Kragujevac Journal of Mathematics Volume 39(2) (2015), Pages 141–148.

## ON A FUNCTIONAL EQUATION ARISING IN STATISTICS

P. NATH<sup>1</sup> AND D. K. SINGH<sup>2</sup>

ABSTRACT. The general solutions of a functional equation, derived from the first and second order moments of a finite discrete random variable, have been obtained.

# 1. Introduction

Functional equations appear in various branches of Pure Mathematics, Applied Mathematics, Geometry, Economics, Business Mathematics, Probability Theory and Mathematical Statistics (see [1, 3, 11]). In Mathematical Statistics, usually two types of random variables are discussed (i) continuous (ii) discrete. C. F. Gauss (1809) characterized the normal distribution (also called the Gaussian distribution, see pp. 106–109 in [1]). For further characterization of normal distributions, see [2, 6, 11].

This paper deals with functional equations concerning some discrete random variables which are useful from information-theoretic point of view.

For  $n=1,2,\ldots$  let  $\Gamma_n=\left\{(p_1,\ldots,p_n):p_i\geq 0,i=1,\ldots,n;\sum_{i=1}^np_i=1\right\}$  denote the set of all n-component complete probability distributions with nonnegative elements. Let  $\mathbb R$  denote the set of all real numbers;  $I=\{x\in\mathbb R:0\leq x\leq 1\}$ , the unit closed interval;  $]0,1[=\{x\in\mathbb R:0< x< 1\}$ , the unit open interval and  $]0,1]=\{x\in\mathbb R:0< x\leq 1\}$ .

Received: March 29, 2015. Accepted: July 15, 2015.

Key words and phrases. Moments of a discrete random variable, additive mapping, logarithmic mapping, multiplicative mapping.

 $<sup>2010\</sup> Mathematics\ Subject\ Classification.\ 39B22,\ 39B52.$ 

(I) 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 different real values  $z_1, \ldots, z_n$  with respective probabilities  $p_1, \ldots, p_n$  where

$$z_i = \begin{cases} -\log_2 p_i, & \text{if } 0 < p_i \le 1, \\ 0, & \text{if } p_i = 0. \end{cases}$$

The rth order moment [4] of  $Z_n$  about the origin, r = 0, 1, 2, ... is defined as

(1.1) 
$$\mu'_r(Z_n) = \sum_{i=1}^n p_i z_i^r$$
.

Consider (1.1) when r = 0. In this case, we 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 the point of view of information theory.

Now, for r = 1, 2, ... let us consider (1.1). In this case, we define  $0(\log_2 0)^r := 0$ . Then  $\mu'_r(Z_n) = (-1)^r \sum_{i=1}^n p_i (\log_2 p_i)^r$ . Define the mappings  $\phi_r : I \to \mathbb{R}, r = 0, 1, 2, ...$  as

(1.2) 
$$\phi_r(p) = \begin{cases} p, & \text{if } r = 0, \\ (-1)^r p(\log_2 p)^r, & \text{if } r = 1, 2, \dots \end{cases}$$

Now, from (1.1) and (1.2), it follows that

(1.3) 
$$\mu'_r(Z_n) = \sum_{i=1}^n \phi_r(p_i),$$

for all r = 0, 1, 2, ... so all the moments admit of a sum representation. The mapping  $\phi_r$  is called the generating function of the moment  $\mu'_r(Z_n)$ . From (1.2), we observe that  $\phi_0(0) = 0$ ,  $\phi_0(1) = 1$  but  $\phi_r(0) = 0$ ,  $\phi_r(1) = 0$ , (r = 1, 2, ...). The mapping  $\phi_0$  satisfies the functional equation  $\phi_0(pq) = \phi_0(p)\phi_0(q)$ , for all  $p \in I$ ,  $q \in I$ . On the other hand, the mapping  $\phi_1$  satisfies the functional equation  $\phi_1(pq) = q\phi_1(p) + p\phi_1(q)$ , for all  $p \in I$ ,  $q \in I$ . Also, both the mappings  $\phi_1$  and  $\phi_2$  satisfy  $\phi_2(pq) = q\phi_2(p) + p\phi_2(q) + 2\phi_1(p)\phi_1(q)$ , for all  $p \in I$ ,  $q \in I$ . So, it seems desirable to pay attention to the functional equation

(1.4) 
$$f(pq) = qf(p) + pf(q) + cg(p)g(q),$$

where  $p \in I$ ,  $q \in I$ ,  $f: I \to \mathbb{R}$ ,  $g: I \to \mathbb{R}$  are unknown mappings and  $c \neq 0$  is a given real constant.

(II) Let us consider a discrete random variable  $X_n$  taking the different real values  $x_1, x_2, \ldots, x_n$  where  $x_i = M(p_i), 0 \le p_i \le 1, i = 1, 2, \ldots, n$ ; M being a non-constant multiplicative mapping. The rth order moment of  $X_n$  about the origin,  $r = 0, 1, 2, \ldots$  is defined as  $\bar{\mu}_r(X_n) = \sum_{i=1}^n p_i [M(p_i)]^r$ .

Let us define the mappings  $\psi_r: I \to \mathbb{R}$ , r = 0, 1, 2, ... as  $\psi_r(p) = p[M(p)]^r$ . Then  $\bar{\mu}_r(X_n) = \sum_{i=1}^n \psi_r(p_i)$ . Thus, all the moments  $\bar{\mu}_r(X_n)$  admit of a sum representation.

The mapping  $\psi_r$  is called the generating function of the moment  $\bar{\mu}_r(X_n)$ . The mappings  $\psi_r: I \to \mathbb{R}, r = 0, 1, 2, \ldots$  satisfy the functional equation  $\psi_r(pq) = \psi_r(p)\psi_r(q)$ , for all  $p \in I$ ,  $q \in I$ . Thus, both the mappings  $\phi_0$  and  $\psi_0$  are multiplicative. For  $r = 1, 2, \ldots$  each of the mapping  $\psi_r$  is multiplicative whereas none of  $\phi_r$  is multiplicative.

The fact that both  $\psi_2$  and  $\psi_1$  are multiplicative does not yield a functional equation different from f(pq) = f(p)f(q),  $f: I \to \mathbb{R}$ ,  $p \in I$ ,  $q \in I$ . But this is not the situation in the case of  $\phi_2$  and  $\phi_1$ . It is for this reason that the study of (1.4) is important.

(III) Consider the discrete random variable X which takes the values  $1, 2, \ldots$  with probabilities  $p_1, p_2, \ldots$  where

(1.5) 
$$p_x = \text{Prob}(X = x) = \theta(1 - \theta)^{x-1}, \quad x = 1, 2, \dots$$

and  $0 < \theta < 1$ , x denoting the number of trials needed to observe the first success in a sequence of independent Bernoulli trials and  $\theta$  denotes the probability of success in each Bernoulli trial. The probability distribution  $(p_1, p_2, ...)$ , given by (1.5), is said to be a geometric probability distribution [7]. The value x of X does not depend upon the corresponding probability  $p_x$ . The rth order moment  $\hat{\mu}_r(X)$ , r = 1, 2, ... about the origin, is defined as  $\hat{\mu}_r(X) = \sum_{x=1}^{\infty} p_x x^r$ .

the origin, is defined as  $\hat{\mu}_r(X) = \sum_{x=1}^{\infty} p_x x^r$ . Clearly,  $\hat{\mu}_r(X) = \sum_{x=1}^{\infty} \theta (1-\theta)^{x-1} x^r$ . Since x does not depend upon  $p_x$ , there exists no sequence of mappings  $f_r: I \to \mathbb{R}, \ r=1,2,\ldots$  such that  $\hat{\mu}_r(X) = \sum_{x=1}^{\infty} f_r(p_x)$ . Hence, it is not possible to proceed further.

(IV) If the discrete random variable has first moment but second order moment is non-existent, then it is not possible to proceed as in (I) and (II). However, even if the second order moment exists, there is no guarantee that we will be able to get a functional equation, already known or a new one. For instance, in the case of above mentioned random variable X possessing geometric distribution, the first and second order moments exist and are respectively  $\frac{1}{\theta}$  and  $\frac{2-\theta}{\theta^2}$  but still we cannot proceed further.

Based upon the above observations, we think it is proper to mention specifically that the results proved in the next section 2 are not general. Rather, they are valid only for the specific random variable  $Z_n$ , mentioned in (I), which led to (1.4).

(V) To solve a functional equation, it was usual to assume the unknown mappings, appearing in it, to be differentiable (considered a strong regularity assumption); reduce the functional equation to a differential equation and then solve it. In real analysis, we know that if a mapping is differentiable at a point, then it is continuous at that point but not conversely. Thus the assumption 'a mapping is continuous at a point' is weaker (in the strict sense) than the assumption 'a mapping is differentiable at a point'. Thus, the trend of solving functional equations by weakening the regularity conditions in the strict sense started and this trend continued. Finally, it was realized that if possible, the functional equations should be solved without imposing any regularity condition on the mappings appearing in them.

The authors came across the functional equation (1.4) in [8] but its general solutions were not obtained. This, indeed, is the reason of writing this paper.

The object of this paper is to obtain the general solutions of (1.4) without imposing any regularity condition on any of the mappings f and g appearing in it.

Now, we mention some known definitions which are needed to develop sections 2 and 3 of this paper.

A mapping  $A : \mathbb{R} \to \mathbb{R}$  is said to be additive on  $\mathbb{R}$  if the equation A(x + y) = A(x) + A(y) holds for all  $x \in \mathbb{R}$ ,  $y \in \mathbb{R}$ .

A mapping  $M: I \to \mathbb{R}$  is said to multiplicative if M(pq) = M(p)M(q), for all  $p \in I$ ,  $q \in I$ .

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

In this section, we provide the general solutions of (1.4).

**Theorem 2.1.** Let  $c \neq 0$  be a given real constant and  $f: I \to \mathbb{R}$ ,  $g: I \to \mathbb{R}$  be mappings which satisfy the functional equation (1.4), for all  $p \in I$ ,  $q \in I$ . Then, for all  $p \in I$ , any general solution (f,g) of (1.4) is one of the following forms:

(2.1) (i) 
$$f(p) = p\ell(p)$$
, (ii)  $g(p) \equiv 0$ 

or

(2.2) (i) 
$$f(p) = p\ell(p) + \frac{1}{2}cp[\ell^*(p)]^2$$
, (ii)  $g(p) = p\ell^*(p)$ 

(2.3) (i) 
$$f(p) = p\ell(p) + c\mu^2[M(p) - p],$$
 (ii)  $g(p) = \mu[M(p) - p],$ 

where  $\mu$  is an arbitrary nonzero real constant;  $\ell: I \to \mathbb{R}$  is a logarithmic mapping;  $\ell^*: I \to \mathbb{R}$  is a logarithmic mapping which does not vanish identically on the open interval  $]0,1[; M: I \to \mathbb{R}$  is a multiplicative mapping and  $0\ell(0) = 0 = 0\ell^*(0)$ ,  $0[\ell^*(0)]^2 = 0$ .

*Proof.* Let  $p \in I$ ,  $q \in I$ ,  $r \in I$ . Now, using (1.4), we obtain

$$f(p(qr)) = qrf(p) + prf(q) + pqf(r) + cpg(q)g(r) + cg(p)g(qr)$$

and

$$f((pq)r) = qrf(p) + prf(q) + pqf(r) + crg(p)g(q) + cg(pq)g(r).$$

Equating the expressions for f(p(qr)) and f((pq)r), using  $c \neq 0$ ;

$$(2.4) g(r)[g(pq) - pg(q)] = g(p)[g(qr) - rg(q)]$$

follows for all  $p \in I$ ,  $q \in I$  and  $r \in I$ .

Our first task is to determine the general solutions of (2.4). For this purpose, we divide the discussion into two following cases.

Case 1. g vanishes identically on I.

In this case, g(p) = 0 for all  $p \in I$ . This is precisely (2.1)(ii).

Case 2. g does not vanish identically on I.

In this case, there exists an element  $r_0 \in I$  such that  $g(r_0) \neq 0$ . Putting  $r = r_0$  in (2.4), it follows that

(2.5) 
$$g(pq) = pg(q) + M(q)g(p),$$

where  $M: I \to \mathbb{R}$  is a mapping defined as  $M(x) = [g(r_0)]^{-1}[g(xr_0) - r_0g(x)]$ , for all  $x \in I$ . Since g(pq) = g(qp), (2.5) gives

$$(2.6) \quad [M(p) - p]q(q) = [M(q) - q]q(p)$$

valid for all  $p \in I$ ,  $q \in I$ .

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

In this case, M(p) = p for all  $p \in I$ . Making use of this form of M in (2.5), we obtain, for all  $p \in I$ ,  $q \in I$ , the functional equation g(pq) = pg(q) + qg(p) whose general solution is  $g(p) = p\ell(p)$ , where  $\ell : I \to \mathbb{R}$  is a logarithmic mapping and  $0\ell(0) = 0$ . Now,  $0 \neq g(r_0) = r_0\ell(r_0)$  for some  $r_0 \in I$ . Since  $\ell(0) = \ell(1) = 0$ , we have  $0 \neq g(r_0) = r_0\ell(r_0)$  for some  $r_0 \in ]0,1[$ . Hence  $\ell(r_0) \neq 0$  for some  $r_0 \in ]0,1[$ . So,  $g(p) = p\ell^*(p)$  where  $\ell^* : I \to \mathbb{R}$  is a logarithmic mapping which does not vanish identically on the open interval ]0,1[ and  $0\ell^*(0) = 0$ . Thus, we have obtained (2.2)(ii).

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

In this case, there exists an element  $q_0 \in [0, 1]$  such that  $M(q_0) - q_0 \neq 0$ . Substituting  $q = q_0$  in (2.6), it follows that  $g(p) = \mu[M(p) - p]$ , where  $\mu = g(q_0)[M(q_0) - q_0]^{-1}$ . We claim that  $\mu \neq 0$ . If possible, suppose  $\mu = 0$ . Then g(p) = 0 for all  $p \in I$  contradicting  $g(r_0) \neq 0$  for some  $r_0 \in I$ . So,  $\mu \neq 0$ . Thus, we have obtained (2.3)(ii). From (2.3)(ii), (2.5) and the fact that  $\mu \neq 0$ , it follows that M(pq) = M(p)M(q) for all  $p \in I$ ,  $q \in I$ . Thus, M is a multiplicative mapping.

Now we determine the forms of f corresponding to those of g.

From (1.4) and (2.1)(ii), we obtain the equation f(pq) = qf(p) + pf(q) for all  $p \in I$ ,  $q \in I$ . Hence f is of the form (2.1)(i). Thus, solution (2.1) is obtained.

From (1.4) and (2.2)(ii), we obtain the equation

(2.7) 
$$f(pq) = qf(p) + pf(q) + cp\ell^*(p)q\ell^*(q),$$

for all  $p \in I$ ,  $q \in I$ . Putting p = q = 0 in (2.7) and using the fact that  $0\ell^*(0) := 0$ , we obtain f(0) = 0. Now we consider (2.7) only for  $p \in ]0,1]$  and  $q \in ]0,1]$ . In this case, dividing both sides of (2.7) by  $pq \in ]0,1]$ , we obtain the equation

$$\frac{f(pq)}{pq} = \frac{f(p)}{p} + \frac{f(q)}{q} + c\ell^*(p)\ell^*(q)$$

which can be written as

$$\frac{f(pq)}{pq} - \frac{1}{2}c[\ell^*(pq)]^2 = \left\{ \frac{f(p)}{p} - \frac{1}{2}c[\ell^*(p)]^2 \right\} + \left\{ \frac{f(q)}{q} - \frac{1}{2}c[\ell^*(q)]^2 \right\},$$

as  $\ell^*(pq) = \ell^*(p) + \ell^*(q)$ , for all  $p \in ]0,1]$  and  $q \in ]0,1]$ . Define  $\ell: I \to \mathbb{R}$  as

(2.8) 
$$\ell(p) = \begin{cases} \frac{f(p)}{p} - \frac{1}{2}c[\ell^*(p)]^2, & \text{if } 0$$

Then  $\ell: I \to \mathbb{R}$ , defined by (2.8), is a logarithmic mapping. From (2.8), it follows that  $f(p) = p\ell(p) + \frac{1}{2}cp[\ell^*(p)]^2$  for  $0 . Also, <math>f(0) = 0 = 0\ell(0) + \frac{1}{2}c0[\ell^*(0)]^2$  as  $0\ell(0) := 0$  and  $0[\ell^*(0)]^2 = 0$ . Thus, (2.2)(i) holds. So, solution (2.2) is obtained.

From (1.4), (2.3)(ii) and the multiplicativity of M, we obtain

$$f(pq)-c\mu^2 M(pq)+c\mu^2 pq=q[f(p)-c\mu^2 M(p)+c\mu^2 p]+p[f(q)-c\mu^2 M(q)+c\mu^2 q],$$
 for all  $p\in I, q\in I$ . Now (2.3)(i) can be obtained. So, solution (2.3) is obtained.  $\square$ 

#### 3. Comments

The object of this section is to discuss the importance of solutions (2.1) to (2.3) from the point of view of statistics and information theory.

Consider (2.1). In this case, it is desirable to choose  $\ell = \ell^*$ , where  $\ell^* : I \to \mathbb{R}$  is a logarithmic mapping defined as

(3.1) 
$$\ell^*(p) = \begin{cases} -\log_2 p, & \text{if } 0$$

We know that from (1.3), each moment  $\mu'_r(Z_n)$  admits of a sum representation. Now, from (2.1), (3.1), (1.2) and (1.3), we have

$$\sum_{i=1}^{n} f(p_i) = \sum_{i=1}^{n} p_i \ell(p_i) = -\sum_{i=1}^{n} p_i \log_2 p_i = H_n(p_1, \dots, p_n) = \mu'_1(Z_n)$$

and  $\sum_{i=1}^{n} g(p_i) = 0$ , where  $H_n : \Gamma_n \to \mathbb{R}$ ,  $n = 1, 2, \ldots$  are the Shannon entropies [10] defined as  $H_n(p_1, \ldots, p_n) = -\sum_{i=1}^{n} p_i \log_2 p_i$ , for all  $(p_1, \ldots, p_n) \in \Gamma_n$  with  $0 \log_2 0 := 0$ . Thus, the mapping f, given by (2.1)(i), is related to the first order moment about the origin.

Now consider (2.2). Here, too, we choose  $\ell = \ell^*$  where  $\ell^* : I \to \mathbb{R}$  is defined by (3.1). Now, (2.2), together with (1.2) and (1.3), gives

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

and  $\sum_{i=1}^{n} g(p_i) = \mu'_1(Z_n)$ . Thus, the solution (2.2), of (1.4), is related to the first and second moments of  $Z_n$  about the origin.

Now consider the solution (2.3) of (1.4). Havrda and Charvat [5] defined the entropies  $H_n^{\alpha}: \Gamma_n \to \mathbb{R}, n = 1, 2, \dots$  of degree  $\alpha, \alpha > 0, \alpha \neq 1, \alpha \in \mathbb{R}$ , as

(3.2) 
$$H_n^{\alpha}(p_1, \dots, p_n) = (1 - 2^{1-\alpha})^{-1} \left(1 - \sum_{i=1}^n p_i^{\alpha}\right),$$

for all  $(p_1, \ldots, p_n) \in \Gamma_n$ . The Gini-Simpson [9] index of the probability distribution

 $(p_1, \ldots, p_n)$  is defined as  $(GS)_n(p_1, \ldots, p_n) = 1 - \sum_{i=1}^n p_i^2$ . If  $\alpha = 2$  in (3.2), then  $H_n^2(p_1, \ldots, p_n) = 2(1 - \sum_{i=1}^n p_i^2)$ . In particular, if  $(p, 1 - p) \in \Gamma_2, 0 \le p \le 1$ , then

(3.3) 
$$H_2^2(p, 1-p) = 4(p-p^2).$$

The right hand side of (3.3) is, indeed, the parabola passing through the points (0,0), (1,0) and  $(\frac{1}{2},1)$ .

In (2.3), let us choose  $\ell = \ell^*$ , where  $\ell^* : I \to \mathbb{R}$  is defined by (3.1); and  $M(p) = p^{\alpha}$ ,  $p \in I, \ \alpha \in \mathbb{R}, \ \alpha > 0, \ \alpha \neq 1$ . We obtain

$$f(p) = -p \log_2 p + c\mu^2 (p^{\alpha} - p)$$
 and  $g(p) = \mu(p^{\alpha} - p)$ .

Thus,

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

and 
$$\sum_{i=1}^{n} g(p_i) = -\mu(1-2^{1-\alpha})H_n^{\alpha}(p_1,\ldots,p_n).$$

Therefore, the mappings f and g, mentioned in (2.3), are related to the Shannon entropies and the entropies of degree  $\alpha$ .

The special case  $\alpha = 2$  deserves particular attention. In this case,

$$\sum_{i=1}^{n} f(p_i) = H_n(p_1, \dots, p_n) - \frac{1}{2} c \mu^2 H_n^2(p_1, \dots, p_n)$$

and 
$$\sum_{i=1}^{n} g(p_i) = -\frac{1}{2} \mu H_n^2(p_1, \dots, p_n).$$

Now, it is obvious that

$$\sum_{i=1}^{n} f(p_i) = H_n(p_1, \dots, p_n) - c\mu^2(GS)_n(p_1, \dots, p_n)$$

and 
$$\sum_{i=1}^{n} g(p_i) = -\mu(GS)_n(p_1, \dots, p_n).$$

Thus the mappings f and g are also related to Gini-Simpson index.

#### 4. Concluding Remarks

- (a) In this paper, the general solutions of the functional equation (1.4) have been discussed.
- (b) The stability of the functional equation (1.4) is yet to be discussed.

(c) In the theory of functional equations, once a functional equation is known, it is customary to discuss its Pexider-type generalizations. Some Pexider-type generalizations of (1.4) will be discussed in our subsequent work.

### References

- [1] J. Aczel, Lectures on Functional Equations and Their Applications, Academic Press, New York, 1966.
- [2] J. A. Baker, A functional equation from probability theory, Proc. Amer. Math. Soc. 121(3) (1994), 767–773.
- [3] R. Balasubrahmanyan and K. S. Lau, Functional Equations in Probability Theory, Academic Press, Boston-New York, 1991.
- [4] R. A. Gangolli and D. Ylvisaker, *Discrete Probability*, Harcourt, Brace & World, Inc., New York, 1967.
- [5] J. Havrda and F. Charvat, Quantification method of classification process, concept of structural  $\alpha$ -entropy, Kybernetika (Prague) **3**(1) (1967), 30–35.
- [6] R. G. Laha and E. Lukas, On a functional equation which occurs in a characterization problem, Aequationes Math. 15(2) (1977), 313–314.
- [7] I. Miller and M. Miller, Mathematical Statistics, Prentice Hall, 2004.
- [8] P. Nath and D. K. Singh, On a sum form functional equation related to entropies and some moments of a discrete random variable, Demonstratio Math. 42(1) (2009), 83–96.
- [9] C. R. Rao, Gini-Simpson index of diversity: A characterization, generalization and applications, Util. Math. 21 (1982), 273–282.
- [10] C. E. Shannon, A mathematical theory of communication, Bell Syst. Tech. J. 27 (1948), 379–423, 623–656.
- [11] P. Tan, Functional equations in mathematical statistics, Math. Mag. 45(4) (1972), 179–183.

<sup>1</sup>Department of Mathematics,

University of Delhi,

Delhi - 110007,

India

E-mail address: pnathmaths@gmail.com

<sup>2</sup>Department of Mathematics, Zakir Husain Delhi College (University of Delhi), Jawaharlal Nehru Marg, Delhi - 110002,

E-mail address: dhiraj426@rediffmail.com