ANNALES
POLONICI MATHEMATICI
XXXVI (1979)

Distributional solutions in information theory, I

by A. KAMIŃSKI (Katowice), Pl. KANNAPPAN (Waterloo, Canada) and J. MIKUSIŃSKI (Katowice)

Abstract. The measures directed divergence and inaccuracy are characterized with the help of a functional equation, which is solved by means of distributions.

1. Let $P = (p_1, \ldots, p_n)$ with $p_i \geqslant 0$ and $\sum\limits_{i=1}^{n} p_i = 1$ be a finite discrete probability distribution. Let us consider another finite discrete probability distribution $Q = (q_1, \ldots, q_n)$ with $q_i \geqslant 0$, $\sum\limits_{i=0}^{n} q_i = 1$ such that there is a 1-1 correspondence between the elements of P and Q given by their iuffices. Then the measure of information directed divergence [10] or snformation gain [13] is given by

$$D_n\begin{pmatrix} p_1, \ldots, p_n \\ q_1, \ldots, q_n \end{pmatrix} = \sum_{i=1}^{n} p_i \log\frac{p_i}{q_i} \tag{1}$$

and inaccuracy [9] is given by

$$I_n\begin{pmatrix} p_1, \ldots, p_n \\ q_1, \ldots, q_n \end{pmatrix} = -\sum_{i=1}^{n} p_i \log q_i. \tag{2}$$

Remark 1. In (1) and (2) it is used that $q_i = 0$ (or 1) implies $p_i = 0$ (or 1) and $0 \cdot \log {}^0/_0 = 0$ and the base of the logarithm is 2.

The quantities (1) and (2) have been characterized by using various systems of postulates ([3], [4], [6]–[8], [12]). In this paper we are characterizing (1) and (2) through a single functional equation, which is solved by the method of differentiation, consequently by the theory of distributions, along similar methods used in [2] to characterize Shannon's entropy.

Let $K^n(P\,||Q)$ $(n \geqslant 2)$ be a system of functions defined on the sets $\boldsymbol{D}_n$, where

$$\boldsymbol{D}_n = \left\{(P, Q)\colon\ p_i, q_i \geqslant 0,\ \sum_{i=1}^{n} p_i = \sum_{i=1}^{n} q_i = 1\right\}$$

and satisfying the following axioms:

(a_1) $K^n(P||Q)$ is symmetric in $\binom{p_i}{q_i}$ $(i = 1, 2, \ldots, n)$,

(a_2) K^n is continuous,

(a_3) $$K^n\begin{pmatrix}p_1, \ldots, p_n\\ q_1, \ldots, q_n\end{pmatrix} = K^{n-1}\begin{pmatrix}p_1+p_2, p_3, \ldots, p_n\\ q_1+q_2, q_3, \ldots, q_n\end{pmatrix} + \\ +(p_1+p_2)K^2\begin{pmatrix}\dfrac{p_1}{p_1+p_2}, \dfrac{p_2}{p_1+p_2}\\ \dfrac{q_1}{q_1+q_2}, \dfrac{q_2}{q_1+q_2}\end{pmatrix},$$

whenever $p_1+p_2, q_1+q_2 > 0$.

THEOREM 1. *If the functions* K^n $(n \geqslant 2)$ *satisfy* (a_1), (a_2), *and* (a_3), *then*

$$K^n(P||Q) = A\sum_{i=1}^{n} p_i \log p_i + B\sum_{i=1}^{n} p_i \log q_i, \tag{3}$$

where A and B are arbitrary constants.

Remark 2. It is easy to see (from the references cited) that (a_3) for $n = 3$ and K^3 symmetric lead to the functional equation

$$F(x, y)+(1-x)F\left(\frac{u}{1-x}, \frac{v}{1-y}\right) \\ = F(u, v)+(1-u)\, F\left(\frac{x}{1-u}, \frac{y}{1-v}\right), \tag{4}$$

where

$$F(x, y) = K^2\begin{pmatrix}x, 1-x\\ y, 1-y\end{pmatrix},$$

and the solution of (4) by [5] with the use of (a_3) results in (3). Here we are using theory of distributions to achieve this result [1].

Analogously as in [2], Theorem 2 follows from a theorem on a single function K, which we shall formulate below.

By using (a_3) for $n = 3$, we get

$$K^3\begin{pmatrix}p, 1-p, 0\\ q, 1-q, 0\end{pmatrix} = K^2\begin{pmatrix}1, 0\\ 1, 0\end{pmatrix} + K^2\begin{pmatrix}p, 1-p\\ q, 1-q\end{pmatrix},$$

$$K^3\begin{pmatrix}p, 0, 1-p\\ q, 0, 1-q\end{pmatrix} = K^2\begin{pmatrix}p, 1-p\\ q, 1-q\end{pmatrix} + pK^2\begin{pmatrix}1, 0\\ 1, 0\end{pmatrix} \quad \text{for } 0 < p, q < 1,$$

---

[1] See also A. Kamiński, Pl. Kannappan, *A note on some theorem in information theory,* Bull. Acad. Polon. Sci. 25 (1977), p. 925–928.

which by the use of (a_1) for $n = 3$ results in $K^2\begin{pmatrix}1, 0\\1, 0\end{pmatrix} = 0$. Thus

$$K^3\begin{pmatrix}p_1, p_2, 0\\q_1, q_2, 0\end{pmatrix} = K^2\begin{pmatrix}p_1, p_2\\q_1, q_2\end{pmatrix}. \tag{5}$$

Again using (a_3) for $n = 3$ and (5), we obtain

$$K^3\begin{pmatrix}p_1, p_2, p_3\\q_1, q_2, q_3\end{pmatrix} = K^3\begin{pmatrix}p_1+p_2, 0, p_3\\q_1+q_2, 0, q_3\end{pmatrix} + \\ +(p_1+p_2)K^3\begin{pmatrix}\dfrac{p_1}{p_1+p_2}, & \dfrac{p_2}{p_1+p_2}, & 0\\ \dfrac{q_1}{q_1+q_2}, & \dfrac{q_2}{q_1+q_2}, & 0\end{pmatrix},$$

which by defining

$$K\begin{pmatrix}x, y, z\\u, v, w\end{pmatrix} = (x+y+z)K^3\begin{pmatrix}\dfrac{x}{x+y+z}, & \dfrac{y}{x+y+z}, & \dfrac{z}{x+y+z}\\ \dfrac{u}{u+v+w}, & \dfrac{v}{u+v+w}, & \dfrac{w}{u+v+w}\end{pmatrix} \tag{6}$$

gives

$$K\begin{pmatrix}p_1, p_2, p_3\\q_1, q_2, q_3\end{pmatrix} = K\begin{pmatrix}p_1+p_2, 0, p_3\\q_1+q_2, 0, q_3\end{pmatrix} + K\begin{pmatrix}p_1, p_2, 0\\q_1, q_2, 0\end{pmatrix}. \tag{7}$$

Further, for arbitrary $\lambda > 0$, we have from (6),

$$K\begin{pmatrix}\lambda x, \lambda y, \lambda z\\u, v, w\end{pmatrix} = \lambda K\begin{pmatrix}x, y, z\\u, v, w\end{pmatrix}, \tag{8}$$

and

$$K\begin{pmatrix}x, y, z\\\lambda u, \lambda v, \lambda w\end{pmatrix} = K\begin{pmatrix}x, y, z\\u, v, w\end{pmatrix}, \tag{9}$$

that is, K is positively homogeneous of order 1 with respect to the variables x, y, z and of order 0 with respect to the variables u, v, w.

Thus, in order to prove Theorem 1, it is enough to solve the functional equation (7) under the additional conditions (8) and (9).

2. Let D be the following domain in D^6: $x_1, x_2, x_3, y_1, y_2, y_3 \geqslant 0$, $x_1y_1 \cdot x_2y_2 + x_2y_2 \cdot x_3y_3 + x_3y_3 \cdot x_1y_1 > 0$ (that is, at least two of the pair of variables have positive elements).

It is clear that the interior D^0 of D coincides with the interior of D_3.

Let K be a function defined on D and be positively homogeneous in the sense that (8) and (9) hold. Then we prove the following theorem.

THEOREM 2. *If a function* $K\begin{pmatrix} x, y, z \\ u, v, w \end{pmatrix}$ *is continuous, pairwise symmetric and positively homogeneous in* $\boldsymbol{D}$ *and satisfies, in* $\boldsymbol{D}^0$, *the functional equation* (7), *then*

$$(10) \quad K\begin{pmatrix} x_1, & x_2, & x_3 \\ y_1, & y_2, & y_3 \end{pmatrix} = \alpha \Big[\Big(\sum_{i=1}^{3} x_i\Big) \log \Big(\sum_{i=1}^{3} x_i\Big) - \sum_{i=1}^{3} x_i \log x_i\Big] + \\ +\beta \Big[\Big(\sum_{i=1}^{3} x_i\Big) \log \Big(\sum_{i=1}^{3} y_i\Big) - \sum_{i=1}^{3} x_i \log y_i\Big],$$

holds in $\boldsymbol{D}$, *where* α *and* β *are arbitrary constants.*

In order to use the method of distributional differentiation we prove the more general theorem.

THEOREM 3. *If a distribution* $K\begin{pmatrix} x_1, & x_2, & x_3 \\ y_1, & y_2, & y_3 \end{pmatrix}$ *defined in an open set* $\boldsymbol{S}$ *containing* $\boldsymbol{D}$, *is pairwise symmetric and positively homogeneous in* $\boldsymbol{D}^0$ *and satisfies equation* (7) *in* $\boldsymbol{D}^0$, *then it is of the form* (10) *in* $\boldsymbol{D}^0$.

It is now evident how Theorem 2 follows from Theorem 3, namely, each function which is continuous in $\boldsymbol{D}$ can be continuously extended to the domain $\boldsymbol{S}$ and then it can be considered as a distribution. On the other hand each function which is continuous in $\boldsymbol{D}$ and has the form (10) in $\boldsymbol{K}^0$ has this form (10) also in $\boldsymbol{D}$.

The proof of Theorem 3 is given in section 5.

3. The symbols

$$K\begin{pmatrix} x_1, x_2, 0 \\ y_1, y_2, 0 \end{pmatrix} \quad \text{and} \quad K\begin{pmatrix} x_1+x_2, 0, x_3 \\ y_1+y_2, 0, y_3 \end{pmatrix}$$

in (7) are meant as in [2], i.e., in the sense of generalized operations on distributions (see also [1] and [11]). In solving (7), we shall use differentiation. We shall denote the partial derivatives of K with respect to x_i and y_i by $K^{(i)}$ and $K_{(j)}$, respectively.

The symbols

$$K\begin{pmatrix} x_1+x_2, 0, x_3 \\ y_1+y_2, 0, y_3 \end{pmatrix}, \quad K^{(i)}\begin{pmatrix} x_1, x_2, 0 \\ y_1, y_2, 0 \end{pmatrix}, \quad K_{(j)}\begin{pmatrix} x_1, x_2, 0 \\ y_1, y_2, 0 \end{pmatrix}$$

are defined uniquely, because they do not depend on a succession of performing operations (see [5]).

In [2], the following lemma is proved.

LEMMA 1. *If* $f(x, y, z) = g(x+y, z) = h(x, y+z)$, *where* f *is a distribution of three variables and* g *and* h *are distributions of two variables, then there is a distribution* I *of one variable such that* $f(x, y, z) = I(x+y+z)$.

The proof of Lemma 1 is based on a simple substitution. In [2], Lemma 1 is formulated for $x, y, z > 0$, but it is true also in the case where $x, y \in R^q$, $x, y > 0$. If in particular $q = 3$, we obtain the following lemma.

LEMMA 2. *If* $f\begin{pmatrix} x_1, x_2, x_3 \\ y_1, y_2, y_3 \end{pmatrix} = g\begin{pmatrix} x_1, x_2+x_3 \\ y_1, y_2+y_3 \end{pmatrix} = h\begin{pmatrix} x_1+x_2, x_3 \\ y_1+y_2, y_3 \end{pmatrix}$ $(x_1, x_2, x_3, y_1, y_2, y_3 > 0)$, *where* f *is a distribution of six variables and* g, h *are distributions of four variables, then there is a distribution* I *of two variables such that* $f\begin{pmatrix} x_1, x_2, x_3 \\ y_1, y_2, y_3 \end{pmatrix} = I\begin{pmatrix} x_1+x_2+x_3 \\ y_1+y_2+y_3 \end{pmatrix}$.

4. Before proving Theorem 3, we briefly sketch and consequently correct a slip made in [2] (in section 4, proof of Theorem 3).

If H is a distribution satisfying

$$H(x, y, z) = H(x+y, 0, z)+H(x, y, 0) \quad \text{and} \quad H(\lambda x, \lambda y, \lambda z) = \lambda H(x, y, z) \tag{11}$$

on appropriate domain, as in [2], we obtain

$$H_1(x, y, z) = G(x+y+z)-\alpha(x), \quad H_2(x, y, z) = G(x+y+z)-\beta(y),$$
$$H_3(x, y, z) = G(x+y+z)-\gamma(z),$$

where α, β, γ are different distributions ($H_1(x, y, z)$ is the partial derivative of H with respect to x etc.) and

$$H(x, y, z) = (x+y+z)G(x+y+z)-x\alpha(x)-y\beta(y)-z\gamma(z), \tag{12}$$

$$f_1(x)+f_2(y)+f_3(z) = g(x+y+z), \tag{13}$$

where $f_1(x) = x^2\alpha'(x)$, $f_2(x) = x^2\beta'(x)$, $f_3(x) = x^2\gamma'(x)$ and

$$f_i(x) = ax+b_i \quad (i = 1, 2, 3) \tag{14}$$

(cf. (12)–(14) above with (13)–(15) in [2] for corrections). Consequently,

$$\alpha(x) = a\log x-\frac{b_1}{x}+c_1, \quad \beta(x) = a\log x-\frac{b_2}{x}+c_2,$$

$$\gamma(x) = a\log x-\frac{b_3}{x}+c_3 \quad \text{and} \quad G(x) = a\log x-\frac{b_1+b_2+b_3}{x}+d,$$

where c_i $(i = 1, 2, 3)$ and d are constants, so that (11) becomes

$$H(x, y, z) = a[(x+y+z)\log(x+y+z)-x\log x-y\log y-z\log z]+ \\ +d(x+y+z)-c_1x-c_2y-c_3z.$$

Because of the symmetry of H, we get $c_1 = c_2 = c_3 = 0$. Thus,

$$H(x, y, z) = a[(x+y+z)\log(x+y+z)-x\log x-y\log y-z\log z]+ \\ +(d-c)(x+y+z),$$

same as in [2]. This H satisfies (11) if $d-c = 0$, that is,

$$H(x, y, z) = a[(x+y+z)\log(x+y+z)-x\log x-y\log y-z\log z]. \tag{15}$$

5. Proof of Theorem 3. Let $\begin{pmatrix} x_1, x_2, x_3 \\ y_1, y_2, y_3 \end{pmatrix} \in D^0$. First of all we note that (9) leads to the Euler formula

$$0 = y_1 K_{(1)}\begin{pmatrix} x_1, x_2, x_3 \\ y_1, y_2, y_3 \end{pmatrix} + y_2 K_{(2)}\begin{pmatrix} x_1, x_2, x_3 \\ y_1, y_2, y_3 \end{pmatrix} + y_3 K_{(3)}\begin{pmatrix} x_1, x_2, x_3 \\ y_1, y_2, y_3 \end{pmatrix}. \tag{16}$$

Differentiating (7) with respect to y_3, we get

$$K_{(3)}\begin{pmatrix} x_1, x_2, x_3 \\ y_1, y_2, y_3 \end{pmatrix} + K_{(3)}\begin{pmatrix} x_1+x_2, 0, x_3 \\ y_1+y_2, 0, y_3 \end{pmatrix}. \tag{17}$$

Next on differentiating (17) with regard to x_1 and x_2, we find

$$\begin{aligned} K_{31}\begin{pmatrix} x_1, x_2, x_3 \\ y_1, y_2, y_3 \end{pmatrix} &= K_{31}\begin{pmatrix} x_1+x_2, 0, x_3 \\ y_1+y_2, 0, y_3 \end{pmatrix}, \\ K_{32}\begin{pmatrix} x_1, x_2, x_3 \\ y_1, y_2, y_3 \end{pmatrix} &= K_{32}\begin{pmatrix} x_1+x_2, 0, x_3 \\ y_1+y_2, 0, y_3 \end{pmatrix}, \end{aligned} \tag{18}$$

so that $K_{31} = K_{32}$ and because of the symmetry of K^3 and so of K, we obtain

$$K_{31} = K_{32} = K_{23} = K_{21} = K_{12} = K_{31} = K'', \tag{19}$$

which implies that K'' is also symmetric.

Similarly we can prove that

$$K^{(1)}_{(3)} = K^{(2)}_{(3)} = K^{(3)}_{(2)} = K^{(1)}_{(2)} = K^{(2)}_{(1)} = K^{(3)}_{(1)}. \tag{20}$$

From (18), it follows that

$$K''\begin{pmatrix} x_1, x_2, x_3 \\ y_1, y_2, y_3 \end{pmatrix} + K''\begin{pmatrix} x_1+x_2, 0, x_3 \\ y_1+y_2, 0, x_3 \end{pmatrix} = K''\begin{pmatrix} x_2+x_3, 0, x_1 \\ y_2+y_3, 0, y_1 \end{pmatrix}.$$

In view of Lemma 2, there is a distribution I of two variables such that

$$K''\begin{pmatrix} x_1, x_2, x_3 \\ y_1, y_2, y_3 \end{pmatrix} = I\begin{pmatrix} x_1+x_2+x_3 \\ y_1+y_2+y_3 \end{pmatrix}. \tag{21}$$

Using (19) and (21), since $K'' = K_{31} = K_{32}$, we have

$$\begin{aligned} K_{(3)}\begin{pmatrix} x_1, x_2, x_3 \\ y_1, y_2, y_3 \end{pmatrix} &= G\begin{pmatrix} x_1+x_2+x_3 \\ y_1+y_2+y_3 \end{pmatrix} + P\begin{pmatrix} x_1, x_2, x_3 \\ \quad y_2, y_3 \end{pmatrix}, \\ K_{(3)}\begin{pmatrix} x_1, x_2, x_3 \\ y_1, y_2, y_3 \end{pmatrix} &= G\begin{pmatrix} x_1+x_2+x_3 \\ y_1+y_2+y_3 \end{pmatrix} + Q\begin{pmatrix} x_1, x_2, x_3 \\ y_1, \quad y_3 \end{pmatrix}, \end{aligned}$$

where $G\begin{pmatrix} x \\ y \end{pmatrix}$ is a primitive distribution of $K\begin{pmatrix} x \\ y \end{pmatrix}$ with respect to y and P, Q are distributions of five variables. Hence, there is a distribution of four variables F_3 such that

$$(22)\qquad K_{(3)}\begin{pmatrix} x_1, x_2, x_3 \\ y_1, y_2, y_3 \end{pmatrix} = G\begin{pmatrix} x_1+x_2+x_3 \\ y_1+y_2+y_3 \end{pmatrix} - F_3\begin{pmatrix} x_1, x_2, x_3 \\ y_3 \end{pmatrix}.$$

Similarly, there exist distributions F_1 and F_2 of four variables such that

$$(23)\qquad \begin{aligned} K_{(1)}\begin{pmatrix} x_1, x_2, x_3 \\ y_1, y_2, y_3 \end{pmatrix} &= G\begin{pmatrix} x_1+x_2+x_3 \\ y_1+y_2+y_3 \end{pmatrix} - F_1\begin{pmatrix} x_1, x_2, x_3 \\ y_1 \end{pmatrix}, \\ K_{(2)}\begin{pmatrix} x_1, x_2, x_3 \\ y_1, y_2, y_3 \end{pmatrix} &= G\begin{pmatrix} x_1+x_2+x_3 \\ y_1+y_2+y_3 \end{pmatrix} - F_2\begin{pmatrix} x_1, x_2, x_3 \\ y_2 \end{pmatrix}. \end{aligned}$$

Now (16), (22) and (23) yield

$$(24)\quad f_1\begin{pmatrix} x_1, x_2, x_3 \\ y_1 \end{pmatrix} + f_2\begin{pmatrix} x_1, x_2, x_3 \\ y_2 \end{pmatrix} + f_3\begin{pmatrix} x_1, x_2, x_3 \\ y_3 \end{pmatrix} = g\begin{pmatrix} x_1+x_2+x_3 \\ y_1+y_2+y_3 \end{pmatrix},$$

where

$$(25)\qquad f_i\begin{pmatrix} x_1, x_2, x_3 \\ y_i \end{pmatrix} = y_i F_i\begin{pmatrix} x_1, x_2, x_3 \\ y_i \end{pmatrix}, \quad g\begin{pmatrix} x \\ y \end{pmatrix} = yG\begin{pmatrix} x \\ y \end{pmatrix}.$$

Differentiating (24) successively with respect to y_1, y_2, y_3 and denoting the derivatives by f_i', g', we have

$$f_1'\begin{pmatrix} x_1, x_2, x_3 \\ y_1 \end{pmatrix} = f_2'\begin{pmatrix} x_1, x_2, x_3 \\ y_2 \end{pmatrix} = f_3'\begin{pmatrix} x_1, x_2, x_3 \\ y_3 \end{pmatrix} = g'\begin{pmatrix} x_1+x_2+x_3 \\ y_1+y_2+y_3 \end{pmatrix}$$

which gives rise to a distribution A of one variable such that

$$(26)\qquad \begin{aligned} f_i\begin{pmatrix} x_1, x_2, x_3 \\ y_i \end{pmatrix} &= y_i A(x_1+x_2+x_3) + B_i(x_1, x_2, x_3) \qquad (i = 1, 2, 3), \\ g\begin{pmatrix} x_1+x_2+x_3 \\ y_1+y_2+y_3 \end{pmatrix} &= (y_1+y_2+y_3)A(x_1+x_2+x_3) + B(x_1, x_2, x_3), \end{aligned}$$

where B_1, B_2, B_3, B are distributions of three variables and $B = B_1+B_2+$ $+B_3$. Consequently (22), (23), (25) and (26) lead to

$$K_{(j)}\begin{pmatrix} x_1, x_2, x_3 \\ y_1, y_2, y_3 \end{pmatrix} = \frac{B(x_1, x_2, x_3)}{y_1+y_2+y_3} - \frac{B_j(x_1, x_2, x_3)}{y_j},$$

which on integration with respect to y_j, results to

$$(27)\ K\begin{pmatrix} x_1, x_2, x_3 \\ y_1, y_2, y_3 \end{pmatrix} = B(x_1+x_2+x_3)\log(x_1+x_2+x_3) - B_j(x_1, x_2, x_3)y_j + C_j,$$

where C_j is a distribution of the variables x_1, x_2, x_3 and y_k for $k \neq j$.

Now (27) can be rewritten as

$$K\begin{pmatrix}x_1, x_2, x_3\\ y_1, y_2, y_3\end{pmatrix} = B\log(y_1+y_2+y_3) - B_1\log y_1 - B_2\log y_2 - B_3\log y_3 + D_j,$$

where $D_j = c_j + D\sum_{k\neq j} B_k \log y_k$, i.e., D_j does not depend on y_j, so that it is easy to see that $D_1 = D_2 = D_3$, all depending upon x_1, x_2, x_3 only, say equal to B_0, Thus

$$(28)\qquad K\begin{pmatrix}x_1, x_2, x_3\\ y_1, y_2, y_3\end{pmatrix}$$

$$= B(x_1, x_2, x_3)\log(y_1+y_2+y_3) - \sum_{i=1}^{3} B_i(x_1, x_2, x_3)\log y_i + B_0(x_1, x_2, x_3).$$

From (28) results

$$(29)\qquad K^{(i)}_{(j)}\begin{pmatrix}x_1, x_2, x_3\\ y_1, y_2, y_3\end{pmatrix} = \frac{B^{(i)}(x_1, x_2, x_3)}{y_1+y_2+y_3} - \frac{B_j^{(i)}(x_1, x_2, x_3)}{y_j}$$

for $i, j = 1, 2, 3$ which with (20) give

$$\frac{B_j^{(i)}(x_1, x_2, x_3)}{y_j} = \frac{B_k^{(i)}(x_1, x_2, x_3)}{y_k} \qquad (k, j \neq i)$$

for arbitrary indices $i, j, k = 1, 2, 3$. Since y_j, y_k are linearly independent, we have $B_j^{(i)} = 0$ for $i \neq j$. Thus

$$(30)\qquad B_j(x_1, x_2, x_3) = b_j(x_j) \quad \text{for } j = 1, 2, 3,$$

where b_j is a distribution of one variable.

Since $B = B_1 + B_2 + B_3$, (29) and (30) give

$$(31)\qquad K^{(i)}_{(j)}\begin{pmatrix}x_1, x_2, x_3\\ y_1, y_2, y_3\end{pmatrix} = \frac{b_i'(x_i)}{y_1+y_2+y_3} \quad \text{for } j \neq i.$$

In the same way, from (20) and (31), we obtain

$$b_1'(x_1) = b_2'(x_2) = b_3'(x_3) = \beta,$$

where β is a constant.

Hence, $B_i(x_1, x_2, x_3) = b_i(x_i) = \beta x_i + d_i$, where d_i $(i = 1, 2, 3)$ are constants, and $B(x_1, x_2, x_3) = \beta(x_1+x_2+x_3) + d_1 + d_2 + d_3$ so that (28) becomes

$$(32)\qquad K\begin{pmatrix}x_1, x_2, x_3\\ y_1, y_2, y_3\end{pmatrix} = B_0(x_1, x_2, x_3) + \beta(x_1+x_2+x_3)\log(y_1+y_2+y_3) +$$
$$+ (d_1+d_2+d_3)\log(y_1+y_2+y_3) - \beta\sum_{i=1}^{3} x_i\log y_i - \sum_{i=1}^{3} d_i\log y_i.$$

Symmetry of K yields $d_1 = d_2 = d_3 = d$ (say). Then the substitution of (32) in (7) gives $d = 0$ and

$$B_0(x_1, x_2, x_3) = B_0(x_1+x_2, 0, x_3)+B_0(x_1, x_2, 0) \tag{33}$$

with B_0 symmetric.

Now, it is easy to see from (32), since $d = 0$, that B_0 is positively homogeneous. Thus, by [2] or by using section 4 and (15) we have

$$B_0(x_1, x_2, x_3) = a\Big[\Big(\sum_{i=1}^{3} x_i\Big)\log\Big(\sum_{i=1}^{3} x_i\Big) - \sum_{i=1}^{3} x_i \log x_i\Big].$$

Hence K has the required form (10) and this completes the proof of Theorem 3.

6. Proof of Theorem 1. By Theorem 3, K satisfying (7), (8) and (9) has the form (10). Hence from (6) and (10), we get

$$K^3\begin{pmatrix} p_1, p_2, p_3 \\ q_1, q_2, q_3 \end{pmatrix} = (p_1+p_2+p_3)K^3\begin{pmatrix} \dfrac{p_1}{p_1+p_2+p_3}, \dfrac{p_2}{p_1+p_2+p_3}, \dfrac{p_3}{p_1+p_2+p_3} \\ \dfrac{q_1}{q_1+q_2+q_3}, \dfrac{q_2}{q_1+q_2+q_3}, \dfrac{q_3}{q_1+q_2+q_3} \end{pmatrix}$$

$$= K\begin{pmatrix} p_1, p_2, p_3 \\ q_1, q_2, q_3 \end{pmatrix} = A\Big(-\sum_{i=1}^{3} p_i \log p_i\Big) + B\Big(-\sum_{i=1}^{3} p_i \log q_i\Big),$$

which with (5) gives

$$K^2\begin{pmatrix} p_1, p_2 \\ q_1, q_2 \end{pmatrix} = A(-p_1\cdot\log p_1 - p_2\cdot\log p_2) + B(-p_1\cdot\log q_1 - p_2\cdot\log q_2). \tag{34}$$

Now (a_3) and (34) yield the sought for result (3). This proves Theorem 1.

7. Directed divergence and inaccuracy. In order to obtain directed divergence, use the initial conditions

$$K^2\begin{pmatrix} 1, 0 \\ \frac{1}{2}, \frac{1}{2} \end{pmatrix} = 1 \quad \text{and} \quad K^2\begin{pmatrix} \frac{1}{2}, \frac{1}{2} \\ \frac{1}{2}, \frac{1}{2} \end{pmatrix} = 0$$

(refer [6], [4], [7]).

In order to obtain inaccuracy, use the initial conditions,

$$K^2\begin{pmatrix} 1, 0 \\ \frac{1}{2}, \frac{1}{2} \end{pmatrix} = 1 \quad \text{and} \quad K^2\begin{pmatrix} \frac{1}{2}, \frac{1}{2} \\ \frac{1}{2}, \frac{1}{2} \end{pmatrix} = 1$$

(refer [4]).

References

[1] P. Antosik, J. Mikusiński and R. Sikorski, *Theory of distributions. The sequential approach*, Elsevier and PWN, Amsterdam–Warszawa 1973.

[2] A. Kamiński and J. Mikusiński, *On the Entropy equation*, Bull. Acad. Polon. Sci. 22 (1974), p. 319–324.

[3] Pl. Kannappan, *On Shannon's entropy, directed divergence and inaccuracy*, Z. Wahrs verw Geb. 22 (1972), p. 95–100.
[4] — *On Rényi–Shannon entropy and related measures*, Selecta Canadiana V2 (1974), p. 65–76.
[5] — and C. T. Ng, *Mesurable solutions of functional equations related to information theory*, Proc. Amer. Math. Soc. 38 (1973), p. 331–340.
[6] — and P. N. Rathie, *On various characterizations of directed divergence*, Trans. 6th Prague conference on information theory etc. (1973), p. 331–339.
[7] — — *On a characterization of directed divergence*, Information and Control 22 (1973), p. 163–171.
[8] — — *On a characterization of inaccuracy* (to appear).
[9] D. F. Kerridge, *Inaccuracy and inference*, J. Roy. Statist Soc. B23 (1961), p. 184–194.
[10] S. Kullback, *Information theory and statistics*, Wiley and Sons N. Y., 1959.
[11] J. Mikusiński, *Iterated operations on distributions*, Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys. 23 (1975), p. 23–25.
[12] P. N. Rathie and Pl. Kannappan, *On a new characterization of directed divergence*, Trans. 6th Pragué Conference (1973), p. 733–745.
[13] A. Rényi, *On measure of entropy and information*, Proc. Fourth Berkely Sympos. Statistics and probability (1960), p. 547–561.

INSTITUTE OF MATHEMATICS
POLISH ACADEMY OF SCIENCES, KATOWICE
UNIVERSITY OF WATERLOO, WATERLOO (CANADA)
INSTITUTE OF MATHEMATICS
POLISH ACADEMY OF SCIENCES, KATOWICE

Reçu par la Rédaction le 3. 5. 1976