Professor Ion PURCARU, PhD
Professor Gabriela BEGANU, PhD
Professor Felicia ALEXANDRU, PhD
Professor Ion VERBONCU, PhD
The Bucharest Academy of Economic Studies

INTERDEPENDENCE AND DIVERSITY MEASURES FOR A JOINT ECOSYSTEM WITH APPLICATIONS IN SYSTEMS MANAGEMENT

Abstract. *In this paper, using the well known classical concepts of the Shannon entropy and the Simpson diversity index, some interesting considerations on the measuring and on the measures of the interdependence and the diversity for a joint ecosystem are presented.*

Key words: *entropy, diversity, interdependence, ecosystem.*

MSC 2000: 62P10; 62P12; 62P20; 62P25; 91B82; 92D40; 94A17.
JEL Classification: C02, C44, C61, C63

1. INTRODUCTION

Generally, it is well known that, any probabilistic experiment is characterized by a set of possible outcomes (also, called elementary events) and a probability distribution of the possible outcomes. Let us consider a arbitrary finite joint probabilistic experiment (X,Y) denoted by

$$(X,Y)=\begin{pmatrix} (x_1,y_1) & \dots & (x_i y_j) & \dots & (x_m,y_n) \\ \pi_{11} & \dots & \pi_{ij} & \dots & \pi_{mn} \end{pmatrix} \tag{1.1}$$

where, in this case, (x_i,y_j) is a pair of possible outcomes x_i in experiment X and possible outcomes y_j in experiment Y and the numbers

$$\pi_{ij}\geq 0,(i=1,...,m),(j=1,...,n),\sum_{i=1}^{m}\sum_{j=1}^{n}\pi_{ij}=1 \tag{1.2}$$

define *a joint probability distribution* of (X,Y). As a result, the following numbers

$$p_i = \sum_{j=1}^{n} \pi_{ij}, (i = 1,...,m), \quad q_j = \sum_{i=1}^{m} \pi_{ij}, (j = 1,...,n) \tag{1.3}$$

define *the marginal probability distributions* of the joint probability distribution and the numbers

$$p(i/j) = \frac{\pi_{ij}}{q_j}, \ (i = 1,...,m), \quad q(j/i) = \frac{\pi_{ij}}{p_i}, \quad (j = 1,...,n) \tag{1.4}$$

define *the conditional probability distributions* of the joint probability distribution. As a result of the relations (1.3)-(1.4), the number $p_i = P(X = x_i)$ is the occurrence probability of x_i in X, whatever happens in Y, the number $p(i/j) = P(X = x_i / Y = y_i)$ is the occurrence probability of x_i in X, if the outcome y_j occurs in Y and similarly for the numbers $q_j = P(Y = y_j)$ and $q(j/i) = P(Y = y_j / X = x_i)$, $(i = 1,2,...,m)$, $(j = 1,2,...,n)$.
Before a probabilistic experiment to take place, there is an uncertainty about its actual outcomes. This uncertainty essentially depends on the probability distribution of the outcomes. According to Shannon (1948), the amount of uncertainty (information) contained by a probabilistic experiment is measured by a certain positive number called, generally, Shannon entropy.

Definition 1.1. *For a finite probabilistic experiment*

$$X = \begin{pmatrix} x_1 & \cdots & x_i & \cdots & x_m \\ p_1 & \cdots & p_i & \cdots & p_m \end{pmatrix} \tag{1.5}$$

the following number

$$H(X) = H(p) = -\sum_{i=1}^{m} p_i \log p_i \tag{1.6}$$

is called Shannon absolute discrete entropy of X *(or of the probability distribution p) where, in general, log represent the natural logarithm (sometimes denoted by ln and usually by* $\log_2$*).*

Proposition 1.1. *The following properties of the Shannon entropy* $H(X)$ *defined by (1.6) are immediately (see: [3], [4], [8], [14]):*

1) $H(X)$ *is a nonnegative, symmetric and continuous function of the variables* (p_k).

2) $H(p_1, p_2, ..., p_m, 0) = H(p_1, p_2, ..., p_m)$.

3) *if* $p_s = 1$ *and* $p_i = 0$ *for* $i \neq s$ *, then* $H(X) = 0$.

4) $H(p_1,...,p_m) \leq H\left(\frac{1}{m},...,\frac{1}{m}\right) = \log m$.

Definition 1.2. *For a joint probabilistic experiment* (X,Y) *given by (1.1), the number*

$$H(X,Y) = H(\pi) = -\sum_{i=1}^{m}\sum_{j=1}^{n}\pi_{ij}\log\pi_{ij} \tag{1.7}$$

is called Shannon absolute discrete entropy of (X,Y) *(or of the joint distribution* π *).*

Definition 1.3. *For a finite joint probabilistic experiment* (X,Y) *given by (1.1), the number*

$$H(X/y_j) = -\sum_{i=1}^{m} p(x_i/y_j)\log p(x_i/y_j) = -\sum_{i=1}^{m} p(i/j)\log p(i/j) \tag{1.8}$$

is called the relative (conditional) entropy of the experiment X *, if* y_j *occurs in* Y *, the number*

$$H(X/Y) = -\sum_{j=1}^{n} q_j H(x/y_j) = -\sum_{i=1}^{m}\sum_{j=1}^{n}\pi_{ij}\log p(i/j) \tag{1.9}$$

is called the relative entropy of the experiment X *conditioned by the experiment* Y *,the number*

$$H(Y/x_i) = -\sum_{j=1}^{n} q(j/i)\log q(j/i) \tag{1.10}$$

is called the relative (conditional) entropy of the experiment Y *if* x_i *occurs in* X *and the number*

$$H(Y/X) = \sum_{i=1}^{m} p_i H(Y/x_i) = -\sum_{i=1}^{m}\sum_{j=1}^{n}\pi_{ij}\log q(j/i) \tag{1.11}$$

is called the relative entropy of the experiment Y *conditioned by the experiment* X**.**

Proposition 1.2. *The following properties of the entropies of the joint probabilistic experiment* (X,Y) *given by (1.6)-(1.11) are immediately* (see: [3], [4], and [8]):

1) $H(X,Y) = H(Y,X)$

2) $H(X,Y) = H(X) + H(Y/X) = H(Y) + H(X/Y)$

3) $H(X/Y) \leq H(X)$ *and* $H(Y/X) \leq H(Y)$

4) $H(X) - H(X/Y) = H(Y) - H(Y/X)$ (*entropic or informational balance*)

5) $H(X,Y) \le H(X) + H(Y)$

6) if X *and* Y *are independent probabilistic experiments, then we have the relations:*

$$H(X/Y) = H(X) \quad ; \quad H(Y/X) = H(Y) \quad ; \quad H(X,Y) = H(X) + H(Y)$$

Remarks. It is well known that any finite probabilistic experiments X and Y are probabilistic independent if $\pi_{ij} = p_i q_j$ (or, equivalently, if $p(i/j) = p_i$ and $q(j/i) = q_j$), for all pair (i,j).

Definition 1.4. *For any finite probabilistic experiment* X *given by (1.5), the number*

$$C(X) = C(p) = \sum_{i=1}^{m} p_i^2 \tag{1.12}$$

is called the concentration index of Gini (1912), Simpson (1949) and Onicescu (1966) of the probabilistic distribution $p = (p_i)$ *and the number*

$$D(X) = D(p) = 1 - C(p) = 1 - \sum_{i=1}^{m} p_i^2 = \sum_{i=1}^{m} p_i(1-p_i) \tag{1.13}$$

is called the diversity index of Gini, Simpson and Onicescu of the probabilistic distribution p.

Proposition 1.3. *The following properties of Gini-Simpson-Onicescu diversity index (1.13) are immediately (see:[2], [3], [6], [9], [10], [11], [12], [13], [15]):*

$$0 \le D(p) \le 1 - \frac{1}{m} \tag{1.14}$$

with the equalities $D(p) = 0$, *if and only if there is* $p_s = 1$ *and* $p_i = 0$ *for* $i \ne s$, $1 \le i \ne s \le m$ *and* $D(p) = \frac{m-1}{m}$, *if and only if* p *is an uniform probability distribution* ($p_i = \frac{1}{m}$, $1 \le i \le m$).

$$D(p) \le H(p) \tag{1.15}$$

where $H(p)$ *is the Shannon entropy (1.6), with equality if and only if there is* $p_s = 1$ *and* $p_i = 0$ *for* $i \ne s$, $1 \le i \ne s \le m$.

Remarks. The two measures (1.6) and (1.13) are not independent. Indeed, if from the series development of the function $\log x$, namely,

$$\log x = (x-1) - \frac{1}{2}(x-1)^2 + \frac{1}{3}(x-1)^3 - ..., \text{ for } x \in (0,1)$$

we keep only the linear term,

$$\log x \approx x - 1$$

then we get the following very important approximation (see: [3], [9], [10], [13]):

$$H(p) = -\sum_{i=1}^{m} p_i \log p_i \approx \sum_{i=1}^{m} p_i(1-p_i) = 1 - \sum_{i=1}^{m} p_i^2 = D(p)$$

and this equality shows the possibility to use, sometimes, the entropy as a measure of diversity.

2. INTERDEPENDENCE AND DIVERSITY MEASURES

Let us consider an arbitrary larger ecosystem which consists of two sub-ecosystems X and Y, also called joint ecosystem and denoted by (X,Y), which is defined by (1.1).

Definition 2.1. *The positive number (see:[3],[4],[8], [16])*

$$W\left[(X,Y);X,Y\right] = H(X) + H(Y) - H(X,Y) = W(X,Y)$$
$$= \sum_{i=1}^{m}\sum_{j=1}^{n} \pi_{ij} \log \frac{\pi_{ij}}{p_i q_j} \qquad (2.1)$$

is called Watanabe interdependence index of the joint system (or ecosystem) (X,Y).

Proposition 2.1. *From the properties of Shannon entropies of the joint system (X,Y), the following properties of the interdependence (2.1) are immediately (see: [3], [4], [8], [16]):*

1) $W(X,Y) = W(Y,X)$*:*

2) $W(X,Y) \geq 0$*, with equality if and only if the sub-ecosystems* X *and* Y *are independent;*

3) $W(X,Y) = H(X) - H(X/Y) = H(Y) - H(Y/X)$*;*

4) $W(X,Y) \leq \min\{H(X),H(Y)\} \leq \frac{1}{2}[H(X)+H(Y)] \leq \max\{H(X),H(Y)\}$

Remarks. Let us notice that the measure W given by (2.1) may be generalized to any finite number of probability experiments (see: [3], [8], and [16]). For example, if we consider the ecosystem (X,Y,Z), then we have, by similarity to (2.1), the following interdependence measures:

$$W\left[(X,Y,Z);X,Y,Z\right]=H(X)+H(Y)+H(Z)-H(X,Y,Z) \tag{2.2}$$

$$W\left[(X,Y,Z);(X,Y,Z)\right]=H(X,Y)+H(Z)-H(X,Y,Z) \tag{2.3}$$

with a certain interpretation in the analysis of the organization of the systems and we notice that

$$W[(X,Y,Z);(X,Y),Z]\leq W[(X,Y,Z);X,Y,Z] \tag{2.4}$$

with equality if and only if X , Y and Z are independent (see [8]).

Definition 2.2. *For any ecosystem (X,Y) given by (1.1),the number*

$$D(X)=D(p)=1-\sum_{i=1}^{m}p_i^2 \tag{2.5}$$

is called the Simpson index of diversity of the ecosystem X, the number

$$D(X/y_j)=1-\sum_{i=1}^{m}p^2(i/j) \tag{2.6}$$

is called the Simpson index of diversity of the ecosystem X conditioned by the specie y_j of the ecosystem Y, the number

$$D(X/Y)=\sum_{j=1}^{n}q_jD(X/y_j)=1-\sum_{i=1}^{m}\sum_{j=1}^{n}\pi_{ij}p(i/j) \tag{2.7}$$

is called the Simpson index of diversity of the ecosystem X conditioned by the ecosystem Y and

$$D(X,Y)=1-\sum_{i=1}^{m}\sum_{j=1}^{n}\pi_{ij}^2 \tag{2.8}$$

is called the Simpson index of diversity of joint ecosystem (X,Y) and, similarly, we can introduce the corresponding numbers $D(Y/x_i)$ and $D(Y/X)$.

Proposition 2.2. *By direct calculus, the following properties of the Simpson indices given by the relations (2.5)-(2.8) are immediately:*

1) $D(X/Y)\leq D(X)\leq D(X,Y)\leq D(X)+D(Y)$

2) $D(X,Y)\geq \max\left\{D(X),D(Y)\right\}\geq \frac{1}{2}(D(X)+D(Y))\geq \min\left\{D(X),D(Y)\right\}$

3) *if the ecosystems X and Y are independent, then we have:*

$$D(X/Y)=D(X);\;\; D(Y/X)=D(Y);\;\; D(X,Y)=D(X)+D(Y)-D(X)\times D(Y)$$

4) *if the joint probabilistic distribution of the ecosystem (X,Y) is a diagonal or a pseudo-diagonal probabilistic distribution, then we have the equalities:*

$$D(X/Y) = D(Y/X) = 0 \quad ; \quad D(X) = D(Y) = D(X,Y)$$

Remarks. Many very interesting interpretations of the presented relations are possible. Let us consider the following number (see [3]):

$$V(X,Y) = \sum_{i=1}^{m}\sum_{j=1}^{n}\frac{\pi_{ij}^2}{p_i q_j} - 1 = \sum_{i=1}^{m}\sum_{j=1}^{n} p(i/j)q(j/i) - 1 \tag{2.9}$$

which is defined for any finite ecosystem (X,Y) given by (1.1) and we remark the followings:

1) $V(X,Y) = V(Y,X)$; **2)** $V(X,Y) \geq 0$; **3)** if X and Y are independent, then $V(X,Y) = 0$.

Let us notice that, as a result of the relation $\log x \leq x - 1$, we can write the followings:

$$0 \leq W(X,Y) \leq V(X,Y) \tag{2.10}$$

and, if the considered ecosystems X and Y are independent, i.e., $\pi_{ij} = p_i q_j$, for all pair (i,j), $1 \leq i \leq m$ and $1 \leq j \leq n$, then we have the interesting equality $W(X,Y) = V(X,Y) = 0$.

Let us consider too the following approximation $\log x \approx x - 1$, for any $x \in (0,1)$. Then we obtain the approximation $W(X,Y) \approx V(X,Y)$ and, as a result, in this situation, we can consider the number $V(X,Y)$ given by (2.9) as a measure of interdependence between the ecosystem X and Y, corresponding to the entropic measure of interdependence $W(X,Y)$ defined by (2.1).

Example. If the ecosystem X has three elements (x_1, x_2, x_3) and the ecosystem Y has four elements (y_1, y_2, y_3, y_4), we assume that, taken together, they form the joint ecosystem, given by the following table:

$X_{\downarrow}$; $Y_{\rightarrow}$	y_1	y_2	y_3	y_4	p_i
x_1	0.15625	0.0625	0.015625	0.015625	0,25
x_2	0.0625	0.3125	0.046875	0.078125	0.50
x_3	0.03125	0.125	0.0625	0.03125	0.25
q_j	0.25	0.50	0,125	0.125	1.00

where, for example, we can read the followings: $\pi_{11} = P(X = x_1, Y = y_1) = 0.15625$; $p_1 = \pi_{11} + \pi_{12} + \pi_{13} + \pi_{14} = 0{,}25$; $q_1 = \pi_{11} + \pi_{21} + \pi_{31} = 0.25$ and also for the other numbers. As a result, we obtain:

$$C(X) = 0.375;\ C(Y) = 0.34375\ ;\ C(X,Y) = 0.16015625$$

$$D(X) = 0.625;\ D(Y) = 0.65625;\ D(X,Y) = 0.83984375$$

We also get the following conditional distributions:

1) for the conditional distributions $p(x_i / y_j) = p(i/j) = \pi_{ij} \div q_j$, we have:

p(1/1)=0.625; p(2/1)=0.250; p(3/1)=0.125
p(1/2)=0.125; p(2/2)=0.625; p(3/2)=0.250
p(1/3)=0.125; p(2/3)=0.375; p(3/3)=0.500
p(1/4)=0.125; p(2/4)=0.625; p(3/4)=0.250

and as a result of (2.6) and (2.7), we obtain:

$$C(X/y_1) = 0.46875 = C(X/y_2) = C(X/y_4) \text{ and } C(X/y_3) = 0.40625$$

$$D(X/y_1) = 0.53125 = D(X/y_2) = D(X/y_4) \text{ and } D(X/y_3) = 0.59375$$

$$C(X/Y) = 0.4609375 \text{ and } D(X/Y) = 0.5390625$$

2) for the conditional distributions $q(y_j / x_i) = q(j/i) = \pi_{ij} \div p_i$, we have:

q(1/1)=0.6250; q(2/1)=0.2500; q(3/1)=0.06250; q(4/1)=0.06250
q(1/2)=0.1250; q(2/2)=0.6250; q(3/2)=0.09375; q(4/2)=0.15625
q(1/3)=0.1250; q(2/3)=0.5000; q(3/3)=0.25000; q(4/3)=0.12500

and as a result of (2.6) and (2.7), we obtain:

$$C(Y/x_1) = 0.4609375;\ C(Y/x_2) = 0.439453125 \text{ and } C(Y/x_3) = 0.34375$$

$$D(Y/x_1) = 0.5390625;\ D(Y/x_2) = 0.560546875 \text{ and } D(Y/x_3) = 0.65625$$

$$C(Y/X) = 0.4208984375 \text{ and } D(Y/X) = 0.5791015625$$

Finally, the amount of interdependence between the ecosystems X and Y measured by (2.9) is

$$V(X,Y) = \frac{37}{128} = 0.2890625$$

Similarly, from (1.6), (1.7) and (2.1), we get the entropies:

$$H(X) = \frac{3\log 2}{2};\ H(Y) = \frac{7\log 2}{4};\ H(X,Y) = \frac{317\log 2 - 3\log 3 - 35}{64}$$

$$H(X/Y) = \frac{205\log 2 - 3\log 3 - 35}{64};\ H(Y/X) = \frac{221\log 2 - 3\log 3 - 35}{64}$$

and as a result, the entropic amount of interdependence between the ecosystems X and Y (or the amount of organization of the joint ecosystem (X,Y)) measured by the expression (2.1) is

$$W(X,Y) = \frac{35 + 3\log 3 - 109\log 2}{64} \cong 0.05655 < V(X,Y)$$

and some interesting practical interpretations are possible in the ecosystems study.

3. SOME CONCLUSIONS

Diversity is an essential characteristic of the living world (Guiaşu and Guiaşu, 2003). An understanding of diversity is required for the study of ecosystems and the analysis of the relationships among the species found within these ecosystems. The topic of diversity is of special interst to ecologists, conservation biologists and biogeographers. But the living world is a complexe systeme with a certain degree of interdependence betweens the components of system and a certain organization of system. There are several definitions for systems. In spite of differences in opinion, all experts agree that a system represent something more than the mere sum oh its parts or components. This "something more" refers to the interdependence and interaction among the components of the respective system. As a result, the interdependence is a very important concept and the measuring of interdependence and of organization of system is a very interesting problem. Interdependence can now be measured as a result of a combination of Shannon entropies (see (2.1)-(2.4)) or by an approximation of a certain diversity (see (2.9)-(2.10). The measure of interdependence $V(X,Y)$ given by (2.9) has the advantage of being easier to be calculated than $W(X,Y)$ given by (2.1) because it does not require to use the logarithm function. The generalizations of these measures for a finite system $X = (X_1, X_2, ..., X_k)$, $k \geq 3$, with a certain structure is a very interesting problem with many interpretation possibilities in ecology, in economy, in financial operations, in social sciences or in the study of the systems management.

REFERENCES

[1] **Baumgärtner, S. (2006):"** ***Why the Measurement of Species Diversity Requires Prior Value Judgment?*** ". Working paper, University of Heidelberg, Germany;
[2] **Gini, C. (1912) ,"*Variabilità e mutabilità*".** Studi Economico-Giuridici dela Facoltà di Giurisprundenza dell' Universita di Cagliari, Italy;
[3] **Guiasu, R. C. and Guiasu, S. (2003) ," *Entropy in Ecology and Ethology*"**. Nova Science Publishers, New York;
[4] **Guiasu, S. (2001) , *"Quantum Mechanics".*** Nova Sciences Publishers, New York;
[5] **Magurran, A.E. (1988), "*Ecological Diversity and Its Measurement*"**. Princeton University Press, Princeton, USA;

[6] **Onicescu, O. (1966) ,"** ***Energia informațională*** ". Mathematical Reports 18, 1419-1421;
[7] **Pielou, E.C. (1977) ,"*Mathematical Ecology***". John Willey & Sons, New York;
[8] **Purcaru, I. (1988),"*Informație şi corelație***". Ştiințifică şi Enciclopedică Publishing House, Bucharest;
[9] **Purcaru, I. (2005), "*Asupra unor măsuri informaționale ale studiului şi analizei ecosistemelor".*** Journal SCCECE, ASE Publishing House, no. 4, pp. 59-70, Bucharest;
[10] **Purcaru, I. (2006) , "*Considérations sur la mesure de la diversité de Guiaşu*"** Journal ECECSR, no. 3-4, pp. 67-72, ASE Publishing House, Bucharest ;
[11] **Purcaru, I. and Beganu, G. (2007) ,"** ***Some Considerations on a Generalized Measure of Diversity and the Maximum Diversity Principle of Guiasu***". Journal of Economic Computation and Economic Cybernetics Studies and Research no. 1-2, pp.47-59, Bucharest;
[12] **Purcaru, I. and Beganu, G. (2007), *„Diversitate de ordin α şi tip β a unui ecosistem***. Journal SCCECE , ASE Publishing House, no. 2 (41), pp. 27-38, Bucharest;
[13] **Purcaru, I., Beganu, G. and Verboncu, I. (2007), *"Some Considerations on the Information and Diversity Measuring".*** Journal Economic Computation and Economic Cybernetics Studies and Research, ASE Publishing House , no. 3-4, pp. 73-86, Bucharest;
[14] **Shannon, C.E. (1948),"** ***A Mathematical Theory of Communications***". Bell System Technical Journal, 27, 379-423, 623-656;
[15] **Simpson, E.H. (1949) ,"** ***Measurement of Diversity***". Nature, no. 163, 688.
[16] **Watanabe, S. (1969), *"Knowing and Guessing".*** Wiley, New York.