KYBERNETIKA — VOLUME 7 (1971), NUMBER 5

On Some New Measures of Uncertainty, Inaccuracy and Information and their Characterizations

P. N. RATHIE

Various measures of uncertainty, inaccuracy and information are defined in this paper which generalize several well known measures of uncertainty, inaccuracy and information. These generalized measures are given in (1.1), (1.3), (1.5) and (1.7) respectively. These are characterized by means of functional equations.

1. INTRODUCTION AND DEFINITIONS

Consider a finite discrete probability distribution $P = (p_1, \ldots, p_n)$ with $p_i > 0$, $W(P) = \sum_{i=1}^{n} p_i \leqq 1$. Let Δ denote the set of all finite discrete distributions.

Throughout this paper, Σ will stand for the sum $\sum_{i=1}^{n}$ and logarithms will be taken to the base 2 unless otherwise specified.

Now we define the various measures of uncertainty, inaccuracy and information one by one. First of all we define the β-entropy for the distribution P as

$$H^\beta\langle P\rangle = -\sum p_i^{\beta+1} \log p_i / \sum p_i , \tag{1.1}$$

which is clearly a generalization of the well known entropy,

$$H(P) = -\sum p_i \log p_i / \sum p_i , \tag{1.2}$$

introduced earlier in [5, 6].

Consider $P = (p_1, \ldots, p_n) \in \Delta$ and $Q = (q_1, \ldots, q_n) \in \Delta$ as the two generalized probability distributions, the correspondence between the elements of P and Q is that given by their subscripts. Then we define the β-inaccuracy as

$$H^\beta\langle P : Q\rangle = -\sum p_i^{\beta+1} \log q_i / \sum p_i , \tag{1.3}$$

which when $\beta = 0$ reduces to the inaccuracy [2],

(1.4) $$H(P \mid Q) = -\sum p_i \log q_i / \sum p_i \, .$$

Also we define the β-information as,

(1.5) $$I^\beta\langle P : Q\rangle = \sum p_i^{\beta+1} \log (p_i/q_i)/\sum p_i \, ;$$

which yields the information-gain [5] or the directed divergence [3] for $\beta = 0$ as given below

(1.6) $$I(P \mid Q) = \sum p_i \log (p_i/q_i)/\sum p_i \, .$$

Lastly, we define another generalized measure of information and inaccuracy as follows:

(1.7) $$\hat{I}_\alpha^\beta\langle P : Q\rangle = (1 - \sum p_i^{\beta+1} q_i^{1-\alpha}/\sum p_i)/(1 - 2^{\alpha-1}) \, , \quad \alpha \neq 1 \, .$$

We shall call the measure $\hat{I}_\alpha^\beta\langle P : Q\rangle$ as (α, β)-inaccuformation. Clearly for $\beta = \alpha - 1$, (1.7) gives the non-additive information of order α defined by,

(1.8) $$\hat{I}_\alpha^{\alpha-1}\langle P : Q\rangle = (1 - \sum p_i^\alpha q_i^{1-\alpha}/\sum p_i)/(1 - 2^{\alpha-1}) \, , \quad \alpha \neq 1 \, .$$

Also for $\beta = 0$, (1.7) gives the non-additive inaccuracy of order α, defined by

(1.9) $$\hat{I}_\alpha^0\langle P : Q\rangle = (1 - \sum p_i q_i^{1-\alpha}/\sum p_i)/(1 - 2^{\alpha-1}) \, , \quad \alpha \neq 1 \, .$$

When $p_i = q_i$ for all i, then (1.9) where α is replaced by $2 - \alpha$ is the non-addrtive entropy introduced earlier in [7].

The non-additive information of order α defined in (1.8) was recently characterized by the author [4].

If p_i's are allowed to take zero values and the convention $0 \log 0 = 0$ is followed, then we have to impose the restriction $\beta > -1$ in (1.1), (1.3), (1.5) and (1.7) and the restriction $\alpha > 0$ in (1.8).

The object of this paper is to prove characterization theorems for the various generalized measures defined in (1.1), (1.3), (1.5) and (1.7) to (1.9) by means of functional equations. Section 2 deals with the functional equations useful for the characterization purposes and their solutions. Section 3 deals with the characterization theorems for (1.1), (1.3) and (1.5). In section 4, characterization theorems for (1.7), (1.8) and (1.9) are proved. All the characterization theorems proved in this paper are for $n = 2$ from which the corresponding results for any $n > 2$ can be derived.

This section deals with the solutions of the following functional equations:

$$f(pq) = p^\beta f(q) + q^\beta f(p), \tag{2.1}$$

where $p, q \in (0, 1]$ and $f(p)$ is a continuous function of $p \in (0, 1]$.

$$f(p_1 p_2, q_1 q_2) = p_1^\beta f(p_2, q_2) + p_2^\beta f(p_1, q_1), \tag{2.2}$$

where $p_1, p_2, q_1, q_2 \in (0, 1]$; $f(p, 1)$ and $f(1, q)$ are continuous functions of $p \in (0, 1]$ and $q \in (0, 1]$ respectively.

$$f(p_1 p_2, q_1 q_2) = f(p_1, q_1) + f(p_2, q_2) + (2^{\alpha-1} - 1) f(p_1, q_1) f(p_2, q_2), \tag{2.3}$$

where $\alpha \neq 1$, $p_1, p_2, q_1, q_2 \in (0, 1]$; $f(p, 1)$ and $f(1, q)$ are continuous functions of $p \in (0, 1]$ and $q \in (0, 1]$ respectively.

Solution of (2.1)

Rewriting (2.1) in the form,

$$(pq)^{-\beta} f(pq) = q^{-\beta} f(q) + p^{-\beta} f(p) \tag{2.4}$$

and setting

$$F_\beta(p) = p^{-\beta} f(p), \quad p \in (0, 1], \tag{2.5}$$

we have

$$F_\beta(pq) = F_\beta(p) + F_\beta(q). \tag{2.6}$$

Since $f(p)$ is a continuous function of $p \in (0, 1]$ and therefore $F_\beta(p)$ is also a continuous function of p. Thus the continuous solution [1, p. 41] of (2.6) are given by

$$F_\beta(p) = c \log p, \quad p \in (0, 1]: \tag{2.7}$$

Hence from (2.5) and (2.7), we get

$$f(p) = cp^\beta \log p, \tag{2.8}$$

where c is an arbitrary real constant.

Solution of (2.2)

Taking $p_1 = 1$, $p_2 = p$, $q_1 = q$, $q_2 = 1$ in (2.2), we have

$$f(p, q) = f(p, 1) + p^\beta f(1, q). \tag{2.9}$$

Now we have to find expressions for $f(p, 1)$ and $f(1, q)$ respectively so that $f(p, q)$ may be completely known from (2.9). For this, let us take $p_1 = p_2 = 1$ in (2.2) so that we have

$$f(1, q_1 q_2) = f(1, q_1) + f(1, q_2). \tag{2.10}$$

By [1, p. 41], the continuous solution of (2.10) is given by

$$f(1, q) = b \log q, \quad q \in (0, 1], \tag{2.11}$$

where b is an arbitrary real constant.

Again, taking $q_1 = q_2 = 1$ in (2.2), we have

$$f(p_1 p_2, 1) = p_1^\beta f(p_2, 1) + p_2^\beta f(p_1, 1), \tag{2.12}$$

which on multiplying throughout by $(p_1 p_2)^{-\beta}$ and setting

$$F_\beta(p) = p^{-\beta} f(p, 1),\ p \in (0,1] \tag{2.13}$$

gives

$$G_\beta(p_1 p_2) = G_\beta(p_1) + G_\beta(p_2). \tag{2.14}$$

As $f(p, 1)$ is continuous for $p \in (0, 1]$, therefore $G_\beta(p)$ is also continuous for $p \in$ $\in (0, 1]$. Hence the continuous solution [1, p. 41] of (2.14) is given by

$$G_\beta(p) = a \log p. \tag{2.15}$$

Hence (2.13) with the help of (2.15) gives

$$f(p, 1) = a p^\beta \log p, \quad p \in (0, 1], \tag{2.16}$$

where a is an arbitrary real constant. Thus (2.9), (2.11) and (2.16) give

$$f(p, q) = p^\beta [a \log p + b \log q], \tag{2.17}$$

where a and b are arbitrary real constants.

Solution of (2.3)

Taking $p_1 = p$, $p_2 = q_1 = 1$, $q_2 = q$ in (2.3), we have

$$f(p, q) = f(p, 1) + f(1, q) + (2^{\alpha-1} - 1) f(p, 1) f(1, q). \tag{2.18}$$

Thus in order to determine $f(p, q)$ we must find $f(p, 1)$ and $f(1, q)$ separately.

Again taking $q_1 = q_2 = 1$ in (2.3), we get

$$f(p_1 p_2, 1) = f(p_1, 1) + f(p_2, 1) + (2^{\alpha-1} - 1) f(p_1, 1) f(p_2, 1). \tag{2.19}$$

Let

$$g_\alpha(p) = 1 + (2^{\alpha-1} - 1)f(p, 1), \quad p \in (0, 1]. \tag{2.20}$$

Then (2.19) takes the following form

$$g_\alpha(p_1 p_2) = g_\alpha(p_1)\, g_\alpha(p_2). \tag{2.21}$$

Since $f(p, 1)$ is a continuous function of $p \in (0, 1]$, therefore $g_\alpha(p)$ is continuous. Hence by [1, p. 41] the continuous non-zero solutions of (2.21) are given by

$$g_\alpha(p) = p^\beta, \tag{2.22}$$

where β is an arbitrary real constant.

Thus (2.20) by the use of (2.22) gives

$$f(p, 1) = (p^\beta - 1)/(2^{\alpha-1} - 1). \tag{2.23}$$

Similarly on taking $p_1 = p_2 = 1$ in (2.3) and following the procedure given above, we get

$$f(1, q) = (q^\nu - 1)/(2^{\alpha-1} - 1), \tag{2.24}$$

where ν is an arbitrary real constant.

Hence (2.18) with the help of (2.23) and (2.24) gives

$$f(p, q) = (p^\beta q^\nu - 1)/(2^{\alpha-1} - 1), \tag{2.25}$$

where β and ν are arbitrary real constants.

3. CHARACTERIZATION THEOREMS FOR (1.1), (1.3) AND (1.5) FOR $n = 2$

In this section some characterization theorems for (1.1), (1.3) and (1.5) when $n = 2$ are proved by using the functional equations (2.1) and (2.2) and their solutions described in the last section.

We assume the following four postulates to prove a characterization theorem for (1.1) when $n = 2$.

Postulate 1. $H^\beta\langle p\rangle$ is a continuous function of $p \in (0, 1]$.

Postulate 2. $H\langle\frac{1}{2}\rangle = (\frac{1}{2})^\beta$.

Postulate 3. For $p, q \in \Delta$, we have

$$H^\beta\langle pq\rangle = p^\beta H^\beta\langle q\rangle + q^\beta H^\beta\langle p\rangle.$$

Postulate 4. For $P = (p_1, p_2) \in \Delta$, we have

$$H^\beta\langle P\rangle = \sum_{i=1}^{2} p_i H^\beta\langle p_i\rangle / \sum_{i=1}^{2} p_i .$$

Theorem 1. *The function satisfying the postulate* 1, 2, 3 *and* 4 *is the β-entropy,*

$$H^\beta\langle P\rangle = -\sum_{i=1}^{2} p_i^{\beta+1} \log p_i / \sum_{i=1}^{2} p_i . \tag{3.1}$$

Proof. Postulates 1 and 3 are equivalent to (2.1) and hence from (2.8), we have

$$H^\beta\langle p\rangle = c p^\beta \log p . \tag{3.2}$$

Thus (3.2) on using the postulate 2 yields $c = -1$ giving

$$H^\beta\langle p\rangle = -p^\beta \log p . \tag{3.3}$$

Hence (3.3) and the postulate 4 proves theorem 1.

Now, let us assume that $H^\beta\langle P : Q\rangle$ satisfies the following five postulates:

Postulate 1. $H^\beta\langle p : 1\rangle$ and $H^\beta\langle 1 : q\rangle$ are continuous functions of $p \in (0, 1]$ and $q \in (0, 1]$ respectively.

Postulate 2. $H^\beta\langle 1 : \frac{1}{2}\rangle = 1$.

Postulate 3. $H^\beta\langle \frac{1}{2} : 1\rangle = 0$.

Postulate 4. For $p_1, p_2, q_1, q_2 \in \Delta$, we have

$$H^\beta\langle p_1 p_2 : q_1 q_2\rangle = p_1^\beta H^\beta\langle p_2 : q_2\rangle + p_2^\beta H^\beta\langle p_1 : q_1\rangle .$$

Postulate 5. For $P = (p_1, p_2) \in \Delta$ and $Q = (q_1, q_2) \in \Delta$, we have

$$H^\beta\langle P : Q\rangle = \sum_{i=1}^{2} p_i H^\beta\langle p_i : q_i\rangle / \sum_{i=1}^{2} p_i .$$

Theorem 2. *The function satisfying the postulates* 1, 2, 3, 4 *and* 5 *is the β-inaccuracy,*

$$H^\beta\langle P : Q\rangle = -\sum_{i=1}^{2} p_i^{\beta+1} \log q_i / \sum_{i=1}^{2} p_i . \tag{3.4}$$

Proof. Postulates 1 and 4 are equivalent to (2.2) and hence from (2.17), we have

$$H^\beta\langle p : q\rangle = p^\beta[a \log p + b \log q] . \tag{3.5}$$

On using postulates 2 and 3 in (3.5), we get $b = -1$ and $a = 0$ respectively giving,

$$H^\beta\langle p : q\rangle = -p^\beta \log q . \tag{3.6}$$

Hence the use of (3.6) in postulate 5 proves theorem 2.

In order to prove a characterization theorem for β-information (1.5) let us assume the following postulates:

Postulate 1. $I^\beta\langle p : 1\rangle$ and $I^\beta\langle 1 : q\rangle$ are continuous functions of $p \in (0, 1]$ and $q \in (0, 1]$ respectively.

Postulate 2. $I^\beta\langle 1 : \frac{1}{2}\rangle = 1$.

Postulate 3. $I^\beta\langle \frac{1}{2} : \frac{1}{2}\rangle = 0$.

Postulate 4. For $p_1, p_2, q_1, q_2 \in \Delta$, we have

$$I^\beta\langle p_1 p_2 : q_1 q_2\rangle = p_1^\beta I^\beta\langle p_2 : q_2\rangle + P_2^\beta I^\beta\langle p_1 : q_1\rangle .$$

Postulate 5. For $P = (p_1, p_2) \in \Delta$ and $Q = (q_1, q_2) \in \Delta$, we have

$$I^\beta\langle P : Q\rangle = \sum_{i=1}^{2} p_i I^\beta\langle p_i : q_i\rangle / \sum_{i=1}^{2} p_i .$$

Theorem 3. *The function satisfying the postulates* 1, 2, 3, 4 *and* 5 *is the β-information,*

$$I^\beta\langle P : Q\rangle = \sum_{i=1}^{2} p_i^{\beta+1} \log (p_i/q_i) / \sum_{i=1}^{2} p_i . \tag{3.7}$$

Proof. The postulates 1 and 4 are equivalent to (2.2) and therefore from (2.17) we get

$$I^\beta\langle p : q\rangle = p^\beta[a \log p + b \log q] . \tag{3.8}$$

Now using postulates 2 and 3 in (3.8), we get $b = -1$ and $a = 1$. Thus

$$I^\beta\langle p : q\rangle = p^\beta \log (p/q) . \tag{3.9}$$

Hence the use of (3.9) in postulate 5 proves theorem 3.

4. CHARACTERIZATION THEOREMS FOR (1.7), (1.8) AND (1.9) FOR $n = 2$

This section deals with the characterization theorems for (1.7), (1.8) and (1.9) when $n = 2$ by using the solution to the functional equation (2.3). We start by assuming the following postulates:

Postulate 1. $\hat{I}_\alpha\langle p : 1\rangle$ and $\hat{I}_\alpha\langle 1 : q\rangle$ are continuous functions of $p \in (0; 1]$ and $q \in (0, 1]$ respectively.

Postulate 2. $\hat{I}_\alpha\langle 1 : \frac{1}{2}\rangle = 1$.

Postulate 3. $\hat{I}_\alpha\langle \frac{1}{2} : \frac{1}{2}\rangle = 0$.

Postulate 4. $\hat{I}_\alpha\langle \frac{1}{2} : 1\rangle = 0$.

Postulate 5. If $p_1, p_2, q_1, q_2 \in \Delta$, then

$$\hat{I}_\alpha\langle p_1 p_2 : q_1 q_2\rangle = \hat{I}_\alpha\langle p_1 : q_1\rangle + \hat{I}_\alpha\langle p_2 : q_2\rangle + \\ + (2^{\alpha-1} - 1)\hat{I}_\alpha\langle p_1 : q_1\rangle \hat{I}_\alpha\langle p_2 : q_2\rangle .$$

Postulate 6. If $P = (p_1, p_2) \in \Delta$ and $Q = (q_1, q_2) \in \Delta$, then

$$\hat{I}_\alpha\langle P : Q\rangle = \sum_{i=1}^{2} p_i \hat{I}_\alpha\langle p_i : q_i\rangle / \sum_{i=1}^{2} p_i .$$

Now we proceed to prove characterization theorems for (1.7), (1.8) and (1.9) for $n = 2$.

Theorem 4. *The function satisfying the postulates* 1, 2, 5 *and* 6 *is the* (α, β)-*inaccuformation,*

$$\hat{I}_\alpha^\beta\langle P : Q\rangle = \Big(1 - \sum_{i=1}^{2} p_i^{\beta+1} q_i^{1-\alpha} / \sum_{i=1}^{2} p_i\Big)/(1 - 2^{\alpha-1}), \quad \alpha \neq 1 . \tag{4.1}$$

Proof. The postulates 1 and 5 are equivalent to (2.3) and hence from (2.25), we have

$$\hat{I}_\alpha\langle p : q\rangle = (p^\beta q^\nu - 1)/(2^{\alpha-1} - 1) . \tag{4.2}$$

Taking $p = 1$, $q = \frac{1}{2}$ in (4.2) and using the postulate 2, we get $\nu = 1 - \alpha$. Thus

$$\hat{I}_\alpha\langle p : q\rangle = (p^\beta q^{1-\alpha} - 1)/(2^{\alpha-1} - 1) . \tag{4.3}$$

Hence (4.3) and the postulate 6 prove theorem 4.

Theorem 5. *The function satisfying the postulates* 1, 2, 3, 5 *and* 6 *is the non-additive information of order* α *given by,*

$$\hat{I}_\alpha^{\alpha-1}\langle P : Q\rangle = \Big(1 - \sum_{i=1}^{2} p_i^\alpha q_i^{1-\alpha} / \sum_{i=1}^{2} p_i\Big)/(1 - 2^{\alpha-1}), \quad \alpha \neq 1 . \tag{4.4}$$

Proof. As done in the proof of theorem 4, postulates 1, 2 and 5 give (4.3). Now using postulate 3 in (4.3) yields $\beta = \alpha - 1$ giving

$$\hat{I}_\alpha\langle p : q\rangle = (p^{\alpha-1} q^{1-\alpha} - 1)/(2^{\alpha-1} - 1) . \tag{4.5}$$

Hence the use of (4.5) in postulate 6 proves theorem 5.

Theorem 6. *The function satisfying the postulates* 1, 2, 4, 5 *and* 6 *is the non-additive inaccuracy of order* α,

$$(4.6) \qquad \hat{I}^0_\alpha\langle P : Q\rangle = \left(1 - \sum_{i=1}^{2} p_i q_i^{1-\alpha} \Big/ \sum_{i=1}^{2} p_i\right) \Big/ \left(1 - 2^{\alpha-1}\right), \quad \alpha \neq 1 \,.$$

Proof. As done earlier in theorem 4; the postulates 1, 2 and 5 imply (4.3). Hence the use of postulate 4 in (4.3) yields $\beta = 0$, giving

$$(4.7) \qquad \hat{I}_\alpha\langle p : q\rangle = \left(q^{1-\alpha} - 1\right) / \left(2^{\alpha-1} - 1\right) .$$

Now using (4.7) in postulate 6 proves theorem 6.

(Received December 15, 1970.)

REFERENCES

[1] J. Aczél: Lectures on functional equations and their applications. Academic Press, New York 1966.

[2] D. F. Kerridge: Inaccuracy and inference. J. Roy. Statist. Soc., *B 23* (1961), 184—194.

[3] S. Kullback: Information theory and statistics. John Wiley and Sons, New York 1959.

[4] P. N. Rathie: Some characterization theorems for generalized measures of uncertainty and information. Proc. of Meeting on Information Measures, University of Waterloo, Waterloo, Ontario, Canada, April 10—14, 1970. Metrika (to appear).

[5] A. Rényi: On measures of entropy and information, Proc. Fourth Berkeley Symposium Math. Statist. and Probability *1* (1961), 547—561.

[6] C. E. Shannon: A mathematical theory of communication. Bell System System Tech. J. *27* (1948), 378—423 and 623—656.

[7] I. Vajda: Axiomy *a*-entropie zobecněného pravděpodobnostního schématu, Kybernetika *4* (1968), 105—112.

O některých nových mírách nejistoty, nepřesnosti a informace a o jejich charakteristikách

P. N. RATHIE

Autor zavádí v první části článku β-entropii (1.1), β-nepřesnost (1.3), β-informaci (1.5) a (α, β)-informaci (1.7). Takto zavedené míry zobecňují některé známé míry nejistoty, nepřesnosti a informace studované již dříve v teorii informace. Zavedené zobecněné míry jsou charakterizovány funkcionálními rovnicemi, jejich řešení je probráno ve druhé části článku. Ve třetí části jsou axiomaticky definovány první tři zavedené míry, v poslední části je axiomaticky definována čtvrtá zavedená míra a některé její speciální varianty.

Pushpa Narayan Rathie, M. Sc., Ph. D., Visiting Scientist, Department of Mathematics, McGill University, Montreal, Canada.