The Annals of Statistics
1977, Vol. 5, No. 3, 505-515

A CONCEPT OF POSITIVE DEPENDENCE FOR EXCHANGEABLE RANDOM VARIABLES[1]

BY MOSHE SHAKED

University of New Mexico

An n-variate distribution function is said to be positive dependent by mixture (PDM) if it is a mixture of independent n-variate distributions with equal marginals. PDM distributions arise in various contexts of reliability and other areas of statistics. We give a necessary and sufficient condition, by means of independent random variables, for an n-variate distribution function to be PDM. The distributions and the expectations of the order statistics of PDM and of independent n-variate distributions which have the same marginals, are compared and the results applied to obtain bounds for the reliability of certain "k out of n" systems. A characterization of vectors of expectations of order statistics of PDM distribution is shown. Surprisingly many exchangeable distributions are found to be PDM. We prove a closure property of the class of PDM distributions and list some examples.

1. Introduction. In this paper we consider n-variate distribution functions (df's) which admit the representation

$$F(x_1, \cdots, x_n) = \int_\Omega \prod_{i=1}^n F^{(\omega)}(x_i)\, d\tau(\omega) \tag{1.1}$$

where $\{F^{(\omega)}, \omega \in \Omega\}$ is a family of univariate df's, Ω is a subset of a finite dimensional Euclidean space and τ is a df on Ω. Such df's, which are mixtures of independent n-variate df's with equal marginals, will be called *positive dependent by mixture*. A random vector will be called PDM if its df is PDM.

By definition PDM df's are exchangeable (that is $F(x_1, \cdots, x_n) = F(x_{\Pi(1)}, \cdots, x_{\Pi(n)})$ for every permutation Π of the integers $1, 2, \cdots, n$) and all the marginals are equal to

$$\tilde{F}(x) = \int_\Omega F^{(\omega)}(x)\, d\tau(\omega)\ . \tag{1.2}$$

Note that the family of n-variate PDM df's contains the joint df's of n i.i.d. random variables and also contains the other extremity: the df's whose mass is concentrated on the line $x_1 = x_2 = \cdots = x_n$. We do not consider the more general family of df's that admit the representation $\int_\Omega \prod_{i=1}^n F_i^{(\omega)}(x_i)\, d\tau(\omega)$ because every n-variate df admits such a representation.

PDM df's arise in a variety of circumstances. Tong (1970, 1977) and Šidák (1973) observed that in some cases the joint df's of test statistics of some dependent hypotheses are PDM. PDM df's have been suggested as a useful class

Received June 1975; revised September 1976.

[1] This work was partially supported by National Science Foundation Grant GP-30707X1. It is a part of the author's Ph.D. dissertation prepared at the University of Rochester.

AMS 1970 *subject classifications*. Primary 62G30, 62E10; Secondary 62N05.

Key words and phases. Mixtures, random environment, de Finetti's theorem, positive dependence, order statistics, majorization.

Institute of Mathematical Statistics is collaborating with JSTOR to digitize, preserve, and extend access to *The Annals of Statistics.* www.jstor.org

of priors by Ericson (1969a, b), Lindley and Smith (1972) and Shaked (1975b). A well-known theorem of de Finetti (1937), which was generalized by Hewitt and Savage (1955) (see also Olshen (1974)), states that a necessary and sufficient condition for n random variables (rv's) to be embedded in an infinite sequence of exchangeable rv's is that they are PDM. Various inequalities which are satisfied by PDM df's are discussed in Dykstra et al. (1973) and some closure properties of the family of PDM df's are given in Shaked (1974, 1975a).

Of particular interest for our applications are PDM df's that arise in reliability theory. If $X_1, \cdots, X_n$ are the lifelengths of n identical components of a complex system which operates in a random environment, and if, given the "environment" of the system (i.e., the user of the system, the weather conditions, etc.), $X_1, \cdots, X_n$ are i.i.d. then $(X_1, \cdots, X_n)$ is a PDM vector. Often system life distributions are computed under the assumption that the components lifelengths are independent. In Section 2 we show that in some circumstances it is possible to determine whether under or over estimates result from the assumption of independence when in fact the component lifelengths are PDM.

2. Inequalities for the df's of order statistics of PDM rv's. Let $(X_1, \cdots, X_n)$ be a PDM random vector with the df F. Let $\tilde{F}$ be the marginal df of X_i. In this section we compare the df's and the expectations of the order statistics $X_{(1)}, \cdots, X_{(n)}$ to those of i.i.d. rv's $Y_1, \cdots, Y_n$ having $\tilde{F}$ as their common df. As usual, denote their order statistics by $Y_{(1)}, \cdots, Y_{(n)}$. These comparisons enable one to determine whether over (or under) estimates occur in various probabilistic computations when a set of rv's is PDM but one acts as if the rv's are independent. Exact expressions for the df of subsets of order statistics when the joint distribution of $(X_1, \cdots, X_n)$ is known can be found in Maurer and Margolin (1976).

2.1. *Comparison of the tails of the df's of* $X_{(k)}$ *and* $Y_{(k)}$. The following results indicate that the df of $X_{(k)}$, $1 \leqq k \leqq n$, has in many useful cases heavier tails than the df of $Y_{(k)}$. Theorem 2.1 is useful when there is a positive mass on the endpoints of $S(\tilde{F})$—support of $\tilde{F}$ (this is the case, e.g., if the X's are discrete rv's with support which is bounded from below and/or from above). Corollaries 2.1 and 2.2 are useful when there is not a positive mass on the endpoints of $S(\tilde{F})$ (e.g., when $F_{X_{(k)}}$ is absolutely continuous).

To introduce the notation of Theorem 2.1 recall (see, e.g., Barlow and Proschan (1965), pages 216–217) that for every two integers k and n ($2 \leqq k \leqq n-1$, $n \geqq 2$) the function

$$(2.1) \qquad h_{k,n}(p) \equiv \sum_{i=k}^{n} \binom{n}{i} p^i (1-p)^{n-i} = \binom{n}{k} k \int_0^p u^{k-1}(1-u)^{n-k}\, du$$

is increasing on $[0, 1]$, is convex on $[0, p_0]$ and concave on $[p_0, 1]$ where $p_0 = (k-1)/(n-1)$ (see Figure 2.1). Let $\tilde{\tilde{p}}$ be the (unique) point in $(0, 1)$ which maximizes $(1 - h(p))/(1-p)$ (we will not specify the subscipts k and n when it is obvious or unnecessary). Similarly let $\tilde{p}$ be the (unique) point in $(0, 1)$ that

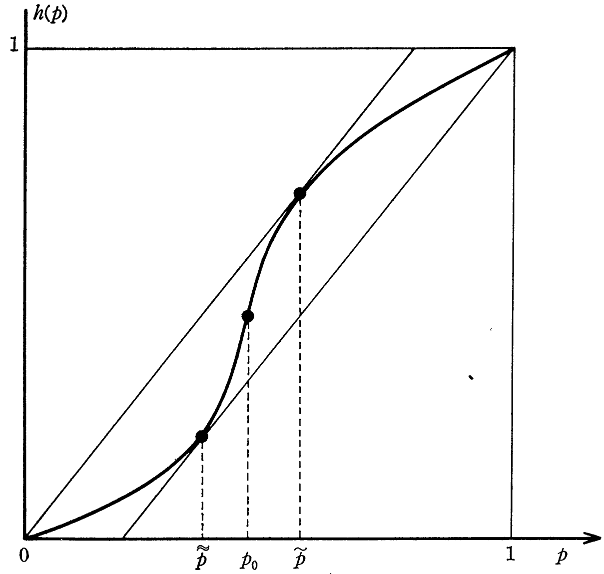

FIG. 2.1. The function $h(p)$.

that maximizes $h(p)/p$. For $k = n (n \geqq 2)$ let $\tilde{p} = \tilde{\tilde{p}} = 1$ and for $k = 1$ let $\tilde{p} = \tilde{\tilde{p}} = 0$.

THEOREM 2.1. *Let x be a real number.* (a) *If* $\tilde{F}(x) \leqq \tilde{\tilde{p}}$ *then* $F_{X_{(k)}}(x) \geqq F_{Y_{(k)}}(x)$. (b) *If* $\tilde{F}(x-) \geqq \tilde{p}$ *then* $F_{X_{(k)}}(x-) \leqq F_{Y_{(k)}}(x-)$.

The proof is given in the Appendix.

Of particular interest are the cases when $x = L \equiv \inf S(\tilde{F})$ and $x = R \equiv \sup S(\tilde{F})$. Thus if $\tilde{F}(L) \leqq \tilde{\tilde{p}}$ then $F_{X_{(k)}}(L) \geqq F_{Y_{(k)}}(L)$ and if $\tilde{F}(R-) \geqq \tilde{p}$ then $F_{X_{(k)}}(R-) \leqq F_{Y_{(k)}}(R-)$. Note that for some $\tilde{F}$'s there does not exist an x such that $0 < \tilde{F}(x) \leqq \tilde{\tilde{p}}$; then it may happen that $\tilde{F}_{X_{(k)}}(L) < \tilde{F}_{Y_{(k)}}(L)$, as the following example (in which $L = 0$) shows. Consider the following df's:

$$F^{(1)}(x) = \begin{cases} 0 & x < 0 \\ 0.4 & 0 \leqq x < 1 \\ 1 & 1 \leqq x \end{cases} \quad \text{and} \quad F^{(2)}(x) = \begin{cases} 0 & x < 0 \\ 0.7 & 0 \leqq x < 1 \\ 1 & 1 \leqq x . \end{cases}$$

Then $F(x_1, x_2, x_3) = \frac{2}{3}F^{(1)}(x_1)F^{(1)}(x_2)F^{(1)}(x_3) + \frac{1}{3}F^{(2)}(x_1)F^{(2)}(x_2)F^{(2)}(x_3)$ is PDM according to formula (1.1). Taking $k = 2$ ($n = 3$) one can verify that $\tilde{F}_{X_{(k)}}(L) < \tilde{F}_{Y_{(k)}}(L)$. A similar remark applies also to the right endpoint.

The following corollaries are useful for comparison of the tails when there is no mass on the endpoints. We exclude from Corollaries 2.1 and 2.2 the cases $k = 1$ and $k = n$. It is easy to verify that $F_{X_{(1)}}[F_{X_{(n)}}]$ is stochastically larger [smaller] than $F_{Y_{(1)}}[F_{Y_{(n)}}]$ and the comparison of the tails is trivial.

COROLLARY 2.1. *Let* $2 \leqq k \leqq n-1$. (a) *If there exists a real number* $c > -\infty$ *such that* $F_{X_{(k)}}(x) \neq F_{Y_{(k)}}(x)$ *on* $(-\infty, c]$ *and* $F_{X_{(k)}}$ *and* $F_{Y_{(k)}}$ *are continuous on*

$(-\infty, c]$ *then*

$$F_{X_{(k)}}(x) > F_{Y_{(k)}}(x)\,, \qquad x \leqq c\,. \tag{2.2}$$

(b) *If there exists a real number* $d < \infty$ *such that* $F_{X_{(k)}}(x) \neq F_{Y_{(k)}}(x)$ *on* $[d, \infty)$ *and* $F_{X_{(k)}}$ *and* $F_{Y_{(k)}}$ *are continuous on* $[d, \infty)$ *then*

$$F_{X_{(k)}}(x) < F_{Y_{(k)}}(x)\,, \qquad x \geqq d\,. \tag{2.3}$$

PROOF. To obtain (2.2) choose $x_0 \leqq c$ such that $\tilde{F}(x_0) \leqq \tilde{\tilde{p}}$ so by Theorem 2.1 $F_{X_{(k)}}(x_0) \geqq F_{Y_{(k)}}(x_0)$. But, by assumption, $F_{X_{(k)}}(x_0) \neq F_{Y_{(k)}}(x_0)$, so (2.2) follows for x_0, and hence for all $x \leqq c$ by continuity and the assumption that $F_{X_{(k)}}(x) \neq F_{Y_{(k)}}(x)$ for all $x \leqq c$. Part (b) is proved similarly. □

COROLLARY 2.2. *Let* $2 \leqq k \leqq n-1$. (a) *If* $F_{X_{(k)}}$ *and* $F_{Y_{(k)}}$ *are absolutely continuous and* $F_{X_{(k)}}(x) \neq F_{Y_{(k)}}(x)$ *on* $(-\infty, c] \cap S(\tilde{F})$ *for some* $c > -\infty$ *then* (2.2) *holds.* (b) *If* $F_{X_{(k)}}$ *and* $F_{Y_{(k)}}$ *are absolutely continuous and* $F_{X_{(k)}}(x) \neq F_{Y_{(k)}}(x)$ *on* $[d, \infty) \cap S(F)$ *for some* $d < \infty$ *then* (2.3) *holds.*

PROOF. By making a $1-1$ order preserving transformation of the support of $\tilde{F}$ onto the extended real line that takes L into $-\infty$ we reduce the conditions of this corollary to the conditions of Corollary 2.1 and the result follows. □

We remark that if τ of representation (1.1) gives mass to at most two points, then a stronger result than Theorem 2.1 and Corollary 2.1 can be achieved. In this case $F_{X_{(k)}}(x)$ crosses $F_{Y_{(k)}}(x)$ at most once and if it does it crosses from above (Shaked (1975a)).

2.2. *An application in reliability theory.* A typical application of the previous results is the following. Consider a "k out of n" system (that is, a system which functions if and only if at least k of its n components function) with identical components, which operates in a random environment (see discussion in Section 1). If $X_1, \cdots, X_n$ are the component lifelengths then the system reliability $\Psi(t) = P(X_{(n-k+1)} > t)$ is often approximated by assuming the X_i's are independent. If $\hat{\Psi}(t)$ is such an approximation then (recalling from Section 1 that $(X_1, \cdots, X_n)$ is PDM) there exist t_1 and t_0 such that for $t \geqq t_1$, $\Psi(t) \geqq \hat{\Psi}(t)$ and for $t \leqq t_0$, $\Psi(t) \leqq \hat{\Psi}(t)$.

2.3. *Comparison of the closeness of* df'*s of order statistics.* The next theorem shows that the df's of $X_{(k)}$, $k = 1, \cdots, n$ are "closer to each other" than the df's of $Y_{(k)}$, $k = 1, \cdots, n$. In its corollary we obtain the intuitively clear result that the vector $(EX_{(1)}, \cdots, EX_{(n)})$ is "more homogeneous" than $(EY_{(1)}, \cdots, EY_{(n)})$. The following definition enables us to compare the homogeneity of two vectors (for more details see, e.g., Hardy, Littlewood and Pólya (1952), page 45, or Marshall and Olkin (in preparation)). A vector $(a_1, \cdots, a_n)$ is said to *majorize* a vector $(b_1, \cdots, b_n)$ written $(a_1, \cdots, a_n) \succ (b_1, \cdots, b_n)$ if, after the components have been ordered such that $a_1 \geqq a_2 \geqq \cdots \geqq a_n$, $b_1 \geqq b_2 \geqq \cdots \geqq b_n$, the relations

$$\textstyle\sum_{i=1}^{k} a_i \geqq \sum_{i=1}^{k} b_i\,, \qquad k = 1, 2, \cdots, n-1 \tag{2.4.i}$$

and

(2.4.ii) $$\sum_{i=1}^n a_i = \sum_{i=1}^n b_i$$

prevail.

THEOREM 2.2. *Let* $X_1, \cdots, X_n$ *and* $Y_1, \cdots, Y_n$ *be, respectively,* PDM *and* i.i.d. rv's *having the same univariate marginal* $\tilde{F}$. *Then for every* $x \in R$

(2.5) $$(F_{X_{(1)}}(x), \cdots, F_{X_{(n)}}(x)) \prec (F_{Y_{(1)}}(x), \cdots, F_{Y_{(n)}}(x)) .$$

The proof is given in the Appendix.

COROLLARY 2.3. *Let* $X_1, \cdots, X_n$ *and* $Y_1, \cdots, Y_n$ *be as in Theorem* 2.2, *then for every measurable function* g *which is monotonic on the support of* X_1

(2.6) $$(Eg(X_{(1)}), \cdots, Eg(X_{(n)})) \prec (Eg(Y_{(1)}), \cdots, Eg(Y_{(n)})) .$$

PROOF. Using the fact that for every rv Z with df F_Z, $EZ = \int_0^\infty (1 - F_Z(x))\, dx - \int_{-\infty}^0 F_Z(x)\, dx$ and using Theorem 2.2 it is easy to see that

(2.7) $$\sum_{i=1}^k EX_{(i)} \geqq \sum_{i=1}^k EY_{(i)} , \qquad k = 1, 2, \cdots, n .$$

Since $\sum_{i=1}^n EX_{(i)} = \sum_{i=1}^n EY_{(i)}$, (2.7) implies $\sum_{i=n-j}^n EX_{(i)} \leqq \sum_{i=n-j}^n EY_{(i)}$, $j = 0, 1, \cdots, n-2$. Noting that $EX_{(n)} \geqq \cdots \geqq EX_{(1)}$ and $EY_{(n)} \geqq \cdots \geqq EY_{(1)}$, one obtains

(2.8) $$(EX_{(1)}, \cdots, EX_{(n)}) \prec (EY_{(1)}, \cdots, EY_{(n)}) .$$

This proves (2.6) for $g(x) = x$. Denote $X_i' = g(X_i)$ and $Y_i' = g(Y_i)$, $i = 1, 2, \cdots, n$. Note that $(X_1', \cdots, X_n')$ is a PDM vector and that Y_i', $i = 1, \cdots, n$ are i.i.d. rv's having the same common df as X_1'. If g is nondecreasing then

(2.9) $$X'_{(i)} = g(X_{(i)}) \quad \text{and} \quad Y'_{(i)} = g(Y_{(i)}) , \qquad i = 1, 2, \cdots, n$$

and (2.6) is obtained from (2.8) and (2.9). If g is nonincreasing then

(2.10) $$X'_{(i)} = g(X_{(n-i+1)}) \quad \text{and} \quad Y'_{(i)} = g(Y_{(n-i+1)}) , \qquad i = 1, 2, \cdots, n$$

and (2.6) is obtained from (2.8) and (2.10). □

A particular case of Corollary 2.3 is $(EX_{(1)}^{2k+1}, \cdots, EX_{(1)}^{2k+1}) \prec (EY_{(1)}^{2k+1}, \cdots, EY_{(n)}^{2k+1})$, $k = 0, 1, 2, \cdots$. If X_1 is a nonnegative rv then $(EX_{(1)}^k, \cdots, EX_{(n)}^k) \prec (EY_{(1)}^k, \cdots, EY_{(n)}^k)$, $k = 0, 1, 2, \cdots$.

3. A characterization of vectors of expectations of order statistics of PDM random vectors. Recently Kadane (1971, 1974) and Mallows (1973) have found necessary and sufficient conditions for a given vector to belong to $M_I(n)$, where $M_I(n) = \{(EY_{(1)}, \cdots, EY_{(n)}); Y_i, i = 1, 2, \cdots, n \text{ are i.i.d}\}$. Mallows applies the results to obtain bounds on df's and Kadane applies them to show that a particular structure for the expectations of the order statistics is satisfied only by a degenerate df.

It may be of interest to characterize vectors in $M_D(n) \equiv \{(EX_{(1)}, \cdots, EX_{(n)}); (X_1, \cdots, X_n) \text{ is PDM}\}$. We do it in the next theorem. The proof was suggested by J. H. B. Kemperman.

THEOREM 3.1. $M_D(n) = M_I(n)$.

PROOF. From (1.1) it is easy to see that every vector of $M_D(n)$ is a convex combination of vectors in $M_I(n)$, hence the statement of the theorem is equivalent to the assertion that $M_I(n)$ is convex. To see that $M_I(n)$ is convex consider the set $A = \{T: T(y) = \sup\{x: F(x) < y\}, 0 < y < 1$, for some univariate df $F\}$ of monotone nondecreasing and right continuous functions on (0, 1). Clearly A is convex, hence from the proof of Theorem 1 of Kadane (1974) it is easy to see that $M_I(n)$ is convex. □

Theorem 3.1 shows that when we are given a vector $\mathbf{a} \in M_D(n)$, it is possible to find a PDM vector $(Y_1, \cdots, Y_n)$ such that $a_i = EY_{(i)}$, $i = 1, \cdots, n$ and $\text{corr}(Y_i, Y_j) = 0$ when $i \neq j$ (in fact $Y_1, \cdots, Y_n$ can be assumed to be i.i.d.). One can ask whether $\mathbf{a}$ determines an upper bound for $\text{corr}(X_i, X_j)$ where $(X_1, \cdots, X_n)$ is a PDM vector such that $a_i = EX_{(i)}$. The answer is negative, provided $EY_i^2 < \infty$. To see this define $X_i = Y_i + W$, $i = 1, 2, \cdots, n$ where W is a rv independent of $Y_1, \cdots, Y_n$ with $EW = 0$ and $\text{Var } W > \varepsilon^{-1}(1 - \varepsilon) \text{ Var } Y_1$. Then $(X_1, \cdots, X_n)$ is PDM and it is easily verified that $\text{corr}(X_i, X_j) > 1 - \varepsilon$ and $EX_{(k)} = EY_{(k)}$.

4. Examples and applications. A useful way to identify PDM vectors is given by the following proposition, the proof of which is omitted.

PROPOSITION 4.1. *The n-variate* df F *is* PDM *if and only if there exist* i.i.d. rv's U_i $(i = 1, 2, \cdots, n)$, *a random vector* W *independent of the* U_i*'s and a Borel measurable function* g *such that*

$$X_i = g(U_i, W)\,, \qquad i = 1, 2, \cdots, n$$

have joint df F.

Thus, equicorrelated normals with nonnegative correlation, some exchangeable multivariate exponential (Marshall–Olkin (1967)), some exchangeable multivariate geometric (Esary–Marshall (1974)) and exchangeable multivariate F distributions (discussed by Hewett and Bulgren (1971)) are PDM. From representation (1.1) it can be seen that the multivariate logistic (Malik and Abrahams (1973)) and the exchangeable Tallis' (1962) distributions are PDM. Shaked (1975b) has shown that some of the Johnson–Kotz (1975) distributions are PDM.

Jensen (1969) defines a bivariate χ^2 df in the following way: Let $\mathbf{Y}_1$ and $\mathbf{Y}_2$ be two $m \times 1$ vectors such that

$$\begin{pmatrix} \mathbf{Y}_1 \\ \mathbf{Y}_2 \end{pmatrix} \sim N\left(\begin{pmatrix} \mathbf{0} \\ \mathbf{0} \end{pmatrix}, \begin{pmatrix} \Psi_{11} & \Psi_{12} \\ \Psi_{12} & \Psi_{22} \end{pmatrix}\right)$$

where $\Psi_{11} = \sigma_1^2 I_m$ and $\Psi_{22} = \sigma_2^2 I_m$, $\sigma_1 > 0$, $\sigma_2 > 0$, and Ψ_{12} is an $m \times m$ matrix such that

$$\begin{pmatrix} \Psi_{11} & \Psi_{12} \\ \Psi_{12} & \Psi_{22} \end{pmatrix}$$

is a positive semidefinite matrix. Then $X_i \equiv \mathbf{Y}_i'\mathbf{Y}_i/2\sigma_i^2$, $i = 1, 2$, are said to have bivariate chi-square df which m degrees of freedom. We proceed now to show that (X_1, X_2) is PDM. First note that without loss of generality we can assume $\sigma_1^2 = \sigma_2^2 = 1$ because the df of X_i is independent of σ_i^2, $i = 1, 2$. Next observe that we can assume

$$\Psi_{12} = \operatorname{diag}(\rho_1, \cdots, \rho_m), \qquad \rho_j \geqq 0, j = 1, 2, \cdots, m, \tag{4.1}$$

otherwise there exists orthogonal matrices M and N such that $M\Psi_{12}N' = R = \operatorname{diag}(r_1, \cdots, r_m)$, $r_j \geqq 0, j = 1, 2, \cdots, m$, and then the transformation

$$\begin{pmatrix} \tilde{\mathbf{Y}}_1 \\ \tilde{\mathbf{Y}}_2 \end{pmatrix} = \begin{pmatrix} M & 0 \\ 0 & N \end{pmatrix} \begin{pmatrix} \mathbf{Y}_1 \\ \mathbf{Y}_2 \end{pmatrix}$$

has $2m$-variate normal df with zero mean and covariance matrix

$$\begin{pmatrix} I_m & R \\ R & I_m \end{pmatrix}$$

and $\frac{1}{2}\tilde{\mathbf{Y}}_i'\tilde{\mathbf{Y}}_i = \frac{1}{2}\mathbf{Y}_i'\mathbf{Y}_i = X_i$, $i = 1, 2$,.

Now let Y_{ij} be the jth element of $\mathbf{Y}_i$ ($i = 1, 2, j = 1, 2, \cdots, m$) then (when $\sigma_1 = \sigma_2 = 1$) $X_i = \frac{1}{2}\sum_{j=1}^m Y_{ij}^2$, $i = 1, 2$, and by (4.1) (Y_{1j}, Y_{2j}), $j = 1, 2, \cdots, m$ are independent, each has a bivariate normal df with zero mean unit variance and a nonnegative correlation, and (Y_{i1}, Y_{i2}) is PDM. It can be shown now by Proposition 4.1 that (X_1, X_2) is PDM.

For more examples, we refer the reader to Šidák (1973) and Jensen (1971) which discuss some bivariate df's, all of which are shown to be PDM in Shaked (1974).

PDM distributions can be identified also by the following:

PROPOSITION 4.2. *The limit in distribution of* PDM df'*s is a* PDM df.

PROOF. Let $\bar{R} = [-\infty, \infty]$ with the usual topology so $\bar{R}$ is a compact metric space and hence $\bar{R}^n$ is a compact metric space. The separable Banach space of bounded continuous functions of $\bar{R}^n$ with the sup-norm is denoted by $C(\bar{R}^n)$. As is well known, $C^*(\bar{R}^n)$, the dual space of $C(\bar{R}^n)$, is isomorphic to E, the space of all bounded regular countable additive measures on $\bar{R}^n$. Equip C^* with the usual weak*-topology and E with the (corresponding by isomorphism) weak topology. For $A \subseteq E$ denote by $\hat{A}$ the corresponding set in C^*.

Consider

$$Z = \{P^n \mid P \text{ is a probability measure on } \bar{R}\}.$$

Clearly $\hat{Z}$ is w^*-compact. Let $\overline{CO}(\hat{Z})[\overline{CO}(Z)]$ denote the closed (w^*) [(weak)] convex hull of $\hat{Z}[Z]$. From Proposition 1.2 of Phelps (1966) and the isomorphism, each $\alpha \in \overline{CO}(Z)$ has the representation

$$\alpha = \int_Z P^n \tau(P^n) \tag{4.2}$$

where τ is a probability measure on Z.

By definition, α corresponds to a PDM distribution if and only if (4.2) holds and $\alpha(R^n) = 1$. Clearly, $\alpha(R^n) = 1$ if, and only if, $\tau\{P^n \mid P^n(R^n) = 1\} = 1$.

If $\alpha_m \to_w \alpha$ (α_m is converging in distribution to α) where α is a probability measure on R^n and each $\alpha_m \in \overline{CO}(Z)$, then $\alpha \in \overline{CO}(Z)$ so it has the representation (4.2). Since $\alpha(R^n) = 1$ by assumption, α corresponds to a PDM distribution. □

Some applications of Proposition 4.2 are discussed below.

Let $\mathbf{X}_m = (X_{m1}, \cdots, X_{mn})$, $m = 1, 2, \cdots$ be independent random vectors with PDM distributions F_m. It is easy to verify that the random vector $(h_m(X_{11}, \cdots, X_{m1}), \cdots, h_m(X_{1n}, \cdots, X_{mn}))$ has a PDM distribution G_m, say, for every Borel measurable function h_m, $m = 1, 2, \cdots$. Thus by Proposition 4.2 the limiting distribution G, say, is PDM.

Of particular interest is the case where $F_1 = F_2 = \cdots$ and $h_m(X_1, \cdots, X_m) = b_m(\sum_{i=1}^m X_i) + a_m$ for some a_m and $b_m > 0$. Then if a limiting distribution G exists it is a multivariate PDM distribution with stable marginals. Similarly by taking $h_m(X_1, \cdots, X_m) = b_m \max(X_1, \cdots, X_m) + a_m$ a PDM distribution with extreme value marginals is obtained. This fact is of interest in contrast to a result of Campbell and Tsokos (1973) which gives a general form for limits of extremes of bivariate random vectors. Their result is not restricted to exchangeable distributions, but it applies only to smooth (φ^2-bounded) distributions. Distributions with singular part along the main diagonal are (in general) not φ^2-bounded (Lancaster (1958)) but they can be PDM. Then the Campbell and Tsokos result cannot be applied but the fact that G is PDM can be used to obtain bounds for G.

Another possible application of Proposition 4.2 is the following. Multivariate distributions with specified parametric marginals are sometimes defined as limits in analogy to the univariate case (e.g., bivariate Poisson is defined as a limit of bivariate binomial). Thus, when the defining distributions are PDM it follows that the defined distribution is PDM.

APPENDIX

Proof of theorems. The following lemma is used in the proof of Theorem 2.1.

LEMMA A.1. *Let G be a* df *on the unit interval. If $\int_0^1 p\, dG(p) \leqq \tilde{\tilde{p}}$ ($\geqq \tilde{p}$), then for $1 \leqq k \leqq n$,*

$$\int_0^1 h_{k,n}(p)\, dG(p) \geqq (\leqq)\, h_{k,n}\left(\int_0^1 p\, dG(p)\right), \tag{A.1}$$

where $h_{k,n}$ is defined in (2.1) *and $\tilde{p}$ and $\tilde{\tilde{p}}$ are described in Figure* 2.1.

PROOF. First note that if $q \leqq \tilde{\tilde{p}}$ then (omitting the subscripts of h), since the tangent to h at $(q, h(q))$ is is under h along $[0, 1]$ the following inequality holds:

$$h(p) \geqq h(q) + h'(q)(p - q), \qquad 0 \leqq p \leqq 1. \tag{A.2}$$

Denote $q = \int_0^1 p\, dG(p)$ and assume $q \leqq \tilde{\tilde{p}}$. Integrating (A.2) with respect to $dG(p)$ we obtain $\int_0^1 h(p)\, dG(p) \geqq h(\int_0^1 p\, dG(p))$. Similarly if $q \geqq \tilde{p}$ we obtain the second inequality of (A.1). □

PROOF OF THEOREM 2.1. By representation (1.1)

$$F_{X_{(k)}}(x) = \int_\Omega h(F^{(\omega)}(x))\, d\tau(\omega) \tag{A.3}$$

and

$$F_{Y_{(k)}}(x) = h(\tilde{F}(L)) = h(\int_\Omega F^{(\omega)}(x))\, d\tau(\omega)\,. \tag{A.4}$$

Substituting $p = F^{(\omega)}(x))$ in (A.3) we get $F_{X_{(k)}}(x) = \int_0^1 h(p)\, dG(p)$ where $G = G(\tau, F^{(\cdot)}(x))$ is a df on the unit interval. Similarly $\tilde{F}(x) = \int_0^1 p dG(p)$. As $\tilde{F}(x) \leqq \tilde{\tilde{p}}$ we can apply Lemma A.1 to complete the proof of (a). The proof of (b) is similar. □

The following lemma is used in the proof of Theorem 2.2.

LEMMA A.2. *Let n and k be integers*, $1 \leqq k \leqq n$, *then*

$$\tilde{\tilde{h}}_{k,n}(p) \equiv \sum_{i=1}^k h_{i,n}(p) = \sum_{i=1}^k \sum_{j=i}^n \binom{n}{j} p^j(1-p)^{n-j}$$

is concave on $[0, 1]$.

PROOF. The function $\tilde{\tilde{h}}_{n,n}(p) = np$ is clearly concave on $[0, 1]$. For $1 \leqq k \leqq n-1$, $\tilde{\tilde{h}}'_{k,n}(p) = n(1 - h_{k,n-1}(p))$, but $h_{k,n-1}(p)$ increases on $[0, 1]$, hence $\tilde{\tilde{h}}_{k,n}$ is concave. □

PROOF OF THEOREM 2.2. For $1 \leqq i \leqq n$, $F_{X_{(i)}}(x) = \int_\Omega h_{i,n}(F^{(\omega)}(x))\, d\tau(\omega)$ by (1.1). Hence for $1 \leqq k \leqq n$

$$\begin{aligned}\sum_{i=1}^k F_{X_{(i)}}(x) &= \int_\Omega \tilde{\tilde{h}}_{k,n}(F^{(\omega)}(x))\, d\tau(\omega)\\ &\leqq \tilde{\tilde{h}}_{k,n}(\int_\Omega F^{(\omega)}(x)\, d\tau(\omega)) \quad \text{by Lemma A.2 and Jensen's inequality}\\ &= \sum_{i=1}^k F_{Y_{(i)}}(x)\,.\end{aligned}$$

Also

$$\sum_{i=1}^n F_{X_{(i)}}(x) = n\tilde{F}(x) = \sum_{i=1}^n F_{Y_{(i)}}(x)$$

and the proof is complete. □

Acknowledgments. It is my greatest pleasure to thank A. W. Marshall for his constant advice and his will to devote unlimited time in improving this work. I also thank Z. Šidák for very useful comments on an earlier version of this paper; and a referee and J. H. B. Kemperman for suggesting, respectively, the simple proofs of Corollary 2.1 and Theorem 3.1.

REFERENCES

[1] BARLOW, R. and PROSCHAN, F. (1965). *Mathematical Theory of Reliability*. Wiley, New York.

[2] CAMPBELL, J. W. and TSOKOS, C. P. (1973). The asymptotic distribution of maxima in bivariate samples. *J. Amer. Statist. Assoc.* **68** 734–739.

[3] DE FINETTI, B. (1937). La prévision: ses lois logiques, ses sources subjectives. *Ann. Inst. H. Poincaré* (*Paris*), **7** 1–68. (English translation in *Studies in Subjective Probability* ((1964), H. E. Kyburg and H. E. Smokler, eds.). Wiley, New York.)

[4] DYKSTRA, R. L., HEWETT, J. E. and THOMPSON, W. A. (1973). Events which are almost independent. *Ann. Statist.* **1** 674-681.

[5] ERICSON, W. A. (1969a). Subjective Bayesian models in sampling finite populations (with discussion). *J. Roy. Statist. Soc. Ser. B* **31** 195-233.

[6] ERICSON, W. A. (1969b). Subjective Bayesian models in sampling finite populations: stratification. In *New Developments in Survey Sampling* (N. L. Johnson and H. Smith, eds.) 326-357. Wiley, New York.

[7] ESARY, J. D. and MARSHALL, A. W. (1974). Multivariate geometric distributions generated by a cumulative damage process. Report NPS55EY73041A, Naval Postgraduate School, Monterey, California.

[8] HARDY, G. H., LITTLEWOOD, J. E. and PÓLYA, G. (1952). *Inequalities.* Cambridge Univ. Press.

[9] HEWETT, J. E. and BULGREN, W. G. (1971). Inequalities for some multivariate F-distributions with applications. *Technometrics* **13** 397-402.

[10] HEWITT, E. and SAVAGE, L. J. (1955). Symmetric measures on Cartesian products. *Trans. Amer. Math. Soc.* **80** 470-501.

[11] JENSEN, D. R. (1969). An inequality for a class of bivariate chi-square distributions. *J. Amer. Statist. Assoc.* **64** 333-336.

[12] JENSEN, D. R. (1971). A note of positive dependence and the structure of bivariate distributions. *SIAM J. Appl. Math.* **20** 749-753.

[13] JOHNSON, N .L. and KOTZ, S. (1975). On some generalized Farlie-Gumbel-Morgenstern distributions. *Comm. Statist.* **4** 415-427.

[14] KADANE, J. B. (1971). A moment problem for order statistics. *Ann. Math. Statist.* **42** 745-751. Correction, *Ann. Statist.* **1** 797.

[15] KADANE, J. B. (1974). A characterization of triangular arrays which are expectations of order statistics. *J. Appl. Probability* **11** 413-416.

[16] LANCASTER, H. O. (1958). The structure of bivariate distributions. *Ann. Math. Statist.* **29** 719-736.

[17] LINDLEY, D. V. and SMITH, A. F. M. (1972). Bayes estimates for the linear model (with discussion). *J. Roy. Statist Soc. Ser. B* **34** 1-41.

[18] MALIK, H. J. and ABRAHAM, B. (1973). Multivariate logistic distributions. *Ann. Statist.* **1** 588-590.

[19] MALLOWS, C. L. (1973). Bounds on distribution functions in terms of expectations of order statistics. *Ann. Probability* **1** 297-303.

[20] MARSHALL, A. W. and OLKIN, I. (1967). A multivariate exponential distribution. *J. Amer. Statist. Assoc.* **62** 30-44.

[21] MARSHALL, A. W. and OLKIN, I. *Inequalities via Majorization, with Applications to Matrix Theory, Combinatorics and Probability.* In preparation.

[22] MAURER, W. and MARGOLIN, B. H. (1976). The multivariate inclusion-exclusion formula and order statistics from dependent variables. *Ann. Statist.* **4** 1190-1199.

[23] OLSHEN, R. (1974). A note on exchangeable sequences. *A. Wahrasch. Geb.* **28** 317-321.

[24] PHELPS, R.R. (1966). *Lectures on Choquet's Theorem.* Van Nostrand, New York.

[25] SHAKED, M. (1974). Some concepts of positive dependence for exchangeable bivariate distributions. Technical report, Department of Statistics, Univ. of Rochester.

[26] SHAKED, M. (1975a). On concepts of dependence for multivariate distributions. Ph.D. dissertation, Univ. of Rochester.

[27] SHAKED, M. (1975b). A note on the exchangeable generalized Farlie-Gumbel-Morgenstern distributions. *Comm. Statist.* **4** 711-721.

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

[29] TALLIS, G. M. (1962). The use of a generalized multinomial distribution in the estimation of correlation in discrete data. *J. Roy. Statist. Soc. Ser. B* **24** 530-534.

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

[31] TONG, Y. L. (1977) An ordering theorem for the moments of non-negative random variables with applications. *Ann. Statist.* **5** 274–277.

DEPARTMENT OF MATHEMATICS AND STATISTICS
UNIVERSITY OF NEW MEXICO
ALBUQUERQUE, NEW MEXICO 87131