The Annals of Probability
1982, Vol. 10, No. 3, 765–772

SOME CONCEPTS OF NEGATIVE DEPENDENCE

BY HENRY W. BLOCK[1], THOMAS H. SAVITS[2], AND MOSHE SHAKED[3]

University of Pittsburgh, University of Pittsburgh and Indiana University

The theory of positive dependence notions cannot yield useful results for some widely used distributions such as the multinomial, Dirichlet and the multivariate hypergeometric. Some conditions of negative dependence that are satisfied by these distributions and which have practical meaning are introduced. Useful inequalities for some widely used distributions are obtained.

1. Introduction. Concepts of positive dependence of sets of random variables (rv's) have received a lot of attention recently. Their study was found to yield a better understanding of the structure of some widely used multivariate distribution functions (df's). In addition to this, various useful inequalities were obtained with applications in many areas of probability and statistics. Barlow and Proschan (1975), Chapter 5, include a review of most of the work done prior to 1972. A list of more recent references can be found in Block and Ting (1981).

On the other hand, notions of negative dependence have received very little attention in the literature. Some negative dependence analogs of positive dependence concepts have been mentioned by some authors (Lehmann, 1966; Brindley and Thompson, 1972; Dykstra, Hewett and Thompson, 1973; and Shaked, 1977, among others). In the bivariate setting the random vector (T_1, T_2) is usually said to satisfy some negative dependence condition if $(T_1, -T_2)$ satisfies the analogous positive dependence condition. Many of the negative dependence results of Lehmann (1966) were obtained in this way. Mallows (1968) examined some properties of the multinomial distribution and Jogdeo and Patil (1975) extended Mallows' technique to other specific distributions. In this paper we attempt a systematic study of negatively dependent distributions.

While the first draft of this paper was being written, two related works were brought to our attention. The first work by Ebrahimi and Ghosh (1980) discusses some negative dependence analogs of well known positive dependence concepts. Some of our definitions overlap those of Ebrahimi and Ghosh (1980); however, our main results differ from theirs. The second related paper is by Karlin and Rinott (1980). They introduce a negative dependence notion which is closely related to one of ours and they obtain some results which are similar to ours. Basically, we have one condition which is stronger than theirs but is easier to check and a second condition which is weaker than theirs (but of a much simpler form) which yields many of their results.

The main motivation for our definitions is to try to formulate the intuitive requirement that if a set of negatively dependent random variables is split into two subsets in some manner, then one subset will tend to be "large" when the other subset is "small" and vice versa. Our condition N introduced in Section 4 accomplishes this. In Section 2 we define some conditions to be discussed. We derive some properties in Section 3 and show that condition N implies the other properties in Section 4. Examples are given in Section 5.

In the following, "increasing" stands for "nondecreasing" and "decreasing" for "nonincreasing". Vectors in $\mathbb{R}^n$ are denoted by $\mathbf{t} = (t_1, \cdots, t_n)$ and $\mathbf{t} \le \mathbf{t}'$ means $t_i \le t_i'$, i

Received July 1980; revised May 1981.

[1] Supported by ONR Contract N00014-76-C-0839.

[2] Supported by NSF Grant MCS77-01458 and ONR Contract N00014-76-C-0839.

[3] Supported by NSF Grant MCS79-27150.

AMS 1980 *subject classification*. Primary 62H20, 62H05, 62E10; secondary 60E05.

Key words and phrases. Multinomial, Dirichlet, multivariate hypergeometric, multivariate normal, PF_2 densities, total positivity, negative and positive dependence, inequalities.

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

$= 1, \cdots, n$. Similarly $\mathbf{t} < \mathbf{t}'$ means $t_i < t_i'$, $i = 1, \cdots, n$, and $\mathbf{0} = (0, \cdots, 0)$. A real function on $\mathbb{R}^n$ will be called increasing if it is increasing in each variable when the other variables are held fixed.

A bivariate function $K(\cdot, \cdot)$ which is defined on $S_1 \times S_2$ (where S_1 and S_2 are subsets of $\mathbb{R}$) is said to be *totally positive of order 2* (TP_2) on $S_1 \times S_2$ if $K(x, y) \geq 0$ and if

$$K(x, y)K(x', y') \geq K(x, y')K(x', y) \quad \text{whenever } x \leq x', y \leq y' \tag{1.1}$$

(see Karlin, 1968). The function K is said to be *reverse regular of order 2* (RR_2) on $S_1 \times S_2$ if $K(x, y) \geq 0$ and if

$$K(x, y)K(x', y') \leq K(x, y')K(x', y) \quad \text{whenever } x \leq x', y \leq y', \tag{1.2}$$

(see Karlin (1968), page 12).

2. Negative Dependence Concepts.

One of the fundamental problems of positive dependence has been to obtain conditions on a multivariate random vector $\mathbf{T} = (T_1, \cdots, T_n)$ such that the condition (or conditions similar to this)

$$P(T_1 > t_1, \cdots, T_n > t_n) \geq \prod_{i=1}^{n} P(T_i > t_i)$$

holds for all real t_i. A particular problem has been to find conditions on the covariance structure of the multivariate normal distribution so that conditions of the above type hold. In reliability theory the problem of obtaining upper and lower bounds on joint probabilities of dependent lifetimes in terms of the marginal lifetimes has also led to investigations of problems of this type. A sufficient condition that the above inequality hold is that the random vector have a joint density (discrete or continuous) which is TP_2 in pairs (see Barlow and Proschan (1975) for a simple proof of this).

In many cases, such as the case of the multinomial distribution, the opposite inequality holds (Mallows, 1968). For distributions like the multinomial where correlations are nonpositive it is important to have checkable conditions which imply the reverse inequality

$$P(T_1 > t_1, \cdots, T_n > t_n) \leq \prod_{i=1}^{n} P(T_i > t_i)$$

for all real t_i. A natural condition is to assume the reverse TP_2 condition, i.e. that the joint density is RR_2 in pairs. Unfortunately, unlike the TP_2 case, the marginal distributions need not have the same property. A simple 3×2 discrete example suffices to show this. In fact it is known that Theorem 5.1, page 173 of Karlin (1968) does not apply for RR_2 kernals. It should also be remarked here that even in the TP_2 case, one must make some assumptions on the nature of the set $\{f > 0\}$ before one can use the above Theorem 5.1 to conclude that the marginal densities are also TP_2 in pairs: see Kemperman (1977). A possible alternative then is to assume that not only f, but all its marginal densities, are RR_2 in pairs. But even under this assumption we have not been able to show that our weakest condition (see Definition 2.2) is consequently satisfied. Ebrahimi and Ghosh (1980) claim to have proven this result; however, their proof is based on an incorrect implication.

Karlin and Rinott (1980) have given a stronger definition of RR_2 called *strongly* MRR_2 (S-MRR_2) which implies the previous inequality. This definition however is not easy to check directly (see pages 503–513 of that paper) and the form of the definition does not lend itself to intuitive interpretation. We give here two new concepts of negative dependence. One of them has a quite simple form (i.e., the measure is RR_2 in pairs) and is a weaker definition than S-MRR_2 (as is shown in Remark vi, following Definition 2.1). However, it possesses many of the same properties as the S-MRR_2 class.

The second new definition which we call condition N (see Section 4) is stronger than S-MRR_2 but it is intuitive and is easier to verify for specific distributions than S-MRR_2. In fact, all of the examples of Karlin and Rinott (1980) satisfy this stronger condition, as we show in Section 5.

Unlike other RR_2 definitions, we do not need to assume the existence of a density. We work directly with the measure itself.

Let μ be a probability measure on the Borel sets in $\mathbb{R}^n$. If $I_1, \cdots, I_n$ are intervals in $\mathbb{R}^1$, we define the set function $\tilde{\mu}(I_1, \cdots, I_n)$ by $\tilde{\mu}(I_1, \cdots, I_n) = \mu(I_1 \times \cdots \times I_n)$. By abuse of notation, we write μ instead of $\tilde{\mu}$. If I and J are intervals in $\mathbb{R}^1$, we write $I < J$ if $x \in I$, $y \in J$ implies $x < y$, that is, if I lies to the left of J.

DEFINITION 2.1. Let μ be a probability measure on $\mathbb{R}^2$. We say that μ *is reverse regular of order two* (RR_2) if

$$\mu(I_1, I_2)\mu(I_1', I_2') \le \mu(I_1, I_2')\mu(I_1', I_2) \tag{2.1}$$

for all intervals $I_1 < I_1'$, $I_2 < I_2'$ in $\mathbb{R}^1$. We also say that $\mu(I_1, I_2)$ is RR_2 in the variables I_1, I_2. If μ is a probability measure on $\mathbb{R}^n$ ($n \ge 2$), we say that μ is RR_2 *in pairs* if $\mu(I_1, \cdots, I_n)$ is RR_2 in the pairs I_i, I_j for all $1 \le i < j \le n$ when the remaining variables are held fixed. The random variables $T_1, \cdots, T_n$ (or the random vector $\mathbf{T}$ or its distribution function F) are said to be RR_2 *in pairs* if its corresponding probability measure on $\mathbb{R}^n$ is.

REMARKS.

(i) The obvious TP_2 definitions for μ are obtained by reversing the inequality in (2.1).

(ii) Clearly, if μ is $\mathrm{RR}_2(\mathrm{TP}_2)$ in pairs, then so are all marginals. Furthermore, it is not difficult to show that if F is the distribution function associated with μ and if $\bar{F}(t_1, \cdots, t_n) = \mu((t_1, \infty), (t_2, \infty), \cdots, (t_n, \infty))$ is the survival function, then the functions F and $\bar{F}$ are $\mathrm{RR}_2(\mathrm{TP}_2)$ in pairs in the sense of (1.2) ((1.1)).

(iii) It is easy to show by a simple limiting argument that if μ is $\mathrm{RR}_2(\mathrm{TP}_2)$ in pairs, and if μ has a density f with respect to a product measure $m = m_1 \times \cdots \times m_n$ of σ-finite measures such that f is continuous on the support of m and zero off the support of m, then f is $\mathrm{RR}_2(\mathrm{TP}_2)$ in pairs.

(iv) In the $n = 2$ case, we have the stronger converse, namely, if μ has a density f with respect to a product measure $m = m_1 \times m_2$ of σ-finite measures which is $\mathrm{RR}_2(\mathrm{TP}_2)$ on $S_1 \times S_2$, where S_i is the support of $m_i (i = 1, 2)$, then μ is $\mathrm{RR}_2(\mathrm{TP}_2)$.

(v) In the TP_2 case, one can generalize to higher dimensions if one makes some assumption on the set $\{f > 0\}$. Let μ have a density f with respect to a product measure $m = m_1 \times \cdots \times m_n$ of σ-finite measures. Let S_i be the support of m_i. Then the support of m is $S = S_1 \times \cdots \times S_n$. We assume that there exists $\tilde{S} = \tilde{S}_1 \times \cdots \times \tilde{S}_n$ such that $\{f > 0\} \cap S = \tilde{S}$ and that f is TP_2 in pairs on $\tilde{S}$. Then μ is TP_2 in pairs. Just use Theorem 5.1, page 123 of Karlin (1968) repeatedly to show, e.g., that for fixed intervals $I_3, \cdots, I_n$ in $\mathbb{R}^1$,

$$g(x_1, x_2) = \int_{I_3} \cdots \int_{I_n} f(x_1, x_2, x_3, \cdots, x_n)\, dm_3(x_3) \cdots dm_n(x_n)$$

is TP_2 in x_1 and x_2 on $\tilde{S}_1 \times \tilde{S}_2$. The result then follows by a simple integration using the TP_2 inequality for g.

(vi) The generalization to the RR_2 case is not as simple. If one assumes, however, that μ has a density f with respect to a product measure $m = m_1 \times \cdots \times m_n$ of σ-finite measures such that the density f when integrated over any $n - 2$ intervals in $\mathbb{R}^1$ is RR_2 in the remaining unintegrated variables, then μ is RR_2 in pairs. In terms of random variables, this can be paraphrased as follows. Let $T_1, \cdots, T_n$ be random variables with a density f (with respect to a product measure of σ-finite measures). Then μ is RR_2 in pairs if and only if for every $1 \le i < j \le n$ the conditional density of

$$(T_i, T_j) \mid \cap_{k \ne i,j} \{T_k \in I_k\}$$

is RR_2 in t_i and t_j for all choices of intervals $I_k (k \ne i, j)$ in $\mathbb{R}^1$. Equivalently, if χ_I denotes the indicator function of I, then μ is RR_2 in pairs if and only if

$$\int_{\mathbb{R}} \cdots \int_{\mathbb{R}} [\textstyle\prod_{k \ne i,j} \chi_{I_k}(t_k)] f(t_1, \cdots, t_n) [\prod_{k \ne i,j} dt_k]$$

is RR_2 in the unintegrated variables t_i and t_j for all choices of intervals $I_k (k \neq i, j)$ in $\mathbb{R}^1$. By replacing χ_{I_k} by ϕ_k in the above integral, and requiring it to be RR_2 in t_i and t_j whenever $\{\phi_k\}_{k\neq i,j}$ is a set of PF_2 functions, one obtains the negative dependence condition of Karlin and Rinott (1980). It is a condition which is stronger than the one of Definition 2.1, as can be easily seen by recalling that the indicator function of an interval is PF_2. To see that it is strictly stronger, consider the $2 \times 2 \times 2$ discrete distribution

$$P(X = x, Y = y, Z = z) = \begin{cases} 1/4 & \text{if} \quad (x, y, z) \in \{(1, 0, 0), (0, 1, 0), (1, 1, 0), (0, 0, 1)\}, \\ 0 & \text{otherwise,} \end{cases}$$

which satisfies Definition 2.1, but does not satisfy the S-MRR_2 definition for the PF_2 function $\phi(z) = e^z$ (this example is due to M. L. Lee).

(vii) Clearly, if $\{\mu_n\}$ is a sequence of probability measures which are RR_2 in pairs and if μ_n converges weakly to μ then μ is RR_2 in pairs.

DEFINITION 2.2. The rv's $T_1, \cdots, T_n$ (or the random vector $\mathbf{T}$ or its df) are said to be *negatively upper orthant dependent* (NUOD) if for every $\mathbf{t}$,

$$P(\mathbf{T} > \mathbf{t}) \leq \pi_{i=1}^{n} P(T_i > t_i). \tag{2.2}$$

They are said to be *negatively lower orthant dependent* (NLOD) if for every $\mathbf{t}$,

$$P(\mathbf{T} \leq \mathbf{t}) \leq \pi_{i=1}^{n} P(T_i \leq t_i). \tag{2.2'}$$

When $n = 2$, (2.2) and (2.2′) are equivalent, but not when $n \geq 3$ (see, for example, Ebrahimi and Ghosh, 1980).

It is true that if μ is RR_2 in pairs then (2.2) and (2.2′) hold. This is clear since from the Remark (ii) following Definition 2.1, we have that both F and $\bar{F}$ are RR_2 in pairs. The result then follows by a simple argument. It also follows that the S-MRR_2 condition implies (2.2) and (2.2′) because of Remark (vi) following Definition 2.1. Finally, since any of these conditions implies that $P(T_i > t_i, T_j > t_j) \leq P(T_i > t_i)P(T_j > t_j)$ for $1 \leq i < j \leq n$, it follows, as in Lehmann (1966), that

$$\operatorname{cov}(T_i, T_j) \leq 0. \tag{2.3}$$

This justifies the name negative dependence for any of these concepts.

3. Properties.

This section is devoted to the derivation of some inequalities and properties which may be of special interest. The inequalities are similar to those found in Karlin and Rinott (1980), but are derived under weaker assumptions.

Suppose that μ is RR_2 in pairs. For every i, let $I_i = J_i \cup K_i$, all intervals, with $J_i < K_i$.

THEOREM 3.1. *If* $1 \leq k \leq n$, *then*

$$\mu(J_1, \cdots, J_n)\mu(I_1, \cdots, I_n) \leq \mu(J_1, \cdots, J_k, I_{k+1}, \cdots, I_n) \times \mu(I_1, \cdots, I_k, J_{k+1}, \cdots, J_n). \tag{3.1}$$

The result (3.1) is also true if we replace all J's by K's.

PROOF. We proceed by induction. If $n = 2$, then

$$\mu(J_1, J_2)\mu(I_1, I_2) \leq \begin{cases} \mu(J_1, I_2)\mu(I_1, J_2) & \text{if} \quad k = 1. \\ \mu(J_1, J_2)\mu(I_1, I_2) & \text{if} \quad k = 2. \end{cases}$$

The case $k = 1$ follows from the RR_2 assumption since

$$0 \geq \begin{vmatrix} \mu(J_1, J_2) & \mu(J_1, K_2) \\ \mu(K_1, J_2) & \mu(K_1, K_2) \end{vmatrix} = \begin{vmatrix} \mu(J_1, J_2) & \mu(J_1, J_2 \cup K_2) \\ \mu(J_1 \cup K_1, J_2) & \mu(J_1 \cup K_1, J_2 \cup K_2) \end{vmatrix},$$

and the case $k = 2$ is an identity.

Now suppose that (3.1) is true whenever ν is a probability measure on $\mathbb{R}^n$ which is RR_2 in pairs. Let μ be a probability measure on $\mathbb{R}^{n+1}$ which is RR_2 in pairs and let $1 \leq k \leq$

$n + 1$. Since there is nothing to prove if $k = n + 1$, we may assume that $1 \leq k \leq n$. Similarly, we may assume that $\mu(J_1, \cdots, J_n, J_{n+1}) \neq 0$. It suffices then to prove that

$$\frac{\mu(I_1, \cdots, I_n, I_{n+1})}{\mu(J_1, \cdots, J_k, I_{k+1}, \cdots, I_n, I_{n+1})} \leq \frac{\mu(I_1, \cdots, I_n, J_{n+1})}{\mu(J_1, \cdots, J_k, I_{k+1}, \cdots, I_n, J_{n+1})} \tag{3.2}$$

since by the induction hypothesis, we have that

$$\frac{\mu(I_1, \cdots, I_n, J_{n+1})}{\mu(J_1, \cdots, J_k, I_{k+1}, \cdots, I_n, J_{n+1})} \leq \frac{\mu(I_1, \cdots, I_k, J_{k+1}, \cdots, J_n, J_{n+1})}{\mu(J_1, \cdots, J_n, J_{n+1})}.$$

But,

$$\frac{\mu(I_1, \cdots, I_n, I_{n+1})}{\mu(I_1, \cdots, I_n, J_{n+1})} \leq \frac{\mu(J_1, I_2, \cdots, I_n, I_{n+1})}{\mu(J_1, I_2, \cdots, I_n, J_{n+1})} \leq \cdots \leq \frac{\mu(J_1, \cdots, J_k, I_{k+1}, \cdots, I_n, I_{n+1})}{\mu(J_1, \cdots, J_k, I_{k+1}, \cdots, I_n, J_{n+1})}$$

which is another way of writing (3.2). The jth inequality above follows from the fact that μ is RR_2 in the pair j and $(n + 1)$. □

Suppose that $(T_1, \cdots, T_n)$ is RR_2 in pairs.

COROLLARY 1. *If α, β partition $\{1, \cdots, n\}$, then*

$$\begin{aligned} &P(T_i \in J_i, i \in \alpha \cup \beta) P(T_i \in I_i, i \in \alpha \cup \beta) \\ &\quad \leq P(T_i \in J_i, i \in \alpha; T_j \in I_j, j \in \beta) P(T_i \in I_i, i \in \alpha; T_j \in J_j, j \in \beta). \end{aligned} \tag{3.3}$$

It also holds true if we replace all J's with the K's.

COROLLARY 2. *If α, β, γ partition $\{1, \cdots, n\}$, then*

$$\begin{aligned} &P(T_i \in L_i, i \in \alpha; T_j \in J_j, j \in \beta \cup \gamma) P(T_i \in L_i, i \in \alpha; T_j \in I_j, j \in \beta \cup \gamma) \\ &\leq P(T_i \in L_i, i \in \alpha; T_j \in J_j, j \in \beta; T_k \in I_k, k \in \gamma) \\ &\cdot P(T_i \in L_i, i \in \alpha; T_j \in I_j, j \in \beta; T_k \in J_k, k \in \gamma) \end{aligned} \tag{3.4}$$

for any intervals L_i, $i \in \alpha$. It also holds true if we replace all J's with the K's.

If we take $J_i = (-\infty, b_i]$ and $I_i = (-\infty, \infty)$ for $i \in \alpha \cup \beta$ in Corollary 1, we get (3.5) of Karlin and Rinott (1980). If we take, in Corollary 2, $L_i = [a_i, b_i]$, $i \in \alpha$, $J_j = (-\infty, b_j]$ and $I_j = (-\infty, \infty)$ for $j \in \beta \cup \gamma$ we get (1.7) of Karlin and Rinott (1980).

Also, note that as soon as we have an inequality of the form

$$P(T_1 \leq b_1, \cdots, T_n \leq b_n) \leq P(T_1 \leq b_1, \cdots, T_k \leq b_k) P(T_{k+1} \leq b_{k+1}, \cdots, T_n \leq b_n) \tag{3.5}$$

it follows easily that

$$E[\pi_{i=1}^n \phi_i(T_i)] \leq E[\pi_{i=1}^k \phi_i(T_i)] E[\pi_{j=k+1}^n \phi_j(T_j)] \tag{3.6}$$

whenever ϕ_ι are nonnegative and decreasing. Similarly, if we have

$$P(T_1 > a_1, \cdots, T_n > a_n) \leq P(T_1 > a_1, \cdots, T_k > a_k) P(T_{k+1} > a_{k+1}, \cdots, T_n > a_n) \tag{3.7}$$

then (3.6) holds for ϕ_i nonnegative and increasing.

In particular, (3.5) holds if F is RR_2 in pairs and (3.7) holds of $\bar{F}$ is RR_2 in pairs. In fact, if F (or $\bar{F}$) is RR_2 in pairs, then it is MRR_2 in the sense of Karlin and Rinott (1980); i.e.,

$$F(\mathbf{x} \wedge \mathbf{y}) F(\mathbf{x} \vee \mathbf{y}) \leq F(\mathbf{x}) F(\mathbf{y}).$$

The following results are useful for constructing new negatively dependent df's from known ones.

THEOREM 3.2. *If $T_1, \cdots, T_n$ are (*) and if $\psi_1, \cdots, \psi_n$ are strictly increasing functions then $\psi_1(T_1), \cdots, \psi_n(T_n)$ are (*) where (*) is one of the following:* RR_2 *in pairs,* NUOD *or* NLOD.

THEOREM 3.3. *If* $(T_1, \cdots, T_m)$ *and* $(S_1, \cdots, S_n)$ *are independent and are* (*) *then* $(T_1, \cdots, T_m, S_1, \cdots, S_n)$ *is* (*) *where* (*) *is the same as in Theorem* 3.2.

The proofs of these theorems are straightforward and will be omitted.

4. The structural condition. In this section we give our main result. The result is that if a joint distribution satisfies a certain structural condition, then it is negatively dependent according to all of the definitions discussed so far, as well as in the sense of Karlin and Rinott. The condition intuitively states that if a distribution is like the multinomial, i.e., essentially the sum of the random variables is fixed, then it satisfies the other negative dependence conditions. This structural condition conforms to our original idea that if the random variables are split into two subsets, one subset will be "large" when the other subset is "small".

We first need the following definition. A univariate density f is said to be a *Polya frequency function of order 2* (PF_2) if $f(x - y)$ is TP_2 on $\mathbb{R} \times \mathbb{R}$. A probability function f is PF_2 if $f(x - y)$ is TP_2 on $\mathbb{N} \times \mathbb{N}$ where $\mathbb{N} = \{\cdots, -1, 0, 1, \cdots\}$. A thorough discussion of PF_2 densities and many examples can be found in Karlin (1968).

DEFINITION 4.1. The random vector $(T_1, \cdots, T_n)$ satisfies *condition N* if there exist $n + 1$ independent rv's $S_0, S_1, \cdots, S_n$ each with a PF_2 density (or each with a PF_2 probability function) and a real number s such that

$$(T_1, \cdots, T_n) =_{st} [(S_1, \cdots, S_n) \mid S_0 + S_1 + \cdots + S_n = s]. \tag{4.1}$$

THEOREM 4.1. *Let* $(T_1, \cdots, T_n)$ *satisfy condition N. Then* $(T_1, \cdots, T_n)$ *is* RR_2 *in pairs and consequently* NUOD *and* NLOD.

PROOF. Let μ be the probability measure of $(T_1, \cdots, T_n)$ on $\mathbb{R}^n$. Then by assumption $\mu(I_1, \cdots, I_n) = P(S_1 \in I_1, \cdots, S_n \in I_n \mid S_0 + S_1 + \cdots + S_n = s)$. Now the joint density of $[(S_1, \cdots, S_n) \mid S_0 + S_1 + \cdots + S_n = s]$ is given by

$$c\, \pi_{i=1}^{n} f_i(s_i) f_0(s - s_1 - \cdots - s_n)$$

where c is a normalizing constant. We first show that μ is RR_2 in the variables I_1, I_2 when the remaining intervals $I_3, \cdots, I_n$ are held fixed. According to the Remark (vi) following Definition 2.1, we need only show that

$$g(s_1, s_2) = c f_1(s_1) f_2(s_2) f_0'(s - s_1 - s_2)$$

is RR_2 in s_1 and s_2, where

$$f_0'(\xi) = \int_{I_3} \cdots \int_{I_n} f_3(s_3) \cdots f_n(s_n) f_0(\xi - s_3 - \cdots - s_n)\, dm(s_3) \cdots dm(s_n)$$

and m is either the Lebesgue measure or the counting measure. However, the above is nothing but the convolution of the PF_2 functions $f_3\chi_{I_3}, \cdots, f_n\chi_{I_n}$ and f_0, where χ_A is the indicator function of the set A, and so f_0' is PF_2. It easily follows then that g is RR_2. The proof that μ is RR_2 in the variables I_i, I_j for all $1 \leq i < j \leq n$ is similar. □

REMARK. By changing the indicator functions to PF_2 functions in the previous proof, it follows that the S-MRR_2 definition of Karlin and Rinott (1980) is satisfied.

5. Examples. It is now shown that many of the standard examples of distributions which are considered to be negative dependent in some sense actually satisfy the structural condition N. Therefore they must satisfy all of the negative dependence conditions mentioned in this paper. In particular they satisfy (2.2) and (2.2′), so that we obtain as special cases the results of Jogdeo and Patil (1975).

5.1 *Multinomial.* Let $(T_1, \cdots, T_n)$ have the joint probability function with parameters $(N, p_1, \cdots, p_n)$,

$$P(T_1 = t_1, \cdots, T_n = t_n) = \frac{N!}{t_1! \cdots t_n!(N - \sum_{i=1}^n t_i)!} (\pi_{i=1}^n p_i^{t_i}) \times (1 - \sum_{i=1}^n p_i)^{N - \sum_{i=1}^n t_i}, \quad t_i \geq 0, \sum_{i=1}^n t_i \leq N,$$

where $p_i \geq 0$ $(i = 1, \cdots, n)$ and $0 < \sum_{i=1}^n p_i < 1$.

The multinomial df is the conditional df of independent Poisson rv's given their sum. Thus, by Theorem 4.1 the multinomial df is RR_2 in pairs and hence it is also NUOD and NLOD. By Remark (iii) the joint probability function of $(T_1, \cdots, T_n)$ is RR_2 in pairs. By the discussion after Theorem 4.1 the multinomial df satisfies the S-MRR_2 condition.

5.2 *Multivariate normal.* Let $\mathbf{T} = (T_1, \cdots, T_n)$ be a multivariate symmetric normal random vector with $\mathrm{Corr}(T_i, T_j) = \rho \leq 0$, $1 \leq i < j \leq n$. Then $\rho \geq -(n-1)^{-1}$. We will show that $\mathbf{T}$ is RR_2 in pairs.

Using Theorem 3.2 assume, without loss of generality, that $ET_i = 0$ and $\mathrm{Var}(T_i) = 1$, $i = 1, \cdots, n$. Let $Y_1, \cdots, Y_n$ be independent identically distributed normal rv's such that $EY_i = 0$ and $\mathrm{Var}(Y_i) = 1 - \rho$ $(i = 1, \cdots, n)$ and let Y_0 be an independent normal rv with $EY_0 = 0$ and $\mathrm{Var}(Y_0) = (-\rho)^{-1}(1 - \rho)(1 + (n-1)\rho)$. Then

$$(T_1, \cdots, T_n) =_{st} [(Y_1, \cdots, Y_n) \mid Y_0 + Y_1 + \cdots + Y_n = 0].$$

Since any normal density is PF_2 it follows, from Theorem 4.1, that $\mathbf{T}$ is RR_2 in pairs.

In fact we can obtain a stronger result. If the correlation matrix of $\mathbf{T}$ is of the form

$$\begin{bmatrix} r_1 & & 0 \\ & \ddots & \\ 0 & & r_n \end{bmatrix} - \begin{bmatrix} (r_1 - 1)^{1/2} \\ \vdots \\ (r_n - 1)^{1/2} \end{bmatrix} ((r_1 - 1)^{1/2}, \cdots, (r_n - 1)^{1/2}) \tag{5.2.1}$$

where $r_i \geq 1$, $i = 1, \cdots, n$ and $\sum_{i=1}^n r_i^{-1} \geq n - 1$, then $\mathbf{T}$ is RR_2 in pairs.

To show this, note that every matrix of the form (5.2.1) can be the correlation matrix of a multinomial random vector, $\mathbf{X} = (X_1, \cdots, X_n)$, say. Let $\mathbf{X}^{(\ell)}$, $\ell = 1, 2, \cdots$, be a sequence of independent random vectors distributed as $\mathbf{X}$. Clearly $\mathbf{Y}^{(m)} = \sum_{\ell=1}^m \mathbf{X}^{(\ell)}$ is a multinomial random vector with correlation matrix (5.2.1). Normalizing $\mathbf{Y}^{(m)}$ such that it has zero means and unit variances, it converges in distribution, by the multivariate central limit theorem, to a multivariate normal random vector with correlation matrix (5.2.1). By Remark (vii) the limit in distribution of RR_2 in pairs random vectors is RR_2 in pairs. The assertion in the preceding paragraph now follows. The previous result that deals with the symmetric multivariate normal df with negative correlations ρ is obtained by taking $r_i = 1 - \rho$ in (5.2.1).

A simple transformation of parameters shows that the Karlin-Rinott normal example is a special case of the above in that ours contains a singular normal distribution and theirs does not.

5.3. *Multivariate hypergeometric.* Let $(T_1, \cdots T_n)$ have the probability function

$$P(T = t_1, \cdots, T_n = t_n) = \binom{M}{N}^{-1} \left[\pi_{i=1}^n \binom{M_i}{t_i}\right] \binom{M - \sum_{i=1}^n M_i}{N - \sum_{i=1}^n t_i},$$
$$t_i \geq 0, \quad \sum_{i=1}^n t_i \leq N, \quad \sum_{i=1}^n M_i \leq M,$$

with positive integer-valued parameter vector $(N, M_1, \cdots, M_n, M)$ [see Johnson and Kotz, 1969].

The multivariate hypergeometric df is the conditional df of independent binomial rv's given their sum. Thus, by Theorem 4.1 the hypergeometric df is RR_2 in pairs and hence it is also NUOD and NLOD. By Remark (iii) the joint probability function of $(T_1, \cdots, T_n)$

is RR_2 in pairs. A special case of this fact was observed by Lehmann (1966), page 1144. See also Ebrahimi and Ghosh (1980) and Gastwirth (1980). By the discussion after Theorem 4.1 it follows that this df satisfies the S-MRR_2 condition.

5.4 *The Dirichlet Distribution.* Let $\mathbf{T} = (T_1, \cdots, T_n)$ have the density

$$f(t_1, \cdots, t_n) = \frac{\Gamma(\sum_{j=0}^{n} \theta_j)}{\pi_{j=0}^{n} \Gamma(\theta_j)} (1 - \sum_{j=1}^{n} t_j)^{\theta_0 - 1} \pi_{j=1}^{n} t_j^{\theta_j - 1}, \qquad t_j \geq 0, \qquad \sum_{j=1}^{n} t_j \leq 1,$$

where the parameter vector $(\theta_0, \theta_1, \cdots, \theta_n)$ satisfies $\theta_j \geq 1, j = 0, 1, \cdots, n$.

The Dirichlet df is the conditional df of independent gamma rv's given their sum is equal to 1. Thus, by Theorem 4.1 the Dirichlet df is RR_2 in pairs and hence it is also NUOD and NLOD. By Remark (iii) f is RR_2 in pairs. A special case of this fact is Example 10 (iii) of Lehmann (1966). See also Ebrahimi and Ghosh (1980). By the discussion after Theorem 4.1 it follows that this df satisfies the S-MRR_2 condition.

5.5 *Dirichlet compound multinomial df.* Let $\mathbf{T} = (T_1, \cdots, T_n)$ have the probability function

$$P(T_1 = t_1, \cdots, T_n = t_n) = \frac{N!\Gamma(\sum_{j=0}^{n} \theta_j)}{\Gamma(N + \sum_{j=0}^{n} \theta_j)} \pi_{j=1}^{n} \frac{\Gamma(t_j + \theta_j)}{t_j!\Gamma(\theta_j)} \times \frac{\Gamma(N - \sum_{j=1}^{n} t_j + \theta_0)}{(N - \sum_{j=1}^{n} t_j)!\Gamma(\theta_0)}, \qquad t_j \geq 0, \qquad \sum_{j=1}^{n} t_j \leq N,$$

where N is a positive integer and $\theta_j \geq 1, j = 0, 1, \cdots, n$ [see Johnson and Kotz, 1969].

The Dirichlet compound multinomial df is the conditional df of independent Pascal (negative binomial) rv's given their sum. Thus, by Theorem 4.1 this df is RR_2 in pairs and hence it is also NUOD and NLOD, and by the remark after Theorem 4.1 it satisfies the S-MRR_2 condition.

REFERENCES

[1] Barlow, R. E. and Proschan, F. (1975). *Statistical Theory of Reliability and Life Testing: Probability Models.* Holt, Rinehart and Winston, New York.

[2] Block, H. W. and Ting, M. L. (1981). Some concepts of multivariate dependence. *Comm. Statist.* **A10(8)** 749–762.

[3] Brindley, E. C. and Thompson, W. A. (1972). Dependence and aging aspects of multivariate survival. *J. Amer. Statist. Assoc.* **67** 822–830.

[4] Dykstra, R. L., Hewett, J. E. and Thompson, W. A. (1973). Events which are almost independent. *Ann. Statist.* **1** 674–681.

[5] Ebrahimi, N. and Ghosh, M. (1980). Multivariate negative dependence. Technical Report, Department of Statistics, Iowa State University, Ames, Iowa.

[6] Gastwirth, J. L. (1980). A statistical problem arising in jury discrimination. Technical Report, Department of Statistics, George Washington University.

[7] Jogdeo, K. and Patil, G. P. (1975). Probability inequalities for certain multivariate discrete distributions. *Sankhyā, Ser. B,* **37** 158–164.

[8] Johnson, N. L. and Kotz, S. (1969). *Distribution in Statistics, Discrete Distributions.* Wiley, New York.

[9] Karlin, S. (1968). *Total Positivity.* Stanford University Press, Stanford.

[10] Karlin, S. and Rinott, Y. (1980). Classes of orderings of measures and related correlation inequalities. II. multivariate reverse rule distributions. *J. Multivariate Anal.* **10** 499–516.

[11] Kemperman, J. H. B. (1977). On the FKG-inequality measures on a partially ordered space. *Proc. Kon. Ned. Akad. Wet., Amsterdam, Series A,* **80 (4)** 313–331.

[12] Lehmann, E. L. (1966). Some concepts of dependence. *Ann. Math. Statist.* **43** 1137–1153.

[13] Mallows, C. L. (1968). An inequality involving multinomial probabilities. *Biometrika* **55** 422–424.

[14] Shaked, M. (1977). A family of concepts of positive dependence for bivariate distributions. *J. Amer. Statist. Assoc.* **72** 642–650.

Henry W. Block
Thomas H. Savits
Department of Mathematics and Statistics
University of Pittsburgh
Pittsburgh, Pennsylvania 15260

Moshe Shaked
Department of Mathematics
Indiana University
Bloomington, Indiana 47404