Chin. Ann. of Math.
4B (3) 1983

WEAK MONOTONE REGRESSION AND WEAK REGRESSION DEPENDENCE

LIANG WENGI (梁文祺) LI ZHIJIE (李志杰)
(*Jinan University*)

Abstract

This paper gives the exhaustive relations among several concepts of positive dependence for two dimensional r. v.,and two weak types of positive dependence are studied in detail

§ 1. Introduction and summary

In 1966, E. L. Lehmann discussed rather systematically the positive dependence between two random variables X and Y, i. e. roughly speaking, large values of Y tend to be associated with large values of X and vice versa. Later, Esary, Proschan, Walkup, Jogdeo and Shea continued the research on various possible definitions of positive dependence, the generalizations to the multivariate case, their properties and statistical applications. The main concepts in their investigations involved the following four definitions (Def. 1—4): (The expectations concerned are always assumed to be finite)

Definition 1. (*Regression dependence*) *X is said to be regression dependent on Y iff $P(X>x|Y)$ is a non-decreasing function of Y for any fixed x.*

Definition 2. (*Quadrant dependence*) *X and Y are said to be quadrant dependent iff $P(X>x, Y>y) \geqslant P(X>x)P(Y>y)$ for all x and y.*

Definition 3. (*Association*) *X and Y are said to be associated iff*

$$\operatorname{cov}\ (f(X, Y),\ g(X, Y)) \geqslant 0,$$

where f and g are any non-decreasing function (i. e. they are non-decreasing in each of the independent variables whenever the other one is fixed).

Definition 4. (*Monotone regression*) *X is said to have a monotone regression on Y iff $E(X|Y)$ is a non-decreasing function of Y.*

In the present paper we shall discuss the following two concepts (Definition 5 and 6) in detail.

Denote the support of Y by S_Y.

Definition 5. (*Weak regression dependence*) *X is said to be weak regression dependent on Y iff for every fixed x, $P(X>x|Y>y)$ is a non-decreasing function of y, where $y<\sup S_Y$.*

Manuscript received November 16, 1981.

Definition 6. (*Weak monotone regression*) *X is said to have a weak monotone regression on Y iff it possesses one of the following properties:*

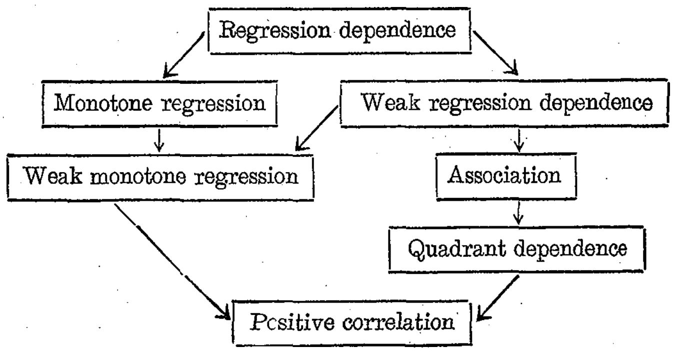
(a) *$E(X|Y<y)$ is a non-decreasing function of y, where $y>\inf S_Y$.*

(b) *$E(X|Y>y)$ is a non-decreasing function of y, where $y<\sup S_Y$.*

Finally we mention the most well-known concept (Definition 7).

Definition 7. (*Positive correlation*) *X and Y are said to be positive correlated iff* $\operatorname{cov}(X, Y)\geqslant 0$.

The relations among the foregoing concepts can be represented by the following lattice:



This lattice is complete, i. e., no other arrows can be put in.

§ 2. Weak Monotone Regression

Theorem 1. *If $E(X|Y)$ is non-decreasing in Y, then $E(X|Y<y)$ and $E(X|Y>y)$ are both non-decreasing in y.*

For a proof, it suffices to notice that when A belongs to some Borel field F and $P(A)>0$, then $E(X|A)$ is the mean value of integral of X or $E(X|F)$ over A.

An inequality of Chebyshev. If

(a) μ is a measure on R;

(b) C is a measurable set in R;

(c) u and v are non-decreasing functions on C, then

$$\mu(C)\int_C uv\,d\mu\geqslant\int_C u\,d\mu\int_C v\,d\mu.$$

This will be useful in proving Theorem 2.

Theorem 2. *$E(X|Y>y)$ is a non-decreasing function of $y\Rightarrow$*

$$\operatorname{cov}(X, Y)\geqslant 0.$$

Proof Let

$$\eta=\begin{cases} Y, & \text{if } |Y|\leqslant K, \\ K, & \text{if } Y>K, \\ -K, & \text{if } Y<-K, \end{cases}$$

where $K>0$, and denote $\alpha=\inf S_\eta$, $\beta=\sup S_\eta$. Then $[\alpha, \beta]\subset[-K, K]$. Divide

$[\alpha, \beta]$ into m subintervals of equal length by means of points

$$\alpha=\alpha_1<\alpha_1<\cdots<\alpha_m=\beta.$$

If $P(\eta\in(\alpha_{i-1}, \alpha_i])=0(i=2, 3, \cdots, m-1)$, then we delete α_{i-1} and redenote the points remained by $\alpha=a_0<a_1<\cdots<a_n=\beta$. Let

$$A_1=(Y\leqslant a_1)=(\eta\leqslant a_1),$$
$$A_i=(Y\in(a_{i-1}, a_i])=(\eta\in(a_{i-1}, a_i]), \ i=2, 3, \cdots, n-1,$$
$$A_n=(Y>a_{n-1})=(\eta>a_{n-1}).$$

Clearly all these sets have positive probabilities.

By virtue of Chebyshev's inequality and the condition of the theorem, we have

$$E(X)E(\eta)\leqslant\int_{-\infty}^{\infty} y\left\{I_{(-\infty, a_1]}(y)\frac{E(I_{A_1}X)}{P(A_1)}+I_{(a_1, \infty)}(y)\frac{E(I_{A_1^c}X)}{P(A_1^c)}\right\}dF_\eta(y),$$

where F_η is the probability distribution function of η and I stands for indicator. Similarly

$$\int_{(a_1, \infty)} y\, dF_\eta(y)\frac{E(I_{A_1^c}X)}{P(A_1^c)}\leqslant\int_{(a_1, \infty)} y\left\{I_{(a_1, a_2]}(y)\frac{E(I_{A_2}X)}{P(A_2)}\right.$$
$$\left.+I_{(a_2, \infty)}(y)\frac{E(I_{(A_1\cup A_2)^c}X)}{P(A_1\cup A_2)^c}\right\}dF_\eta(y).$$

and so on. Finally we obtain

$$E(X)E(\eta)\leqslant\sum_1^n E(I_{A_i}\eta)E\left(\frac{I_{A_i}X}{P(A_i)}\right).$$

On the other hand

$$E(X\eta)=\lim_{m\to\infty}E\left[\left(\sum_1^n I_{A_i}a_{i-1}\right)X\right]=\lim_{m\to\infty}\sum_1^n a_{i-1}E(I_{A_i}X),$$

and

$$\left|\sum_1^n E(I_{A_i}\eta)E\left(\frac{I_{A_i}X}{P(A_i)}\right)-\sum_1^n a_{i-1}E(I_{A_i}X)\right|\leqslant\sum_1^n\left|E\left(\frac{I_{A_i}\eta}{P(A_i)}\right)\right.$$
$$\left.-a_{i-1}\right|E(I_{A_i}|X|)\leqslant\frac{2K}{m}E|X|.$$

Therefore, $E(X)E(\eta)\leqslant E(X\eta)$. Let $K\to\infty$, we obtain $E(X)E(Y)\leqslant E(XY)$. This completes the proof.

Obviously, Theorem 2 can be proven in the same way if the assumption of the theorem is replaced by that $E(X|Y<y)$ is a non-decreasing function of y.

§ 3. Weak Regression Dependence

Theorem 3. *$P(X>x|Y)$ is non-decreasing in $Y\Rightarrow P(X>x|Y>y)$ is non-decreasing in y.*

Proof Since $P(A|\cdot)=E(I_A|\cdot)$, Theorem 3 is merely a special case of Theorem 1.

Theorem 4. *For each fixed x, $P(X>x|Y>y)$ is non-decreasing in $y\Leftrightarrow$ For any non-decreasing function $f(X)$, $E(f(X)|Y>y)$ is non-decreasing in y.*

Proof Note that $P(X\geqslant x|Y>y)=\lim_{n\to\infty}P\left(X>x-\frac{1}{n}\middle|Y>y\right)$, which is non-decreasing in y. Let $A_i=\left(x:f(x)>\frac{i}{k}\right)$ and $f_i(X)=I_{A_i}(X)$, $i=1, 2, \cdots$, where k is any positive integer. If f is non-negative, then $E(f(X)\,|Y>y)$ is non-decreasing in y since

$$E(f(X)\,|Y>y)=\lim_{k\to\infty}E\left(\frac{1}{k}\sum_{i=1}^{k^2}f_i(X)\,|Y>y\right).$$

For a general function f, we put

$$f_c(x)=\begin{cases}f(x), & \text{if } f(x)\geqslant c,\\ c, & \text{if } f(x)<c.\end{cases}$$

Since f_c-c is non-negative and non-decreasing, $E(f_c(X)-c|Y>y)$ is non-decreasing in y, and therefore $E(f_c(X)\,|Y>y)$ is non-decreasing in y. Hence $E(f(X)\,|Y>y)$ $=\lim_{c\to-\infty}E(f_c(X)\,|Y>y)$ is also non-decreasing in y. This completes the proof.

Theorem 4 says that weak regression dependence implies weak monotone regression.

Theorem 5. *$P(X>x|Y>y)$ is a non-decreasing function of y when x is fixed $\Rightarrow$ X and Y are associated.*

As shown in [2], in order to prove that X and Y are associated, it suffices to prove that

$$\operatorname{cov}\ (f(X,\ Y),\ g(X,\ Y))\geqslant 0$$

for any non-decreasing functions f and g which only take values 0 and 1. In fact, it can easily be seen from Hoeffding's identity

$$\operatorname{cov}\ (X_1,\ X_2)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\operatorname{cov}\ (I_{\{X_1>x_1\}},\ I_{\{X_2>x_2\}})\,dx_1\,dx_2.$$

Any non-decreasing function $f(x,\ y)$ which takes only values 0 and 1 can be written as an indicator I_A. We shall denote the class of all these sets A by $\mathscr{A}$ and denote the boundary of A by A_B, and $A_{B_1}=A\cap A_B$, $A_{B_2}=A_B-A_{B_1}$. For each point $a(x_0,\ y_0)$ we set

$$Q_a=((x,\ y):x>x_0,\ y>y_0),$$
$$\bar{Q}_a=((x,\ y):x\geqslant x_0,\ y\geqslant y_0).$$

Before proving Theorem 5, let us deduce some lemmas. For the sake of brevity, we shall write $P(A)$ for $P((X,\ Y)\in A)$ in what follows.

Lemma 1. *$A\in\mathscr{A}$ iff one of the following conditions holds:*

(a) *A can be written as $A=(\bigcup_a\bar{Q}_a)\cup(\bigcup_{a'}Q_{a'})$, in particular*

$$A=\bigcup_{a\in A}\bar{Q}_a=(\bigcup_{a\in A_{B_1}}\bar{Q}_a)\cup(\bigcup_{a'\in A_{B_2}}Q_{a'}),$$

(b) *If the coordinate system x_y is rotated through an angle of $-45°$, then with respect to the resulting coordinate system $x'-y'$ we have*

(i) *A_B is a single valued continuous curve: $y'=\varphi(x')$, $-\infty<x'<\infty$;*

(ii) *Each chord of this curve has a slope $\in[-1,\ 1]$;*

(iii) *All points above this curve must belong to A and those below the curve must not belong to A.*

The proof is straightforward.

Lemma 2. *For any $A\in\mathscr{A}$, there must be some open set $A_k\in\mathscr{A}$ such that*

$$I_{A_k}(X, Y)\xrightarrow{p} I_A(X, Y).$$

Proof Take an open set $G\supset A_{B_1}$, such that $P(G-A_{B_1})<\frac{1}{k}$. Clearly, A_{B_1} contains at most countably many straight segments $l_1, l_2, \cdots$, which are parallel to the x-axis or y-axis. Let $Q_{a_n}\supset l_n$ such that

$$P(\bigcup_n(Q_{a_n}-A))<\frac{1}{k}.$$

For each $b\in A_{B_1}-\bigcup_n l_n$, take a circula r neighbourhood U_b of b such that $U_b\subset G$. Let b_1 and b_2 be the intersection points of A_B and the boundary of U_b. Choose point b'_1 from $\widehat{b_1b}$ and b'_2 from $\widehat{bb_2}$ on A_B and point b' such that

(i) b'_1 and b'_2 are on the boundary of $Q_{b'}$;

(ii) $Q_{b'}\supset\bar{Q}_b$ and $Q_{b'}-A\subset U_b$.

Now put

$$A_k=(\bigcup_n Q_{a_n})\cup(\bigcup_{b\in A_{B1}-\cup l_n}Q_{b'})\cup(\bigcup_{a\in A_{B_2}}Q_a),$$

then obviously A_k is an open set in $\mathscr{A}$ and

$$I_{A_k}(X, Y)\xrightarrow{p} I_A(X, Y).$$

This completes the proof.

Lemma 3. *For any open set $A\in\mathscr{A}$, there is an open set $A^{(n)}\in\mathscr{A}$ such that*

(a) $A^{(n)}\subset A$;

(b) $A^{(n)}_B$ *consists of finite segments which are parallel to the x-axis or y-axis;*

(c) $P(A-A^{(n)})\to 0$.

Proof Choose $a'\in A_B$ and $a''\in A_B$ such that a' and a'' are contained in the boundary of Q_a for which

$$P(A-Q_a)<\frac{1}{n}.$$

Take $a'=a_0, a_1, a_2, \cdots, a_N=a''$ on the arc $\widehat{a'a''}$ of A_B. When the division is fine enough

$$P(A-\bigcup_i Q_{a_i})<\frac{2}{n}.$$

Clearly the open set $A^{(n)}=\bigcup_i Q_{a_i}\subset A$ and $A^{(n)}\in\mathscr{A}$. This completes the proof.

The following two lemmas can be easily proved.

Lemma 4. *Let*

$$A=((x, y):x>x_0,\ y>y_0),$$
$$B=((x, y):x\leqslant x_0,\ y>y_0),$$
$$C=((x, y):x>x_0,\ y_1<y\leqslant y_0),$$
$$D=((x, y):x\leqslant x_0,\ y_1<y\leqslant y_0),$$

where $y_1 < y_0$. *Then* $P(X>x|Y>y)$ *is non-decreasing in* y *when* x *is fixed iff* $P(A)P(D) \geqslant P(B)P(C)$ *for all* x_0 *and* $y_0 > y_1$.

Lemma 5. *X and Y are associated iff* $P(A)P(D) \geqslant P(B)P(C)$ *for*

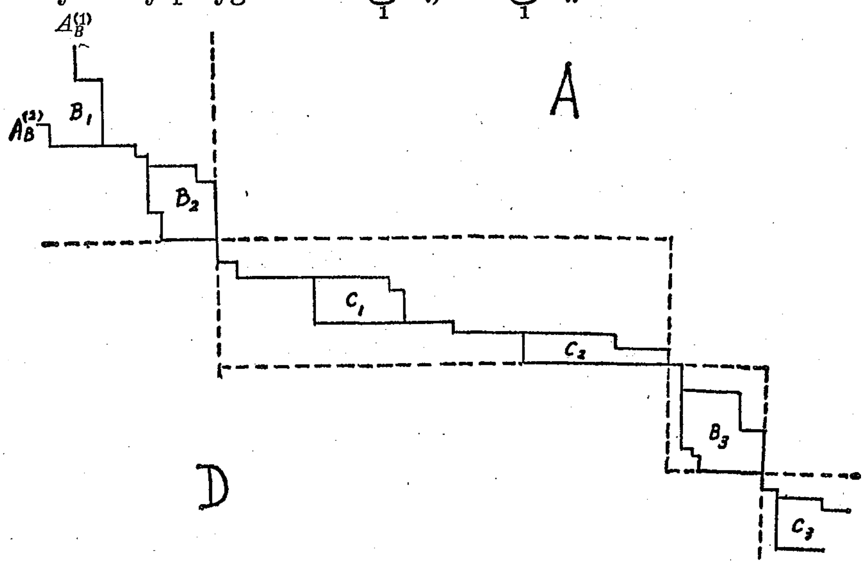
$$A = A^{(1)} \cap A^{(2)},\ B = A^{(2)} - A^{(1)},$$
$$C = A^{(1)} - A^{(2)},\ D = (A^{(1)} \cup A^{(2)})^c,$$

where $A^{(1)}$ *and* $A^{(2)}$ *are open sets in* $\mathscr{A}$.

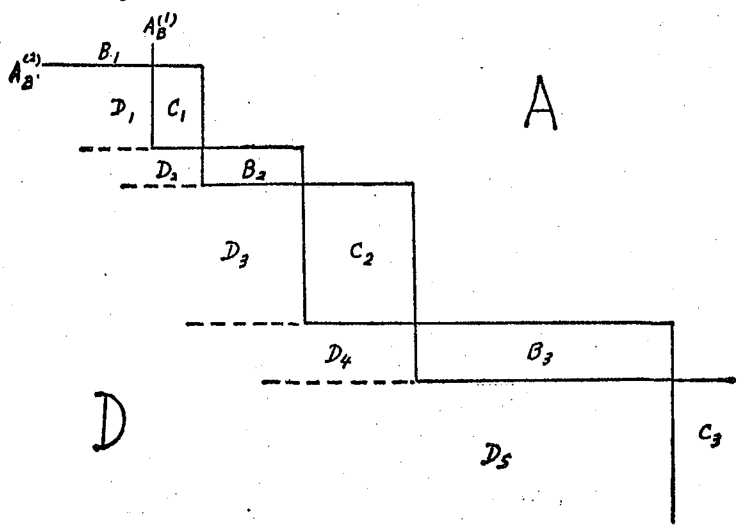
Proof of Theorem 5 In order to prove that X and Y are associated, we only need to prove the inequality in Lemma 5, i. e.

$$P(A)P(D) \geqslant P(B)P(C).$$

According to Lemmas 2 and 3, we can only pay attention to the case that $A^{(1)}$, $A^{(2)}$ are both open sets and $A_B^{(1)}$, $A_B^{(2)}$ are both broken lines consisting of finitely many segments parallel to x-axis or y-axis. Generally speaking, in this case, both B and C consist of finitely many polygons: $B = \bigcup_1^n B_i$, $C = \bigcup_1^m C_i$.



In order to simplified the structure, the B_i 's are appropriately combined and enlarged into several rectangles and so are the C_i 's. Thus B, C are enlarged and A, D are shrunk as shown below by the broken lines.



If the inequality remains valid in this case, a fortiori, the theorem is true. Therefore we only need to consider this normal case. By virtue of Lemma 4, it is easy to see that

$$P(A)P(D_1)\geqslant P(B_1)P(C_1),$$
$$P(A)P(D_2)\geqslant (P(B_1)+P(C_1))P(B_2),$$
$$P(A)P(D_3)\geqslant (P(B_1)+P(C_1)+P(B_2))P(C_2),$$
$$\cdots\cdots\cdots\cdots\cdots\cdots$$

By adding these inequalities togather we obtain

$$P(A)P(D)\geqslant\sum_{i,j}P(B_i)P(C_j)=P(B)P(C).$$

The proof is completed.

§ 4. Counterexamples

Consider the probability distributions of (X, Y) given in the following tables.

$y \backslash x$	0	1	2
2			0.8
1	0.1		
0		0.1	

Table 1

$y \backslash x$	0	1	2
2		0.1	
1			0.1
0	0.8		

Table 2

$y \backslash x$	0	1	2
2			0.1
1		0.1	
0	0.7		0.1

Table 3

$y \backslash x$	0	1	2
2	0.1		0.5
1		0.1	
0	0.2		0.1

Table 4

$y \backslash x$	0	1	2	3
3				0.4
2		0.05		
1		0.1	0.05	
0	0.4			

Table 5

$y \backslash x$	0	1	2	3	4	5
3	0.1					0.3
2					0.1	
1		0.1				
0	0.3					0.1

Table 6

(1) In Table 1, $E(X|Y>y)$ is non-decreasing in y, but $E(X|Y<y)$ is not monotone in y. In Table 2, $E(X|Y<y)$ is non-decreasing in y, but $E(X|Y>y)$ is not monotone in y. These demonstrate that the two conditions in the definition of weak monotone regression are not identical.

(2) In Table 3, X and Y are weak regression dependent, and X has a monotone regression on Y, but X is not regression dependent on Y.

(3) In Table 4, X and Y are associated, and X has a monotone regression on Y,

but X is not weak regression dependent on Y.

(4) In Table 5, X is weak regression dependent on Y, and the two conditions of weak monotone regression are both satisfied, but X has no monotone regression on Y.

(5) In Table 6, X and Y are associated, but X has no monotone regression on Y.

(6) It has already been proved in [2, 5] that monotone regression does not imply quadrant dependence and quadrant dependence does not imply assotiation. The fact that positive correlation does not imply weak monotone regression or quadrant dependence is evident.

The foregoing counterexamples demonstrate that the lattice given in section 1 is complete.

References

[1] Lehmann, E. L., Some Concapts of Dependence *Ann. Math. Statist*, **37** (1966), 1137—1153.

[2] Esary, J. D., Proschan, F. and Walkup, D. W., Association of random variables with Applications, *Ann. Math. . Statist*, **38** (1967), 1466—1474.

[3] Jogdeo, K., Characterization of Independence in Certain Families of Bivariate and Multivariate Distributions, *Ann. Math. Statist*, **39** (1968), 433—441.

[4] Jogdeo, K., Association and Probability Inequalities, *Ann. Statist*, **5** (1977), 496—504.

[5] Shea, G., Monotone Regression and Covarirnce Structure, *Ann. Statist*, **7** (1979) 1121—1126.