78. *Information and Statistics. I*

By Yukiyosi KAWADA
Department of Mathematics, Faculty of Science,
University of Tokyo

(Communicated by Shokichi IYANAGA, M. J. A., Sept. 14, 1987)

This short note is a summary of an axiomatic consideration of an information and its applications to statistics.[1)]

I. Axioms of an information **1.** Under an information we usually understand the *Kullback-Leibler information* (1951):

$$(1) \qquad I_{KL}(\boldsymbol{p}, \boldsymbol{q}) = \sum_{k=1}^{m} p_k \log (p_k / q_k)$$

where $\boldsymbol{p}=(p_1, \cdots, p_m)$, $\boldsymbol{q}=(q_1, \cdots, q_m)$ are two finite probability distributions. There are, however, similar known functions $I(\boldsymbol{p}, \boldsymbol{q})$ which may be also called informations. For example,

$$(2) \qquad I_P(\boldsymbol{p}, \boldsymbol{q}) = \left(\sum_{k=1}^{m} p_k^2 q_k^{-1}\right) - 1 \quad (\textit{Pearson's information}, 1900)$$

which can be expressed as $(\sum_{k=1}^{m} (n_k - nq_k)^2 / nq_k)/n$ when $\boldsymbol{p}=(n_1/n, \cdots, n_m/n)$ $(n=n_1+\cdots+n_m)$, and

$$(3) \qquad I_K(\boldsymbol{p}, \boldsymbol{q}) = 2\left(1 - \sum_{k=1}^{m} p_k^{1/2} q_k^{1/2}\right) \quad (\textit{Kakutani's information}, [5], 1948).$$

These are included as special cases of the family

$$(4) \qquad I^{\lambda}(\boldsymbol{p}, \boldsymbol{q}) = \frac{1}{\lambda}\left\{\left(\sum_{k=1}^{m} p_k^{1+\lambda} q_k^{-\lambda}\right) - 1\right\}, \quad -\frac{1}{2} \leqq \lambda < \infty, \quad \lambda \neq 0,$$

namely, $I_P = I^1$, $I_K = I^{-1/2}$, and we define $I^0 = I_{KL}$.

We can easily see that

$$I^{\lambda}(\boldsymbol{p}, \boldsymbol{q}) \leqq I^{\mu}(\boldsymbol{p}, \boldsymbol{q}) \qquad \text{for } \lambda < \mu$$

and

$$\lim_{n\to\infty} I^{\lambda_n}(\boldsymbol{p}, \boldsymbol{q}) = I^{\lambda_0}(\boldsymbol{p}, \boldsymbol{q}) \qquad \text{for } \lim_{n\to\infty} \lambda_n = \lambda_0.$$

We call I^0 the *parabolic* information, I^{λ} $(\lambda>0)$ a *hyperbolic* information and $I^{-\mu}(1/2 \geqq \mu > 0)$ an *elliptic* information.

Remark. (i) The function $I^{-\mu}(\boldsymbol{p}, \boldsymbol{q})$ for $1/2 \geqq \mu > 0$ was introduced by several authors [4], [7] and the general case was also considered in [9].

(ii) In the definition (4) we can extend the value λ for $\lambda < -1/2$ formally. Then we have

$$I^{-\lambda}(\boldsymbol{p}, \boldsymbol{q}) = -\frac{\lambda-1}{\lambda} I^{\lambda-1}(\boldsymbol{q}, \boldsymbol{p}), \quad \lambda > 1$$

$$I^{-1}(\boldsymbol{p}, \boldsymbol{q}) = 0$$

$$I^{-\mu}(\boldsymbol{p}, \boldsymbol{q}) = \frac{1-\mu}{\mu} I^{\mu-1}(\boldsymbol{q}, \boldsymbol{p}), \quad 1/2 < \mu < 1.$$

1) The details will be published in the Proceedings of the Institute of Statistical Mathematics (Tōkei Sūri) in Japanese.