KYBERNETIKA — VOLUME 21 (1985), NUMBER 2

ON MEASURES OF RELATIVE 'USEFUL' INFORMATION

H. C. TANEJA

Taneja and Tuteja [7] have introduced the measure of relative 'useful' information which satisfies the property of additivity. A characterization of this measure depending upon the additivity postulate and the mean value property has been provided. Then by choosing a particular type of non-additive law in place of additivity, a measure of non-additive relative 'useful' information has been characterized which may be considered as a quantitative-qualitative measure corresponding to the directed-divergence function of type β studied by Rathie and Kannappan [5].

1. INTRODUCTION

Let $P = (p_1, p_2, \ldots, p_n)$, $0 < p_i \leqq 1$, $\sum_{i=1}^{n} p_i = 1$ be a finite discrete probability distribution of a set of n events $E = (E_1, E_2, \ldots, E_n)$ on the basis of an experiment whose predicted probability distribution is $P^0 = (p_1^0, p_2^0, \ldots, p_n^0)$, $0 < p_i^0 \leqq 1$, $\sum_{i=1}^{n} p_i^0 = 1$. Then the Kullback's measure of directed-divergence [3], or information gain [6], is defined as

$$I(P \mid P^0) = \sum_{i=1}^{n} p_i \log (p_i/p_i^0)\,. \tag{1.1}$$

The quantity (1.1) measures the average 'information gain' in predicting the experiment $E = (E_1, E_2, \ldots, E_n)$. However, this measure does not take into account the effectiveness or, importance of the events. This is because (1.1) depends only on the probabilities associated with the events $E_1, E_2, \ldots, E_n$. In a practical situation of probabilistic nature, quite often subjective consideration get involved with the study. These considerations take into account the effectiveness of the outcomes also. Motivated by this idea, Belis and Guiasu [2] introduced a 'utility distribution' $U = (u_1, u_2, \ldots, u_n)$ where each u_i is a non-negative real number accounting for the utility of the ith event E_i. Thus further study of relative information measures may be based on the utility information schemes given by

(1.2) $$S_n = [E, P, U, n], \quad u_i > 0, \quad 0 < p_i \leqq 1, \quad \sum_{i=1}^{n} p_i = 1,$$

of a set of n events after an experiment, and

(1.3) $$S_n^0 = [E, P^0, U, n], \quad u_i > 0, \quad 0 < p_i^0 \leqq 1, \quad \sum_{i=1}^{n} p_i^0 = 1,$$

of the same set of n events before the experiment.

In both the schemes (1.2) and (1.3) the utility distribution is the same because it is assumed that the utility u_i of an outcome E_i is independent of its probability of occurrence p_i, or predicted probability p_i^0 (refer Longo [4]).

A measure of relative 'useful' information that the scheme (1.2) provides about the scheme (1.3) has been suggested and characterized by Taneja and Tuteja [7]. It is given by

(1.4) $$I(S_n \mid S_n^0) = \sum_{i=1}^{n} u_i p_i \log (p_i/p_i^0), \quad u_i > 0, \quad 0 < p_i, \quad p_i^0 \leqq 1.$$

An interpretation for (1.4) may be given as follows:

The quantity $-u_i \log p_i$ in literature is usually referred to as the 'useful' self-information associated with the event E_i whose probability of occurrence is p_i and utility is u_i. Thus $u_i \log (p_i/p_i^0) = (-u_i \log p_i^0) - (-u_i \log p_i)$ may be taken as the 'useful' information gain in predicting the experiment $E = (E_1, E_2, \ldots, E_n)$.

Now if $T_m = [F, Q, V, m]$ is another utility information scheme, independent of S_n, then (1.4) satisfies the following property, called the weighted additivity:

(1.5) $$I(S_n * T_m \mid S_n^0 * T_m^0) = \bar{V} I(S_n \mid S_n^0) + \bar{U} I(T_m \mid T_m^0),$$

where $S_n * T_m = [E * F, P * Q, U * V, nm]$ etc., and

$$\bar{U} = \sum_{i=1}^{n} u_i p_i, \quad \bar{V} = \sum_{j=1}^{m} v_j q_j.$$

In [7], Taneja and Tuteja have characterized the measure (1.4) for finite discrete complete probability schemes under a set of axioms mainly consisting of additivity and monotonicity law expressed by the utilities. In practical situation, not all the outcomes of an experiment or market situation are needed to be relevant, significant or observable. Under such circumstances the physical phenomenon can be described statistically by a finite discrete incomplete probability scheme with $\sum_{i=1}^{n} p_i < 1$ rather than the ordinary complete one with $\sum_{i=1}^{n} p_i = 1$. In Section 2, we characterize the measure of relative 'useful' information for incomplete probability schemes by using weighted additivity and the mean value property.

Secondly, with ever increasing applications of the informational approach, sub-additivity rather than additivity is becoming an acceptable basis. In social and physical systems additivity does not quite prevail. For instance, in biological systems the interaction between various drugs call for non-additivity of the individual effects. In Section 3, we consider a simple model based on a suitable modification of the additivity property (1.5) and thus formulate a generalized measure of relative 'useful' information satisfying that property.

2. MEASURE OF RELATIVE 'USEFUL' INFORMATION FOR INCOMPLETE PROBABILITY SCHEMES

Let p, p^0 and u be the probability of occurrence, predicted probability and the utility of a random event E. Then the relative 'useful' information conveyed by the occurrence of the E must be a function of p, p^0 and u and let it be denoted by $I(p, p^0, u)$. We now determine the form of $I(p, p^0, u)$ under the following set of postulates:

Postulate 1 (Additivity). Given two independent events E_1, E_2 with probabilities of occurrence p_1, p_2, predicted probabilities p_1^0, p_2^0 and utilities u_1, u_2, then the relative 'useful' information provided by their joint occurrence is given by

$$I(p_1 p_2, p_1^0 p_2^0, u_1 u_2) = u_2\, I(p_1, p_1^0, u_1) + u_1\, I(p_2, p_2^0, u_2)\,, \tag{2.1}$$

where $p_1 p_2$, $p_1^0 p_2^0$ and $u_1 u_2$ are respectively the probability of occurrence, the predicted probability and the utility of their joint occurrence.

Postulate 2 (Unity). An event with unit probability of occurrence, unit utility and having $\frac{1}{2}$ as predicted probability conveys unit amount of relative 'useful' information, that is,

$$I(1, \tfrac{1}{2}, 1) = 1\,. \tag{2.2}$$

Postulate 3 (Zero). An event with p as the probability of occurrence, p^0 as the predicted probability conveys no relative 'useful' information if $p = p^0$, what-so-ever the utility u may be, that is,

$$I(p, p^0, u) = 0\,, \tag{2.3}$$

for all $p = p^0$ and $u > 0$. In particular, we have

$$I(1, 1, u) = 0\,, \tag{2.4}$$

for all $u > 0$.

Theorem 1. The relative 'useful' information $I(p, p^0, u)$ conveyed by the occurrence of a single event with probability p, predicted probability p^0 and utility u ($u > 0$, $0 < p, p^0 \leqq 1$) satisfying Postulates 1 to 3, can be only of the form

$$I(p, p^0, u) = u \log (p/p^0)\,, \quad u > 0\,, \quad 0 < p\,,\ p^0 \leqq 1\,. \tag{2.5}$$

Proof. Since $u_1 u_2 \neq 0$, from (2.1) we get

$$\frac{I(p_1 p_2, p_1^0 p_2^0, u_1 u_2)}{u_1 u_2} = \frac{I(p_1, p_1^0, u_1)}{u_1} + \frac{I(p_2, p_2^0, u_2)}{u_2}\,,$$

or

(2.6)
$$F(p_1 p_2, p_1^0 p_2^0, u_1 u_2) = F(p_1, p_1^0, u_1) + F(p_2, p_2^0, u_2),$$

where

$$F(p_i, p_i^0, u_i) = \frac{I(p_i, p_i^0, u_i)}{u_i} \quad \text{for} \quad i = 1, 2.$$

Substituting $F(p_i, p_i^0, u_i) = G(\log p_i, \log p_i^0, \log u_i)$, for $i = 1, 2$, we can rewrite (2.6) as

(2.7)
$$G(\log p_1 + \log p_2, \log p_1^0 + \log p_2^0, \log u_1 + \log u_2) = \\ = G(\log p_1, \log p_1^0, \log u_1) + G(\log p_2, \log p_2^0, \log u_2).$$

The most general continuous solution of (2.7), refer [1, p. 215], is given by

$$G(\log p, \log p^0, \log u) = c_1 \log p + c_2 \log p^0 + c_3 \log u,$$

and hence

(2.8)
$$I(p, p^0, u) = c_1 u \log p + c_2 u \log p^0 + c_3 u \log u,$$

where c_1, c_2 and c_3 are arbitrary constants.

Using Postulate 2 and Postulate 3 in (2.8), we get (2.5). □

Remark. In deriving the above result, we have taken $u, p, p^0 > 0$. This is to avoid the obvious mathematical complexities. However, if we assume that whenever $p^0 = 0$, p is also zero and further $0 \log 0 = 0 \log (0/0) = 0$, then (2.5) also holds for $u = 0$ and $p = p^0 = 0$.

Next, we derive an expression for the relative 'useful' information measure $I(S_n \mid S_n^0)$. We derive this expression under the following set of postulates for the generalized utility information schemes, viz.

(2.9)
$$S_n = [E, P, U, n]; \quad u_i > 0, \quad 0 < p_i \leqq 1, \quad \sum_{i=1}^{n} p_i \leqq 1,$$

(2.10)
$$T_m = [F, Q, V, m]; \quad v_j > 0, \quad 0 < q_j \leqq 1, \quad \sum_{j=1}^{m} q_j \leqq 1,$$

(2.11)
$$S_n^0 = [E, P^0, U, n]; \quad u_i > 0, \quad 0 < p_i^0 \leqq 1, \quad \sum_{i=1}^{n} p_i^0 \leqq 1,$$

(2.12)
$$T_m^0 = [F, Q^0, V, m]; \quad v_j > 0, \quad 0 < q_j^0 \leqq 1, \quad \sum_{j=1}^{m} q_j^0 \leqq 1,$$

where $\sum_{i=1}^{n} p_i + \sum_{j=1}^{m} q_j \leqq 1$ and $\sum_{i=1}^{n} p_i^0 + \sum_{j=1}^{m} q_j^0 \leqq 1$.

Postulate 4 (Additivity). Given the utility information schemes S_n and T_m and having the predicted utility information schemes S_n^0 and T_m^0 respectively, the following holds

(2.13)
$$I(S_n * T_m \mid S_n^0 * T_m^0) = \bar{V} I(S_n \mid S_n^0) + \bar{U} I(T_m \mid T_m^0),$$

where

$$\bar{U} = (\sum_{i=1}^{n} u_i p_i)/(\sum_{i=1}^{n} p_i) \quad \text{and} \quad \bar{V} = (\sum_{j=1}^{m} v_j q_j)/(\sum_{j=1}^{m} q_j).$$

Postulate 5 (Mean-Value). The relative 'useful' information conveyed by the utility information scheme $S_n U T_m$ on the basis of an experiment and having predicted utility information scheme $S_n^0 U T_m^0$, is given by

$$(2.14) \qquad I(S_n U T_m \mid S_n^0 U T_m^0) = \frac{(\sum_{i=1}^{n} p_i) I(S_n \mid S_n^0) + (\sum_{j=1}^{m} q_j) I(T_m \mid T_m^0)}{\sum_{i=1}^{n} p_i + \sum_{j=1}^{m} q_j},$$

where

$S_n U T_m = [EUF, PUQ, UUV, n + m]$; $u_i, v_j > 0$, $\sum_{i=1}^{n} p_i + \sum_{j=1}^{m} q_j \leqq 1$, etc.

We prove the following theorem:

Theorem 2. The relative 'useful' information provided by the utility information scheme S_n about the utility information scheme S_n^0 and satisfying the Postulates 2, 3, 4 and 5, is given by

$$(2.15) \qquad I(S_n \mid S_n^0) = \frac{\sum_{i=1}^{n} u_i p_i \log (p_i/p_i^0)}{\sum_{j=1}^{n} p_i},$$

where $u_i > 0$, $0 < p_i, p_i^0 \leqq 1$ and $\sum_{i=1}^{n} p_i \leqq 1$.

Proof. We may consider

$$S_n \mid S_n^0 = s_1 U s_2 U \ldots U s_n \mid s_1^0 U s_2^0 U \ldots U s_n^0,$$

where

$$s_i = [E_i, p_i, u_i, 1] \quad \text{and} \quad s_i^0 = [E_i, p_i^0, u_i, 1],$$

for all $i = 1, 2, \ldots, n$.

Using Postulate 5, we have

$$I(S_n \mid S_n^0) = \frac{\sum_{i=1}^{n} p_i I(s_i \mid s_i^0)}{\sum_{i=1}^{n} p_i},$$

where $I(s_i \mid s_i^0)$ is the relative 'useful' information conveyed by the occurrence of a single event E_i with probability of occurrence p_i, predicted probability p_i^0 and utility u_i.

Postulate 4 reduces to Postulate 1 for s_i and s_i^0 and, therefore, from Theorem 1 we have

$$I(s_i \mid s_i^0) = u_i \log (p_i/p_i^0),$$

and hence

$$I(S_n \mid S_n^0) = \frac{\sum_{i=1}^{n} u_i p_i \log (p_i/p_i^0)}{\sum_{i=1}^{n} p_i},$$

which is (2.15). □

In case that the distribution is complete, that is $\sum_{i=1}^{n} p_i = 1$, (2.15) reduces to (1.4), the relative 'useful' information for complete probability distribution.

3. GENERALIZED RELATIVE 'USEFUL' INFORMATION FOR INCOMPLETE PROBABILITY SCHEMES

In this section we consider a non-additive generalization of the additivity property (2.13). There can be various generalization of (2.13), we consider the one given by

(3.1) $$I(S_n * T_m \mid S_n^0 * T_m^0) = \bar{V} I(S_n \mid S_n^0) + \bar{U}\, I(T_m \mid T_m^0) + K\, I(S_n \mid S_n^0)\, I(T_m \mid T_m^0),$$

where $K \neq 0$ is an arbitrary constant.

Obviously when $K = 0$, (3.1) reduces to (2.13). Now we derive a generalized measure of relative 'useful' information satisfying this non-additivity postulate.

In order to determine the form of generalized relative 'useful' information measure based upon (3.1), the additivity Postulate 1 is replaced by the following non-additivity postulate.

Postulate 6 (Non-additivity). Given two independent events E_1, E_2 with probabilities of occurrence p_1, p_2, predicted probabilities p_1^0, p_2^0 and utilities u_1, u_2 then the relative 'useful' information provided by their joint occurrence is given by

(3.2) $$I(p_1 p_2, p_1^0 p_2^0, u_1 u_2) = u_2\, I(p_1, p_1^0, u_1) + u_1\, I(p_2, p_2^0, u_2) + K\, I(p_1, p_1^0, u_1)\, I(p_2, p_2^0, u_2),$$

where $u_1, u_2 > 0$, $0 < p_1, p_2, p_1^0, p_2^0 \leqq 1$ and $K \neq 0$ is an arbitrary constant.

Theorem 3. The relative 'useful' information $I(p, p^0, u)$ conveyed by the occurrence of a single event with probability of occurrence p, predicted probability p^0 and utility u, satisfying the non-additivity Postulate 6 and Postulates 2, 3, can be only of the form

(3.3) $$I(p, p^0, u) = u \left[\frac{(p/p^0)^c - 1}{2^c - 1} \right],$$

where $c \neq 0$ is an arbitrary constant.

Proof. Since $u_1 u_2 \neq 0$, from (3.2) we have

$$F(p_1 p_2, p_1^0 p_2^0, u_1 u_2) = F(p_1, p_1^0, u_1) + F(p_2, p_2^0, u_2) + K\, F(p_1, p_1^0, u_1)\, F(p_2, p_2^0, u_2)\,, \quad K \neq 0\,, \tag{3.4}$$

where

$$F(p_i, p_i^0, u_i) = \frac{I(p_i, p_i^0, u_i)}{u_i} \quad \text{for} \quad i = 1, 2\,.$$

Setting $1 + K\, F(p_i, p_i^0, u_i) = G(p_i, p_i^0, u_i)$ in (3.4), we get

$$G(p_1 p_2, p_1^0 p_2^0, u_1 u_2) = G(p_1, p_1^0, u_1)\, G(p_2, p_2^0, u_2)\,. \tag{3.5}$$

The most general continuous solution of (3.5) is given by

$$G(p, p^0, u) = p^{c_1} p^{0c_2} u^{c_3}\,,$$

and hence

$$I(p, p^0, u) = u\left[\frac{p^{c_1} p^{0c_2} u^{c_3} - 1}{K}\right], \quad K \neq 0\,, \tag{3.6}$$

where $u > 0$, $0 < p$, $p^0 \leqq 1$ and, c_1, c_2 and c_3 are arbitrary constants.

Using Postulates 2 and 3 in (3.6), we get (3.3). □

Measure (3.3) is expected to give the additive 'useful' information gain when the constant c tends to zero. To maintain the convention in the existing literature, it is desirable to take $c = \beta - 1$, so that additivity can follow when $\beta \to 1$. Thus, we have

$$I(p, p^0, u) = u\left[\frac{(p/p^0)^{\beta-1} - 1}{2^{\beta-1} - 1}\right], \quad u > 0,\ \ 0 < p,\ \ p^0 \leqq 1\,, \quad \beta \neq 1\,, \tag{3.7}$$

and the non-additivity postulate (3.2) then takes the form

$$I(p_1 p_2, p_1^0 p_2^0, u_1 u_2) = u_2 I(p_1^0, p_1, u_1) + u_1\, I(p_2, p_2^0, u_2) + (2^{\beta-1} - 1)\, I(p_1, p_1^0, u_1)\, I(p_2, p_2^0, u_2)\,. \tag{3.8}$$

We refer (3.7) as the generalized relative 'useful' self-information of degree β, conveyed by an event with probability of occurrence p, predicted probability p^0 and utility u and we denote this expression by $I^\beta(p, p^0, u)$.

Next, we determine the non-additive form of the relative 'useful' information which the generalized utility information scheme (2.9) provides about the generalized utility information scheme (2.11). We give the following postulates for the generalized utility information schemes:

Postulate 7 (Non-additivity). For the generalized utility information schemes S_n and T_m with the predicted utility information schemes S_n^0 and T_m^0 respectively, the following holds

$$I(S_n * T_m \mid S_n^0 * T_m^0) = \bar{V} I(S_n \mid S_n^0) + \bar{U}\, I(T_m \mid T_m^0) + (2^{\beta-1} - 1)\, I(S_n \mid S_n^0)\, I(T_m \mid T_m^0)\,, \quad \beta \neq 1\,. \tag{3.9}$$

Theorem 4. The non-additive relative 'useful' information conveyed by the utility information scheme S_n whose predicted utility information scheme is S_n^0 and satisfying the Postulates 2, 3, 5 and 7, can be only of the form

$$I^\beta(S_n \mid S_n^0) = \frac{\sum_{i=1}^n u_i p_i[(p_i/p_i^0)^{\beta-1} - 1]}{(2^{\beta-1} - 1)\sum_{i=1}^n p_i}, \quad \beta \neq 1, \tag{3.10}$$

where $u_i > 0$, $0 < p_i, p_i^0 \leqq 1$ and $\sum_{i=1}^n p_i \leqq 1$.

Proof. Considering as earlier

$$S_n \mid S_n^0 = s_1 U s_2 U \dots U s_n \mid s_1^0 U s_2^0 U \dots U s_n^0,$$

where

$$s_i = [E_i, p_i, u_i, 1] \quad \text{and} \quad s_i^0 = [E_i, p_i^0, u_i, 1].$$

for $i = 1, 2, \dots, n$.

Using Postulate 5, we have

$$I^\beta(S_n \mid S_n^0) = \frac{\sum_{i=1}^n p_i I^\beta(s_i \mid s_i^0)}{\sum_{i=1}^n p_i}, \quad \beta \neq 1, \quad \sum_{i=1}^n p_i \leqq 1,$$

where $I^\beta(s_i \mid s_i^0)$ is the generalized relative 'useful' information conveyed by a single event E_i with probability of occurrence p_i, predicted probability p_i^0 and utility u_i.

Using Theorem 3, we have

$$I^\beta(s_i \mid s_i^0) = u_i \left[\frac{(p_i/p_i^0)^{\beta-1} - 1}{2^{\beta-1} - 1}\right], \quad \beta \neq 1,$$

and hence

$$I^\beta(S_n \mid S_n^0) = \frac{\sum_{i=1}^n u_i p_i[(p_i/p_i^0)^{\beta-1} - 1]}{(2^{\beta-1} - 1)\sum_{i=1}^n p_i}, \quad \beta \neq 1, \quad \sum_{i=1}^n p_i \leqq 1,$$

which is (3.10). □

In case when $\sum_{i=1}^n p_i = 1$, then (3.10) reduces to

$$I^\beta(S_n \mid S_n^0) = \frac{\sum_{i=1}^n u_i p_i[(p_i/p_i^0)^{\beta-1} - 1]}{(2^{\beta-1} - 1)}, \quad \beta \neq 1, \tag{3.11}$$

which may be considered as a quantitative-qualitative measure corresponding to the directed-divergence function of type β studied by Rathie and Kannappan [5].

We call (3.11) the generalized relative 'useful' information measure of degree β. It is noted that when $\beta \to 1$, (3.10) and (3.11) reduce to (2.15) and (1.4) respectively.

ACKNOWLEDGEMENTS

The author is thankful to Dr. R. K. Tuteja, Reader in Mathematics, M. D. University, Rohtak for his guidance in the preparation of this paper. Thanks are also due to the referee for suggesting various useful improvements over the earlier version of this paper.

(Received February 8, 1983.)

REFERENCES

[1] J. Aczel: Lectures on Functional Equations and Their Applications. Academic Press, New York 1966.

[2] M. Belis and S. Guiasu: A quantitative-qualitative measure of information in cybernetic systems. IEEE Trans. Inform. Theory *14* (1968), 593—594.

[3] S. Kullback: Information Theory and Statistics. Wiley, New York 1959.

[4] G. Longo: Quantitative-Qualitative Measures of Information. Springer-Verlag, New York 1972.

[5] P. N. Rathie and Pl. Kannappan: A directed-divergence function of type β. Inform. and Control *20* (1972), 38—45.

[6] A Renyi: On measures of entropy and information. In: Proc. Fourth Berkely Symposium on Mathem. Statistics and Probability, University of California Press, 1961, pp. 547—561.

[7] H. C. Taneja and R. K. Tuteja: Characterization of a quantitative-qualitative measure of relative information. Inform. Sciences *33* (1984).

H. C. Taneja, Department of Mathematics, Maharshi Dayanand University, Rohtak - 124001, India. Permanent address: *Department of Mathematics, Guru Nanak Dev University, Amritsar - 143005, India.*