MOMENTS OF AN M.L.E. OF ENTROPY AND 'USEFUL' INFORMATION OF DEGREE β*

by BHU DEV SHARMA *and* MAN MOHAN, *Faculty of Mathematics, University of Delhi, Delhi* 110007

(*Received* 14 *March* 1978)

The paper contains exact expressions for the first two moments of Havrda-Charvat entropy of degree β. Further if an utility scheme is also associated with the probability scheme, then one gets generalized 'useful' information of degree β. Exact expressions for the first two moments of this measure are also derived.

1. INTRODUCTION

Let a utility distribution $\mathcal{U} = (u_1, u_2, ..., u_k)$ be attached with a probability distribution $\mathcal{P} = (p_1, p_2, ..., p_k)$, where $u_i > 0$ is the utility of the ith event which occurs with probability $p_i > 0$, $i = 1, 2, ..., k$ ($\sum_{i=1}^{k} p_i = 1$). The Belis-Guiaşu (1968) measure of 'useful' information (refer also Longo 1972, Sharma *et al.* 1978) is given by

$$I(\mathcal{U}; \mathcal{P}) = -\sum_{i=1}^{k} u_i p_i \log p_i. \qquad ...(1.1)$$

This measure reduces to Shannon (1948) entropy if we take each $u_i = 1$. Recently, Sharma *et al.* (1978) and Emptoz (1976) have characterized a non-additive generalization of $I(\mathcal{U}; \mathcal{P})$, called "generalized 'useful' information of degree β", given by

$$I^{\beta}(\mathcal{U}; \mathcal{P}) = (2^{1-\beta} - 1)^{-1} \sum_{i=1}^{k} u_i(p_i^{\beta-1} - 1), \quad \beta > 0. \qquad ...(1.2)$$

Again, when the utility considerations are ignored by taking $u_i = 1$ for each i, $I^{\beta}(\mathcal{U}; \mathcal{P})$ reduces to the Havrda-Charvat (1967) entropy,

$$I^{\beta}(\mathcal{P}) = (2^{1-\beta} - 1)^{-1} (\sum_{i=1}^{k} p_i^{\beta} - 1), \quad \beta > 0. \qquad ...(1.3)$$

Moments of the statistical estimates of Shannon entropy and $I(\mathcal{U}; \mathcal{P})$ have been studied by several authors. The work in this direction started with the papers

*This work was supported by a Teacher Fellowship awarded to the second author by the University Grants Commission, New Delhi, India. This paper was presented at the second conference of the Forum for Interdisciplinary Mathematics, held at Jaipur, July 21-22, 1978.

of Miller (1955) and Basharin (1959), who derived asymptotic mean and variance of an M.L.E. of Shannon entropy. Exact expressions for these moments have been obtained by Rogers and Green (1955) and Hutcheson and Shenton (1974). In a recent work, Sharma *et al.* (1977) derived asymptotic mean and variance of M.L E. of $I(\mathcal{U}; \mathcal{P})$. Exact expressions for these moments have recently been obtained by Sharma and Mohan (1978a, 1978b).

Applications of $I^{\beta}(\mathcal{P})$ and $I^{\beta}(\mathcal{U}; \mathcal{P})$ have been made recently in the theory of questionnaires (refer Picard 1972) and in the analysis of Business and Accounting Data (refer Sharma *et al.* 1976, 1978). Thus there arose a need of studying statistical estimates of $I^{\beta}(\mathcal{P})$ and $I^{\beta}(\mathcal{U}; \mathcal{P})$.

Sections 2 and 3 of this communication contain first two moments of the M.L.E. of $I^{\beta}(\mathcal{P})$. The results are extended in Section 4 to obtain first two moments of the M.L.E. of $I^{\beta}(\mathcal{U}; \mathcal{P})$.

2. First Moment of the M.L.E. of $I^{\beta}(\mathcal{P})$

Let n_i, $i = 1, 2, \ldots, k$, be the frequency of the occurrence of the ith event in a random sample of size N, such that $\sum_{i=1}^{k} n_i = N$, then the M.L.E. of $I^{\beta}(\mathcal{P})$ is given by

$$\hat{I}^{\beta}(\mathcal{P}) = (2^{1-\beta} - 1)^{-1} \left(\sum_{i=1}^{k} \hat{p}_i^{\beta} - 1 \right), \quad \beta > 0 \qquad \ldots(2.1)$$

where $\hat{p}_i = n_i/N$ is the M.L.E. of p_i, $i = 1, 2, \ldots, k$, and the n_i's follow the multinomial distribution.

Consider the moment-generating function of this multinomial distribution, given by

$$M(t) = M(t_1, t_2, \ldots, t_k; n_1, n_2, \ldots, n_k) = (p_1 e^{t_1} + p_2 e^{t_2} + \ldots + p_k e^{t_k})^N \qquad \ldots(2.2)$$

where $\sum_{i=1}^{k} p_i = 1$ and $\sum_{i=1}^{k} n_i = N$.

The result for first moment $E[\hat{I}^{\beta}(\mathcal{P})]$ may now be stated as :

Theorem 1 — For an integer β (> 0),

$$E[\hat{I}^{\beta}(\mathcal{P})] = (2^{1-\beta} - 1)^{-1} \left[N^{-\beta} \sum_{i=1}^{k} (1 + p_i \Delta)^N 0^{\beta} - 1 \right] \qquad \ldots(2.3)$$

where Δ is the usual difference operator; and $\Delta^r 0^{\beta}$ is the rth difference of 0^{β}.

Proof : Letting $q_i = \sum_{\substack{r=1 \\ r \neq i}}^{k} p_r e^{t_r}$ in (2.2), we obtain

$$M(t) = (q_i + p_i e^{t_i})^N$$
$$= \sum_{s=0}^{N} \binom{N}{s} (p_i e^{t_i})^s \, q_i^{N-s}. \qquad \ldots(2.4)$$

Differentiating (2.4) β times w.r.t. t_i, we get

$$\frac{\partial^\beta M(t)}{\partial t_i^\beta} = \sum_{s=0}^{N} \binom{N}{s} p_i^s \, q_i^{N-s} \, s^\beta e^{t_i s}. \qquad \ldots(2.5)$$

Setting each $t_i = 0$ in (2.5) gives

$$E(n_i^\beta) = \sum_{s=0}^{N} \binom{N}{s} p_i^s \, (1 - p_i)^{N-s} \, . \, s^\beta \qquad \ldots(2.6)$$

$$= (1 - p_i)^N \sum_{s=0}^{N} p_i^s \, (1 - p_i)^{N-s} E^s 0^\beta$$

(E is the shift operator defined by $E^r u_x = u_{x+r}$)

$$= (1 - p_i)^N \left(1 + \frac{p_i}{1 - p_i} E \right)^N 0^\beta$$

$$= (1 + p_i \, \triangle)^N \, 0^\beta. \qquad \ldots(2.7)$$

Taking mathematical expectation on both sides of (2.1) and using (2.7) we now get the result.

3. The Second Moment of $\hat{I}^\beta(\mathscr{P})$

An exact expression for the second moment of $\hat{I}^\beta(\mathscr{P})$ can be obtained and we state the result in the following :

Theorem 2 — For an integer β (> 0), the second moment of $\hat{I}^\beta(\mathscr{P})$ is given by

$$E\,\{[\hat{I}^\beta(\mathscr{P})]^2\} = N^{-2\beta}(2^{1-\beta} - 1)^{-2} \sum_{i=1}^{k} \Bigg\{ (1 + p_i \, \triangle)^N \, 0^{2\beta} + \frac{N^{2\beta}}{k}$$
$$- 2N^\beta (1 + p_i \, \triangle)^N \, 0^\beta + \sum_{\substack{j=1 \\ j \neq i}}^{k} \sum_{s=0}^{N} \binom{N}{s}$$
$$\times \, p_i^s \, (1 - p_i - p_j + p_j E^s)^{N-s} \, 0^\beta \Bigg\} \qquad \ldots(3.1)$$

where $\triangle$ is the usual difference operator; and $\triangle^r 0^\beta$ is the rth difference of 0^β.

PROOF : We have

$$[\hat{I}^\beta(\mathscr{P})]^2 = N^{-2\beta}(2^{1-\beta} - 1)^{-2}\left[\sum_{i=1}^{k} n_i^\beta - N^\beta\right]^2$$

$$= N^{-2\beta}(2^{1-\beta} - 1)^{-2}\left[\sum_{i=1}^{k} n_i^{2\beta} + \sum_{\substack{i=1\\ i\neq j}}^{k} \sum_{j=1}^{k} n_i^\beta n_j^\beta + N^{2\beta} - 2N^\beta \sum_{i=1}^{k} n_i^\beta\right]. \qquad ...(3.2)$$

Taking expectation on both sides of (3.2), we get

$$E\,[I^\beta(\mathscr{P})]^2 = N^{-2\beta}(2^{1-\beta} - 1)^{-2}\,[\sum_{i=1}^{k} E(n_i^{2\beta}) + \sum_{\substack{i=1\\ i\neq j}}^{k} \sum_{j=1}^{k} E(n_i^\beta n_j^\beta) + N^{2\beta} - 2N^\beta \sum_{i=1}^{k} E(n_i^\beta)]. \qquad ...(3.3)$$

Now differentiating (2.5) β times w.r.t. $t_j\,(i \neq j)$, we have

$$\frac{\partial^{2\beta} M(t)}{\partial t_i^\beta\, \partial t_j^\beta} = \sum_{s=0}^{N} \binom{N}{s} p_i^s\, s^\beta e^{t_i s} \frac{\partial^\beta(q_i^{N-s}}{\partial t_j^\beta}$$

$$= \sum_{s=0}^{N} \binom{N}{s} p_i^s\, s^\beta e^{t_i s} \sum_{r=0}^{N-s} \binom{N-s}{r} q_j^{N-s-r}\, p_j^r\, r^\beta e^{t_j r} \qquad ...(3.4)$$

where $\quad q_j = \sum_{r=1,\, r\neq i,\, j}^{k} p_r\, e^{t_r}.$

Setting each $t_i = 0$ in (3.4) we get

$$E(n_i^\beta n_j^\beta) = \sum_{s=0}^{N} \binom{N}{s} p_i^s \sum_{r=0}^{N-s} \binom{N-s}{r} q_{j*}^{N-s-r}\, p_j^r\, (rs)^\beta \; (q_j^* = 1 - p_i - p_j)$$

$$= \sum_{s=0}^{N} \binom{N}{s} p_i^s\, (q_j^* + p_j E^s)^{N-s}\, 0^\beta$$

$$= \sum_{s=0}^{N} \binom{N}{s} p_i^s\, (1 - p_i - p_j + p_j E^s)^{N-s}\, 0^\beta. \qquad ...(3.5)$$

Now, (3.3) with (2.7) and (3.5) gives (3.1).

4. Moments of the M.L.E. of $I^{\beta}(\mathcal{U}; \mathcal{P})$

As explained earlier for $I^{\beta}(\mathcal{P})$, the M.L.E. of $I^{\beta}(\mathcal{U}; \mathcal{P})$ is given by

$$\hat{I}^{\beta}(\mathcal{U}; \mathcal{P}) = (2^{1-\beta} - 1)^{-1} \left[N^{-\beta} \sum_{i=1}^{k} u_i n_i^{\beta} - \frac{1}{N} \sum_{i=1}^{k} u_i n_i \right] \qquad \ldots(4.1)$$

where $n_1, n_2, \ldots, n_k, \sum_{i=1}^{k} n_i = N$ follow multinomial distribution.

Taking mathematical expectation on both sides of (4.1) we obtain

$$E[\hat{I}^{\beta}(\mathcal{U}; \mathcal{P})] = (2^{1-\beta} - 1)^{-1} [N^{-\beta} \sum_{i=1}^{k} u_i E(n_i^{\beta}) - \bar{u}] \qquad \ldots(4.2)$$

where $\bar{u} = \sum_{i=1}^{k} u_i p_i.$

Invoking (2.7) this gives

$$E[\hat{I}^{\beta}(\mathcal{U}; \mathcal{P})] = (2^{1-\beta} - 1)^{-1} [N^{-\beta} \sum_{i=1}^{k} u_i(1 + p_i \triangle)^N 0^{\beta} - \bar{u}]. \qquad \ldots(4.3)$$

Further, we have

$$[\hat{I}^{\beta}(\mathcal{U}; \mathcal{P})]^2 = N^{-2\beta}(2^{1-\beta} - 1)^{-2} \sum_{i=1}^{k} \left\{ u_i^2 n_i^{2\beta} + \sum_{\substack{j=1 \\ j \neq i}}^{k} u_i u_j n_i^{\beta} n_j^{\beta} + \frac{N^{2\beta}}{k} \bar{u}^2 - 2N^{\beta} \bar{u} u_i n_i^{\beta} \right\}. \qquad \ldots(4.4)$$

Taking expectation on both the sides and using (2.7) and (3.5), eqn. (4.4) becomes

$$E[\hat{I}^{\beta}(\mathcal{U}; \mathcal{P})]^2 = N^{-2\beta}(2^{1-\beta} - 1)^{-2} \sum_{i=1}^{k} \left\{ u_i^2 (1 + p_i \triangle)^N 0^{2\beta} + \frac{N^{2\beta}}{k} \bar{u}^2 - 2N^{\beta} \bar{u} u_i (1 + p_i \triangle)^N 0^{\beta} + \sum_{s=0}^{N} \binom{N}{s} \sum_{\substack{j=1 \\ j \neq i}}^{k} u_i u_j p_i^s (1 - p_i - p_j + p_j E^s)^{N-s} 0^{\beta} \right\}. \qquad \ldots(4.5)$$

Remark : Setting $u_i = 1$ for all i we observe that the results of (4.3) and (4.5) reduce to those of (2.3) and (3.1) respectively.

REFERENCES

Basharin, G. P. (1959). On a statistical estimate for the entropy of a sequence of independent random variables. *Theory Probab. Applic.*, **4**, 333-36.

Belis, M., and Guiaşu, S. (1968). A quantitative-qualitative measure of information in cybernetic systems. *IEEE Trans. Inform. Theory*, **14**, 593-94.

Emptoz, H. (1976). Information de type β intigrant un concept d'utilite. *C.R. Acad. Sci. Paris*, **282**A, 911-14.

Havrda, J., and Charvat, F. (1967). Quantification method of classification processes: Concepts of structural β entropy. *Kybernetika*, **3**, 30-35.

Hutcheson, K., and Shenton, L. R. (1974). Some moments of an estimate of Shannon's measure of information. *Communications in Statistics*, **3**, 89-94.

Longo, G. (1972). A Quantitative-Qualitative Measure of Information. Springer-Verlag, Berlin.

Miller, G. A. (1955). Note on the bias of information estimates, *'Information Theory in Psychology'*, *H. Questler ed.* The Free Press, Glencoe, pp. 95-100.

Picard, C. F. (1972). Graphes et Questionnaires. Gautheir-Villars, Paris.

Rogers, M. S., and Green, B. F. (1955). The moments of sample information when the alternatives are equally likely, *'Information Theory in Psychology'*, *H. Questler ed.* The Free Press, Glencoe, pp. 101-107.

Shannon, C. E. (1948). The mathematical theory of communications. *Bell System Tech. J.*, **27**, 379-423.

Sharma, B. D., Mitter, J., and Mohan, Man (1978). On measures of 'useful' information. To appear in *Inform. Control.*

Sharma, B. D., Mohan, Man, and Dev, J. (1976). Financial statements : An informational approach. *Indian J. Accounting*, **6**, 29-48.

Sharma, B. D., Mohan, Man, and Mitter, J. (1977). On the mean and variance of an estimate of 'useful' information. *Statistica*, **37**, 427-38.

Sharma, B. D., and Mohan, Man (1978a). An exact expression for the mean of an M.L.E. of 'useful' information. *Pure Appl. Math. Sci.*, **7**, 29-33.

——————— (1978b). Exact variance of an M.L.E. of 'useful' information. To appear in *Statistica.*