Ing. SRBOLJUB STOJANOVIĆ

# JEDNA NESIMETRIČNA MERA STOHASTIČKE ZAVISNOSTI
## - MOGUĆNOST IDENTIFIKACIJE DOMINACIJE -

Center za proučevanje sodelovanja z deželami v razvoju,
Ljubljana

Ljubljana, december 1973

---

SRBOLJUB STOJANOVIĆ

# AN ASSYMMETRIC MEASURE FOR STOHASTIC DEPENDENCE
## - A POSSIBILITY FOR REVEALING DOMINATION -

Research Center for Co-operation with Developing Countries,
Ljubljana

Ljubljana, December 1973

| Serija: Aktuelne teme<br>Program serije u g. 1973/74: | Series: Topics of Our Time<br>Program of the Series in 1973/74: |
|---|---|
| 1. mag Milica Žiberna | "Opšta šema carinskih preferencijala i položaj zemalja u razvoju"<br>(The Generalized System of Preferences and the Position of Developing Countries) |
| 2. dr Zoran Trputec | "Spoljna trgovina i nacionalna privreda - dileme uključivanja u medjunarodnu podelu rada"<br>(Foreign Trade and National Economy - Dilemmas Concerning International Division of Labour) |
| 3. dr Miodrag Sukijasović | "Privredni ugovori Jugoslavije - Opšte medjunarodno privredno pravo primenjeno na privredne ugovore Jugoslavije"<br>(Economic Treaties of Yugoslavia - General International Economic Law Applied to Economic Treaties of Yugoslavia) |
| 4. Srboljub Stojanović (bilingual: Serbian-English) | "Jedna nesimetrična mera stohastičke zavisnosti - mogućnost identifikacije dominacije"<br>"An Asymmetric Measure of Stochastic Dependence - A Possibility for Revealing Domination" |
| 5. Srboljub Stojanović, dr Zoran Trputec | "Opća teorija razvoja i koncept ekonomskog razvoja i nerazvoja"<br>(General Development Theory and the Concepts of Economic Development and Under-development |

- Studies under Nos 1, 2, 3, and 5 have comprehensive English summaries -

Centar zahvaljuje Jugobanci na sufinansiranju ove studije.

A grant by Jugobanka for this study is hereby gratefully acknowledged.

| | |
|---|---|
| Izdavač:<br>Publisher: | Center za proučevanje sodelovanja z deželami v razvoju, Ljubljana<br>61 000 LJUBLJANA, Hajdrihova 16, tel.: 61-591 |
| Žiro račun: | 50103-603-50753, SDK Ljubljana |
| Foreign currency account: | 50100-620-107-858, LB Ljubljana |
| Za izdavača:<br>Director: | v.d. direktora Jože Zemljak |
| Urednik publikacija:<br>Managing editor: | Boris Cizelj |

Recenzent: prof dr Ljubomir Martić
Reviewer:

Prevodioc: dr Branko Kronfeld
Translation:

Korektor: V. Kovačič
Proof-reader:



Razmnoževanje:
Off-set multiplication: Janez Pleško, Rožna dolina c. IV/36, Ljubljana

CONTENTS - SADRŽAJ

# P R E F A C E

The folowing methodological research is rather general, giving mathematical interpretation to relations of asymmetry feasible between elements in any system. Being based on the theory of information, it may find applications in all social and natural sciences.

The science of economy, in particular problems related to economic development and underdevelopment, give ample opportunities for applications of these methodological instruments. From the very beginnig the Marxist approach to economic phenomena developed within a frame of treating economical processes, in given social and historical conditions (in the scheme of capitalist mode of production), as being cumulative and polarizatory, permanently emphasizing opposing interests of various economic agents - producers, consumers, managers. Precisely that unity of contradictions, characterized by asymmetry among economic agents, makes possible an appropriation and utilization of surplus value produced by others. Thus capitalist system presents itself as a socio-economic system in which processes of concentration and polarization are constantly present.

The bourgeois economic science however developed later in another direction. Mostly just in order to counteract the marxist economic theory, and trying to construct theoretical solutions logically explaining operation of contemporary capitalist economies as well, gradually the neoclassical approach has been formulated. In this approach, based on a sequence of scientifically anacceptible axioms[1], a distinguished place has the principle of circular determinizm, based on symmetrical and reversible relations among elements of the system[2]. Such a nature of relations among elements of the system expressed itself in the mathematical instruments used, through equations of general equilibrium.

---

1) For a systematical exposition of critiques of the neoclassical approach and in particular of the model of general economic equilibrium, see: Z. Trputec, Foreign trade and national economy, Developing countries research center, Ljubljana, 1973, part II.

2) Principle of circular determinism is expounded in: V. Pareto, "Le nuove teorie economiche", Giornale degli Economisti, Vol. II, 1901, p. 238.

For a theoretical construction in a bourgeois economic science to be accepted as scientific it is not necessary that it has been verified empirically but it is enough that it meets principles of formally logical exactness, rigour and validity.[1] In such a way the hypothesis of symmetrical relations between economically relevant elements became scientific. This theoretical premise is closely related to the concept of equilibrium spontaneously being established in an economic system satisying certain assumptions, in particular the assumption of perfect competition.

Combining these elements within frame of a system in general equilibrium, a basical assumption of bourgeois economy, the assumption of harmony, emerged as completely scientific. Motivated by individual and egoistic interests, economic agents in a capitalist system guided by an "invisible hand" attain the optimal state of the system. The idea of symmetry in the process of exchange lies in the basis of the well known Pareto's optimum representing even today one of the basic principles of the modern theory of allocation of resources. Long ago, however, even some bourgeois economists pointed out that Pareto's optimum requires accepting the principle of symmetrical relations between partners in the process of exchange, and if this principle in reality does not work, the perfect competition does not ensure attaining the optimal state.[2] This objection was however ignored since exactly the perfect competition was always declared to ensure the symmetry in relations among partners in various economic processes. The symmetry in relations among elements in a system of general economic equilibrium guarantees at the same time that there are no cumulative effects of these processes; this instigating the persuasion of bourgeois economists that embedding economic processes through necessary changes of assumptions into a statical scheme of general equilibrium leads to "dynamic equilibrium". The dynamiza-

1) A critique of bourgeois interpretation of the process of founding sciences explaining why the above numbered principles are considered sufficient may be found in: S. Stojanović and Z. Trputec, "Model as a part of scientifically acquired knowledge", Economic institute Zagreb, Zagreb 1972.

2) A. Marchal shows in his critiques of Pareto's approach how an unequal influence of partners in the process of exchange may lead to results entirely different from those which predicts Pareto (A. Marshal, Principles of Economics, London, 1962, p. 652-654).

tion of models of general equilibrium by introducing the economic time leads in fact to inconsistencies within the model itself, in fact to its destruction, even accepting symmetrical relations among its elements. However, this problem is only indirectly connected with the problem of symmetry in economic theory.[1] The central question of a possibility of application of the principle of symmetry of relations among elements in an economic system remains still unsolved.

Severe atacks on this principle came at first in the economic theory indirectly. G. Myrdal, in the forties of this century, atacked the principle of stable equilibrium in a social or economic system pointing out that the really scientific approach must treat economic processes as cumulative.[2] De facto this implicitely insinuates economic relations not being symmetric. It is not at all a matter of chance that external critiques of the neoclassical approach, in particular critiques of premises foundating the theory of general equilibrium, became more frequent and more serious precisely with the appearance of the problem of economic development and underdevelopment. They come from those theoreticians which in their field of exploration find the reality astonishingly discrepant with theoretical assumptions of symmetry, harmony and equilibrium. Here we have in mind first and foremost those theoreticians which began studying these phenomena after the underdeveloped countries had entered the world scene. G. Myrdal himself was one of those economists whose scientific attention was attracted by the phenomenon of the underdeveloped countries.

Approximately at the same time critical objections to the neoclassical approach came from F. Perroux.[3] He shows that the contemporary

1) If introducing dynamic elements into the theory of general equilibrium destroys this theoretical construction, as shown by the Italian economic school (especially in the works of G. di Nardi, A. Graziani, P. Garegnani and B. Tosato), then in creating an alternative approach considerably more premises are to be changed then their critics require. This relates in particular to the principle of symmetrical relations among economically relevant phenomena.
2) A critique of the "vicious circle" and emphasis on the principle of cumulativity in social processes can be found in: G. Myrdal, An American Dilemma, the Negro Problem and Modern Democracy, New York, 1944, and in later works of the same author relating the problem of underdeveloped countries.
3) Of a special importance with respect to these problems is the paper: F. Perroux, "Esquisse d´une théorie de l´économie dominante", Economie appliquée, No. 2-3, 1948.

economic relations are characterised by domination and that the economic relations are asymmetric and not reciprocal, contrary to assumptions of the neoclassical school, and that the economic processes are cumulative and ireversible. In this author´s work a certain measure of tolerance and a desire for a better understanding with the marxist economic thought is felt, but also a certain displeasure, especially regarding methodology, with the marcistic stock of knowledge.[1) F. Perroux and his followers in France developed an extremely interesting and a broad spectrum of methodological instruments in trying to explain contemporary economic reality in a way entirely different from that present in the remaining western economic theory.

Especially interesting is Perroux´s concept of domination which he explains as follows: "...we can say that A accomplishes an effect of domination over B if, disregarding any special purpose of A, A has influence upon B while there is no reciprocity or there is not enough reciprocity."[2)

Theoretical assumptions connected with concepts of asymmetry, cumulativity and polarization were developed in his aproach in great detail and even applied in numerous empirical investigations. A large number of objections on Perroux´s approach is possible from the marxist point of view, but the majority of his suppositions do not contradict marxist point of view on contemporary socio-economic relations and processes. Thus many young marxists in underdeveloped countries, free from dogmatic prejudices, accepted Perroux´s reasonings as an integral part of advanced modern economic theory.[3) F. Perroux states that the influence the dominant unit of production exercises upon the dominated unit of production is realized through material flows, prices, innovations and anticipations, with help of which factors the dominating unit modifies the relevant surrounding of the do-

---

1) Some observations regarding mode and extent to which marxists treat this problem gave this author in his next work. Marxistic philosopher G. Granger pointed once out that a number of his assumptions already much earlier were formulated in the works of Marx (see G. Granger, "Sur la définition do progres par F. Perroux", Cahiers de l' I.S.E.A., no. 119, november 1961).

2) F. Perroux, op.cit., p. 30.

3) See the arguments of S. Amin on manv places in his book <u>L´accumulation à l´échelle mondiale,</u> Paris, 1970.

minated unit as well as its internal structure. The general language of the information theory S. Stojanović uses hits exactly these problems and therefore enriches the whole economic theory with mathematical instruments it did not posess yet. Verbally and by way of simbols the concept of domination in the above approaches is already developed. However, what the neoclassical theory was supplied with - a rigorously formulated mathematical model of symmetric relations - the theory of domination was lacking. Hence the frequent objection of such theories not being scientifically founded. In psychometrical research the methods revealing asymmetry are already in use, but in most social scientific explorations the methods reflecting symmetric relations are still prevailing. The present short methodological work guided by marxist ideas not accepting apriori equivalence, order and harmony as basic features of social relations in general, making use of the newest concepts of the mathematical theory of information represents definitely a contribution to the modern theory of domination in socio-economic field. Its applications, as we already said, are not limited with this and there are wide possibilities in other domains as well. For us on this place it is however especially the field of economic relations which is important, and in particular the problems of economic development and underdevelopment; rhe present study supplying methodological instruments which might enable further improvement and progress in the theory of economic development.

Ljubljana, June, 1973.                                        dr Zoran Trputec

# AN ASYMMETRIC MEASURE FOR STOCHASTIC DEPENDENCE

## A possibility for revealing domination

## 1. Introduction

A frequent subject of exploration in various scientific disciplines are states, phenomena or processes with mutual dependence being inherently stochastic, i. e. states, phenomena or processes which are said to be stochastically dependent. Various measures for stochastic dependence in applications, like coefficients of correlation, coefficients of range correlations, coefficients of contingency etc. are in general symmetrical; which reflects an implicite assumption that one state, phenomenon or process influences the other in exactly the same measure as the second one influences the first.

In this paper we shall introduce an asymmetrical measure for stochastic dependence and we shall show on an example how it is evaluated in practice. We shall use some known facts and concepts from the statistical theory of information, in particular the concept of entropy and the concept of quantity of information.

## 2. Notation and preliminary definitions

In defining our measure for stochastic dependence between two states, phenomena or processes, we shall assume that the states, phenomena or processes we are interested in can be described by finite sequences of discrete random variables.

We shall use the following notation:

X - random variable with possible values

$x_i = x_1, x_2, \ldots, x_m$.

Y - random variable with possible values

$y_j = y_1, y_2, \ldots, y_n$.

$p(x_i)$ - marginal probability for X to assume the value $x_i$, idependently of what happens with Y; $(i = 1, 2, \ldots, m)$

$p(y_j)$ - marginal probability for Y to assume the value $y_j$, independently of what happens with X; $(j = 1, 2, \ldots, n)$

$p(y_j/x_i)$ - conditional probability of Y assuming the value $y_j$, under condition $X = x_i$, $(i = 1, 2, \ldots, m;\ j = 1, 2, \ldots, n)$.

$p(x_i/y_j)$ - conditional probability of X assuming the value $x_i$, under condition $Y = y_j$, $(i = 1, 2, \ldots, m;\ j = 1, 2, \ldots, n)$.

The joint probability will be denoted by

$p(x_i, y_j)$ - the probability of $X = x_i$ and simultaneously $Y = y_j$ $(i = 1, 2, \ldots, m,\ j = 1, 2, \ldots, n)$.

These probabilities define the joint probability law for the pair (X, Y).

It follows from the definitions that

(1) $$p(x_i, y_j) = p(y_j, x_i)$$

There always hold the following relations:

(2) $$\sum_{i=1}^{m} p(x_i) = 1$$

(3) $$\sum_{j=1}^{n} p(y_j) = 1$$

(4) $$\sum_{j=1}^{n} p(y_j/x_i) = 1, \quad (i = 1, 2, \ldots, m)$$

(5) $$\sum_{i=1}^{m} p(x_i/y_j) = 1, \quad (j = 1, 2, \ldots, n)$$

(6) $$p(x_i) = \sum_{j=1}^{n} p(x_i, y_j), \quad (i = 1,2,\ldots,m)$$

(7) $$p(y_j) = \sum_{i=1}^{m} p(x_i, y_j), \quad (j = 1,2,\ldots,n)$$

Combing (6) and (2), or (7) and (3) we get

(8) $$\sum_{i=1}^{m} \sum_{j=1}^{n} p(x_i, y_j) = 1$$

There hold also the following relations:

(9) $$p(y_j/x_i) = \frac{p(x_i, y_j)}{p(x_i)}; \quad (i = 1,2,\ldots,m; \; j = 1,2,\ldots,n);$$

(10) $$p(x_i/y_j) = \frac{p(x_i, y_j)}{p(y_j)}; \quad (i = 1,2,\ldots,m; \; j = 1,2,\ldots,n);$$

or

(11) $$p(x_i, y_j) = p(x_i) \cdot p(y_j/x_i); \quad (i = 1,2,\ldots,m; \; j = 1,2,\ldots,n);$$

(12) $$p(x_i, y_j) = p(y_j) \cdot p(x_i/y_j); \quad (i = 1,2,\ldots,m; \; j = 1,2,\ldots,n).$$

A usual measure for indefiniteness of a random variable, i.e. for uniformity of probability distribution of that random variable, is entropy. It is defined as the sum of products of probabilities with the logarithms of their reciprocal values. The common unity for this measure is bit, corresponding to the logarithms with the base 2, binary logarithms. In this paper such logarithms will be denoted by ld, and cologarithms by cold; obviously we'll have:

(13) $$\text{cold } a = \text{ld } \frac{1}{a}$$

(14) $$\text{cold } a = -\text{ld } a$$

Entropies of $X$ and $Y$ will be denoted by:

$H(X)$ - unconditional entropy of $X$.

$H(Y)$ - unconditional entropy of $Y$.

They are defined by

(15) $$H(X) = \sum_{i=1}^{m} p(x_i) \operatorname{ld} \frac{1}{p(x_i)} = - \sum_{i=1}^{m} p(x_i) \operatorname{ld} p(x_i);$$

(16) $$H(Y) = \sum_{j=1}^{n} p(y_j) \operatorname{ld} \frac{1}{p(y_j)} = - \sum_{j=1}^{n} p(y_j) \operatorname{ld} p(y_j);$$

or, with regard to (13) or (14):

(17) $$H(X) = \sum_{i=1}^{m} p(x_i) \operatorname{cold} p(x_i);$$

(18) $$H(Y) = \sum_{j=1}^{n} p(y_j) \operatorname{cold} p(y_j).$$

## 3. Quantity of information as a measure for stochastic dependence

If the random variables $X$ and $Y$ were functionally dependent in a unique way, then knowing which value $(x_i)$ is assumed by $X$, should enable us to determine with certainty the corresponding value $(y_i)$ which is assumed by $Y$. The conditional probability for $Y$ to assume the value $y_i$ would then equal 1:

(19) $$p(y_i/x_i) = 1; \quad (i = 1,2,\dots,m).$$

Analogously, in case of a functional dependence the conditional probability for $X$ to assume the value $x_i$, under the condition $Y = y_i$, is also equal one:

(20) $$p(x_i/y_i) = 1; \quad (i = 1,2,\dots,m).$$

It follows that in case of a functional one-one dependence between two random variables X and Y, knowing the value which X assumes gives enough information to eliminate all entropy (indefiniteness) of Y. In this case the quantity of information about Y which is contained in X equals the unconditional entropy of Y:

(21) $I(X, Y) = H(Y).$

I(X,Y) here denotes the quantity of information contained in the random variable X which is relevant to the random variable Y.

On this place we have to note that in case of functional dependence the sequences of (marginal) probabilities for X and Y are identical and correspond to the joint probability distribution of X and Y. Both sequences then have the same number of terms:

(22) $m = n$

It follows also:

(23) $p(x_i) = p(y_i) = p(x_i, y_i); \; (i = 1, 2, \ldots, m);$

(24) $H(X) = H(Y).$

We shall return to this very special case later and in the following section, after having defined our measure for stochastic dependence; now we shall reconsider the case of stochastic dependence.

If we know that X assumes the value $x_i$, the entropy of Y in general does not vanish and is equal to:

(25) $$H(Y/x_i) = \sum_{j=1}^{n} p(y_j/x_i) \text{ cold } p(y_j/x_i); \; (i = 1, 2, \ldots, m).$$

Thus the conditional entropy is expressed in terms of (known) conditional probabilities $p(y_j/x_i)$.

$H(Y/x_i)$ is said to be the conditional entropy of Y, under condition $X = x_i$.

This entropy can be smaller, equal to or even greater than the unconditional entropy:

(26) $H(Y/x_i) \gtreqless H(Y); \; (i = 1,2,\ldots,m).$

Since the random variable X can assume any of values $x_i$ with corresponding probabilities $p(x_i)$, as a measure for indefiniteness of Y, given the value $x_i$ assumed by X in some process or in some phenomenon, we shall take the following expression:

(27) $$H(Y/X) = \sum_{i=1}^{m} p(x_i) \, . \, H(Y/x_i),$$

H(Y/X) is said to be the conditional entropy of the random variable Y, if it is known what value $x_i$ is assumed by the random variable X.

Combining (25) and (27) gives

(28) $$H(Y/X) = \sum_{i=1}^{m} \sum_{j=1}^{n} p(x_i) \, p(y_j/x_i) \text{ cold } p(y_j/x_i)$$

Since the conditional entropy cannot exceed the unconditional entropy, there always holds the unequality

(29) $H(Y/X) \leqslant H(Y)$

The equality holds if and only if the random variables X and Y are independent, so that knowing the value assumed by one of the variables does not decrease the indefiniteness of the other one. In case of a functional dependence the conditional entropy will vanish, as deducible from (19) and (28):

(30) $$p(y_i/x_i) = 1 \Rightarrow \text{cold } p(y_i/x_i) = 0 \Longrightarrow H(Y/X) = 0; (i = 1,2,\ldots,m).$$

The difference between the unconditional and the conditional entropy gives, by definition, the amount of information:

(31) $$I(X,Y) = H(Y) - H(Y/X)$$

Hence the amount of information contained in the random variable X relevant to the random variable Y is defined as the average decrease of entropy of the random variable Y if the value assumed by X is known.

Inserting (18) and (28) into (31) after a few simple algebraic transformations gives:

$$(32) \qquad I(X,Y) = \sum_{i=1}^{m} \sum_{j=1}^{n} p(x_i, y_j) \cdot \text{cold} \frac{p(x_i) \cdot p(y_j)}{p(x_i, y_j)}$$

The quantity of information has the following properties:

$$(33) \qquad I(X,Y) \geqslant 0$$

$$(34) \qquad I(X,Y) \equiv I(Y,X)$$

$$(35) \qquad I(X,Y) \leqslant H(Y)$$

$$(36) \qquad I(X,Y) \leqslant H(X)$$

The relation (33) shows that the quantity of information cannot be negative but must be either positive or zero; the last case occuring only with the random variables independent, one of them containing no information about the other. Then, as we already saw, the unconditional entropy is equal to the conditional one.

The relation (34) shows the symmetry in the amount of information. This means that the quantity of information contained in the random variable X about the random variable Y is equal to the quantity of information contained in the random variable Y which is relevant to the random variable X.

The relations (35) and (36) show that the quantity of information can never exceed any of the unconditional entropies. It follows that in case the unconditional entropies of the two random variables are different the maximal amount of information which one of them contains about the other can be at most equal to the smaller of the unconditional entropies. In this case the random variable with greater unconditional entropy contains the complete information regarding the other variable, but the converse is not true. Thus in this case knowing the value which assumes the first random variable gives us complete information about which value assumed the other one and to any value of the first variable thus corresponds just one value of the second one; but knowing the value which assumed the second variable gives us only a partial information on what happened with the first one, since to every value of the second variable there may correspond several values of the first one. For example, if the unconditional entropy of the random variable X exceeds the one of the random variable Y:

(37) $$H(X) > H(Y)$$

then the following hold

(38) $$I(X,Y)\ max = H(Y)$$

(39) $$I(X,Y)\ max < H(X)$$

Of course, if the unconditional entropies are equal:

(40) $$H(X) = H(Y)$$

then the quantity of information contained in one of these variables can, as shown by (33), (35) and (36) assume any value between zero (which is the case when the variables are not even stochasticaly dependent) and the common value of the unconditional entropies (when the variables are connected by a functional one-one relation).

The above discussion shows that the quantity of information contained in one random variable relevant to the other one does represent some measure for stochastic dependence. This measure too, just as the other ones mentioned in the introduction, is symmetrical. As a measure for stochastic dependence, this measure is in many respects less useful then the other ones. First of all, it is not suitable for comparing stochastical dependencies of several pairs of random variables, since its upper bound depends on the number of values the variables assume and on corresponding distributions of probabilities, while the above mentioned measures are normed either to the interval (0,1) or to the interval (-1,1). Nevertheless the quantity of information can be used as a basis for derivation of a more convenient measure for stochastic dependence, which will be normed to the interval (0,1) and which will loose the feature of being symmetrical; thus it will not necessarily give the same value for both directions of stochastic dependence.

## 4. Relative decrease of entropy as a measure for stochastic dependence

Since the quantity of information defined as a difference between the unconditional and the conditional entropies (31) is symmetric (34), we have:

(41) $$H(Y) - H(Y/X) = H(X) - H(X/Y),$$

whence it follows

(42) $$H(X/Y) - H(Y/X) = H(X) - H(Y)$$

As we see the difference between the unconditional entropies is the same as the difference between the conditional ones. The decrease in entropy given by the quantity of information is the same for both variables, and the differences between the entropies remain the same. However, if we devide the left side in (41) by H(Y) and the right side by H(X), there will be no more equality. Assuming again (37):

$$H(X) > H(Y)$$

we obtain the following relation:

$$(43) \qquad \frac{H(Y) - H(Y/X)}{H(Y)} > \frac{H(X) - H(X/Y)}{H(X)} ,$$

or

$$(44) \qquad \frac{I(X,Y)}{H(Y)} > \frac{I(X,Y)}{H(X)} .$$

The left sides in (43) and (44) represent the relative decrease of entropy of the random variable Y, as a consequence of knowing X, and the right sides represent the relative decrease of entropy of X as a consequence of knowing Y. These relations show that the relative decrease of entropy as a measure for indefiniteness, or the relative increase in the quantity of information, are greater at the random variable with smaller unconditional entropy, given the other variable, then vice versa.

From the preceding it follows that the right and the left sides in (43) and (44), giving relative dereases of entropies of the random variables X and Y, can serve as special asymmetrical measures for stochastic dependence of the variables. The left sides in (43) and (44) will be called coefficients for stochastic dependence, and they will be denoted by:

K(Y/X) - coefficient for stochastic dependence of the random variable Y with respect to the random variable X;

K(X/Y) - coefficient for stochastic dependence of the random variable X with respect to the random variable Y.

These coefficients are consequently defined by

(45) $$K(Y/X) = \frac{H(Y) - H(Y/X)}{H(Y)}$$

(46) $$K(X/Y) = \frac{H(X) - H(X/Y)}{H(X)}$$

or

(47) $$K(Y/X) = \frac{I(X,Y)}{H(Y)}$$

(48) $$K(X/Y) = \frac{I(X,Y)}{H(X)}$$

Taking into account (33), (35) and (36) it is clear that these coefficients lie between zero (in case of no stochastic dependence) and one (in case of a one-one functional dependence between the variables):

(49) $$0 \leqslant K(Y/X) \leqslant 1$$

(50) $$0 \leqslant K(X/Y) \leqslant 1$$

Of course in case the unconditional entropies are different, only one of the coefficients can attain the maximum value 1. In this case the variable whose coefficient for stochastic dependence with respect to the other one is one must be functionally dependent in a unique manner from the other variable, with the inverse not being true. This means that to every value of the second random variable there must correspond just one value of the first one, but conversely to some values of the first variable there must correspond several values of the second one. So, in case the unconditional entropy of the random variable $X$ exceeds the one of $Y$, then the maximal value of the coefficients for stochastic dependence will be:

(51) $$K(Y/X)\ \max = 1$$

(52) $$K(X/Y)\ \max = \frac{H(Y)}{H(X)}$$

This means that the random variable $Y$ will be in this case functionally dependent from the random variable $X$, i.e. to every value

X can assume there will correspond just one value Y can assume; while to a given value of a random variable Y there may correspond several values of the random variable X.

We shall now show on a numerical example how the above defined coefficients for stochastic dependence are evaluated and we shall compare the result with the corresponding Pearson's contingency coefficient.

## 5. Evaluating of the coefficient for stochastic dependence

In order to evaluate the coefficient for stochastic dependence as defined by (47) and (48) we have to know the joint probability distribution of the pair X, Y, given by $p(x_i, y_j)$. If the joint probability distribution is not known but instead the conditional probability distributions of one of the variables and the marginal probability distribution of the other are given, then at first the joint probability distribution is to be evaluated, using (11), resp. (12).

We shall assume in our example that the joint probability distribution is given. The numerical values for various joint probabilities are given in the Table 1. Also are evaluated and recorded in the corresponding column, resp. row, in the same table, the marginal probabilities of X, resp. Y, and the last column, resp. the last row, contains the products of marginal probabilities with their binary cologarithms for X and Y resp. In Tables 2 and 3 the conditional probabilities are recorded. Using (17) we get the unconditional entropy of X:

$$H(X) = \sum_{i=1}^{m} p(x_i) \text{ cold } p(x_i) = 1,3538 \text{ bits};$$

(Using (18) we get the unconditional entropy of Y:

$$H(Y) = \sum_{j=1}^{n} p(y_j) \text{ cold } p(y_j) = 3,000 \text{ bits.}$$

To evaluate the information, we use (32):

$$I(X,Y) = \sum_{i=1}^{m} \sum_{j=1}^{n} p(x_i, y_j) \text{ cold } \frac{p(x_i).p(y_j)}{p(x_i, y_j)}$$

$$= 8.0,118 \text{ . cold } \frac{0,476 \text{ . } 0,125}{0,118} + 8 \text{ . } 0,001 \text{ . cold}$$

$$\frac{0,476 \text{ . } 0,125}{0,001} + 48 \text{ . } 0,001 \text{ . cold } \frac{0,008 \text{ . } 0,125}{0,001} =$$

$$= 0,944 - \text{cold } 0,504 - 0,008 \text{ . cold } 0,017 + 0 =$$

$$= 0,9331 - 0,0470 = 0,8861 \text{ bits}$$

Coefficients for stochastic dependence are then:

$$K(X/Y) = \frac{I(X,Y)}{H(X)} = \frac{0,8861}{1,3538} = 0,6545$$

$$K(Y/X) = \frac{I(X,Y)}{H(Y)} = \frac{0,8861}{3,0000} = 0,2954$$

We see that in the present example

$$K(X/Y) > K(Y/X),$$

which means that the stochastic dependence of the random variable X with respect to the random variable Y is greater than that of Y with respect to X. In this example this is evident also from the conditional probability laws tabulated in the Tables 2 and 3. In the Table 2 we see that the conditional probability for X to assume any value out of the sequence $x_i$, given $Y = y_j$, does not depend on $y_j$ and is equal to

$$p(x_i/y_j) \max = 0,944; \quad (j = 1,2,\ldots,8);$$

This shows great stochastic dependence of X with respect to Y.

Tabel 1

$p(x_i, y_j)$; $p(x_i)$; $p(y_j)$; $p(x_i)$ cold $(plx_i)$; $p(y_j)$ cold $p(y_j)$

| X \ Y | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $y_5$ | $y_6$ | $y_7$ | $y_8$ | $p(x_i)$ | $p(x_i)$ cold p $(x_i)$ (bit) |
|---|---|---|---|---|---|---|---|---|---|---|
| $x_1$ | 0,118 | 0,118 | 0,118 | 0,118 | 0,001 | 0,001 | 0,001 | 0,001 | 0,476 | 0,5098 |
| $x_2$ | 0,001 | 0,001 | 0,001 | 0,001 | 0,118 | 0,118 | 0,118 | 0,118 | 0,476 | 0,5098 |
| $x_3$ | 0,001 | 0,001 | 0,001 | 0,001 | 0,001 | 0,001 | 0,001 | 0,001 | 0,008 | 0,0557 |
| $x_4$ | 0,001 | 0,001 | 0,001 | 0,001 | 0,001 | 0,001 | 0,001 | 0,001 | 0,008 | 0,0557 |
| $x_5$ | 0,001 | 0,001 | 0,001 | 0,001 | 0,001 | 0,001 | 0,001 | 0,001 | 0,008 | 0,0557 |
| $x_6$ | 0,001 | 0,001 | 0,001 | 0,001 | 0,001 | 0,001 | 0,001 | 0,001 | 0,008 | 0,0557 |
| $x_7$ | 0,001 | 0,001 | 0,001 | 0,001 | 0,001 | 0,001 | 0,001 | 0,001 | 0,008 | 0,0557 |
| $x_8$ | 0,001 | 0,001 | 0,001 | 0,001 | 0,001 | 0,001 | 0,001 | 0,001 | 0,008 | 0,0557 |
| $p(y_j)$ | 0,125 | 0,125 | 0,125 | 0,125 | 0,125 | 0,125 | 0,125 | 0,125 | 1,000 | 1,3538 |
| $p(y_j)$ cold $p(y_j)$ (bit) | 0,3750 | 0,3750 | 0,3750 | 0,3750 | 0,3750 | 0,3750 | 0,3750 | 0,3750 | 3,000 | |

Tabel 2

$$p(x_i/y_j) = \frac{p(x_i, y_j)}{p(y_j)}$$

| X \ Y | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $y_5$ | $y_6$ | $y_7$ | $y_8$ |
|---|---|---|---|---|---|---|---|---|
| $x_1$ | 0,944 | 0,944 | 0,944 | 0,944 | 0,008 | 0,008 | 0,008 | 0,008 |
| $x_2$ | 0,008 | 0,008 | 0,008 | 0,008 | 0,944 | 0,944 | 0,944 | 0,944 |
| $x_3$ | 0,008 | 0,008 | 0,008 | 0,008 | 0,008 | 0,008 | 0,008 | 0,008 |
| $x_4$ | 0,008 | 0,008 | 0,008 | 0,008 | 0,008 | 0,008 | 0,008 | 0,008 |
| $x_5$ | 0,008 | 0,008 | 0,008 | 0,008 | 0,008 | 0,008 | 0,008 | 0,008 |
| $x_6$ | 0,008 | 0,008 | 0,008 | 0,008 | 0,008 | 0,008 | 0,008 | 0,008 |
| $x_7$ | 0,008 | 0,008 | 0,008 | 0,008 | 0,008 | 0,008 | 0,008 | 0,008 |
| $x_8$ | 0,008 | 0,008 | 0,008 | 0,008 | 0,008 | 0,008 | 0,008 | 0,008 |
| $\sum_{i=1}^{m} p(x_i/y_j)$ | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 |

T a b e l 3

$$p(y_j/x_i) = \frac{p(x_i, y_j)}{p(x_i)}$$

| X \ Y | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $y_5$ | $y_6$ | $y_7$ | $y_8$ | $\sum_{j=1}^{n} p(y_j/x_i)$ |
|---|---|---|---|---|---|---|---|---|---|
| $x_1$ | 0,248 | 0,248 | 0,248 | 0,248 | 0,002 | 0,002 | 0,002 | 0,002 | 1,000 |
| $x_2$ | 0,002 | 0,002 | 0,002 | 0,002 | 0,248 | 0,248 | 0,248 | 0,248 | 1,000 |
| $x_3$ | 0,125 | 0,125 | 0,125 | 0,125 | 0,125 | 0,125 | 0,125 | 0,125 | 1,000 |
| $x_4$ | 0,125 | 0,125 | 0,125 | 0,125 | 0,125 | 0,125 | 0,125 | 0,125 | 1,000 |
| $x_5$ | 0,125 | 0,125 | 0,125 | 0,125 | 0,125 | 0,125 | 0,125 | 0,125 | 1,000 |
| $x_6$ | 0,125 | 0,125 | 0,125 | 0,125 | 0,125 | 0,125 | 0,125 | 0,125 | 1,000 |
| $x_7$ | 0,125 | 0,125 | 0,125 | 0,125 | 0,125 | 0,125 | 0,125 | 0,125 | 1,000 |
| $x_8$ | 0,125 | 0,125 | 0,125 | 0,125 | 0,125 | 0,125 | 0,125 | 0,125 | 1,000 |

In the Table 3 we see that the conditional probabilities for $Y$ to assume any value out of the sequence $y_j$ under the condition $X = x_i$ do not depend on $x_i$ either, and are equal to

$$p(y_j/x_i) \max = 0,248; \quad (i = 1,2);$$

showing a smaller stochastic dependence of $Y$ with respect to $X$ than vice versa. Obviously in a more complex case with the conditional probabilities not being all equal it would be difficult to tell which stochastic dependence should be greater. The above introduced coefficients solve this problem quite generally.

We shall now evaluate the Pearson's contingency coefficient for the example above and we shall compare this symmetric measure for stochastic dependence with the one we just introduced.
The Pearson's contingency coefficient is given by

$$\varphi^2 = \sum_{i=1}^{m} \sum_{j=1}^{n} \frac{p(x_i, y_j)^2}{p(x_i) \cdot p(y_j)} - 1$$

which gives in the example above

$$\varphi^2 = 8 \frac{0,118^2}{0,476 \cdot 0,125} + 8 \frac{0,001^2}{0,476 \cdot 0,125} + 48 \frac{0,001^2}{0,008 \cdot 0,125} -$$

$$-1 = 1,872 + 0,048 - 1 = 0,920$$

As we see, the Pearson's contingency coefficient, in addition to being symmetrical, assumes here a disproportionally great value compared with the conditional probabilities, as recorded in the Tables above.

## 6. Conclusion

In this paper we showed that relative decreases of entropies of two stochastically dependent random variables achieved through the information one of them contains about the other one can serve as a convenient asymmetrical measure for stochastic dependence. These measures we defined as coefficients for stochastic dependence. In the numerical example in the last paragraph we showed how the coefficients are partically evaluated. (In the Appendix there are Tables of binary cologarithms for probabilities ranging from 0,001 to 1, as well as products of probabilities with corresponding cologarithms).

We did not discuss the possibility of application the introduced coefficients for stochastic dependence in various explorations, this being beyond author's intentions in the present paper. Author beleives the practical usefulness of the above coefficients and the possibility for their application should be examined and discovered in practice, in particular regarding their interpretation in different fields. However, we have to indicate some possibilities for the interpretation of the coefficients for stochastical dependence with regard to their asymmetry, this distinguishing them from other measures for dependence presently in use. Thus if we have two processes taking place at the same time and interacting one with the other and if the processes can be described by finite sequences of random variables and their corresponding probability laws, then the coefficients for stochastic dependence may show us which of the processes dominates upon the other, which was not possible with the symmetrical coefficients for dependence. At the same time these coefficients indicate which of the two stochasticaly dependent processes or phenomena gives more complete information, hence being a better representant, about the other.

They also show where there is a greater degree of freedom. When the connestion between two temporally distant processes is discussed, then the coefficients for stochastic dependence give relations between cause and consequence, the dependence of the consequence from its cause; throwing additional light on the interaction between the cause and the consequence. They also may indicate what is the connection from present and future, showing whether some definite present situation more reflects the anticipation of the future or it does less.

Finally, we shall emphasize an important point. We beleive that the use of various symmetric measures for stochastic dependence in various social, organizational, economical and other empirical explorations roots in ideology. Already at their construction the symmetry was apriori considered necessary; this reflecting a wiev according to which the world is in general a world of harmony, of equilibrium and of mutual reciprocity, where no dominance or unequality is accepted. Of cours the measures constructed in this way do not make it possible to reveal any asymmetrical relation, dominance, unequality, difference in the degree of freedom in behaviour etc. The proposed coefficients for stochastic dependence do give a possibility for discovering such asymmetry. The success will of course depend upon the actual interpretation in various investigations. Author shall be very grateful for any information regarding practical applications of above coefficients no matter in which scientific discipline they should be used.