Tôhoku Math. Journ.
27 (1975), 207-212.

ON GENERALIZED INFORMATION FUNCTION*

PL. KANNAPPAN AND P. N. RATHIE

(Received August 10, 1973)

Abstract. In the characterization theorems for generalized information measures, we come across some functional equations which are fundamental in nature. These functional equations define entropy, directed-divergence and inaccuracy functions, all of type β. Measurable solutions of the functional equations concerning these measures are derived in this paper.

1. Introduction. Let $P = (p_1, \cdots, p_n)$, $p_i \geqq 0$, $\sum_{i=1}^{n} p_i = 1$ be a finite discrete probability distribution, then the entropy of type β ($\beta \neq 1$) for the distribution P is defined [Havrda and Charvát (1967), Daróczy (1970)] by the expression

$$H^{\beta}(p_1, \cdots, p_n) = \left(\sum_{i=1}^{n} p_i^{\beta} - 1\right)\Big/(2^{1-\beta} - 1) . \tag{1.1}$$

For $\beta \to 1$, (1.1) reduces to Shannon's entropy [Shannon (1948)]. Also (1.1) has a relation to the entropy of order β defined in [Rényi (1961)].

While characterizing (1.1), Daróczy (1970) came across the following functional equation:

$$f(x) + (1 - x)^{\beta} f[y/(1 - x)] = f(y) + (1 - y)^{\beta} f[x/(1 - y)] , \tag{1.2}$$

for $x, y \in [0, 1[$ with $x + y \in [0, 1]$.

For solving the equation (1.2), Daróczy (1970) also assumed $f(0) = f(1)$ and $f(1/2) = 1$.

Here we will give the measurable solutions for the following more general functional equation

$$f(x) + (1 - x)^{\beta} g[y/(1 - x)] = h(y) + (1 - y)^{\beta} k[x/(1 - y)] , \tag{1.3}$$

for $x, y \in [0, 1[$, $x + y \in [0, 1]$, without the initial conditions.

In (1.3), we assume $g, k\colon [0, 1] \to \mathscr{R}$ and $f, h\colon [0, 1[\to \mathscr{R}$ where $\mathscr{R}$ stands for the real numbers.

With the help of (1.3), we will also obtain the measurable solutions for the following functional equation,

$$\begin{aligned} F(x, y) &+ (1 - x)^{\beta}(1 - y)^{\gamma} F[u/(1 - x), v/(1 - y)] \\ &= F(u, v) + (1 - u)^{\beta}(1 - v)^{\gamma} F[x/(1 - u), y/(1 - v)] , \end{aligned} \tag{1.4}$$

* This work is partially supported by a NRC of Canada Grant.

for $x, y, u, v \in [0, 1[$ with $x+u,\ y+v \in [0, 1]$, the motivation for the consideration of this equation is given below.

The functional equation of the type (1.4) comes very often in the characterization of directed-divergence of type β [Rathie and Kannappan (1972)] and inaccuracy of type β [Rathie and Kannappan (1973)]. For the two finite discrete probability distributions $(p_1, \cdots, p_n)$, $p_i \geqq 0$, $\sum_{i=1}^n p_i = 1$ and $(q_1, \cdots, q_n)$, $q_i \geqq 0$, $\sum_{i=1}^n q_i = 1$, the directed-divergence of type β $(\beta \neq 1)$ is defined by

$$I_n^\beta\begin{pmatrix} p_1, \cdots, p_n \\ q_1, \cdots, q_n \end{pmatrix} = \left(\sum_{i=1}^n p_i^\beta q_i^{1-\beta} - 1\right) \Big/ (2^{\beta-1} - 1)\,, \tag{1.5}$$

and the inaccuracy of type β $(\beta \neq 1)$ by the following expression:

$$H_n^\beta\begin{pmatrix} p_1, \cdots, p_n \\ q_1, \cdots, q_n \end{pmatrix} = \left(\sum_{i=1}^n p_i q_i^{\beta-1} - 1\right) \Big/ (2^{1-\beta} - 1)\,. \tag{1.6}$$

For $\beta \to 1$ (1.5) reduces to the directed-divergence [Kullback (1959)] or the information-gain [Rényi (1961)] and (1.6) yields the inaccuracy [Kerridge (1961)].

For the various applications and characterization theorems of the information measures mentioned above, the reader may see the references cited at the end of this paper.

2. Measurable solutions of (1.3). The measurable solutions of (1.3) will be given in the form of a theorem towards the end of this section. First we will prove a few lemmas which will be required in the proof of the theorem later on.

LEMMA 1. *If f, h: $[0, 1[\to \mathscr{R}$ and g, k: $[0, 1] \to \mathscr{R}$ satisfy the functional equation* (1.3), *then*

$$f(x) = g(1-x) + [k(1) - k(0)]x^\beta - g(1)\,(1-x)^\beta + f(0)\,, \quad \text{for } x \in]0, 1[\,, \tag{2.1}$$

$$h(y) = g(y) - k(0)(1-y)^\beta + f(0)\,, \qquad \text{for } y \in [0, 1[\,, \tag{2.2}$$

$$k(x) = g(1-x) + [k(1) - k(0)]x^\beta + [g(0) - g(1)](1-x)^\beta + f(0) - h(0)\,, \quad \text{for } x \in]0, 1[\,, \tag{2.3}$$

and

$$g(x) = u(x) - [g(0) - g(1)]x^\beta + g(0)\,, \tag{2.4}$$

where u satisfies the functional equation

$$u(1-x) + (1-x)^\beta u[y/(1-x)] = u(y) + (1-y)^\beta u[(1-x-y)/(1-y)] \quad \text{for } x \in]0, 1[,\ y \in [0, 1[\quad \text{with } x+y \in]0, 1]\,. \tag{2.5}$$

PROOF. For $x = 0$, (1.3) yields (2.2). Taking $y = 1 - x$ in (1.3) and using (2.2) we get (2.1). Putting $y = 0$ in (1.3) and utilizing (2.1) we arrive at (2.3).

Substituting the expressions for f, h and k from (2.1), (2.2) and (2.3) respectively in (1.3), we have

$$\begin{aligned}(2.6)\quad & g(1-x) + (1-x)^\beta g[y/(1-x)] \\ &= g(y) + (1-y)^\beta g[(1-x-y)/(1-y)] + [g(0) - g(1)](1-x-y)^\beta \\ &\quad + [f(0) - h(0) - k(0)](1-y)^\beta + g(1)(1-x)^\beta\end{aligned}$$

for $x \in]0, 1[$, $y \in [0, 1[$ with $x + y \in]0, 1]$.

It is easy to see that for $y = 0$, the equation (2.6) gives $g(0) + f(0) - h(0) - k(0) = 0$ and hence the equation (2.6) takes the following form:

$$\begin{aligned}(2.7)\quad & g(1-x) + (1-x)^\beta g[y/(1-x)] \\ &= g(y) + (1-y)^\beta g[(1-x-y)/(1-y)] \\ &\quad + g(1)(1-x)^\beta - g(0)(1-y)^\beta + [g(0) - g(1)](1-x-y)^\beta ,\end{aligned}$$

for $x \in]0, 1[$, $y \in [0, 1[$ with $x + y \in]0, 1[$. Define

$$u(x) = g(x) + [g(0) - g(1)]x^\beta - g(0) .$$

Then we obtain (2.4) and by (2.7) that of (2.5). This completes the proof of Lemma 1.

The proof of the following Lemma 2 is analogous to that of Lemma 3 in Kannappan and Ng (1973).

LEMMA 2. *If u is measurable in $]0, 1[$ and satisfies (2.5) for all x, $y \in]0, 1[$ with $x + y \in]0, 1[$, then u is locally bounded in $]0, 1[$ and hence locally integrable.*

Now we determine the measurable solution of (2.5) for $x, y \in]0, 1[$ with $x + y \in]0, 1[$.

LEMMA 3. *The most general measurable solution of (2.5) for $x, y \in]0, 1[$ with $x + y \in]0, 1[$, is given by*

$$(2.8)\qquad u(x) = AS_\beta(x) , \qquad x \in]0, 1[$$

where

$$(2.9)\qquad S_\beta(x) = x^\beta + (1-x)^\beta - 1 , \qquad \textit{for } \beta(>0) \neq 1 ,$$

where A is an arbitrary constant.

PROOF. Following [Kannappan and Ng (1973)] it is easy to see that u is differentiable infinitely in $]0, 1[$.

Now differentiating (2.5) first with respect to x and then the resulting expression with respect to y and then replacing $y/(1-x)$ by t and $x/(1-y)$ by $1-s$, we get

$$(2.1)\quad (1-t)^{2-\beta}[tu''(t)-(\beta-1)u'(t)] = s^{2-\beta}[(1-s)u''(s)+(\beta-1)u'(s)]\,,$$

for $t, s \in]0, 1[$.

Now from (2.10) we have

$$(2.11)\qquad u(t) = [\lambda/\beta(\beta-1)](1-t)^{\beta} + c_2 t^{\beta} - c_1/\beta\,, \qquad \text{for } t \in]0,1[\,,$$

where λ, c_1 and c_2 are constants.

Substituting the value of u from (2.11) in (2.5) we get

$$\lambda/\beta(\beta-1) = c_1/\beta = c_2 = A \quad \text{(say)}\,.$$

This proves Lemma 3.

Now we prove the following Theorem for the functional equation (1.3).

THEOREM 1. *The most general measurable solutions of* (1.3) *are given by*

$$(2.12)\qquad \left.\begin{aligned} f(x) &= AS_\beta(x) + d_1x^\beta - c_2(1-x)^\beta + c_1 \\ g(y) &= AS_\beta(y) + d_2y^\beta + c_2 \\ h(x) &= AS_\beta(x) + d_2x^\beta - c_4(1-x)^\beta + c_1 \\ k(y) &= AS_\beta(y) + d_1y^\beta + c_4 \end{aligned}\right\}$$

for $x \in [0, 1[$, $y \in [0, 1]$, *where* S_β *is given by* (2.9) *and* A, c_1, c_2, c_4, d *and* d_2 *are arbitrary constants.*

PROOF. From (2.1), (2.2), (2.3), (2.4), (2.8) and (2.9) it is easy to see that the functions f, g, h and k have the following forms:

$$(2.13)\qquad \left.\begin{aligned} f(x) &= AS_\beta(x) + d_1x^\beta + c_1 + b_1(1-x)^\beta \\ g(x) &= AS_\beta(x) + d_2x^\beta + c_2 \\ h(x) &= AS_\beta(x) + d_2x^\beta + c_1 + b_3(1-x)^\beta \\ k(x) &= AS_\beta(x) + d_1x^\beta + c_4 + b_4(1-x)^\beta \end{aligned}\right\}$$

for $x \in]0, 1[$ where $A, d_1, d_2, c_1, c_2, c_4, b_1, b_3$ and b_4 are arbitrary constants.

Direct substitution of (2.13) in (1.3) yields $b_1 = -c_2$, $b_3 = -c_4$ and $b_4 = 0$ giving the expressions in Theorem 1.

An examination at the boundary reveals that f, g, h, k have the form (2.12) on the respective domains. This completes the proof of Theorem 1.

When f, g, h and k are the same, (1.3) reduces to (1.2), the measurable solution of which is given in the following corollary to Theorem 1.

COROLLARY. *The most general measurable solution of* (1.2) *is given by,*

$$(2.14)\qquad f(x) = AS_\beta(x) + Bx^\beta\,, \qquad \text{for } x \in [0, 1]\,,$$

where A and B are arbitrary constants.

REMARK. For $f(0) = f(1)$ and $f(1/2) = 1$, (2.14) reduces to the information function given in [Daróczy (1970)].

3. Measurable solutions of (1.4). Using the results of the previous section, we will establish the following theorem concerning the functional equation (1.4).

THEOREM 2. *The function $F: [0, 1] \times [0, 1] \to \mathscr{R}$, measurable in each variable and satisfying the functional equation* (1.4), *is given by*

$$(3.1)\qquad F(x, y) = bx^\beta y^\gamma - A[x^\beta y^\gamma + (1-x)^\beta(1-y)^\gamma - 1]\,,$$

for $x, y \in [0, 1]$, $\beta(>0) \neq 1$, where A and b are arbitrary constants.

PROOF. For each fixed $y, v \in [0, 1[$ with $y + v \in [0, 1]$, the functional equation (1.4) is of the form (1.3) in the variables x and u. Hence we can apply the results of Theorem 1. Thus, there exist constants $A(y, v)$, $d_1(y, v)$, $d_2(y, v)$, $c_1(y, v)$, $c_2(y, v)$ and $c_4(y, v)$ such that

$$(3.2)\qquad F(x, y) = A(y, v)S_\beta(x) + d_1(y, v)x^\beta + c_1(y, v) - c_2(y, v)(1-x)^\beta$$

$$(3.3)\qquad (1-y)^\nu F[x, v/(1-y)] = A(y, v)S_\beta(x) + d_2(y, v)x^\beta + c_2(y, v)$$

$$(3.4)\qquad F(x, v) = A(y, v)S_\beta(x) + d_2(y, v)x^\beta + c_1(y, v) - c_4(y, v)(1-x)^\beta$$

$$(3.5)\qquad (1-v)^\gamma F[x, y/(1-v)] = A(y, v)S_\beta(x) + d_1(y, v)x^\beta + c_4(y, v)\,.$$

From (3.2) it is easy to see that $A(y, v) + d_1(y, v)$, $A(y, v) - c_2(y, v)$ and $A(y, v) - c_1(y, v)$ are all functions of y alone. Also from (3.4) we find that $A(y, v) - c_1(y, v)$ is a function of v alone. Hence $-A(y, v) + c_1(y, v)$ is a constant, say A. Denoting $A(y, v) + d_1(y, v)$ and $A(y, v) - c_2(y, v)$ by $B(y)$ and $c(y)$ respectively, the equation (3.2) takes the following form

$$(3.6)\qquad F(x, y) = A + B(y)x^\beta + c(y)(1-x)^\beta\,, \qquad \text{for } x, y \in [0, 1]\,.$$

Now, substituting the expression for F from (3.6) in (1.4) we get $c(y) + A(1-y)^\gamma = 0$ and $B(y) - (1-v)^\gamma B[y/(1-v)] = 0$ giving

$$(3.7)\qquad B(y) = \lambda y^\gamma$$

and

$$(3.8)\qquad c(y) = -A(1-y)^\gamma\,,$$

where λ is an arbitrary constant.

Finally, taking $b = \lambda + A$ and substituting for $B(y)$ and $c(y)$ from (3.7) and (3.8) in (3.6) we arrive at (3.1). This completes the proof of Theorem 2.

The following three cases are interesting from the point of view of information theory.

(i) For $\gamma = 1 - \beta$, $F(1, 1) = F(0, 0)$, $F(0, 1/2) = 1$, the function F given in (3.1) reduces to the directed-divergence function of type β.

(ii) For $\beta = 1$, $\gamma = \delta - 1$, $F(1/2, 1/2) = 1$, $F(0, 0) = F(1, 1)$, the function F given in (3.1) reduces to the inaccuracy function of type β.

(iii) For $x = y$, (3.1) reduces to (2.14).

In the end, it is interesting to point out that similar results can be easily derived for functional equations corresponding to (1.4) in three or more variables.

References

[1] Z. DARÓCZY, Generalized information functions, Information and Control, 16 (1970), 36–51.

[2] J. HAVRDA AND F. CHARVÁT, Quantification method of classification processes, the concept of structural α-entropy, Kybernetika, 3 (1967), 30–35.

[3] PL. KANNAPPAN AND C. T. NG, Measurable solutions of functional equations related to information theory, Proc. Amer. Math. Soc. 38 (1973), 303–310.

[4] D. F. KERRIDGE, Inaccuracy and inference, J. Roy, Statist. Soc. B 23 (1961), 184–194.

[5] S. KULLBACK, Information Theory and Statistics, John Wiley and Sons, New York, (1959).

[6] P. N. RATHIE AND PL. KANNAPPAN, A directed-divergence function of type β, Information and Control, 20 (1972), 38–45.

[7] P. N. RATHIE, AND PL. KANNAPPAN, An inaccuracy function of type β, Ann. Inst. Statist. Math., V 25 (1973), 205–214.

[8] A. RÉNYI, On measures of entropy and information, Proc. Fourth Berkeley Symp. Math. Statist. Probability., 1 (1961), 547–561.

[9] C. E. SHANNON, A mathematical theory of communication, Bell System Tech. J., 27 (1948), 378–423 and 623–656.

FACULTY OF MATHEMATICS
UNIVERSITY OF WATERLOO
WATERLOO, ONTARIO
AND
DEPARTMENT OF MATHEMATICS
M. R. ENGG. COLLEGE
JAIPUR, INDIA
AND
IIT, BOMBAY