

Note on information measurement

TARALD O. KVÅLSETH

University of Minnesota, Minneapolis, Minnesota

The information measures by Hartley (1928) and Shannon (1948) have been extensively used in human performance studies. An alternative and most general information measure is briefly introduced in this paper, as a potentially useful measure in behavioral research. This new generalized measure contains a number of arbitrary parameters, making it a parameterized family of measures whose members include Hartley's and Shannon's measures and various others. Some of the well-known measures are shown to be special cases of the generalized measure, called the (α, B, δ) information measure. Some of its properties are also pointed out.

For measuring human performance in a variety of task situations, and for analyzing the capabilities and limitations of various basic human processes, certain measures from information theory have been extensively used (see, e. g., Sheridan & Ferrell, 1974; Wickens, 1984). In particular, the entropy (information) measures of Hartley (1928) and Shannon (1948) have been widely used. For a set of n random events, these measures are, respectively, defined by

$$I(n) = \log n \quad (1)$$

and

$$I(P) = -\sum p_i \log p_i, \quad (2)$$

where $P = (p_1, \dots, p_n)$ is the probability distribution of n events; the logarithm is to the base 2 as is customary in information theory (and hence the unit of measurement is bits); and Σ stands for the summation from $i = 1$ to $i = n$ as used throughout this paper. Of course, if all n events have equal probabilities, that is, if $p_1 = \dots = p_n = 1/n$, then Equations 1 and 2 are equivalent.

In behavioral research, the random events are frequently considered to be occurrences of certain stimuli, with p_i being the probability that the i th stimulus will occur during any replication of the experiment. The words "uncertainty," "surprise," and "information" are often used interchangeably, with the entropy $I(P)$ in Equation 2 considered as a measure of their amount or extent. Thus, $I(P)$ is used as a measure of the average amount of (a) uncertainty about the set of events before any one of the events occurs, (b) surprise as one of the events occurs, and (c) information gained after an event has occurred.

The use of such measurements in the behavioral sciences appears to be confined to the measures in Equations 1 and 2. There are, however, other information measures that may be equally appropriate for such applications. The purpose of this communication is to identify some of the best-known alternative measures, and in par-

ticular, to define a new generalized information measure that includes most other proposed measures as special cases.

GENERALIZED INFORMATION MEASURE

A number of alternative measures have been proposed as parameterized generalizations of Shannon's entropy. These have generally been developed axiomatically. That is, starting with a set of axioms or propositions that seem to be reasonable and desirable properties of a measure of information, it is then proved mathematically that there is one and only one function that satisfies those given axioms. Thus, each set of axioms yields its unique measure. For reviews of such developments, the interested reader may refer to Aczél and Daróczy (1975), Kapur (1983), and Mathai and Rathie (1974) (see also Klir & Folger, 1988, especially concerning Equations 1 and 2).

By means of such axiomatic characterization, a new measure of information has recently been formulated by this author, which appears to be the most general measure developed to date. While the detailed mathematical developments and proofs of properties are better suited for a different journal publication, this communication will simply present the expression for the measure, a few of its important properties, and some of its special cases. The measure may be defined as follows:

$$I_B^{\alpha\delta}(P) = \lambda \left[(\sum p_i^{\alpha-1+\beta_i} / \sum p_i^{\beta_i})^{(\delta-1)/(\alpha-1)} - 1 \right], \quad (3a)$$

with

$$\lambda = (2^{1-\delta} - 1)^{-1} \quad (3b)$$

and where α , δ , and $B = (\beta_1, \dots, \beta_n)$ are arbitrary parameters subject only to the restrictions that α and δ are different from 1, δ is nonnegative, and the β_i and $\alpha - 1 + \beta_i$ are positive for all i . This measure also applies to the so-called generalized distribution P , that is, when the sum of the probabilities p_1, \dots, p_n is less than or equal to unity, although in human performance experiments this sum is typically equal to 1.

Correspondence may be addressed to Tarald O. Kvålseth, Department of Mechanical Engineering, University of Minnesota, Minneapolis, MN 55455.

It is customary in information theory to use a normalization such that an information measure has the value of 1 for the case of two events with equal probabilities ($p_1 = p_2 = 1/2$). The expression for λ in Equation 3b comes from this normalization condition. However, such normalization is not strictly necessary, and λ in Equation 3a may be treated as an arbitrary parameter for which we may select any real value as long as $\lambda \geq 1$ for $0 \leq \delta < 1$ and $\lambda < 0$ for $\delta > 1$. In this case, we should use $I_{B\lambda}^{\alpha\delta}(P)$ to denote the measure in Equation 3a.

The new measure may be called the (α, B, δ) information measure, or (α, B, δ) entropy measure, and the $(\alpha, B, \delta, \lambda)$ information measure if λ is treated as a separate arbitrary parameter (i.e., if normalization is not considered). According to this measure, information cannot be negative. It is equal to zero when one of the p_i values equals 1 (and the other probabilities are all 0). It achieves its maximum value when all probabilities are equal ($p_1 = \dots = p_n = 1/n$). These are also the conditions under which Shannon's measure in Equation 2 achieves its limiting values. While the maximum value of Shannon's entropy is $\log n$, the maximum value of the new measure is $\lambda(n^{1-\delta}-1)$ or $(2^{1-\delta}-1)^{-1}(n^{1-\delta}-1)$ for the λ in Equation 3b. Other properties of the generalized measure include continuity and expansibility. That is, the measure is a continuous function of the probabilities $p_i, i = 1, \dots, n$; and if a component with zero probability is added to the distribution P , the value of the information measure remains unchanged. Also, when the β_i parameters are all equal ($\beta_1 = \dots = \beta_n = \beta$, where $\beta > 0$), the measure is symmetric in the p_i , that is, the measure is invariant with respect to any permutations of p_i, \dots, p_n .

Let us now consider some of the special cases of the (α, B, δ) information measure. If we choose $\delta = \alpha \neq 1$ and $\beta_1 = \dots = \beta_n = 1$, then Equations 3a and 3b reduce to, for $\Sigma p_i = 1$, the expression

$$I_1^{\alpha\alpha}(P) = (2^{1-\alpha} - 1)^{-1}(\Sigma p_i^\alpha - 1), \tag{4}$$

which is the Havrda-Charvat-Daróczy information measure (Daróczy, 1970; Havrda & Charvat, 1967). For $\beta_1 = \dots = \beta_n = \beta, \alpha \neq 1$, and when we consider the limit of $I_B^{\alpha\delta}(P)$ in Equations 3a and 3b as δ approaches 1, and using the Bernoulli-1'Hospital theorem, we find that

$$\lim_{\delta \rightarrow 1} I_B^{\alpha\delta}(P) = I_B^{\alpha 1}(P) = (1-\alpha)^{-1} \log(\Sigma p_i^{\alpha-1+\beta} / \Sigma p_i^\beta), \tag{5}$$

which is the information measure proposed by Kapur (1967). With $\beta=1$ in Equation 5 and $\Sigma p_i=1$, we get

$$I_1^{\alpha 1}(P) = (1-\alpha)^{-1} \log \Sigma p_i^\alpha, \tag{6}$$

which is the well-known measure by Rényi (1961). The fact that Shannon's measure in Equation 2 is also a special case of $I_B^{\alpha\delta}(P)$ is seen from Equation 6 when one takes the limit as α goes to 1. Thus, for $\beta_1 = \dots = \beta_n = 1$ and $\Sigma p_i=1$ in Equation 3a, we have that

$$\lim_{\alpha \rightarrow 1} \lim_{\delta \rightarrow 1} I_1^{\alpha\delta}(P) = I_1^1(P) = -\Sigma p_i \log p_i. \tag{7}$$

Finally, Hartley's measure in Equation 1 is a limiting case of $I_B^{\alpha\delta}(P)$, since

$$\lim_{\alpha \rightarrow 0} \lim_{\delta \rightarrow 1} I_1^{\alpha\delta}(P) = I_1^{01}(P) = \log n. \tag{8}$$

It should also be mentioned that, whenever one or more of the p_i are zero, we shall use the usual conventions $0 \log 0 = 0, 0^a = 0$ for $a > 0$ and $0^a = 1$ for $a = 0$.

CONCLUDING COMMENTS

While measurements of information in the behavioral sciences have almost exclusively been based on Hartley's and Shannon's classical entropies, other measures may be equally appropriate in many cases. This paper has defined a most general information measure that includes many of the proposed measures, with Hartley's and Shannon's as particular cases among them. The generalized measure $I_B^{\alpha\delta}(P)$ in Equations 3a and 3b, or $I_{B\lambda}^{\alpha\delta}(P)$ given by Equation 3a with λ as a separate arbitrary parameter when normalization is not imposed, has the advantage of offering the considerable flexibility provided by the various parameters. Such flexibility may prove useful in some behavioral research applications. A very simple special case of $I_B^{\alpha\delta}(P)$ would, for example, be (when $\Sigma p_i=1$)

$$I_1^{22}(P) = 2(1-\Sigma p_i^2),$$

or the nonnormalized measure

$$I_1^{22}(P) = 1-\Sigma p_i^2.$$

Some such applications of the new information measure are currently being explored by this investigator in studies of human information processing during tasks involving choice reaction time, process monitoring, and subjective versus objective uncertainty.

REFERENCES

ACZÉL, J., & DARÓCZY, Z. (1975). *On measures of information and their characterizations*. New York: Academic Press.
 DARÓCZY, Z. (1970). Generalized information functions. *Information and control*, **16**, 36-51.
 HARTLEY, R. V. L. (1928). Transmission of information. *Bell System Technical Journal*, **7**, 535-563.
 HAVRDA, J., & CHARVAT, F. (1967). Quantification method of classification processes: The concept of structural α -entropy. *Kybernetika*, **3**, 30-35.
 KAPUR, J. N. (1967). Generalized entropies of order α and type β . *The Mathematics Seminar*, **4**, 78-94.
 KAPUR, J. N. (1983). A comparative assessment of various measures of entropy. *Journal of Information & Optimization Sciences*, **4**, 207-232.
 KLIR, G. J., & FOLGER, T. A. (1988). *Fuzzy sets, uncertainty, and information*. Englewood Cliffs, NJ: Prentice-Hall.
 MATHAI, A. M., & RATHIE, P. N. (1974). *Basic concepts of information theory and statistics*. New Delhi, India: Wiley Eastern.

RÉNYI, A. (1961). On measures of entropy and information. In *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability* (Vol. 1, pp. 541-561). Berkeley: University of California Press.

SHANNON, C. E. (1948). The mathematical theory of communication. *Bell System technical journal*, 27, 379-423, 623-656.

SHERIDAN, T. B., & FERRELL, W. R. (1974). *Man-machine systems:*

Information, control, and decision models of human performance. Cambridge, MA: MIT Press.

WICKENS, C. D. (1984). *Engineering psychology and human performance*. Columbus, OH: Charles E. Merrill.

(Manuscript received August 29, 1988.)