Negative Association of Random Variables, With Applications

by

Kumar Jogdeo and Frank Proschan

The Florida State University
Department of Statistics
Tallahassee, Florida 32306

October, 1981

FSU Statistics Report M-590
AFOSR Technical Report No. 81-127


Key Words: Negative association, negative dependence, negative orthant dependence, association, conditionally decreasing in sequence, negatively dependent in sequence, reverse rule of order 2, multivariate inequalities, negatively correlated normal random variables.

Research performed under the support of Air Force Office of Scientific Research, AFSC, USAF, under Grant-AFOSR-81-0038.
ABSTRACT

Negative Association of Random Variables, With Applications

by

Kumar Jogdeo and Frank Proschan

Random variables \( X_1, \ldots, X_k \) are said to be negatively associated (NA) if for every pair of disjoint subsets \( A_1, A_2 \) of \( \{1, 2, \ldots, k\} \),

\[
\text{Cov}[f(X_i, i \in A_1), g(X_j, j \in A_2)] \geq 0
\]

for all nondecreasing functions \( f, g \). The basic properties of NA random variables are derived. Especially useful is the property that nondecreasing functions of mutually exclusive subsets of NA random variables are NA. This property is shown not to hold for several other types of negative dependence recently proposed.

One application yields the inequality

\[
\frac{1}{k} \sum_{i=1}^{k} \Pr[X_i \leq x_i] \leq \Pr[X_1, \ldots, X_k \leq x_1]^{1/k}
\]

for NA random variables \( X_1, \ldots, X_k \), and the dual inequality resulting from reversing the inequalities inside the square brackets. In another application, it is shown that negatively correlated normal random variables are NA. Other NA distributions are the (a) multinomial, (b) convolution of unlike multinomials, (c) multivariate hypergeometric, (d) Dirichlet, and (e) Dirichlet compound multinomial. Other applications are to multiple ranking of individuals and ranking of populations.

Negative Association of Random Variables, With Applications.

1. Introduction and Summary. The concept of (positively) associated random variables was introduced into the statistical literature by Esary, Proschan, and Walkup (1967). Since then a great many papers have been written on the subject and its extensions, and numerous multivariate inequalities have been obtained.
In this paper we introduce the notion of negatively associated (NA) random variables (see Section 2 for the definition), derive basic theoretical properties, and develop applications in multivariate statistical analysis. The theory and application of NA are not simply the duals of the theory and application of positive association, but differ in important respects.

Actually, NA is but one qualitative version of negative dependence among random variables. Other versions are treated in Block, Savits, and Shaked (1980) (BSS(1980), Ebrahimi and Ghosh (1980), Jogdeo and Patil (1975), and Karlin and Rinott (1980). (See Section 2 for definitions of certain types of negative dependence.)

NA has one distinct advantage over the other known types of negative dependence. Increasing functions of disjoint sets of NA random variables are also NA. This type of closure property does not hold for the three other types of negative dependence described in Section 2.

In Section 2, we define NA and develop its basic properties. We define other types of negative dependence, such as negative upper (lower) orthant dependence, reverse regular of order two (RR$_2$) in pairs, conditionally decreasing in sequence (CDS), and negatively dependent in sequence (NDS), introduced by other statisticians working in negative dependence. We prove that among these types of negative dependence, only the NA class enjoys the important property of being closed under formation of increasing functions of disjoint sets of random variables, as mentioned above.
A very useful theorem is proved in Section 2 which gives simple sufficient conditions for the conditional random variables \( (X_i | I, \ldots, X_i | I) \) to be NA (Theorem 2.9).

In Section 3, we show that permutation distributions are NA. (See Def. 3.1 of a permutation distribution.) Next we give applications of the result. For example, we show that the ranks of a random sample of observations are NA. In the remaining two applications, we apply this last result to problems in multiple rankings (say m judges ranking n individuals) and population rankings.

In Section 4, we present applications of the theory developed in Section 2. An aesthetically appealing result states very simply that negatively correlated normal random variables are NA. Next we point out that a number of well known multivariate distributions possess the NA property, such as the (a) multinomial, (b) convolution of unlike multinomials, (c) multivariate hypergeometric, (d) Dirichlet, and (e) Dirichlet compound multinomial. As a consequence, they satisfy the inequalities (4.4) of Section 4.

2. Definition and Basic Properties.

2.1. Definition. Random variables \( X_1, \ldots, X_k \) are said to be negatively associated (NA) if for every pair of disjoint subsets \( A_1, A_2 \) of \( \{1, \ldots, k\} \),

\[
\text{Cov} \left( f_1(X_i, i \in A_1), f_2(X_j, j \in A_2) \right) \leq 0
\]  

(2.1)
whenever $f_1$ and $f_2$ are increasing. "NA" may also refer to the vector $X = (X_1, \ldots, X_k)$ or to the underlying distribution of $X$. Additionally, "NA" may denote negative association.

(Throughout, we use increasing in place of nondecreasing, decreasing in place of nonincreasing, positive in place of nonnegative, and negative in place of nonpositive.)

Clearly (2.1) holds if both $f_1$, and $f_2$ are decreasing. Also, without loss of generality, we may assume that $A_1 \cup A_2 = \{1, \ldots, k\}$.

Some basic properties of NA follow.

2.2. Definition (Lehmann). Random variables $X$ and $Y$ are negative quadrant dependent (NQD) if

$$P[X \leq x, Y \leq y] \leq P[X \leq x]P[Y \leq y]$$

(2.2)

for all real $x$ and $y$.

($P_1$) For a pair of random variables, NQD is equivalent to NA.

The proof is immediate from Def.'s 2.1 and 2.2.

($P_2$) Let $A_1, \ldots, A_m$ be disjoint subsets of $\{1, \ldots, k\}$ and $f_1, \ldots, f_m$ be increasing positive functions. Then $X_{A_1}, \ldots, X_{A_k}$ NA implies

$$E_{j=1}^{m} f_j(X_j, j \in A_j) \leq \prod_{j=1}^{m} E f_j(X_j, j \in A_j).$$

(2.3)

Proof. $E_{j=1}^{m} f_j(X_j, j \in A_j) \leq Ef_j(X_j, j \in A_j) \times \prod_{j=1}^{m} f_j(X_j, j \in A_j)$

[by Def. 2.1]. Repeated application of Def. 2.1 yields (2.3). \[\]

2.3. Definitions. The random variables $X_1^+, \ldots, X_k^+$ are said to be negatively upper orthant dependent (NUOD) if

$$P[X_i > x_i, i=1, \ldots, k] \leq \prod_{i=1}^{k} P[X_i > x_i]$$

(2.4a)

for all $x_1, \ldots, x_k$. $X_1^-, \ldots, X_k^-$ are said to be negatively lower orthant dependent (NLOD) if
\[ P \left( X_i \leq x_i, \ i=1, \ldots, k \right) \leq \prod_{i=1}^{k} P \left( X_i \leq x_i \right) \]  \hspace{1cm} (2.4b)

for all \( x_1, \ldots, x_k \). Random variables \( X_1, \ldots, X_k \) are said to be negatively orthant dependent (NOD) if they are both NUOD and NLOD.

(P$_3$) Let \( X_1, \ldots, X_k \) be NA. Then (a) for disjoint sets \( A_1, A_2 \) whose union is \( \{1, \ldots, k\} \), we have
\[ P \left( X_i \leq x_i, \ i=1, \ldots, k \right) \leq P \left( X_i \leq x_i, \ i \in A_1 \right) P \left( X_i > x_i, \ i \in A_2 \right) \]  \hspace{1cm} (2.5a)
and
\[ P \left( X_i > x_i, \ i=1, \ldots, k \right) \leq P \left( X_i > x_i, \ i \in A_1 \right) P \left( X_i \leq x_i, \ i \in A_2 \right) \]  \hspace{1cm} (2.5b)
for all \( x_1, \ldots, x_k \). As a consequence,

(b) \( X_1, \ldots, X_k \) are NOD.

These results follow immediately from Def. 2.1.

Next we show

2.4. Neither NUOD nor NLOD implies NA. We present an example in which \( X = (X_1, X_2, X_3, X_4) \) possesses NOD, but does not possess NA.

Let \( X_4 \) be a binary random variable such that \( P(X_4 = 1) = .5 \) for \( i = 1, 2, 3, 4 \). Let \( (X_1, X_2) \) have joint distribution:

\[
\begin{array}{c|cc}
X_1 & 0 & 1 \\
\hline
0 & .24 & .26 \\
1 & .26 & .24 \\
\end{array}
\]

Let \( (X_3, X_4) \) have the same joint distribution. Let \( (X_1, X_2, X_3, X_4) \) have joint distribution:
\[(X_1, X_2)\]

\[
\begin{array}{cccc}
(0,0) & (0,1) & (1,0) & (1,1) \\
(0,0) & 0.0577 & 0.0623 & 0.0623 & 0.0577 \\
(0,1) & 0.0623 & 0.0677 & 0.0677 & 0.0623 \\
(1,0) & 0.0623 & 0.0677 & 0.0677 & 0.0623 \\
(1,1) & 0.0577 & 0.0623 & 0.0623 & 0.0577 \\
\end{array}
\]

\[(X_3, X_4)\]

It can be verified that all the NLCD and all the NUDD conditions hold (with strict inequalities in some cases). However,

\[P[X_1 = 1, i = 1, 2, 3, 4] > P[X_1 = 1 = X_2]P[X_3 = 1 = X_4],\]

violating NA.

\[(P_u)\] The union of independent sets of NA random variables is NA.

We will first need the following two elementary identities.

2.5 Identities. Let \(U, V, W\) be random vectors, with \(U, W\) independent. Then for arbitrary measurable functions \(f, g\),

(a) \(\text{Cov} [f(U), g(V)] = \text{Cov} [f(U), E(g(V)|U)]\), \hspace{1cm} (2.6)

(b) \(\text{Cov} [f(U), g(V, W)] = E[\text{Cov}(f(U), g(V, W)|W)]\) \hspace{1cm} (2.7)

Proof. (a) is easily verified; we omit the details.

(b) follows from the fact that the left side of (2.7) is equal to the right side plus the 0-term: \(\text{Cov} [E(f(U)|W), E(g(V, W)|W)]\).

This last term is 0 since the independence of \(U\) and \(W\) implies that \(E[f(U)|W] = E[f(U)]\), a constant. \[\|

Proof of \((P_u)\): Let \(X\) and \(Y\) be independent vectors, each NA. We shall show that the vector \((X, Y)\) is NA.
Let \((X_1, X_2)\) and \((Y_1, Y_2)\) denote arbitrary partitions of \(X\) and \(Y\) respectively. Let \(f\) and \(g\) be arbitrary increasing functions.

Write

\[
\text{Cov} [f(X_2, Y_1), g(X_2, Y_2) - f(X_1, Y_1)] = \text{Cov} [(f, g) | Y_1] + \text{Cov} [E(f|Y_1), E(g|Y_1)],
\]

where we now omit the arguments of \(f\) and \(g\). Since \(X\) is NA and independent of \(Y_2\) given \(Y_1\), the first term on the right is negative by Identity 2.5(b).

We may rewrite the second term on the right as:

\[
\text{Cov} [E(f|Y_1), E(g|Y_1)] = \text{Cov} [h(Y_1), E(g|Y_1)]
\]

[where \(h(Y_1) = E(f|X_1, Y_1|Y_1)\)]

= \text{Cov} [h(Y_1), g];

the last equality follows from Identity 2.5(a). By using the fact that \(X_1\) is independent of \((Y_1, Y_2)\), the fact that \(Y\) is NA, and 2.5(b) again, we deduce that the last covariance is negative.

Thus we have shown that \((X, Y)\) is NA. ||

(P_5) Increasing functions defined on disjoint subsets of a set of NA random variables are NA.

The proof is similar to that of the corresponding property of positively associated random variables.

2.6 Remark. Properties (P_4) and (P_5) broaden the scope of application of NA considerably. As a simple illustration, consider a multinomial distribution with \(k\) cells and with only one observation. Thus the multinomial outcome vector \((X_1, \ldots, X_k)\) consists of \(k-1\) elements 0 and one element 1. It is easy to check the distribution
But given $S_1$, $S_2$, $f_1(X_i, i \in A_1)$ and $f_2(X_j, j \in A_2)$ are independent. Thus $E \{[\text{Cov}(f_1, f_2) | S_1, S_2] | S \} = 0$. Furthermore, $E[f_1(X_i, i \in A_1) | S_1, S_2] = E[f_1(x_i, i \in A_1) | S_1]$ is increasing in $S_1$ by hypothesis. Similarly, $E[f_2(x_j, i \in A_2) | S_2]$ is increasing in $S_2$ by hypothesis, and hence is decreasing in $S_1$, since $S_1 + S_2 = S$ is given. Thus the second term of the expression set off just above is the covariance of two functions of $S_1$, one increasing and the other decreasing, and hence is negative, as follows from Tchebychev's inequality. ||

Theorem 2.7 takes on added interest when considered in conjunction with the following theorem of Efron (1965):

2.8 Theorem (Efron). Let $X_1, \ldots, X_k$ be $k$ independent random variables with PF$_2$ densities, let $S = \sum_{1}^{k} X_i$ be their sum, and let $\phi(x_1, \ldots, x_k)$ be increasing in each argument. Then $E [ \phi(X_1, \ldots, X_k) | S_k = s]$ is increasing in $s$.

As a consequence of Theorems 2.7 and 2.8, we immediately obtain:

2.9. Theorem. Let $X_1, \ldots, X_k$ be $k$ independent random variables with PF$_2$ densities. Then the conditional random variables $(X_1, \sum_{1}^{k} X_i), \ldots, (X_k, \sum_{1}^{k} X_i)$ are NA.

BSS(1980) obtain a related result. We need some definitions first, developed by BSS (1980) as modifications of Karlin and Rinott (1980) concepts.

Let $\mu$ be a probability measure on the Borel sets in $R^k$. For intervals $I_1, \ldots, I_k$ in $R^k$, define set function $\tilde{\mu}(I_1, \ldots, I_k)$