $\int_{\mathscr{X}} f(\mathbf{x})\, d\sigma(\mathbf{x}) = \int f(\mathbf{x})\, d\mathbf{x}$, etc. Whenever integrals or expectations appear, requirements of measurability and integrability are tacitly assumed without further mention. For $\mathbf{x} = (x_1,\dots, x_n)$, $\mathbf{y} = (y_1,\dots, y_n)$ in $\mathscr{X}$, define the lattice operations

$$\mathbf{x} \vee \mathbf{y} = (\max(x_1, y_1),\dots, \max(x_n, y_n)),$$
$$\mathbf{x} \wedge \mathbf{y} = (\min(x_1, y_1),\dots, \min(x_n, y_n)).$$

The following theorem is essentially due to Ahlswede and Daykin [1]. Their formulation consists of embedding the lattice $\mathscr{X}$ in a lattice of subsets of a finite set. Their proof proceeds by a delicate induction on the number of elements in the finite set followed by an approximation argument to extend the conclusion to the continuous case.

Our proof is related to the approach in Karlin [24], Preston [14], and Kemperman [32] operating via induction on the dimension of $\mathscr{X}$.

THEOREM 2.1. *Let f_1, f_2, f_3, and f_4 be nonnegative functions on $\mathscr{X}$ satisfying for all* $\mathbf{x}$, $\mathbf{y} \in \mathscr{X}$

$$f_1(\mathbf{x})\, f_2(\mathbf{y}) \leqslant f_3(\mathbf{x} \vee \mathbf{y})\, f_4(\mathbf{x} \wedge \mathbf{y}). \tag{2.1}$$

Then

$$\int f_1(\mathbf{x})\, d\mathbf{x} \int f_2(\mathbf{x})\, d\mathbf{x} \leqslant \int f_3(\mathbf{x})\, d\mathbf{x} \int f_4(\mathbf{x})\, d\mathbf{x}. \tag{2.2}$$

Proof. We proceed by induction on the dimension of $\mathscr{X}$. Since (2.1) entails

$$f_1(\mathbf{x})\, f_2(\mathbf{x}) \leqslant f_3(\mathbf{x})\, f_4(\mathbf{x}), \tag{2.3}$$

the comparison of (2.2) for $n = 1$ will follow provided we confirm the inequality

$$\begin{aligned} &\int_{x<y}\int (f_1(x)\, f_2(y) + f_1(y)\, f_2(x))\, dx\, dy \\ &\quad \leqslant \int_{x<y}\int (f_3(x)\, f_4(y) + f_3(y)\, f_4(x))\, dx\, dy. \end{aligned} \tag{2.4}$$

To this end, abbreviate for $x < y$ fixed, $f_1(x)\, f_2(y) = a$, $f_1(y)\, f_2(x) = b$, $f_3(x)\, f_4(y) = c$, and $f_3(y)\, f_4(x) = d$. It suffices to establish that $a + b \leqslant c + d$. By (2.1) we know $d \geqslant a, b$ and, by (2.3), $ab \leqslant cd$. Then $a + b \leqslant c + d$ holds since $c + d - (a + b) = (1/d)[(d - a)(d - b) + (cd - ab)] \geqslant 0$.

In order to advance the induction step we next prove that if f_1, f_2, f_3, f_4 satisfy (2.1), then the (marginals) functions $\varphi_j(\mathbf{x}) = \int_{\mathscr{X}_n} f_j(\mathbf{x}, x)\, d\sigma_n(x)$, $j = 1, 2, 3, 4$, $\mathbf{x} \in \prod_{i=1}^{n-1} \mathscr{X}_i$ continue to satisfy (2.1) on $\prod_{i=1}^{n-1} \mathscr{X}_i$, i.e.,

$$\varphi_1(\mathbf{x})\, \varphi_2(\mathbf{y}) \leqslant \varphi_3(\mathbf{x} \vee \mathbf{y})\, \varphi_4(\mathbf{x} \wedge y). \tag{2.5}$$

Indeed, (2.5) written out reduces to

$$\begin{aligned} &\int_{x<y}\int [f_1(\mathbf{x}, x)\, f_2(\mathbf{y}, y) + f_1(\mathbf{x}, y)\, f_2(\mathbf{y}, x)]\, d\sigma_n(x)\, d\sigma_n(y) \\ &\leqslant \int_{x<y}\int [f_3(\mathbf{x} \vee \mathbf{y}, x)\, f_4(\mathbf{x} \wedge \mathbf{y}, y) + f_3(\mathbf{x} \vee \mathbf{y}, y)\, f_4(\mathbf{x} \wedge \mathbf{y}, x)]\, d\sigma_n(x)\, d\sigma_n(y). \end{aligned} \tag{2.6}$$

Setting $a = f_1(\mathbf{x}, x)\, f_2(\mathbf{y}, y)$, $b = f_1(\mathbf{x}, y)\, f_2(\mathbf{y}, x)$, $c = f_3(\mathbf{x} \vee \mathbf{y}, x)$ $f_4(\mathbf{x} \wedge \mathbf{y}, y)$, and $d = f_3(\mathbf{x} \vee \mathbf{y}, y)\, f_4(\mathbf{x} \wedge \mathbf{y}, x)$ we observe again that $d \geqslant a, b$

and $ab \leqslant cd$ by (2.1). Therefore (2.6) obtains as before. The induction hypotheses can be applied to yield

$$\int \varphi_1(\mathbf{x})\, d\mathbf{x} \int \varphi_2(\mathbf{x})\, d\mathbf{x} \leqslant \int \varphi_3(\mathbf{x})\, d\mathbf{x} \int \varphi_4(\mathbf{x})\, d\mathbf{x}.$$

Substituting for $\varphi_i(\mathbf{x})$, the above passes into (2.2). The induction step is advanced completing the proof.

EXAMPLE 2.1. Let A, $B \subseteq \mathscr{X}$ and define

$$A \vee B = \{\mathbf{a} \vee \mathbf{b} : \mathbf{a} \in A, \mathbf{b} \in B\},$$

$$A \wedge B = \{\mathbf{a} \wedge \mathbf{b} : \mathbf{a} \in A, \mathbf{b} \in B\}.$$

Denote the indicator function of a set A by

$$I_A(\mathbf{x}) = \begin{cases} 1, & \mathbf{x} \in A \\ 0, & \mathbf{x} \notin A \end{cases}$$

and set $f_1 = I_A$, $f_2 = I_B$, $f_3 = I_{A \vee B}$, $f_4 = I_{A \wedge B}$. Then direct examination of cases verifies that f_1, f_2, f_3, f_4 satisfy (2.1).

Remark 2.1. If f_1, f_2, f_3, f_4 satisfy (2.1) and the same holds for a set of functions g_1, g_2, g_3, g_4, then the products $f_1 g_1, f_2 g_2, f_3 g_3, f_4 g_4$ also satisfy (2.1). The conjunction of Theorem 2.1 applied to $f_i g_i$, $i = 1, 2, 3, 4$, where g_i are indicator functions of the type described in Example 2.1, yields

COROLLARY 2.1. *If the nonnegative functions* f_1, f_2, f_3, f_4 *satisfy* (2.1), *then*

$$\int_A f_1(\mathbf{x})\, d\mathbf{x} \int_B f_2(\mathbf{x})\, d\mathbf{x} \leqslant \int_{A \vee B} f_3(\mathbf{x})\, d\mathbf{x} \int_{A \wedge B} f_4(\mathbf{x})\, d\mathbf{x}. \tag{2.7}$$

DEFINITION 2.1. A function $\varphi: \mathscr{X} \to R$ is said to be increasing (decreasing) if $x \leqslant y$ implies $\varphi(\mathbf{x}) \leqslant \varphi(\mathbf{y})$ ($\varphi(\mathbf{x}) \geqslant \varphi(\mathbf{y})$).

Let f_1, f_2 be nonnegative functions on $\mathscr{X}$, satisfying

$$f_1(\mathbf{x})\, f_2(\mathbf{y}) \leqslant f_2(\mathbf{x} \vee \mathbf{y})\, f_1(\mathbf{x} \wedge \mathbf{y}) \tag{2.8}$$

i.e., $f_2 \succ_{\mathrm{TP}_2} f_1$, see (1.18a) and let φ be an increasing nonnegative function. Define

$$f_1^* = f_1 \varphi, \qquad f_2^* = f_2, \qquad f_3^* = f_2 \varphi, \qquad f_4^* = f_1.$$

Then $f_1^*, f_2^*, f_3^*, f_4^*$ satisfy (2.1) and Theorem 2.1 can be applied. When $\int f_1(\mathbf{x})\,d\sigma(\mathbf{x}) = \int f_2(\mathbf{x})\,d\sigma(\mathbf{x}) = 1$ the assumption that φ is nonnegative can be omitted by subtracting a constant, truncating first if necessary and using a standard limiting argument. These considerations establish.

THEOREM 2.2 (Holley [18], Preston [43], Kemperman [32]). *Let $f_1(\mathbf{x})$ and $f_2(\mathbf{x})$ be probability densities on $\mathscr{X}$ satisfying for all $\mathbf{x}, \mathbf{y} \in \mathscr{X}$*

$$f_1(\mathbf{x})\, f_2(\mathbf{y}) \leqslant f_2(\mathbf{x} \vee \mathbf{y})\, f_1(\mathbf{x} \wedge \mathbf{y}). \tag{2.9}$$

(i.e., $f_2 \succ_{\mathrm{TP}_2} f_1$). Then for any increasing φ on $\mathscr{X}$

$$\int \varphi(\mathbf{x})\, f_1(\mathbf{x})\, d\mathbf{x} \leqslant \int \varphi(\mathbf{x})\, f_2(\mathbf{x})\, d\mathbf{x}. \tag{2.10}$$

THEOREM 2.3 (Sarkar [48], Fortuin, Ginibre, and Kasteleyn [14], Preston [43]). *Let f be a probability density with respect to $d\sigma(\mathbf{x})$ on $\mathscr{X}$, satisfying for $\mathbf{x}, \mathbf{y} \in \mathscr{X}$.*

$$f(\mathbf{x})\, f(\mathbf{y}) \leqslant f(\mathbf{x} \vee \mathbf{y})\, f(\mathbf{x} \wedge \mathbf{y}). \tag{2.11}$$

Then for any pair of increasing (or decreasing) functions φ and ψ on $\mathscr{X}$ we have

$$\int \varphi(\mathbf{x})\, \psi(\mathbf{x})\, f(\mathbf{x})\, d\sigma(\mathbf{x}) \geqslant \left(\int \varphi(\mathbf{x})\, f(\mathbf{x})\, d\sigma(\mathbf{x})\right)\left(\int \psi(\mathbf{x})\, f(\mathbf{x})\, d\sigma(\mathbf{x})\right). \tag{2.12}$$

Proof. Without loss of generality we assume $\psi(\mathbf{x})$ positive and define

$$f_2(\mathbf{x}) = \frac{\psi(\mathbf{x})\, f(\mathbf{x})}{\gamma}, \qquad f_1(\mathbf{x}) = f(\mathbf{x}), \tag{2.13}$$

where $\gamma = \int \psi(\mathbf{x})\, f(\mathbf{x})\, d\sigma$. When ψ is increasing, the functions f_1 and f_2 defined in (2.13) clearly satisfy (2.9) and since f is a probability density, $\int f_1(\mathbf{x})\, d\sigma = \int f_2(\mathbf{x})\, d\sigma = 1$. Inequality (2.10) for the case at hand provides (2.12).

In the following discussion we briefly describe the relationship between the results described above and the theory of total positivity.

PROPOSITON 2.1. *Let $f(\mathbf{x}) = f(x_1, \ldots, x_n)$, $\mathbf{x} \in \mathscr{X}$ be TP_2 in every pair of arguments when the remaining arguments are held constant, and suppose that* $f(\mathbf{x})\, f(\mathbf{y}) \neq 0$ implies $f(\mathbf{u})\, f(\mathbf{v}) \neq 0$ for any $\mathbf{x}, \mathbf{y} \in \mathscr{X}$, $\mathbf{x} \wedge \mathbf{y} \leqslant \mathbf{u}$, $\mathbf{v} \leqslant \mathbf{x} \vee \mathbf{y}$. Then for all $\mathbf{x}, \mathbf{y} \in \mathscr{X}$

$$f(\mathbf{x})\, f(\mathbf{y}) \leqslant f(\mathbf{x} \vee \mathbf{y})\, f(\mathbf{x} \wedge \mathbf{y}).$$

Proof. (This result is essentially due to Lorentz [31]; see also Rinott [46, inequality (2.4)], and Kemperman [32, assertion (i)]).

Suppose (without loss of generality) $\mathbf{x} = (x_1^*,..., x_k^*, x_{k+1},..., x_n)$, $\mathbf{y} = (y_1,..., y_k, y_{k+1}^*,..., y_n^*)$, where $x_i^* \geqslant y_i$, $i = 1,..., k$, and $x_j \leqslant y_j^*$, $j = k+1,..., n$. Then

$$\begin{aligned}
&\frac{f(\mathbf{x} \vee \mathbf{y}) f(\mathbf{x} \wedge \mathbf{y})}{f(\mathbf{x}) f(\mathbf{y})} \\
&\quad = \frac{f(x_1^*,..., x_k^*, y_{k+1}^*,..., y_n^*) f(y_1,..., y_k, x_{k+1},..., x_n)}{f(x_1^*,..., x_k^*, x_{k+1},..., x_n) f(y_1,..., y_k, y_{k+1}^*,..., y_n^*)} \\
&\quad = \frac{f(x_1^*,..., x_k^*, y_{k+1}^*,..., y_n^*) f(x_1^*, y_2,..., y_k, x_{k+1},..., x_n)}{f(x_1^*,..., x_k^*, x_{k+1},..., x_n) f(x_1^*, y_2,..., y_k, y_{k+1}^*,..., y_n^*)} \\
&\qquad \times \frac{f(x_1^*, y_2,..., y_k, y_{k+1}^*,..., y_n^*) f(y_1,..., y_k, x_{k+1},..., x_n)}{f(x_1^*, y_2,..., y_k, x_{k+1},..., x_n) f(y_1,..., y_k, y_{k+1}^*,..., y_n^*)}.
\end{aligned} \tag{2.14}$$

The last expression is a product of two terms which either exceed or equal to one by the induction hypothesis: the first term by fixing x_1^* and applying the induction hypothesis to the remaining $n-1$ variables, and the second by fixing $y_2,..., y_k$. (Note that the denominators are nonzero since they contain terms of the form $f(\mathbf{u})$ with $\mathbf{x} \wedge \mathbf{y} \leqslant \mathbf{u} \leqslant \mathbf{x} \vee \mathbf{y}$.)

The analysis of Theorem 2.1 shows that when f_1 and f_2 satisfy (2.8) then the corresponding marginals continue to satisfy (2.8). These facts lead to the following theorem.

THEOREM 2.4. *If $f \succ_{TP_2} g$ and $K(\mathbf{x}, \mathbf{y})$ is MTP_2, then*

$$\int K(\mathbf{x}, \mathbf{y}) f(\mathbf{y})\, d\mathbf{y} \succ_{TP_2} \int K(\mathbf{x}, \mathbf{y})\, g(\mathbf{y})\, dy. \tag{2.15}$$

More generally, if $f(\mathbf{y}) \succ_{TP_2} g(\mathbf{y})$ and $K(\mathbf{x}, \mathbf{y}) \succ_{TP_2} L(\mathbf{x}, \mathbf{y})$, then

$$\int K(\mathbf{x}, \mathbf{y}) f(\mathbf{y})\, d\mathbf{y} \succ_{TP_2} \int L(\mathbf{x}, \mathbf{y})\, g(\mathbf{y})\, \mathbf{dy}. \tag{2.16}$$

3. Multivariate Total Positivity of Order 2. Some Properties and Examples

DEFINITION 3.1. *A function $f: \mathscr{X} \to [0, \infty)$ will be called multivariate totally positive of order 2 (MTP_2) if for all $\mathbf{x}, \mathbf{y} \in \mathscr{X}$*

$$f(\mathbf{x}) f(\mathbf{y}) \leqslant f(\mathbf{x} \vee \mathbf{y}) f(\mathbf{x} \wedge \mathbf{y}). \tag{3.1}$$

In this develop some properties of MTP_2 functions, examples, related inequalities, and applications.

The following proposition is confirmed easily by induction, taking cognizance of relation (2.5) from the proof of Theorem 2.1.

PROPOSITION 3.1. *Let f_1, f_2, f_3, f_4 be nonnegative functions on $\mathscr{X}$ satisfying for all* $\mathbf{x}, \mathbf{y} \in \mathscr{X}$

$$f_1(\mathbf{x})\, f_2(\mathbf{y}) \leqslant f_3(\mathbf{x} \vee \mathbf{y})\, f_4(\mathbf{x} \wedge \mathbf{y}). \tag{3.2}$$

Let $\varphi_i(x_1,\dots,x_k) = \int_{\mathscr{X}_n} \cdots \int_{\mathscr{X}_{k+1}} f(x_1,\dots,x_k, x_{k+1},\dots,x_n)\, dx_{k+1} \cdots dx_n$, $i = 1,\dots,4$ *(recall $dx_i = d\sigma_i(x_i)$). Then for all* $\mathbf{x}, \mathbf{y} \in \prod_{i=1}^k \mathscr{X}_i$, *we have*

$$\varphi_1(\mathbf{x})\, \varphi_2(\mathbf{y}) \leqslant \varphi_3(\mathbf{x} \vee \mathbf{y})\, \varphi_4(\mathbf{x} \wedge \mathbf{y}).$$

PROPOSITION 3.2. *Let f be a MTP_2 function on $\mathscr{X}$. Then the marginal function φ defined on $\prod_{i=1}^k \mathscr{X}_i$ by*

$$\varphi(x_1,\dots,x_k) = \int_{\mathscr{X}_n} \cdots \int_{\mathscr{X}_{k+1}} f(x_1,\dots,x_k, x_{k+1},\dots,x_n)\, dx_{k+1} \cdots dx_n \tag{3.3}$$

is MTP_2. (This is the statement of (1.16) *in the Introduction.)*

The following proposition is checked in a straightforward manner.

PROPOSITION 3.3. (Property (1.14)). *Let f and g be MTP_2 functions. Then fg is MTP_2.*

The conjunction of Propositions 3.2 and 3.3 imply

PROPOSITION 3.4. *Let $\mathscr{X} = \prod_{i=1}^n \mathscr{X}_i$, $\mathscr{Y} = \prod_{i=1}^m \mathscr{Y}_i$, $\mathscr{Z} = \prod_{i=1}^k \mathscr{Z}_i$, where $\mathscr{X}_i$, $\mathscr{Y}_i$, and $\mathscr{Z}_i$ are totally ordered spaces. Let f be MTP_2 on $\mathscr{Y} \times \mathscr{X}$, and g be MTP_2 on $\mathscr{X} \times \mathscr{Z}$. Define*

$$h(\mathbf{y}, \mathbf{z}) = \int_{\mathscr{X}} f(\mathbf{y}, \mathbf{x})\; g(\mathbf{x}, \mathbf{z})\, d\sigma(\mathbf{x}), \tag{3.4}$$

where as before $\sigma = \sigma_1 \times \cdots \times \sigma_n$.

Then h is MTP_2 on $\mathscr{Y} \times \mathscr{Z}$.

Propositon 3.3 also readily implies:

PROPOSITION 3.5. *Independent random variables have a joint MTP_2 density.*

Generally a product of nonnegative functions $f(\mathbf{x}) = \prod_{i=1}^{n} f_i(x_i)$ is MTP_2 with equality in (3.1).

The following proposition is elementary.

PROPOSITION 3.6. *If $f(\mathbf{x})$, $\mathbf{x} \in \mathscr{X}$ is MTP_2, and $\varphi_1,\ldots,\varphi_n$ are all increasing (or all decreasing) functions on $\mathscr{X}_1,\ldots,\mathscr{X}_n$, respectively, then the function*

$$\psi(\mathbf{x}) = \psi(x_1,\ldots,x_n) = f(\varphi_1(x_1),\ldots,\varphi_n(x_n))$$

is MTP_2 on $\mathscr{X}$.

This implies property (1.13) in the Introduction.

Propositions 3.2–3.6 provide methods for generating new MTP_2 functions from other MTP_2 functions.

We next highlight several important classes of MTP_2 probability densities.

EXAMPLE 3.1. *Normal distribution. Let* $\mathbf{X} = (X_1,\ldots,X_n) \sim N(\mu, \Sigma)$, i.e., $\mathbf{X}$ follows the density

$$f(\mathbf{x}) = (2\pi)^{-n/2} |\Sigma|^{-1/2} \exp[-\tfrac{1}{2}(\mathbf{x}-\boldsymbol{\mu})' B(\mathbf{x}-\boldsymbol{\mu})], \tag{3.5}$$

where $\Sigma^{-1} = B = \|b_{ij}\|_{i,j=1}^{n}$. This density is manifestly TP_2 in each pair of arguments, and hence MTP_2 if and only if $b_{ij} \leqslant 0$ for all $i \neq j$ (Sarkar [48], Barlow and Proschan [4]). Such a matrix B is known as an M-matrix, Leontief matrix in the economic literature.

Consider the partitions $\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$, $B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}$ where Σ_{11} and B_{11} are $k \times k$ matrices. Then, $(X_1,\ldots,X_k)$ have a normal distribution with covariance Σ_{11}. If $X_1,\ldots,X_n$ have a MTP_2 density, or equivalently $B = \Sigma^{-1}$ is an M-matrix, then by Proposition 3.2 the (marginal) density of $(X_1,\ldots,X_k)$ is MTP_2, and therefore the matrix $\Sigma_{11}^{-1} = B_{11} - B_{12}B_{11}^{-1}B_{21}$ is an M-matrix.

EXAMPLE 3.2. *Absolute-value multinormal variables.* Let $\mathbf{X} = (X_1,\ldots,X_n)$ be a random vector having density $f(\mathbf{x})$ of (3.5) with $\boldsymbol{\mu} = \mathbf{0}$. It is proved by Karlin and Rinott [28] that the joint density of $(|X_1|,\ldots,|X_n|)$ is MTP_2 if and only if there exists a diagonal matrix D with elements ± 1 such that the off-diagonal elements of $-D\Sigma^{-1}D$ are all nonnegative. (The case of $n = 3$ is due to Abdel-Hameed and Sampson [3]). For example, if $\mathbf{X} \sim N(0, \Sigma)$ and $\Sigma = \Gamma + \boldsymbol{\alpha}\boldsymbol{\alpha}'$, $\boldsymbol{\alpha}' = (\alpha_1,\ldots,\alpha_n) \in R^n$, Γ diagonal, or if

$$\Sigma = \Gamma + \begin{pmatrix} \boldsymbol{\alpha}\boldsymbol{\alpha}' & \rho\boldsymbol{\alpha}\boldsymbol{\beta}' \\ \rho\boldsymbol{\beta}\boldsymbol{\alpha}' & \boldsymbol{\beta}\boldsymbol{\beta}' \end{pmatrix},$$

where $\boldsymbol{\alpha}' = (\alpha_1,\ldots,\alpha_k)$, $\boldsymbol{\beta}' = (\beta_{k+1},\ldots,\beta_n)$, then $(|X_1|,\ldots,|X_n|)$ has an MTP_2 density (Karlin and Rinott [28, Sect. 4]).

EXAMPLE 3.3 (Dykstra and Hewett [9]). *Characteristic roots of random Wishart matrices.* Consider $\mathbf{S}$ distributed as a Wishart random matrix with identity covariance matrix. Let $\mathbf{S}_1$ and $\mathbf{S}_2$ be two independent Wishart random matrices having the same covariance matrix. The joint density of the characteristic roots of $\mathbf{S}$ is of the form $c \prod_{i=1}^{p} g_i(x_i) \prod_{i<j} (x_i - x_j)_+$, $c > 0$, where $(x-y)_+ = x - y$ for $x > y$ and $(x-y)_+ = 0$ for $x \leqslant y$ and is therefore MTP_2. The same is true for the characteristic roots of $\mathbf{S}_1 \mathbf{S}_2^{-1}$ and $\mathbf{S}_1(\mathbf{S}_1 + \mathbf{S}_2)^{-1}$. This fact was used by Perlman and Olkin [40] in proving unbiasedness of certain tests for Manova, whose statistics are functions of such characteristic roots.

EXAMPLE 3.4. *Multivariate logistic distribution.* Gumbel [15] defined the multivariate logistic density as

$$f(x_1,..., x_n) = n! \exp\left\{-\sum_{i=1}^{n} x_i\right\}\left\{1 + \sum_{i=1}^{n} e^{-x_i}\right\}^{-(m+1)}. \tag{3.6}$$

The generalized Cauchy kernel $k(\mathbf{y}) = (1/(1 + \sum_{i=1}^{n} y_i)^{\alpha})$, $y_i > 0$, $i = 1,..., n$ is TP_2 in every pair of variables. We then invoke Propositions 3.3 and 3.6 to deduce that (3.6) is MTP_2.

The following propositions provide further methods for generating MTP_2 densities. Applications to the multivariate Γ and F distributions will be given in the sequel.

PROPOSITION 3.7. *Let $\mathbf{X} = (X_1,..., X_n)$ be a random vector of independent components $X_1,..., X_n$, each X_i, $i = 1,..., n$, governed by a PF_2 density function f_{X_1} (see [24]). Let $\mathbf{Y} = (Y_1,..., Y_n)$ have a joint MTP_2 density $f_{\mathbf{Y}}$ on R^n, and suppose $\mathbf{X}$ and $\mathbf{Y}$ are independent. Then $\mathbf{Z} = \mathbf{X} + \mathbf{Y}$ has a MTP_2 density.*

Proof. The joint density of $\mathbf{Z}$ is given by

$$f_{\mathbf{Z}}(z_1,..., z_n) = \int \left(\prod_{i=1}^{n} f_{X_i}(z_i - y_i)\right) f_{\mathbf{Y}}(y_1,..., y_n)\, d\mathbf{y}. \tag{3.7}$$

Since f_{X_i} is PF_2 $f_{X_i}(z_i - y_i)$ is TP_2 in (z_i, y_i) and therefore the integrand in (3.7) is MTP_2 in $(z_1,..., z_n, y_1,..., y_n)$.

We know by Proposition 3.2 that $f_{\mathbf{Z}}(\mathbf{z})$ is MTP_2 as claimed.

PROPOSITION 3.8. *Let $\mathbf{X}$ be as in Proposition 3.7 and let X_0 be a random variable independent of $\mathbf{X}$ having a density f_{X_0}. Define $Z_i = X_i + X_0$, $i = 1,..., n$. Then the joint density of $\mathbf{Z} = (Z_1,..., Z_n)$ is MTP_2.*

Proof. The proof is similar to that of Proposition 3.7 with

$$f_{\mathbf{Z}}(z_1,..., z_n) = \int \prod_{i=1}^{n} f_{X_i}(z_i - x)\, f_{X_0}(x)\, dx.$$

A similar argument also yields the following:

PROPOSITION 3.9. *Let* $\mathbf{X} = (X_1,..., X_n)$ *have independent components* $X_1,..., X_n$ *and assume that each* X_i, $i = 1,..., n$, *has a density function* f_{X_i}. *Let* X_0 *be a positive random variable. If for* $i = 1,..., n$, *either* $f_{X_i}(u/v)$ *is* TP_2 *in* $-\infty < u < \infty$ *and* $v > 0$, *or* $f_{X_i}(uv)$ *is* TP_2 *in* $-\infty < u < \infty$, $v > 0$, *then both random vectors* $\mathbf{Z} = (X_1X_0,..., X_nX_0)$ *and* $\mathbf{Z} = (X_1/X_0,..., X_n/X_0)$ *have* MTP_2 *densities.*

EXAMPLE 3.5. *The multivariate gamma distribution.* Let $X_1,..., X_n$ be independent, $X_i \sim \Gamma(\alpha_i, \beta_i)$ $\alpha_i \geqslant 1$, $\beta_i > 0$, $i = 1,..., n$. Then $f_{X_i}(x) = c_i x^{\alpha_i - 1} e^{-\beta_i x}$, $x > 0$ is a PF_2 density (see Karlin [24]). If X_0 is independent of $X_1,..., X_n$, $X_0 \sim \Gamma(\alpha_0, \beta_0)$, then the vector $\mathbf{Z} = (X_1 + X_0,..., X_n + X_0)$ is said to have a multivariate gamma distribution (Ramabhadran [45], see also Johnson and Kotz [20, p. 217]). By Proposition 3.8, $\mathbf{Z}$ has an MTP_2 joint density.

EXAMPLE 3.6. *Multivariate F distribution.* Let $X_i \sim \Gamma(\alpha_i, \beta_i)$, $\alpha_i > 0$, $\beta_i > 0$, $i = 1,..., n$. Then $f_{X_i}(u/v) = c_i(u/v)^{\alpha_i - 1} e^{-\beta_i u/v}$ is TP_2 in u and v positive. Proposition 3.9 implies, for example, that if $X_0, X_1,..., X_n$ are independent, with $X_i \sim \chi^2_{\nu_i}$ (chi square with ν degrees of freedom), then $\mathbf{Z} = ((X_1/\nu_1)(X_0/\nu_0)^{-1},..., (X_n/\nu_n)(X_0/\nu_0)^{-1})$ has a joint MTP_2 density. $\mathbf{Z}$ is said to follow a multivariate F distribution (Finney [13]; see also Johnson and Kotz [20, p. 240]).

EXAMPLE 3.7. *Absolute value multivariate Cauchy distribution.* Let $\mathbf{X} = (X_1,..., X_n) \sim N(0, I)$ and $S \sim \chi^2$, $\mathbf{X}$ and S independent. Set

$$\mathbf{Z} = (Z_1,..., Z_n) = \left(\frac{X_1}{S}, \frac{X_2}{S},..., \frac{X_n}{S}\right)$$

$\mathbf{Z}$ is said to have a multivariate Cauchy distribution (Johnson and Kotz [20, p. 123]).

Then $(|Z_1|,..., |Z_n|)$ have a joint MTP_2 density. This follows by the same argument as the previous example since e^{-u^2/v^2} is TP_2 in u, $v > 0$.

The following proposition presents a class of MTP_2 densities.

PROPOSITON 3.10 (Barlow and Proschan [4]). *Let* $\mathbf{Y}=(Y_1,..., Y_n)$ *describe the evolution of a Markov chain (i.e., the successive time realizations) with* TP_2 *transition probability densities. Then* $\mathbf{Y}$ *has a* MTP_2 *joint density.*

Proof. Let $P(Y_i \in A \mid Y_{i-1}=y)=\int_A f_i(x, y)\, d\sigma_i(x)$, $i=2,..., n$ and $P(Y_1 \in A)=\int_A f_1(x)\, d\sigma_1(x)$. The Markovian hypothesis implies that the joint density of $\mathbf{Y}$ with respect to $\sigma=\sigma_1 \times \cdots \times \sigma_n$ is given by the product $f_1(y_1)\, f_2(y_2, y_1) \cdots f_n(y_n, y_{n-1})$, which is MTP_2 by Proposition 3.3 since each $f_i(y_i, y_{i-1})$, $i=2,..., n$, is assumed to be TP_2.

The scope of applications of Proposition 3.10 is enhanced in view of a theorem by Karlin and McGregor [26] (see also Karlin [24, p. 38]) which indicates that the transition density of a temporally homogeneous strong Markov process on the real line, having continuous sample paths (with probability one) is totally positive. The same applies to any birth–death process.

PROPOSITION 3.11. *Let* $X_1,..., X_n$ *be a sample of i.i.d. random variables* X_i *having a density function* f. *Then the joint density of the order statistics* $X_{(1)},..., X_{(n)}$ *is* MTP_2.

Proof. The joint density of $X_{(1)},..., X_{(n)}$ is given by $n!\, g(x_1,..., x_n) \prod_{i=1}^n f(x_i)$, where

$$g(x_1,..., x_n)=\begin{cases} 1, & x_1 \leqslant \cdots \leqslant x_n \\ 0 & \text{otherwise.} \end{cases}$$

A direct verification shows that g is MTP_2 and the result follows.

Remark. Note that if $f(x+y)$ is TP_2, then the "interval times" $D_i=X_{(i)}-X_{(i-1)}$, $i=1,..., n$, where $X_{(0)}=0$ have the joint MTP_2 density $f(d_1,..., d_n)=n!\, f(d_1)\, f(d_1+d_2) \cdots f(d_1+\cdots+d_n)$.

We conclude this section with a discrete example.

EXAMPLE 3.8. *The negative multinomial distribution.* Consider the density

$$f(k_1,..., k_n)=\frac{\Gamma(N+\sum_{i=1}^n k_i)}{(\prod_{i=1}^n k_i!)\, \Gamma(N)}\, p_0^N \prod_{i=1}^n p_i^{k_i}, \tag{3.9}$$

$k_i=0, 1, 2,...$, $i=1,..., n$. Since the classical gamma function $\Gamma(x+y)$ is TP_2, it follows that $f(k_1,..., k_n)$ is MTP_2.

4. Probability Inequalities of MTP_2 Variables and Concepts of Positive Dependence

The relations between multivariate total positivity and concepts of positive dependence with attendant probability inequalities has been dealt with from varying viewpoints and interests by many authors (e.g., Sarkar [48], Esary and Proschan [10, 11], Fortuin, Kasteleyn, and Ginibre [14], Barlow and Proschan [4], Sidak [50]).

We do not attempt to provide a comprehensive account of the subject in the present paper, but rather to emphasize some basic results and new aspects.

THEOREM 4.1 (Sarkar [48]). *Let* $\mathbf{X} = (X_1, \ldots, X_n)$ *be a random vector having a* MTP_2 *joint density* f *with respect to some product measure* $\sigma = \sigma_1 \times \cdots \times \sigma_n$ *on* R^n. *Then for any increasing function* $\varphi: R^k \to R$, $1 \leqslant k \leqslant n$ *we have that*

$$E\{\varphi(X_1, \ldots, X_k) \mid X_{k+1} = x_{k+1}, \ldots, X_n = x_n\}$$

is increasing in $x_{k+1}, \ldots, x_n$.

Remark 4.1. This property is called the *conditional monotone regression* endowment by Lehman [36].

Proof. Let $x_{k+1} \leqslant x^*_{k+1}, \ldots, x_n \leqslant x^*_n$ and define

$$f_1(x_1, \ldots, x_k) = \frac{f(x_1, \ldots, x_k, x_{k+1}, \ldots, x_n)}{\hat{f}(x_{k+1}, \ldots, x_n)},$$

$$f_2(x_1, \ldots, x_k) = \frac{f(x_1, \ldots, x_k, x^*_{k+1}, \ldots, x^*_n)}{\hat{f}(x^*_{k+1}, \ldots, x^*_n)},$$

where $\hat{f}(x_{k+1}, \ldots, x_n) = \int \cdots \int f(x_1, \ldots, x_n)\, d\sigma_1(x_1) \cdots d\sigma_k(x_k)$.

The result now follows by Theorem 2.2 once it is verified that f_1 and f_2 satisfy condition (2.9) on R^k. This task is straightforward.

In probabilistic terms Theorem 2.3 can be expressed as follows.

THEOREM 4.2. *Let* $\mathbf{X} = (X_1, \ldots, X_n)$ *have a joint* MTP_2 *density. Let* φ *and* ψ *be increasing (or both decreasing) on* R^n. *Then*

$$E\{\varphi(\mathbf{X})\, \psi(\mathbf{X})\} \geqslant (E[\varphi(\mathbf{X})])(E[\psi(\mathbf{X})]). \tag{4.1}$$

Equivalently, $\operatorname{Cov}\{\varphi(\mathbf{X})\, \psi(\mathbf{X})\} \geqslant 0$.

DEFINITION 4.1 (Esary, Proschan, and Walkup [12]). Let $\mathbf{X} = (X_1, \ldots, X_n)$ be a random vector satisfying

$$\operatorname{Cov}\{\varphi(\mathbf{X})\, \psi(\mathbf{X})\} \geqslant 0 \tag{4.2}$$

for any pair of increasing (or decreasing) function φ and ψ. The components of $\mathbf{X}$, $X_1, \ldots, X_n$ are said to be associated.

Thus, existence of a MTP_2 density implies association.

COROLLARY 4.1. *Let f and g be MTP_2 functions with respect to $\boldsymbol{\sigma} = \sigma_1 \times \cdots \times \sigma_n$ on $\mathscr{X}$. Then for any pair of increasing (or decreasing) functions φ and ψ*

$$\begin{aligned}\left(\int g(\mathbf{x})\, f(\mathbf{x})\, d\sigma(\mathbf{x})\right)&\left(\int \varphi(\mathbf{x})\, \psi(\mathbf{x})\; g(\mathbf{x})\, f(\mathbf{x})\, d\sigma(\mathbf{x})\right)\\ &\geqslant \left(\int \varphi(\mathbf{x})\; g(\mathbf{x})\, f(\mathbf{x})\, d\sigma(\mathbf{x})\right)\left(\int \psi(\mathbf{x})\; g(\mathbf{x})\, f(\mathbf{x})\, d\sigma(\mathbf{x})\right).\end{aligned} \tag{4.3}$$

Proof. The product fg is MTP_2 by Proposition 3.3. Therefore, by (4.1)

$$\begin{aligned}&\left(\frac{\int \varphi(\mathbf{x})\, \psi(\mathbf{x})\, f(\mathbf{x})\; g(\mathbf{x})\, d\sigma(\mathbf{x})}{\int f(\mathbf{x})\; g(\mathbf{x})\, d\sigma(\mathbf{x})}\right)\\ &\quad\geqslant \left(\frac{\int \varphi(\mathbf{x})\, f(\mathbf{x})\; g(\mathbf{x})\, d\sigma(\mathbf{x})}{\int f(\mathbf{x})\; g(\mathbf{x})\, d\sigma(\mathbf{x})}\right)\left(\frac{\int \psi(\mathbf{x})\, f(\mathbf{x})\; g(\mathbf{x})\, d\sigma(\mathbf{x})}{\int f(\mathbf{x})\; g(\mathbf{x})\, d\sigma(\mathbf{x})}\right)\end{aligned}$$

and (4.3) ensues.

EXAMPLE 4.1. *Let* $\mathbf{X} = (X_1, \ldots, X_n)$ have a joint MTP_2 density f. Setting

$$g(x_1, \ldots, x_n) = \begin{cases} 1, & a_i \leqslant x_i \leqslant b_i \\ 0 & \text{otherwise}, \end{cases}$$

$i = 1, \ldots, k$, where $k < l < n$ we obtain, in particular,

$$\begin{aligned}P(a_1 &\leqslant X_1 \leqslant b_1, \ldots, a_k \leqslant X_k \leqslant b_k, X_{k+1} \leqslant b_{k+1}, \ldots, X_n \leqslant b_n)\\ &\times P(a_1 \leqslant X_1 \leqslant b_1, \ldots, a_k \leqslant X_k \leqslant b_k)\\ &\geqslant P(a_1 \leqslant X_1 \leqslant b_1, \ldots, a_k \leqslant X_k \leqslant b_k, X_{k+1} \leqslant b_{k+1}, \ldots, X_l \leqslant b_l)\\ &\times P(a_1 \leqslant X_1 \leqslant b_1, \ldots, a_k \leqslant X_k \leqslant b_k, X_{l+1} \leqslant b_{l+1}, \ldots, X_n \leqslant b_n).\end{aligned}$$

Many applications are based on the following:

COROLLARY 4.2. *Let* $\mathbf{X} = (X_1, \ldots, X_n)$ *be associated random variables* *and let $\varphi_1,\ldots,\varphi_k$ be nonnegative functions on R^n all increasing (or all decreasing). Then*

$$E\left[\prod_{i=1}^{k} \varphi_i(\mathbf{X})\right] \geqslant \prod_{i=1}^{k} E|\varphi_i(\mathbf{X})|. \tag{4.4}$$

In particular, if $\varphi_i(\mathbf{X}) = \varphi_i(X_i)$, $i = 1,\ldots, n$, (4.4) becomes

$$E[\varphi_1(X_1) \cdots \varphi_n(X_n)] \geqslant \prod_{i=1}^{n} E[\varphi_i(X_i)]. \tag{4.5}$$

Proof. This follows easily by induction from the definition of association.

Remark 4.1. Specialization of (4.5) yields

$$P(X_1 \geqslant c_1,\ldots, X_n \geqslant c_n) \geqslant \prod_{i=1}^{n} P(X_i \geqslant c_i), \tag{4.6a}$$

$$P(X_1 \leqslant c_1,\ldots, X_n \leqslant c_n) \geqslant \prod_{i=1}^{n} P\{X_i \leqslant c_i\}. \tag{4.6b}$$

Remark 4.2. It is well known (see Esary, Proschan, and Walkup [12]) that the union of independent sets of associated random produces an enlarged set of associated random variables. Clearly increasing functions of associated random variables are again associated. It follows, for example, that if $\mathbf{X} = (X_1,\ldots, X_n)$ and $\mathbf{Y} = (Y_1,\ldots, Y_n)$ are independent vectors each with associated components, then the components of $\mathbf{Z} = (Z_1,\ldots, Z_n) = \mathbf{X} + \mathbf{Y}$ or $\mathbf{W} = \min(\mathbf{X}, \mathbf{Y})$ are associated so that inequalities of the type (4.1) or (4.4) continue to hold.

Thus, *in particular if $\mathbf{X}$ and $\mathbf{Y}$ both have MTP_2 densities, then association of $(Z_1,\ldots, Z_n)$ is retained.* However, $\mathbf{Z}$ need *not* have a joint MTP_2 density. In order to see that the convolution of MTP_2 densities need not be MTP_2 consider two trivariate normal distributions with zero mean vectors and covariance matrices

$$A = \begin{pmatrix} 1 & \rho & \rho \\ \rho & 1 & \rho \\ \rho & \rho & 1 \end{pmatrix} \rho > 0$$

and

$$B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & k \\ 0 & k & k^2+1 \end{pmatrix}.$$

Direct calculations shows that $-A^{-1}$ and $-B^{-1}$ exhibit nonnegative off-diagonal elements and by Example 3.1 the corresponding densities are MTP_2. For large enough k the matrix $-(A+B)^{-1}$ has some negative off-diagonal terms implying that the convolution is not MTP_2. Similarly, it can be shown that the convolution of negative multinomial variables with different sets of parameters (and $r \geqslant 3$) need not be MTP_2. The latter example be discussed in Section 5.

PROPOSITION 4.1. *Let $X_1,\dots,X_n$ be exchangeable random variables having a joint MTP_2 density. Designate*

$$c_m(a) = P(X_1 \leqslant a,\dots,X_m \leqslant a), \qquad m = 1,\dots,n, \quad c_0(a) \equiv 1. \tag{4.7}$$

Then for all $-\infty < a < \infty$

$$c_{m-1}(a)\, c_{m+1}(a) \geqslant c_m^2(a), \qquad m = 1,\dots,n-1. \tag{4.8}$$

Proof. For $\mathbf{x} \in R^n$ define

$$g(\mathbf{x}) = \begin{cases} 1 & \text{if } x_1,\dots,x_{m-1} \leqslant a \\ 0 & \text{otherwise,} \end{cases}$$

$$\varphi(\mathbf{x}) = \begin{cases} 1 & \text{if } x_m \leqslant a \\ 0 & \text{otherwise,} \end{cases}$$

$$\psi(\mathbf{x}) = \begin{cases} 1 & \text{if } x_{m+1} \leqslant a \\ 0 & \text{otherwise.} \end{cases}$$

Since φ and ψ are decreasing and g is MTP_2, Corollary 4.1 applies. Substituting the prescribed functions in (4.3), inequality (4.8) ensues.

COROLLARY 4.3. *Under the conditions of Proposition 4.1, $(c_m(a))^{1/m}$ is decreasing in m.*

EXAMPLE 4.2 (Tong [54], Sidak [51]). Let $\mathbf{X} = (X_1,\dots,X_n) \sim N(0,\Sigma)$ where

$$\Sigma = \begin{pmatrix} 1 & \rho & \cdots & \rho \\ \rho & \ddots & & \rho \\ \vdots & & \ddots & \\ \rho & \cdots & & 1 \end{pmatrix}.$$

and $\rho \geqslant 0$. Then the conditions of Proposition 4.1 hold (see Example 3.1). In particular, $P^{1/m}(X_1 \leqslant a,\dots,X_m \leqslant a)$ is decreasing in m. As a special case, we obtain

$$P(X_1 \leqslant a,..., X_m \leqslant a) \geqslant [P(X_1 \leqslant a)]^m = \prod_{i=1}^{m} P(X_i \leqslant a),$$

which can also be obtained directly on the basis of Corollary 4.2 (see Remark 4.1).

PROPOSITION 4.2. *Let $X_1,..., X_n$ be as in Proposition* 4.1 *and let S be a positive random variable. Denote*

$$d_m(a) = P(X_1/S \leqslant a,..., X_m/S \leqslant a), \qquad m = 1,..., n, \quad d_0(a) = 1.$$

Then

$$d_{m-1}(a)\, d_{m+1}(a) \geqslant d_m^2(a), \qquad m = 1,..., n-1. \tag{4.9}$$

Proof. Let F be the distribution function of S. Then

$$\begin{aligned} d_{m+1}(a) &= \int P(X_1 \leqslant as,..., X_{m+1} \leqslant as)\, dF(s) \\ &= \int c_{m+1}(as)\, dF(s). \end{aligned}$$

Reference to Proposition 4.1 and application of Schwarz's inequality yields

$$\begin{aligned} d_{m+1}(a) &\geqslant \int \frac{c_m^2(as)}{c_{m-1}(as)}\, dF(s) = \int \left(\frac{c_m(as)}{c_{m-1}(as)} \right)^2 c_{m-1}(as)\, dF(s) \\ &\geqslant \left(\int \left(\frac{c_m(as)}{c_{m-1}(as)} \right) c_{m-1}(as)\, dF(s) \right)^2 \left(\int c_{m-1}(as)\, dF(s) \right)^{-1} \\ &= d_m^2(a)\, d_{m-1}^{-1}(a). \end{aligned}$$

EXAMPLE 4.3. Take $X_1,..., X_n$ as in Example 4.2 and let S^2 be an independent χ_ν^2 (chi square) variable. Then $\mathbf{T} = (T_1,..., T_n) = \sqrt{\nu}(X_1/S,..., X_n/S)$ has a multivariate t distribution (Dunnett and Sobel [7]). The inequalities resulting from the monotonicity of $d_m^{1/m}(a)$ were obtained by Tong [54] and Sidak [51].

Remark 4.3. Let $S \sim W_n(\nu, \Sigma)$ ($n \times n$ Wishart matrix with ν degrees of freedom and covariance matrix Σ). It was shown by Karlin and Rinott [28] that a sufficient condition for association (Definition 4.1) of the diagonal elements of S, the terms $S_1,..., S_n$ is the existence of a diagonal matrix D whose diagonal elements are ± 1, such that the off-diagonal elements of

$-D\Sigma^{-1}D$ are nonnegative (see Example 3.2). Under this condition we obtain

$$P\Big(\underset{(\geqslant)}{S_1 \leqslant} a_1,..., \underset{(\geqslant)}{S_n \leqslant} a_n\Big) \geqslant \prod_{i=1}^{n} P\Big(\underset{(\geqslant)}{S_i \leqslant} a_i\Big) \tag{4.10}$$

and if $\mathbf{S} = (S_1, S_2,..., S_n)$ is associated and independent of $\mathbf{X}$ where $\mathbf{X} \sim N(0, \psi)$, then for any covariance matrix ψ

$$P(|X_1|/S_1 \leqslant a_1,..., |X_n|/S_n \leqslant a_n) \geqslant \prod_{i=1}^{n} P(|X_i|/S_i \leqslant a_i). \tag{4.11}$$

Proof of (4.11). Sidak [49] proved for *any* ψ, $\mathbf{X} \sim N(0, \psi)$ implies

$$P(|X_1| \leqslant a_1,..., |X_n| \leqslant a_n) \geqslant \prod_{i=1}^{n} P(|X_i| \leqslant a_i). \tag{4.12}$$

Let $\mathbf{S} = (S_1,..., S_n)$ be associated positive random variables. Then

$$\begin{aligned} P(|X_1| \leqslant a_1 S_1,..., |X_n| \leqslant a_n S_n) &= \int P(|X_1| \leqslant a_1 s_1,..., |X_n| \leqslant a_n s_n)\, dF_{\mathbf{S}}(\mathbf{s}) \\ &\geqslant \int \prod_{i=1}^{n} P(|X_i| \leqslant a_i s_i)\, dF_{\mathbf{S}}(\mathbf{s}) \qquad \text{(by (4.12))} \\ &\geqslant \prod_{i=1}^{n} \int P(|X_i| \leqslant a_i s_i)\, dF_{S_i}(s_i) \qquad \text{(by (4.4)} \end{aligned}$$

since $S_1,..., S_n$ are associated and $P(|X_i| \leqslant a_i s_i)$ are increasing in s_i).

$$= \prod_{i=1}^{n} P(|X_i| \leqslant a_i S_i).$$

Remark 4.4. The inequality

$$P(|X_1| \geqslant a_1,..., |X_n| \geqslant a_n) \geqslant \prod_{i=1}^{n} P(|X_i| \geqslant a_i), \tag{4.13}$$

where $\mathbf{X} \sim N(0, \Sigma)$ is not true for all covariance matrices Σ (Sidak [50]). Conditions for the density of $(|X_1|,..., |X_n|)$ to be MTP_2 were given in Example 3.2. Under these conditions, (4.13) clearly holds and inequalities of the type

$$P(|X_1|/S_1 \geqslant a_1,..., |X_n|/S_n \geqslant a_n) \geqslant \prod_{i=1}^{n} P(|X_i|/S_i \geqslant a_i) \tag{4.14}$$

with $S_1,..., S_n$ as in Remark 4.3 are valid adapting the same arguments. The inequalities (4.13) and (4.14) were first established by Khatri [33] under certain covariance structures.

For a discussion of related inequalities see Sidak [50, 51].

Remark 4.5. One-sided inequalities of the type

$$P\left\{\frac{X_1}{S_1}\leqslant a_1,\dots,\frac{X_n}{S_n}\leqslant a_n\right\}\geqslant\prod_{i=1}^{n}P\left\{\frac{X_i}{S_i}\leqslant a_i\right\} \tag{4.15a}$$

and

$$P\left\{\frac{X_1}{S_1}\geqslant a_1,\dots,\frac{X_n}{S_n}\geqslant a_n\right\}\geqslant\prod_{i=1}^{n}P\left\{\frac{X_i}{S_i}\geqslant a_i\right\} \tag{4.15b}$$

can be derived, but subject to certain restrictions.

We state the result formally.

Let $\mathbf{X}\sim N(0,\Sigma)$ *where* $\sigma_{ij}\geqslant 0$ *for all* i, j *and take* a_i *all of one sign. Assume* $(S_1, S_2,\dots, S_n)$ *are associated positive random variables. Then* (4.15) *holds.* The proof paraphrases the analysis of (4.11) using a well-known inequality of Slepian [53] pertaining to one-sided multinormal probabilities.

5. Multiparameter Negative Multinomial Distribution

The richness, and perhaps complexity, of the subject will be illustrated in this section by a more detailed discussion of the multiparameter negative multinomial distribution. Related problems arise in the multinomial case and will not be discussed here. Some results in this latter case appear in Karlin and Rinott [29].

Let

$$\mathbf{X}^{(\nu)}=(X_1^{(\nu)},\dots,X_n^{(\nu)}) \tag{5.1}$$

be independent vectors, where

$$P(X_1^{(\nu)}=k_1,\dots,X_n^{(\nu)}=k_n)=\frac{(k_1+\cdots+k_n)!}{\prod_{i=1}^{n}k_i!}\,p_{(\nu)0}\prod_{i=1}^{n}p_{(\nu)i}^{k_i} \tag{5.2}$$

for k_i nonnegative integers, $i=1,\dots,n$, and $\sum_{i=0}^{n}p_{(\nu)i}=1$, $\nu=1,\dots,r$.

Note that (5.2) coincides with (3.9) for $N=1$ and then (5.1) is MTP_2. Observe also that the N-fold convolution of the densities of (5.2) with $p_{(\nu)i}=p_i$, $i=0,\dots,n$, independent of ν generates the negative multinomial distribution (3.9). Now consider the random vector

$$\mathbf{Y}=(Y_1,\dots,Y_n)=\sum_{\nu=1}^{r}\mathbf{X}^{(\nu)} \tag{5.3}$$

for the multiparameter context of (5.1) and (5.2).

If the parameter sets $\mathbf{p}_{(\nu)} = (p_{(\nu)0},..., p_{(\nu)n})$ are not coincident with respect to $\nu = 1,..., r$, the distribution of $\mathbf{Y}$ is now a convolution of nonidentical densities of type (3.9). In general, $\mathbf{Y}$ does not have a MTP_2 density, as will be indicated below. However, recall that the MTP_2 property of $\mathbf{X}^{(\nu)}$ implies association of all the components of $\mathbf{X}^{(1)},..., \mathbf{X}^{(r)}$ and hence of the sums $Y_1,..., Y_n$. (See Remark 4.2) Therefore, inequalities of the nature of

$$E[\varphi(\mathbf{Y})\, \psi(\mathbf{Y})] \geqslant E[\varphi(\mathbf{Y})]\, E[\psi(\mathbf{Y})] \tag{5.4}$$

when φ and ψ are monotone in the same direction hold.

The integral representation given in the next theorem, of value in itself, provides some access in assessing the nature of MTP_2 property.

THEOREM 5.1.

$$\begin{aligned} P(Y_1 = k_1,..., Y_n = k_n) &= \left(\prod_{\nu=0}^{r} p_{(\nu)0}\right) \frac{1}{\prod_{i=1}^{n} k_i!} \int_0^\infty \cdots \int_0^\infty \left(\sum_{\nu=1}^{r} p_{(\nu)1} t_\nu\right)^{k_1} \cdots \left(\sum_{\nu=1}^{r} p_{(\nu)n} t_\nu\right)^{k_n} \\ &\quad \times \exp\left\{-\sum_{\nu=1}^{r} t_\nu\right\} dt_1 \cdots dt_r. \end{aligned} \tag{5.5}$$

Remark 5.1. This integral representation expresses a set of multiparameter negative multinomial probabilities as a compounding of a multiparameter multinomial with a generalized multivariate gamma density. This generalizes the formula which calculates the standard univariate negative binomial distribution as a mixture of the binomial with a gamma density.

Proof of (5.5). First note that

$$\begin{aligned} P(Y_1 = k_1,..., Y_n = k_n) &= \left(\prod_{\nu=1}^{r} p_{(\nu)0}\right) \sum \left\{ \frac{\Gamma(l_{(1)1} + \cdots + l_{(1)n} + 1)}{\Gamma(l_{(1)l} + 1) \cdots \Gamma(l_{(1)n} + 1)} \right. \\ &\quad \left. \times \left(\prod_{i=1}^{n} p_{(1)i}^{l_{(1)i}}\right) \cdots \frac{\Gamma(l_{(r)1} + \cdots + l_{(r)n} + 1)}{\Gamma(l_{(r)1} + 1) \cdots \Gamma(l_{(r)n} + 1)} \left(\prod_{i=1}^{n} p_{(r)i}^{l_{(r)i}}\right) \right\}, \end{aligned} \tag{5.6}$$

where the sum extends over the collection of all integers $l_{(\nu)i}$, $\nu = 1,..., r$, $i = 1,..., n$, satisfying

$$\sum_{\nu=1}^{r} l_{(\nu)1} = k_1,..., \sum_{\nu=1}^{r} l_{(\nu)n} = k_n. \tag{5.7}$$

Expanding each term $(\sum_{\nu=1}^{r} p_{(\nu)i} t_\nu)^{k_i}$ gives

$$\int_0^\infty \cdots \int_0^\infty \left(\sum_{\nu=1}^{r} p_{(\nu)1} t_\nu\right)^{k_1} \cdots \left(\sum_{\nu=1}^{r} p_{(\nu)n} t_\nu\right)^{k_n} \exp\left\{-\sum_{\nu=1}^{r} t_\nu\right\} dt_1 \cdots dt_r$$

$$= \sum \int_0^\infty \cdots \int_0^\infty \left(\frac{k_1!}{l_{(1)1}! \cdots l_{(r)1}!}\right)\left(\prod_{\nu=1}^{r} p_{(\nu)1}^{l_{(\nu)1}}\right)\left(\prod_{\nu=1}^{r} t_\nu^{l_{(\nu)1}}\right) \cdots \left(\frac{k_n!}{l_{(1)n}! \cdots l_{(r)n}!}\right)$$

$$\times \left(\prod_{\nu=1}^{r} p_{(\nu)n}^{l_{(\nu)n}}\right)\left(\prod_{\nu=1}^{r} t_\nu^{l_{(\nu)n}}\right) \exp\left\{-\sum_{j=1}^{r} t_j\right\} dt_1 \cdots dt_r, \tag{5.8}$$

where the sum is governed by prescriptions (5.7). In view of (5.6) Eq. (5.5) results after rearranging terms in (5.8) and recognition of the gamma function

$$\int_0^\infty t^{l_{(\nu)1} + \cdots + l_{(\nu)n}} e^{-t}\, dt = \Gamma(l_{(\nu)1} + \cdots + l_{(\nu)n} + 1).$$

In representation (5.5) (assume for definiteness $r \geqslant n$) consider the linear change of variables

$$z_i = \sum_{\nu=1}^{r} p_{(\nu)i} t_\nu, \qquad i = 1, 2,..., n,$$

appended by a judicious choice of other linear combinations $z_i = \sum_{\nu=1}^{r} b_{(\nu)i} t_\nu$ such that the augmented $r \times r$ matrix $\tilde{P} = \|P, B\|$ is of full rank where $P = \| p_{(\nu)i} \|_{\nu=1,i=1}^{r\ \ n}$ and $B = \| b_{(\nu)i} \|_{\nu=1,i=n+1}^{r\ \ r}$.

Then the change of variables $z_i = \sum_{\nu=1}^{r} \tilde{p}_{(\nu)i} t$, $i = 1, 2,..., n$, in (5.8) yields

$$P(Y_1 = k_1,..., Y_n = k_n) \tag{5.9}$$

$$= \frac{c}{\prod_{i=1}^{n} k_i!} \int_0^\infty \cdots \int_0^\infty z_1^{k_1} \cdots z_n^{k_n} \exp\left\{-\sum_{i=1}^{r} a_i z_i\right\} \psi(z_1,..., z_r)\, dz_1 \cdots dz_n,$$

where $c > 0$ and $\psi(z_1,..., z_r) = 1$ on the conical set

$$L = \{\mathbf{z} = (z_1,..., z_r) = \mathbf{t}\tilde{P} : \mathbf{t} = (t_1,..., t_r) \geqslant 0\}, \tag{5.10}$$

and zero otherwise. The function ψ is MTP_2 if and only if L is a lattice. When ψ is MTP_2 then the integrand in (5.9) is MTP_2 and hence, by Proposition 3.2, $P\{Y_1 = k_1,..., Y_n = k_n\}$ is MTP_2 in $\mathbf{k}$.

The characterization of matrices $\tilde{P}$ for which the image of the positive orthant L of (5.10) is a lattice and of parameter matrices P which can be appropriately augmented to such $\tilde{P}$ is still open. Several examples can be constructed satisfying the requirements that L is a lattice. This problem and further examples will be elaborated in Karlin and Rinott [29].

We now show that $P(Y_1 = k_1, \ldots, Y_n = k_n)$ need not always be MTP_2. Let $n = 3$, $r = 2$, and $p_{(1)0} = 1/6$, $p_{(1)1} = 1/2$, $p_{(1)2} = 0$, $p_{(1)3} = 1/3$, $p_{(2)0} = 1/6$, $p_{(2)1} = 0$, $p_{(2)2} = 1/2$, and $p_{(2)3} = 1/3$, then by direct calculation

$$\begin{aligned} &P(X_1 = 0, X_2 = 0, X_3 = k_3)\, P(X_1 = 1, X_2 = 1, X_3 = k_3) \\ &\quad < P(X_1 = 0, X_2 = 1, X_3 = k_3)\, P(X_1 = 1, X_2 = 0, X_3 = k_3), \end{aligned} \tag{5.11}$$

indicating that $P(Y_1 = k_1, Y_2 = k_2, Y_2 = k_3)$ is not TP_2 in the variables k_1 and k_2.

We close this section with another formula for calculating the probabilities $P\{\mathbf{Y} = \mathbf{k}\}$ normalized differently. Specifically, we present an integral representation of the quantities

$$\frac{\prod_{i=1}^{n} k_i!}{(k_1 + \cdots + k_n + r)} P\{Y_1 = k_1, Y_2 = k_2, \ldots, Y_n = k_n\} = Q_{\mathbf{k}} \tag{5.12}$$

generalizing a formula of Hardy, Littlewood, Pólya [16, p. 164] of the univariate case.

We state this formally.

THEOREM 5.2. *With $Q_{\mathbf{k}}$ defined in* (5.12), *we have*

$$\begin{aligned} Q_{\mathbf{k}} = \prod_{\nu=1}^{r} p_{(\nu)0} \int \cdots \int \left(\sum_{\nu=1}^{r} p_{(\nu)1} s_\nu \right)^{k_1} \\ \times \left(\sum_{\nu=1}^{r} p_{(\nu)2} s_\nu \right)^{k_2} \cdots \left(\sum_{\nu=1}^{r} p_{(\nu)n} s_\nu \right)^{k_n} ds_1 \cdots ds_{r-1}, \end{aligned} \tag{5.13}$$

where $s_r = 1 - \sum_{\nu=1}^{r-1} s_\nu$ and the region of integration is $s_\nu \geqslant 0$, $\nu = 1, \ldots, r-1$, $\sum_{\nu=1}^{r-1} s_\nu \leqslant 1$.

Formula (5.13) can be regarded as a compounding of a *multiparameter multinomial* with the *Dirichlet density*.

The proof of (5.13) is smilar to that of Theorem 5.1 relying on the formula for the multivariate beta function.

Formula (5.13) can be viewed presenting $Q_{\mathbf{k}}$ as a multidimensional moment sequence. It follows on the basis of (5.13) that $\log[Q_{\mathbf{k}}]$ is strictly convex with respect to $\mathbf{k}$. This fact and related matters will be discussed more extensively elsewhere.

6. SOME MTP_2 PROPERTIES OF STOCHASTIC PROCESSES

In this section we briefly discuss the relevance of MTP_2 functions in the context of monotone stochastic processes. Characterizations of monotone processes and corresponding limit theorems were investigated in recent papers, e.g., see Harris [17], Kamae, Krengel, and O'Brien [22], Kirstein, Franken, and Stoyan [34].

Let $\{\mathbf{X}_t,\ t \geqslant 0\}$ be a stationary Markov process whose state space is a lattice $\mathscr{X} = \mathscr{X}_1 \times \cdots \times \mathscr{X}_n$ where $\mathscr{X}_i$ are totally ordered sets. Let $p(t, \mathbf{x}, \mathbf{y})$, $t \geqslant 0$, $\mathbf{x}, \mathbf{y} \in \mathscr{X}$ be the transition density of the process, with respect to a product measure on $\mathscr{X}$, denoted by dy. Set $p(t, \mathbf{x}, A) = \int_A p(t, \mathbf{x}, \mathbf{y})\, dy$. Given a distribution Q on $\mathscr{X}$, let $U_t Q(A) = \int_{\mathscr{X}} p(t, \mathbf{x}, A)\, Q(d\mathbf{x})$, that is, $U_t Q$ is the distribution of $\mathbf{X}_t$ where $\mathbf{X}_0 \sim Q$. Let $\mathscr{F}$ denote the class of increasing functions defined on $\mathscr{X}$. The following problems are of interest.

(i) Let p_1, p_2 be transition densities and let Q_1, Q_2 be distributions on $\mathscr{X}$. Find conditions on p_1, p_2 and Q_1, Q_2 such that $Q_2 \succ^{\mathscr{F}} Q_1$ (see (1.2)) entails

$$U_t^{(2)} Q_2 \succ^{\mathscr{F}} U_t^{(1)} Q_1, \tag{6.1}$$

where $U_t^{(i)} Q(A) = \int_{\mathscr{X}} p_i(t, \mathbf{x}, A)\, Q(d\mathbf{x})$. That is, when is the stochastic ordering at the initial time maintained for all later times t?

(ii) Determine conditions under which

$$U_{t_2} Q \succ^{\mathscr{F}} U_{t_1} Q \qquad \text{for} \quad t_2 \geqslant t_1, \tag{6.2}$$

i.e., the process is stochastically increasing with time.

(iii) Suppose $\mathbf{X}_0 \sim Q$ satisfies

$$E\{\varphi(\mathbf{X}_0)\, \psi(\mathbf{X}_0)\} \geqslant (E\{\varphi(\mathbf{X}_0)\})(E\{\psi(\mathbf{X}_0)\}) \tag{6.3}$$

for any φ, $\psi \in \mathscr{F}$ (the property of association, see (1.8)). Let $\mathbf{X}_t \sim U_t Q$. Find conditions under which the association property is preserved as the process evolves. (For a complete characterization with a finite state space, see Harris [17].)

Pertaining to problem (i) the following proposition comes out directly from Theorem 2.4:

PROPOSITION 6.1. *Let Q_1 and Q_2 have density functions q_1 and q_2, respectively, and suppose for a fixed t that*

$$q_2(\mathbf{x}) \succ_{TP_2} q_1(\mathbf{x}) \qquad (\textit{see}\ (1.18\text{a}))$$

and (6.4)

$$p_2(t, \mathbf{x}, \mathbf{y}) \succ_{TP_2} p_1(t, \mathbf{x}, \mathbf{y}) \qquad \textit{in} \quad (\mathbf{x}, \mathbf{y}).$$

Then $U_t^{(2)} Q_2 \succ_{TP_2} U_t^{(1)} Q_2 \succ_{TP_2} U_t^{(1)} Q_1$ *and* (6.1) *follows.* (In the last comparison $\succ_{\mathrm{TP}_2}$ refers to the densities of the distributions $U_t^{(i)} Q_j$.)

Concerning problem (iii), we have

PROPOSITION 6.2. *Let $\mathbf{X}_0 \sim Q$, and suppose $\mathbf{X}_0$ is associated (i.e., (6.3) holds), and $p(t, \mathbf{x}', \mathbf{y}) \succ_{TP_2} p(t, \mathbf{x}, \mathbf{y})$ as a function of $\mathbf{y}$ for all $\mathbf{x}' \geqslant \mathbf{x}$ (t fixed). Then $\mathbf{X}_t$ is associated.*

Proof. Observe that under the conditions prescribed $p(t, \mathbf{x}, \mathbf{y})$ is MTP_2 in $\mathbf{y}$. Therefore by Theorem 2.3

$$\int \varphi(\mathbf{y})\, \psi(\mathbf{y})\, p(t, \mathbf{x}, \mathbf{y})\, d\mathbf{y} \geqslant \left(\int \varphi(\mathbf{y})\, p(t, \mathbf{x}, \mathbf{y})\, d\mathbf{y}\right)\left(\int \psi(\mathbf{y})\, p(t, \mathbf{x}, \mathbf{y})\, d\mathbf{y}\right) \tag{6.5}$$

for φ, $\psi \in \mathscr{F}$. Theorem 2.2 affirms for $\varphi \in \mathscr{F}$ that $\int \varphi(\mathbf{y})\, p(t, \mathbf{x}, \mathbf{y})\, d\mathbf{y}$ is increasing in $\mathbf{x}$. Therefore, with φ, $\psi \in \mathscr{F}$, we have

$$\begin{aligned}
&E[\varphi(\mathbf{X}_t)\, \psi(\mathbf{X}_t)] \\
&\quad = \int \varphi(\mathbf{y})\, \psi(\mathbf{y}) \left\{\int p(t, \mathbf{x}, \mathbf{y})\, Q(d\mathbf{x})\right\} d\mathbf{y} \\
&\quad = \int \left\{\int \varphi(\mathbf{y})\, \psi(\mathbf{y})\, p(t, \mathbf{x}, \mathbf{y})\, d\mathbf{y}\right\} Q(d\mathbf{x}) \\
&\quad \geqslant \int \left(\int \varphi(\mathbf{y})\, p(t, \mathbf{x}, \mathbf{y})\, d\mathbf{y}\right)\left(\int \psi(\mathbf{y})\, p(t, \mathbf{x}, \mathbf{y})\, d\mathbf{y}\right) Q(d\mathbf{x}) \qquad \text{(by (6.5))} \\
&\quad \geqslant \left(\iint \varphi(\mathbf{y})\, p(t, \mathbf{x}, \mathbf{y})\, d\mathbf{y}\, Q(d\mathbf{x})\right)\left(\iint \psi(\mathbf{y})\, p(t, \mathbf{x}, \mathbf{y})\, d\mathbf{y}\, Q(d\mathbf{x})\right) \\
&\qquad\qquad \text{(by association)} \\
&\quad = (E[\varphi(\mathbf{X}_t)])(E[\psi(\mathbf{X}_t)]).
\end{aligned}$$

The principal known examples of monotone processes occur with a one-dimensional state space. We present some examples (cf. Karlin [23]).

(1) Suppose P_{ij} is a TP_2 Markov transition matrix (for $i, j = 0, 1, 2, \ldots$). Then the n step transition density $P_{0j}^{(n)}$ is TP_2 in the time variable n and state variable j [loc. cit.]. Accordingly, $X_{n+1} \succ_{\mathrm{TP}_2} X_n$, where $X_0 = 0$ and X_n has the density of $\{P_{0j}^{(n)}\}$, $j = 0, 1, 2, \ldots$. The same conclusion obtains for a Markov TP_2 transition kernel $p(x, y)$ on $x, y \geqslant 0$. Then defining recursively $p_n(x, z) = \int p(x, y)\, p_{n-1}(y, z)\, dy$ there evolves the monotone stochastic ordering $X_{n+1} \succ X_n$ where X_n has the density $p_n(0, y)$ provided $X_0 = 0$. Analogous results obtain for diffusion processes on the line and birth–death processes since the transition density is always TP_2 in these cases.

Other examples in this vein involve absorption and first passage probability comparisons for TP_2 processes (see Karlin [23]).

(2) Let $f(\xi)$ be a symmetric PF_∞ density, i.e., $f(\xi)$ has a Laplace transform

$$\int_{-\infty}^{\infty} f(\xi)\, e^{-s\xi}\, d\xi = \frac{e^{\gamma s^2}}{\prod_{i=1}^{\infty} (1 - a_i^2 s^2)}$$

with $\gamma \geqslant 0$, $a_i^2 \geqslant 0$, and $0 < \sum_{i=0}^{\infty} a_i^2 + \gamma < \infty$. Let $X_1, X_2 \cdots$ be i.i.d. random variables distributed according to $f(\xi)$. Form the process $(X_1 + X_2 + \cdots + X_n) = S_n$. Then S_n is a symmetrically stochastically increasing process, meaning that for any increasing $\varphi(|x|)$ in $|x|$, we have

$$\int \varphi(|\xi|)\, f_{S_n}(\xi)\, d\xi \leqslant \int \varphi(|\xi|)\, f_{S_{n+1}}(\xi)\, d\xi,$$

where $f_{S_n}(\xi)$ is the density function of S_n.

(3) The direct product of marginal TP_2 processes with transition density $P(\mathbf{x}, \mathbf{y}) = \prod_{i=1}^{n} P_i(x_i, y_i)$ is MTP_2. The process governed by the kernel $P(\mathbf{x}, \mathbf{y})$ inherits the properties of Proposition 6.2.

Further results and examples of MTP_2 and monotone processes will be discussed elsewhere.

REFERENCES

[1] AHLSWEDE, R. AND DAYKIN, D. E. (1979). An inequality for weights of two families of sets, their unions and intersections. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, **93** 183–185.

[2] ALFSEN, E. M. (1971). *Compact Convex Sets and Boundary Integrals*. Springer-Verlag, Berlin New York.

[3] ABDEL-HAMEED, M. AND SAMPSON, A. R. (1978). Positive dependence of the bivariate and trivariate absolute normal, t, χ^2 and F distributions. *Ann. Statist.* **6** 1360–1368.

[4] BARLOW, R. E., AND PROSCHAN, F. (1975). Statistical Theory of Reliability and Life Testing. Holt, Rinehart Winston, New York.

[5] BORELL, C. (1974). Convex measures on locally convex spaces. *Ark. Mat.* **12** 239–252.

[6] DAS GUPTA, S., EATON, M. L., OLKIN, I., PERLMAN, M. D., SAVAGE, L. J., AND SOBEL, M. (1972). Inequalities on the probability content of convex regions for eliptically contoured distributions. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability*, Vol. II, pp. 241–265. Univ. of California Press, Berkeley.

[7] DUNNETT, C. W., AND SOBEL, M. (1954). A bivariate generalization of student's t distribution with tables for certain cases. *Biometrika* **41** 153–169.

[8] DYKSTRA, R. L. (1979). Product inequalities involving the multivariate normal distribution. Techn. Report No. 85, Dept. of Statist., University of Missouri–Columbia.

[9] DYKSTRA, R. L., AND HEWETT, J. E. (1978). Positive dependence of the roots of a Wishart matrix. *Ann. Statist.* **6** 235–238.

[10] ESARY, J. D., AND PROSCHAN, F. (1968). Generating associated random variables, Boeing Scientific Research Laboratories, Document D1-82-0696, Seattle, Wash.

[11] ESARY, J. D., AND PROSCHAN, F. (1972). Relationships among some notions of bivariate dependence. *Ann. Math. Statist.* **43** 651–655.

[12] ESARY, J. D., PROSCHAN, F., AND WALKUP, D. (1967). Association of random variables, with applications. *Ann. Math. Statist.* **38** 1466–1474.

[13] FINNEY, D. J. (1941). The joint distribution of variance ratios based on a common error mean square. *Ann. Eigencies* **11** 136–140.

[14] FORTUIN, C. M., GINIBRE, J., AND KASTELEYN, P. W. (1971). Correlation inequalities on some partially ordered sets. *Comm. Math. Phys.* **22** 89–103.

[15] GUMBEL, E. J., (1961). Bivariate logistic distributions. *J. Amer. Statist. Assoc.* **56** 335–349.

[16] HARDY, G. H., LITTLEWOOD, J. E., AND PÓLYA, G. (1934). *Inequalities*. Cambridge Univ. Press, London/New York.

[17] Harris, T. E. (1977). A correlation inequality for Markov processes in partially ordered state spaces. *Ann. Probability* **5** 451–454.

[18] HOLLEY, R. (1974). Remarks on the FKG inequalities. *Comm. Math. Phys.* **36** 227–231.

[19] JOGDEO, K. (1977). Association and probability inequalities, *Ann. Statist.* **5** 495–504.

[20] JOHNSON, N. L., AND KOTZ, S. (1972). *Distributions in Statistics: Continuous Multivariate Distributions.* Wiley, New York.

[21] KAMAE, T., AND KRENGEL, U. (1978). Stochastic partial ordering. *Ann. Probability* **6** 1044–1049.

[22] KAMAE, T., KRENGEL, U., AND O'BREIN, G. L., (1977). Stochastic inequalities on partially ordered spaces. *Ann. Probability* **5** 899–912.

[23] KARLIN, S. (1964). Total positivity, absorption probabilities and applications. *Trans. Amer. Math. Soc.* **111** 33–107.

[24] KARLIN, S. (1968). *Total Positivity.* Stanford Univ. Press, Stanford, Calif.

[25] KARLIN, S. (1974). Inequalities for symmetric sampling plans I. *Ann. Statist.* **2** 1065–1094.

[26] KARLIN, S., AND MCGREGOR, J. L. (1959). Coincidence probabilities. *Pacific J. Math.* **9** 1141–1164.

[27] KARLIN, S., AND RINOTT, Y. (1980). Classes of orderings of measures and related correlation inequalities. II. Multivariate reverse rule distribution. *J. Multivariate Anal.* **10** 499–516.

[28] KARLIN, S. AND RINOTT, Y. (1979). Total positivity properties of absolute value multinormal variables with applications to confidence interval estimates and related probabilistic inequalities. To appear.

[29] KARLIN, S., AND RINOTT, Y. (1980). Multivariate monotonicity, generalized convexity and ordering relations, with applications in analysis and statistics. To appear.

[30] KARLIN, S., AND TRUAX, D. R. (1960). Slippage problems. *Ann. Math. Statist.* **31** 296–323.

[31] KEMPERMAN, J. H. B. (1973). Moment problems for sampling without replacement I, II, III. *Indag. Math.* **35** 149–188.

[32] KEMPERMAN, J. H. B. (1977). On the FKG inequality for measures on a partially ordered space. *Indag. Math.* **39** 313–331.

[33] KHATRI, G. G. (1967). On certain inequalities for normal distributions and their applications to simulatneous confidence bounds. *Ann. Math. Statist.* **38** 1853–1867.

[34] KIRSTEIN, B. M., FRANKEN, P., AND STOYAN, D. (1977). Comparability and monotonicity of Markov processes. *Theor. Probability Appl.* **12** 40–52.

[35] KRISHNAIAH, P. R. (1979). Some developments on simultaneous test procedures. In *Developments in Statistics*, Vol. 2 (P. R. Krishnaiah, Ed.). Academic Press, New York.

[36] LEHAMNN, E. L. (1966). Some concepts of dependence. *Ann. Math. Statist.* **37** 1137–1153.

[37] LORENTZ, G. G. (1953). An inequality for rearrangement. *Amer. Math. Monthly* **60** 176–179.

[38] MARSHALL, A. W., AND OLKIN, I. (1979). "Inequalities: Theory of Majorization and Its Applications," Academic Press, New York.

[39] MEYER, P. A. (1966). *Probability and Potentials*. Ginn/Blaisdell, Waltham, Mass.

[40] PERLMAN, M. D., AND OLKIN, I. (1978). Unbiasedness of invariant tests for MANOVA and other multivariate problems, submitted for publication.

[41] PHELPS, R. R. (1966). *Lectures on Choquet's Theorem*, Van Nostrand Mathematical Studies No. 7. Van Nostrand, Princeton, N. J.

[42] PRÉKOPA, A. (1973). On logarithmic concave measures and functions. *Acta. Sci. Math. (Szeged)* **34** 335–343.

[43] PRESTON, C. J. (1974). A generalization of the KFG inequalities. *Comm. Math. Phys.* **36** 223–241.

[44] PRESTON, C. J. (1976). *Random Fields*, Lecture Notes in Mathematics No. 434, Springer-Verlag, New York.

[45] RAMABHADRAN, V. R. (1951). A multivariate gamma-type distribution, *Sankhya* **11** 45–46.

[46] RINOTT, Y. (1973). Multivariate majorization and rearrangement inequalities with some applications to probability and statistics. *Israel J. Math.* **15** 60–77.

[47] RINOTT, Y., AND SANTNER, T. J. (1977). An inequality for multivariate normal probabilities with applications to a design problem. *Ann. Statist.* **5** 1228–1234.

[48] SARKAR, T. K. (1969). Some lower bounds of realiability. Tech. Report, No. 124, Dept. of Operations Research and Statistics, Stanford University.

[49] SIDAK, Z. (1967). Rectangular confidence regions for the means of multivariate normal distributions. *J. Amer. Statist. Assoc.* **68** 734–739.

[50] SIDAK, Z. (1971). On probabilities of ractangles in multivariate student distributions: Their dependence on correlations. *Ann. Math Statist.* **42** 169–175.

[51] SIDAK, Z. (1973). A chain of inequalities for some types of multivariate distributions, with nine special cases. *Apl. Mat.* **18** 110–118.

[52] STRASSEN, V. (1965). The existence of probability measures with given arginals. *Ann. Math. Statist.* **36** 423–439.

[53] SLEPIAN, D. (1962). On one-sided barrier problem for Gaussian noise. *Bell System Tech. J.* **41** 463–501.

[54] TONG, Y. L. (1970). Some probability inequalities of multivariate normal and multivariate *t*. *J. Amer. Statist. Assoc.* **65** 1243–1247.