KYBERNETIKA — VOLUME 19 (1983), NUMBER 5

SUB-ADDITIVE MEASURES OF INFORMATION IMPROVEMENT

D. S. HOODA

Additivity plays a great role in the study of information theoretic measures. However, it is very interesting to consider sub-additivity. Starting from sub-additivity for measures associated with three probability distributions of a discrete random variable and using another function of three probability distributions, it has been changed into generalized additivity. Using sum property of the functions and the generalized additivity, a functional equation and its complex solutions are obtained. In terms of the real continuous solutions of this functional equation, three sub-additive measures of information improvement have been defined and characterized. Particular cases and some simple properties including convexity of these new measures have also been studied.

1. INTRODUCTION

Let X be a random variable taking n values $x_1, x_2, \ldots, x_n$ having prediction probability distribution $Q = (q_1, q_2, \ldots, q_n)$, $\sum_{i=1}^{n} q_i \leqq 1$, $q_i > 0$ which is revised as $R = (r_1, r_2, \ldots, r_n)$, $\sum_{i=1}^{n} r_i \leqq 1$, $r_i > 0$ on the basis of a distribution $P = (p_1, p_2, \ldots, p_n)$, $\sum_{i=1}^{n} p_i = 1$, $p_i \geqq 0$ supposed to have been realized after some experiment, then the information theoretic measure associated with these three probability distributions P, Q and R is given by

$$I(P; Q; R) = \sum_{i=1}^{n} p_i \log_2 (r_i/q_i) . \tag{1.1}$$

The measure (1.1) is called Theil's [7] measure of information improvement and it has many applications in economics. The measure (1.1) satisfies the property of additivity which can be expressed as

$$I(P*P'; Q*Q'; R*R') = I(P; Q; R) + I(P'; Q'; R') \tag{1.2}$$

where $$P = (p_1, p_2, \ldots, p_n); \quad P' = (p'_1, p'_2, \ldots, p'_m);$$
$$P*P' = (p_1p'_1, \ldots, p_1p'_m, \ldots; p_np'_1, \ldots, p_np'_m) \quad \text{etc.}$$

Using sum property given by

$$I(P; Q; R) = \sum_{i=1}^{n} h(p_i, q_i, r_i), \tag{1.3}$$

some generalizations of the measure (1.1) have been studied by Sharma and Soni [5] and by Taneja [6].

Sharma and Taneja [4] have studied three measures of entropy satisfying the sub-additivity

$$H(P_1*P_2) \leqq H(P_1) + H(P_2) \tag{1.4}$$

and using another function G of a probability distribution such that

$$H(P_1*P_2) = H(P_1)\,G(P_2) + H(P_2)\,G(P_1), \tag{1.5}$$

where $G(P_1)$ and $G(P_2)$ both take values not exceeding unity. The property (1.5) can be said as generalized additivity. The three measures of inaccuracy and relative-information associated with a pair of probability distributions and satisfying the generalized additivity

$$H(P_1*P_2; Q_1*Q_2) = H(P_1; Q_1)\,G(P_2; Q_2) + H(P_2; Q_2)\,G(P_1; Q_1) \tag{1.6}$$

have been studied by Sharma and Gupta [3] and by Gupta [2].

In this communication, we study three sub-additive measures associated with three discrete probability distributions. Simple properties including convexity of these measures and particular cases have also been studied.

2. GENERALIZED ADDITIVITY AND FUNCTIONAL EQUATION

Let $I(P; Q; R)$ be an information theoretic measure satisfying

$$I(P_1*P_2; Q_1*Q_2; R_1*R_2) \leqq I(P_1; Q_1; R_1) + I(P_2; Q_2; R_2) \tag{2.1}$$

Next let G be another function of three probability distributions satisfying

$$I(P_1*P_2; Q_1*Q_2; R_1*R_2) = I(P_1; Q_1; R)\,G(P_2; Q_2; R_2) + \\ + I(P_2; Q_2; R_2)\,G(P_1; Q_1; R_1) \tag{2.2}$$

The relation (2.2) can be said as generalized additivity of information improvement. Now we suppose that

$$I(P; Q; R) = \sum_{i=1}^{n} h(p_i, q_i, r_i) \tag{2.3}$$

$$G(P; Q; R) = \sum_{i=1}^{n} g(p_i, q_i, r_i). \tag{2.4}$$

Using (2.3) and (2.4) in (2.2) we have the functional equation

$$\sum_{i=1}^{n}\sum_{j=1}^{m} h(p_{1i}, p_{2j}; q_{1i}q_{2j}; r_{1i}r_{2j}) = \sum_{i=1}^{n}\sum_{j=1}^{m} h(p_{1i}, q_{1i}, r_{1i}) \cdot g(p_{2j}, q_{2j}, r_{2j}) + \sum_{i=1}^{n}\sum_{j=1}^{m} h(p_{2j}, q_{2j}, r_{2j})\, g(p_{1i}, q_{1i}, r_{1i}), \tag{2.5}$$

where

$$q_{1i}, q_{2j}, r_{1i}, r_{2j} \in (0, 1] \quad \text{and} \quad p_{1i}, p_{2j} \in [0, 1].$$

The continuous functions h and g that satisfy the functional equation (2.5) are the continuous solutions of the functional equation

$$h(xx', yy', zz') = h(x, y, z)\, g(x', y', z') + g(x, y, z)\, h(x', y', z') \tag{2.6}$$

where

$$y, y', z, z' \in (0, 1] \quad \text{and} \quad x, x' \in [0, 1].$$

Therefore, we find the real continuous solutions of (2.6) in the following theorem:

Theorem 1. The most general complex solutions of (2.6) are given by

$$h(x, y, z) = 0, \quad g(x, y, z) \text{ arbitrary} \tag{2.7}$$

$$h(x, y, z) = e_0(x, y, z)\, a(x, y, z); \quad g(x, y, z) = e_0(x, y, z) \tag{2.8}$$

and

$$h(x, y, z) = \frac{1}{2k}\left[e_1(x, y, z) - e_2(x, y, z)\right]; \quad g(x, y, z) = \tfrac{1}{2}\left[e_1(x, y, z) + e_2(x, y, z)\right], \tag{2.9}$$

where $k \neq 0$ is an arbitrary complex constant and $a(x, y, z)$, $e_j(x, y, z)$ $(j = 0, 1, 2)$ are arbitrary functions satisfying respectively

$$a(xx', yy', zz') = a(x, y, z) + a(x', y', z') \tag{2.10}$$

and

$$e_j(xx', yy', zz') = e_j(x, y, z)\, e_j(x', y', z') \quad (j = 0, 1, 2). \tag{2.11}$$

The proof when functions are of single variable will be found in Aczél [1], p. 205. The above result also follows on the same lines with suitable modifications.

Real Continuous Solutions of (2.6)

The real continuous solutions of (2.6) depend on solutions of the well-known in auxiliary equations (2.10) and (2.11). If we substitute the solutions of (2.10) and (2.11) in the solutions given by (2.8) and (2.9) respectively, these take the form

$$(2.12)\qquad h(x, y, z) = x^{\alpha}y^{\beta}z^{\gamma}(c_1 \log x + c_2 \log y + c_3 \log z)\,,$$
$$g(x, y, z) = x^{\alpha}y^{\beta}z^{\gamma}\,,$$

where α, β, γ, c_1, c_2, c_3 are arbitrary complex constants.

$$(2.13)\qquad h(x, y, z) = \frac{1}{2k}\,(x^{\alpha}y^{\beta}z^{\gamma} - x^{\delta}y^{\mu}z^{\nu})\,;$$
$$g(x, y, z) = \tfrac{1}{2}(x^{\alpha}y^{\beta}z^{\gamma} + x^{\delta}y^{\mu}z^{\nu})\,,$$

where α, β, γ, δ, μ, ν and k are arbitrary complex constants. Further, we see that $g(x, y, z)$ in (2.12) would be real iff α, β, γ are real and it would be continuous if α, β and γ are non-negative. It follows that corresponding $h(x, y, z)$ would be real iff c_1, c_2, c_3 are real and α, β, γ are non-negative. Thus one set of real and continuous solutions of (2.6) is given by

$$(2.14)\qquad h(x, y, z) = x^{\alpha}y^{\beta}z^{\gamma}(c_1 \log x + c_2 \log y + c_3 \log z)\,,$$
$$g(x, y, z) = x^{\alpha}y^{\beta}z^{\gamma}\,,$$

where $\alpha > 0$, $\beta \geqq 0$, $\gamma \geqq 0$ and c_1, c_2, c_3 are arbitrary real constants.

Now $g(x, y, z)$ in (2.13) would be real only under the following sets of conditions:

(i) α, β, γ, δ, μ, ν are all real or

(ii) α, β, γ, are complex conjugate of δ, μ, ν respectively.

The continuity of $g(x, y, z)$ requires that α, β, γ, δ, μ, ν are all non-negative. When $g(x, y, z)$ in (2.13) is real, corresponding $h(x, y, z)$ is also real iff k is real. Thus one of the other two sets of real continuous solutions of (2.6) obtained from (2.13) is given by

$$(2.15)\qquad h(x, y, z) = \frac{1}{2k}\,(x^{\alpha}y^{\beta}z^{\gamma} - x^{\delta}y^{\mu}z^{\nu})\,,$$
$$g(x, y, z) = \tfrac{1}{2}(x^{\alpha}y^{\beta}z^{\gamma} + x^{\delta}y^{\mu}z^{\nu})\,,$$

where α, β, γ, δ, μ, ν (all non-negative) and k are real arbitrary constants.

For second set of solutions, let $\alpha = \alpha_1 + i\alpha_2$; $\beta = \beta_1 + i\beta_2$; $\gamma = \gamma_1 + i\gamma_2$; $\delta = \alpha_1 - i\alpha_2$; $\mu = \beta_1 - i\beta_2$; $\nu = \gamma_1 - i\gamma_2$; $k = iR$, then (2.13) gives

$$(2.16)\qquad h(x, y, z) = \frac{1}{R}\, y^{\alpha_1}y^{\beta_1}z^{\gamma_1} \sin(\alpha_2 \log x + \beta_2 \log y + \gamma_2 \log z)\,,$$
$$g(x, y, z) = x^{\alpha_1}y^{\beta_1}z^{\gamma_1} \cos(\alpha_2 \log x + \beta_2 \log y + \gamma_2 \log z).$$

Taking α, β, γ, δ, μ, ν for α_1, β_1, γ_1, α_2, β_2, γ_2 respectively in (2.16), we have the third set of solutions given by

$$(2.17) \qquad h(x, y, z) = \frac{1}{R} x^\alpha y^\beta z^\gamma \sin(\delta \log x + \mu \log y + \nu \log z),$$

$$g(x, y, z) = x^\alpha y^\beta z^\gamma \cos(\delta \log x + \mu \log y + \nu \log z),$$

where $\alpha(>0)$, $\beta(\geqq 0)$, $\gamma(\geqq 0)$, δ, μ, ν and R are real constants. Hence (2.14), (2.15) and (2.17) are the only three non-trivial sets of real and continuous solutions of the functional equation (2.6) for $x \in [0, 1]$ and $y, z \in (0, 1]$.

3. CHARACTERIZATION OF INFORMATION IMPROVEMENT UNDER GENERALIZED ADDITIVITY

We adopt the following definition:

Information Improvement. The measure of information improvement $I(P; Q; R)$ associated with three discrete probability distributions P, Q and R is given by

$$(3.1) \qquad I(P; Q; R) = \sum_{i=1}^{n} h(p_i, q_i, r_i)$$

where $h(p, q, r)$ is a real continuous solution of (2.5) under the conditions

$$(3.2) \qquad h(\tfrac{1}{2}, \tfrac{1}{2}, \tfrac{1}{2}) = 0, \quad h(1, \tfrac{1}{2}, \tfrac{1}{2}) = 0 \quad \text{and} \quad h(1, 1, \tfrac{1}{2}) = -1.$$

Now we characterize sub-additive measures of information improvement in the next theorem which follow from Theorem 1 and sum property.

Theorem 2. Corresponding to the real continuous solutions (2.14), (2.15) and (2.17), the three sub-additive measures of information improvement satisfying (2.2) can be only one of the following three forms:

$$(3.3) \qquad I^l(P; Q; R : \alpha, \beta, \gamma) = 2^\gamma \sum_{i=1}^{n} p_i^\alpha q_i^\beta r_i^\gamma \log_2 (r_i/q_i),$$

$$\alpha > 0, \quad \beta \geqq 0, \quad \gamma \geqq 0,$$

$$(3.4) \qquad I^p(P; Q; R : \alpha, \beta, \gamma, \delta) = (2^{\delta-\gamma} - 2^{\beta-\gamma})^{-1} \sum_{i=1}^{n} p_i^\alpha (q_i^\beta r_i^{\gamma-\beta} - q_i^\delta r_i^{\gamma-\delta}),$$

$$\alpha > 0, \quad \beta \geqq 0, \quad \delta \geqq 0, \quad \beta \neq \gamma, \quad \delta \neq \gamma$$

and

$$(3.5) \qquad I^s(P; Q; R : \alpha, \beta, \gamma, \delta) = \frac{2^\gamma}{\sin \delta} \sum_{i=1}^{n} p_i^\alpha q_i^\beta r_i^\gamma \sin\left(\delta \log_2 \frac{r_i}{q_i}\right),$$

$$\alpha > 0, \quad \beta \geqq 0, \quad \gamma \geqq 0, \quad \delta \neq 0.$$

4. PARTICULAR CASES

(a) Taking $\alpha = 1$, $\beta = 0$, $\gamma = 0$ in (3.3), we get

$$I^I(P; Q; R : 1, 0, 0) = \sum_{i=1}^{n} p_i \log_2 (r_i/q_i),$$

which is Theil's [7] measure of information improvement.

(b) Taking $\beta = \gamma = \alpha - 1$ and $\delta = 0$ in (3.4), we have

$$I^P(P; Q; R : \alpha, \alpha - 1, \alpha - 1, 0) = (2^{1-\alpha} - 1)^{-1} \sum_{i=1}^{n} p_i^\alpha (q_i^{\alpha-1} - r_i^{\alpha-1})$$

which is information improvement of order α. Further we have

$$\lim_{\alpha \to 1} I^P(P; Q; R : \alpha, \alpha - 1, \alpha - 1, 0) = \sum_{i=1}^{n} p_i \log_2 (r_i/q_i),$$

which is Theil's [7] measure of information improvement.

(c) We see that

$$\lim_{\delta \to 0} I^s(P; Q; R : \alpha, \beta, \gamma, \delta) = 2^\gamma \sum_{i=1}^{n} p_i^\alpha q_i^\beta r_i^\gamma \log_2 (r_i/q_i)$$

which is (3.3).

5. PROPERTIES

Some of the common simple properties of the three subadditive measures of information improvement are enlisted below:

(a) Generalized additivity
(b) Sub-additivity
(c) Sum property
(d) Symmetry with respect to its arguments
(e) $I_n(P; Q; Q) = 0$.

Next we discuss the convexity of the sub-additive measure $I^P(P; Q; R; \alpha, \beta, \gamma, \delta)$ with respect to the probability distributions Q and R.

Theorem 3. The sub-additive measure of information improvement $I^P(P; Q; R : : \alpha, \beta, \gamma, \delta)$ is a convex $\cap$ function of the probability distribution Q whenever $\beta < 1 < < \delta$ or $\delta < 1 < \beta$.

Proof. Let us consider r probability distributions

$$Q_j(X) = \{q_j(x_1), \ldots, q_j(x_n)\}, \quad q_j(x_i) > 0, \quad \sum_{i=1}^{n} q_j(x_i) = 1,$$

$j = 1, 2, \ldots, r$ and a probability distribution

$$Q_0(X) = \{q_0(x_1), \ldots, q_0(x_n)\} \quad \text{of } X \text{ such that} \quad q_0(x_i) = \sum_{j=1}^{r} a_j\, q_j(x_i),$$

$i = 1, 2, \ldots, n$, where a_j's are non-negative numbers such that $\sum_{j=1}^{r} a_j = 1$. The probability distribution $Q_0(X)$ is a bonafide probability distribution of X since $\sum_{i=1}^{n} q_0(x_i) = \sum_{i=1}^{n} \sum_{j=1}^{r} a_j q_j(x_i) = 1$. Let

$$\Delta = I^p(P(X);\ Q_0(X);\ R(X) : \alpha, \beta, \gamma, \delta) - \sum_{j=1}^{r} a_j I^p(P(X);\ Q_j(X);\ R(X);\ \alpha, \beta, \gamma, \delta)\,.$$

Then $I^p(P; Q; R : \alpha, \beta, \gamma, \delta)$ will be a convex $\cap$ or $\cup$ function of the probability distribution Q according as $\Delta \gtrless 0$.

Now we have

$$(5.1) \quad \Delta = (2^{\delta-\gamma} - 2^{\beta-\gamma})^{-1} \Big[\sum_{i=1}^{n} p^\alpha(x_i) \{q_0^\beta(x_i)\, r^{\gamma-\beta}(x_i) - q_0^\delta(x_i)\, r^{\gamma-\delta}(x_i)\} - \sum_{j=1}^{r} a_j \sum_{i=1}^{n} p^\alpha(x_i) \{q_j^\beta(x_i)\, r^{\gamma-\beta}(x_i) - q_j^\delta(x_i)\, r^{\gamma-\delta}(x_i)\}\Big] =$$

$$= (2^{\delta-\gamma} - 2^{\beta-\gamma})^{-1} \Big[\sum_{i=1}^{n} p^\alpha(x_i) \{(\sum_{j=1}^{r} a_j q_j(x_i))^\beta\, r^{\gamma-\beta}(x_i) - (\sum_{j=1}^{r} a_j q_j(x_i))^\delta\, r^{\gamma-\delta}(x_i)\} - \sum_{i=1}^{n} p^\alpha(x_i) \{ \sum_{j=1}^{r} a_j q_j^\beta(x_i)\, r^{\gamma-\beta}(x_i) - \sum_{j=1}^{r} a_j q_j^\delta(x_i)\, r^{\gamma-\delta}(x_i)\}\Big] =$$

$$= (2^{\delta-\gamma} - 2^{\beta-\gamma})^{-1} \sum_{i=1}^{n} p^\alpha(x_i) \Big[\{(\sum_{j=1}^{r} a_j q_j(x_i))^\beta - \sum_{j=1}^{r} a_j q_j^\beta(x_i)\}\, r^{\gamma-\beta}(x_i) - \{(\sum_{j=1}^{r} a_j q_j(x_i))^\delta - \sum_{j=1}^{r} a_j q_j^\delta(x_i)\}\, r^{\gamma-\delta}(x_i)\Big]\,.$$

Now by Jensen's inequality

$$(5.2) \qquad (\sum_{j=1}^{r} a_j q_j(x_i))^k \gtreqless \sum_{j=1}^{r} a_j q_j^k(x_i)\,,$$

according as $k \lessgtr 1$ with equality iff $q_j(x_i)$ are constants. Further we have

$$(5.3) \qquad (2^{\delta-\gamma} - 2^{\beta-\gamma})^{-1} \gtrless 0$$

according as $\beta \lessgtr \delta$.

By taking $\beta < 1 < \delta$ or $\delta < 1 < \beta$ it follows from (5.1), (5.2) and (5.3) that $\Delta > 0$. The result of the theorem is now obvious. □

Theorem 4. The sub-additive measure of information improvement $I^p(P; Q; R : \alpha, \beta, \gamma, \delta)$ is a convex $\cap$ function of the probability distribution R whenever $\gamma - \beta < 1 < \gamma - \delta$ or $\gamma - \delta < 1 < \gamma - \beta$.

The proof is exactly similar to that of Theorem 3.

Theorem 5. The sub-additive measures of information improvement $I^t(P; Q; R : : \alpha, \beta, \gamma)$, $I^p(P; Q; R : \alpha, \beta, \gamma, \delta)$ and $I^s(P; Q; R : \alpha, \beta, \gamma, \delta)$ are convex $\cap$ or $\cup$ functions of the probability distribution P according as $\alpha \lessgtr 1$.

ACKNOWLEDGEMENTS

The author is thankful to Dr. R. K. Tuteja, Department of Mathematics, M. D. University, Rohtak. Thanks are also due to Professor O. P. Srivastava, Dean, College of Basic Science, H.A.U., Hissar for his constant encouragement.

(Received November 11, 1981.)

REFERENCES

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

[2] H. C. Gupta: A Study on Sub-Additive and Non-Additive Measures of Information. Ph. D. Thesis, University of Delhi, India 1977.

[3] B. D. Sharma and H. C. Gupta: Sub-additive measures of relative information and inaccuracy. Metrika *23* (1976), 155—165.

[4] B. D. Sharma and I. J. Taneja: Three generalized-additive measures of entropy. Elektron. Informationsverarb. Kybernet. *13* (1977), 419—433.

[5] B. D. Sharma and R. S. Soni: Additive and non-additive measures of information of comparison and information improvement. Elektron. Informationsverarb. Kybernet. *11* (1975), 489—504.

[6] I. J. Taneja: A generalized functional equation of three variables in information theory. Funkcial. Ekvac. *19* (1976), 139—147.

[7] H. Theil: Economics and Information Theory. North Holland, Amsterdam 1967.

Dr. D. S. Hooda, Department of Mathematics and Statistics, Haryana Agricultural University, Hissar-125004. India.