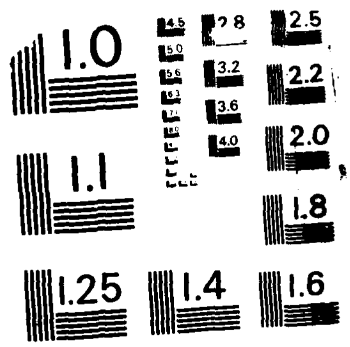
AD-A190 329 SOME COMMENTS ON QUADRANT DEPENDENCE IN HIGHER DIMENSIONS(U) PITTSBURGH UNIV PA CENTER FOR MULTIVARIATE ANALYSIS K SUBRAMANYAM SEP 87 TR-87-32 1/1

UNCLASSIFIED N00014-85-C-0008 F/G 12/2 NL



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

AD-A190 329

DTIC FILE COPY

(2)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER AFOSR-TR- 87-1938	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Some comments on positive quadrant dependence in higher dimensions		5. TYPE OF REPORT & PERIOD COVERED ~~Technical~~ - September 1987 journal
		6. PERFORMING ORG. REPORT NUMBER 87-32
7. AUTHOR(s) K. Subramanyam		8. CONTRACT OR GRANT NUMBER(s) ~~N00014-85-K-0292~~ F49620-85-C-0008
9. PERFORMING ORGANIZATION NAME AND ADDRESS Center for Multivariate Analysis, 515 Thackeray Hall, University of Pittsburgh, Pittsburgh, PA 15260		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 61102F 2304 A5
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research; Air Force Office of Scientific Research		12. REPORT DATE September 1987
		13. NUMBER OF PAGES 14
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) AFOSR Bldg 410 BAFB DC 20332 NM		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

DTIC ELECTE JAN 05 1988 S D E

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Upper positive quadrant dependence, lower positive quadrant dependence, convex set, extreme points

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

An extreme point analysis has been performed on two natural definitions of positive quadrant dependence of three random variable. This analysis helps us to understand how much these two notions of dependence are different.

DD FORM 1 JAN 73 1473

Unclassifed

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

AFOSR-TR- 87-1938

Center for Multivariate Analysis

University of Pittsburgh

87 12 29 245

SOME COMMENTS ON POSITIVE QUADRANT DEPENDENCE
IN HIGHER DIMENSIONS

K. Subramanyam

Center for Multivariate Analysis
University of Pittsburgh
Pittsburgh PA 15260

September 1987

Technical Report Number 87-32

Center for Multivariate Analysis
University of Pittsburgh
Pittsburgh PA 15260

Research sponsored by Contract N00014-85-K-0292 of the Office of Naval Research and Contract F49620-85-C-0008 of the Air Force Office of Scientific Research. The United States Government is authorized to reproduce and distribute reprints for governmental purposes notwithstanding any copyright notation hereon.

ABSTRACT

An extreme point analysis has been performed on two natural definitions of positive quadrant dependence of three random variables. This analysis helps us to understand how much these two notions of dependence are different.

Key words and phrases. Upper positive quadrant dependence, lower positive quadrant dependence, convex set, extreme points

1. INTRODUCTION

Let X and Y be two random variables with some joint probability distribution function F. X and Y or F are/is said to be positive quadrant department (PQD) if

$$\Pr(X \leq x,\ Y \leq y) \geq \Pr(X \leq x)\ \Pr(Y \leq y) \tag{1.1}$$

for all real numbers x and y. The condition (1.1) is equivalent to each of

$$\Pr(X \geq x,\ Y \geq y) \geq \Pr(X \geq x)\Pr(Y \geq y) \tag{1.2}$$

for all x and y,

$$\Pr(X \leq x,\ Y \geq y) \leq \Pr(X \leq x,\ Y \geq y) \tag{1.3}$$

for all x and y,

$$\Pr(X \geq x,\ Y \leq y) \leq \Pr(X \geq x)\ \Pr(Y \leq y) \tag{1.4}$$

for all x and y. See Lehmann (1966, p.1138).

One faces problems if one wishes to extend the notion of positive quadrant dependence to more than two random variables. If X,Y, and Z are three random variables, one could say that X, Y, and Z are PQD by adapting any one of the conditions (1.1), (1.2), (1.3), or (1.4) in a natural way To be more precise, say that X, Y, and Z are positive lower orthant dependent (PLOD) if

$$\Pr(X \leq x,\ Y \leq y,\ Z \leq z) \geq \Pr(X \leq x)\Pr(Y \leq y)\ \Pr(Z \leq z) \tag{1.5}$$

for all x, y, and z. Say that X, Y, and Z are positive upper orthant dependent (PUOD) if

$$\Pr(X \geq x,\ Y \geq y,\ Z \geq z) \geq \Pr(X \geq x)\ \Pr(Y \geq y)\ \Pr(Z \geq z) \tag{1.6}$$

for all x, y, and z.

In this paper, we discuss the ramifications of the definitions of PLOD and PUOD which are analogues of conditions (1.1), and (1.2), respectively. These two notions of PLOD and PUOD are not equivalent. Ahmed, Langberg, Leon and Proschan (1978) gave an example of a trivariate distribution which is PUOD, but not PLOD.

The main goal of this paper is to examine how different are these two notions of dependence. More precisely, we want to perform extreme point analysis on these two notions of dependence. In some special cases, extreme point analysis helps us to characterize all trivariate distributions which are both PLOD and PUOD.

2. EXTREME POINT ANALYSIS

To simplify the problem, we consider the case where each of X, Y, and Z assumes only two values 1 and 2, say. Let $p_{ijk} = Pr(X = i, Y = j, Z = k)$, $i = 1,2$; $j = 1,2$; $k = 1,2$. The joint probability law of X,Y, and Z is written, for convenience,

$$P = \begin{bmatrix} p_{111} & p_{112} & p_{121} & p_{122} \\ p_{211} & p_{212} & p_{221} & p_{222} \end{bmatrix}$$

In terms of this new notation, P is PLOD if

$$p_{111} \geq p_1 q_1 r_1 \tag{2.1}$$

$$p_{111} + p_{112} \geq p_1 q_1 \tag{2.2}$$

$$p_{111} + p_{121} \geq p_1 r_1 \tag{2.3}$$

$$p_{111} + p_{211} \geq q_1 r_1 \tag{2.4}$$

and P is PUOD IS

$$p_{222} \geq p_2 q_2 r_2 \tag{2.5}$$

$$p_{222} + p_{221} \geq p_2 q_2 \tag{2.6}$$

$$p_{222} + p_{212} \geq p_2 r_2 \tag{2.7}$$

$$p_{222} + p_{122} \geq q_2 r_2 \qquad (28$$

where $p_1 = Pr(X = 1)$; $q_1 = Pr(Y=1)$; $r_1 = Pr(Z=1)$; $p_2 = 1-p_1$; $q_2 = 1 - q_1$; and $r_2 = 1 - r_1$.

The extreme point analysis consists of looking at these two notions of dependence from a global point of view. Let $0 < p_1 < 1$, $0 < q_1 < 1$, and $0 < r_1 < 1$ be three fixed numbers. Let $M_{PLOD}(P_1,q_1,r_1)$ be the collection of all trivariate distributions $P = (p_{ijk})$ with support contained in $\{(i,j,k);$

$i = 1,2$, $j = 1,2$, and $k = 1,2\}$ such that P is PLOD, and the marginal distributions of X, Y, and Z under P are $p_1, 1-p_1$; $q_1, 1-q_1$; and $r_1, 1-r_1$ respectively. The set $M_{PUOD}(p_1,q_1,r_1)$ is defined analogously. The following result is obvious.

THEOREM 1. The sets $M_{PLOD}(p_1,q_i,r_1)$ and $M_{PUOD}(p_1,q_1,r_1)$ are compact and convex. More strongly, they are simplexes, i.e., each of these sets is bounded and a finite intersection of hyperplanes.

Nguyen and Sampson (1985) have looked into properties of sets of the above type for bivariate distributions with fixed marginals. Subramanyam and Bhaskara Rao (1986) have developed an algebraic method for identifying the extreme points of sets of the above type in the context of bivariate distributions.

Being simplexes, *the sets* $M_{PLOD}(p_1,q_1,r_1)$ *and* $M_{PUOD}(p_1,q_1,r_1)$ have each a finite number of extreme points. Once we identify the extreme points of the set $M_{PLOD}(p_1,q_1,r_1)$ say, we can express every member of $M_{PLOD}(p_1,q_1,r_1)$ as a convex combination of the extreme points of $M_{PLOD}(p_1,q_1,r_1)$. We describe now a method of identifying the extreme points of $M_{PLOD}(p_1,q_1,r_1)$ as well as $M_{PUOD}(p_1,q_1,r_1)$. First, we take up the case of $M_{PLOD}(p_1,q_1,r_1)$. Any $P = (p_{ijk}) \in M_{PLOD}(p_1,q_1,r_1)$ will have to satisfy the inequalities (2.1), (2.2), (2.3), and (2.4). Also, due to marginality restrictions, we should have

$$p_{111} + p_{112} + p_{121} \le p_1 \tag{2.9}$$

$$p_{111} + p_{112} + p_{211} \le q_1 \tag{2.10}$$

$$p_{111} + p_{121} + p_{211} \le r_1 \tag{2.11}$$

$$p_{112} \geq 0 \tag{2.12}$$

$$p_{121} \geq 0 \tag{2.13}$$

$$p_{211} \geq 0 \tag{2.14}$$

All these inequalities (2.1) to (2.4) and (2.9) to (2.14) involve p_{111}, p_{112}, p_{121}, p_{211} only. If some four numbers p_{111}, p_{112}, p_{121}, p_{211} satisfy the inequalities (2.1) to (2.4) and (2.9) to (2.14), then one could define

$$p_{122} = p_1 - (p_{111} + p_{112} + p_{121}), \tag{2.15}$$

$$p_{212} = q_1 - (p_{111} + p_{112} + p_{211}), \tag{2.16}$$

$$p_{221} = r_1 - (p_{111} + p_{121} + p_{211}), \tag{2.17}$$

and

$$p_{222} = 1 - p_1 - q_1 - r_1 + p_{111} + (p_{111} + p_{112} + p_{121} + p_{211}). \tag{2.18}$$

The numbers p_{122}, p_{212}, and p_{221} will be nonnegative. If $p_{222} \geq 0$, then

$$P = (p_{ijk}) \in M_{PLOD}(p_1, q_1, r_1)$$

A standard method of identifying the extreme points of $M_{PLOD}(p_1, q_1, r_1)$ is as follows. Select 4 inequalities from (2.1) to (2.4) and (2.9) to (2.14). Replace the inequality signs by equality signs. Solve the resultant system of 4 linear equations in 4 unknowns p_{111}, p_{112}, p_{121} and p_{211}. If there is a solution, and this solution satisfies the remaining inequalities, determine p_{122}, p_{212}, p_{221}, and p_{222} as per the equations (2.15), (2.16), (2.17), and (2.10). If $p_{222} \geq 0$, then

$$P = (p_{ijk})$$

is an extreme point of $M_{PLOD}(p_1,q_1,r_1)$. A computer program is easy to write which will identify the extreme points of $M_{PLOD}(p_1,q_1,r_1)$.

Pursuing this approach, we have isolated the extreme points of $M_{PLOD}(p_1,q_1,r_1)$ and $M_{PUOD}(p_1,q_1,r_1)$ when $p_1 = q_1 = r_1 = 1/2$, given in Table 1.

The above extreme point analyses of the sets $M_{PLOD}(\frac{1}{2},\frac{1}{2},\frac{1}{2})$ and $M_{PUOD}(\frac{1}{2},\frac{1}{2},\frac{1}{2})$ reveal the following insights.

1. The extreme points of $M_{PLOD}(\frac{1}{2},\frac{1}{2},\frac{1}{2})$ and $M_{PUOD}(\frac{1}{2},\frac{1}{2},\frac{1}{2})$ fall into three distinct categories. The first five extreme points are common to both the sets. Observe that

$$P_6 = \frac{1}{2} P_4 + \frac{1}{2} P_{15}$$

$$P_8 = \frac{1}{2} P_2 + \frac{1}{2} P_{15}$$

$$P_{10} = \frac{1}{2} P_3 + \frac{1}{2} P_{15}$$

$$P_{12} = \frac{1}{4} P_5 + \frac{3}{4} P_{15}$$

Consequently, P_6, P_8, P_{10}, P_{12} ε $M_{PUOD}(\frac{1}{2},\frac{1}{2},\frac{1}{2})$. Also observe that

$$P_7 = \frac{1}{2} P_4 + \frac{1}{2} P_{14}$$

$$P_9 = \frac{1}{2} P_2 + \frac{1}{2} P_{14}$$

$$P_{11} = \frac{1}{2} P_3 + \frac{1}{2} P_{14}$$

$$P_{13} = \frac{1}{4} P_5 + \frac{3}{4} P_{14}$$

Consequently P_7, P_9, R_{11}, P_{13} ε $M_{PLOD}(\frac{1}{2},\frac{1}{2},\frac{1}{2})$, and

Table 1 Extreme Points of $M_{PLOD}(\frac{1}{2},\frac{1}{2},\frac{1}{2})$ and $M_{PUOD}(\frac{1}{2},\frac{1}{2},\frac{1}{2})$

Serial No.	$M_{LPQD}(\frac{1}{2},\frac{1}{2},\frac{1}{2})$	$M_{UPOD}(\frac{1}{2},\frac{1}{2},\frac{1}{2})$
1.	$P_1 = \frac{1}{8}\begin{bmatrix}1 & 1 & 1 & 1\\ 1 & 1 & 1 & 1\end{bmatrix}$	$P_1 = \frac{1}{8}\begin{bmatrix}1 & 1 & 1 & 1\\ 1 & 1 & 1 & 1\end{bmatrix}$
2.	$P_2 = \frac{1}{8}\begin{bmatrix}2 & 0 & 2 & 0\\ 0 & 2 & 0 & 2\end{bmatrix}$	$P_2 = \frac{1}{8}\begin{bmatrix}2 & 0 & 2 & 0\\ 0 & 2 & 0 & 2\end{bmatrix}$
3.	$P_3 = \frac{1}{8}\begin{bmatrix}2 & 2 & 0 & 0\\ 0 & 0 & 2 & 2\end{bmatrix}$	$P_3 = \frac{1}{8}\begin{bmatrix}2 & 2 & 0 & 0\\ 0 & 0 & 2 & 2\end{bmatrix}$
4.	$P_4 = \frac{1}{8}\begin{bmatrix}2 & 0 & 0 & 2\\ 2 & 0 & 0 & 2\end{bmatrix}$	$P_4 = \frac{1}{8}\begin{bmatrix}2 & 0 & 0 & 2\\ 2 & 0 & 0 & 2\end{bmatrix}$
5.	$P_5 = \frac{1}{8}\begin{bmatrix}4 & 0 & 0 & 0\\ 0 & 0 & 0 & 4\end{bmatrix}$	$P_5 = \frac{1}{8}\begin{bmatrix}4 & 0 & 0 & 0\\ 0 & 0 & 0 & 4\end{bmatrix}$
6.	$P_6 = \frac{1}{8}\begin{bmatrix}1 & 1 & 1 & 1\\ 2 & 0 & 0 & 2\end{bmatrix}$	$P_7 = \frac{1}{8}\begin{bmatrix}2 & 0 & 0 & 2\\ 1 & 1 & 1 & 1\end{bmatrix}$
7.	$P_8 = \frac{1}{8}\begin{bmatrix}1 & 1 & 2 & 0\\ 1 & 1 & 0 & 2\end{bmatrix}$	$P_9 = \frac{1}{8}\begin{bmatrix}2 & 0 & 1 & 1\\ 0 & 2 & 1 & 1\end{bmatrix}$
8.	$P_{10} = \frac{1}{8}\begin{bmatrix}1 & 2 & 1 & 0\\ 1 & 0 & 1 & 2\end{bmatrix}$	$P_{11} = \frac{1}{8}\begin{bmatrix}2 & 1 & 0 & 1\\ 0 & 1 & 2 & 1\end{bmatrix}$
9.	$P_{12} = \frac{1}{8}\begin{bmatrix}1 & \frac{3}{2} & \frac{3}{2} & 0\\ \frac{3}{2} & 0 & 0 & \frac{5}{2}\end{bmatrix}$	$P_{13} = \frac{1}{8}\begin{bmatrix}\frac{3}{2} & 0 & 0 & \frac{3}{2}\\ 0 & \frac{3}{2} & \frac{3}{2} & 1\end{bmatrix}$
10.	$P_{14} = \frac{1}{8}\begin{bmatrix}2 & 0 & 0 & 2\\ 0 & 2 & 2 & 0\end{bmatrix}$	$P_{15} = \frac{1}{8}\begin{bmatrix}0 & 2 & 2 & 0\\ 2 & 0 & 0 & 2\end{bmatrix}$

$$P_i \in M_{PLOD}(\tfrac{1}{2},\tfrac{1}{2},\tfrac{1}{2}) \cap M_{PUOD}(\tfrac{1}{2},\tfrac{1}{2},\tfrac{1}{2})$$

for $i = 1,2,\ldots, 12,13$. The extreme point trivariate distribution P_{14} of $M_{PLOD}(\tfrac{1}{2},\tfrac{1}{2},\tfrac{1}{2})$ is not PUOD. The extreme point trivariate distribution P_{15} of $M_{PUOD}(\tfrac{1}{2},\tfrac{1}{2},\tfrac{1}{2})$ is not PLOD.

2. Because of the symmetry present in the probabilities $p_1 = \tfrac{1}{2} = p_2$, $q_1 = \tfrac{1}{2} = q_2$, and $r_1 = \tfrac{1}{2} = r_2$, the extreme points of $M_{PUOD}(\tfrac{1}{2},\tfrac{1}{2},\tfrac{1}{2})$ can be obtained from those of $M_{PLOD}(\tfrac{1}{2},\tfrac{1}{2},\tfrac{1}{2})$ by flipping 1 and 2 among the indices of p_{ijk}'s of p_i's, $i = 1,2,3,4,5,6,8,10,12,14$.

3. The distributions p_i's, $i = 1,2,\ldots,12,13$ are extreme points of $M_{PLOD}(\tfrac{1}{2},\tfrac{1}{2},\tfrac{1}{2}) \cap M_{PUOD}(\tfrac{1}{2},\tfrac{1}{2},\tfrac{1}{2})$.

4. If one wishes to construct a trivariate distribution P which is PLOD but not PUOD, one could use P_{14} as a building block. Look for convex combinations of P_{14} and some or all of $P_1,P_2,P_3,P_4,P_5,P_6,P_8,P_{10},P_{12}$. Any convex combination $\lambda P_1 + (1-\lambda)P_{14}$ with $0 \leq \lambda < 1$ is PLOD but not PUOD.

3. CONCLUDING REMARKS

The extreme point analysis of two natural definitions of positive quadrant dependence in three dimensions reveals that these two notions of dependence are not violently different. Extreme point of analysis is useful in evaluating the power function of any test proposed for testing independence of X,Y, and Z against strict positive quadrant dependence of X,Y, and Z. For details, in the case of 2 dimensions, see Subramanyam and Bhaskara Rao (1985). Also, certain measures of dependence can be shown to be affine functions over the sets M_{PLOD} and M_{PUOD}. This affine function property is useful to evaluate asymptotic power of tests based on these measures of dependence. All these ideas and an algebraic method for isolating extreme points of the sets M_{PLOD} and M_{PUOD} will be the subject matter of a forthcoming report.

REFERENCES

[1] AHMED, A. N., LANGBERG, N. A., LEON, R. V. and PROSCHAN, F., (1978). Two concepts of positive dependence, with Applications in Multivariate Analysis. Department of Statistics, The Florida State University, AFOSR Technical Report No: 78-6.

[2] LEHMANN, E. L., (1966). Some concepts of dependence. *Ann. Math. Stat.*, **37**, 1137-1153.

[3] NGUYEN, T. T. & SAMPSON, A. R., (1985). The geometry of certain fixed marginal probability distributions, *Linear Algebra And Its Applications*, **70**, 73-87.

[4] SUBRAMANYAM, K and BHASKARA Rao, M., (1986). Extreme point methods in the study of classes of bivariate distributions and some applications to contingency tables. *Technical Report* No. 86-12. Center for Multivaraite Analysis, University of Pittsburgh, Pittsburgh.

END
DATE
FILMED
DTIC
4/88