INFORMATION OF ORDER α AND TYPE β

By J. N. Kapur, F.A.Sc.

(Indian Institute of Technology, Kanpur)

Received February 18, 1967

Abstract

Information $I_\alpha^\beta(Q/P)$ of order α and type β is introduced and it is shown that for every fixed β, this information is a monotonic increasing function of α. It is also shown that information of order α and type 1 is non-negative when $\sum_{k=1}^{N} q_k = \sum_{k=1}^{N} p_k$, where $(q_1, q_2, \ldots, q_N)$ and $(p_1, p_2, \ldots, p_N)$ are generalised probability distributions for Q and P respectively.

1. Introduction

Let P denote the original unconditional distribution of a random variable ξ and let Q denote the conditional distribution of ξ under the condition that an event E, connected in some way with ξ, has taken place. The measure of the amount of information concerning the random variable contained in this observation is denoted by $I(Q/P)$. If

$$P(\xi = x_k) = p_k; \quad P(\xi = x_k \mid E) = q_k; \quad (k = 1, 2, \ldots, N) \tag{1}$$

then a possible measure of the amount of information in question is

$$I_1\left(\frac{Q}{P}\right) = \sum_{k=1}^{N} q_k \log_2 \frac{q_k}{p_k}. \tag{2}$$

Renyi[1] considered the problem of finding a suitable measure for generalised probability distributions for which

$$\sum_{k=1}^{N} p_k \leqslant 1; \quad \sum_{k=1}^{N} q_k \leqslant 1 \tag{3}$$

and obtained a more general measure for $I(Q/P)$, *viz.*,

$$I_\alpha\left(\frac{Q}{P}\right) = \frac{1}{\alpha - 1} \log_2 \frac{\sum_{k=1}^{N} \frac{q_k^{\alpha}}{p_k^{\alpha-1}}}{\sum_{k=1}^{N} q_k}, \quad (\alpha \neq 1). \tag{4}$$

When $\alpha \to 1$, we get as a limit

$$I_1\left(\frac{Q}{P}\right) = \frac{\sum_{k=1}^{N} q_k \log_2 \frac{q_k}{p_k}}{\sum_{k=1}^{N} q_k}. \tag{5}$$

For a complete distribution, (5) reduces to (2). $I_\alpha(Q/P)$ may be called the information of order α when the distribution P is replaced by the distribution Q.

In the present paper we shall study some of the properties of $I_\alpha(Q/P)$ and also extend these to the more general case of information of order α and type β. Renyi's information comes out to be a particular case of this when $\beta = 1$.

2. Renyi's Postulates for Information of Order α

(i) $I(Q/P)$ is unchanged if the elements of P and Q are arranged in the same way so that the one-to-one correspondence between them is unchanged.

(ii) If $P = (p_1, p_2, \ldots, p_N)$ and $Q = (q_1, q_2, \ldots, q_N)$ and $p_k \leqslant q_k$ for $k = 1, 2, \ldots, N$, then $I(Q/P) \geqslant 0$, while if $p_k \geqslant q_k$ for all k, then $I(Q/P) \leqslant 0$.

(iii) $I(\{1\} \mid \{2\}) = 1.$ (6)

(iv) If $I(Q_1/P_1)$ and $I(Q_2/P_2)$ are defined and $P = P_1{*}P_2$, $Q = Q_1{*}Q_2$ and the correspondences between the elements of P and Q is that induced by the correspondence between the elements of P_1 and Q_1 and those of P_2 and Q_2, then

$$I\left(\frac{Q}{P}\right) = I\left(\frac{Q_1}{P_1}\right) + I\left(\frac{Q_2}{P_2}\right). \tag{7}$$

(v) There exists a continuous and strictly increasing function $y = g(x)$ defined for all real x, such that if $I(Q_1/P_1)$ and $I(Q_2/P_2)$ are defined and

$$0 < W(P_1) + W(P_2) \leqslant 1 \quad \text{and} \quad 0 < W(Q_1) + W(Q_2) \leqslant 1, \tag{8}$$

and the correspondence between the elements of $P_1 U P_2$, $Q_1 U Q_2$ is that induced between the elements of P_1 and Q_1 and between those of P_2 and Q_2, then we have

$$g\left[I\left\{\frac{(Q_1 U Q_2)}{(P_1 U P_2)}\right\}\right]$$

$$= \frac{W(Q_1)\, g\left[I\left(\frac{Q_1}{P_1}\right)\right] + W(Q_2)\, g\left[I\left(\frac{Q_2}{P_2}\right)\right]}{W(Q_1) + W(Q_2)}. \qquad (9)$$

Under these postulates Renyi showed that $g(x)$ is either a linear or an exponential function and these functions give rise to (5) and (4) respectively.

3. Our Postulates for Information of Order α and Type β

The first four postulates are the same as those of Renyi. In the fifth postulate, the weights are replaced by

$$W_\beta(P) = p_1^\beta + p_2^\beta + \ldots + p_N^\beta; \quad W_\beta(Q) = q_1^\beta + q_2^\beta + \ldots + q_N^\beta, \qquad (10)$$

so that (8) and (9) are replaced respectively by

$$0 < W_\beta(P_1) + W_\beta(P_2) \leq 1 \quad \text{and} \quad 0 < W_\beta(Q_1) + W_\beta(Q_2) \leq 1 \qquad (11)$$

and

$$g\left[I\left\{(Q_1 U Q_2) \mid (P_1 U P_2)\right\}\right]$$

$$= \frac{W_\beta(Q_1)\, g\left[I\left(\frac{Q_1}{P_1}\right)\right] + W_\beta(Q_2)\, g\left[I\left(\frac{Q_2}{P_2}\right)\right]}{W_\beta(Q_1) + W_\beta(Q_2)}. \qquad (12)$$

The advantage of using $W_\beta(P)$ instead of $W(P)$ has been explained earlier.[2]

With this modification, and proceeding in the same way as Renyi,[1] we get

$$I_\alpha^\beta\left(\frac{Q}{P}\right) = \frac{1}{\alpha - 1} \log_2 \frac{\sum_{k=1}^{N} q_k^\beta \left(\frac{q_k}{p_k}\right)^{\alpha - 1}}{\sum_{k=1}^{N} q_k^\beta}; \quad (\alpha \neq 1) \qquad (13)$$

and

$$I_1^\beta\left(\frac{Q}{P}\right) = \frac{\sum_{k=1}^{N} q_k^\beta \log_2 \frac{q_k}{p_k}}{\sum_{k=1}^{N} q_k^\beta}. \qquad (14)$$

When $\beta = 1$, these reduce to Renyi's measure of information of order α and 1 respectively. Postulates (i), (ii), (iii) are easily verified. For verifying (iv), we have

$$I_\alpha^\beta\left(\frac{Q}{P}\right) = \frac{1}{\alpha-1}\log_2 \frac{\sum_{k=1}^{N}\sum_{j=1}^{M}(q_k^2)^\beta (q_j')^\beta \left(\frac{q_k^2}{p_k^2}\frac{q_j'}{p_j'}\right)^{\alpha-1}}{\sum_{k=1}^{N}\sum_{j=1}^{M}(q_k^2)^\beta (q_j')^\beta}$$

$$= \frac{1}{\alpha-1}\log_2 \frac{\sum_{k=1}^{N}(q_k^2)^\beta \left(\frac{q_k^2}{p_k^2}\right)^{\alpha-1} \sum_{j=1}^{M}(q_j')^\beta \left(\frac{q_j'}{p_j'}\right)^{\alpha-1}}{\sum_{k=1}^{N}(q_k^2)^\beta \sum_{j=1}^{M}(q_j')^\beta}$$

$$= I_\alpha^\beta\left(\frac{Q_2}{P_2}\right) + I_\alpha^\beta\left(\frac{Q_1}{P_1}\right). \tag{15}$$

Similarly it is verified that (12) will be satisfied. The postulates are also easy to verify in the limit when $\alpha \to 1$.

4. Non-Negative Character of $I_\alpha(Q/P)$ when $\alpha > 1$

According to postulate (ii), $I_\alpha(Q/P)$ can be both positive and negative. We show however that if

$$\sum_{k=1}^{N} p_k \leq \sum_{k=1}^{N} q_k \tag{16}$$

then

$$I_\alpha\left(\frac{Q}{P}\right) \geqslant 0. \tag{17}$$

This includes the case when P and Q are complete probability distributions.

We use Jensen's theorem which states[3]:

"Let $a_1, a_2, \ldots, a_N$ be arbitrary positive numbers, then if $\phi''(t)$ exists and $\phi(t)$ is convex in $\alpha \leqslant t \leqslant \beta$, then

$$\phi\left(\frac{a_1 t_1 + a_2 t_2 + \ldots + a_N t_N}{a_1 + a_2 + \ldots + a_N}\right) \leqslant \frac{a_1\phi(t_1) + \ldots + a_N\phi(t_N)}{a_1 + a_2 + \ldots + a_N}. \tag{18}$$

Further if $\phi(t)$ is continuous and convex, sign of equality can hold only if all t's are equal or $\phi(t)$ is linear.

The inequality is reversed if the function $\phi(t)$ is concave." We apply the theorem to the function

$$\phi(t) = \log x^{\alpha-1} \tag{19}$$

which is convex if $\alpha < 1$ and concave if $\alpha > 1$.

Case (*i*).—If $\alpha > 1$, this gives

$$\frac{\sum_{k=1}^{N} q_k \log \left(\frac{q_k}{p_k}\right)^{\alpha-1}}{\sum_{k=1}^{N} q_k} \leqslant \log \left(\frac{\sum_{k=1}^{N} q_k \left(\frac{q_k}{p_k}\right)^{\alpha-1}}{\sum_{k=1}^{N} q_k}\right)^{\alpha-1}$$

or

$$(\alpha-1) \frac{\sum_{k=1}^{N} q_k \log \frac{q_k}{p_k}}{\sum_{k=1}^{N} q_k} \leqslant (\alpha-1) \frac{\sum_{k=1}^{N} q_k \left(\frac{q_k}{p_k}\right)^{\alpha-1}}{\sum_{k=1}^{N} q_k} \tag{20}$$

or

$$I_\alpha \left(\frac{Q}{P}\right) \geqslant \frac{1}{\alpha-1} \frac{\sum_{k=1}^{N} q_k \log \frac{q_k}{p_k}}{\sum_{k=1}^{N} q_k}. \tag{21}$$

Again from the same theorem (since $\log x$ is concave),

$$\frac{\sum_{k=1}^{N} q_k \log \frac{p_k}{q_k}}{\sum_{k=1}^{N} q_k} \leqslant \log \left(\frac{\sum_{k=1}^{N} q_k \frac{p_k}{q_k}}{\sum_{k=1}^{N} q_k}\right) = \log \frac{\sum_{k=1}^{N} p_k}{\sum_{k=1}^{N} q_k} \leqslant 0, \tag{22}$$

since we are assuming

$$\sum_{k=1}^{N} p_k \leqslant \sum_{k=1}^{N} q_k.$$

From (21) and (22)

$$I_\alpha \left(\frac{Q}{P}\right) \geqslant 0 \text{ when } \alpha > 1. \tag{23}$$

The inequality sign holds when all q_k/p_k are equal. In this case $I_\alpha(Q/P) = 0$, *i.e.*, no information is obtained by the observation of E.

Case (*ii*).—The proof for $\alpha \leqslant 1$ will be given in the next section.

5. MONOTONIC CHARACTER OF $I_\alpha(Q/P)$

$$\frac{d}{d\alpha}\left[I_\alpha\left(\frac{Q}{P}\right)\right] = \frac{d}{d\alpha}\left[\frac{1}{\alpha-1}\log\frac{\sum_{k=1}^{N} q_k\left(\frac{q_k}{p_k}\right)^{\alpha-1}}{\sum_{k=1}^{N} q_k}\right]$$

$$= \frac{1}{\alpha-1}\,\frac{\sum_{k=1}^{N} q_k\left(\frac{q_k}{p_k}\right)^{\alpha-1}\log\left(\frac{q_k}{p_k}\right)}{\sum_{k=1}^{N} q_k\left(\frac{q_k}{p_k}\right)^{\alpha-1}} - \frac{1}{(\alpha-1)^2}\log\frac{\sum_{k=1}^{N} q_k\left(\frac{q_k}{p_k}\right)^{\alpha-1}}{\sum_{k=1}^{N} q_k}$$

$$= \frac{1}{(\alpha-1)^2}\left\{\frac{\sum_{k=1}^{N} q_k\left(\frac{q_k}{p_k}\right)^{\alpha-1}\log\left(\frac{q_k}{p_k}\right)^{\alpha-1}}{\sum_{k=1}^{N}\left(\frac{q_k}{p_k}\right)^{\alpha-1}} - \log\frac{\sum_{k=1}^{N} q_k\left(\frac{q_k}{p_k}\right)^{\alpha-1}}{\sum_{k=1}^{N} q_k}\right\}. \tag{24}$$

Applying Jensen's theorem to the convex function $x \log x$ and putting

$$a_k = q_k, \quad t_k = \left(\frac{q_k}{p_k}\right)^{\alpha-1},$$

we get

$$\frac{\sum_{k=1}^{N} q_k\left(\frac{q_k}{p_k}\right)^{\alpha-1}}{\sum_{k=1}^{N} q_k}\log\frac{\sum_{k=1}^{N} q_k\left(\frac{q_k}{p_k}\right)^{\alpha-1}}{\sum_{k=1}^{N} q_k} \leq \frac{\sum_{k=1}^{N} q_k\left(\frac{q_k}{p_k}\right)^{\alpha-1}\log\left(\frac{q_k}{p_k}\right)^{\alpha-1}}{\sum_{k=1}^{N} q_k}$$

or

$$\log\frac{\sum_{k=1}^{N} q_k\left(\frac{q_k}{p_k}\right)^{\alpha-1}}{\sum_{k=1}^{N} q_k} \leqslant \frac{\sum_{k=1}^{N} q_k\left(\frac{q_k}{p_k}\right)^{\alpha-1}\log\left(\frac{q_k}{p_k}\right)^{\alpha-1}}{\sum_{k=1}^{N} q_k\left(\frac{q_k}{p_k}\right)^{\alpha-1}}. \tag{25}$$

Thus

$$\frac{d}{d\alpha}\left[I_\alpha\left(\frac{Q}{P}\right)\right] \geqslant 0,$$

so that $I_\alpha(Q/P)$ is a monotonic increasing function of α for all α. Now

$$I_0\left(\frac{Q}{P}\right) = -\log_2\frac{\sum_{k=1}^{N} q_k\left(\frac{p_k}{q_k}\right)}{\sum_{k=1}^{N} q_k}$$

$$= -\log_2 \frac{\sum_{k=1}^{N} p_k}{\sum_{k=1}^{N} q_k} = \log_2 \frac{\sum_{k=1}^{N} q_k}{\sum_{k=1}^{N} p_k}. \tag{26}$$

If

$$\sum_{k=1}^{N} p_k \leqslant \sum_{k=1}^{N} q_k, \tag{27}$$

$$I_0\left(\frac{Q}{P}\right) \geqslant 0. \tag{28}$$

Since $I_\alpha(Q/P)$ is a monotonic increasing function of α, $I_\alpha(Q/P) \geqslant 0$ for all α. This completes the proof of the result of the last section for $\alpha \leqslant 1$.

6. The Case when $\sum_{k=1}^{N} p_k \neq \sum_{k=1}^{N} q_k$

It is obvious from (26) that if

$$\sum_{k=1}^{N} q_k > \sum_{k=1}^{N} p_k,$$

then $I_0(Q/P) > 0$ and as such the non-negative character of $I_\alpha(Q/P)$ is maintained for all α whenever

$$\sum_{k=1}^{N} q_k \geqslant \sum_{k=1}^{N} p_k.$$

If

$$\sum_{k=1}^{N} q_k < \sum_{k=1}^{N} p_k,$$

then $I_0(Q/P)$ is negative and there will be a range of values of α for which $I_\alpha(Q/P)$ will be negative.

To examine whether it will be positive somewhere, we consider the variation as $\alpha \to \infty$.

Since

$$\sum_{k=1}^{N} q_k < \sum_{k=1}^{N} p_k,$$

two possibilities arise.

(i) $q_k < p_k$ for all k. In this case

$$I_1\left(\frac{Q}{P}\right) = \frac{\sum_{k=1}^{N} q_k \log\left(\frac{q_k}{p_k}\right)}{\sum_{k=1}^{N} q_k}. \tag{29}$$

Since $q_k < p_k$ for all k, $\log q_k/p_k < 0$ for all k and as such

$$I_1\left(\frac{Q}{P}\right) < 0. \tag{30}$$

For $\alpha > 1$,

$$\left(\frac{q_k}{p_k}\right) < 1 \text{ implies} \left(\frac{q_k}{p_k}\right)^{\alpha-1} < 1$$

so that

$$\sum_{k=1}^{N} q_k \left(\frac{q_k}{p_k}\right)^{\alpha-1} < \sum_{k=1}^{N} q_k$$

and

$$\frac{1}{\alpha-1} \log_2 \frac{\sum_{k=1}^{N} q_k \left(\frac{q_k}{p_k}\right)^{\alpha-1}}{\sum_{k=1}^{N} q_k} < 0. \tag{31}$$

Thus $I_\alpha(Q/P)$ is negative for all α. This is of course one of the postulates for Renyi's information of order α.

(ii) $\sum_{k=1}^{N} q_k < \sum_{k=1}^{N} p_k$,

but at least for one k, $q_k > p_k$.

Let q_M/p_M be the largest of $\{q_k/p_k\}$, and let q_M be the largest of $\{q_k\}$,

$$\underset{\alpha\to\infty}{Lt} \frac{1}{\alpha-1} \log \frac{\sum_{k=1}^{N} q_k \left(\frac{q_k}{p_k}\right)^{\alpha-1}}{\sum_{k=1}^{N} q_k}$$

$$= \underset{\alpha \to \infty}{Lt} \frac{1}{\alpha - 1} \log \frac{q_M \left(\frac{q_M}{p_M}\right)^{\alpha - 1}}{q_M}$$

$$= \log \left(\frac{q_M}{p_M}\right). \tag{32}$$

Since $q_M/p_M > 1$, $I_\infty(Q/P) > 0$ and as such $I_\alpha(Q/P) > 0$ for some value of α, but if q_M is not the largest, even $I_\infty(Q/P)$ may be negative.

We can see the result in another way.

Let

$$\sum_{k=1}^{N} q_k = A, \quad \sum_{k=1}^{N} p_k = B$$

$$q_k' = \frac{q_k}{A}, \quad p_k' = \frac{p_k}{B},$$

so that

$$\sum_{k=1}^{N} q_k' = 1 = \sum_{k=1}^{N} p_k'$$

$$I_\alpha\left(\frac{Q}{P}\right) = \frac{1}{\alpha - 1} \log \frac{\sum_{k=1}^{N} Aq_k' \left(\frac{Aq_k'}{Bp_k'}\right)^{\alpha - 1}}{\sum_{k=1}^{N} Aq_k'}$$

$$= \frac{1}{\alpha - 1} \log \left(\frac{A}{B}\right)^{\alpha - 1} + \frac{1}{\alpha - 1} \log \frac{\sum_{k=1}^{N} q_k' \left(\frac{q_k'}{p_k'}\right)^{\alpha - 1}}{\sum_{k=1}^{N} q_k'}.$$

From what we proved in Section 5, the second term is positive. If $A \geqslant B$, the first term is also non-negative and as such $I_\alpha(Q/P)$ is always positive except when $p_k = q_k$. If $A < B$, the first term is negative and $I_\alpha(Q/P)$ may be positive or negative.

7. Monotonic Character of $I_\alpha^\beta(Q/P)$

For a fixed β, the proof of Section 5 applies except that here we replace a_k by q_k^β and so we have

$$\frac{d}{d\alpha}\left[I_\alpha^\beta\left(\frac{Q}{P}\right)\right] > 0$$

so that for every fixed value of β, the information of order α and type β is monotonic increasing. Again

$$I_0^{\beta}\left(\frac{Q}{P}\right) = -\log \frac{\sum_{k=1}^{N} q_k^{\beta}\left(\frac{p_k}{q_k}\right)}{\sum_{k=1}^{N} q_k^{\beta}}.$$

If all $p_k/q_k > 1$, this is negative; if all $p_k/q_k < 1$, this expression is positive, but if some are greater than and other less than unity, then this expression can be positive or negative. When

$$\sum_{k=1}^{N} p_k q_k^{\beta-1} < \sum_{k=1}^{N} q_k^{\beta},$$

the information of order α and type β will always be positive.

Again consider

$$I_1^{\beta}\left(\frac{Q}{P}\right) = \frac{\sum_{k=1}^{N} q_k^{\beta}\log\frac{q_k}{p_k}}{\sum_{k=1}^{N} q_k^{\beta}}$$

$$\frac{d}{d\beta}\left[I_1^{\beta}\left(\frac{Q}{P}\right)\right]$$

$$= \frac{1}{\left(\sum_{k=1}^{N} q_k^{\beta}\right)^2}\left\{\sum_{k=1}^{N} q_k^{\beta} \sum_{k=1}^{N} q_k^{\beta}(\log q_k)^2 - \left(\sum_{k=1}^{N} q_k^{\beta}\log q_k\right)^2\right.$$

$$- \sum_{k=1}^{N} q_k^{\beta} \sum_{k=1}^{N} q_k^{\beta}\log q_k \log p_k$$

$$\left. + \sum_{k=1}^{N} q_k^{\beta}\log p_k \sum_{k=1}^{N} q_k^{\beta}\log q_k\right\}.$$

The first two terms are $\geqslant 0$, but the last two terms can change sign.

If $\log p_k = A \log q_k$ and $A < 1$, then the whole expression is positive and the information is a monotonic increasing function of β. If $A > 1$, then this expression is negative and the information in a monotonic decreasing function of β. No completely general statement about the monotonic character of information of order 1 and type β can however be made.

References

1. Renyi, A. .. "On measures of information and entropy," *Proc. 4th Berkeley Symp. Math. Stat. and Prob.*, 1961, **1**, 547–61.
2. Kapur, J. N. .. "Generalised entropy of order α and type β," *The Maths. Seminar*, 1967, **4**.
3. Phillips, E. G. .. *Analysis*, 1948, Chapter VI.