Z. Wahrscheinlichkeitstheorie verw. Geb. 22, 95–100 (1972)

On Shannon's Entropy, Directed Divergence and Inaccuracy

Pl. Kannappan

1. Introduction

In this paper, we are dealing with two functional equations arising in information theory. The first one concern with Shannon's entropy and the second one with directed divergence or information gain and inaccuracy. The Shannon's entropy [12],

$$H(p_1, \dots, p_n) = -\sum_{i=1}^{n} p_i \log p_i, \tag{1}$$

where $0 \leqq p_i \leqq 1$, $\sum_{i=1}^{n} p_i = 1$ was extensively studied and was characterized by many authors. Detailed proofs can be found in the original papers [2, 3, 5, 8, 10] and a useful survey of known results in this field can be found in [1]. In particular, in [3], it was studied as a problem of obtaining the most general real valued continuous function f, which for all positive integers m, n satisfies the functional equation

$$\sum_{i=1}^{m} \sum_{j=1}^{n} f(x_i y_j) = \sum_{i=1}^{m} f(x_i) + \sum_{j=1}^{n} f(y_j), \tag{2}$$

where $x_i, y_j \geqq 0$, $\sum_{i=1}^{m} x_i = 1 = \sum_{j=1}^{n} y_j$. A new and simple proof of obtaining the solution of (2) is given in Section 2.

Another quantity called directed-divergence [8] or information gain [11]

$$I_n \begin{pmatrix} p_1, p_2, \dots, p_n \\ q_1, q_2, \dots, q_n \end{pmatrix} = \sum_{i=1}^{n} p_i \log \frac{p_i}{q_i}, \tag{3}$$

with $p_i, q_i \geqq 0$, $\sum_{i=1}^{n} p_i = 1 = \sum_{j=1}^{n} q_j$, was characterized among others in [5, 10]. Here a new characterization of (3) based on (2) is given through a functional equation in Section 3.

One other quantity known as inaccuracy [7],

$$H_n \begin{pmatrix} p_1, p_2, \dots, p_n \\ q_1, q_2, \dots, q_n \end{pmatrix} = -\sum_{i=1}^{n} p_i \log q_i \tag{4}$$

was characterized among others in [5]. This quantity (4) is characterized in Section 4 by the same functional equation appearing in Section 3 under different boundary conditions.

Remark 1. The usual convention with regard to (1), (3) and (4) are adhered to: there is a one-one correspondence between p_i's and q_i's given by their suffices; $q_i=0$ implies the corresponding $p_i=0$, $0\log 0=0$; $p_i \log \frac{p_i}{q_i}=p_i(\log p_i-\log q_i)$ whenever $p_i=0$, $q_i=0$; the logarithm is with respect to base 2.

2. Shannon's Entropy

Here we describe all the continuous solutions of (2) and prove the following theorem.

Theorem 1. *A necessary and sufficient condition that a continuous function f satisfies* (2) *is that,*

$$f(x)=A x \log x,$$

for all $x\in I=[0,1]$, where A is an arbitrary constant.

Proof. The sufficient part is a mere verification. To prove the necessary part, define a function

$$g(x)=x f\left(\frac{1}{x}\right), \qquad \text{for all real } x\geqq 1. \tag{5}$$

Evidently g is continuous. Let m, n be any integers $\geqq 1$. Then putting $x_i=1/m$ $(i=1,2,\ldots,m)$ and $y_j=1/n$ $(j=1,2,\ldots,n)$ in (2) and using (5), we get

$$g(mn)=g(m)+g(n), \qquad \text{for positive integers } m,n\geqq 1. \tag{6}$$

It is known that g in (6) has a unique extension to the multiplicative group of positive rational numbers. The question is whether this extension is the same as g in (5). It is indeed the case is shown as follows.

Taking any rational $r\in(0,1)$ as m/n $(m<n)$ and letting $x_1=m/n$, $x_2=\cdots=x_{n-m+1}=1/n$ and $y_1=\cdots=y_m=1/m$ in (2), we obtain (by taking m as $n-m+1$ and n as m)

$$m f\left(\frac{1}{n}\right)+(n-m)\, m f\left(\frac{1}{mn}\right)=f\left(\frac{m}{n}\right)+(n-m) f\left(\frac{1}{n}\right)+m f\left(\frac{1}{m}\right). \tag{7}$$

Using (5), (6) and (7), we find that

$$g\left(\frac{n}{m}\right)=g(n)-g(m), \qquad \text{for } m<n. \tag{8}$$

From (6) and (8) results that

$$g(xy)=g(x)+g(y), \qquad \text{for all rationals } x,y\geqq 1. \tag{9}$$

The continuity of g implies that (9) holds for all real $x,y\geqq 1$. From [2], it follows that g has a unique extension to all positive reals which is continuous and satisfies (9) for all real $x,y>0$. Thus $g(x)=-A\log x$, for some real constant A. This, together with (5) gives the required solution f of (2).

3. Directed-Divergence

Now, we consider a functional equation connected with directed-divergence (3). Let F be a real valued, continuous function, defined on

$$J=]0,1[\times]0,1[\cup\{(0,y)\}\cup\{(1,y')\}$$

with $y\in[0,1)$ and $y'\in(0,1]$ such that F, for all positive integers m and n satisfies the functional equation

$$\sum_{i=1}^{m}\sum_{j=1}^{n}F(x_i y_j, u_i v_j)=\sum_{i=1}^{m}F(x_i,u_i)+\sum_{j=1}^{n}F(y_j,v_j), \tag{10}$$

for $x_i, u_i, y_j, v_j\geqq 0$ and $\sum_{i=1}^{m}x_i=1=\sum_{j=1}^{m}y_j$, $\sum_{i=1}^{m}u_i\leqq 1$ and $\sum_{j=1}^{n}v_j\leqq 1$. Now, we determine all the solutions of (10) under the boundary conditions

$$F(1,\tfrac{1}{2})=1, \tag{11}$$

and

$$F(\tfrac{1}{2},\tfrac{1}{2})=0. \tag{12}$$

As a matter of fact, we prove the following theorem.

Theorem 2. *The most general continuous solution F of the functional equation* (10), *satisfying further* (11) *and* (12) *is given by* $F(x,y)=x\log\frac{x}{y}$, *for* $(x,y)\in J$.

Proof. Let us define a function G,

$$G(x,y)=xF\left(\frac{1}{x},\frac{1}{y}\right),\qquad \text{for all real } x,y\geqq 1, \tag{13}$$

(of course with the restriction that whenever $y=1$, then $x=1$). Obviously G is continuous. First, we note that, if we take $u_i=x_i$ $(i=1,2,\ldots,m)$ and $v_j=y_j$ $(j=1,2,\ldots,n)$ in (10), (10) is same as (2). Then from Theorem 1, it is easy to see that

$$G(x,x)=xF\left(\frac{1}{x},\frac{1}{x}\right)=-A\log x,\qquad \text{for all real } x\geqq 1.$$

With the help of (12), we get $A=0$ and hence

$$G(x,x)=xF\left(\frac{1}{x},\frac{1}{x}\right)=0,\qquad \text{for all real } x\geqq 1. \tag{14}$$

Let m,n,r,s be any integers for which $1\leqq m\leqq r$, $1\leqq n\leqq s$. Then setting $x_i=1/m$ $(i=1,2,\ldots,m)$, $y_j=1/n$ $(j=1,2,\ldots,n)$, $u_i=1/r$ $(i=1,2,\ldots,m)$ and $v_j=1/s$ $(j=1,2,\ldots,n)$ in (10) and utilizing (13), we have

$$G(mn,rs)=G(m,r)+G(n,s),\qquad \text{for } 1\leqq m\leqq r,\ 1\leqq n\leqq s. \tag{15}$$

For $m=1$ and $n=1$, (15) reduces to

$$G(1,rs)=G(1,r)+G(1,s),\qquad \text{for integers } r,s\geqq 1. \tag{16}$$

For any rational $x=p/q$ $(0<p\leqq q)$, allowing $x_1=1$, $y_1=1$, $u_1=p/q$ and $v_1=1/p$ in (10), we get

$$F\left(1,\frac{1}{q}\right)=F\left(1,\frac{p}{q}\right)+F\left(1,\frac{1}{p}\right), \tag{17}$$

from which and (13), that

$$G\left(1,\frac{q}{p}\right)=G(1,q)-G(1,p), \qquad \text{for } 0<p\leqq q. \tag{18}$$

Now (16) and (18) yield

$$G(1,xy)=G(1,x)+G(1,y), \qquad \text{for all rationals } x,y\geqq 1. \tag{19}$$

As before, since G is continuous, it follows immediately using (11) that

$$G(1,x)=\log x, \qquad \text{for all real } x\geqq 1. \tag{20}$$

Use (15) for $s=n$, $m=1$, (14) and (20), to get

$$G(n,rn)=\log r, \qquad \text{for all integers } r,n\geqq 1. \tag{21}$$

Now, (15) for $s=mn$ and (21) imply

$$G(mn,rmn)=G(m,r)+G(n,mn), \qquad \text{for } m\leqq r,$$

that is,

$$\log r=G(m,r)+\log m,$$

or

$$mF\left(\frac{1}{m},\frac{1}{r}\right)=G(m,r)=\log\frac{r}{m}, \qquad \text{for } 1\leqq m\leqq r \text{ (integers)}. \tag{22}$$

Choose any two rationals $x,y\in(0,1)$. Then for $x=m/n$ $(m<n)$, $y=p/q$ $(p<q)$, choose an integer k sufficiently large such that $kp\geqq m$, $kq\geqq n$ and $k\geqq q(n-m)/n(q-p)$. Now, putting $x_1=m/n$,

$$x_2=\cdots=x_{n-m+1}=\frac{1}{n}, \qquad y_1=\cdots=y_m=\frac{1}{m}, \qquad u_1=\frac{p}{q},$$

$$u_2=\cdots=u_{n-m+1}=\frac{1}{kn} \quad \text{and} \quad v_1=v_2=v_m=\frac{1}{pk} \quad \text{in (10)},$$

(10) can be rewritten as,

$$\begin{aligned} &mF\left(\frac{1}{n},\frac{1}{qk}\right)+(n-m)\,mF\left(\frac{1}{mn},\frac{1}{pnk^2}\right)\\ &\quad=F\left(\frac{m}{n},\frac{p}{q}\right)+(n-m)\,F\left(\frac{1}{n},\frac{1}{kn}\right)+mF\left(\frac{1}{m},\frac{1}{pk}\right). \end{aligned} \tag{23}$$

Thus, (22) and (23) yield

$$F\left(\frac{m}{n}, \frac{p}{q}\right) = \frac{m}{n} \log \frac{m q}{n p},$$

that is,

$$F(x, y) = x \log \frac{x}{y}, \qquad \text{for all rationals } x, y \in (0, 1). \tag{24}$$

Now, use the continuity of F to get (24) for all $(x, y) \in J$. This completes the proof of this theorem.

Remark 2. Theorem 2 is proved for the case $\sum_{i=1}^{m} u_i \leqq 1$ and $\sum_{j=1}^{n} v_j \leqq 1$. The same result can be obtained for the case $\sum_{i=1}^{m} u_i = 1$ and $\sum_{j=1}^{n} v_j = 1$, provided we assume that F in addition to (10), (11) and (12) also satisfies the condition $F(0, q) = 0$ for for $q \in [0, 1)$. This condition is needed to get the Eqs. (15), (16), (17) and (23).

4. Inaccuracy

In order to characterize the quantity inaccuracy given by (4), we use the functional equation (10) in Section 3, under the boundary conditions (11) and

$$F(\tfrac{1}{2}, \tfrac{1}{2}) = \tfrac{1}{2}, \tag{25}$$

and prove the following theorem.

Theorem 3. *If a real valued function F which is continuous in J satisfies the functional equation* (10) *and the conditions* (11) *and* (25), *then* $F(x, y) = -x \log y$.

Proof. The proof of this theorem runs parallel to that of Theorem 2. As in the proof of Theorem 2, we get

$$G(x, x) = x F\left(\frac{1}{x}, \frac{1}{x}\right) = -A \log x, \qquad \text{for all real } x \geqq 1,$$

which by using (25) gives $A = -1$ and hence,

$$G(x, x) = x F\left(\frac{1}{x}, \frac{1}{x}\right) = \log x, \qquad \text{for all real } x \geqq 1. \tag{26}$$

The Eqs. (15) and (20) also hold good in this case.

Now (15) with $s = n$, $m = 1$, (26) and (20) give

$$G(n, r n) = \log r n, \qquad \text{for all natural numbers } r \text{ and } n. \tag{27}$$

Thus (27) and (15) for $s = m n$ yield

$$m F\left(\frac{1}{m}, \frac{1}{r}\right) = G(m, r) = \log r, \qquad \text{for } 1 \leqq m \leqq r. \tag{28}$$

From (28) and (23), we find that

$$F\left(\frac{m}{n}, \frac{p}{q}\right) = \frac{m}{n} \log \frac{q}{p}, \quad \text{for } m<n, \quad p<q;$$

that is,

$$F(x, y) = -x \log y, \quad \text{for all rationals } x, y \in (0, 1). \tag{29}$$

With the help of continuity of F, it is easy to see that (29) is indeed true for all $(x, y) \in J$. The proof of the theorem is thus complete.

Remark 3. Theorem 3 remains true for the complete distributions $\sum_{i=1}^{m} u_i = 1$ and $\sum_{j=1}^{n} v_j = 1$ also, provided $F(0, q) = 0$ for $q \in [0, 1)$, in addition to F satisfying (10), (11) and (25). This condition is used to obtain (15), (16), (17) and (23).

I express my thanks to the referee for his useful remarks. As is pointed out by the referee, it is highly probable that the solutions of the functional equations (2) and (10) are unique among Lebesgue measurable functions.

References

1. Aczel, J.: On different characterizations of entropies, in Probability and Information theory, Proc. Internat. Symp. McMaster Univ., pp. 1–11. Berlin-Heidelberg-New York: Springer Lecture Notes in Math. No. **89**, 1969.
2. Aczel, J., Baker, J.A., Djokovic, D.Z., Kannappan, Pl., Rado, F.: Extensions of certain homomorphisms of subsemigroups to homomorphisms of groups. Aeq. Math. V 6, 2/3, 263–271 (1971).
3. Chaundy, T.W., Mcleod, J.B.: On a functional equation, Edin. Math. Notes **43**, 7–8 (1960).
4. Daroczy, Z.: Über die Charakterisierung der Shannonschen Entropie, Proc. Colloq. of Information Theory, **1**, János Bolyai Math. Soc., Budapest, 135–139 (1968).
5. Kannappan, Pl., Rathie, P.N.: On some axiomatic characterization of measures of inaccuracy and directed-divergence (to appear).
6. Kendall, D.G.: Functional Equation in Information Theory. Z. Wahrscheinlichkeitstheorie verw. Geb. **2**, 225–229 (1963).
7. Kerridge, D.F.: Inaccuracy and inference. J. Roy. Statist. Soc. B **23**, 184–194 (1961).
8. Kullback, S.: Information Theory and Statistics. New York: Wiley 1959.
9. Pintacuda, N.: Shannon entropy: A more general derivation. Statistica (Bologna) **26**, 509–524 (1965).
10. Rathie, P.N., Kannappan, Pl.: On a new characterization of directed divergence in information theory (to appear).
11. Rényi, A.: On measure of entropy and information. Proc. Fourth Berkeley Sympos. Statistics and Probability, 547–561 (1960).
12. Shannon, C.E.: A mathematical theory of communication. Bell System Tech. J. **27**, 379–423, 623–656 (1948).

Pl. Kannappan
Department of Applied Analysis
and Computer Science
Faculty of Mathematics
University of Waterloo
Waterloo, Ontario, Canada

(Received July 22, 1971)