ON GENERALIZED INFORMATION IMPROVEMENT

by RAM AUTAR,* *Faculty of Mathematics, University of Delhi, Delhi*

(Communicated by F. C. Auluck, F.N.A.)

(*Received 10 February 1975*)

Theil (1967) has introduced an information theoretic measure associated with three probability distributions known as 'Information Improvement' which has been of proved interest in applications. The object of this note is to study some new additive generalizations of this measure along with their important properties like monotonicity, non-negativity, etc.

1. INTRODUCTION

Let there be n allocations having the prediction probability distribution $Q = (q_1, \ldots, q_n)$ which is revised as $R = (r_1, \ldots, r_n)$ on the basis of a distribution $P = (p_1, \ldots, p_n)$ realized after some experiment E. Then Theil's (1967) measure of Information Improvement (I.I.) is

$$I(P; Q; R) = \sum_{i=1}^{n} p_i \log (r_i/q_i) \qquad \ldots(1.1)$$

where

$$p_i, q_i, r_i \geq 0$$

and

$$\sum_{i=1}^{n} p_i = \sum_{i=1}^{n} q_i = \sum_{i=1}^{n} r_i = 1.$$

In what follows, Σ will stand for the sum $\sum_{i=1}^{n}$ unless stated otherwise.

An extensive use of this measure has been made in economic-analysis. In this paper our aim is to characterize some additive generalizations of (I.I.) and to study their properties.

Remarks : For obvious reasons we assume that whenever either p_i or q_i or both zero for a particular i then corresponding p_i is also zero and further we take $0 \log 0 = 0 \log (0/0) = 0$ and all logarithms are considered to the base 2.

---

* *Present address*: N.C.E.R.T., Aurobindo Marg, New Delhi 110016.

2. Generalized Measures of I.I.

$I(P; Q; R)$ given in (1.1) is an ordinary arithmetic average of elementary information improvements $\log (r_i/q_i)$ with weights p_i. Now we define information improvement as an generalized average of elementary information improvements with weights as function f of p_i, that is,

$$I'_\psi (P; Q; R) = \psi^{-1}\left(\frac{\sum f(p_i)\,\psi(\log (r_i/q_i))}{\sum f(p_i)}\right) \qquad \ldots(2.1)$$

$$\left(0 < \sum_{i=1}^{n} p_i, \sum_{i=1}^{n} q_i, \sum_{i=1}^{n} r_i \leq 1\right)$$

where

(A.1) ψ is a strictly monotonic and continuous function;

(A.2) f is a positive valued bounded weight function defined in $(0, 1]$ such that $f(xy) = f(x) f(y)$;

(A.3) $I'_\psi (P_1 * P_2; Q_1 * Q_2; R_1 * R_2) = I'_\psi (P_1; Q_1; R_1) + I'_\psi (P_2; Q_2; R_2)$

where for $P_1 = (p_{11}, \ldots, p_{1n})$, $P_2 = (p_{21}, \ldots, p_{2m})$

$$P_1 * P_2 = (p_{11}p_{21}, \ldots, p_{11}p_{2m}, \ldots, p_{1n}p_{21}, \ldots, p_{1n}p_{2m})$$

etc., and there is one-one correspondence in $P_1 * P_2, Q_1 * Q_2, R_1 * R_2; P_1, Q_1, R_1; P_2, Q_2, R_2$; as induced by the indices of the elements;

(A.4) $I'_\psi (P_1 \cup P_2;\ Q_1 \cup Q_2;\ R_1 \cup R_2)$

$$= \psi^{-1}\left(\frac{\sum_{i=1}^{n} f(p_{1i})\,\psi[I'_\psi (P_1; Q_1; R_1)] + \sum_{j=1}^{m} f(p_{2j})\,\psi[I'_\psi (P_2; Q_2; R_2)]}{\sum_{i=1}^{n} f(p_{1i}) + \sum_{j=1}^{m} f(p_{2j})}\right)$$

where $P_1 \cup P_2 = (p_{11}, \ldots, p_{1n}, p_{21}, \ldots, p_{2m})$

and

$$\sum_{i=1}^{n} p_{1i} + \sum_{j=1}^{m} p_{2j} \leq 1, \text{ etc.}$$

Theorem 1—Under the postulates (A.1) to (A.4) above the function $I'(P; Q; R)$ defined in (2.1) can be only one of the following two forms:

$$I'_1 (P; Q; R) = [\sum f(p_i) \log (r_i/q_i)] \,/\, \sum f(p_i) \qquad \ldots(2.2)$$

$$I'_\alpha (P; Q; R) = (\alpha - 1)^{-1} \log [(\sum f(p_i)\,(r_i/q_i)^{\alpha-1}) \,/\, \sum f(p_i)] \qquad \ldots(2.3)$$

$$(\alpha > 0, \quad \alpha \neq 1)$$

PROOF : By taking $n = 1$ in (2.1) using (A.3), (A.2) and (A.4) we get,

$$\psi^{-1}\left(\frac{\sum f(p_i)\,\psi\,(\log(r_i/q_i)+\log(r/q))}{\sum f(p_i)}\right) = \psi^{-1}\left(\frac{\sum f(p_i)\,\psi\,(\log(r_i/q_i)}{\sum f(p_i)}\right) + \log(r/q) \quad \ldots(2.4)$$

Setting $\log(r_i/q_i) = x_i$, $\log(r/q) = y$ and $\psi(x+y) = \psi_y(x)$ in (2.4), we obtain

$$\psi_y^{-1}\left(\frac{\sum f(p_i)\,\psi_y(x_i)}{\sum f(p_i)}\right) = \psi^{-1}\left(\frac{\sum f(p_i)\,\psi(x_i)}{\sum f(p_i)}\right) \quad \ldots(2.5)$$

which gives (refer Aczél 1966)

$$\psi_y(x) = \psi(x+y) = A(y)\,\psi(x) + B(y), \qquad A \neq 0, \quad \ldots(2.6)$$

For $\psi(x)$ strictly monotonic, there are exactly two solutions of (2.6) given by

$$\psi(x) = ax + b, \qquad a \neq 0 \quad \ldots(2.7)$$

and

$$\psi(x) = a\,e^{cx} + b,\ a \neq 0,\ c \neq 0. \quad \ldots(2.8)$$

Now (2.2) and (2.3) correspond to the above two solutions when $c = (\alpha - 1)\log_e 2$.

Particular cases

(1) For $f(p) = p^\beta$ or p and $p_i = r_i$ for each i, the quantities (2.2) and (2.3) reduce to measures of information studied by Sharma (1972) and Kullback (1959), Rényi (1961) respectively.

(2) If $f(p) = p^\beta$ and $q_i = p_i r_i$ for each i then $I_1^I(P; Q; R)$ and $I_\alpha^I(P; Q; R)$ reduce to the entropies studied by Aczél and Daróczy (1963) and Kapur (1967).

(3) When $q_i = r_i^2$ then we obtain generalized measures of inaccuracy from (2.2) and (2.3).

We shall call $I_\alpha^I(P; Q; R)$ and $I_\alpha^I(P; Q; R)$ as I.I. of order (α, β) and of order $(1, \beta)$ respectively when $f(x) = x^\beta$, $\beta > 0$ in (2.2) and (2.3) and denote as $I_\alpha^\beta(P; Q; R)$ and $I_1^\beta(P; Q; R)$ respectively.

Properties of $I_1^\beta(P; Q; R)$ *and* $I_\alpha^\beta(P; Q; R)$

(1) $I_\alpha^\beta(P; Q; R)$ is a monotonically increasing function of α (for a fixed β).

(2) $I_1^\beta(\{1, 0\}; \{\frac{1}{2}, \frac{1}{2}\}; \{1, 0\}) = I_\alpha^\beta(\{1, 0\}; \{\frac{1}{2}, \frac{1}{2}\}; \{1, 0\}) = 1$.

(3) Both $I_1^\beta(P; Q; R)$ and $I_\alpha^\beta(P; Q; R)$ remain unchanged if the elements in P, Q and R are rearranged in the way that there is one-one correspondence in p_i, q_i and r_i as induced by indices of the elements.

(4) (*a*) If $I^{\beta}(P_1; Q_1; R_1)$ and $I^{\beta}(P_2; Q_2; R_2)$ are defined and there is one-one correspondence in the elements of the generalized probability distributions $P_1 \cup P_2$, $Q_1 \cup Q_2$, and $R_1 \cup R_2$ then

$$I^{\beta}(P_1 \cup P_2; Q_1 \cup Q_2; R_1 \cup R_2)$$

$$= \frac{W_{\beta}(P_1) I^{\beta}(P_1; Q; R_1) + W_{\beta}(P_2) I^{\beta}(P_2; Q_2; R_2)}{W_{\beta}(P_1) + W_{\beta}(P_2)}.$$

(*b*) Also

$$I^{\beta}(P_1 \cup P_2; Q_1 \cup Q_2; R_1 \cup R_2)$$

$$= g^{-1}\left(\frac{W_{\beta}(P_1) g[I^{\beta}(P_1; Q_1; R_1)] + W_{\beta}(P_2) g(I^{\beta}(P_2; Q_2; R_2)]}{W_{\beta}(P_1) + W_{\beta}(P_2)}\right)$$

where $g(x) = a\,2^{(\alpha-1)x} + b$, is a strictly monotonic and continuous function and $W_{\beta}(P) = p_1^{\beta} + p_2^{\beta} + \ldots + p_n^{\beta}$, etc. Also I^{β} in (*a*) and (*b*) is I_1^{β} and I_{α}^{β} respectively.

Next we find that $I_{\alpha}^{\beta}(P; Q; R)$ is positive, zero, negative according as $r_i \gtreqless q_i$ for each i, for all values of $\alpha\,(>1)$ and β. It is easy to establish the following result:

Theorem 2—$I_{\alpha}^{\beta}(P; Q; R)$ is non-negative for all values of $\alpha\,(>0)$ and a fixed β, provided:

$$\Sigma P_i^{\beta} \geq \Sigma\,(p_i^{\beta} q_i \,|\, r_i) \qquad \ldots(2.9)$$

for each i. Further $I_{\alpha}^{\beta}(P; Q; R)$ is negative for all α if

$$\Sigma P_i^{\beta} < \Sigma\,(p_i^{\beta} q_i \,|\, r_i) \qquad \ldots(2.10)$$

for each i, and if for some value of α, say α_1,

$$\Sigma P_i^{\beta} = \Sigma P_i^{\beta}(r_i \,|\, q_i)^{\alpha_1 - 1}$$

then $I_{\alpha}^{\beta}(P; Q; R) > 0$ if $\alpha > \alpha_1$.

Further study of (1.1), refer Autar (1973, 1975) and Sharma-Autar (1974, 1975).

ACKNOWLEDGEMENTS

The author expresses his sincere gratitude to Dr. B. D. Sharma, University of Delhi, for his help in the preparation of this paper.

Thanks are also due to the referee for suggestions to improve the paper.

References

Aczél, J. (1966). Lectures on Functional Equations and their Applications. Academic Press, New York.

Aczél, J., and Daróczy, Z. (1963). Über Verallgemeinerte quasilineare Mittelwerte die mit Gewichtsfunktionen Gebildet Sind. *Pub. Math. Debrecen*, **10**, 171–90.

Aczél, J., and Nath, P. (1972). Axiomatic characterization of some measures of divergence in information. *Z. Wahrs Verw, Geb.*, **21**, 215–24.

Autar, R. (1975). Generalized information improvement functions. *Indian J. pure appl. Math.*, **6** (5), 553–63.

———— (1975). On a characterization of information Improvement. *J. appl. prob.*, **12** (2), 845–51.

Kapur, J. N. (1967). Generalized entropy of order α and type β. *Math. Seminar*, **4**, 78–94.

Kullback, S. (1959). Information Theory and Statistics. John Wiley and Sons, Inc., New York.

Rényi, A. (1961). On Measures of Entropy and Information. Proc. 4th Berk. Symp. Math. Stat. and Prob., Vol. 1, 547–61.

Sharma, B. D. (1972). On amount of information of type β and other measures. *Metrika*, **19**, 1–10.

Sharma, B. D., and Autar, R. (1974). Information improvement functions. *Econometrica*, **42** (1), 103–12.

———— (1975). On information improvement. *Metrika*, **22** (2), 27–34.

Theil, H. (1967). Economics and Information Theory. North-Holland Pub. Co., Amsterdam.