ON SHARMA - TANEJA'S ENTROPY OF TYPE (α, β)

by Ishwar Singh, *Department of Mathematics, Govt. Degree College, Jaora, Dist. Ratlam (M.P.)*

(*Received* 19 *November* 1977)

Starting from a suitable set of axioms, which are modifications of those considered by Fadeev and later by Sharma-Taneja for studying Shannon's entropy and degree α-entropy, Sharma-Taneja's entropy of type (α, β) is characterized in this paper.

1. Introduction

Shannon's measure of entropy for a discrete probability distribution

$$P = (p_1, \ldots, p_n), p_i \geqslant 0, \sum_{i=1}^{n} p_i = 1,$$

given by

$$H(P) = - \sum_{i=1}^{n} p_i \log p_i$$

has been characterized in several ways (*see* Aczel 1975). Out of the many ways of characterization the two elegant approaches are to be found in the work of

(i) Fadeev (1956), who uses branching property viz.,

$$H^n(P_1, \ldots, P_n) = H^{n-1}(p_1 + p_2, p_3, \ldots, p_n) + (p_1 + p_2) \times H^2\left(\frac{p_1}{p_1 + p_2}, \frac{p_2}{p_1 + p_2}\right) \qquad \ldots(1)$$

$n = 3, 4, \ldots$ for the above distribution P, as the basic postulate, and

(ii) Chaundy and Mcleod (1960), who studied the functional equation

$$\sum_{i=1}^{n} \sum_{j=1}^{m} f(p_i q_j) = \sum_{i=1}^{n} f(p_i) + \sum_{j=1}^{m} f(q_j) \qquad \ldots(2)$$

for $p_i \geqslant 0, q_j \geqslant 0$.

Both the above mentioned approaches have been extensively exploited and generalized. The most general form of (2) has been studied by Sharma and Taneja (1974) who considered the functional equation

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

$$\sum_{i=1}^{n} p_i = \sum_{j=1}^{m} q_j = 1, \quad p_i \geqslant 0, \quad q_j \geqslant 0. \qquad \ldots(3)$$

In terms of real continuous solutions of (3) Sharma and Taneja (1974) obtained three new measures given by

$$H_l^{\alpha}(P) = -2^{\alpha-1} \sum_{i=1}^{n} p_i^{\alpha} \log p_i, \quad \alpha > 0 \qquad \ldots(4)$$

$$H_p^{(\alpha,\beta)}(P) = (2^{1-\alpha} - 2^{1-\beta})^{-1} \left(\sum_{i=1}^{n} p_i^{\alpha} - \sum_{i=1}^{n} p_i^{\beta} \right); \alpha \neq \beta; \alpha, \beta > 0 \quad \ldots(5)$$

and

$$H_s^{(\alpha,\beta)}(P) = -\frac{2^{\alpha-1}}{\sin \beta} \sum_{i=1}^{n} p_i^{\alpha} \sin(\beta \log p_i); \beta \neq 0, \alpha > 0. \qquad \ldots(6)$$

Our aim in this paper is to consider a suitable branching property for characterizing a measure which gives (5). The characterization theorem is given in the next section.

2. Characterization Theorem

Given a discrete probability distribution $P = (p_1, \ldots, p_n); p_i \geqslant 0, \sum_{i=1}^{n} p_i = 1$, let us assume that there is a bi-parametric measure $H(P, \alpha, \beta)$ which satisfies the following set of axioms.

(i) $H_n(p_1, \ldots, p_n; \alpha, \beta)$ is continuous in the region $p_1 \geqslant 0$,

$$\sum_{i=1}^{n} p_i = 1, \alpha, \beta > 0, \alpha \neq \beta,$$

(ii) $H_{n+1}(p_1, \ldots, p_{i-1}, v_{i_1}, v_{i_2}; p_{i+1}, \ldots, p_n; \alpha, \beta)$

$$= H_n(p_1, \ldots, p_{i-1}, p_i, p_{i+1}, \ldots, p_n; \alpha, \beta) + \mu(v_{i_1}^{\alpha} + v_{i_2}^{\alpha} - p_i^{\alpha})$$

$$+ \nu(v_{i_1}^{\beta} + v_{i_2}^{\beta} - p_i^{\beta});$$

for every $v_{i_1} + v_{i_2} = p_i > 0$, $i = 1, 2, \ldots, n$; where μ and ν are same constants.

(iii) $H_n(p_1, \ldots, p_{i-1}, 0, p_{i+1}, \ldots, p_n; \alpha, \beta) = H_{n-1}(p_1, \ldots, p_{i-1}, \ldots, p_n; \alpha, \beta)$ for every $i = 1, 2, \ldots, n$; and

(iv) $H_2(1, 0 ; \alpha, \beta) = 0 ; \alpha, \beta > 0.$

Theorem — If $\alpha \neq \beta, \alpha, \beta > 0$, then axioms (i) to (iv) determine a measure given by

$$H_n(p_1, \ldots, p_n ; \alpha, \beta) = \mu(\sum_{i=1}^{n} p_i^{\alpha} - 1) + \nu(\sum_{i=1}^{n} p_i^{\beta} - 1). \quad \ldots(7)$$

Before proving the theorem we prove some intermediary results based on the above set of axioms in the lemmas below.

Lemma 1 — If $v_k \geqslant 0 ; k = 1, 2, \ldots, m ; \sum_{k=1}^{m} v_k = p_i > 0$ then

$$H_{n+m-1}(p_1, \ldots, p_{i-1}, v_1, v_2, \ldots, v_m ; p_{i+1}, \ldots, p_n ; \alpha, \beta)$$
$$= H_n(p_1, \ldots, p_n, \alpha, \beta) + \mu(v_1^{\alpha} + v_2^{\alpha} + \ldots + v_m^{\alpha} - p_i^{\alpha})$$
$$+ \nu(v_1^{\beta} + v_2^{\beta} + \ldots + v_m^{\beta} - p_i^{\beta}). \quad \ldots(8)$$

PROOF : To prove the lemma, we proceed by induction. From axiom (ii) the statement holds for $m = 2$. Let us suppose the result is true for less than or equal to m. We shall prove it for $m + 1$. For which we have that

$$H_{n+m}(p_1, \ldots, p_{i-1}, v_1, v_2, \ldots, v_{m+1}, p_{i+1}, \ldots, p_n ; \alpha, \beta)$$
$$= H_{n+1}(p_1, \ldots, p_{i-1}, v_1, u, p_{i+1}, \ldots, p_n, \alpha, \beta)$$
$$+ \mu(v_2^{\alpha} + \ldots + v_{m+1}^{\alpha} - u^{\alpha}) + \nu(v_2^{\beta} + \ldots + v_{m+1}^{\beta} - u^{\beta}) \quad \ldots(9)$$

where

$$u = v_2 + \ldots + v_{m+1}$$
$$= H_n(p_1, \ldots, p_n; \alpha, \beta) + \mu(v_1^{\alpha} + u^{\alpha} - p_i^{\alpha}) + \nu(v_1^{\beta} + u^{\beta} - p_i^{\beta})$$
$$+ \mu(v_2^{\alpha} + \ldots + v_{m+1}^{\alpha} - u^{\alpha}) + \nu(v_2^{\beta} + \ldots + v_{m+1}^{\beta} - u^{\beta})$$
$$= H_n(p_1, \ldots, p_n ; \alpha, \beta) + \mu(v_1^{\alpha} + \ldots + v_{m+1}^{\alpha} - p_i^{\alpha})$$
$$+ \nu(v_1^{\beta} + \ldots + v_{m+1}^{\beta} - p_i^{\beta}). \quad \ldots(10)$$

Hence it is true for any value of m since it is true for $m + 1$.

Lemma 2 — If $v_{ij} \geqslant 0, j = 1, 2, \ldots, m_i, \sum_{j=1}^{m_i} v_{ij} = p_i > 0, i = 1, 2, \ldots, n,$

$\sum_{i=1}^{n} p_i = 1$, then

$$H_{m_1+m_2+\cdots+m_n}(v_{1_1}, v_{1_2}, \ldots, v_{1_{m_1}}, \ldots, v_{n_1}, v_{n_2}, \ldots, v_{n_{m_n}} ; \alpha, \beta)$$

$$= H_n(p_1, \ldots, p_n ; \alpha, \beta) + \mu \sum_{i=1}^{n} (v_{i_1}^{\alpha} + \ldots + v_{i_{m_i}}^{\alpha} - p_i^{\alpha})$$

$$+ \nu \sum_{i=1}^{n} (v_{i_1}^{\beta} + v_{i_2}^{\beta} + \ldots + v_{i_{m_i}}^{\beta} - p_i^{\beta}). \quad \ldots(11)$$

This follows by repeated application of above lemma.

Lemma 3 — If

$$F(n ; \alpha, \beta) = H_n\left(\frac{1}{n}, \ldots, \frac{1}{n} ; \alpha, \beta\right) \quad \ldots(12)$$

then

$$F(n ; \alpha, \beta) = \mu(n^{1-\alpha} - 1) + \nu(n^{1-\beta} - 1) \text{ where } \alpha, \beta \neq 1. \quad \ldots(13)$$

PROOF: Replace in Lemma 2, m_i by m and $v_{ij} = 1/mn$, $i = 1, 2, \ldots, n$; $j = 1, 2, \ldots, m$; where m and n are positive integers, we then have

$$F(mn ; \alpha, \beta) = F(n ; \alpha, \beta) + \mu n^{1-\alpha}(m^{1-\alpha} - 1) + \nu n^{1-\beta}(m^{1-\beta} - 1). \quad \ldots(14)$$

Also

$$F(mn; \alpha, \beta) = F(m ; \alpha, \beta) + \mu m^{1-\alpha}(n^{1-\alpha} - 1) + \nu m^{1-\beta}(n^{1-\beta} - 1). \quad \ldots(15)$$

Putting $m = 1$ in (15) and using $F(1 ; \alpha, \beta) = 0$ (axiom iv), we get

$$F(n ; \alpha, \beta) = \mu(n^{1-\alpha} - 1) + \nu(n^{1-\beta} - 1) \quad \alpha, \beta \neq 1 \quad \ldots(16)$$

which is (15).

Proof of the Theorem — We first prove the theorem when p_i's are rational numbers. On account of axiom (i) the result will then hold for real p_i's. Let m and γ_i's be positive integers such that $\sum_{i=1}^{n} \gamma_i = m$. Let us put $p_i = \gamma_i/m$, where $i = 1, 2, \ldots, n$, then an application of Lemma 2 gives

$$H_m\Big(\underbrace{\frac{1}{m}, \ldots, \frac{1}{m}}_{\gamma_1} ; \ldots ; \underbrace{\frac{1}{m}, \ldots, \frac{1}{m}}_{\gamma_m} ; \alpha, \beta\Big) =$$

(*equation continued on p.* 568)

$$= H_n(p_1, \ldots, p_n ; \alpha, \beta) + \mu \sum_{i=1}^{n} \Big\{ \underbrace{\frac{1}{m^\alpha} + \ldots + \frac{1}{m^\alpha}}_{\gamma_i} - \Big(\frac{\gamma_i}{m}\Big)^\alpha \Big\}$$

$$+ \nu \sum_{i=1}^{n} \Big\{ \underbrace{\frac{1}{m^\beta} + \ldots + \frac{1}{m^\beta}}_{\gamma_i} - \Big(\frac{\gamma_i}{m}\Big)^\beta \Big\}.$$

By Lemma 3 we have

$$F(m ; \alpha, \beta) = H_n(p_1, \ldots, p_n ; \alpha, \beta) + \mu \sum_{i=1}^{n} \Big\{ \frac{\gamma_i}{m^\alpha} - \frac{\gamma_i^\alpha}{m^\alpha} \Big\}$$

$$+ \nu \sum_{i=1}^{n} \Big\{ \frac{\gamma_i}{m^\beta} - \frac{\gamma_i^\beta}{m^\beta} \Big\},$$

or

$$H_n(p_1, \ldots, p_n ; \alpha, \beta)$$

$$= \mu(m^{1-\alpha} - 1) + \nu(m^{1-\beta} - 1) - \mu \cdot \frac{1}{m^{\alpha-1}} \sum_{i=1}^{n} p_i$$

$$+ \mu \sum_{i=1}^{n} p_i^\alpha - \nu \frac{1}{m^{\beta-1}} \sum_{i=1}^{n} p_i + \nu \sum_{i=1}^{n} p_i^\beta .$$

$$= \mu(m^{1-\alpha} - 1) + \nu(m^{1-\beta} - 1) - \mu m^{1-\alpha} + \mu \sum_{i=1}^{n} p_i^\alpha$$

$$- \nu m^{1-\beta} + \nu \sum_{i=1}^{n} p_i^\beta$$

$$= \mu \sum_{i=1}^{n} p_i^\alpha + \nu \sum_{i=1}^{n} p_i^\beta - \mu - \nu.$$

$$H_n(p_1, \ldots, p_n ; \alpha, \beta)$$

$$= \mu \Big(\sum_{i=1}^{n} p_i^\alpha - 1 \Big) + \nu \Big(\sum_{i=1}^{n} p_i^\beta - 1 \Big) \qquad \ldots(17)$$

which is (16). This completes the proof of the theorem.

Remarks : It is customary to take that a unit of information obtained from the measure when the distribution is $(\frac{1}{2}, \frac{1}{2})$. So we may have an additional requirement given by

$$H(\tfrac{1}{2}, \tfrac{1}{2} ; \alpha, \beta) = 1 \qquad \ldots(18)$$

then from (17) we get

$$1 = \mu\,[2^{1-\alpha} - 1] + \nu\,[2^{1-\beta} - 1].$$

The type (α, β) measure in eqn. (5) now follows if we make an assumption $\mu + \nu = 0$ so that in that case

$$\mu = -\nu = (2^{1-\alpha} - 2^{1-\beta})^{-1}$$

and (17) reduces to $H_p^{(\alpha,\beta)}(P)$.

References

Aczel, J., and Daroczy, Z. (1975). Measures of Information and their Characterizations. Academic Press, New York.

Chaundy, T. W., and Mcleod, J. B. (1960). On a functional equation. *Proc. Edinb. math. Soc., Edinb. Math. Notes*, **43**, 7-8.

Fadeev, D. K. (1956). On the concept of entropy of a finite probabilistic scheme [in Russian]. *Uspekhi Math. Nauk.*, **11** (1(67)), 227-31.

Sharma, B. D., and Taneja, I. J. (1974). Three new measures of entropy. To appear in *E.I.K. (Germany)*. Abstract in *Proceedings of the Indian Science Congress Association held at Nagpur*, Jan. 3-7, 1974.