DEMONSTRATIO MATHEMATICA
Vol. XLII No 1 2009

P. Nath, D. K. Singh

ON A SUM FORM FUNCTIONAL EQUATION RELATED TO ENTROPIES AND SOME MOMENTS OF A DISCRETE RANDOM VARIABLE

Abstract. The general solutions of a sum form functional equation have been obtained. The importance of its solutions in relation to the entropies and some moments of a discrete random variable has been discussed.

1. Introduction

For $n = 1, 2, \ldots$; let $\Gamma_n = \left\{(p_1, \ldots, p_n) : p_i \geq 0, i = 1, \ldots, n; \sum_{i=1}^{n} p_i = 1\right\}$ denote the set of all n-component complete discrete probability distributions with nonnegative elements. For any probability distribution $(p_1, \ldots, p_n) \in \Gamma_n$, the entropies

$$H_n(p_1, \ldots, p_n) = -\sum_{i=1}^{n} p_i \log_2 p_i \tag{1.1}$$

are known as the Shannon entropies [7] with $H_n : \Gamma_n \to \mathbb{R}$, $n = 1, 2, \ldots$; $\mathbb{R}$ denoting the set of all real numbers and $0 \log_2 0 := 0$.

Given any probability distribution $(p_1, \ldots, p_n) \in \Gamma_n$, let us consider a discrete real-valued random variable Z_n taking the values $z_1, \ldots, z_n$ with respective probabilities $p_1, \ldots, p_n$ where

$$z_i = \begin{cases} -\log_2 p_i & \text{if } 0 < p_i \leq 1 \\ 0 & \text{if } p_i = 0. \end{cases} \tag{1.2}$$

Let

$$\mu_r'(Z_n) = \sum_{i=1}^{n} p_i z_i^r \tag{1.3}$$

denote the rth order moment of Z_n about the origin, $r = 0, 1, 2, \ldots$.

2000 *Mathematics Subject Classification*: 39B22, 39B52.
Key words and phrases: The Shannon entropy, the entropies of degree α, moments of a discrete random variable, additive mapping, logarithmic mapping, multiplicative mapping.

Consider (1.3) when $r = 0$. In order that $\mu_0'(Z_n)$ is unambiguously defined, we must define $0^0 := 1$. Then $\mu_0'(Z_n) = 1$. Since $\mu_0'(Z_n)$ does not depend upon the probabilities $p_1, \ldots, p_n$; it is not of much importance from information-theoretic point of view.

Now consider (1.3) when $r = 1$. In this case, (1.1), (1.2), (1.3), together with $0 \log_2 0 := 0$, give

$$\mu_1'(Z_n) = H_n(p_1, \ldots, p_n). \tag{1.4}$$

Now consider (1.3) when $r = 2, 3, \ldots$. In this case, as a generalization of $0 \log_2 0 := 0$, we define

$$0(\log_2 0)^r := 0. \tag{1.5}$$

Then

$$\mu_r'(Z_n) = \sum_{i=1}^{n} p_i(-\log_2 p_i)^r. \tag{1.6}$$

Let $I = \{x \in \mathbb{R} : 0 \leq x \leq 1\} = [0, 1]$. Define the mappings $\varphi_r : I \to \mathbb{R}$, $r = 1, 2, \ldots$ as

$$\varphi_r(p) = p(-\log_2 p)^r \tag{1.7}$$

for all $p \in I$. Then, keeping in view $0 \log_2 0 := 0$ and (1.5) (for $r = 2, 3, \ldots$), it is obvious that

$$\mu_r'(Z_n) = \sum_{i=1}^{n} \varphi_r(p_i) \tag{1.8}$$

for $r = 1, 2, \ldots$. So, all the moments $\mu_r'(Z_n)$, $r = 1, 2, \ldots$ admit of a sum representation. The mapping $\varphi_r : I \to \mathbb{R}$ is called the generating function of the moment $\mu_r'(Z_n)$ whenever $r = 1, 2, \ldots$. Since, for all $r = 1, 2, \ldots$,

$$\varphi_r(0) = 0, \quad \varphi_r(1) = 0 \tag{1.9}$$

it follows that the values of the random variable Z_n taking with probabilities 0 and 1 do not contribute anything to the value of $\mu_r'(Z_n)$. Also, the right hand side of (1.8) is a symmetric function of $p_1, \ldots, p_n$.

Making use of (1.5) and (1.9) for $r = 1, 2$ it is easy to verify that the mappings φ_1 and φ_2 satisfy the functional equation

$$\varphi_2(pq) = q\varphi_2(p) + p\varphi_2(q) + 2\varphi_1(p)\varphi_1(q)$$

for all $p \in I$, $q \in I$. Consequently, for all $(p_1, \ldots, p_n) \in \Gamma_n$, $(q_1, \ldots, q_m) \in \Gamma_m$, the functional equation

$$\sum_{i=1}^{n} \sum_{j=1}^{m} \varphi_2(p_i q_j) = \sum_{i=1}^{n} \varphi_2(p_i) + \sum_{j=1}^{m} \varphi_2(q_j) + 2 \sum_{i=1}^{n} \varphi_1(p_i) \sum_{j=1}^{m} \varphi_1(q_j) \tag{1.10}$$

holds. So, it seems desirable to pay attention to the functional equation

$$\sum_{i=1}^{n}\sum_{j=1}^{m} f(p_i q_j) = \sum_{i=1}^{n} f(p_i) + \sum_{j=1}^{m} f(q_j) + c\sum_{i=1}^{n} g(p_i)\sum_{j=1}^{m} g(q_j) \tag{1.11}$$

in which $f : I \to \mathbb{R}$, $g : I \to \mathbb{R}$, $(p_1, \dots, p_n) \in \Gamma_n$, $(q_1, \dots, q_m) \in \Gamma_m$ and $c \neq 0$ is a given constant. Clearly $f = \varphi_2$ and $g = \varphi_1$ satisfy the functional equation (1.11) when $c = 2$. Keeping in view (1.9), we shall assume that

$$f(0) = f(1) = 0 \tag{1.12}$$

and

$$g(0) = g(1) = 0. \tag{1.13}$$

If $\sum\limits_{i=1}^{n} g(p_i) = 0$ or $\sum\limits_{j=1}^{m} g(q_j) = 0$, then (1.11) reduces to the functional equation

$$\sum_{i=1}^{n}\sum_{j=1}^{m} f(p_i q_j) = \sum_{i=1}^{n} f(p_i) + \sum_{j=1}^{m} f(q_j). \tag{1.14}$$

The functional equation (1.14) was first considered by T.W. Chaundy and J.B. McLeod [1] who came across it while studying some problems in statistical thermodynamics.

The object of this paper is to determine all possible solutions (f, g) of equation (1.11) satisfying initial conditions (1.12) and (1.13), assuming that $f : I \to \mathbb{R}$, $g : I \to \mathbb{R}$, $(p_1, \dots, p_n) \in \Gamma_n$, $(q_1, \dots, q_m) \in \Gamma_m$, $n \geq 3$ and $m \geq 3$ being fixed integers. This has been done in section 3. In section 4, we have discussed the importance of various obtained solutions of (1.11) in relation to statistics and various entropies in information theory.

2. Some Preliminary Results

In this section, we mention some definitions and known results.

A mapping $a : I \to \mathbb{R}$ is said to be additive on I or on the unit triangle $\Delta = \{(x, y) : 0 \leq x \leq 1, 0 \leq y \leq 1, 0 \leq x + y \leq 1\}$ if it satisfies the equation $a(x + y) = a(x) + a(y)$ for all $(x, y) \in \Delta$. A mapping $A : \mathbb{R} \to \mathbb{R}$ is said to be additive on $\mathbb{R}$ if $A(x + y) = A(x) + A(y)$ holds for all $x \in \mathbb{R}$, $y \in \mathbb{R}$. It is known (see Z. Daróczy and L. Losonczi [2]) that if a mapping $a : I \to \mathbb{R}$ is additive on I, then it has a unique additive extension $A : \mathbb{R} \to \mathbb{R}$ in the sense that A is additive on $\mathbb{R}$ and $A(x) = a(x)$ for all $x \in I$.

Result 1. (See [5]) *Let $\psi : I \to \mathbb{R}$ be a mapping which satisfies the equation $\sum\limits_{i=1}^{k} \psi(x_i) = c$ for all $(x_1, \dots, x_k) \in \Gamma_k$; c a given constant and $k \geq 3$ a* *fixed integer. Then there exists an additive mapping* $b : \mathbb{R} \to \mathbb{R}$ *such that* $\psi(x) = b(x) - \frac{1}{k}b(1) + \frac{c}{k}$ *for all* $x \in I$.

Definition 1. A mapping $M : I \to \mathbb{R}$ is said to be multiplicative on I if $M(0) = 0$, $M(1) = 1$ and $M(pq) = M(p)M(q)$ for all $p \in \,]0,1[$, $q \in \,]0,1[$ where $]0,1[\, = \{x \in \mathbb{R} : 0 < x < 1\}$.

Definition 2. A mapping $\ell : I \to \mathbb{R}$ is said to be logarithmic on I if $\ell(0) = 0$ and $\ell(pq) = \ell(p) + \ell(q)$ for all $p \in]0,1], q \in]0,1]$ where $]0,1] = \{x \in \mathbb{R} : 0 < x \le 1\}$.

Result 2. ([5]) *Suppose a mapping* $f : I \to \mathbb{R}$ *satisfies the functional equation* (1.14) *for all* $(p_1, \ldots, p_n) \in \Gamma_n$, $(q_1, \ldots, q_m) \in \Gamma_m$, $n \ge 3$, $m \ge 3$ *being fixed integers. If* $f(1) = f(0) = 0$, *then* f *is of the form*

$$f(p) = \begin{cases} a(p) + D(p,p) & \text{if } 0 < p \le 1 \\ 0 & \text{if } p = 0 \end{cases}$$

where $a : \mathbb{R} \to \mathbb{R}$ *is additive;* $D : \mathbb{R} \times \,]0,1] \to \mathbb{R}$ *is additive in the first variable and there exists a mapping* $E : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ *additive in both variables such that* $a(1) = E(1,1)$ *and*

$$D(pq, pq) = D(pq, p) + D(pq, q) + E(p,q) \tag{2.1}$$

for all $p \in \,]0,1]$ *and* $q \in \,]0,1]$.

Using the fact that $a(1) = E(1,1)$, it can be easily deduced from (2.1) that

$$a(1) + D(1,1) = 0. \tag{2.2}$$

3. On the Functional Equation (1.11)

The main result of this paper is the following

Theorem. *Let* c *be a nonzero given constant and* $f : I \to \mathbb{R}$, $g : I \to \mathbb{R}$ *be mappings which satisfy the equation* (1.11) *for all* $(p_1, \ldots, p_n) \in \Gamma_n$, $(q_1, \ldots, q_m) \in \Gamma_m$, $n \ge 3$, $m \ge 3$ *being fixed integers. If* $f : I \to \mathbb{R}$ *and* $g : I \to \mathbb{R}$ *satisfy respectively* (1.12) *and* (1.13), *then for all* $p \in I$, *any general solution of* (1.11) *is of the form*

$$\begin{cases} \text{(i)} \quad f(p) = \begin{cases} a(p) + D(p,p) & \text{if } 0 < p \le 1 \\ 0 & \text{if } p = 0 \end{cases} \\ \text{(ii)} \quad g(p) = A_1(p) \end{cases} \tag{3.1}$$

or

$$(3.2)\qquad \begin{cases} \text{(i)} \quad f(p) = \begin{cases} a(p) + D(p,p) + \frac{c}{2}p[\ell(p)]^2 & \textit{if } 0 < p \leq 1 \\ 0 & \textit{if } p = 0 \end{cases} \\ \text{(ii)} \quad g(p) = p\,\ell(p) \quad \textit{if } 0 \leq p \leq 1 \end{cases}$$

or

$$(3.3)\qquad \begin{cases} \text{(i)} \quad f(p) = \begin{cases} a(p) + D(p,p) + c\lambda^2(M(p) - p) & \textit{if } 0 < p \leq 1 \\ 0 & \textit{if } p = 0 \end{cases} \\ \text{(ii)} \quad g(p) = \lambda(M(p) - p) \end{cases}$$

where $A_1 : \mathbb{R} \to \mathbb{R}$ *is an additive mapping with* $A_1(1) = 0$*;* $a : \mathbb{R} \to \mathbb{R}$ *and* $D : \mathbb{R}\times]0,1] \to \mathbb{R}$ *are mappings meeting all the requirements as stated in Result* 2*;* $\ell : I \to \mathbb{R}$ *is a mapping logarithmic on* I *in the sense of Definition* 2*;* $M : I \to \mathbb{R}$ *is a nonconstant nonadditive mapping which is multiplicative on* I *in the sense of Definition* 1*; and* λ *is an arbitrary nonzero real constant.*

The proof of this theorem needs the following

Lemma. *Let* $g : I \to \mathbb{R}$ *be a mapping which satisfies the equation*

$$(3.4)\qquad \left[\sum_{j=1}^{m} g(xq_j) - g(x)\right]\sum_{t=1}^{m} g(r_t) = \left[\sum_{t=1}^{m} g(xr_t) - g(x)\right]\sum_{j=1}^{m} g(q_j)$$

for all $x \in I$ *and* $(q_1, \dots, q_m) \in \Gamma_m$*,* $(r_1, \dots, r_m) \in \Gamma_m$*,* $m \geq 3$ *being a fixed integer. Suppose* $g : I \to \mathbb{R}$ *also satisfies* (1.13)*. Then for all* $p \in I$*, any general solution* g *of* (3.4) *is of the form* (3.1)*(ii) or* (3.2)*(ii) or* (3.3)*(ii) with mappings* A_1*,* ℓ*,* M *and the constant* λ *as mentioned in the statement of the Theorem stated above.*

Proof of the Lemma. We divide our discussion into two cases:

Case 1. $\sum_{t=1}^{m} g(r_t) \equiv 0$ on Γ_m.

In this case, $\sum_{t=1}^{m} g(r_t) = 0$ for all $(r_1, \dots, r_m) \in \Gamma_m$. By Result 1, there exists an additive mapping $A_1 : \mathbb{R} \to \mathbb{R}$ such that

$$(3.5)\qquad g(p) = A_1(p) - \frac{1}{m}A_1(1)$$

for all $p \in I$. Putting $p = 0$ in (3.5), using (1.13) and $A_1(0) = 0$, it follows that $A_1(1) = 0$. Then (3.5) reduces to (3.1)(ii) with $A_1(1) = 0$.

Case 2. $\sum_{t=1}^{m} g(r_t)$ does not vanish identically on Γ_m.

In this case, there exists a probability distribution $(r_1^*, \ldots, r_m^*) \in \Gamma_m$ such that $\sum_{t=1}^{m} g(r_t^*) \neq 0$. Substituting $r_t = r_t^*$ in (3.4) and using $\sum_{t=1}^{m} g(r_t^*) \neq 0$, we obtain the equation

$$\sum_{j=1}^{m} g(xq_j) = M(x) \sum_{j=1}^{m} g(q_j) + g(x) \tag{3.6}$$

where $M : I \to \mathbb{R}$ is defined as

$$M(x) = \left[\sum_{t=1}^{m} g(r_t^*)\right]^{-1} \sum_{t=1}^{m} [g(xr_t^*) - g(x)r_t^*] \tag{3.7}$$

for all $x \in I$. From (3.7) and (1.13), it follows that

$$M(0) = 0, \quad M(1) = 1. \tag{3.8}$$

Thus, $M : I \to \mathbb{R}$ is a nonconstant mapping. Now let us write (3.6) in the form

$$\sum_{j=1}^{m} [g(xq_j) - M(x)g(q_j) - q_j g(x)] = 0.$$

By Result 1, there exists an additive mapping $\bar{E} : \mathbb{R} \times I \to \mathbb{R}$, additive in the first variable, such that

$$g(xq) - M(x)g(q) - qg(x) = \bar{E}(q; x) - \frac{1}{m}\bar{E}(1; x) \tag{3.9}$$

for all $x \in I$, $q \in I$. Since $\bar{E}$ is additive in the first variable, so $\bar{E}(0; x) = 0$ for all $x \in I$. Putting $q = 0$ in (3.9) and making use of (1.13), it also follows that $\bar{E}(1; x) = 0$ for all $x \in I$. Consequently, (3.9) reduces to the equation

$$g(xq) - M(x)g(q) - qg(x) = \bar{E}(q; x) \tag{3.10}$$

for all $q \in I$, $x \in I$. Making use of (1.13) and (3.8), it can be concluded from (3.10) that $\bar{E}(q; 0) = 0$ and $\bar{E}(q; 1) = 0$ for all $q \in I$. Now we prove that, indeed,

$$\bar{E}(q; x) = 0 \tag{3.11}$$

for all $q \in I$, $x \in I$. To the contrary, suppose that there exists a pair (q^*, x^*), $q^* \in I$, $x^* \in I$ such that $\bar{E}(q^*; x^*) \neq 0$. Keeping in view the information already obtained, we must have $0 < q^* < 1$ and $0 < x^* < 1$. To proceed further, we prove that

$$\begin{aligned} r - [\bar{E}(q^*; x^*)]^{-1}\{M(x^*)\bar{E}(r; q^*) + \bar{E}(q^* r; x^*) - \bar{E}(r; x^* q^*)\} \\ = \bar{E}(q^*; x^*)[M(x^*)M(q^*) - M(x^* q^*)]g(r) \end{aligned} \tag{3.12}$$

holds for all r, $0 \le r \le 1$. Indeed, (3.12) holds for $r = 0$ and $r = 1$ because of (1.13) and $\bar{E}(0; x) = 0$, $\bar{E}(1; x) = 0$ for all $x \in I$. Now we prove (3.12) for all r, $0 < r < 1$.

Let $0 < r < 1$. Since $0 < q^* < 1$, $0 < x^* < 1$, we have

$$g(x^*(q^*r)) = g((x^*q^*)r). \tag{3.13}$$

Making use of (3.10), it can be proved that

$$\begin{aligned} g(x^*(q^*r)) = {} & M(x^*)M(q^*)g(r) + rM(x^*)g(q^*) \\ & + M(x^*)\bar{E}(r; q^*) + q^*rg(x^*) + \bar{E}(q^*r; x^*) \end{aligned} \tag{3.14}$$

and

$$\begin{aligned} g((x^*q^*)r) = {} & M(x^*q^*)g(r) + rM(x^*)g(q^*) + rq^*g(x^*) \\ & + r\bar{E}(q^*; x^*) + \bar{E}(r; x^*q^*). \end{aligned} \tag{3.15}$$

From (3.13), (3.14), (3.15) and $\bar{E}(q^*; x^*) \neq 0$, (3.12) follows for all r, $0 < r < 1$. Since the mappings $r \mapsto r$ and $r \mapsto \bar{E}(r\,;\cdot)$ are additive on I, the left hand side of (3.12) is a mapping additive on I. Now we prove that the right hand side of (3.12) is a mapping not additive on I. To the contrary, suppose the right hand side of (3.12) is a mapping additive on I. This is possible only if $g : I \to \mathbb{R}$ is additive. Then, for all $(r_1, \ldots, r_m) \in \Gamma_m$, $\sum_{t=1}^{m} g(r_t) = g(1) = 0$ contradicting $\sum_{t=1}^{m} g(r_t^*) \neq 0$. So, (3.11) holds for all $q \in I$, $x \in I$. Now (3.10) reduces to

$$g(xq) = M(x)g(q) + qg(x) \tag{3.16}$$

valid for all $x \in I$, $q \in I$. The left hand side of (3.16) is symmetric in x and q. Hence, so must also be the right hand side of (3.16). This gives us the equation

$$[M(x) - x]g(q) = [M(q) - q]g(x) \tag{3.17}$$

for all $x \in I$, $q \in I$.

Case 2.1. The mapping $x \mapsto M(x) - x$ vanishes identically on I.

In this case, (3.16) reduces to

$$g(xq) = xg(q) + qg(x) \tag{3.18}$$

for all $x \in I$ and $q \in I$. The most general solution of (3.18) is of the form (3.2) (ii) in which $\ell : I \to \mathbb{R}$ is a mapping logarithmic in the sense of Definition 2.

Case 2.2. The mapping $x \mapsto M(x) - x$ does not vanish identically on I.

In this case, keeping in view (3.8), there exists an element $x_0 \in \,]0, 1[$ such that $(M(x_0) - x_0) \neq 0$. Putting $x = x_0$ in (3.17), using $(M(x_0) - x_0) \neq 0$ and performing necessary calculations, (3.3)(ii) follows with $\lambda = [M(x_0) -$

$x_0]^{-1}g(x_0)$. We prove that $\lambda \neq 0$. To the contrary, suppose that $\lambda = 0$. Then (3.3)(ii) reduces to $g(p) = 0$ for all $p \in I$. Consequently, $\sum_{t=1}^{m} g(r_t) = 0$ for all $(r_1, \ldots, r_m) \in \Gamma_m$ contradicting $\sum_{t=1}^{m} g(r_t^*) \neq 0$. So, $\lambda \neq 0$. Now elimination of g from (3.16) and (3.3)(ii) (with $\lambda \neq 0$) gives rise to the equation $M(xq) = M(x)M(q)$ valid for all $q \in I$, $x \in I$. In particular, $M(xq) = M(x)M(q)$ for all $x \in]0,1[$, $q \in]0,1[$. Thus, $M : I \to \mathbb{R}$ is multiplicative on I in the sense of Definition 1. Now we prove that $M : I \to \mathbb{R}$ is nonadditive. To the contrary, suppose $M : I \to \mathbb{R}$ is additive. Then, for all $(r_1, \ldots, r_m) \in \Gamma_m$, (3.3)(ii) gives

$$\sum_{t=1}^{m} g(r_t) = \lambda\Big(\sum_{t=1}^{m} M(r_t) - 1\Big) = \lambda\Big(M\Big(\sum_{t=1}^{m} r_t\Big) - 1\Big) = \lambda(M(1) - 1) = 0$$

contradicting $\sum_{t=1}^{m} g(r_t^*) \neq 0$. ∎

Proof of the Theorem. Let us write (1.11) in the form

$$\sum_{i=1}^{n} \Big\{ \sum_{j=1}^{m} f(p_i q_j) - f(p_i) - p_i \sum_{j=1}^{m} f(q_j) - c g(p_i) \sum_{j=1}^{m} g(q_j) \Big\} = 0.$$

By Result 1, there exists a mapping $A : \mathbb{R} \times \Gamma_m \to \mathbb{R}$, additive in the first variable, such that

$$(3.19)\quad \sum_{j=1}^{m} f(pq_j) - f(p) - p\sum_{j=1}^{m} f(q_j) - cg(p)\sum_{j=1}^{m} g(q_j) = A(p; q_1, \ldots, q_m) - \frac{1}{n}A(1; q_1, \ldots, q_m)$$

for all $p \in I$ and $(q_1, \ldots, q_m) \in \Gamma_m$. Putting $p = 0$ in (3.19), using (1.12), (1.13) and $A(0; q_1, \ldots, q_m) = 0$, it follows that $A(1; q_1, \ldots, q_m) = 0$. Consequently, (3.19) reduces to

$$(3.20)\quad \sum_{j=1}^{m} f(pq_j) - f(p) - p\sum_{j=1}^{m} f(q_j) - cg(p)\sum_{j=1}^{m} g(q_j) = A(p; q_1, \ldots, q_m)$$

valid for all $p \in I$ and $(q_1, \ldots, q_m) \in \Gamma_m$.

Let $x \in I$ and $(r_1, \ldots, r_m) \in \Gamma_m$ be any probability distribution. Putting $p = xr_1, \ldots, xr_m$ in (3.20), adding the resulting m equations and using the additivity of A in the first variable, the equation

$$(3.21)\quad \sum_{t=1}^{m}\sum_{j=1}^{m} f(xr_tq_j) - \sum_{t=1}^{m} f(xr_t) - x\sum_{j=1}^{m} f(q_j) - c\sum_{t=1}^{m} g(xr_t)\sum_{j=1}^{m} g(q_j) = A(x; q_1, \ldots, q_m)$$

arises. Also, from (3.20),

$$(3.22)\quad \sum_{t=1}^{m} f(xr_t) = f(x) + x\sum_{t=1}^{m} f(r_t) + cg(x)\sum_{t=1}^{m} g(r_t) + A(x; r_1, \ldots, r_m).$$

From (3.21) and (3.22), the equation

$$(3.23)\quad \sum_{t=1}^{m}\sum_{j=1}^{m} f(xr_tq_j) - f(x) - x\sum_{t=1}^{m} f(r_t) - x\sum_{j=1}^{m} f(q_j) - A(x; r_1, \ldots, r_m) - A(x; q_1, \ldots, q_m) = cg(x)\sum_{t=1}^{m} g(r_t) + c\sum_{t=1}^{m} g(xr_t)\sum_{j=1}^{m} g(q_j)$$

follows. The left hand side of (3.23) is symmetric in r_t and q_j, $t = 1, \ldots, m$; $j = 1, \ldots, m$. Hence, so should be the right hand side of (3.23). This gives rise to the equation (3.4) for all $x \in I$ and $(r_1, \ldots, r_m) \in \Gamma_m$ as $c \neq 0$. Also, (1.13) holds by assumption. So, by the Lemma proved above, $g : I \to \mathbb{R}$ is of the form (3.1)(ii) or (3.2)(ii) or (3.3)(ii) with mappings A_1, M and the constant λ as stated in the statement of the Theorem.

From (3.1)(ii) with $A_1(1) = 0$ and the additivity of $A_1 : \mathbb{R} \to \mathbb{R}$, it follows that $\sum_{i=1}^{n} g(p_i) = 0$ for all $(p_1, \ldots, p_n) \in \Gamma_n$. Making use of this fact in (1.11), we observe that $f : I \to \mathbb{R}$ satisfies the equation (1.14) for all $(p_1, \ldots, p_n) \in \Gamma_n$, $(q_1, \ldots, q_m) \in \Gamma_m$, $n \geq 3$, $m \geq 3$ being fixed integers. Moreover, (1.12) holds by assumption.

From (1.11) and (3.2)(ii), the equation

$$\sum_{i=1}^{n}\sum_{j=1}^{m} f(p_iq_j) = \sum_{i=1}^{n} f(p_i) + \sum_{j=1}^{m} f(q_j) + c\sum_{i=1}^{n}\sum_{j=1}^{m} p_iq_j\, \ell(p_i)\ell(q_j)$$

follows. Since $\ell : I \to \mathbb{R}$ is logarithmic, the above equation can be written as

$$\sum_{i=1}^{n}\sum_{j=1}^{m}\left\{f(p_iq_j)-\frac{1}{2}c\,p_iq_j[\ell(p_iq_j)]^2\right\} \tag{3.24}$$
$$=\sum_{i=1}^{n}\left\{f(p_i)-\frac{1}{2}c\,p_i[\ell(p_i)]^2\right\}+\sum_{j=1}^{m}\left\{f(q_j)-\frac{1}{2}c\,q_j[\ell(q_j)]^2\right\}.$$

Define a mapping $f_1 : I \to \mathbb{R}$ as

$$f_1(p) = f(p) - \frac{1}{2}c\,p[\ell(p)]^2 \tag{3.25}$$

for all $p \in I$. Then, making use of (1.12), it follows that $f_1(0) = 0$ and $f_1(1) = 0$. Also, from (3.24) and (3.25), one can infer that $f_1 : I \to \mathbb{R}$ also satisfies the equation (1.14).

From (1.11) and (3.3)(ii) with $\lambda \neq 0$, we obtain the equation

$$\sum_{i=1}^{n}\sum_{j=1}^{m} f(p_iq_j) = \sum_{i=1}^{n} f(p_i) + \sum_{j=1}^{m} f(q_j)$$
$$+ c\lambda^2\left[\sum_{i=1}^{n} M(p_i) - 1\right]\left[\sum_{j=1}^{m} M(q_j) - 1\right]$$

which can be written in the form (using the multiplicativity of M)

$$\sum_{i=1}^{n}\sum_{j=1}^{m}\{f(p_iq_j) - c\lambda^2 M(p_iq_j) + c\lambda^2 p_iq_j\} \tag{3.26}$$
$$= \sum_{i=1}^{n}\{f(p_i) - c\lambda^2 M(p_i) + c\lambda^2 p_i\} + \sum_{j=1}^{m}\{f(q_j) - c\lambda^2 M(q_j) + c\lambda^2 q_j\}.$$

Define $f_2 : I \to \mathbb{R}$ as

$$f_2(p) = f(p) - c\lambda^2 M(p) + c\lambda^2 p \tag{3.27}$$

for all $p \in I$. From (3.27), (1.12) and (3.8), one can infer that $f_2(0) = 0$ and $f_2(1) = 0$. Also, from (3.26) and (3.27), it follows that $f_2 : I \to \mathbb{R}$ also satisfies (1.14).

Making use of Result 2, it follows that

$$f(p) = f_1(p) = f_2(p) = \begin{cases} a(p) + D(p,p) & \text{if } 0 < p \leq 1 \\ 0 & \text{if } p = 0 \end{cases} \tag{3.28}$$

where $a : \mathbb{R} \to \mathbb{R}$ and $D : \mathbb{R} \times]0,1] \to \mathbb{R}$ are as stated in Result 2. From (3.28) and respectively (3.1)(ii), (3.2)(ii), (3.3)(ii), the equations (3.1)(i), (3.2)(i) and (3.3)(i) follows. ■

Remark. In our subsequent research work, the Theorem proved above has proved to be useful in obtaining the general solutions of the functional equation ([4], [6])

$$\sum_{i=1}^{n}\sum_{j=1}^{m} F(p_i q_j) = \sum_{i=1}^{n} G(p_i) + \sum_{j=1}^{m} H(q_j) + \sum_{i=1}^{n} K(p_i) \sum_{j=1}^{m} L(q_j)$$

in which $F : I \to \mathbb{R}$, $G : I \to \mathbb{R}$, $H : I \to \mathbb{R}$, $K : I \to \mathbb{R}$, $L : I \to \mathbb{R}$, $(p_1, \ldots, p_n) \in \Gamma_n$, $(q_1, \ldots, q_m) \in \Gamma_m$ and $n \geq 3$, $m \geq 3$ being fixed integers. The details are complicated and will be published elsewhere.

4. Comments

In this section, we discuss the importance of solutions (3.1), (3.2) and (3.3) in information theory and statistics.

Let $(p_1, \ldots, p_n) \in \Gamma_n$ and $S = \{i : 1 \leq i \leq n, 0 < p_i \leq 1\}$.

Then S is a nonempty set.

Let us consider (3.1). Using equations (2.1), (2.2), $a(1) = E(1,1)$, $A_1(1) = 0$ and the additivity of (i) $a : \mathbb{R} \to \mathbb{R}$ (ii) $D : \mathbb{R} \times]0,1] \to \mathbb{R}$ in the first variable (iii) $E : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ in both variables and (iv) $A_1 : \mathbb{R} \to \mathbb{R}$, equation (3.1) gives

$$\sum_{i=1}^{n} f(p_i) = -D(1,1) + \sum_{i \in S} D(p_i, p_i) \tag{4.1}$$

$$\sum_{i=1}^{n} g(p_i) = 0.$$

Keeping into consideration the form of the Shannon entropies given by (1.1), it seems desirable to consider the mapping $D : \mathbb{R} \times]0,1] \to \mathbb{R}$ defined as

$$D(x,y) = dx \log_2 y \tag{4.2}$$

for all $x \in \mathbb{R}$, $y \in]0,1]$, d being an arbitrary real constant. The case $d = 0$ is not of much importance. So, we restrict to $d \neq 0$. Now $D(p,p) = dp \log_2 p$ for all $p \in]0,1]$. So, $D(1,1) = 0$. To accommodate the 0-probabilities, it seems natural to assume $\lim\limits_{p \to 0^+} D(p,p) = 0$, that is, $0 \log_2 0 = 0$ as $d \neq 0$. Now, (4.1) gives

$$\sum_{i=1}^{n} f(p_i) = -\, dH_n(p_1, \ldots, p_n).$$

Thus, the summand $\sum\limits_{i=1}^{n} f(p_i)$ represents the Shannon entropy or the first order moment $\mu_1'(Z_n)$ up to nonzero multiplicative constant.

Now we discuss (3.2). Here, proceeding as in the case of solution (3.1) and using the fact that $\ell : I \to \mathbb{R}$ is logarithmic, it follows that

$$\sum_{i=1}^{n} f(p_i) = -D(1,1) + \sum_{i\in S} D(p_i, p_i) + \frac{c}{2}\sum_{i\in S} p_i[\ell(p_i)]^2 \tag{4.3}$$

$$\sum_{i=1}^{n} g(p_i) = \sum_{i\in S} p_i \ell(p_i) \quad \text{because} \quad 0\,\ell(0) = 0. \tag{4.4}$$

Keeping in view the form of the Shannon entropies given by (1.1), it seems desirable to choose $D : \mathbb{R} \times \,]0,1] \to \mathbb{R}$ as defined by (4.2) and assume $0 \log_2 0 = 0$ as justified above but as regards the logarithmic mapping $\ell : I \to \mathbb{R}$ is concerned, it seems appropriate to choose $\ell : I \to \mathbb{R}$ defined as

$$\ell(p) = \begin{cases} \lambda_1 \log_2 p & \text{if } 0 < p \leq 1 \\ 0 & \text{if } p = 0 \end{cases} \tag{4.5}$$

where λ_1 is an arbitrary, real constant. Here, too, the case $\lambda_1 = 0$ is not of much importance. So, we restrict to $\lambda_1 \neq 0$. To accommodate 0-probabilities, it seems desirable to assume

$$0 \log_2 0 = 0 \quad \text{and} \quad 0(\log_2 0)^2 = 0.$$

Since

$$\mu_2'(Z_n) = \sum_{i=1}^{n} p_i z_i^2 = \sum_{i\in S} p_i(-\log_2 p_i)^2 = \frac{1}{\lambda_1^2}\sum_{i\in S} p_i[\ell(p_i)]^2,$$

it follows that

$$\sum_{i=1}^{n} f(p_i) = -d\mu_1'(Z_n) + \frac{c\lambda_1^2}{2}\mu_2'(Z_n) \tag{4.6}$$

and

$$\sum_{i=1}^{n} g(p_i) = -\lambda_1 \mu_1'(Z_n) \tag{4.7}$$

where $\lambda_1 \neq 0$ and $d \neq 0$ are arbitrary constants and $c \neq 0$ is a given constant. Thus the summand $\sum_{i=1}^{n} f(p_i)$ represents a suitable linear combination of the first two moments of the random variable Z_n. On the other hand, the summand $\sum_{i=1}^{n} g(p_i)$ represents the first moment of Z_n up to a nonzero multiplicative constant.

Now we discuss (3.3). Here, proceeding as in the case of solution (3.1) and using the fact that $M : I \to \mathbb{R}$ is a nonadditive and multiplicative mapping, it follows that

$$(4.8)\qquad \sum_{i=1}^{n} f(p_i) = -D(1,1) + \sum_{i\in S} D(p_i,p_i) - c\lambda^2\left[1 - \sum_{i=1}^{n} M(p_i)\right]$$

and

$$(4.9)\qquad \sum_{i=1}^{n} g(p_i) = -\lambda\left[1 - \sum_{i=1}^{n} M(p_i)\right]$$

where λ is an arbitrary nonzero constant.

For any probability distribution $(p_1,\dots,p_n)\in\Gamma_n$, the entropies

$$(4.10)\qquad H_n^{\alpha}(p_1,\dots,p_n) = (1-2^{1-\alpha})^{-1}\left(1 - \sum_{i=1}^{n} p_i^{\alpha}\right)$$

with $H_n^{\alpha}:\Gamma_n\to\mathbb{R}$, $n=1,2,\dots$; $\alpha\in\mathbb{R}$, $\alpha>0$, $\alpha\neq 1$, $1^{\alpha}:=1$, $0^{\alpha}:=0$ are called the entropies of degree α, $\alpha>0$, $\alpha\neq1$, $\alpha\in\mathbb{R}$. These entropies are nonadditive and were given by J. Havrda and F. Charvát [3].

Keeping into consideration the forms of entropies given by (1.1) and (4.10), it is desirable to choose $D:\mathbb{R}\times\,]0,1]\to\mathbb{R}$ as in the case of solutions (3.1) and (3.2) but choose $M: I\to\mathbb{R}$ defined as $M(p)=p^{\alpha}$ for all $p\in I$ with $\alpha\in\mathbb{R},\alpha>0,\alpha\neq1,0^{\alpha}:=0$ and $1^{\alpha}:=1$. Then (4.8) and (4.9) give

$$(4.11)\qquad \sum_{i=1}^{n} f(p_i) = -dH_n(p_1,\dots,p_n) - c\lambda^2(1-2^{1-\alpha})H_n^{\alpha}(p_1,\dots,p_n)$$

and

$$(4.12)\qquad \sum_{i=1}^{n} g(p_i) = -\lambda(1-2^{1-\alpha})H_n^{\alpha}(p_1,\dots .p_n).$$

Thus we see that in (3.3), the mapping g is connected only with the nonadditive entropy $H_n^{\alpha}(p_1,\dots,p_n)$ whereas f is connected with both the entropies $H_n^{\alpha}(p_1,\dots,p_n)$ and $H_n(p_1,\dots,p_n)$ which is also the first moment $\mu_1'(Z_n)$.

References

[1] T. W. Chaundy, J. B. McLeod, *On a functional equation*, Edinburgh Math. Notes 43 (1960), 7–8.

[2] Z. Daróczy, L. Losonczi, *Über die Erweiterung der auf einer Punktmenge additiven Funktionen*, Publ. Math. (Debrecen) 14 (1967), 239–245.

[3] J. Havrda, F. Charvát, *Quantification method of classification process. Concept of structural α-entropy*, Kybernetika (Prague) 3 (1967), 30–35.

[4] Pl. Kannappan, *On some functional equations from additive and nonadditive measures II*, Second International Conference on Information Sciences and Systems (University of Patras, Patras, 1979), Vol. I, 45–50, Reidel—Dordrecht—Boston—Mass., 1980.

[5] L. Losonczi, Gy. Maksa, *On some functional equations of the information theory*, Acta Math. Acad. Sci. Hung. 39 (1982), 73–82.
[6] P. Nath, *On some sum form functional equations related to information theory*, Journal of Interdisciplinary Mathematics 9 (1) (2006), 1–15.
[7] C. E. Shannon, *A mathematical theory of communication*, Bell Syst. Tech. Jour. 27 (1948), 378–423; 623–656.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF DELHI
DELHI 110 007, INDIA
E-mail: pnathmaths@gmail.com,
dksingh@maths.du.ac.in, dhiraj426@rediffmail.com

Received June 6, 2008.